

PARTICLE ACCELERATION
AND DETECTION

R. Hellborg (Ed.)

Electrostatic Accelerators

Fundamentals and
Applications

With a Foreword by Kai Siegbahn

 Springer

Electrostatic Accelerators

Particle Acceleration and Detection

www.springeronline.com/series/5267

The series *Particle Acceleration and Detection* is devoted to monograph texts dealing with all aspects of particle acceleration and detection research and advanced teaching. The scope also includes topics such as beam physics and instrumentation as well as applications. Presentations should strongly emphasise the underlying physical and engineering sciences. Of particular interest are

- contributions which relate fundamental research to new applications beyond the immediate realm of the original field of research
- contributions which connect fundamental research in the aforementioned fields to fundamental research in related physical or engineering sciences
- concise accounts of newly emerging important topics that are embedded in a broader framework in order to provide quick but readable access of very new material to a larger audience

The books forming this collection will be of importance for graduate students and active researchers alike.

Series Editors:

Professor Alexander Chao
SLAC
2575 Sand Hill Road
Menlo Park, CA 94025
USA

Professor Christian W. Fabjan
CERN
PPE Division
1211 Genève 23
Switzerland

Professor Rolf-Dieter Heuer
DESY
Gebäude 1d/25
22603 Hamburg
Germany

Professor Takahiko Kondo
KEK
Building No. 3, Room 319
1-1 Oho, 1-2 1-2 Tsukuba
1-3 1-3 Ibaraki 305
Japan

Professor Francesco Ruggiero
CERN
SL Division
1211 Genève 23
Switzerland

R. Hellborg (Ed.)

Electrostatic Accelerators

Fundamentals and Applications

With a Foreword by Kai Siegbahn

 Springer

Professor Ragnar Hellborg
Department of Physics
Lund University
Sölvegatan 14
22362 Lund
Sweden
E-mail: ragnar.hellborg@nuclear.lu.se

Cover picture courtesy of CERN.

Library of Congress Control Number: 2005922605

ISSN 1611-1052

ISBN-10 3-540-23983-9 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-23983-3 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable for prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media
springeronline.com

© Springer-Verlag Berlin Heidelberg 2005
Printed in The Netherlands

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typesetting: by the editor and TechBooks using a Springer $\text{\LaTeX}$ macro package
Cover design: eStudio Calamar Steinen
Cover production: *design & production* GmbH, Heidelberg

Printed on acid-free paper SPIN: 10946114 54/3141/JVG 5 4 3 2 1 0

To the memory of those pioneers of the 1920s and 1930s who
laid the ground for today's accelerators

Foreword

It is interesting to note from this book how intimately achievements in science and technology are connected to each other in the advancement of our knowledge in atomic and nuclear physics. They are so strongly related that it is often difficult to find out which really shows the direction of research. The field of accelerators and the associated new physics around it are a particularly good example of this. The fairly high energies which are required to penetrate into atomic electron core shells to produce hard X-rays for medical purposes and, even more, into nuclear matter were basic conditions for progress in nuclear and elementary-particle physics. On the other hand, one should remember that the very first information about atoms and nuclei was obtained by means of quite modest equipment without access to accelerators. Nuclear physics had its beginning with the discovery of radioactivity by Henri Becquerel and was investigated by Marie and Pierre Curie in Paris about a hundred years ago using very simple detectors to record the radioactive radiation. Later on, the actual discovery of the atomic nucleus and even of nuclear transmutations of nuclei was also made without accelerators. The pioneering work during this period was to a large extent performed in Ernest Rutherford's laboratory in Cambridge, which indicated that new technology had to be invented to be able to continue along this new scientific route. The "golden age" of nuclear physics during the 1930s was characterized by the invention of several powerful accelerator designs based on essentially new principles. This research resulted in a new area in nuclear physics, often called "big science". Electrostatic machines were invented and gradually developed to produce particle currents with energies of many MeV. Even more spectacular were the new ideas leading to the construction of accelerators with multiple acceleration stages under resonance conditions. A large part of this technological research and development was performed in the USA, in particular in Berkeley, by the inventor of the cyclotron Ernest Lawrence and his coworkers. Through an interesting interplay between theoretical particle physics and ingenious new accelerator concepts, the Bevatron was constructed there, with the direct purpose of producing a new theoretically anticipated nuclear particle, the antiproton (and the antineutron). It is well known that this new area of big science was extremely successful and resulted in a rather

detailed description of nuclear substructure and nuclear forces, leading to a deep theoretical understanding.

Electrostatic accelerators of much more modest size were developed in parallel for alternative uses in nuclear physics. It is not possible to reach very high energies with these machines, but they have other advantages. Because of their general design, they are often called belt generators; the most well-known of them is due to Van de Graaff and is therefore called the Van de Graaff accelerator. The energy regime for these is in the MeV region. They are stable and precise in their performance, and much less expensive. There is, nowadays, a great number of different types of accelerators used for various purposes. Altogether there are presently more than 15 000 accelerators in operation. Only a rather small fraction of them are actually used for nuclear research, as was the original purpose of their design. This book gives a very comprehensive account of the different design principles and also the different scientific fields where accelerators are of basic importance. A great emphasis in this book is given to the clear and pedagogic presentation of this vast amount of material. The book is well organized and up to date. The authors are all well-known experts in their fields. This excellent book will therefore serve a very useful purpose. The authors have successfully tried to give a historical perspective on the development, which makes the reading much more interesting. The emphasis on the intimate coupling between new technological ideas and new fields of scientific progress is another feature of the book. This should stimulate young people at the beginning of their career to look for new experimental methods and ways to do research in various promising fields. There are several examples in the book that illustrate how accelerator techniques and associated methods have opened up new scientific fields. Such examples are found, for example, in atomic spectroscopy and accelerator-based mass spectrometry (AMS), applied to biology and archaeology, for example. In the latter field, there are already more than 50 accelerators now in use especially designed for such purposes. The number of new, powerful synchrotron accelerators for work in materials science, surface science, biology etc. is steadily increasing and is presently around 70.

Uppsala, February 2004

A handwritten signature in dark ink, reading 'Kai Siegbahn'. The signature is fluid and cursive, with the first name 'Kai' and last name 'Siegbahn' clearly distinguishable.

Kai Siegbahn

Editor's Note: Kai Siegbahn has been a key figure in many of the developments in atomic and molecular physics, nuclear physics, and electron optics that are described herein. His research contributions are manifold, and he was awarded the Nobel Prize in Physics in 1981 for his work on electron spectroscopy for chemical analysis.

Preface

This book is intended to provide an introduction to the field of accelerator techniques and to give overviews of some of the more important research fields. The book is mainly concentrated on electrostatic accelerators. In these accelerators, the acceleration takes part in one large voltage step and the high voltage is obtained by a mechanical charging system consisting of an insulating belt or a metal chain. Several other types of accelerators are also discussed in some detail. Many topics are of course common to all types of accelerators, for example ion sources, vacuum, high voltage, ion optics, equipment for beam diagnostics, computer control and radiation protection. The book will therefore also be of value to readers interested in other types of accelerators.

The book is divided into three parts. In the first part of the book, an introduction to the field of accelerators is given, as well as chapters about two fields of great importance today in the exploitation of accelerator technology, namely, medicine and synchrotron radiation. The second part deals with the technique involved in the construction, operation and maintenance of the accelerator. The third part covers several examples of research utilizing an accelerator.

The electrostatic accelerator, either single-stage or double-stage (i.e. tandem), is one of the most versatile tools when one is using ion beams of comparatively low (i.e. a few MeV) energies. The high quality of the beam, as concerns its energy stability and emittance, the possibility of accelerating virtually every kind of ion and also continuously varying the energy over a wide range, combine to permit detailed investigations in many interesting fields of research. Among these we can mention nuclear and atomic spectroscopy, heavy-ion reactions, accelerator mass spectrometry, and ion beam analysis and modification. However, many other fields, including applied research, benefit from the use of an electrostatic accelerator.

Throughout the book, the general recommendations of the International Union of Pure and Applied Physics (IUPAP) for the writing of physical quantities, symbols, mathematical expressions, units, prefixes etc. have been followed. The SI system has been used with very few exceptions.

Many people have contributed to the production of this book. The authors of the various chapters/boxes – each person a world-known expert in his/her

field – have played the most important role in this project. But it is not only the authors who have made considerable contributions. The Editor and many of the authors have received support and contributions from a number of people around the world. Two of these people are Professor Kohei Furuno, University of Tsukuba, and Ms. Martha J. Meigs, Holifield Radioactive Ion Beam Facility, Oak Ridge, USA. Max Strandberg has converted a number of graphical figures to a suitable format; he has also produced all of the drawings in those chapters written by the Editor. Carl Erik Magnusson and the members of the AMS group in Lund have carefully gone through all of the manuscripts and given valuable suggestions for improvements. Professor Kai Siegbahn has reviewed a number of chapters and given valuable comments; he has also suggested a few new chapters and kindly written the Foreword to the book.

Throughout this book, several registered trademarks are used. These include Dynamitron, which is a registered trademark of Radiation Dynamics Inc.; Pelletron, which is a registered trademark of National Electrostatic Corp.; ConFlat, which is a registered trademark of Varian Associates Inc.; Tandetron, which is a registered trademark of High Voltage Engineering Europe; and VIVIRAD-ITC, which is a registered trademark of VIVIRAD.

During the last few decades, accelerators have undergone a considerable amount of development and improvement. As a result, a number of fascinating problems, not only in the field of physics, but also in many other fields such as technology, chemistry, biology and medicine, can be studied today with the help of accelerators. These developments and improvements have also led to a much more user-friendly apparatus. In many cases the users of today can start up and operate the machine themselves – tasks that earlier required a well-trained support staff.

A few comments concerning frequently used acronyms and proprietary nomenclature may be helpful. A few commercial suppliers of electrostatic accelerators are mentioned a number of times throughout the book, sometimes with their full name, and sometimes abbreviated. Some of them are as follows. HVEC (High Voltage Engineering Corporation, Burlington, USA) was founded in 1947 by Robert Van de Graaff, Denis Robinson and John Trump. HVEC was taken over by VIVIRAD in 1987 and was operated under the name VIVIRAD-HIGH VOLTAGE until 2000. NEC (National Electrostatic Corporation, Middleton, Wisconsin, USA) was founded in 1965 by Ray Herb. HVEE (High Voltage Engineering Europe, Amersfoort, The Netherlands) was originally a branch of HVEC but today is a separate company. At General Ionex Corporation (founded by Ken Purser in 1969), Purser began in 1978 to produce Tandetrans. This production was taken over by HVEE in April 1989. VIVIRAD, Handschuheim, France, was founded by Michel Letournel in 1983.

In the literature, a somewhat confusing use of names for those accelerators charged by a belt or a chain can be found. In this book, the name

“electrostatic accelerator” is used throughout. The name “Van de Graaff” or “VdG,” which not so seldom can be found in the literature, is in this book reserved for accelerators produced by HVEC (and later HVEE). The name “Pelletron” is used when talking about an accelerator produced by NEC. The name “Tandetron” is used for accelerators (of the cascade type) originally produced by General Ionex Corporation and today produced by HVEE.

Lund
May 2005

Ragnar Hellborg

Contents

Part I Accelerators

Introduction to Part I – Accelerators

R. Hellborg, J. McKay 3

1 Accelerators – an Introduction

R. Hellborg 4

2 Accelerators for Medicine

R. Hellborg, S. Mattsson 24

3 Accelerators for Synchrotron Radiation

M. Eriksson, S. Sorensen 36

Part II The Electrostatic Accelerator

Introduction to Part II – the Electrostatic Accelerator

R. Hellborg, J. McKay 51

4 History of the Electrostatic Accelerator

J. McKay 52

5 Electrostatics

H.R.McK. Hyder 64

Box 1: Calculation Technique for High-Voltage Equipment in Gas

K.A. Rezvykh, V.A. Romanov, R. Hellborg 84

6 Charging Systems

C. Westerfeldt 89

Box 2: Development of Charging Belts in Russia

V.A. Romanov 101

Box 3: Cascade Generators

R. Hellborg 104

**7 Voltage Distribution Systems – Resistors
and Corona Points**

D. Weisser 110

8 Accelerator Tubes

H.R.McK. Hyder 123

**Box 4: Development of Tubes
in Obninsk, Russia**

V.A. Romanov 147

9 Stabilization

L. Rohrer, H. Schnitter 152

10 Stripper Systems

D. Weisser 166

Box 5: Charge Exchange and Electron Stripping

H.J. Whitlow, H. Timmers 181

Box 6: Carbon Stripper Foils – Preparation and Quality

V. Liechtenstein 187

11 Positive-Ion Sources

L. Bartha 192

**12 Negative-Ion Formation Processes
and Sources**

G.D. Alton 222

Box 7: Tandem Terminal Ion Source

G.C. Harper 274

13 Ion Optics and Beam Transport

J.D. Larson 278

14 Beam Envelope Techniques for Ion-Optical Calculations

S. Bazhal, R. Hellborg 299

15 Equipment for Beam Diagnostics

M. Friedrich 317

16 Computer Control

M.L. Roberts 328

17 Radiation Protection at an Accelerator Laboratory

R. Hellborg, C. Samuelsson 337

Box 8: Nonradiation Hazards and Safety Considerations*R. Hellborg, C. Samuelsson* 365**Box 9: Confined-Space Maintenance***G.A. Norton* 369**Box 10: Earthquake Protection – for Pelletrons***G.A. Norton* 372**Box 11: Earthquake Protection
of the Bucharest FN Tandem Accelerator***S. Dobrescu, L. Marinescu* 374**18 Electrostatic-Accelerator Free-Electron Lasers***A. Gover, Y. Pinhasi* 378

**Part III Research Fields
and Their Technical Requirements**

**Introduction to Part III – Research Fields
and Their Technical Requirements***R. Hellborg, J. McKay* 393**19 Isotope Production
for Medical Applications***A.D. Roberts, T.E. Barnhart, R.J. Nickles* 395**20 Nuclear Structure***C. Fahlander, D. Rudolph* 413**21 Nuclear Reactions***L. Corradi* 429**22 Detection of Explosives and Other Threats
Using Accelerator-Based Neutron Techniques***T. Gozani* 445**23 Accelerator Mass Spectrometry***L.K. Fifield* 461**24 Atomic Collisions in Matter***J. Keinonen* 486**25 Modification of Materials
by MeV Ion Beams***Y. Zhang, H.J. Whitlow* 506

26 Ion Beam Analysis

K.G. Malmqvist 530

27 Atomic Structure

L.J. Curtis, I. Martinson 560

28 Industrial Electron Accelerators

M. Letournel 581

**A Appendix: Electrostatic Accelerators – Production
and Distribution**

H.R.McK. Hyder, R. Hellborg 595

B Appendix: SI Units and Other Units

R. Hellborg 604

Index 608

Part I

Accelerators

Introduction to Part I – Accelerators

R. Hellborg¹ and J. McKay²

¹ Department of Physics, Lund University, Sölvegatan 14, 223 62 Lund, Sweden
`ragnar.hellborg@nuclear.lu.se`

² Box 463, Deep River, Ontario, Canada K0J 1P0
`mckay@magma.ca`

In this first part of the book, the readers are introduced to the field of accelerators. It is fascinating to follow the development from the very simple machines constructed around 1930, a few meters in size, to the enormous machines of today with dimensions of several kilometers.

The focus of this book is the electrostatic accelerator, but to put these devices into perspective, it will be useful to examine two other fields of great importance today in the development and exploitation of accelerator technology. These examples are described in two chapters of this introductory section.

Some of the earliest applications of accelerator technology were in medicine. Medical accelerators are used primarily for treatment of cancer patients. This usually, but not exclusively, is done through the generation of X-rays by the acceleration of electrons. Today, accelerators are also important for the production of isotopes used for diagnostic imaging and for sterilization of medical supplies.

Synchrotron radiation facilities provide a new, high-tech apparatus for investigations in the fields of physics, chemistry, biology, crystallography and medicine, as well as for industrial applications. The product of these machines is not a particle beam, but photons. The photon beam can be high-energy, high-brilliance, coherent, tunable and monochromatic. These “light sources” cover most of the electromagnetic spectrum and open new possibilities in the use of photon–matter interactions, complementing the particle–matter interactions of other accelerator techniques.

The greatest significance of these two examples may be their importance beyond basic particle research, but their origins are to be found in the development of electrostatic and other particle accelerators. There is great interplay amongst the various disciplines of accelerator science. Material-science and AMS techniques are being used in medical research, and synchrotron radiation is being used in medical imaging. Synchrotron radiation is being used across the complete range of research and applications. However, the electrostatic-accelerator laboratories have been and still are the source of much of the technology used in all the other accelerator fields.

1 Accelerators – an Introduction

R. Hellborg

Department of Physics, Lund University, Sölvegatan 14, 223 62 Lund, Sweden
`ragnar.hellborg@nuclear.lu.se`

1.1 Introduction

One of the first uses of energetic charged particles – a “beam” of ions – was when the inner structure of atoms was investigated to verify predictions by Ernest Rutherford in 1911. α -particles obtained from Ra and Th natural sources were used, and in this famous experiment the existence of a positively charged nucleus having a diameter of less than 10^{-13} m was demonstrated. Today, the same well-known technique is used in trace element analysis, surface science, solid state physics, etc. However, now well-defined beams of light ions (such as p, d and He) of a few MeV from an accelerator are used. The technique is named RBS, which is an abbreviation for “Rutherford backscattering.”

The next remarkable step took place in 1919 when Rutherford achieved the disintegration of the nitrogen nucleus by α -particle bombardment [1]. The α -particles – again obtained from a natural α -source – hit a target containing nitrogen and produced the first artificially created nuclear reaction. In the reaction, an α -particle enters a nitrogen nucleus, forming a compound nucleus, which quickly splits up into an oxygen and a hydrogen nucleus, according to the following scheme:



The intensity of the α -radiation from a natural source is of course very weak, and the beam is not collimated at all. As an example, an α -source with an activity of 1 GBq (10^9 disintegrations per second, or 27 mCi in older units) emits the particles into a solid angle of 4π and therefore provides a flux density at a distance of 100 mm from the point source of 8×10^5 particles $\text{cm}^{-2} \text{s}^{-1}$. A beam of $1 \mu\text{A}$ from an accelerator, which may easily be collimated to a cross-sectional area of 1cm^2 , delivers 6×10^{12} singly charged particles per second.

These two famous experiments conceived by Rutherford demonstrated the demand for beams of particles with much higher intensities, a well-defined energy and the possibility to freely choose the particle species and their energy, i.e. the demand for accelerators. The goal was to continue the investigations of Rutherford and collaborators of nuclear reactions induced by well-defined

beams and to obtain particle energies high enough to penetrate the Coloumb barrier which surrounds all nuclei.

During the 1920s the X-ray technique developed rapidly, and DC equipment for producing voltages of a few hundred kV became available. Unfortunately, higher voltages were limited by corona discharging and insulation problems. The MV range seemed at that time to be impossible to reach. At the end of the 1920s, the development of quantum mechanics showed that charged particles could penetrate through the potential wall around an atom and therefore that particle energies of 0.5 MeV or less could be enough for splitting light atoms. This was a more moderate goal, and accelerator development started in different laboratories.

The first persons to reach the goal of initiating a nuclear reaction by use of a beam from an accelerator were J.D. Cockcroft and E.T.S. Walton at the Cavendish Laboratory in Cambridge [2]. In 1932 they had a working proton accelerator, and with a beam of 400 keV they induced the reaction



In 1951, Cockcroft and Walton obtained the Nobel Prize; see Table 1.1.

The years around 1930 can be taken as the starting point of the accelerator era, and people at different laboratories did development work following

Table 1.1. Nobel Prizes in Physics awarded to accelerator pioneers

Laureate(s)	Year	Awarded for
E.O. Lawrence	1939	<i>the invention and development of the cyclotron and for results obtained with it, especially with regard to artificial radioactive elements</i>
J.D. Cockcroft and E.T.S. Walton	1951	<i>their pioneer work on the transmutation of atomic nuclei by artificially accelerated atomic particles</i>
E.M. McMillan (in chemistry) (shared with G.T. Seaborg)	1951	<i>their discoveries in the chemistry of the transuranium elements</i>
J. Schwinger (shared with S. Tomonaga and R.P. Feynman)	1965	<i>their fundamental work in quantum electrodynamics, with deep-ploughing consequences for the physics of elementary particles</i>
L.W. Alvarez	1968	<i>his decisive contributions to elementary particle physics, in particular the discovery of a large number of resonance states, made possible through his development of the technique of using hydrogen bubble chamber and data analysis</i>
C. Rubbia and S. Van der Meer	1984	<i>their decisive contributions to the large project, which led to the discovery of the field particles W and Z, communicators of weak interaction</i>

different principles. Ernest Lawrence and colleagues developed the cyclotron; Robert Van de Graaff, Ray Herb and others the electrostatic accelerator; and Cockcroft and Walton the cascade accelerator. Gustaf Ising outlined and Rolf Widerøe built the linear accelerator. During the 1930s and 1940s the development of accelerators was based on the need of nuclear physicists to obtain higher projectile energies to exceed the Coulomb barrier of heavier and heavier elements for nuclear-structure studies, for the production of radioactive isotopes etc. Already in 1937, the first electrostatic accelerator was constructed for clinical use [3]. This machine – an electron accelerator built for the Harvard Medical School – could, in the energy range 0.5–1.2 MeV, produce X-ray intensities up to 40 R/min (or, in modern units, 0.01 C/kg_{air}) per mA electron beam current. The maximum obtainable electron beam current was 3 mA. This was a unique machine at the time and marked the first use of an electrostatic accelerator in clinical work. A schematic drawing of this machine can be seen in Fig. 2.1 in Chap. 2. Beginning in the early 1950s, the community of particle physicists also took part in formulating the goal of accelerator development. The first accelerators were built for protons and electrons. Today, ions from all elements in the periodic table can be accelerated. Furthermore, it is now also possible to handle artificially produced isotopes, i.e. short-lived isotopes far from the stability line, and antiparticles such as positrons and antiprotons. Today accelerators are applied in very diverse fields, such as radiotherapy, isotope production, ion implantation, synchrotron light production, spallation, neutron production, radiography, sterilization and inertial fusion, besides, of course, basic research in nuclear and particle physics. Of the more than 15 000 accelerators in operation around the world, only a handful are used in elementary-particle-physics research, a few hundred are used in physics and applied-physics research, and one-third are involved in medical applications, such as therapy, imaging and the production of short-lived isotopes. The other two-thirds are used for industrial applications, ranging from electron beam processing and micromachining, to food sterilization, and for national-security applications, which include X-ray inspection of cargo containers and nuclear-stockpile stewardship.

The tremendous progress in the construction of accelerators since the 1930s is illustrated in Fig. 1.1, showing an exponential increase of about an order of magnitude in beam energy per seven years! This graph is called a “Livingston plot” after Stanley Livingston, the accelerator physicist who first constructed such a plot in the 1960s. How can this development have happened? Maybe it can be explained in the following way. The progress of each type of accelerator (electrostatic, cyclotron, synchrocyclotron etc.) has saturated fairly quickly, whereas new ideas have been proposed regularly and have been the main contributors to the rapid advance. The development has been made possible by repeated use of the cycle

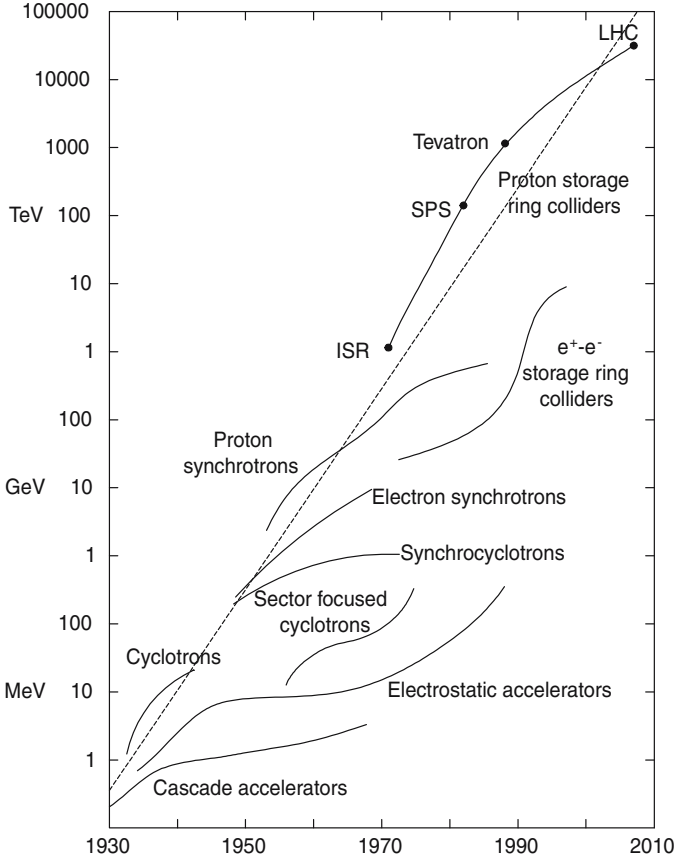


Fig. 1.1. A modified Livingston diagram showing the exponential growth of accelerator beam energy. The extrapolation to 2007–08 ends with the LHC at CERN. The energy of colliders is plotted in terms of the laboratory energy of particles colliding with a proton at rest to reach the same center-of-mass energy (This is described in more detail in Sect. 1.6)

New idea → Improved technology → Until saturation → New idea etc.

In the same time period, the cost per eV beam energy has been drastically reduced, roughly by a factor of one thousand. A few startling examples seen in Fig. 1.1 can be pointed out: the development of the cyclotron and the electrostatic accelerator during the 1930s, the development of the synchrocyclotron during the 1940s and 1950s, the invention of alternating-gradient focusing in the 1950s and the application of the colliding beams in the 1960s and 1970s. The question to be asked is whether there is a different paradigm for building particle accelerators at the energy frontier which will dramatically reduce their size and cost. One approach discussed in the literature

in recent years is to accelerate particles by collective fields in plasmas, by laser-induced acceleration and by using the field from a low-energy beam to accelerate a high-energy beam. A few words about this can be found at the end of this chapter.

As is shown in Fig. 1.1, accelerators can be classified into different principal designs, but all of these designs are of course based on the only known method to accelerate a particle: to charge it and then apply an electrical field. This occurs either in one big step or in several smaller gaps. Below, a very brief overview of the different design principles is given. The details of accelerator technology are not discussed in this chapter; instead, the general categories are presented, and the strengths and weaknesses of each category are discussed.

1.2 Direct Voltage Technique

In accelerators based on this principle (other names are “potential-drop accelerators” and “high-voltage DC accelerators” because the current is DC, contrary to all other accelerator types), the particle (after ionization) is accelerated through an accelerator tube, in one step. The tube is constructed as a long rectilinear drift tube, with a number of electrodes along the axis with a controlled voltage for each electrode, partly to aid in focusing the beam and partly to distribute the voltage gradient uniformly along the insulation surfaces. The positive ions (or electrons) to be accelerated are generated in an ion source located at high voltage (except for tandem accelerators – see below – in which negative ions are generated at ground potential). Direct voltage accelerators are often identified with the type of high-voltage generator used. The high voltage can be generated by rectifying an AC voltage (such a generator is often called a cascade generator) or by using electrostatic charging, in which a mechanical system carries the charge to the high-voltage terminal (these accelerators are called electrostatic accelerators). An open-air accelerator fails above a few MV, mainly because of the moisture in the air, which causes sparks. The voltage available today, if the accelerator is enclosed in a tank with a suitable gas under high pressure, is up to a few tens of MV.

Cascade accelerators. The high-voltage unit consists of a multiplying rectifier–condenser system (first used by Cockcroft and Walton [2]). Accelerators of this type, enclosed in a tank, are today designed for use up to 5 MV. The main advantage is that a cascade accelerator has a large output current of up to several hundred mA. This type of generator has for several years been used as an injector for high-voltage accelerators owing to the high beam current that such an accelerator can handle. In Box 3, more details about the various design principles of cascade generators will be found, as well as schematic drawings of the various principles. In Chap. 5, a photo of Cockcroft and Walton’s first machine is shown.

Electrostatic accelerators. In 1929, Robert Van de Graaff demonstrated the first generator model of this type [4]. An electrostatic charging belt is used to produce the high voltage. The principal design of the charging system is described in Chap. 6, and a photo of one of Van de Graaff's first open-air machines is shown in Chap. 5. In Fig. 1.2, Van de Graaff is demonstrating one of his first test generators for Karl Compton. Two rollers are provided, one at ground potential driven by a motor and the other located in the high-voltage terminal, well insulated from ground. Over the rollers passes an endless belt of insulating material. Charge is sprayed from sharp corona points onto the moving belt. The belt conveys the charge to the insulated high-voltage terminal, within which the charge is removed by collector points and allowed to flow to the surface of the electrode. After development into a working accelerator, this principle was used in thousands of accelerators



Fig. 1.2. Robert Van de Graaff (*on the left*) and Karl Compton in 1931. This test generator was a double unit consisting of positive and negative hollow spheres mounted upon upright Pyrex rods. Each sphere was charged by a silk ribbon belt, running from a grounded motor pulley at the base of the rod to a pulley in the interior of the sphere. Intersphere voltages of 1.5 MV were reported from this simple machine (Reprinted from [5], copyright 1974, with permission from Elsevier)

around the world. The reasons why this type became so popular are that all types of ions can be accelerated, the ion energy can be changed continuously, the high-voltage stability is extremely good and therefore the ion energy has a very low energy spread. An electrostatic accelerator provides an advantage over a cascade accelerator – the terminal voltage is extremely stable and lacks the AC ripple of the cascade accelerator. A disadvantage is the low current output compared with the cascade accelerator. All modern accelerators of the electrostatic type are enclosed in a tank containing gas under high pressure to reduce the size and to be independent of moisture in the air. The first pressure-insulated machine was initiated by Ray Herb in the autumn of 1933. In Fig. 1.3 Herb is seen working on one of his first accelerators, the “Long Tank” machine.



Fig. 1.3. Ray Herb and the “Long Tank” machine in 1936 (Courtesy of Mrs. Ann Herb)

Tandem electrostatic accelerators. During the 1950s negative-ion sources were developed, i.e. beams of ions with an extra electron added to the ordinary electrons became available. This development made it possible to build two-stage (or tandem) accelerators. In a tandem, the high voltage is utilized twice, as can be seen in Fig. 9.2 in Chap. 9. Negative ions are formed at ground potential and injected into the first stage, where acceleration to the positive high-voltage terminal takes place. The energy gain is eU_T eV, where e is the elementary charge and U_T is the terminal voltage. In a stripper system in the high-voltage terminal, the negative ions lose a few electrons and change charge into positive ions. In the second stage, the positive ions once more gain energy. Now the energy gain is qeU_T eV, where q is the charge state

of the ion. Thus a total energy gain of $(q + 1)eU_T$ eV is obtained. For heavy ions and high-voltages, i.e. high-speed ions undergoing stripping, q can be quite high and therefore the final energy of the ions can be hundreds of MeV.

As we shall learn later in this book, the insulating belt for transport of the charge has been replaced in many modern accelerators by a chain of metal cylinders. In this way, a more robust transport system with much more well-defined charge transport and hence a better voltage stability is obtained.

A method similar to the use of a belt or chain has been used by a French company to obtain up to a few hundred kV in a small and compact machine. In this accelerator, often used as a neutron generator to obtain 14 MeV neutrons through the $(p + T)$ reaction, a cylinder rotates and transports charge at a high speed around another fixed cylinder, all enclosed in a tank containing gas at high pressure; see Fig. 1.4.

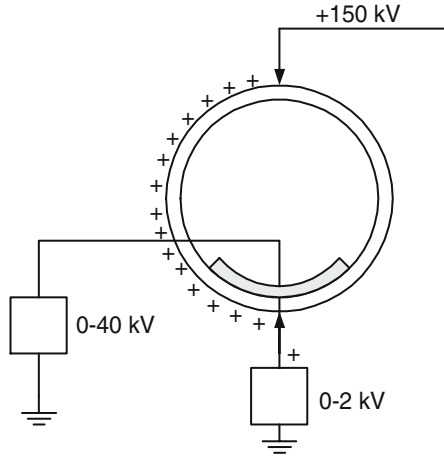


Fig. 1.4. Electrostatic generator using a rotating cylinder to obtain the high voltage

1.3 Resonance Acceleration

The second principle is the use of resonance acceleration by using a radio-frequency field. In the case of this principle, the particle has to pass through a small potential gap several times in resonance with an oscillating electric field. In this way, a much bigger energy gain will be obtained compared with the acceleration voltage. This can be done either in a series of gaps in a straight line called a linear accelerator or with a single gap in a circular machine, i.e. cyclotrons, betatrons, microtrons etc.

Linear accelerators. In 1924, Gustaf Ising from Sweden proposed [6] a method of particle acceleration that would give particles more energy than

Prinzip einer Methode zur Herstellung von Kanalstrahlen hoher Voltzahl.

Von

GUSTAF ISING.

Mit 2 Figuren im Texte.

Mitgeteilt am 12. März 1924 durch C. W. OSEEN und M. SIEGBAHN.

Die folgenden Zeilen beabsichtigen eine Methode zu skizzieren, welche im Prinzip erlaubt, mit einer zu Verfügung stehenden mässigen Spannung Kanalstrahlen (ev. Kathodenstrahlen) beliebiger Voltzahl zu erzeugen. Dies soll dadurch

Fig. 1.5. Ising's article from 1924

that provided by the maximum voltage in the system. The first part of his article can be seen in Fig. 1.5. In 1928, Rolf Widerøe from Norway built the first linear accelerator [7] by using a radio-frequency field over two gaps, and accelerated sodium and potassium. The principle of a linear accelerator (or linac as it often is called today) is shown in Fig. 1.6. The beam travels through a series of hollow, tubular electrodes connected alternately to opposite poles of the RF voltage source. Particles are accelerated as they cross the gaps between the electrodes. Upon entering the interior of an electrode, the particle drifts in a field-free region for a time equal to half the period of the RF voltage. In this way, the polarity of the voltage is reversed during the time the particle is within the drift tube, and the particle is then accelerated as it crosses the next gap. The available RF technology has been decisive for the development of linacs. Today, the highest obtainable linac energy is 40 GeV for electrons and 800 MeV for protons. In 1947, Luis Alvarez built

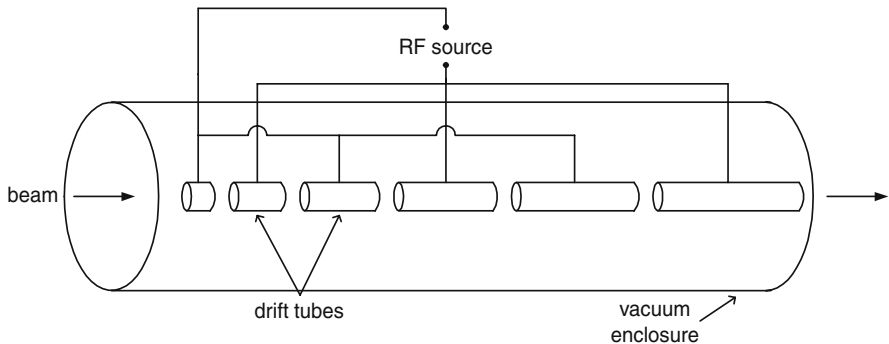


Fig. 1.6. Principle of a linear accelerator

a machine [8] with a structure which differed significantly from the Widerøe structure. The Alvarez structure consists of a set of resonator tubes which have an RF voltage of the same phase applied to them. Inside each resonator tube, a potential distribution exists. Therefore the acceleration takes place in the resonator. A type of standing wave is formed that ensures particle beam acceleration. In 1968 Alvarez obtained the Nobel Prize; see Table 1.1.

Radio-frequency quadrupole (RFQ). A rather new type of low-energy accelerator for very high currents is the RFQ, first proposed by I.M. Kapchinski and V.A. Teplyakov in 1970. The RFQ has a symmetry corresponding to that of an electrostatic quadrupole; see Fig. 1.7. It combines the action of focusing and bunching the beam, in addition to acceleration proper. The acceleration is of a continuous character and therefore it does not occur only in electrode gaps or other structures. The bunching effect is very efficient and close to 100%. Focusing is ensured by a transverse electrical gradient. A 1–2 m long RFQ can accelerate ions from an energy of a few tens of kV up to several MV. RFQs are often used today as part of the injector of big accelerator systems, and in that way are replacing old cascade injectors. The RFQ accelerator is a compact and rather simple accelerator.

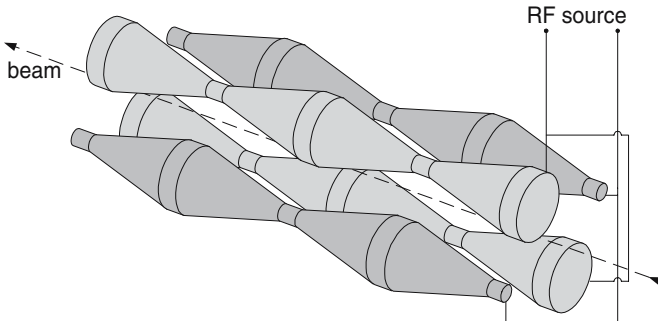


Fig. 1.7. Principle of an RFQ accelerator

Cyclic accelerators. In 1929, Ernest Lawrence at the University of California at Berkeley discovered Widerøe's article [7] and realized – see Fig. 1.8 – that if the ions could somehow be returned to the first gap again, and again, multiple acceleration could take place. In 1930, Lawrence proposed the application of the Widerøe resonance principle, but now inside a homogeneous magnetic field such that the particle would be bent back to the same RF gap twice in each period of the radio-frequency field. The Lawrence type of accelerator is called the cyclotron, and the principle is illustrated in Fig. 1.9. The first demonstration model cyclotron for accelerating particles was built by Lawrence and coworkers (mainly Stanley Livingston) in 1931 [9]. In Fig. 1.10, Lawrence is at the control panel of his 37 inch cyclotron. In 1939, Lawrence was awarded the Nobel Prize; see Table 1.1. Within a flat, cylindrical vacuum

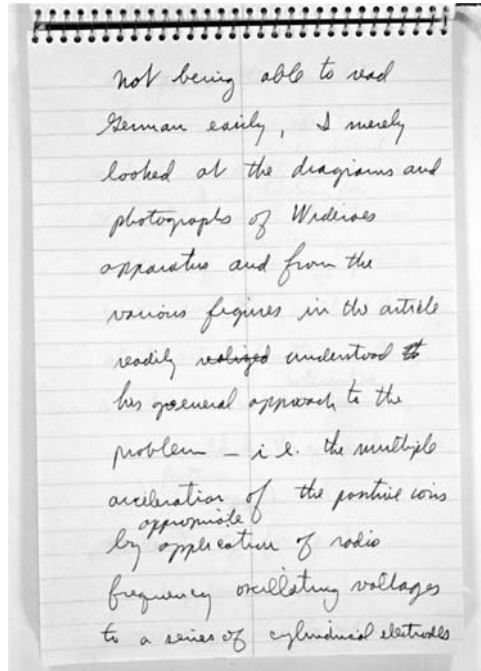


Fig. 1.8. Lawrence's handwritten interpretation of Widerøe's diagram (Reprinted with permission from LBL)

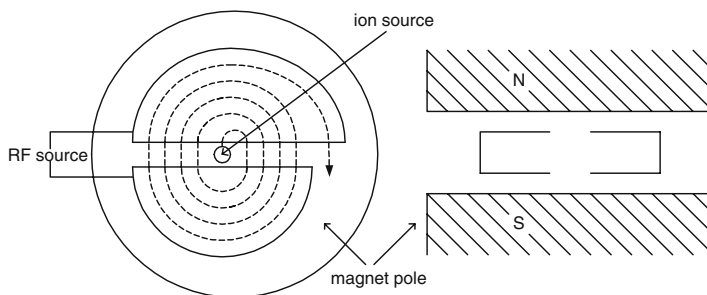


Fig. 1.9. Principle of a cyclotron

chamber placed between the poles of a dipole magnet are two D-shaped electrodes consisting of hollow, flat half-cylinders. In Lawrence and Livingston's first practical cyclotron the "dees" had a radius of 0.125 m, a magnetic field of 1.3 T and frequency of 20 MHz, giving protons of 1.2 MeV. The two dees are connected to the RF source (usually within the range of 10–30 MHz) so that an alternating voltage appears across the gap separating the dees. An ion source, located at the center of the chamber, produces the ions and supplies



Fig. 1.10. Ernest Lawrence at the controls of the 37 inch cyclotron in about 1938 (Reprinted with permission from LBL)

them with a low initial velocity. The path of the ions (with an electric charge q and mass m) is circular in the magnetic field B . The radius R of the circle is given by

$$R = mv/qB \quad (1.3)$$

Since R is proportional to the velocity v , the period of circulation T (and thereby also the frequency f) is constant for all values of R . Once in resonance, the ion will receive an energy gain each time it passes through the acceleration gap between the dees.

The energies possible to obtain with protons are up to 20–30 MeV, corresponding to a magnet diameter of little more than 1 m. As the beam in a cyclotron travels outwards towards the edge of the magnet, the magnetic field lines are diverted somewhat from true straight lines. The curvature of the field lines gives a net force component towards the median plane, which tends to provide focusing and to counteract the tendency of the beam to diverge. At the same time, the field loses its uniformity and the resonance condition can no longer be maintained. The advantage of a cyclotron compared with an electrostatic accelerator is that a much higher beam current (tens of mA) is available from a cyclotron. The disadvantages of the cyclotron are that the beam is pulsed, it is difficult to change the beam energy and, normally, this change cannot be made continuously. The energy resolution of the beam is also much worse compared with the electrostatic accelerator.

It is interesting to notice that subsequent to its initial conception, there has not been a fundamental change in the method of resonance acceleration. Ising's, Widerøe's and Lawrence's approaches can be considered as the beginning of modern RF accelerator development. All later inventions, some of which have been tremendously important as we shall see, have addressed the

problems of efficient guiding and collection of the particles rather than the acceleration itself.

The cyclotron cannot accelerate particles as light as electrons, as they will quickly become relativistic and (1.3) is no longer valid. D.W. Kerst invented and constructed in 1940 [10] the first circular electron accelerator, a 2.35 MeV betatron. Widerøe had formulated the principal design of a betatron already in 1928, but he was unable to make the device work. Kerst realized that the magnetic field must be shaped to provide focusing and prevent the electrons from escaping. A betatron consists of a ring-shaped accelerating chamber located between the poles of an electromagnet whose windings are supplied from a 50 or 60 Hz sinusoidal voltage. The magnetic field plays a dual role. First, it causes the trajectories of the electrons – injected at low energy – to curve and keeps the electrons in a circular orbit. Secondly, it accelerates the electrons by means of the electric field induced by the change in the magnetic flux passing through the circular electron orbit. Kerst built first a 2.3 MeV model that worked immediately. He then built a 20 MeV machine, and then a 300 MeV machine. This was the largest betatron ever built. At General Electric, a 100 MeV betatron was built to produce intense X-ray beams. 100 MeV is a high enough energy to make detectable the relativistic radiation emitted by electrons traveling along curved paths. This radiation, now known as “synchrotron radiation”, is emitted in a small cone directly ahead of the electron. The importance today of synchrotron radiation is outlined in Chap. 3.

Another principle is used in the microtron, which has a constant field and a fixed frequency. The particle moves in a circular orbit between the pole pieces of a magnet. The orbits share one common point, at which an RF acceleration resonator is located. The relativistic limitation of the cyclotron can be overcome if the increase of mass¹ at each gap transit is so large that the revolution time increases by one radio-frequency period. The increase in energy required at each gap transit can be obtained from the cyclotron equation corresponding to one rest mass. This is impossible to achieve for protons or heavy ions but is practicable for electrons. The principle was suggested by Vladimir Veksler and Julian Schwinger in 1944, and is illustrated in Fig. 1.11. In 1965, Schwinger was awarded the Nobel Prize; see Table 1.1. The beam current in a microtron is of the order of μA and the usual operating energy is in the 5–50 MeV range. Microtrons are mostly used as injectors and for industrial radiography.

1.4 Phase-Stabilized Acceleration

The original cyclotron relied on the fact that the revolution frequency of a charged particle in a homogeneous field is independent of the particle energy

¹ Although the concept of speed-dependent mass is not explicit in modern theory, but is inherent in the definition of the relativistic linear momentum, this concept is useful in accelerator technology.

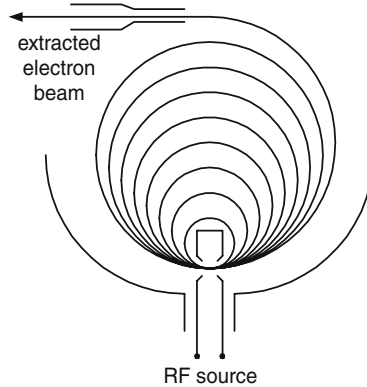


Fig. 1.11. Principle of a microtron

in the nonrelativistic approximation, and the radio frequency used for the acceleration could be kept constant. The increase in particle mass due to relativistic effects limits the energy that can be reached. The remedy is to modulate the applied RF field to keep in step with the cyclotron frequency. In 1945, the third principle, which is to use phase-stabilized acceleration, was proposed independently by Edvin McMillan [11] and Vladimir Veksler [12]. In 1951, McMillan shared (with Glenn Seaborg) the Nobel Prize for chemistry; see Table 1.1. In this type of accelerator, called the *synchrocyclotron*, the frequency f of the applied RF field decreases with increasing particle energy to compensate for the changing mass. This means that the particles travel through the synchrocyclotron in bunches, and the frequency is swept from its maximum value (when the bunch is near the center, the particles are only slightly accelerated and the relativistic increase in mass is slight) to its minimum value (when the bunch is ready to exit the accelerator, the maximum energy has been attained and the mass has its largest value). Technically, the periodic change of f is carried out using a condenser whose capacitance varies when it is rotated. Synchrocyclotrons are sometimes also called frequency-modulated or FM cyclotrons. The limit on synchrocyclotron size is set by magnet cost, which is proportional to $E^{3/2}$, and a maximum energy of up to 1 GeV for protons has been obtained. The disadvantage of a synchrocyclotron compared with a cyclotron is the reduced current. Only one bunch at a time is sent through the synchrocyclotron, compared with lots of pulses through the cyclotron; therefore the beam current is reduced to a mean value of μA or even less. The first synchrocyclotron was built in Berkeley in 1948. It could accelerate protons to 350 MeV and was the first machine used for studies of π -mesons.

Another way of overcoming the problem connected with the lack of resonance due to the increase of the relativistic mass in a homogeneous magnet can be to use an *azimuthally-varying-field* (AVF) cyclotron (also called an

isochronous cyclotron) having an increasing magnetic field with increasing radius. Vertical focusing is obtained by use of radial or spiral ridges built on to the poles to create alternate high- and low-field sectors. Focusing forces giving axial stability arise at each sector boundary. The stable orbits in an AVF cyclotron are not circles; the particles perform radial oscillations about the circular orbit. The maximum energy obtainable with an AVF accelerator is about the same as for synchrocyclotrons. An advantage of the AVF cyclotron is the larger possible beam current (of the order of $100\mu\text{A}$), which depends on the fact that not only one bunch per time is possible, but also many pulses at the same time. In this way it is possible to operate the cyclotron at a fixed frequency even up to relativistic energies.

In the *synchrotron* – also invented by McMillan and Veksler – the massive magnet is replaced by a ring of bending magnets. Electrons are injected at relativistic speed and as the electron energy increases, the magnetic field is also increased at a rate that makes the electron orbit the same at all energies. For protons and heavy particles, it is not so easy to inject at relativistic speed. Therefore both the RF field and the magnetic field are varied for protons and heavy particles, in such a way as to keep the orbital radius constant. The principle is shown in Fig. 1.12. Frank Goward and D.E. Barnes in England were the first to make a synchrotron work. They converted a betatron to an 8 MeV electron synchrotron in 1946. The first proton synchrotron was the 3 GeV Cosmotron in Brookhaven National Laboratory, completed in 1952. Its ring of magnets had a diameter of 21 m and a height of 2.4 m. Its injector was a 4 MV electrostatic accelerator from HVEC. In 1954, the Bevatron at Berkeley was completed. Its energy was 6 GeV, enough to demonstrate the existence of the antiproton. Energies up to 100 GeV can be obtained economically with a conventional synchrotron.

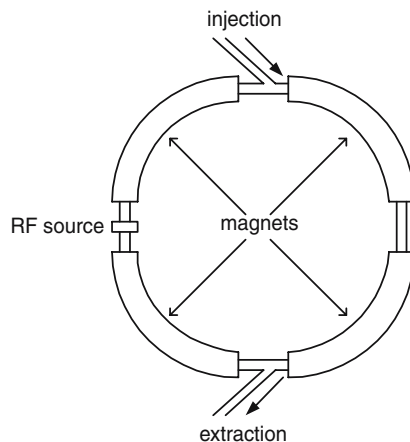


Fig. 1.12. Principle of a synchrotron

1.5 Alternating-Gradient Focusing

The fourth principle uses alternating gradients for magnetic focusing. The use of this principle dramatically reduced the size of the magnets for large accelerators, allowing a much larger energy to become economically achievable. It was Livingston who, at the end of the 1940s, considered the possibility of building a synchrotron with successive magnetic sectors facing inwards toward and outwards from the center in order to compensate for the effects of magnetic leakage. The field at the center of the beam tube has the same value in all sectors, but in one sector it decreases with r and in the neighboring sectors it increases. The variation of the field with r is quite dramatic: $B \sim r^n$ and $B \sim r^{-n}$, with $n \sim 300$, in alternate sectors. This means that in addition to bending the particle trajectories, the magnets have a strong lens effect. The principle is shown in Fig. 1.13. This revolutionized the approach to accelerator design and made it possible to arrive at a compact design; this could also be used for machines with much higher energies than could be envisaged before. The energy range for AG synchrotrons is 100 GeV to 1 TeV. Other than cost of the magnets and the size of the ring, there is no limit on the energy that can be obtained. For synchrotrons from around 1970, a major advance in design has been the separation of the bending and focusing, functions, so that dipole magnets bend the beam and quadrupole magnets do the focusing, as illustrated in Fig. 1.14.

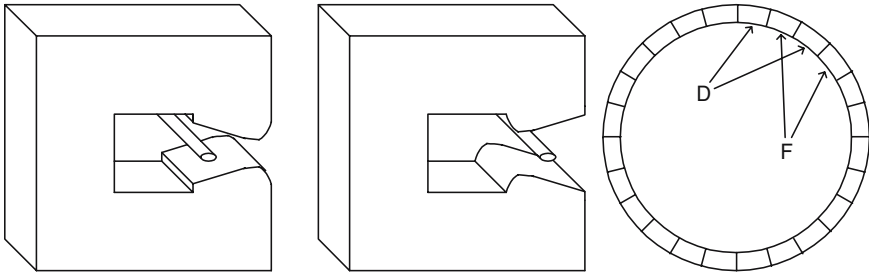


Fig. 1.13. Alternating-gradient (AG) synchrotron having sectors with focusing increasing with r (F) and sectors with focusing decreasing with r (D)

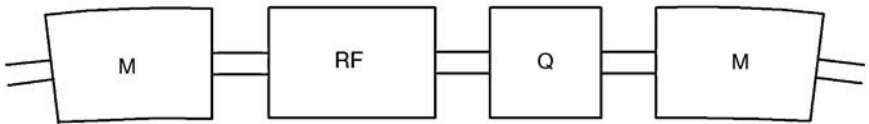


Fig. 1.14. Part of a separate-focusing synchrotron; M = dipole magnet; Q = quadrupole magnet, RF = accelerating gap

1.6 Colliding-Beam System

The next large energy step was taken by a colliding-beam system introduced in the late 1960s. In particle collisions, only the center-of-mass energy is useful. For fixed-target accelerators, this means that the main part of the particle energy will be wasted as kinetic energy of the colliding particles and their reaction products. In contrast, if two particles with the same momentum that move in opposite directions are made to collide, all the available energy can be made use of in the interaction. The first storage rings to operate were the 200 MeV electron ring at Frascati and the 2×500 MeV rings at Stanford. Both started in 1961. Another example of an early proton–proton colliding-beam accelerator is the CERN Intersecting Storage Rings (ISR), in operation during the period 1972–83, with two beams of 28 GeV from the CERN proton synchrotron directed into two interlacing storage rings with opposite directions for the magnetic fields, making collisions at eight positions. The 56 GeV center-of-mass energy is equivalent to a beam energy of 1700 GeV against a fixed target. For the particle density available in a normal accelerator beam, however, the colliding-beam system is unrealistic. The invention of particle stacking changed the picture significantly. From a preaccelerator, a weak current is injected into a storage ring over a long period of time (of the order of a day). At the same time, the beam is focused to occupy a far smaller area than it did upon leaving the preaccelerator. In this way, the particle density is considerably increased, and a circulating current equivalent to several amperes can be obtained. To have beams circulating for long times, an extremely low pressure is needed so as not to lose too much of the intensity by scattering from residual molecules in the vacuum. Storage rings are often equipped with an RF acceleration gap for adjustment of the energy.

In *electron storage rings*, the synchrotron radiation can make the beam shrink in all dimensions, and thus intense electron beams can be accumulated. In *proton storage rings*, the synchrotron radiation is negligible and the required intensity must be obtained by stacking.

The ISR mentioned above was the highest-energy machine in the world until the SPS at CERN started operating as a proton–antiproton collider in 1981 at 2×270 GeV. Antiprotons were produced in a fixed target by irradiation with 26 GeV protons. The beam of $\bar{p}$ is of very low intensity, and the beam also has an extremely low quality as the $\bar{p}$ are produced with a spread in direction as well as in energy. However, one of the exciting developments that originated within the ISR project was “cooling” of the beam. This is a method to reduce the beam dimensions (the phase space) and energy spread, and thus increase the beam density. Two methods for improvement of beam quality were developed during the 1970s, namely electron cooling for proton energies less than 1 GeV and stochastic cooling for proton energies above 1 GeV. Cooling is particularly effective for weak beams. Carlo Rubbia and Simon Van der Meer at CERN cooled antiprotons to dimensions and

intensities comparable to those of a proton beam and accumulated them over long periods. Then they accelerated them to about 300 GeV in the SPS to make them collide with protons of the same energy. This became a success, and the particles W and Z that mediate the weak interaction were identified in 1982–83. As a result Rubbia and Van der Meer shared the Nobel Prize in 1984; see Table 1.1. Another major discovery made by a colliding-beam accelerator was the top quark, at the Tevatron collider at the Fermi Laboratory using 900 GeV protons and 900 GeV antiprotons [13]. Very large colliders are extremely expensive to build as a double-ring collider. Since a particle and its antiparticle have identical masses but opposite signs of electric charge, they rotate in opposite directions in the same magnetic field. In such a case it is possible to use the less expensive single-ring collider concept.

By the use of superconducting magnets with excitation windings operating at the temperature of liquid helium, the mass of the magnets can be radically reduced by at least one order of magnitude. The reduction in the dimensions, weight, cost and supply power provided by superconducting magnets is very attractive, and a much higher magnetic field is available. In this way, the accelerator will be more compact and cheaper.

The present state of the art for proton accelerators is the Large Hadron Collider (LHC) under construction at CERN (Fig. 1.15). It will consist of

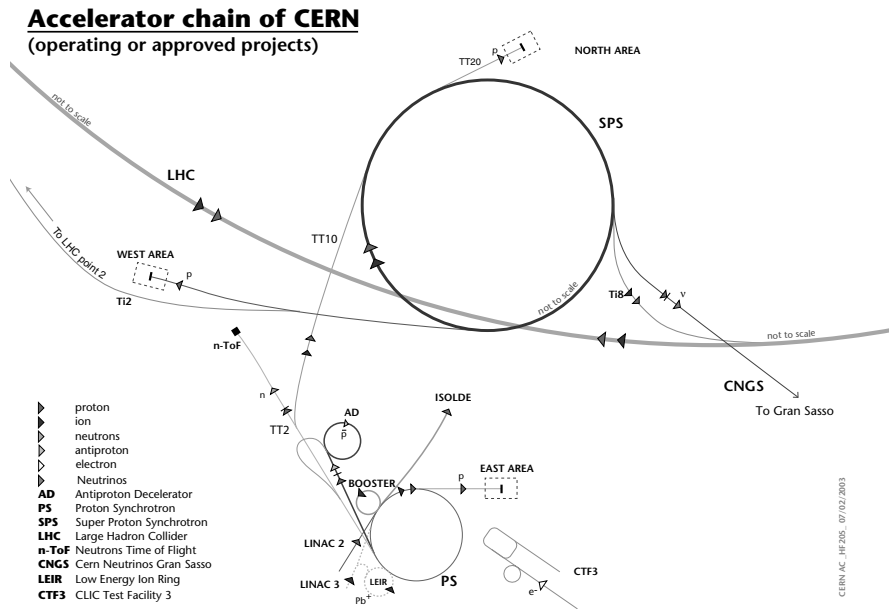


Fig. 1.15. The CERN complex, consisting of various preaccelerators, smaller accelerators, the SPS with a 2.2 km circumference, and the LHC with a 27 km circumference (only part of the LHC is seen in the figure)

more than 1200 superconducting magnets. Two proton beams of 7 TeV each will circulate in the 27 km long circle. The available energy in the LHC is the center-of-mass energy, $2 \times 7 \text{ TeV} = 14 \text{ TeV}$. An equivalent fixed-target accelerator would need to have a beam energy of $E_p = 2 \times 7000^2 \text{ GeV} = 98\,000 \text{ TeV}$, i.e. 14 000 times higher than the 7 TeV of the LHC! The effective constituent collision energy for a hadron collider – i.e. the energy available per quark, which is of most interest as one is looking, primarily, for quark–quark collisions – is 1/6 of the sum of the beam energies, i.e. 1/6 of 14 TeV, or just above 2 TeV.

The development of accelerator technology has been spectacular. However, there is much more to do. The accelerators now in use and under construction are expected to lead the way beyond the present incomplete standard model. They should unearth new classes of particles and enhance our understanding of the asymmetry between matter and antimatter and of the transition to the primordial quark–gluon plasma. With the accelerators of today using RF technology we have come to a limit; they are too big and too expensive. Could it be possible to find a completely different technology, reducing their size and cost? In various articles that have appeared during the last two decades and more, preliminary experiments have been discussed and reported employing different possible approaches. Examples of overview articles are to be found in [14–17]. One approach is to use collective fields in oscillating plasmas. The plasma is excited by lasers or by a driver beam; the accelerating gradients and focusing strengths will be orders of magnitude greater than those achieved thus far by RF accelerators, offering the possibility of smaller, lower-cost accelerators at very high energies in the future. The greater the accelerating gradient, the shorter would be the accelerator to obtain a given energy.

References

1. E. Rutherford: London, Edinburgh and Dublin Philos. Mag. J. Sci., 6th series **37**, 581 (1919)
2. J.D. Cockcroft, E.T.S. Walton: Proc. Roy. Soc. A **136**, 619 (1932) and **137**, 229 (1932)
3. G. Trump, R.J. Van de Graaff: J. Appl. Phys. **8**, 602 (1937)
4. R.J. Van de Graaff: Phys. Rev. **37**, 1919 (1931)
5. D.A. Bromley: Nucl. Instr. Meth. **122**, 1 (1974)
6. G. Ising: Arkiv Mat. Astr. o Fys. **18**, 30 (1924)
7. R. Widerøe: Arch. Electrotechn. **21**, 387 (1928)
8. L.W. Alvarez: Phys. Rev. **70**, 799 (1946)
9. E.O. Lawrence, M.S. Livingston: Phys. Rev. **40**, 19 (1932)
10. D.W. Kerst: Phys. Rev. **60**, 47 (1941)
11. E.M. McMillan: Phys. Rev. **68**, 143 (1945)
12. V. Veksler: J. Phys. (USSR) **9**, 153 (1945)

13. R. Ladbury: Phys. Today June p. 17 (1994)
14. J.S. Wurtele: Phys. Today July p. 33 (1994)
15. M. Tigner: Phys. Today January p. 36 (2001)
16. C. Joshi, T. Katsouleas: Phys. Today June p. 47 (2003)
17. V. Malka: Europhys. News **35**:2, 43 (2004)

2 Accelerators for Medicine

R. Hellborg¹ and S. Mattsson²

¹ Department of Physics, Lund University, Sölvegatan 14, 223 62 Lund, Sweden
`ragnar.hellborg@nuclear.lu.se`

² Department of Radiation Physics, Lund University, Malmö University Hospital,
205 02 Malmö, Sweden
`soren.mattsson@rfa.mas.lu.se`

2.1 Introduction

A few months after Röntgen's discovery of X-rays in the autumn of 1895, this new type of radiation was already being made use of in attempts to treat malignant tumors in patients. Starting during the 1910s, radium was also used for cancer treatment, both in “brachytherapy” with the sources near (“brachy”) the tumor and with the radium sources outside the body (teletherapy). The main gamma rays from ^{226}Ra range between 0.24 and 2.20 MeV, i.e. they were of higher energy than the X-rays available and therefore penetrated deeper into the body of the patient. At the beginning of the 1950s, the use of radioactive ^{60}Co sources for therapy began. ^{60}Co emits gamma rays with energies of 1.17 and 1.33 MeV.

As is mentioned in Chap. 1, a few years after Robert Van de Graaff's first demonstration of the electrostatic accelerator, such a machine had already been installed at Harvard Medical School in Boston, USA [1]. This first medical machine, designed by John Trump and Robert Van de Graaff, was an open-air accelerator; a sketch is shown in Fig. 2.1. The second hospital electrostatic machine came in 1940 and was pressure-insulated. During the 1950s, around 40 electrostatic accelerators of 2–3 MV accelerating potential were delivered to different hospitals.

Starting in 1936, Lawrence used one of his cyclotrons to accelerate deuterons to 8 MeV and in this way provided most of the world's supply of artificial radioactive isotopes at that time. The accelerator was also a good neutron source, and in 1938 the first cancer patient had already been treated with neutrons obtained by use of a (d, n) reaction on a Be target from one of Lawrence's cyclotrons. A photo of his 60 inch cyclotron is shown in Fig. 2.2. Presently, more than 15 000 patients have been treated with neutrons, mainly thermal neutrons, at various places in the world. Today there are, however, only a few indications left for neutron therapy, such as radioresistant and slowly growing tumors. Experience shows that fast neutrons should be used in these cases. As there has not been any great breakthrough in neutron therapy, the number of medical neutron facilities is now decreasing drastically.

At the end of the 1940s, the newly introduced betatron was used for the first time for radiation therapy, with X-rays generated by a beam of 20 MeV

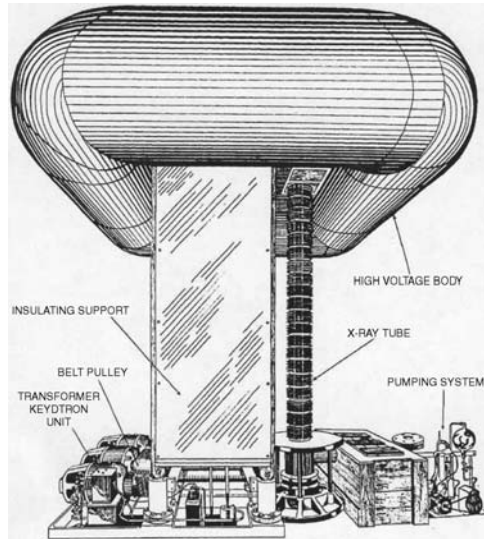


Fig. 2.1. The Trump-Van de Graaff X-ray generator installed at the Harvard Medical School (Reprinted from [1]; copyright 1937, with permission from APS)



Fig. 2.2. The newly completed 60 inch cyclotron. Ernest Lawrence is second from *left* on the floor, and Luis Alvarez and Edvin McMillan are on *top* of the machine (Reprinted with permission from LBL)

electrons. Betatrons played a significant role for several years, since they delivered X-ray beams (up to over 40 MeV) with better properties than those obtained from X-ray tubes and radionuclide sources and since they could also be used for electron beam therapy. At most, about 200 betatrons were in use in hospitals at the beginning of the 1970s. The major disadvantages of betatrons were their relatively low absorbed dose rate, small treatment field size, high weight and cost.

In the early 1950s, a few RF linear accelerators were installed at hospitals. In the 1960s the number increased considerably, and soon the linac became, and still is, the dominant type of hospital-based accelerator for radiotherapy. In Fig. 2.3, a modern linac for patient treatment is shown. The rapid growth and dominant role of linacs depend on the fact that they can deliver a ten times higher dose rate compared with a betatron (several Gy/min, compared with half a Gy/min for a betatron), with a geometrical field size up to $0.4 \times 0.4 \text{ m}^2$, compared with $0.1 \times 0.1 \text{ m}^2$ for the betatron. Today (2004) around 5 000 medical linacs around the world are used for treating several millions of patients yearly.



Fig. 2.3. A linac for radiotherapy equipped with a multileaf collimator. The electron beam is bent in a 270° magnet. After flattening and collimation, the electron beam can be used for treatment. In most cases, however, the electron beam collides with a heavy-metal target, and the high-energy bremsstrahlung produced is used for treatment after flattening and collimation with the multileaf collimator (Reprinted with permission from Varian Medical Systems Inc.)

The first proton beam used for treatment was obtained from a cyclotron in Berkeley, USA, and shortly after that from the synchrocyclotron in Uppsala, Sweden, at the end of the 1950s. Today there are a number of proton facilities around the world, and up to now (December 2004) 40 000 patients have been

treated. Larger use of proton therapy has been limited by the high cost of a dedicated hospital-based proton therapy unit. The cost per treatment fraction is, however, only around 70% higher than the cost of conventional photon and electron therapy [3]. At the beginning of the 1990s, the first superconducting cyclotron was taken into biomedical use. Also at the beginning of the 1990s, the first hospital-based proton synchrotron was used for therapy. Starting in 1975, in Berkeley, ions heavier than protons have also been tested for radiotherapy, with the intention of obtaining more efficient therapy for some patient groups.

Most medical accelerators are used for radiotherapy. Accelerators are also used for production of radionuclides, in the radiopharmaceutical industry or locally at hospitals, and for sterilization of medical equipment. A useful book about biomedical accelerators is found in [2].

2.2 Radiation Therapy

Radiotherapy is, after surgery, the most widely used form of cancer treatment, in Europe and the USA being given to about half of all cancer patients. There is nothing to suggest that other methods for treating cancer can replace radiotherapy in the foreseeable future. The possibility of selectively destroying tumor cells by radiation in the presence of normal cells depends on the fact that tumor cells are more sensitive to radiation than normal cells, and that the repair of malignant cells is less efficient. The difference in the effect on tumor and normal cells is – for photon and electron radiation – increased if the radiation is fractionated in time, for external-beam therapy normally into 2 Gy fractions given once a day, five days per week over a number of weeks. The objective of radiation therapy is to deliver a defined absorbed dose of radiation to a specific tissue volume – including the tumor volume and adjacent tissues where tumor cells might be found – with the intent of killing tumor cells while minimizing irradiation of surrounding, healthy tissue.

External-beam therapy is the most common form of radiotherapy. Externally produced photon beams, or X-rays, are used in more than 80% of all radiation treatments. In addition, electron beam therapy is used in 10–15% of cases. The rest are treated by brachytherapy (either intracavitary or interstitial – with the source either in a cavity or in the tumor itself), external proton therapy, therapy with radiopharmaceuticals or a limited number of other methods.

2.2.1 Electron and Photon Beams

The electron and photon (X-ray) beams are today produced by linacs, which normally offer the possibility to produce two photon energies as well as several electron energies. Typical values are 6 MV and 10, 15 or 18 MV for X-rays, and six electron energies between 6 MeV and 20 or 22 MeV.

In the linac, the accelerated electrons collide with a heavy-metal target. As a result of these collisions, high-energy X-rays are produced in the target. The electron beam itself can also be used for radiotherapy directly, after flattening. The therapy beams are shaped to match the patient's tumor. The beam comes out of a gantry, which can rotate around the patient. The patient lies on a movable treatment couch, and laser beams are used to make sure the patient is in the proper position. Radiation can be delivered to the tumor from any angle by rotating the gantry and moving the treatment couch.

The process of external-beam therapy can be divided into various parts:

- imaging as a basis for treatment
- treatment planning
- simulation of the treatment
- treatment delivery
- follow-up of treatment.

First, the tissue volume to be irradiated and the tissues to be protected must be defined. The imaging needed is done in addition to the diagnostic investigations done in connection with the patient's earlier investigations. Up to now, ordinary radiography and X-ray computed tomography (CT) have been the basis for tumor volume delineation, but positron emission tomography (PET) combined with CT (PET/CT), and magnetic resonance imaging (MRI) are more and more used. The tumor delineation is followed by treatment planning, whereby the direction and shape of the radiation beams are configured to achieve a dose distribution that corresponds as well as possible to that desired. The planning also involves the choice of the appropriate absorbed dose in the respective target volumes, the number of fractions and absorbed dose per fraction, and the total treatment time.

For treatment planning, computers are used to calculate the absorbed dose distribution that will be delivered to the patient's tumor and the surrounding normal tissue. In certain cases, this process may employ such techniques as three-dimensional conformal therapy and intensity-modulated radiation therapy.

During simulation of the treatment, the patient is placed in the treatment position on a special X-ray machine or CT scanner, and simulation X-rays are taken. X-ray images are taken in the directions of the treatment beams, including markings to indicate the size and shape of the field planned for therapy. Fixation or other devices are used to help the patient not to move during the simulation and treatment processes. The beams needed to treat the patient are tested, and small marks on the patients are made to guide the daily treatments. After possible adjustment, the parameters are transferred to the treatment unit. A schematic description of the process leading up to the treatment is shown in Fig. 2.4.

After the simulation, the treatment itself can begin. In the treatment, the patient is placed on the table top of the accelerator, exactly in the same way as in the simulator. Similarly adjusted laser light beams to those used

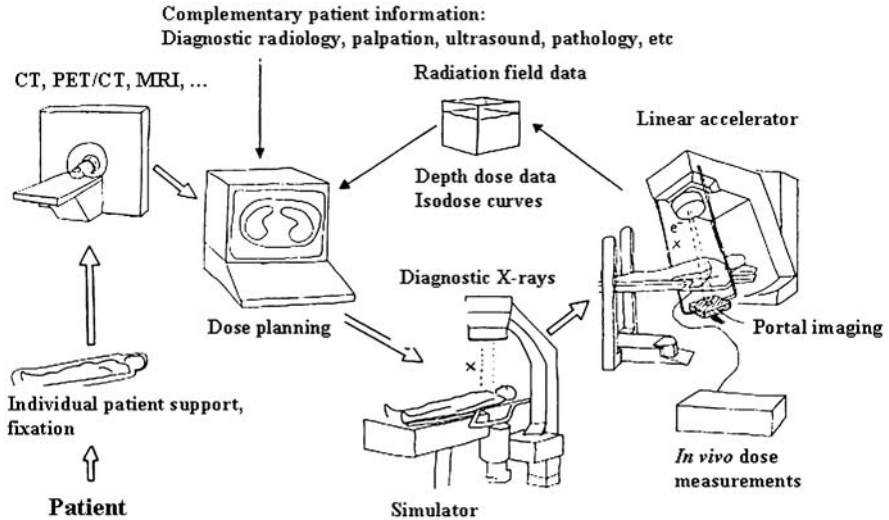


Fig. 2.4. The process of imaging, planning, simulation and treatment

at the simulator and at the diagnostic modalities used are projected onto the patient to determine the position coordinates, and a light field on the patient simulates the radiation field. The setup is checked by imaging the initial therapy radiation transmitted through the patient. New portal imaging devices make image-guided radiotherapy possible. In Fig. 2.5, the preparation just before the start of treatment is shown.

Beams from one or more directions may be used, and the beam may be on for as long as several minutes for each field. The treatment process can take from 10 to 30 min, and most of the time is often spent positioning the patient. Patients usually receive radiation treatments once a day, five days a week, for a total time ranging from two to nine weeks.

The past decade has seen a succession of advances in radiation therapy. New techniques of diagnosis are improving the possibilities for tumor delineation. The possibilities for rapid and adequate 3D treatment planning have been refined. New techniques of external radiation treatment are making possible intensity-modulated radiation beams, confining the high therapeutic dose to the target tissues. These things taken together imply increased possibilities for augmenting the effect on the tumor while at the same time reducing the risk of adverse effects in normal tissue.

The multileaf collimator (MLC) – seen in Fig. 2.3 – offers the possibility to shape the beam in conformity with the target outline (conformal therapy). Recently, intensity-modulated radiotherapy (IMRT) has been developed, giving further improved conformity.

The patient's position has to be reproduced as precisely as possible for each treatment session. Delivery of escalated absorbed doses by IMRT



Fig. 2.5. Preparation before start of treatment by means of a linear accelerator. The patient is placed on the treatment table top and a reproducible positioning is supported by individual casts, which are used during the whole treatment series. Field markers from the simulator are checked against beam indicators. A diode dosimeter for entrance dose measurements is applied. The portal dose-imaging system is not shown. Photo from Malmö University Hospital, Sweden

necessitates knowing the exact tumor position prior to radiation treatment delivery. Tumors located in the chest, abdomen or pelvis can shift their position from day to day over the course of treatment owing to movements of internal organs and volume changes. Geometric precision better than 5 mm and absorbed-dose accuracy of better than 5% are necessary.

2.2.2 Protons and Light Ions

It was R. Wilson who first realized the medical potential of protons and carbon ions for therapy in 1946. Proton irradiation offers better dose distribution than do conventional photon and electron beams. An excellent depth dose can be reached with energy- and intensity-modulated beams from commercially available equipment. Light-ion beams have been used for treatment in two centers in Japan starting in the beginning of the 1990s and in an experimental facility at GSI in Darmstadt starting in 1997. One clinical unit is

under construction in Heidelberg. The cost of light-ion therapy is 2.5–3 times as high as that of proton beam therapy and the clinical rationale for more general use is not yet convincing [3].

2.2.3 Synchrotron Radiation

Synchrotron radiation is used today for research in biology, chemistry, physics and their numerous subfields. As is outlined in Chap. 3, the technique has had a very high impact on research in these fields. Practical applications in medicine seem, however, to be quite a way off.

The interest in synchrotron radiation for medicine is that it can be produced with such a high photon fluence rate that even after an energy selection, the fluence rate is so high that the source can be used for radiography. Another important parameter is that very narrow and parallel beams can be produced.

Several experimental approaches are being explored using synchrotron radiation to enhance the effectiveness of radiation therapy. One is the microbeam therapy technique. It is based on irradiating a tumor with multiple parallel, microscopically narrow, planar beams in the 50–150 keV range. Typically, the width of a microbeam is 25 μm and the distance between the parallel beams 200 μm . In this way, tissue necrosis on the way in to the tumor volume could be prevented. The effects in the target volume can be achieved through cross-firing.

Another approach is photon activation therapy, where a sufficiently high concentration of iodine (or other heavy element) from, for example, an X-ray contrast agent in the tumor creates Auger electrons and photoelectrons close to the tumor cells. The energy of the synchrotron radiation is chosen slightly above the K-absorption edge of the heavy element.

2.3 Production of Radiopharmaceuticals for Medical Imaging

The first radionuclide to be used for nuclear medicine was produced by a cyclotron in 1936. A few years later the first nuclear reactor was demonstrated, and radionuclides could be obtained from reactors as well. Unfortunately, only radionuclides that have an excess of neutrons can be obtained from reactors. On the other hand, all types of nuclides can be obtained from an accelerator. Another advantage of an accelerator compared with a reactor is that the compactness of the accelerator make it possible to have it installed at the hospital. Today more than 200 small compact cyclotrons with proton, deuteron and alpha energies up to 20–40 MeV and beam currents of the order of mA are used for radionuclide production. The activity which can be produced is limited, chemical separation is necessary and the production cost is

relatively high. There are also a number of requirements on the material to be irradiated: thermal conductivity, resistance to overheating and an ability to withstand vacuum. The required chemical amount of material for one investigation is very low (of the order of 10^{-7} g), and much less than that required for chemical synthesis. The production of radionuclides is today dominated by (isochronous) cyclotrons. It is demonstrated in Chap. 19 that dedicated electrostatic accelerators can also play an important role today in the production of short-lived nuclides. If a radioactive material is administered to a patient, the emitted radiation can be detected outside the body (if the energy is high enough) in different directions. Compared with diagnosis with X-rays, this nuclear medical technique is not used so often, even if around 20 million investigations are done per year in the world.

Positron emission tomography (PET) is a powerful diagnostic tool in modern medical imaging. It uses short-lived radionuclides such as ^{18}F (physical half-life 110 min), ^{11}C (20 min), ^{13}N (10 min) and ^{15}O (2 min), which can be produced by low-energy cyclotrons (< 20 MeV) accelerating protons and/or deuterons.

Most cyclotrons either are placed at a hospital or are part of a commercial company. The hospital is interested in a continuous supply of ^{18}F -FDG to use for tumor scintigraphy and delineation. A few large research machines are also employed in nuclide production, such as isochronous cyclotrons, ion RF linacs and tandem accelerators. As an example of the development at a research accelerator, the production of ^{123}I , with half-life 13.2 h, by a large research linac has initiated other large facilities to start a similar production. The ^{123}I competes favorably with the ^{131}I used earlier, with half-life 8.05 d, as the dose to the patient decreases by a factor of 50 to 100 when ^{131}I is replaced with ^{123}I .

2.4 Analytical Applications

The uses of MeV ion beams from electrostatic accelerators for studies of chemical composition, atomic structure, and surface or near-surface layers are analytical techniques which have developed rapidly during the last ten to twenty years. These techniques have mostly been used for studies of materials and solid state physics. Only during the last 10 years have these techniques also been enlarged to include medical investigations. The relative sensitivity of some of these methods is 10^{-6} or even less, and the depth resolution of some of them is down to tens of nm. The analysis time per sample is of the order of tens of minutes and the amount of material necessary is often less than 1 mg. The most well-known techniques are PIXE (particle-induced X-ray emission), RBS (Rutherford backscattering) and NRA (nuclear reaction analysis). These techniques are described in detail in Chaps. 24, 25 and 26. By using a beam with a diameter down to $1\text{ }\mu\text{m}$, microscopy methods can be

applied with these methods. In biomedical applications, this makes it possible to analyze tissue, subcellular and cellular structures.

Accelerator mass spectrometry (AMS) is an ultrasensitive technique which has found its main application in the quantification of very rare long-lived radionuclides. For details about AMS, see Chap. 23. The most well-known example is ^{14}C ($10^{-10}\%$ of the carbon in living organisms is ^{14}C). Unlike a detector for the measurement of radioactivity, a mass spectrometer does not have to wait for the nuclide to decay. Thus, in the case of ^{14}C , AMS is about 1000 times more sensitive than any radiometric method (out of 10^9 ^{14}C atoms, only about 10 will decay during one hour). AMS has the outstanding ability to quantify ^{14}C -labeled substances down to levels of 10^{-18} moles, which is of the order of only one million molecules. The fact that AMS counts atoms and not decays results in some powerful advantages over radiometric techniques, such as highly reduced sample sizes and shortened measuring times. The potential of ^{14}C AMS for biomedical investigations is illustrated in Fig. 2.6, which shows the specific activity in exhaled air after intake of a test amount of ^{14}C -labeled fat [4, 5]. Using AMS, it is possible to follow the exhalation of increased amounts of ^{14}C for years. Using liquid scintillation counting, the exhalation could only be followed for some days. Thanks to AMS, a fine structure in the exhalation pattern due to provocation in the form of 32 h of fasting could also be clearly seen and the increased exhalation quantified.

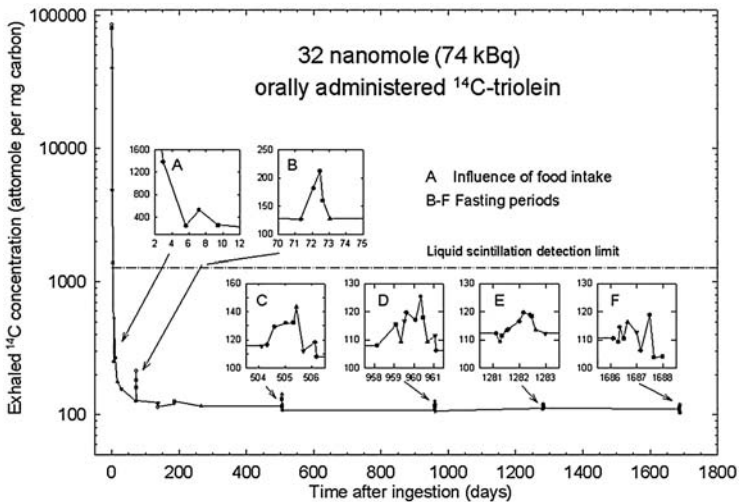


Fig. 2.6. The ^{14}C concentration in exhaled air from one volunteer at various times after oral administration of 32 nanomoles ^{14}C -triolein. The dip in the curve after 6 days (A) was the result of an excessive intake of food prior to sampling. This observation initiated 32 hour-long, controlled fasting periods (B, C, D, E and F). Data from [4, 5]

2.5 Imaging with Synchrotron Light

Over the last few years, several studies have shown the potential of synchrotron radiation for medical imaging. Mammographic images acquired with synchrotron radiation have been reported to provide a better contrast of breast lesions and a relevant reduction of the dose to the breast tissue. In angiography using X-ray contrast, one image is taken with the energy just over the absorption edge of iodine and another image with the energy just below the absorption edge. After subtraction of the images, the vessels are clearly seen. Studies have shown that the same image quality can be produced for half the dose compared with conventional techniques.

The technique can of course never be an alternative to conventional X-ray imaging, owing to the extremely high cost of a synchrotron light source.

2.6 Industrial Applications – Sterilization and Disinfection

In 1896, it had already been found that microorganisms died when irradiated with X-rays. Unfortunately, it took more than 50 years before a radiation facility was put into operation during the 1950s for sterilization of surgical threads. Electrostatic accelerators and linacs, as well as facilities equipped with gamma sources, were used. Syringes, needles, catheters and infusion sets are examples of disposables suitable for radiation sterilization. To bring down the production cost, the disposables are usually manufactured from a plastic which cannot be sterilized by heat, as they do not tolerate the necessary temperature: steam at 150°C or hot air at 200°C. The competitive methods are, of course, various chemical sterilization techniques. A major advantage of radiation sterilization is that the production process and packaging can be performed under nonsterile conditions. The final products in their cartons can be sterilized either by the manufacturer itself or by a separate company. Sterilization is today mainly done by electron beams from accelerators. High-energy bremsstrahlung X-rays from these accelerators can also be used, as well as ^{60}Co gamma sources (up to a few times 10^{16} Bq). Economically, accelerators seem to compete well with gamma sources, as the dose rate and therefore the throughput are considerably higher for an accelerator.

Typically, a dose of 20–30 kGy is used. This is obtained from an accelerator in a fraction of a second, whereas a ^{60}Co source needs to be used for hours. Electron accelerators for industrial use, including sterilization of medical disposables, food sterilization and material treatment (improvement of properties of polymer materials), are discussed in detail in Chap. 28.

References

1. J.G. Trump, R.J. Van de Graaff: J. Appl. Phys. **8**, 602 (1937)
2. W.H. Scharf: *Biomedical Particle Accelerators* (American Institute of Physics, New York 1994)
3. Swedish Proton Therapy Center: Report from an evaluation of a national proton therapy centre for cancer patients (in Swedish, with an English summary) (2003)
4. K. Stenström, S. Leide-Svegborn, B. Erlandsson, R. Hellborg, S. Mattsson, L.-E. Nilsson, B. Nosslin, G. Skog, A. Wiebert: Appl. Radiat. Isot. **47**, 417 (1996)
5. M. Gunnarsson, S. Mattsson, K. Stenström, S. Leide-Svegborn, B. Erlandsson, M. Faarinen, R. Hellborg, M. Kiisk, L.-E. Nilsson, B. Nosslin, P. Persson, G. Skog, M. Åberg: Nucl. Instr. Meth. B **172**, 939 (2000)

3 Accelerators for Synchrotron Radiation

M. Eriksson¹ and S. Sorensen²

¹ MAX-Lab, Lund University, Box 118, 221 00 Lund, Sweden
`mikael.eriksson@maxlab.lu.se`

² Department of Physics, Lund University, Box 118, 221 00 Lund, Sweden
`stacey.sorensen@sljus.lu.se`

3.1 Introduction

The explosion in synchrotron-based research over the last two decades can be attributed to dramatic advances in accelerator technology, but these advances have been to a large degree driven by the demands of an expanding, highly diverse user community. This intense light source covers most of the electromagnetic spectrum. The polarization of the light and the pulsed time structure are being utilized more and more as more complex studies are carried out. These machines have developed from “parasitic” operation during colliding-beam experiments to dedicated machines (TESLA in Hamburg, Germany) where highly coherent pulses of X-rays can be produced.

3.1.1 Dipole Radiation

According to Liénard [1] and Wiechert [2], the electric field originating from a moving charge (see Fig. 3.1) can be written as

$$\mathbf{E}(x, t) = \frac{e}{4\pi\epsilon_0} \left[\frac{\mathbf{n} - \boldsymbol{\beta}}{\gamma^2(1 - \boldsymbol{\beta} \cdot \mathbf{n})^3 R^2} + \frac{1}{c} \frac{\mathbf{n} \times [(\mathbf{n} - \boldsymbol{\beta}) \times \dot{\boldsymbol{\beta}}]}{(1 - \boldsymbol{\beta} \cdot \mathbf{n})^3 R} \right]_{ret}. \quad (3.1)$$

The first term describes the near (velocity) field, which we shall not treat here, while the second term describes the far (acceleration) field (the notation used in the equations are defined in Table 3.1). The latter term refers to the synchrotron radiation (SR) produced when the acceleration is perpendicular

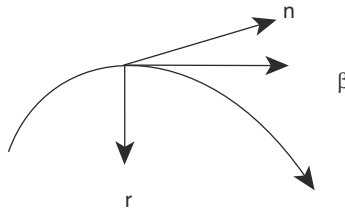


Fig. 3.1. A schematic diagram indicating the relevant vectors for synchrotron radiation produced in a dipole magnet

Table 3.1. Notation and definitions used in the equations

Parameter	Definition
β	Particle velocity in units of c , the speed of light
R	Distance from point where light is emitted to observer
ϵ_0	Permittivity of free space
γ	Lorentz factor
ρ	Bending radius of an electron moving in a magnetic field
$\mathbf{n}$	Unit vector from particle to observer
λ_u	Undulator period length
θ	Angle between particle velocity and radiation direction
$K_u = eB\lambda_u/(2\pi m_e c)$	Undulator parameter
B	Undulator peak magnetic field
N	Number of poles in insertion device
ω_c	Critical frequency for bending-magnet radiation
e	Electron charge
m_e	Electron mass

(when the charge is deflected in a bending magnet) to the velocity β . The far-field radiation has the following characteristics for relativistic particles:

1. The electric vector is in the plane of acceleration.
2. The total power emitted is given by

$$P = \frac{2e^2 c \beta^4 \gamma^4}{3\rho^2} . \quad (3.2)$$

3. The power is highly focused within an angle

$$\frac{1}{\gamma} . \quad (3.3)$$

4. The radiation is emitted in a *broad spectral range* up to the critical frequency

$$\omega_c = \frac{2c\gamma^3}{3\rho} . \quad (3.4)$$

3.1.2 Undulator Radiation

When a relativistic electron passes through a periodic magnet device (period length λ_u), known as an undulator (see Fig. 3.2), we can see from (3.1) that an electromagnetic (EM) wave with the following properties will be generated.

The wavelength of the EM wave is given by

$$\lambda_i = \frac{\lambda_u}{2\gamma^2 i} \left(1 + \frac{K^2}{2} + \frac{\theta^2 \gamma^2}{2} \right) . \quad (3.5)$$

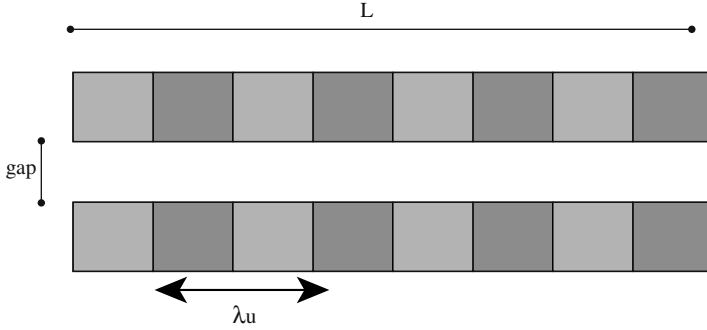


Fig. 3.2. A schematic diagram of an insertion device. The period of the device, λ_u , is indicated

The wavelengths produced are described by $\lambda_i N_u$. The bandwidth of this wave is Fourier transform limited to $1/(nN)$, where n is the harmonic number and N is the number of poles in the magnetic undulator array. Modern storage rings implement insertion devices as the primary source of radiation. See Fig. 3.3 for a diagram comparing the brightness of a bending-magnet

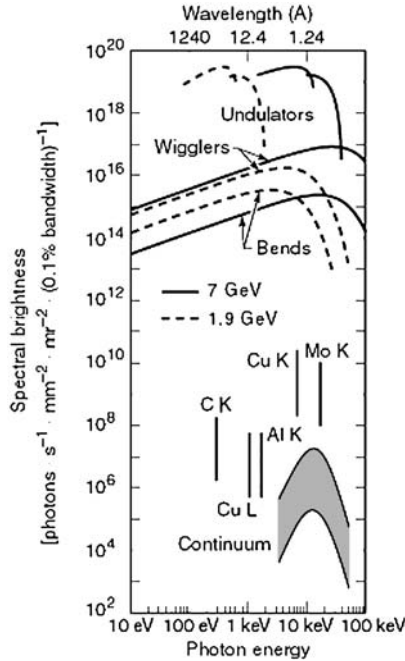


Fig. 3.3. A comparison of the brightness of undulators, wigglers and bending-magnet sources. Metal-anode sources are also indicated in the plot

source with two insertion device sources. It should be mentioned that the full exploitation of the characteristics of undulator radiation puts stringent demands on the electron beam quality in terms of emittance and energy spread in order to preserve the narrowest natural line width. The measures of photon intensity have developed from flux (photons/s) to brightness (photon flux per unit source area per unit solid angle) and brilliance (photon flux per unit phase space volume), where the source size and the small opening angle of undulator radiation are taken into account.

Another variant of the magnetic insertion device is the “wiggler”. The wiggler spectrum is the incoherent superposition of the spectra from N poles, and is generally used as a wavelength shifter which can produce a dipole-like spectrum at a shorter wavelength.

3.1.3 Coherent Radiation

When (3.1) is integrated over all space, it can be seen that for wavelengths larger than or similar to the bunch dimensions, the EM waves add in phase and the power emitted is proportional to the number of particles squared. We now have a coherent radiation source. This is a very important characteristic of undulator radiation and is currently the focus of new developments in machine design.

3.1.4 Synchrotron Radiation Facilities

These radiation characteristics have proved to be quite useful for the study of matter; we have now some 60 dedicated synchrotron radiation facilities in operation around the world, and new ones are being planned and built. The first generation of synchrotron radiation research took place at existing high-energy physics accelerators such as DORIS (Germany) and SPEAR (US). These are called the “first generation” of synchrotron light sources. The second-generation sources were specially built machines with small, intense electron beams dedicated to SR research: SRRF (UK), MAX I (Sweden), and BNL (US), to name a few. The introduction of special insertion devices, undulators and wigglers, increased the brilliance by several orders of magnitude. The third-generation sources take full advantage of undulator radiation. Some examples of X-ray sources are ESRF (France), APS (US), and Spring8 (Japan). Some examples of soft-X-ray sources are ALS (US), MAX II (Sweden), Elettra (Italy), SLS (Switzerland), and BESSY (Germany). The fourth generation of light sources could very well produce mainly coherent light. The two X-ray coherent sources being designed today are XFEL (Germany) and LCLS (US). The performance development of synchrotron radiation sources is demonstrated in Fig. 3.4.

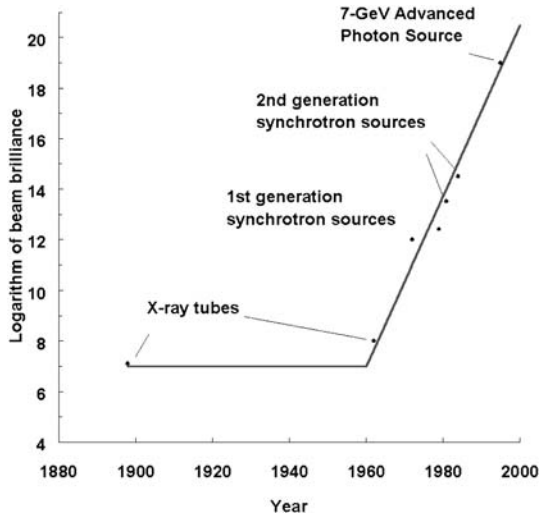


Fig. 3.4. Development of light source performance

3.2 Instrumentation for Synchrotron Radiation Experiments

Although some early experiments used the entire continuous spectrum of light from relativistic electrons more or less directly in, for example, absorption spectroscopy, nearly all modern experiments rely on specialized optical systems to extract the radiation to the experiment in an efficient way. The optical scheme can be designed for a particular purpose, such as high flux throughput, polarization, or high spectral resolution. Microscopy experiments require a highly focused beam, while lithography demands a stable beam of parallel X-rays for accurate imaging. In addition, most experiments require a single, narrow-bandwidth (BW) photon energy, which should be tunable over a wide range of energies. A number of different optical systems have been developed in response to the varied needs of the user community. All of the properties which are important for experiments can be optimized by designing optical systems together with the design of the ring lattice. Achieving the highest performance at the experiment depends upon having the correct emittance, polarization, stability, and pulse structure in the electron beam. These *monochromators* should be able to operate in a laboratory environment where conditions are constantly changing, and they should work with a variety of different kinds of experiments.

A short overview of some common monochromators will be given here. The long-wavelength regions (*visible* and *infrared*) have traditionally been the domain of lasers, but recently synchrotron radiation has proven to be a useful tool even in the infrared region. In the *ultraviolet* region, an unprecedented photon-energy resolution of 150 000 has been obtained in the

10–30 eV region, and new variable-polarization undulators, where the polarization vector at the sample can be rotated by tuning undulator parameters, open up new possibilities for experiments. Dramatic improvements in insertion devices and beam emittance have pushed the performance of *soft-X-ray* monochromators to unprecedented resolution; finally, it has been possible to resolve vibrational levels in molecules even for inner shells. A photon BW substantially smaller than the lifetime broadening of inner-shell levels in light elements (40 MeV BW at 500 eV) is routinely available at many SR laboratories. More information on soft X-ray monochromators can be found in [12].

The quality of the light at the sample depends on all of the optical elements between the source and the experiment. The primary mirror generally images the source at a point several meters from the storage ring. The mirror may also magnify the source, and it functions as a coarse filter owing to the reflection properties of the mirror surface coating. Generally, the mirror has a high-energy cutoff determined by the angle of incidence and the surface material. The mirrors may be subjected to enormous thermal loads, which lead both to deformation of the mirror surface and to photocatalysis, resulting in carbon buildup on the mirror. The former may be handled by cooling schemes, while the latter is more difficult to remedy. Keeping in mind that the storage ring and the optics associated with each beam line are installed in an interconnected vacuum system (see Fig. 3.5), all operations involving movements of mirrors or other optical elements must be designed for ultra-high vacuum (better than 10^{-7} Pa). The vacuum constraints require reliable security systems to guard against accidents in any part of the experiment affecting the operation of the monochromator or storage ring.

The task of the monochromator is to extract a single wavelength from the continuous spectrum after the primary optics. The task is accomplished by a series of mirrors and a dispersive element such as a diffraction grating, which focus the dispersed radiation on an exit slit. See Fig. 3.6 for a schematic diagram of a monochromator for soft X-rays.

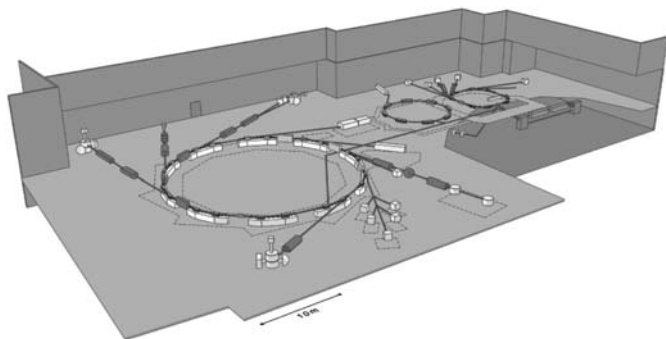


Fig. 3.5. A diagram showing the MAX electron storage rings. The *beam* lines are indicated in the figure

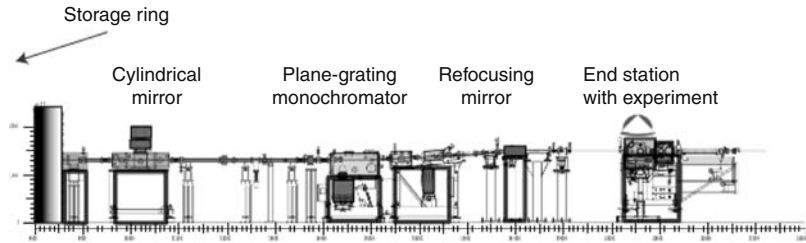


Fig. 3.6. A diagram of a beam line with a grazing-incidence plane-grating monochromator

The final stage in the beam line is the refocusing optics. The purpose of these optics is to make it possible to place an experiment at the final focus of the light. Several meters are needed after the monochromator exit for diagnostic instruments, differential pumping stages, and filter systems and for large experimental vacuum chambers.

The monochromators can be divided into several classes: the dispersive elements can be diffraction gratings, which can be ruled on a focusing mirror surface (spherical or toroidal grating), or crystal diffraction can be used to disperse X-rays. Some common monochromator designs are listed in Table 3.2.

Table 3.2. A list of common monochromator classes used at synchrotron radiation sources. A few typical parameters are given

Monochromator Type	Spectral Range (typical)	Energy Resolution	Energy at which Used (eV)	Applications
FTIR	Infrared	$1000\text{--}10\text{ cm}^{-1}$		
NIM	Ultraviolet	5–40 eV	10^5	Chemical dynamics, valence band
SGM, PGM	Soft X-ray	30–2000 eV	10^5	Inner-shell electrons, NEXAFS
Crystal mono. chromator	Hard X-ray	3–20 keV		Protein crystallography, diffraction, EXAFS

3.3 Research with Synchrotron Radiation

The primary characteristics of synchrotron radiation can be found in short form in [4]. The most important properties are:

1. High brightness.
2. Variable energy.
3. Polarization: linear to circular.
4. Pulsed time structure.
5. Highly focused (monochromator).
6. Narrow bandwidth (monochromator).

Most of these properties depend upon the characteristics of the electron beam, and on the design of the monochromator and beam line optics.

3.4 Experiments

It is impossible within the scope of this chapter to do justice to even a small fraction of the research which is presently being done using synchrotron radiation. The light source is used for spectroscopic studies of atoms, molecules, solids, molecular materials, and hybrid systems; angle-resolved photoemission is a standard tool for studying band structure; and hard X-rays are used for a variety of experiments, including protein structure studies, angiography, and lithography. A short summary of some of these research areas where synchrotron radiation has played a particularly important role is given here, and a few key references are listed at the end of the chapter. In Chap. 2, about “Accelerators for Medicine”, the application of synchrotron radiation in medicine is outlined.

3.4.1 Atomic and Molecular Physics and Chemistry

Early studies exploited the broad spectrum of synchrotron radiation for absorption measurements of atoms and molecules [5]. Angular distributions of photoelectrons can be studied using the highly linearly polarized light from bending magnets or undulators [9]. Fundamental quantities such as total and partial ionization cross sections can be measured using SR if careful measurements of the sample and source fluctuations are made during the spectral measurement. In molecules, it is important to know fragment yield cross sections and to identify the thresholds for dissociation pathways in order to understand photochemical processes taking place in the atmosphere, for example. A series of studies has been made at LURE (France) and Astrid (Denmark) where photoionization cross sections for singly charged atomic ions were measured. Photoionization induced by SR of several hundred eV energy was recently used to study inner-shell photodetachment of negative-ion clusters [11]. Most of these studies are carried out on VUV or soft-X-ray beam lines. A typical end station setup is shown in Fig. 3.7.

The tunability of SR allows selective excitation of electronic states when the photon energy is tuned to a resonance energy. For valence electronic

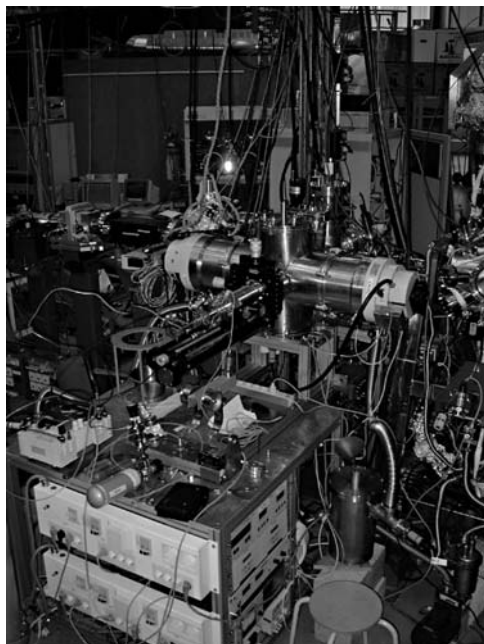


Fig. 3.7. A photograph of the end station at BL I 411 at MAX-Lab. There are a hemispherical electron analyzer, a time-of-flight ion mass spectrometer, and a cluster source mounted on the beam line

levels, this makes it possible to determine the energies of atomic and molecular Rydberg and valence states with high accuracy. The high density of states in the 5–40 eV region often leads to overlapping states, so the narrow bandwidth available from normal-incidence monochromators is necessary to perform selective excitation. For inner-shell states, the electronic states are well separated in energy. In molecules, the molecular orbitals associated with core-excited states are generally localized on a particular atom. Thus selective electronic-state excitation is equivalent to a spatial selectivity in the molecule. Combining the site selectivity of the excitation with spectroscopies such as electron spectroscopy, ion fragment spectroscopy, and fluorescence emission spectroscopy has provided insight into the nature of core molecular orbitals [10]. Another important aspect of core-excited states is the very short lifetime of these states. A core hole in an oxygen molecule will be filled via an Auger transition in a few fs. Since the spectroscopic information pertains to the final states, it is possible to follow the development of core-excited states on the fs timescale by comparing the populations of static and dynamic states in the spectrum. The concepts used in this type of analysis were developed in close collaboration between quantum chemists and experimental groups studying atoms, molecules, and molecular adsorbate systems [6, 7].

3.4.2 Materials Science

The method mentioned in the previous section has been used to study charge transfer between substrates and adsorbates, and between different subunits in a molecule [8]. Although the method probes these dynamical effects in an indirect way, it is one of the few methods able to probe such fast processes. Catalytic materials can be studied by using X-rays for photoelectron spectroscopy, and surface-sensitive methods can illuminate bonding geometries. Near-edge X-ray absorption spectroscopy (NEXAFS) is element-specific, and can provide information about electronic states and bonding. The chemical composition of materials ranging from geological samples to blood and engine exhaust can be obtained. Dynamic information can be obtained by combining XAS with pulsed laser excitation. The degradation of the wood in the excavated Swedish warship *Vasa* has been studied using these methods. In cleaner systems, the bonding of molecules to metallic surfaces can be determined by performing such measurements at different angles with respect to the polarization plane of the SR.

An important tool for structure determination of periodic and semiperiodic solids is extended X-ray absorption fine structure (EXAFS). An inner-shell ionization process produces an electron, which can scatter from ligand neighbors. The scattering is dependent on the distance and on the momentum of the electron. EXAFS can provide distances, coordinations, and even atomic displacements in alloys and under extreme conditions.

Another emerging area exploits the novel polarization properties of SR. Circularly polarized X-rays can be extracted, with the proper optics, from bending-magnet or undulator sources. Magnetic materials can be characterized using circular dichroism, and magnetic domains can be studied in layered systems, for example.

3.4.3 Microscopy

Microscopy is one area which is developing rapidly at synchrotron radiation sources. In photoelectron microscopy schemes, the operation of the ring must be highly stable, and schemes where the ring current is replenished continuously are optimal. The small source size combined with high brilliance makes highly focused beams in the soft-X-ray region possible to achieve. The focused beam is scanned in a raster pattern over the sample surface. Photoelectron spectra are measured at each point in the raster. The surface may be treated by heating or by dosing with reactive agents. More recent spectromicroscopes utilize circularly polarized light for studies of magnetic circular dichroism. Magnetically induced circular dichroism (MCD) is the differential absorption of left and right circularly polarized light in the presence of a magnetic field. Electron detectors which are sensitive to electron spin are employed for spin-resolved photoemission studies. Presently the resolution achieved is 30 nm, but developments in zone plates produced using synchrotron radiation

should push this limit to below 20 nm. Future work will exploit two-photon excitation, where lasers are combined with synchrotron radiation to extract time-dependent information about surface reactions. The coherence of undulator light will also be used for interference lithography.

X-ray microscopy is also used for imaging of biological cells in frozen or aqueous environments. Phase contrast imaging using interferometers for hard X-rays is also used to study medical samples. X-rays penetrate farther into the sample than do electrons; thus X-ray microscopy is able to image thicker samples, and there are no vacuum requirements for X-ray studies, making measurements in realistic environments feasible in many cases. More information on microscopy schemes can be found in [13].

3.4.4 Crystallography

Crystallography is the most important area in synchrotron radiation research in the life sciences. Crystallographic studies provide structural information about macromolecules and complexes [14]. Although it is possible to carry out crystallographic studies using standard rotating-anode sources, the brilliance and small cross section of the beam (microfocused to 10^{-6} m) have allowed determination of much smaller crystals (down to $\sim 10^{-6}$ m) and of samples which are easily damaged by radiation. The multiwavelength anomalous diffraction technique (MAD) requires tunable wavelengths, which is only possible at synchrotron radiation sources. In short, SR permits structural determination of larger, more complex samples than is possible otherwise.

At the frontier of crystallography lie time-resolved studies. There are a number of time ranges of interest for imaging biological processes in action. In some cases it is possible to slow down a process by thermal control, so working on the picosecond timescale will be sufficient for many applications. Conformational changes on the ps timescale in biological macromolecules are highly interesting, since such changes are closely related to function. Enzyme activity in biological systems can change the structure of a macromolecule in several steps; these developments could easily be followed using shot-to-shot imaging methods. Even chemical reactions on the fs-ps scale could be followed using combined spectroscopic and structural studies. Since these experiments require three-dimensional structural information in addition to temporal resolution, the requirements on the light source are stringent.

3.4.5 Summary

The symbiotic development of SR sources and experiments has made the synchrotron radiation laboratory a unique facility. Research is highly cross-disciplinary; experiments in materials science and the life sciences are performed side by side with basic research in atomic and molecular physics and quantum chemistry. Important advances in imaging techniques and crystallography have led to a veritable explosion within biological science at SR

facilities. Future machine developments will focus on short pulses on the fs scale, and on the production of coherent radiation.

References

1. A. M. Liénard: L'Eclairage Electrique **16**, 5 (1898)
2. E. Wiechert: Archives Neerlandaises, 546
3. J. D. Jackson: *Classical Electrodynamics*, Wiley & Sons, New York (1999)
4. Center for X-Ray optics: *X-Ray Data Booklet*, Lawrence Berkeley Laboratory, Berkeley, CA (2001)
5. M. O. Krause: ch. 5, p. 101 in *Synchrotron Radiation Research*, ed. H. Winick and S. Doniach, Plenum, New York (1980)
6. M. N. Piancastelli: J. Electron Spectrosc. Relat. Phenom. **100**, 167 (1999)
7. F. Gel'mukhanov, H. Ågren: Rep. Prog. Phys. **312**, 87 (1999)
8. J. Schnadt, P. A. Bruhwiler et al.: Nature **418**, 620 (2002)
9. V. Schmidt: *Electron Spectrometry of Atoms Using Synchrotron Radiation*, Cambridge University Press, Cambridge(1997)
10. W. Eberhardt: *Applications of Synchrotron Radiation*, Springer, Heidelberg (1995)
11. N. Berrah, R. C. Bilodeau, et al., Phys. Scr. **T110**, 51 (2004)
12. D. Attwood: *Soft X-Rays and Extreme Ultraviolet Radiation: Principles and Applications*, Cambridge University Press, Cambridge (1999)
13. V.V. Aristov, A.I. Erko (eds.): *X-Ray Microscopy IV*, Bogorodski Pechatnik, Chernogolovka Moscow, Russia, (1994)
14. J. C. Philips, K. O. Hodgson: ch 17, p. 565 in *Synchrotron Radiation Research*, ed. H. Winick and S. Doniach, Plenum, New York (1980)

Part II

The Electrostatic Accelerator

Introduction

to Part II – the Electrostatic Accelerator

R. Hellborg¹ and J. McKay²

¹ Department of Physics, Lund University, Sölvegatan 14, 223 62 Lund, Sweden
`ragnar.hellborg@nuclear.lu.se`

² Box 463, Deep River, Ontario, Canada K0J 1P0
`mckay@magma.ca`

It is less that 75 years since Cockcroft and Walton first accelerated an ion beam. The field has passed through many repeated cycles of inspired ideas to practical application to new ideas. The pioneers of electrostatic acceleration are now passing into retirement and beyond, and new generations of accelerator physicists are advancing the technology. This second part of the book is intended to capture some of the wisdom and lore of both the pioneers and the current practitioners of the art.

The authors in this part represent facilities all over the world. Their chapters cover theoretical considerations, historical developments and practical advice. In the spirit of cooperation that has been a hallmark of this field, they have contributed to a resource for users, builders and operators of electrostatic accelerators everywhere.

There may be some inconsistencies between chapters because there are differing ideas and approaches to building machines. Techniques change depending on the needs and resources of the users. Application laboratories usually want reliability, stability and simple operation. In basic research, a need for high-energy or exotic ions may often justify complex operation.

Electrostatic accelerators run at voltages as low as a few hundred kilovolts up to over twenty megavolts. Surprisingly, there are still many elements and principles common to all machines. Ion beam requirements may be simple or exotic. Ion currents vary from almost undetectable to high-powered. Three chapters in this part attempt to deal with the wide range of ion sources.

The authors of this part have contributed chapters on the theory of electrostatics, gas insulation, beam optics and ion generation. Practical advice is given concerning the charging system, resistors and gradient control, the acceleration tube, vacuum, stripping, and ion sources. Operational considerations such as beam diagnostics and safety are also discussed. It cannot be pretended that this part is complete, but it is an attempt to give a fair summary of the current state of knowledge in the field of electrostatic accelerators, and gives the reader pointers toward the vast literature available.

4 History of the Electrostatic Accelerator

J. McKay

Box 463, Deep River, Ontario, Canada K0J 1P0
mckay@magma.ca

4.1 Introduction

The history of electrostatic accelerators might be said to start with early experiments exploring electricity. The first electrostatic machines were constructed in the pursuit of sources of electric charge. Otto von Guericke [1] (1602–1686) may be credited as the inventor of the first electrostatic generator (1663), although he was more famous for his invention of the Magdeburg vacuum hemispheres used to demonstrate the strength of atmospheric pressure. The center of his electrostatic machine was a rotating sulfur sphere that achieved charge separation by friction. The charged ball was then used as a source of charge for experiments exploring the nature of electricity.

In 1784, Walckiers de St. Amand constructed a machine that used an endless band of silk passing over two wooden rollers. Cushions rubbed the silk belt to generate a charge. One version of this machine featured a silk belt 1.5 m wide and 7.6 m long [2]. Other friction machines of increasing complexity and ingenuity were invented throughout the eighteenth and early nineteenth centuries.

Designs that used induction to multiply charge replaced the early friction machines. James Wimshurst [3] built a new machine in about 1883 that was the culmination of these devices. Still manufactured today for use in schools and science displays, the Wimshurst machine was widely used in the late 1800s and early 1900s as a reliable source of high voltage for research.

To this point, electrostatic machines were used to explore the mysteries of electricity and as great parlour or popular lecture demonstrations. They often were beautiful and fantastic machines that appealed to the increasing curiosity of both the scientists and the public. References to and records of the devices are often fragmentary or indirect, making it difficult to give full credit to the many pioneers of the science of electrostatics.

A surprising number of features that became important in the later development of accelerators were considered in the development of these machines. St. Amand's silk belt foreshadowed Van de Graaff. In his graduation thesis (1872), Augusto Righi [4] (1850–1920) built a charge transfer device consisting of metal cylinders mounted on an insulated rope, quite similar in concept to Herb's Pelletron. Righi referred to his machine as a charge amplifier rather than a voltage generator, as it was used to amplify and measure very small

electrical charges. It has been reported, in a reference given only as “an 1911 encyclopaedia”, that F. Tudsbury, in 1900, discovered that an influence machine enclosed in a tank of compressed air or carbon dioxide would produce sparks more than double the length produced at atmospheric pressure. The inventors of this period showed great ingenuity as they built the first foundations of our knowledge of electrostatics. It is not surprising that many of their ideas were independently rediscovered as the technology of accelerators developed.

4.2 The First Accelerators

By the advent of the twentieth century, electrical phenomena were less mysterious, and practical sources of voltage and current became readily available. The interest in electrostatic machines waned. However, the fields of atomic and then nuclear physics began to develop. Spectroscopy and the search for ways to identify elements and ion species brought acceleration into use. In spectroscopy, ionized particles were accelerated across a constant voltage gap and then identified by their charge-to-mass ratio according to their deflection in a transverse magnetic field. Francis Aston (1877–1945), working in the laboratory of J.J. Thomson, invented the magnetic spectrometer and measured the ratio of neon isotopes [5]. The importance of this development is illustrated by the fact that Aston received the Nobel Prize for this work in 1922, just three years after his first measurements.

Ernest Rutherford (1871–1937), working in the Cavendish Laboratory at Cambridge in 1919, transmuted nitrogen atoms into oxygen by bombarding them with alpha particles from “Radium C”, i.e. ^{214}Bi , which decays by a 0.02% branch to ^{210}Tl , producing a 5.617 MeV alpha particle [6]. This transmutation of nitrogen into oxygen was the first artificially induced nuclear reaction. In 1928, Rutherford, in an address to the Royal Society, identified the need for “a source of positive particles more energetic than those emitted from naturally radioactive substances”. Low-energy positive ions were unable to penetrate the repulsive Coulomb barrier surrounding the positive nucleus of the atom. The need for accelerators had been established and the first era of accelerator development began.

J.D. Cockcroft (1897–1967) and E.T.S. Walton (1903–1995) obliged. In 1930, they accelerated protons to 200 keV, the first accelerated particle beam. Needing more energy, they devised and constructed a voltage multiplier circuit. In 1932, they accelerated protons to 600 keV and directed them onto a lithium target. The resulting $^7\text{Li}(p, \alpha)^4\text{He}$ reaction was the first accelerator-induced nuclear reaction [7].

The search for practical ways to accelerate ions was progressing on many fronts. Robert Van de Graaff (1901–1967) began experimenting with a belt-driven voltage generator at Princeton in the fall of 1929 and presented his

concept at a meeting of the American Physical Society in 1931 [8]. He reported a voltage of 1.5 MV between the two spheres, one positive and one negative, of this device. Van de Graaff and his accelerator were transferred to the Department of Terrestrial Magnetism of the Carnegie Institution in Washington. A series of larger machines were built and tested, including one (with a 2 m sphere) that was too large to house indoors! These culminated, by late 1932, in a machine that accelerated protons to 600 keV. These were the first Van de Graaff beams used in a nuclear-physics experiment [9]. In October 1933, a new machine began operation at up to 1.2 MeV and a full nuclear-physics program commenced [10].

Electrostatic generators continued to grow in a most literal sense. Van de Graaff moved to the Massachusetts Institute of Technology in 1931 and began construction of the huge Round Hill double Van de Graaff in a former airship hangar. Two 4.6 m spheres topped a pair of 6.7 m tall Texolite columns. A maximum voltage differential of 5.87 MV was reached, and reliable operation at 5.1 MV differential was achieved [11]. No accelerator tube was installed in this generator. More modest machines, built in better-controlled environments, followed. A 2.75 MV accelerator using the positive Round Hill sphere produced useful beams for nuclear physics. Van de Graaff and John Trump published a paper [12] in 1937 describing a 1.2 MV electron accelerator built for the Harvard Medical School. This powerful source of 1 MeV X-rays was the first electrostatic accelerator used in clinical medicine.

At the same time that Van de Graaff and his associates were building large open-air machines, Ray Herb at the University of Wisconsin began experimentation with a series of enclosed machines. In 1931, after seeing the first primitive cyclotron at Ernest Lawrence's laboratory in Berkeley, Herb had worked with Glen Havens to build a vacuum-insulated belt-driven generator. This device achieved about 300 kV. In 1933, Herb decided to pressurize this 0.75 m diameter, 1.8 m long tank. At about 0.33 MPa air pressure, the generator reached 500 kV. Herb and his group immediately began development of a complete accelerator. In 1934, Herb was able to take data for his Ph.D. thesis. The developments that resulted in this successful accelerator were summarized in a 1935 paper in the *Review of Scientific Instruments* [13].

After completion of his doctorate, Herb spent some time at the Department of Terrestrial Magnetism in Washington working on the 1.2 MV open-air Van de Graaff type accelerator. Returning to Wisconsin in the fall of 1935, he, along with Parkinson and Kerst, developed a new 2.5 MV machine that incorporated many new features [14]. These included potential grading of the column and tube, a field-shaping column ring, feedback voltage control and an insulating gas mixture. This was contained in a 1.7 m diameter by 6.1 m long tank. The tube consisted of a series of metal electrodes separated by 6.4 cm long porcelain cylinders. The porcelain insulators had a corrugated profile. The tank was capable of withstanding a pressure of 0.75 MPa, twice that of the previous machine. While this machine was the first to have many

of the attributes of a modern accelerator, it contained a lot of Texolite, red sealing wax and even some wood. At 0.75 MPa air pressure, the partial pressure of oxygen posed a major hazard. After a number of fires, a permanent CO₂ injection system was installed.

In 1939, this 2.5 MV machine was disassembled and the tank was used to contain a new accelerator. The new machine had a tube almost 4 m long, two potential-grading “intershields” and a single-ended configuration for the column [15]. This highly successful accelerator ran reliably at 4.3 MV and held the record for the highest voltage until the early 1950s.

External events had major effects on the development of accelerators in the period between World Wars I and II. The crash of 1929 and the start of the Great Depression coincided with the first development of accelerators. The next ten years saw a great deal of progress, but budgets were very tight and a premium was placed on sealing wax and ingenuity.

4.3 The Postwar Years

The Second World War changed the face of science. The Herculean efforts of the Manhattan Project, as well as producing the atomic bomb, had the secondary effects of vastly increased knowledge in physics and technology. Peace brought expectations of large research budgets and the hope of limitless progress. This set the stage for the next major period of progress in electrostatic accelerators.

Herb’s 4.3 MV machine had run almost constantly at Los Alamos, complemented by a 2.0 MV machine built by Joe McKibben, another Wisconsin graduate. At MIT, Van de Graaff and his associates designed a vertical 4.0 MV machine that introduced resistor grading to the column and tube structure. This machine was replicated at Chalk River and at other laboratories. 1947 marked the establishment of the High Voltage Engineering Corporation by Trump, Denis Robinson and Van de Graaff and soon began supplying electrostatic generators used in cancer therapy and radiography and in studies of nuclear structure. One of their first products, produced with Ray Herb as a consultant, was a 4 MV electrostatic accelerator required as the injector for the “Cosmotron”, the first proton synchrotron. This combination came on line in 1952. The injector was still running, but for other purposes, in 1999 [16].

In the late 1940s, Trump [17] at MIT and McKibben [18] at Los Alamos constructed new machines with the aim of reaching 12 MV. The Los Alamos machine reached 13 to 14 MV without tubes, but both machines were limited to 8–9 MV in practical operation. These large machines doubled the useful energy achieved by Herb’s prewar machine. The MIT design was used by HVEC as a prototype for its CN series of accelerators. Twenty-six CNs were installed between 1951 and 1966, the first “mass-produced” accelerators. In

the same period, HVEC produced many 3 and 4 MV KN series machines, as both electron and positive-ion accelerators.

These commercially produced accelerators had many common features: resistor grading of the column and tube, PVC-sealed glass and metal tubes, belt charging, field-shaping column hoops and high-pressure (0.7 to 0.8 MPa) N_2 plus CO_2 gas insulation.

The single-ended machines provided practical beams for the growth of nuclear-structure studies but it was clear that much higher energies would be required to explore the vast expanse of the table of isotopes. Voltages on the order of 10 MV were achievable, but tube and column structure limitations inhibited progress above that level. Size could accomplish only so much, and cost and complexity expanded at a higher rate than voltage.

4.4 Tandems

In the 1930s, a number of researchers experimented with charge exchange schemes of acceleration. Otto Peter at the University of Geiszen used multiple stages of positive- and negative-hydrogen acceleration to produce a 100 pA beam [19]. Independently, W.H. Bennett (1903–1987) suggested and later patented [20] the concept of an energy-doubling accelerator. The practical application of these ideas would wait until the mid 1950s and the development of a sufficiently intense source of negative protons. Publications by Luis Alvarez [21] in 1951 and Bennett [22] in 1953 refined the tandem concept. A.C. Whittier at Chalk River measured in 1954 the cross sections for negative-hydrogen-ion production in various gases at various energies [23]. Weinman and Cameron produced a 20 μA beam of H^- at Wisconsin [24] in 1956.

This activity induced the Chalk River Nuclear Laboratory (CRNL), in 1954, to invite HVEC to submit a proposal for a 5 MV tandem. Thus began a period of explosive advance in accelerator technology. September 1956 saw the placing of an order from CRNL to HVEC for the first tandem accelerator. The first test experiments with beams from the tandem were performed at HVEC in Burlington, Massachusetts, on 25 June 1958. The machine was moved north to Chalk River, and the first beam on target was achieved there in February the next year.

The “EN” tandem, as it was designated, had a glass and steel, epoxy-bonded, horizontal column and a single belt. The tank was 2.4 m in diameter and 11 m long. Rated at 5 MV, it ultimately ran as high as 7 MV. This machine was moved to the Université de Montréal in the late 1960s, where it continues to run and has recently been upgraded. EN-1 was the prototype for 30 similar machines produced between 1958 and 1973.

As exciting as the development of the EN was, it was supplemented by a similar but larger machine, the “FN”, first delivered to Los Alamos in 1963. This accelerator was 3.66 m in diameter and 13.4 m long. Rated at a nominal

7 MV, FN accelerators have been known to run relatively reliably at over 11 MV. Nineteen FN accelerators were manufactured.

The next step in HVEC's series of machines was the "MP" model. MP-1 was delivered to Yale University in 1965. This machine represented a major advance in design only six years from the delivery of the first tandem. The simple compressed column was replaced with a truss bridge structure, still glass and steel, still compressed, but twice the length of that of the first tandems. The tank of the MP was over 20 m long and over 4.5 m in diameter. The earliest years of MP operation were beset with many problems as operators learned how to control it. In a few instances, some FNs were operating embarrassingly at the same voltage as the larger and much more expensive machines. The MPs eventually were refined and ran at voltages far above their nominal 10 MV rating. Michel Letournel and his colleagues at Strasbourg ultimately pushed a modified MP to 18 MV with beam [25].

In a period of just 14 years, from the installation of EN-1 in 1959 at Chalk River to the installation of MP-10 at Strasbourg in 1973, HVEC produced 55 tandem accelerators. The MP was the last of the "mass-produced" large tandems.

The advances made by these HVEC machines were made possible by a number of innovations. The inclined field tube design by Van de Graaff, Rose and Wittkower [26] in 1962 greatly reduced the limitations imposed by earlier tube designs. The continued development of better resistors helped to control voltage-gradient variations. The introduction of SF₆ insulating gas, either 100% or as a mixture, extended the voltage range. Improvements in ion sources provided critical advances in beam variety and intensity. Adaptation of Herb's Pelletron charging systems to replace the belt in HVEC machines brought further advances in reliability and performance. Second stripping of ions in the high-energy column produced higher charge states of heavy ions and therefore higher beam energies.

In parallel with the development of the commercially available HVEC machines, there was the construction of other designs. Two 5 MV vertical tandems were built in 1959 by the Metropolitan-Vickers Electrical Company in the UK for the Harwell and Aldermaston Laboratories [27]. In 1965, the Japanese government ordered two 5 MV machines: a vertical tandem from Toshiba that was installed at the University of Tokyo [28] and a horizontal machine by Mitsubishi installed at the University of Kyoto.

A new commercial supplier was founded in 1965 with the formation of the National Electrostatics Corporation (NEC) by Herb, James Ferry and Theodore Pauly. Since his return to the University of Wisconsin in 1946, Herb had directed an extensive program of research into vacuum techniques, ceramic-to-metal bonding, charging systems and other accelerator technologies. This was in addition to his extensive research in nuclear physics. The company's first accelerator order was for a coupled system combining a 4 MV single-ended machine injecting into an 8 MV tandem. These machines were

shipped to São Paulo, Brazil, in November 1970, and in December, NEC received an order from the Australian National University for a 14 MV vertical tandem. The ANU machine was reported to be running at 16.7 MV in 1988 [29]. The ultimate challenge came in 1975 with an order from the Oak Ridge National Laboratory for a 25 MV tandem. The first NEC machines represented a huge risk for the new company and a radical departure from the HVEC designs. The machines were vertical. They featured a modular construction, ceramic insulators, corona point gradient control, high-vacuum ceramic tubes and the new “Pelletron” charging system. This system of linked metal cylinders and insulated sections was devised by James Ferry during the years of work by the University of Wisconsin group. The successful design followed early antecedents such as Righi’s 1872 charge amplifier and many unsatisfactory prototypes until a workable design evolved.

Between 1970 and 1991, NEC manufactured 11 large tandems. In addition to those mentioned above, “14UD” tandems were supplied to the Weizmann Institute in Israel and the Tata Institute in India. 20 MV tandems were built for the Japan Atomic Energy Research Institute (JAERI) and for the Comisión Nacional de Energía Atómica in Buenos Aires, Argentina. A 15UD went to the Nuclear Science Centre in New Delhi, India, in 1991.

The NEC era was marked by a switch to vertical machines and heroic design leaps. The 14 and 15 MV machines were “straight-through” tandems, and the others were “up-down” or “folded”, that is, the beam was deflected 180° in the terminal and returned to ground. There was no production of large series of similar machines, but the NEC modular approach to construction mitigated this to a large extent.

The major production era for the HVEC tandems spanned 1959 to about 1973, and the NEC era spanned 1970 to 1991. Production of large tandems by the two great commercial producers barely overlapped.

4.5 The Big Machines

Efforts to push beyond the 20 MV level began in the early 1970s as the great era of tandem expansion slowed. A number of unique machines were the result. Three were extensions of the standard MP design and three were very large, unique machines.

A modified MP-style accelerator called the XTU was supplied by HVEC to the Laboratori Nazionali di Legnaro, Padova, Italy. Based on the MP structure in an enlarged tank and equipped with a Laddertron charging system, the XTU was designed to inject into a linac booster at 15 MV. This machine was accepted in 1981 and went into full operation in 1984. Commissioning of the ALPI booster started in 1994.

Components of MP-0, the HVEC test machine, were recycled into the HI-13 Beijing accelerator that saw first beam in 1985. The HI-13 features an enlarged tank and a Laddertron charging system. It runs in the 13 MV range.

In 1972, the Science Research Council in Britain approved the development and construction of a 30 MV tandem to be built at Daresbury. The design team came from neither the HVEC nor the NEC tradition and they resolved to make electrostatic design more scientific and less empirical. The history of the project was summarized by T.W. Aitken at the Padova conference in 1992 [30]. Aspects of field distribution, inductive charging systems, breakdown of SF₆ in spark gaps, surge protection of electronics, organic-free tubes, control systems and many more items were studied and tested in detail. Commissioning of the resulting machine began in 1980. It was a “straight-through” tandem with an active tube length of 18 m in each column. The charging system, later marketed by HVEC as the “Laddertron”, consisted of pairs of cylinders interconnected by flat sections reminiscent of the rungs of a ladder. This charging system is used at Stony Brook in an FN, as well as in the Orsay, Beijing and Legnaro accelerators. The tubes were of brazed ceramic-metal construction. The internal electronics were controlled over infrared light links. This great machine reached 29.5 MV during voltage testing but was unable to exceed about 20 MV in full operation [31]. When the facility was shut down in 1992, tube improvements were being considered that could have taken it to a higher operating range.

The Oak Ridge National Laboratory, in 1975, ordered a 25 MV tandem from NEC. Unlike the Daresbury machine, this tandem was built in the “folded” configuration with a large 180° magnet in the terminal. The 25URC Oak Ridge machine was tested, in 1979, at 32 ± 1.5 MV without tubes [16]. This stands as the current record for the maximum voltage produced with any machine. However, like Daresbury, it originally had difficulty running at over 20 MV. Today, it runs at almost 24 MV with beam.

Construction of the “VIVITRON” [32], based on Letournel’s experiments on the Strasbourg MP, began in 1985 and achieved first beam in 1993. The tank was 50 m long but only 8.4 m in diameter at the center. A series of seven “porticos”, or open cages, surrounded the column to control the radial gradient. The assembly was supported with insulating posts, and the column was made up of large plastic plates. A charging belt ran through the full length of the machine. Development of this radical design of this machine produced a greater understanding of electrostatic design. However, a variety of problems inhibited operation over 20 MV. The VIVITRON was shut down in December 2003.

The last of the big machines to be built was the ESTU accelerator at Yale University [33]. This project extended the active structure of MP-1 by 25%, and incorporated a single Letournel “portico” and a 7.6 m diameter tank with a shape optimized for the portico. More modest in its aim than the 25URC or the VIVITRON, the ESTU structure was tested to about 22.5 MV in 1987 and runs consistently in the 19 MV range. This machine is a true hybrid, combining a structure and tubes from HVEC with a charging system from

NEC and resistors and a portico structure from Vivirad, a company founded by Michel Letournel.

The great machine at Oak Ridge is at the pinnacle of electrostatic-accelerator development to date. Successfully running at 24 MV [34], the Oak Ridge machine is now accelerating radioactive beams as its primary role. It was planned to, and did, inject particles into the ORIC cyclotron for boosting its energy. Now that cyclotron is being used to generate radioactive ion species for injection into the tandem. Such versatility is the hallmark of many electrostatic-accelerator laboratories.

In his 1974 review paper [35], Allan Bromley stated, “Looking further into the future, electrostatic accelerator technology has now advanced to the point where it becomes reasonable to at least consider designs for tandem electrostatic accelerators in the 50–60 MV range”. This prediction reflected the optimistic view of many in the field of electrostatics at that time. It was also a period of generous funding fueled by the advances in nuclear-structure research. The explosion of ideas in this field called for ever-greater energies in order to breach the Coulomb barrier in the heaviest nuclear systems.

As the voltage increases, electrostatic accelerators suffer from the fact that the stored energy in a capacitive system increases as the square of the voltage. Further, the capacitance increases as the accelerator gets larger. Thus spectacular sparks in the 20 MV plus range often lead to serious damage in large accelerators.

The difficulties in achieving higher voltages led to the adoption of more complex “afterburner” schemes, where the tandem injected a beam into a linac or cyclotron for final acceleration. This raised the question of whether the tandem was the main accelerator or just a large ion source!

Booster accelerators were being planned in the mid 1970s even though the suggested technologies, such as superconducting linacs or cyclotrons, were very complex. The Chalk River superconducting-cyclotron concept was described at conferences in 1974, but it was not operational until the late 1980s. The superconducting linacs came on line sooner, usually with just a few modules to start. Injection of beam from the Oak Ridge 24URC into the ORIC cyclotron, built in the 1960s, was planned from the start of the project and was achieved in 1982 [36].

Of the four super machines, Daresbury, Oak Ridge, the VIVITRON and Yale, only Oak Ridge and Yale remain in operation in 2004. Five of the eleven MPs have been taken out of service. About 75% of the smaller, first-generation tandems are still in operation, although many have been moved to new institutions. Their versatility and relatively low cost of operation have preserved them. Many have been converted to Pelletron charging and otherwise upgraded. Amongst other useful roles, these machines often serve as test beds for new ion beam application technologies and applications, helping to define the requirements for the next generation of machines.

4.6 Machines for Applications

The emphasis in histories of electrostatic accelerators is on the ever-larger machines built for research in nuclear physics. However, the vast number of accelerators are smaller, low-voltage machines used in a wide variety of applications. The commercial suppliers' survival and prosperity is based on these small machines. The wide variety of charging systems divides these smaller potential-drop accelerators into the classifications of electrostatic accelerators, with mechanical (i.e. belt or chain) charging, and cascade accelerators, including asymmetrical, symmetrical and parallel-driven circuits, insulating-core transformers, etc. The line between a high-voltage electrostatic generator and a power supply has become blurred.

Most accelerators in the world are used as implanters or for accelerator mass spectrometry (AMS), materials research, biology or medical applications. Many of these applications are discussed in Part III of this book. The main interest of the users is in the application of the machine rather than in the machine technology. Whereas the research community could risk some unreliability in the quest for elusive results, the applied users see the accelerator almost as an appliance and expect beam on demand. The manufacturers have been able to fill this need to a large extent, and it is not unusual to find users reporting that they have not had to open their accelerators for service in a year or more. It is interesting to note that tandem machines for applications are now up to 5 MV, the same rating as the first tandem in 1957. There has also been a resurgence of cascade-type machines in the development of more stable and reliable accelerators for applications.

4.7 The Future

The current state of development of electrostatic accelerators poses the question of whether we shall see construction of another large machine. High energies can be obtained from postaccelerator systems and, at the other extreme, astrophysicists seek intense low-energy beams. Fields such as AMS look for beams of great stability and purity. Industrial and medical applications require high reliability and ease of operation.

The changing demands of science and the availability of high-energy beams from booster systems have dulled the quest for higher voltages. That, combined with the technical and financial challenges involved in attaining significantly higher voltages, makes it doubtful that we shall see development of new large machines in the near future.

In the 75 years since Rutherford called for "a source of positive particles more energetic than those emitted from naturally radioactive substances", we have seen periods both of rapid development and of consolidation. Nuclear-structure physics was the dominant driving force for accelerators in the 1960s and 1970s. Slowly, lower-voltage machines started to find uses in applications

of the techniques of nuclear physics. Material analysis and modification, particularly in the manufacture of solid-state devices, became important. AMS pushed analysis techniques to spectacular sensitivities and forced accelerator developers to new understandings of machine operation and the need for ultrastable operation. Part III of this volume discusses the uses of accelerators, and only two of the ten chapters deal with nuclear physics. In 2004, it is the needs of applications that now drive the development of electrostatic accelerators.

A number of review papers have been written over the years describing the development of the electrostatic accelerator [16, 35, 37, 38]. International conference series, such as the Heavy Ion Accelerator Technology Conferences that began in Daresbury in 1973, have chronicled progress in the field. The Symposium of North Eastern Accelerator Personnel started in 1968 and has met annually to discuss the problems in the operation of electrostatic accelerators. These meetings highlight the collaborative efforts of the community and the importance of contributions by a great number of both researchers and operations people in the field.

This chapter relates only a small portion of the countless contributions made to the science of electrostatics. The exchange of ideas through the years and the accumulation of experience by the practitioners of the art and science of electrostatics form a rich story indeed.

References

1. O. von Guericke: *Experimenta Nova (ut vocantur) Magdegurgica de Vacuo Spatio* (Amsterdam, 1672)
2. N. Rouland: *Description des machines electrostatiques a taffetas* (Amsterdam, 1785) pp 45–56
3. J. Wimshurst: A new form of influence-machine, Proc. Phys. Soc. London **12**:1, 403 (October 1892)
4. J. Gray: *Electrical Influence Machines* (Whittaker, London, 1903)
5. F. Aston: Nature **105**, 617 (1920)
6. E. Rutherford: Collisions of alpha particles with light atoms. IV. An anomalous effect in nitrogen, London, Edinburgh and Dublin Philos. Mag. J. Sci., 6th series **37**, 581 (1919)
7. J.D. Cockcroft, E.T.S. Walton: Nature **129**, 242 (1932)
8. R.J. Van de Graaff: Phys. Rev. **38**, 1919A (1931)
9. R.J. Van De Graaff, K.T. Compton, L.C. Van Atta: Phys. Rev. **43**, 149 (1933)
10. M.A. Tuve, L.R. Hafstad, O. Dahl: Phys. Rev. **48**, 315 (1935)
11. L.C. Van Atta, D.L. Northrup, C.M. Van Atta, R.J. Van de Graaff: Phys. Rev. **49**, 761 (1936)
12. J.G. Trump, R.J. Van de Graaff: J. Appl. Phys. **8**, 602 (1937)
13. R.G. Herb, D.B. Parkinson, D.W. Kerst: Rev. Sci. Instr. **6**, 261 (1935)
14. R.G. Herb, D.B. Parkinson, D.W. Kerst: Phys. Rev. **51**, 76 (1937)
15. R.G. Herb, C.M. Turner, C.M. Hudson, R.E. Warren: Phys. Rev. **58**, 579 (1940)

16. G.A. Norton, J.A. Ferry, R.E. Daniel, G.M. Klody: A retrospective of the career of Ray Herb. In: *Proc. Eighth International Conf. on Heavy Ion Technology*, A.I.P. Conference Proceedings, vol. 473 (1999) p. 3
17. J.G. Trump: "Electrostatic Sources of Ionizing Energy", Transactions of the American Institute of Electrical Engineers, 70, Part1, 1021–1027 (1951)
18. R. Woods, J.L. McKibben, R.L. Henkel: Nucl. Instr. Meth. **122**, 81 (1974)
19. O. Peter: Ann. Physik A **27**, 299 (1936)
20. W.H. Bennett: U.S. Patent 2,206,558 (1940)
21. L.W. Alvarez: Rev. Sci. Instr. **22**, 705 (1951)
22. W.H. Bennett: Rev. Sci. Instr. **24**, 915 (1953)
23. A.C. Whittier: Can. J. Phys. **32**, 275 (1954)
24. J.A. Weinman, J.R. Cameron: Rev. Sci. Instr. **27**, 288 (1956)
25. J.C. Oberlin, G. Heng, M. Letournel: Nucl. Instr. Meth. A **244**, 35 (1985)
26. R.J. Van de Graaff, P.A. Rose, A.B. Wittkower: Nature **195**, 1292 (1962)
27. K.W. Allen, F.A. Julian, W.D. Allen, A.E. Pyrah, J. Blears: Nature **184**, 303 (1959)
28. S. Masuda: Direct extraction negative ion source for the tandem accelerator. In: *Proc. Symp. on Ion Sources and Formation of Ion Beams* (Brookhaven National Lab. 1971) p. 289
29. D.C. Weisser: The 16.7 MV upgrade of the Canberra 14 UD. In: *Proc. Symp. of North Eastern Accelerator Personnel*, Yale (1988) p. 37
30. T.W. Aitken: Nucl. Instr. Meth. A **328**, 10 (1992)
31. H.R. McK. Hyder: Nucl. Instr. Meth. A **287**, 1 (1989)
32. M. Letournel et al.: Nucl. Instr. Meth. A **244**, 56 (1985)
33. H.R. McK. Hyder, J. Baris, T.A. Barker, J.W. McKay, P.D. Parker, D.A. Bromley: Status of the ESTU accelerator at Yale. In: *Proc. Symp. of North Eastern Accelerator Personnel Yale* (1988) p. 13
34. M.J. Meigs, R.C. Juras: Oak Ridge 25URC tandem accelerator. In: *Proc. Symp. of North Eastern Accelerator Personnel, Lund*, ed. by R. Hellborg et al. (ISBN 91-631-2676-1, 2001) p. 342
35. D.A. Bromley: Nucl. Instr. Meth. **122**, 1 (1974)
36. C.M. Jones et al.: Nucl. Instr. Meth. A **244**, 7 (1985)
37. H.R. McK. Hyder: Limitations of electrostatic accelerators. In: *Proc. Eighth International Conf. on Heavy Ion Technology*, A.I.P. Conference Proceedings, vol. 473 (1999) p. 47
38. H.R. McK. Hyder: Charging systems in ancient times. In: *Proc. Symp. of North Eastern Accelerator Personnel, Lund*, ed. by R. Hellborg et al. (ISBN 91-631-2676-1, 2001) p. 110

5 Electrostatics

H.R.McK. Hyder

Department of Physics, Oxford University, Denys Wilkinson Building,
Keble Road, Oxford OX1 3RH, England
`r.hyder1@physics.ox.ac.uk`

5.1 Introduction

The first electrostatic accelerators were air-insulated, operating without air-conditioning in humid surroundings, sometimes even in the open air. The designers faced many of the same challenges as high-voltage power-line engineers, whose accumulated experience guided them in specifying the size of buildings and the shape of components needed to avoid sparking. A major challenge was preventing corona discharge from the sharp edges of conductors and leakage currents across the surfaces of damp or dirty insulators.

However, the low dielectric strength of air and the bulky nature of contemporary high-voltage components combined to make compact designs impractical. Photographs of early machines reveal the generous layouts adopted in most laboratories (Fig. 5.1).

They also disclose weaknesses imposed by unsuitable buildings. The roof structure of the Round Hill hangar, in which Van de Graaff achieved 5.1 MV between the twin columns of his large generator, clearly affected the sparking voltage (Fig. 5.2). The overall performance can hardly have been helped by

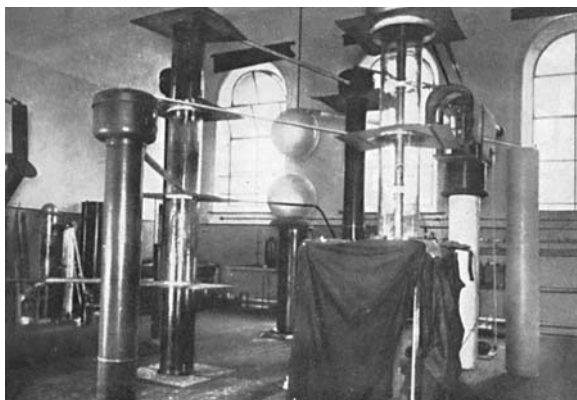


Fig. 5.1. Cockcroft and Walton's 800 kV accelerator in the Cavendish Laboratory, Cambridge, in 1932 (Reprinted from [1], copyright (1932) with permission from the Royal Society)



Fig. 5.2. Van de Graaff's twin-column electrostatic accelerator sparking to the roof and the walls in the Round Hill hangar in Connecticut in 1932 (Reprinted from [2], copyright (1974) with permission from Elsevier)

the incursion of seawater at spring tides. Air at atmospheric pressure and the use of available but inconvenient buildings ensured that the design of these early accelerators was far removed from the ideal electrostatic form. The story of these early developments was summarized by Bromley [2] in a review of large electrostatic accelerators.

The introduction of compressed air, quickly replaced by inert gases at high pressure, required but also made possible the adoption of more sophisticated electrostatic designs. By controlling the gas pressure and composition, it was possible to specify a safe working field with confidence. At the same time, the removal of moisture and dirt delayed the onset of surface leakage currents. Improved high-voltage components became smaller, making it possible to exploit the high dielectric strength of the gas in a compact design. These changes are obvious if one compares the layout of the Round Hill accelerators of Fig. 5.2 with Herb's "Long Tank" accelerator that operated at over 4 MV (Fig. 5.3) in 0.8 MPa of compressed air.

As long as ambient conditions were poorly controlled and imperfectly understood, it was reasonable for designers to rely on precedent and experience. But if the properties of insulators and the presence of water vapor could be measured and controlled, electrostatic-field calculations were needed to enable designers to meet the relevant criteria for avoiding breakdown.

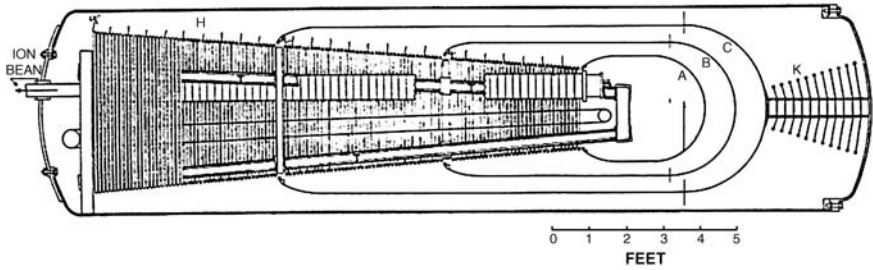


Fig. 5.3. Herb's horizontal accelerator, operating in compressed air with a dash of Freon at over 4 MV (Reprinted from [3], copyright (1940) with permission from APS)

Analytical solutions of Laplace's equation for simple geometries, such as concentric spheres and cylinders, coupled with reliable data for the electrical strength of gases, provide such a basis for designing machines to work at a specified voltage. Unfortunately, the highest fields in real machines occur where the geometry is not simple and the solutions are not analytic, as seen in the field distribution of a typical small accelerator (Fig. 5.4).

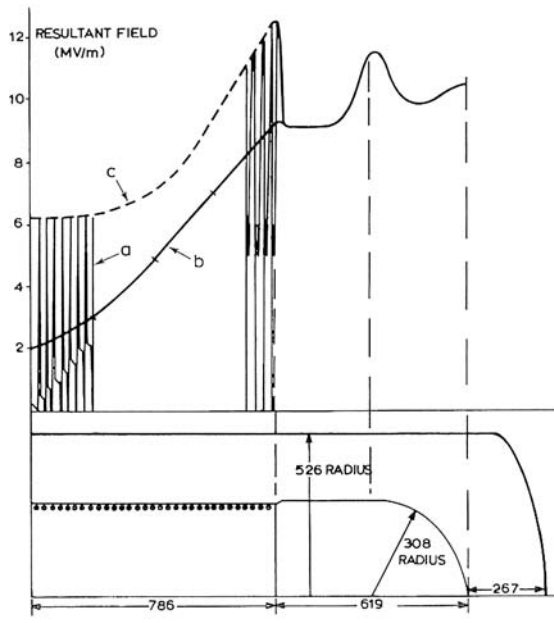


Fig. 5.4. Field distribution on the surfaces of the terminal and column in a 1.5 MV electrostatic accelerator operating in SF_6 at 0.3 MPa [4]. Dimensions in mm. (a) Fine structure due to hoops. (b) Field of equivalent smooth column. (c) Envelope of resultant field (Reprinted from [4])

Numerical methods (or analog measurements with resistance networks or electrolytic tanks) are required to determine the field in and around the column and at transitions between spherical and cylindrical electrodes. Until such techniques became accurate and widely available, it was tempting to neglect the fine structure. A more serious miscalculation in the design of early machines was the assumption that the column components, particularly the accelerator tube, would be able to sustain the same longitudinal field independent of the total voltage. The reasons for this “total-voltage effect” are discussed elsewhere (Chap. 8).

5.2 Field Distributions

Most pressurized electrostatic accelerators conform to one of two basic geometries. Single-ended machines consist of a cylindrical column bounded by toroidal hoops, terminating in a smooth, cylindrical high-voltage terminal with a hemispherical end (Fig. 5.4). The minor diameter of the hoops is small compared with the column diameter. The hoops are separated by gaps somewhat less than their minor diameter. The voltage between adjacent hoops and the macroscopic longitudinal field inside the column are constant. The column and terminal are mounted axially inside a cylindrical tank that has flat or domed ends. Sometimes the cylindrical part of the tank opposite the terminal is joined to a conical section that tapers to a reduced diameter towards the column base, where the field is lower. A field-free region between the base of the column and the tank contains the charging system, motors, controls etc.

Tandem accelerators are symmetrical about the center plane of the terminal. The column, usually constant in radius, extends from end to end of the tank, which is often cylindrical, but may incorporate conical sections in larger machines. The column geometry is similar to that in single-ended machines. In some of the larger machines, one or more intermediate electrodes (or intershields) are installed between the terminal and the tank, as shown in Fig. 5.5. They are supported from and electrically connected to the column and are at the same potential as the points of support. They reduce the peak radial fields and allow an increase in voltage for the same tank diameter. Outside the column, folded tandems are electrostatically similar to single-ended machines.

5.2.1 Macroscopic Field in Cylindrical Geometry

The radial field between concentric cylindrical conductors is given by

$$E(r) = V/(r \ln(r_2/r_1)) \quad (5.1)$$

where V is the potential difference between the two conductors, r_1 is the radius of the inner conductor and r_2 that of the outer. The maximum field,

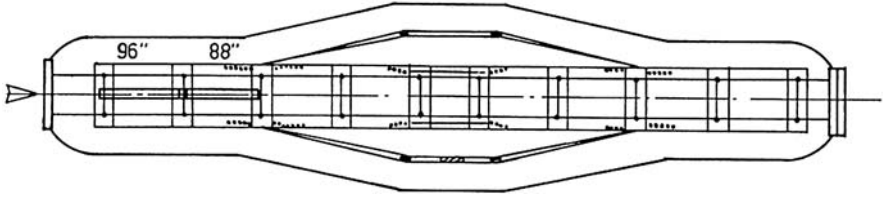


Fig. 5.5. HVEC model MP tandem accelerator with intershield, as modified by Letournel (Reprinted from [5], copyright (1984) with permission from Elsevier)

$$E_{max} = V/(r_1 \ln(r_2/r_1))$$

occurs on the surface of the inner conductor. For fixed values of r_2 and E_{max} , V depends on r_1 :

$$V(r_1) = E_{max} r_1 \ln(r_2/r_1) \quad (5.2)$$

The maximum value of V is obtained by setting the partial differential $\partial V/\partial r_1 = E_{max}[\ln(r_2/r_1) - 1] = 0$, whence $r_2/r_1 = e$ and $V_{max} = E_{max} r_2/e$. However, it is not always possible to achieve this geometry. For example, the ideal terminal radius of a 4 MV accelerator operating in SF_6 at 0.8 MPa, assuming a safe working value of $E_{max} = 16 \text{ MV/m}$ (see Sect. 5.3), is only 0.25 m. This is only just big enough to accommodate a small charging belt and a small accelerator tube. At lower voltages the ratio r_2/r_1 may need to be greater than $1/e$ if the terminal is to be big enough. Fortunately, quite large departures from $r_2/r_1 = e$ have little effect on the maximum voltage, as shown in Fig. 5.6. As the design voltage increases so do the dimensions, and it becomes easier to meet the structural and spatial requirements for column and terminal without prejudicing the electrostatic design.

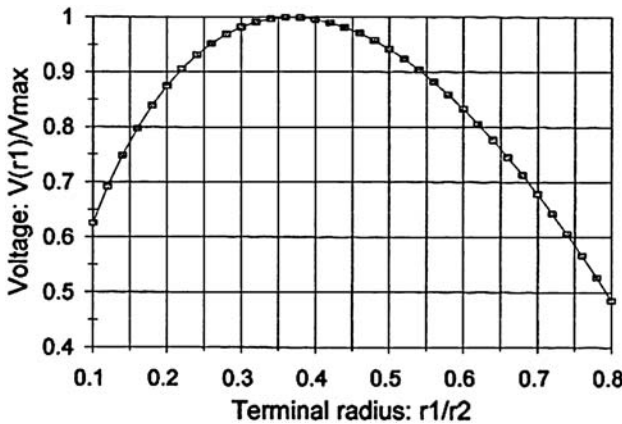


Fig. 5.6. Variation of maximum voltage with terminal radius in cylindrical geometry, normalized to unit maximum field and unit tank radius

5.2.2 Macroscopic Field in Spherical Geometry

The radial field between two concentric spheres is

$$E(r) = V[r_1 r_2 / (r^2(r_2 - r_1))] \quad (5.3)$$

where r_1 and r_2 are the radii of the inner and outer conductors. The field at the surface of the inner conductor is $E(r_1) = V[r_2 / (r_1(r_2 - r_1))]$ and the condition for maximum voltage, E and r_2 constant, is $r_2/r_1 = 2$.

For this value of r_2/r_1 , the spherical surface field is 39% greater than the cylindrical surface field and for the optimum cylindrical geometry, $r_2/r_1 = e$, the ratio is 59%. Such a discrepancy in peak fields is clearly unsatisfactory and calls for a different geometry. Using a flat plate or shallow dome at the end of the tank and increasing the clearance between it and the terminal from $0.5r_2$ to $0.9r_2$ makes the field at the end of the terminal equal to the cylindrical field when $r_2/r_1 = 2$.

5.2.3 Intershields

In the ideal cylindrical geometry, $E(r_1)/E(r_2) = e$. Much of the insulating gas is therefore working at low stress. Especially in large machines, this suggests the use of an intermediate electrode, or intershield, to make the field more uniform and reduce size and cost. The voltage and field are now related by

$$V_t = V_1 + V_2 = E_1 r_1 \ln(r_2/r_1) + E_2 r_2 \ln(r_3/r_2) \quad (5.4)$$

where V_t is the voltage between terminal and tank, V_1 the voltage between terminal and intershield and V_2 the voltage between intershield and tank; E_1 and E_2 are the radial fields on the outer surfaces of the terminal and intershield, and r_1 , r_2 and r_3 are the radii of the terminal, intershield and tank, respectively. Since the gas conditions inside and outside the intershield are the same, the values of V_1 , V_2 and r_2 should be chosen so that E_1 and E_2 are equal. Whatever the values of E_1 , r_2 and r_3 , it is obvious that the maximum value of V_1 will occur when $r_2/r_1 = e$.

It is convenient to normalize the radii by setting $r_3 = 1$. Equation (5.4) then becomes

$$V_t = V_1 + V_2 = E r_2 / e + E r_2 \ln(1/r_2) \quad (5.5)$$

The maximum value of V_t requires that

$$\partial V_t / \partial r_2 = E[1/e + \ln(1/r_2) - 1] = 0$$

and

$$\begin{aligned} r_2/r_3 &= e^{(1-e)/e} = 0.5315 \\ r_1/r_3 &= e^{(1-2e)/e} = 0.1955 \end{aligned}$$

An intershield of the right radius and voltage increases the maximum voltage by 44% compared with a machine of the same tank radius with no intershield and the optimum terminal radius. But this improvement comes at the cost of a very small terminal. Even a 10 MV machine, assuming $E_{max} = 16 \text{ MV/m}$, would be limited to a terminal radius of 0.25 m. A more realistic value of 0.5 m would correspond to a tank radius of 2.5 m and a terminal voltage of 21 MV.

A better guide to the value of an intershield comes from calculating the voltage gain over the range $0.1995 < r_1/r_3 < 1$. (There can be no advantage in making $r_1 < 0.1995$.) For $0.1955 < r_1/r_3 < 0.368$, the improvement is given by the ratio of V_t to $E_{max}r_3/e$, the voltage in the ideal geometry without an intershield. For the range $0.368 < r_1/r_3 < 1$, the comparison must be with $E_{max}r_1 \ln(r_3/r_1)$. The results are shown in Fig. 5.7. Again, the maximum voltage is not very sensitive to small changes in these quantities.

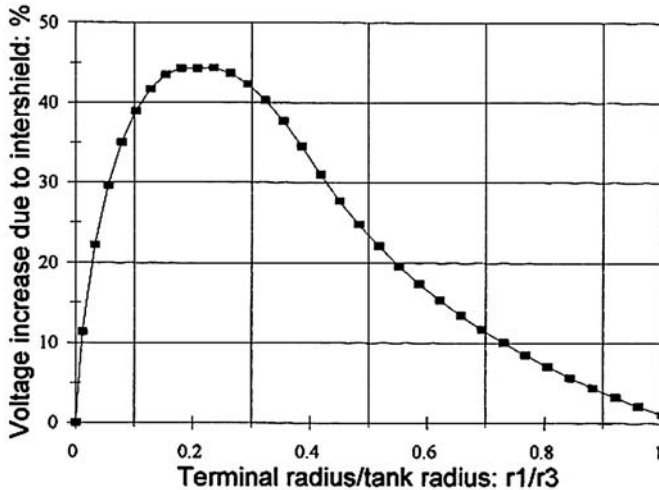


Fig. 5.7. Voltage gain due to a single intershield in cylindrical geometry. For $r_1 < r_3/e$, the gain is relative to the case with no intershield and the ideal geometry ($r_1 = r_3/e$). For $r_1 > r_3/e$, the comparison is with no intershield and the same terminal radius

As the design voltage increases, the economic case for an intershield becomes very strong. For a 20 MV tandem, for example, an ideal intershield would decrease the tank diameter from 3.4 to 2.36 m, the terminal diameter from 1.25 to 0.46 m and the gas inventory by over 50%, from 72 to 35 tonnes. But these cost savings come at a price: access for maintenance is more difficult, additional systems are needed to control the intershield voltage, and the stored energy of the intershield can threaten surge damage to the column.

Some early single-ended machines were designed with multiple intershields, partly in response to the low dielectric strengths of the gases then in use. But the voltage was usually limited by failure of the accelerator tube to work at the required field, and the full advantage was never realized.

Twenty years ago, Letournel [5] introduced the “portico”, a novel design of intershield for tandems, in which the conventional cylinder is replaced by a set of narrow electrodes shaped and placed so as to generate the same radial field pattern. When these have been retrofitted to MP tandems, there has been disagreement as to whether performance has been enhanced. On the other hand, when a single “portico” was fitted to a machine designed for an intershield, a significant voltage gain was achieved [6].

Subsequent three-dimensional finite-element field calculations [7] have revealed that the peak field on the rounded edges of “portico” electrodes is extremely sensitive to small variations in the column gradient, so much so that even small perturbations in the gradient are prone to trigger radial breakdown. When a spark reduces the voltage between portico and terminal to zero, the resulting increase in field on the edges of the portico electrodes is so great that multiple breakdowns occur instantly between tank and portico; see Fig. 5.8. The benefit of a portico is only realized when the column gradient is strictly controlled.

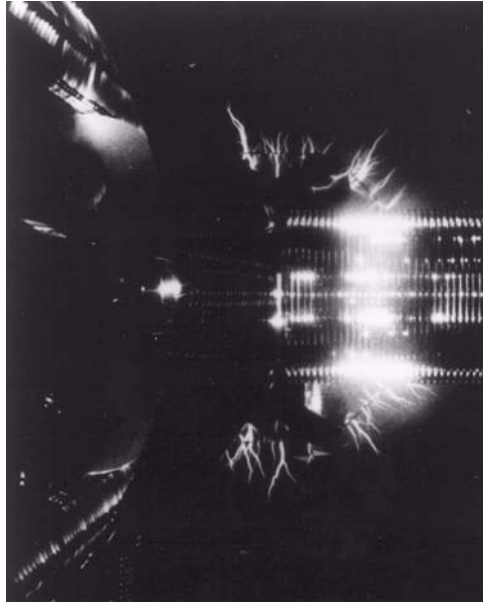


Fig. 5.8. Multiple discharges from the portico electrodes of the Yale ESTU tandem, following a spark between terminal and portico. One of the transverse bracing columns is visible to the right of the main column

5.2.4 The Effect of Terminal Shape

Figure 5.4 shows that the highest field in a simple single-ended machine is near the junction of the column and terminal. A smaller peak occurs where the cylindrical and spherical parts of the terminal join. Finite-element analysis is necessary to determine these fields. A typical value for the peak field at the junction of the cylindrical and spherical sections of the terminal is 15% above the cylindrical value. The field enhancement where the terminal joins the column is caused by the necessary rounding of the end of the terminal and by the small minor radius of the hoops that surround the column. A comparison of the field distributions in several tandem accelerators, using finite-element calculations, was presented by Rabinovitz [8].

Field enhancement at either end of the terminal can be reduced by replacing the abrupt transition from cylindrical to spherical geometry with a more complex shape involving a gradual change in the longitudinal radius of curvature. Replacing a cylindrical terminal with a third-order paraboloid of revolution, Koltay and Kiss [9] improved field uniformity and reduced the peak field enhancement. Their design had the further advantage that the increased radius of the terminal “shadowed” the first few hoops near the terminal so that they no longer carried the peak stress. Complexity and cost have discouraged widespread use of their design, but most large generators have cylindrical terminals greater in radius than the column, thus reducing the peak field at the most vulnerable part of the column.

5.2.5 Hoop Design

The field between adjacent hoops of circular cross section may be obtained from the analytic expression for the field between a pair of infinite cylinders, assuming the hoop minor radius to be much smaller than the column radius. Using the method of images, it can be shown that, for a fixed hoop pitch $2d$ and maximum field E_{max} , the highest voltage between the hoops is achieved when the hoop radius $a = 0.342d$. For this ratio, $V = 0.831 E_{max}d$; see Fig. 5.9.

In a typical accelerator with a column pitch of $2d = 25$ mm and a working field of 16 MV/m, the maximum voltage between adjacent hoops would be 166 kV and the uniform field along the column 6.65 MV/m. But because accelerator tubes are limited to a field of 2.5 MV/m or less, it is preferable to increase the minor radius of the hoops to $a/d = 0.6$ – 0.7 , thus reducing the critical radial field. A further reduction in radial field can be obtained by increasing the hoop pitch to a multiple of the column insulator pitch.

As long ago as 1953, Boag [10] suggested the use of oval hoops to reduce radial-field enhancement. Neglecting the column gradient, he calculated that stress multiplication could be reduced from 1.6 for circular hoops to 1.4 for elliptical hoops when $a/d = 0.66$. Eastham [11], however, has shown that when a realistic longitudinal field is used in finite-element calculations, the

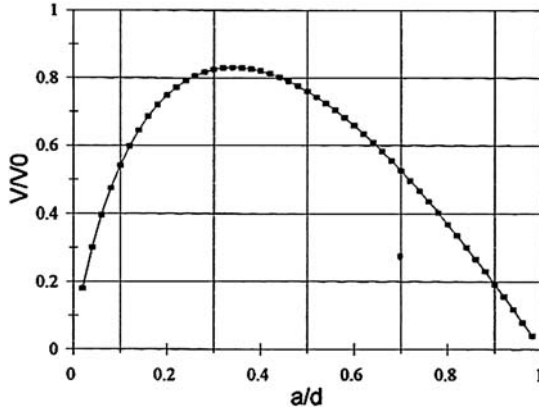


Fig. 5.9. Maximum voltage V between two infinite cylinders of radius a , center separation $2d$, relative to $V_0 = E_{max}d$, where V_0 is the voltage between two parallel planes separated by d and subject to the same limiting field E_{max}

expected improvement is lost. Letournel has fitted a few oval hoops near the terminals of horizontal tandems with intershields, in conjunction with terminals equal in radius to the column. The results are inconclusive. More recently, there has been renewed interest in hoops of noncircular cross section [12], but it is not clear whether the lower field translates into a voltage gain big enough to justify the added cost.

In calculating the fields on hoop surfaces, mechanical tolerances are usually neglected; the hoops are assumed to be truly planar and their supports rigid. Assuming the column gradient to be constant, there is then no net force between adjacent hoops. But in practice hoops are not always absolutely flat, and the supports may allow some movement about the correct position. Variations in the hoop spacing and column gradient then result in unbalanced forces, which may further deflect the hoops, increasing in strength as the displacement increases. In the extreme case hoops may approach close enough to spark or even touch. This is a compelling reason for choosing a pitch such that the hoop spacing is large compared with any mechanical tolerances.

5.2.6 Field Inside the Column

In the interior of an ideal column, there is a uniform longitudinal field. There should be no transverse field. In a real column, the situation is more complicated. Near the hoops, the longitudinal field and the transverse component vary cyclically across each pitch. The presence of electrodes and dead sections, in both the column structure and the accelerator tube, results in the field across the insulators being higher than the average longitudinal field. Variable charges on the belt or chain induce charges on neighboring conductors that result in fluctuating fields; and following breakdown, very large

transient fields, both longitudinal and transverse, travel through the column, leading to secondary sparking and sometimes component damage. All these effects must be taken into account in the column design.

The column structure usually takes the form of a set of two or more legs, each made up of assemblies of insulator disks bonded to thin electrodes. The smaller the column pitch (the distance between each electrode and its neighbor), the greater the field that can be sustained across the insulator surface. In large machines, the hoop pitch is usually a multiple of the column pitch. In some small machines they may be equal. Typical values of column pitches are 40, 25, 20, 12.5 and 10 mm.

A potential divider is required to ensure that the longitudinal field is uniform. This may take the form of a resistor chain or a series of corona points. If the pitches of the column insulators and the tube insulators are the same, and that of the hoops the same or a multiple, a single potential divider may suffice. If the tube pitch is different, a separate divider will be needed. Even if the tube pitch is the same, a separate divider may be desirable to decouple the tube from the column except at the dead sections.

Radial fields inside the column are undesirable. In early machines, every pitch was separated from its neighbor by an equipotential plane. Current practice is to replace these with grading bars that surround the charging system and protect the resistor sticks from surges.

The accelerator tube usually limits the column gradient. This makes it the most critical component in the column. It must be protected from surge damage and decoupled from perturbations that could affect its potential distribution and deflect or defocus the beam. Placing the accelerator tube at the center of the column should minimize the effect of external or radial transients and possibly make it easier to screen. In small machines, however, lack of space may make this impractical. In folded tandems, the two tubes must be symmetrically disposed about the center. Many tandems have operated successfully at very high fields with tubes placed close to the hoops. The actual position seems to be less important than the relative positions of the tube, charging system and potential divider.

In large accelerators, a significant fraction of the column is taken up with dead sections containing mechanical structure, lenses or vacuum pumps. Extending the active length of the accelerator tube into these dead sections can increase the maximum voltage, by as much as 22% for an MP tandem. Care is needed to ensure that adequate clearance is maintained between the tube and the dead section and that suitable grading bars are incorporated to control the field distribution near the dead sections.

Charging systems – belts, chains and ladders (such as Pelletrons and Laddertrons) – are discussed elsewhere, in Chap. 6. They are usually surrounded by grading bars designed so that coupling between the charge they carry and the column components, especially the accelerator tube, is minimized. The field near the charging system is the resultant of the longitudinal field and

the field due to the charge on the belt or chain itself. In the absence of grading bars between the two runs, there will be an attractive force between the opposite charges, and displacement of the belt or chain from a straight line. Belts are often restrained between pairs of grading bars separated by only a few mm. Alternate bars will carry ceramic rods to prevent charge transfer from the belt to the bars. This system overcomes belt flap, but at the cost of increasing surface wear.

5.3 Insulating Gases

The first electrostatic generators were insulated by ambient air, sometimes al fresco. The low dielectric strength of air was known and could be allowed for, but the effects of dirt and humidity were less predictable. Dust, foreign bodies and especially metallic particles can all initiate breakdown. Water vapor actually increases the dielectric strength of air, but water adsorbed by solid insulators leads to volume breakdown and surface tracking. The case for staying indoors, eliminating dust and relying on heating or air-conditioning to reduce moisture was early recognized, as was the fact that insulator performance could be improved by coating permeable materials with special varnishes.

The need for something better than air at one atmosphere was soon apparent. The next development was to increase the gas pressure. Early experiments with dry, compressed air confirmed the expected increase in voltage, but also revealed the benefit of adding electronegative vapor, when a small quantity of CCl_4 was accidentally added to the gas [13]. Subsequent experiments reminded the experimenters of the increased flammability of common materials in compressed oxygen. After a few minor fires, air was replaced with nitrogen to which a small percentage of an electronegative gas such as CO_2 or Freon was added. A mixture of 80% N_2 with 20% CO_2 continues to be a popular choice on the basis of cost, chemical inactivity and environmental acceptability. Freon and chlorofluorocarbons are more effective in trapping electrons but are also more corrosive; their use was declining even before it was prohibited.

The search for more effective gaseous insulators led to the production and use by the power industry of sulfur hexafluoride, a gas that combines the advantages of high dielectric strength, negligible chemical reactivity and nontoxicity. Christophorou [14] has reviewed its insulating properties. Typically used at pressures below 1 MPa, its breakdown strength exceeds that of N_2/CO_2 mixtures at more than twice the pressure. The low operating pressure reduces the cost of the accelerator tank and, coupled with the fact that it can be liquefied at room temperature, the size of the storage system. High cost and environmental regulations require that leakage from the tanks and transfer system must be strictly controlled. Oil-free pumps are required, since oil is soluble in liquid SF_6 and vice versa. Although itself nontoxic,

some of its breakdown products are toxic and reactive, especially when water is present. Fortunately, the gas dryer and circulator needed for successful operation of a high-voltage accelerator are capable of handling this problem and of absorbing and deactivating the toxic breakdown products, such as S_2F_{10} .

The power of SF_6 to quench discharges and prevent breakdown has been extensively studied in the context of its use in power switches and interrupters [15]. These studies do not extend to the very high voltages, long gaps and large surface areas of tandem accelerators. Such measurements as have been made on these machines, usually in haste during the course of commissioning, are often inconsistent, affected by problems unrelated to the properties of clean, dry gas. Nevertheless, the variation of breakdown voltage with pressure, known to be nonlinear, can be estimated from data taken during commissioning and before installation of the accelerator tubes. By relating the pressure not to the voltage but to the field on the terminal, $E = V/(r_1 \ln(r_2/r_1))$, measurements from machines with voltage ratings between 10 and 25 MV and with a variety of terminal shapes and hoop configurations can be compared (Fig. 5.10). The results can be fitted, within a few percent, by the empirical formula $E = 18.6 p^{0.60}$ (E in MV/m, p in MPa). The observed spread in breakdown voltage is less than would be predicted from the range of peak fields occurring in the different designs.

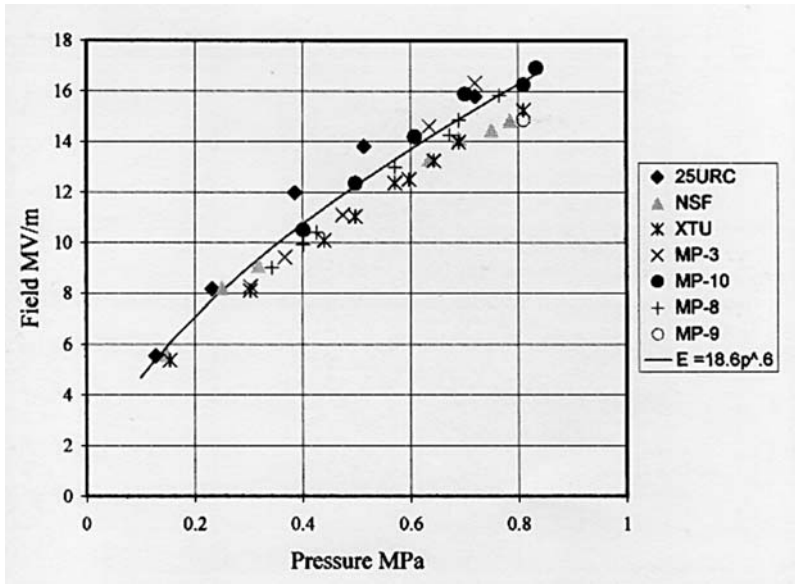


Fig. 5.10. Breakdown field v. pressure of SF_6 for large tandem accelerators

Descriptions of plant for transferring, purifying and drying insulating gas can be found in many reports of accelerator projects, together with accounts of problems experienced in operation, e.g. [16, 17].

5.4 Breakdown and Transient Phenomena

Electrical breakdown in gases is a phenomenon that has been the subject of intense study for over a century. Van Brunt's 1984 bibliography [18] is an important guide to this work. The process is rapid and complex, difficult to measure and sensitive to small changes in gas composition and electrode geometry. Even today, some aspects of breakdown in high-pressure SF_6 are not well understood. Extrapolating the results of laboratory experiments to the voltages and dimensions of large electrostatic accelerators requires a degree of faith and a measure of caution [19].

For the users and designers of electrostatic accelerators, the important questions are:

- (a) what initiates breakdown?
- (b) how do the pressure and composition of the gas affect the critical field at which breakdown takes place?
- (c) what determines the subsequent course of the discharge after initiation?

As the peak field in a gas-insulated high-voltage generator increases, small ionization currents begin to flow. Measurable corona currents are generated by gas multiplication but do not develop into full discharges unless the local field increases above a critical value. Discharge takes place when the ionization coefficient equals or exceeds the attachment coefficient. The critical field at the point of discharge depends on the pressure, on the space charge due to the corona current and on the attachment time, which is very short in SF_6 . A theory proposed by Morrow [20] explains the difference between breakdown in air and in SF_6 in terms of the different values of attachment coefficient and attachment time in these gases. The same model can be used to explain the difference between discharge properties in pure SF_6 and in mixtures with nitrogen and carbon dioxide.

For all these gas mixtures, high voltages can be sustained in the presence of large corona currents. Such currents are induced deliberately when corona stabilizers with sharp needle points are inserted into a region of high field between the tank and the terminal. The current that flows from the needles actually inhibits breakdown. The term "corona stabilization" is applied to this effect. But if the needles are blunt, the current decreases, the field increases, the critical condition is passed and breakdown follows. In a similar way, dust and foreign bodies can also trigger breakdowns. Loose metal objects in contact with a conductor, such as the tank, can become sufficiently charged by induction to levitate, move into a region of high field and initiate

a spark. Horizontal machines especially suffer from this effect, since loose objects tend to migrate to the bottom of the tank below the terminal. Washers, small screws and even wire ends are enough to prevent stable operation. Insulators have much less effect. Dust particles and nonconducting debris acquire smaller induced charges and are less likely to levitate. One may contrast the behavior of well-conditioned accelerators that will run stably, even with large accumulations of high-resistivity belt dust, with the instabilities common in new machines that spark repeatedly at disappointingly low voltages because of contamination with rust, alumina or metallic debris.

Most breakdowns are initiated by particles randomly located near the tank wall and far from the high-voltage electrodes. Their development, as seen in photographs (Figs. 5.2 and 5.8), depends on the subsequent variation in time and space of the field distribution. The primary arc may lead to the nearest conductor, not necessarily to the point of highest static field. Secondary arcs reflect the field distribution at a later time. Loops can form when the field reverses. Machines that have suffered many breakdowns show a pattern of spark marks that extend over the whole terminal and some of the nearest hoops, with a density distribution related to the local field. This is part of the evidence for the area effect that predicts lower breakdown voltages for large electrodes than for small ones [21]. An empirical relation between breakdown voltage and electrode area for a range of SF₆ pressures was given graphically by Aitken and quoted by Joy [22]; see Fig. 5.11.

A spark between terminal and tank results in collapse of the terminal voltage and the flow of a very large current through an arc channel of low resistance and finite inductance. The tank and column act like a transmission line, and the initial voltage collapse is followed by a damped oscillation with a frequency of a few MHz. Attempts have been made to increase the damping by connecting the base of the column to the tank with resistors having the characteristic impedance of the tank and column, regarded as a transmission line. Disintegration of the first set of resistors that were tried in this way at Rochester showed that significant energy could be absorbed, but when more rugged resistors were fitted the overall benefit was marginal.

The energy stored by an accelerator at voltage is proportional to V^2 and to C , the capacity of the terminal and column to the tank. Since C increases with the length of the column, the total stored energy T varies as V^n , where $2 < n < 3$. Typical values for three different accelerators are:

1 MV analytical tandem for RBS	0.09	kJ
6 MV tandem for AMS	5	kJ
25 MV folded tandem for heavy-ion physics	150	kJ.

Putting this into more familiar terms, a tank spark in a 25 MV tandem is equivalent to vaporizing 10 g of aluminum, detonating 10 g of TNT or dropping 1.5 tonnes through 10 meters. Some of this stored energy is dissipated in the arc, some in the walls of the tank and some in the column. The tank is undamaged by the currents which flow through it. The gas recovers, with

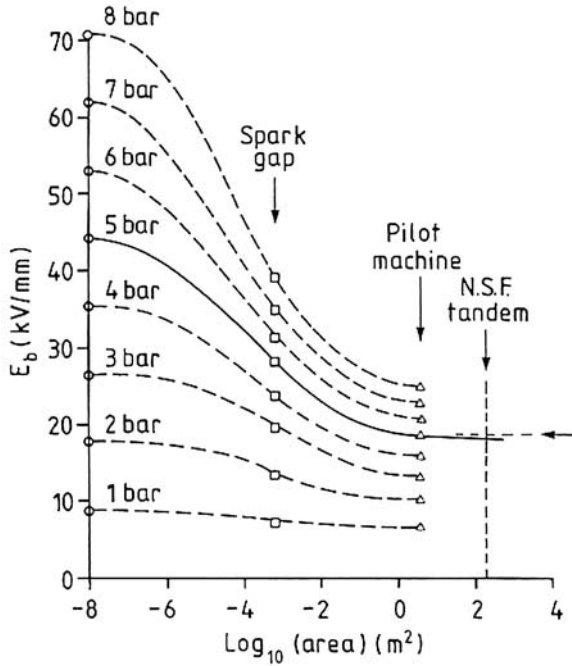


Fig. 5.11. Breakdown voltage v. electrode area at different SF_6 pressures (Reprinted from [22] copyright, (1990) with permission from Elsevier)

some production of unwanted breakdown products. SF_6 has the fortunate property that many of the breakdown products recombine. Sufficient spark energy is dissipated on the terminal or hoops to cause some melting [23], but the resulting craters have a negligible effect on voltage-holding. The dominating concern is the possible damage to column components resulting from the overvoltage that appears between the terminal (or hoop where the spark terminates) and the neighboring hoops.

Even the most cursory glance at Fig. 5.2 suggests that there is a random element in the way sparks develop. The process is also fast and complex, the large transient currents in the arc generating intense electromagnetic fields throughout the tank. There are no instruments capable of measuring these transient voltage distributions along the column. The best that can be done is to make reasonable assumptions about the origin and termination of the arc itself and to combine these with data on the delay times and overvoltages on the column spark gaps in a calculation of the subsequent generation and decay of the electromagnetic radiation. Staniforth [24] has calculated voltage surges in both single-ended and tandem accelerators and has estimated timescales and currents for primary and secondary breakdowns. Such simplified calculations identify the most vulnerable parts of the machine

and underline the importance of transverse oscillations inside the column. They do not yet provide accurate and detailed descriptions of the whole process.

5.5 Protection from Surge Damage

A single component that fails in the terminal of a large tandem will interrupt the experimental program for several days, cost several hundred man-hours and involve the users in tedious repetition of calibrations and checks. Failures in industrial accelerators may have equally serious consequences for critical processes and production schedules. It is difficult to exaggerate the importance of protecting mechanical and electrical components from spark-induced damage.

5.5.1 Insulators

Column insulators, tube insulators, liner supports and electrical standoffs typically operate at voltages below 100 kV. They must be protected from overvoltages high enough to cause volume breakdown, surface tracking or even disintegration. Spherical or annular spark gaps, set to fire at two to three times the maximum operating voltage, are adequate for all but the fastest and most extreme surges, such as can develop in the column at the terminal-to-column and intershield-to-column junctions. Spark gaps must be closely spaced if they are to be fully effective, since the overvoltage rises if a change in direction of the discharge creates an inductive impedance to its path. Annular gaps do not suffer from this disadvantage, but the theoretical improvement in protection is not observed in some laboratory experiments. They do make cleaning and examination of the insulators more difficult.

The relatively low field along the column makes it unnecessary to convolute the external surfaces of column or tube insulators when operating in dry, compressed gas. The inner surfaces of tube insulators can benefit from tapering or convoluting, which helps to control surface charging and ion bombardment and thus improves voltage withstand. However, some convoluted surfaces are prone to physical damage if spark gaps fail to protect against fast-rising surges.

The dielectric constant of ceramics and glasses lies between 5 and 10. Bubbles and voids between insulators and electrodes therefore generate high fields, much greater than the average field in the column. If volume breakdown occurs in insulators, it is very likely to be initiated at these faults.

Many accelerators use long ungraded insulators as drive shafts, control rods and optical fibers for data transmission. A few designs rely on tension rods or plastic plates to support the column. The VIVITRON at Strasbourg depends on molded epoxy-resin posts to support the intershields and column.

With proper design, shafts and rods have been made to work reliably while sustaining voltages up to 5 MV. The choice of material is important. It must be homogeneous, free from defects, not hygroscopic and with consistent electrical characteristics. Methacrylates and polycarbonates are among the plastics of choice. Epoxy resins, loaded with silica or alumina flour, have been used where strength is critical. Glass-fiber-reinforced resins are also used where the anisotropy introduced by the fibers can be tolerated.

The most critical design feature in these components is the connection between the ends of the insulators and the metal structure. Screws and pins must be deeply recessed inside shaped electrodes surrounding the rod or plate. These electrodes must be thick and fully radiused so that the point of contact between insulator and metal is in a low-field region. The insulators must be located away from the spark gaps or other paths taken by discharges during surges. Surface charging must be minimized, and for this reason tubes and sheathed fibers are less reliable than solid rods and monofilaments. In some machines it has been found best to drape fiber-optics along the surface of grading bars or hoops, presumably thus ensuring a controlled field along the surface.

The use of post insulators is controversial. In particular, designs that rely on an internal boss to lower the field at the insulator/electrode junction cannot be relied on to survive surge overvoltages. When failure does occur, it may take the form of an undetectable track through the interior of the insulator, begun during a surge and developing during normal operation either to explosive disintegration of the post or to the point where the leakage current prevents continued operation. It is usually impractical to surround a post designed for several megavolts with protection in the form of a cylindrical metal shield and an annular spark gap.

Charging systems must support the full terminal voltage across their length. Grading may be possible at dead sections where they run over idler rollers or pulleys. In between dead sections, the normal practice is to surround them with grading bars that have the dual function of controlling coupling to the beam and shielding them from transient fields.

Pelletrons and Laddertrons are self-protected because the space between links acts as a spark gap for the insulating link inside. Belts do not have this protection. They usually run in a narrow slot between pairs of grading bars. In this geometry, the longitudinal field is closely controlled and belts that have been properly dried rarely suffer surge damage.

5.5.2 Electrical Components

Resistors or corona points acting as potential dividers are inevitably exposed to the field in the column. Corona points turn into spark gaps when subjected to large overvoltages; the result is to evaporate material from the points and alter the voltage/current characteristic. Resistor specifications rarely allow for overvoltages as large as those seen during tank sparks; when they occur,

resistor values change and, in the end, resistors shatter or go open-circuit. Protection is vital.

Successive levels of resistor protection can be achieved by placing the assemblies between grading bars or equipotential plates; by radiusing the element end caps to prevent surface tracking; by surrounding the end of each resistor with a metal tube to lower the transient field; by incorporating a small inductance between the resistor and the grading bar; and, eventually, by surrounding each resistor string with a metallic tube, one end of which acts as an annular spark gap. Resistors using rare-earth inks on a solid ceramic substrate have survived years of operation undamaged at very high voltages with this degree of protection.

The protection of high-voltage power supplies feeding exposed components such as belt-charging screens, Pelletron inductors and electrostatic lenses requires circuits capable of attenuating surges with rise times of a few nanoseconds and amplitudes of hundreds of kilovolts. The circuits themselves must be immune to damage from the surges. The solution proposed by Johnstone [25] is to use two coaxial RC filters in series, the first preceded by a cylindrical spark gap and consisting of a low, noninductive resistance and a gas-insulated capacitor. This slows down the rising edge of the surge to the point where a second stage with a longer time constant can be used without risk of damage. A smaller version of this device can be used to protect conditioning circuits for column currents, takeoff currents and similar data inputs.

Motors and power supplies operating at line voltage can be protected with conventional screening and semiconductor surge limiters. Power and other leads must be routed so that they do not couple inductively to the high currents flowing during sparks. Connectors and wires need to be well recessed into the terminal or below the column base so as to avoid transient fields that can penetrate into these Faraday cages through apertures.

The most vulnerable components in an electrostatic accelerator are undoubtedly the semiconductors in the control and data transmission systems. Surge voltages in the terminal must be attenuated from megavolts down to volts in order to protect them. Two stages of screening are required to achieve this. One satisfactory design consists of nested aluminum boxes made of thick plate, ventilated by holes with a length-to-diameter ratio of 3:1, with access panels clamped over conductive gauze. In-line filters are fitted to every conductor where it enters or leaves either box. Power lines are fully screened and provided with surge filters before and after entry. Signal and output cables are contained within two coaxial screens, usually made of flexible bellows. Care is taken to avoid multiple earths and to connect the external screens with robust contacts, such as Conflat flanges. With such precautions, the record of reliable operation is impressive, many machines operating for years without a single surge-induced component failure.

References

1. J.D. Cockcroft, E.T.S. Walton: Proc. Roy. Soc. **136**, 619 (1932)
2. D.A. Bromley: Nucl. Instr. Meth. A **122**, 1 (1974)
3. R.G. Herb, C.M. Turner, C.M. Hudson, R.E. Warren: Phys. Rev. **58**, 579 (1940)
4. T.J.L. Greenway et al.: Oxford University report, NPL 37/74 (1974)
5. M. Letournel et al.: Nucl. Instr. Meth. **220**, 10 (1984)
6. H.R.McK. Hyder et al.: Voltage tests of the Yale ESTU with portico. In: *Proceedings of the Symposium of North Eastern Accelerator Personnel, SNEAP XXI*, ed. by K.R. Chapman, Tallahassee, FL, (pub. World Scientific), p. 57 (1987)
7. Y. Thiery et al.: Nucl. Instr. Meth. A **378**, 21 (1996)
8. I.I. Rabinovitz: Proc. 1st Int. Conf. Tech. Electrostatic Accelerators, Daresbury Nuclear Physics Laboratory, DNPL/NSF/R5, p. 179 (1973)
9. E. Koltay, A. Kiss: Proc. 1st Int. Conf. Tech. Electrostatic Accelerators, Daresbury Nuclear Physics Laboratory, DNPL/NSF/R5, p. 200 (1973)
10. J.W. Boag: Proc. IEE IV **100**, 63 (1973)
11. D.A. Eastham: Nucl. Instr. Meth. **108**, 593 (1973)
12. K.A. Rezvykh, V.A. Romanov: Increase of the reliability of an electrostatic accelerator at the highest possible operating potential. In: *Proceedings of Symposium of North Eastern Accelerator Personnel, SNEAP XXXIV*, ed. by R. Hellborg et al. (Lund University, Lund, Sweden, 2001) pp. 89–103
13. D.B. Parkinson, R.G. Herb et al.: Phys. Rev. **53**, 642 (1938)
14. L.G. Christophorou: Nucl. Instr. Meth. A **268** 424 (1998)
15. A.H. Cookson: Proc. IEE A **128**, 303 (1981)
16. K.F. Minati et al.: IEEE Trans. Nucl. Sci. **16**, 109 (1969)
17. R. Hellborg, K. Håkansson: Nucl. Instr. Meth. A **235**, 407 (1985)
18. R.J. Van Brunt: NBS Technical Note 1185 (National Bureau of Standards, Washington, DC, 1984)
19. W. Pfeiffer: Nucl. Instr. Meth. A **220**, 63 (1984)
20. R. Morrow: Nucl. Instr. Meth. A **382**, 57 (1996)
21. C.M. Cooke: Electrode surface effects on large gap breakdown in SF₆. In: *Conference Record of the 1982 IEEE International Symposium on Electrical Insulation*, Philadelphia, PA, (pub. IEEE, New York), p. 215 (1982)
22. T. Joy: Nucl. Instr. Meth. A **287**, 48 (1990)
23. A.N. James: Nucl. Instr. Meth. A **220**, 96 (1984)
24. J.A. Staniforth: Nucl. Instr. Meth. **216**, 1 (1983) and Nucl. Instr. Meth. A **220**, 93 (1984)
25. W.T. Johnstone: Nucl. Instr. Meth. **131**, 549 (1975)

Box 1: Calculation Technique for High-Voltage Equipment in Gas

K.A. Rezvykh¹, V.A. Romanov¹, and R. Hellborg²

¹ State Scientific Center of the Russian Federation, Institute for Physics and Power Engineering, 1 Bondarenko Sq., Obninsk, Kaluga Region, 249033, Russia
rezvykh@ippe.obninsk.ru, romanov@ippe.obninsk.ru

² Department of Physics, Lund University, Sölvegatan 14, 223 62 Lund, Sweden
ragnar.hellborg@nuclear.lu.se

Introduction

The analysis of an electrostatic field and the calculation of breakdown voltage are demonstrated here. A reasonably good prediction can be obtained by a technique developed in Obninsk [1, 2]. The high accuracy (errors within $\pm(2-3)\%$) is a consequence of the physical model used. This model is called “asymptotic breakdown gradient” (ABG) or “the method of base”. The high accuracy is also a consequence of the calculation of the electrostatic field on curvilinear electrode boundaries with correction for the systematic errors in computing a potential gradient [3].

Elementary Phenomenon

When we solve the problem of the electric strength of a gas, the failure of insulation is considered in the form of a spark breakdown or the beginning of a corona discharge. Electric strength is expressed in terms of a statistical average breakdown voltage $\bar{U}$ and the standard deviation of the breakdown voltage σ . The elementary phenomenon is defined exactly.

Elementary Structure

An element of an insulation structure is unambiguously formed by the electrostatic field on the surfaces of the electrodes, and it is characterized by four parameters. An element of the insulation system represents a part of the surface of the electrodes adjoining a point where the potential gradient is equal to its maximum value E_{max} and where the radius of average curvature of the surface has the value R_{av} . The area of an element is limited to the “effective” surface S_{eff} , where the strength of the field, that is, the gradient of potential with the sign reversed, is given by $E \geq 0.8 E_{max}$ [2]. Each element has a value of the breakdown voltage U_{br} . It is also necessary to define two important parameters: the breakdown voltage of the insulation system and the relative strength of its element [1].

The Theoretical Model

A theoretical model of electric breakdown in a gas needs to be based on some sort of affirmation. The “law of similarity of gaseous discharge” (Paschen’s law, in a uniform field) could be such a fundamental and conclusive statement [4]. In a nonuniform field, the law of similarity can have the form

$$E_{br} = E_{uni} (1 + a_{uni}/(p_{20}R_{av})^{m_{uni}}) , \quad (B1.1)$$

where E_{br} and E_{uni} (both in MV/m) are the breakdown gradients in nonuniform and uniform fields, respectively. $a_{uni}=0.061$ and $m_{uni}=0.38$ are constants valid for air and N_2/CO_2 gas, and $a_{uni}=0.0045$ and $m_{uni}=0.54$ are constants valid for SF_6 gas. p_{20} is the gas pressure adjusted to its value at $20^\circ C$, and R_{av} is the radius (in m) of the average curvature of the electrode surface,

$$R_{av} = 2/(1/R_{k1} + 1/R_{k2}) , \quad (B1.2)$$

where R_{k1} and R_{k2} are the principal radii of surface curvature at a point where $E = E_{max}$. Below, the theoretical model of electric breakdown in a gas is outlined:

- Electric strengths of a system and of its elements are estimated from an end result, breakdown voltages. The breakdown voltage is a linear function of the breakdown gradient E_{br} :

$$U_{br} = E_{br}(U_{calc}/E_{max}) , \quad (B1.3)$$

where E_{max} denotes the maximum gradient of the potential on an electrode surface at a potential difference of U_{calc} in the calculation between the electrodes.

- For increasing distance between the electrodes in a uniform field, the breakdown gradient of the potential approaches asymptotically the value E_{asm} .
- This asymptotic breakdown gradient can be calculated from the data of the preliminary base experiment, if two conditions are fulfilled. Firstly, the chemical composition of the gas and the way of conditioning the gaseous insulation and also the electrode material, the finishing of the surface, and the way of preparation of the electrodes to be tested should be identical in the base structure and in the designed structure. Secondly, the maximal gradients of potential in both the structures may not differ by more than

$$0.1 \leq (E_{max}/U_{calc})/(E_{max.base}/U_{calc}) \leq 10 . \quad (B1.4)$$

- As the criterion for the electric strength in the calculations, the parameter “breakdown voltage of a base gap” $U_{br.base}(p_{20})$ is used instead of conditions for self-maintenance of a discharge obtained for some abstract gaps with a uniform field or nonuniform field [5].

The Calculation Procedure and its Verification

The technique has been tested with experimental results obtained for a 3 MV Pelletron tandem with pure SF₆ [6]. For this accelerator, the following value is valid at 0.6 MPa: $U_{exp}(0.6) = 3.673$ MV, as

$$U_{exp} = 5.085(p_{20})^{0.637} \text{ MV}, \quad \sigma = 2.9\%, \quad p_{20} = 0.21 - 0.6 \text{ MPa}. \quad (\text{B1.5})$$

The starting point of the calculation is the computation of the breakdown voltage value obtained from the base experiment on MP tandems [7], $U_{br.base}(0.6) = 13.663$ MV, as

$$U_{br.base} = 19.757(p_{20})^{0.722} \text{ MV}, \quad \sigma = 2.9\%, \quad p_{20} = 0.3 - 0.85 \text{ MPa}. \quad (\text{B1.6})$$

The accuracy of (B1.6) has been checked [2]. The inequality (B1.4) is fulfilled, and therefore the choice of base is valid (with $E_{max.base}/U_{calc} = 2.107$ MV/m, $E_{max}/U_{calc} = 7.47$ MV/m, according to Fig. B1.1). The normalized voltage [1] with 100% SF₆ gas, $k_{rel} = 1$, positive polarity, $k_{pol} = U_{br}/U_{br.pos} = 1$, and 100% of the conditioning, $k_{cdtn} = U_{br}/U_{br.stbl} = 1$, is given by

$$U_{norm.base} = U_{br.base}/(k_{rel}k_{pol}k_{cdtn}) = 13.663 \text{ MV}. \quad (\text{B1.7})$$

The breakdown gradient in the nonuniform field of the base insulation gap is, according to (B1.3) (with $E_{max.base}/U_{calc} = 2.107$ MV/m),

$$E_{br.base} = U_{norm.base}(E_{max.base}/U_{calc}) = 28.788 \text{ MV/m}. \quad (\text{B1.8})$$

The breakdown gradient in a uniform field, according to (B1.1) with $R_{av.base} = 0.0376$ m, is given by

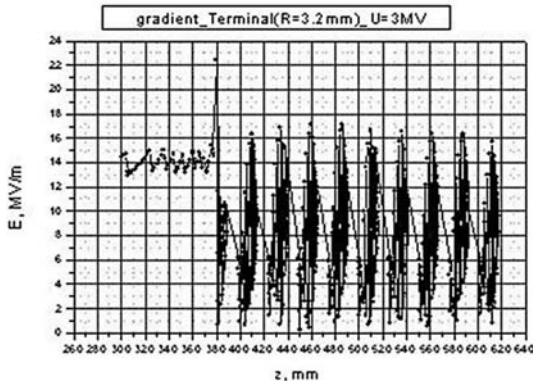


Fig. B1.1. The distribution of the potential gradient along the surfaces of the terminal ($R_{k1} = 3.2$ mm) and of the hoops of the column ($R_{k1} = 8$ mm) for a 3 MV Pelletron. The terminal potential is equal to 3.0 MV. The coordinate z is counted from the middle of the terminal

$$E_{uni.base} = E_{br.base} (1 + a_{uni}/(p_{20}R_{av.base})^{m_{uni}})^{-1} = 27.818 \text{ MV/m} . \quad (\text{B1.9})$$

The asymptotic breakdown gradient in the base experiment, according to (B1.5) in [2] with $L_{base} = 1.8 \text{ m}$, $a_{asm} = 0.000004$ and $m_{asm} = 1$, is given by

$$E_{asm.base} = E_{uni.base} (1 + a_{asm}/(p_{20}L_{base})^{m_{asm}})^{-1} = 27.818 \text{ MV/m} . \quad (\text{B1.10})$$

The asymptotic breakdown gradient, according to (2) in [2] in the 3MV Pelletron structure, is

$$E_{asm} = E_{asm.base} (1 - ((\sigma/\bar{U})/\sigma_N) \ln(S_{eff}/S_{eff.base})) = 31.791 \text{ MV/m} , \quad (\text{B1.11})$$

with $S_{eff} = 0.01 \text{ m}^2$, $S_{eff.base} = 1.42 \text{ m}^2$, $\sigma/\bar{U} = 0.0344$, $\sigma_N = 1.1938$ and $N = 80$. The breakdown gradient in a uniform field for the Pelletron, with $L = 0.416 \text{ m}$, is given by

$$E_{uni} = E_{asm} (1 + a_{asm}/(p_{20}L)^{m_{asm}}) = 31.791 \text{ MV/m} . \quad (\text{B1.12})$$

The breakdown gradient in a nonuniform field, with $R_{av} = 0.00635 \text{ m}$, is given by

$$E_{br} = E_{uni} (1 + a_{uni}/(p_{20}R_{av})^{m_{uni}}) = 34.677 \text{ MV/m} . \quad (\text{B1.13})$$

The breakdown normalized voltage, using $E_{max}/U_{calc} = 7.47 \text{ m}^{-1}$, according to Fig. B1.1, is given by

$$U_{norm} = E_{br}/(E_{max}/U_{calc}) = 4.644 \text{ MV} . \quad (\text{B1.14})$$

For the breakdown voltage using $k_{rel} = 1$ and $k_{pol} = 1$, and considering two values for the coefficient of conditioning $k_{cdtn} = 1$ and $k_{cdtn} = 0.8$ as the real coefficient is unknown, we have

$$U_{br} = U_{norm} (k_{rel}k_{pol}k_{cdtn}) = 4.644 - 3.715 \text{ MV} . \quad (\text{B1.15})$$

Comparing the results from (B1.5) and (B1.15) gives

$$\delta U = U_{br}/U_{exp} = 1.26 - 1.01 . \quad (\text{B1.16})$$

The ratio of 1.26 between the calculated and experimental values is probably a result of an incomplete conditioning of the system. The tandem in Lund is always conditioned without sparks (defined in this calculation to be incomplete), leading to a coefficient of conditioning $k_{cdtn} = 0.8$ in SF_6 . The results of the calculation show also that the column is not a weak element (as was supposed in our preliminary calculations [1]) but that the terminal is. Hence the calculation technique has a satisfactory accuracy of the prediction of the breakdown voltage if a careful numerical calculation of the field is carried out.

References

1. K.A. Rezvykh, R. Hellborg: Electric strength of the gas insulation of the 3UDH Pelletron accelerator at Lund university (Part 1. A technique of the calculation). In: *Proceedings of Symposium of North Eastern Accelerator Personnel, SNEAP XXXIV*, ed. by R. Hellborg et al. (ISBN 91-631-2676-1, 2002) pp. 124–140
2. K.A. Rezvykh, V.A. Romanov: Nucl. Instr. Meth. A **423**, 203 (1999)
3. K.A. Rezvykh, V.A. Romanov: High-voltage accelerators: an analytical estimation of systematic errors of computing a gradient of potential for a finite difference method. In: *Proceedings of the XVIII Conference on Accelerators of Charged Particles, RUPAC-2002*, ed. by V.A. Romanov, SSC RF – IPPE, Obninsk, Russia, 2004, v. 2, pp. 798–810; see also: *Electricity*, Moscow (2004) (6) pp. 17–26 (in Russian)
4. I.M. Bortnik, C.M. Cooke: IEEE Trans. **PAS-91**: 5, 2196 (1972)
5. J.M. Meek, J.D. Craggs: *Electrical Breakdown of Gases* (Clarendon Press, Oxford 1953)
6. R. Hellborg: Nucl. Instr. Meth. A **379**, 185 (1996)
7. S.J. Skorka: Rev. Phys. Appl. **12**:10, 1279 (1977)

6 Charging Systems

C. Westerfeldt

Duke University Physics Department, Triangle Universities Nuclear Laboratory,
Durham, NC 27708-0308, USA
`cwest@tunl.duke.edu`

6.1 Introduction

Charging systems for small electrostatic accelerators fall generally into three different categories: belt-charged, chain-charged and cascade. The belt and chain charging systems are mechanical in nature and can be scaled to multi-megavolt-sized machines, while the cascade system (see Box 3) employs a high-frequency solid-state high-voltage multiplier circuit that has a practical limit of a few MV. Belts are the oldest technology for producing high voltage, and the earliest accelerators were named after the father of this technology – Dr. Robert Van de Graaff of Princeton University. Dr. Van de Graaff received a patent for this technology in 1935. Belts were relatively simple to construct and were able to carry up to 1 mA of charge on their outer surface – more than enough for most applications. They are used in electrostatic accelerators produced by HVEC, which was cofounded by Robert Van de Graaff, Denis Robinson and John Trump in 1947. In the 1960s, Dr. Ray Herb at the University of Wisconsin developed a new mechanical charging system. It consisted of a chain made up of stainless steel cylinders, coupled together with insulating nylon links. These chains were capable of carrying only up to 150 μA each but with greatly improved charging stability. Multiple chains are installed in accelerators requiring greater charging capacity. This system is used in the “Pelletron” accelerators manufactured by NEC.

6.2 Belt Charging Systems

6.2.1 Physical Description

The earliest belts were of simple uncoated-cotton construction. Cotton was chosen for both its electrical and its mechanical properties. It readily accepted a static charge and did not break down in the large electric-field gradient along the accelerator column. Cotton belts also were quite strong and could be used to drive an electric generator in the terminal of the accelerator. Once tensioned, they also typically did not stretch very much with time. Typical modern charging belts manufactured by HVEC consist of a cotton multi-layer carcass, which has been coated with a vulcanized rubber material inside

and out. Belts are produced in narrow (15 cm wide) widths for accelerators that operate up to 1 MV and in wide (52 cm) widths for accelerators up to ~ 20 MV. In the case of a charging belt, the primary reason for variations in the charging current delivered to the terminal is the inhomogeneity of the surface of the belt. The belts are hand made and the rubber coating is vulcanized in sections, producing a variation in thickness where the sections overlap. The charge is applied to the belt via a high-voltage biased screen or shim which contacts the belt at the grounded end of the column. Fig. 6.1 illustrates the principle.

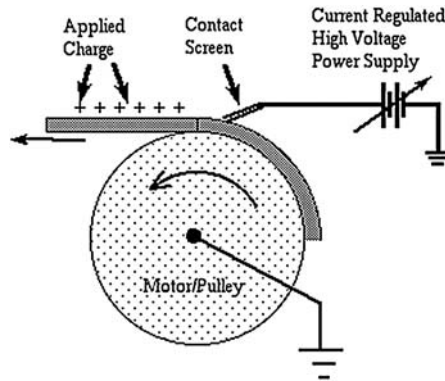


Fig. 6.1. A simplified belt charging system

The belt passes around a grounded pulley at the base of the accelerator. This pulley typically is actually an inverted motor – the armature is stationary and the outer housing rotates. A metal screen or thin shim is pressed lightly onto the surface of the belt that is in contact with the grounded pulley. The screen is, ideally, not perpendicular to the belt but at an acute angle so that wear in the screen does not produce a gap between the screen and the belt. The screen is mounted on an insulating fixture that will withstand up to 50 kV from the charging power supply under pressurized conditions. A current-regulated high-voltage power supply is connected to the charging screen and adjusted to apply the desired charge to the moving belt – up to ~ 1 mA. The charge can of course be positive or negative. Positive charge will produce a positive potential on the accelerator terminal, and this is used for producing positive-ion beams. Negative charge is employed in electron accelerators. The charging screen is usually constructed of stainless steel wire mesh with between 1.5 and 3 wires per mm. The edge of the screen that contacts the belt typically has two transverse wires removed, leaving 1.5–2 mm of wire extending to contact the belt. This screen edge presents a series of sharp corona points to the outside surface of the moving belt at the drive pulley location. The screen is adjusted so that it contacts the belt lightly along the

entire length of the screen to minimize wear on both the belt and the screen. The screen is not as wide as the belt and should be centered on the running belt. Because the pulleys typically are crowned, the belt will change position between rest and running conditions. It is important that the charging screen or shim not be directly exposed to the drive motor during running conditions or an electrical arc will occur – damaging the charging supply and possibly the motor bearings.

In the terminal, a similarly prepared screen which is electrically connected to the terminal contacts the belt ahead of the terminal pulley and removes the charge from the moving belt, transferring it to the terminal. Fig. 6.2 depicts a simple belt charging system for a single-ended electrostatic accelerator.

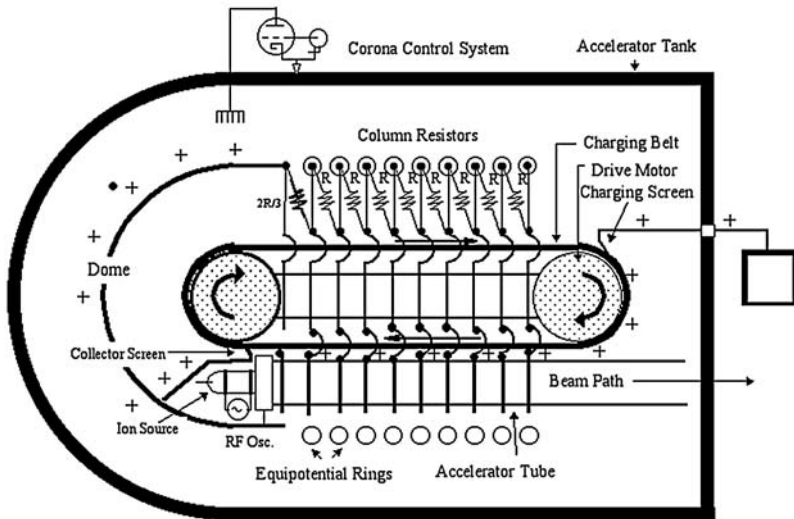


Fig. 6.2. A simple belt-charged accelerator

The DC voltage required to induce sufficient charging current on the surface of the belt is typically 5–10 kV for 100 to 200 μA of charging current. As the belt is not uniform in thickness, the charging and collector screens cannot maintain contact with the insulating rubber surface as it passes by the screen at speeds of 20–25 m/s. The voltage on the terminal will typically exhibit a pattern, which repeats every revolution, with an uncontrolled voltage fluctuation of several kV peak-to-peak, or about 0.1% of the DC terminal voltage. Figure 6.3 shows the capacitive pickoff (CPO) pattern from a model K-3000 belt-charged electrostatic accelerator.

One can observe that the pattern repeats every revolution of the belt. Also evident are about ten large fluctuations, which correspond with the ridges on the belt from the overlapping vulcanized rubber strips contacting the charging screen, causing it to bounce and not collect charge immediately after the ridge

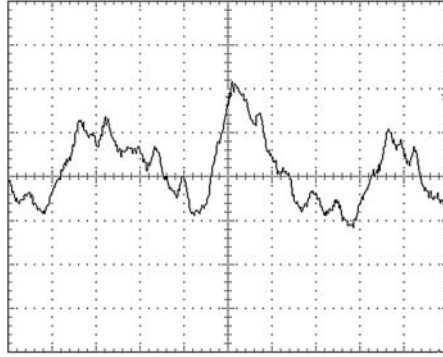


Fig. 6.3. The CPO pattern from a typical belt-charged electrostatic accelerator

passes. For this reason, some laboratories have developed modified screen systems, which utilize multiple screens, or thin metal shim stock in place of screens, to minimize these effects. These multiple screens have independent mechanical mounts that attempt to maintain each small screen in contact with the moving belt, and also independent current-regulated power supplies. These systems – while more complicated mechanically and electrically – can substantially reduce the unregulated voltage ripple from the belt charging system.

6.2.2 Maintenance

The belt charging system requires regular attention to maintain the best charging efficiency and minimum voltage ripple. The primary maintenance item is the tension in the belt. Charging belts stretch with time and as they do, the tension is reduced. This can result in excessive flapping along the column, producing excessive amounts of belt dust and occasionally damage to the belt guides. In machines with a vertically mounted belt, the reduced tension results in the belt dropping on the motor and alternator while it is running (belts normally park high when stopped). In severe cases, the belt can actually drag on the lower column crossbars, resulting in damage to the belt and to the aluminum bars. These events are typically very noticeable in the motor current and by the loss of charge along the column, as when the belt drags it dumps charge. In tandem accelerators, there are typically tension meters installed to permit the monitoring of the upper and lower drive motor tensions from outside the tank. A drop in tension can be compensated for without a tank entry by making adjustments to the motor mounts through pressure-sealed adjusting bolts. A viewer is provided to allow the technician to observe the edges of the belt on the motor. In a single-ended accelerator, adjustments require the tank to be vented and removed. The terminal shell is removed, and the several lock screws on the terminal alternator are loosened.

Typically, a hydraulic jack is provided by the manufacturer to measure the tension on each end of the terminal alternator. The jack can be used to set the tension on the lower pillow block using a manufacturer-recommended pressure reading on the gauge supplied. For an HVEC belt, the recommended tension is approximately 75 N/cm of belt width. The lower tension in a vertically installed 52 cm wide belt would therefore be between 3500 and 4000 N. The upper tension is set approximately 20% lower initially (the slight pitch is necessary to provide a vertical restoring force sufficient to offset the weight of the belt). The drive motor is started momentarily, and the position of the belt is observed. If the belt rises, the tension on the upper pillow block must be increased. If the belt drops precipitously, the upper tension should be reduced further. After several iterations, the belt should then be observed to be fairly stable. The motor can be started and allowed to attain full running speed while fine adjustments to the tension are made to prevent the belt from dropping too low on the alternator. The belt should then be stopped and the locking screws tightened. Note: do not operate the motor for more than twenty minutes in air or it will overheat. The motor is designed to operate continuously only in the pressurized environment inside the accelerator.

After several thousand hours of operation, there may be heavy deposits of belt dust accumulated on the column, which can result in instabilities along the column. To restore stable operation requires removing the belt, outer belt guides and gradient bars and perhaps also the equipotential rings. The column is cleaned with ethanol and lint-free rags. The belt can also be cleaned with unleaded (white) gasoline. This will remove contaminants such as dust and grease from the surfaces of the belt. The belt is reinstalled and tensioned as discussed previously.

6.2.3 Troubleshooting

One of the most useful tools for diagnosing problems internal to an operating accelerator is a well-calibrated CPO system. The CPO is calibrated in air (it has been demonstrated that the calibration is insensitive to pressurized tank gas) with a signal generator or simply the output from a small transformer driven from the mains via a variac in the primary. If one uses a signal of $85 V_{rms}$ as measured with a typical laboratory meter, this will put approximately 100 V peak-to-peak on the terminal of the accelerator if one lead of the secondary winding is connected to the tank. The output of the CPO circuit is observed on a monitor oscilloscope, and either a note of the calibration constant is made or, if adjustment of the CPO amplifier gain is possible, the gain can be adjusted to a convenient value, i.e. $1 V_{pp}$ (on the scope display) = $100 V_{pp}$ (on the terminal). With this tool, one can diagnose small discharges in the column grading resistors, poor performance of a belt due to mechanical properties of the belt, and charging-screen damage. The terminal voltage ripple can be constantly displayed on an oscilloscope at the control console.

In the case of a chain, one can usually see the actual pellet ripple and detect problems in the chain(s) by the frequency of the abnormal signal. In the case of a belt, one can usually detect the strips on the belt where the rubber overlaps and can detect areas of the belt that are not accepting charge as efficiently as the rest.

6.2.4 Troubleshooting Belt Charging Systems

The belt charging system can be divided into four major parts: the external charging power supply and controls, the mechanical charge transfer screens, the charging belt, and the belt guides in the column structure. In order to determine whether a fault lies in the external power supply or controls, one connects a high-value resistor between the output of the charging power supply and a suitable ground. The power supply is energized and the charging current increased to a normal level. If the current is stable then the power supply is working correctly; if not, it will need to be serviced. Typically, these power supplies should current-regulate to at least 0.01%.

The next things to check are the high-voltage feedthrough on the tank, and the charging and collector screens. The feedthrough can be checked with a high-voltage megohmmeter (5000 V DC recommended) for leakage. There should not be any detectable leakage ($R > 2000 \text{ M}\Omega$). The charging screen should have a low-resistance connection to the feed through ($R < 10 \Omega$), and the screen or shim should be lightly contacting the belt all along its exposed edge. If a screen is used, there should be at least one longitudinal wire removed so that there is about 1 mm of wire ends exposed to the belt. The charging screen or shim is made narrower than the belt so that in operation, the screen is not exposed to the grounded drive motor pulley. This would cause charge loss by corona to the motor, and little charge would get onto the belt. The terminal collector screen should be inspected in the same way and adjustments made to insure sufficient but not aggressive contact with the rotating belt. The collector screen should be wider than the belt to insure collection of all of the charge on the belt.

If charge loss from the moving belt is suspected (an abnormally large charging current is required to attain the required terminal voltage), one can connect the column to ground through a microampere meter, and with the terminal grounded check for charge collection at points along the column. Once located, the belt guides in that area can be inspected for proper adjustment or for contamination that is resulting in charge collection at that point. The parts should be cleaned or replaced as necessary. The belt guides (inner and outer) are normally adjusted to attain a clearance of 1.4 mm from the belt. The inner guides are typically mounted onto the column stiffener plates and are not therefore adjustable. If the clearance between the belt and the inner guides is not the same on each side of the column or from top to bottom, the terminal pulley must be moved with respect to the column by

means of adjusting screws in the terminal weldment to attain proper clearance. The outer guides are installed one at a time, and the gap to the belt can be adjusted by loosening the locking screws on the end fittings and sliding them to achieve the 1.4 mm gap. The gap between belt guides should nominally be 5.6 mm, while the gap between the gradient bars (no ceramic insulator installed) should be a nominal 8.75 mm. Gradient bars and belt guides are installed in alternating fashion on each side of the belt throughout the column. The outer guides should be removed and cleaned with a petroleum distillate to remove contamination from the belt and grease whenever they are observed to be excessively dirty. Stubborn stains can be removed from the ceramic on the belt guides using a mild abrasive such as an ink eraser.

Occasionally, the bearings in the motor and alternator will need to be replaced. This will be evident if grease is observed to have contaminated the inside of the belt near either edge. When this is discovered, the belt must be removed and thoroughly cleaned to remove the grease. Continued operation can result in electrical failure of the belt by an energetic spark running the length of the belt. This frequently damages the belt, requiring it to be replaced or the damaged edge trimmed.

6.3 Chain Charging Systems

6.3.1 Introduction

The first chain charging system was developed at the University of Wisconsin in the early 1960s by James Ferry in association with Professor Ray Herb in the Physics Department. In 1965, they founded NEC and began producing Pelletron accelerators. The chains manufactured by NEC are illustrated in Fig. 6.4.

The chain is constructed of 31.75 mm diameter stainless steel tubing that is cut and has the ends rolled inwards to form 31.75 mm long pellets. The pellets are connected by insulating nylon links that are pinned into the pellets with either rivets or threaded pins and screws. There are bushings between

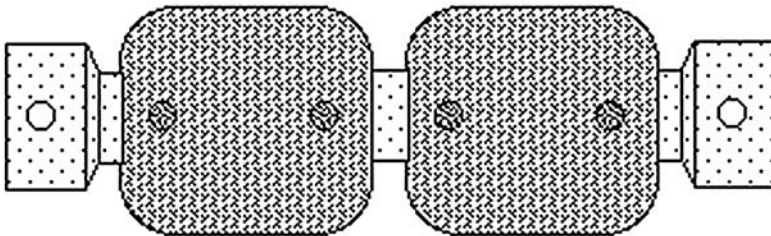


Fig. 6.4. The Pelletron chain

the inner pellet wall and the nylon link that prevent the nylon link from moving side to side and keep it centered in the pellet. The rolled ends act as spark gaps to protect the nylon links from spark damage. While early chains used an asymmetric nylon link that only allowed the chain to bend in one direction, modern chains utilize a symmetric link that permits the chain to bend in either direction. The principle of the chain charging system is illustrated in Fig. 6.5.

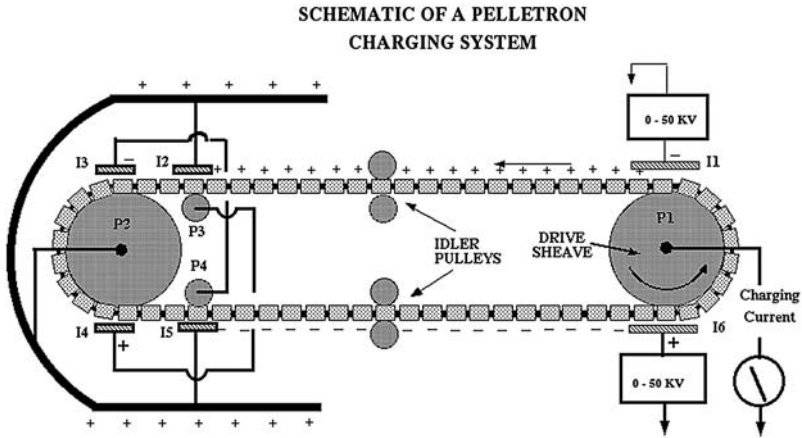


Fig. 6.5. The principle of the chain charging system

The chain rotates on two wheels, typically 30 cm diameter, and travels at about 15 m/s. Charge is induced on the chain as it leaves the grounded end of the column by a negatively charged electrode called the “inductor”. The inductor is biased by a 50 kV high-voltage power supply which is controlled remotely from the accelerator console. As no current is drawn from this supply, it only needs to be voltage regulated. The induced charge on the chain is typically between 3 and 4 $\mu\text{A}/\text{kV}$ of inductor voltage. Positive charge flows through the wheel and the contact bands to the metal link – or pellet – which develops a mirror charge on the surface of the link opposite the negative electrode. As the wheel rotates, the contact between the pellet and the wheel is broken, and the positive charge is trapped on the steel pellet by the insulating nylon connecting link. The charge then redistributes itself on the outside surface of the pellet and is mechanically transported to the terminal by the rotating chain. As the charged pellet arrives in the terminal it passes through another electrode. A mirror charge is developed on this electrode that is negative in sign and equal to the inductor voltage if the gaps are the same as at the inductor. A conductive “charge pickoff” wheel located under this electrode picks up charge from the chain as it passes, and applies it to another inductor that is located on the opposite side of the terminal wheel.

This terminal “inductor” electrode becomes electrostatically positive owing to the electrons that have flowed to the pickoff wheel, and a negative charge is induced on the pellets leaving the terminal. In this way, charge is carried in both directions and the charging efficiency can be doubled. The pellet arriving in the terminal contacts the conductive rim of the terminal pulley and the trapped charge is thereby transferred to the terminal. As mentioned previously, the arriving charge is not a DC current, owing to the nature of the charging system, but it delivers a much more constant charge than belt systems typically can. For chain charging systems, the time-dependent fluctuations in the charging current are substantially less as the charge is applied to individual metal chain elements, which have a high degree of uniformity. During an open-tank maintenance, it is useful to check the charging uniformity and efficiency by grounding the terminal through a resistor, $\cong 1\text{ M}\Omega$ for instance. Charge can be applied to the chain(s) and the terminal voltage observed with an oscilloscope connected to the terminal. This works for charging currents up to $\cong 100\text{ }\mu\text{A}$ and will make apparent any problems in the charging system before closing the accelerator. It is good practice to record the results with a photograph for future reference. Figure 6.6 is an oscilloscope trace showing the magnitude of the arriving charge versus time for one chain in a typical chain charging system.

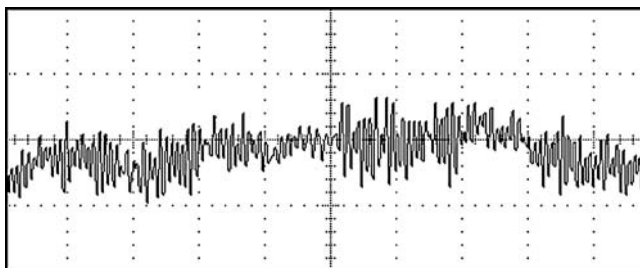


Fig. 6.6. An oscilloscope trace showing the magnitude of the arriving charge versus time for one chain in a typical chain charging system

The high-frequency ripples are due to the arrival of the individual charges on the pellets, while the slower fluctuations are due to nonuniformities in the charging and discharging wheels. The typical time-dependent voltage fluctuations on the terminal observed for a chain charging system are much less than 0.1%, and can, in a well-adjusted system, approach 0.01% or better.

6.3.2 Other Chain Charging Systems

Researchers at Daresbury Laboratory (UK) developed a different type of charging chain in the 1970s. High Voltage Engineering developed its own version of this chain system and named it the Laddertron. This chain consisted of

two chains running side by side with aluminum crossbars connecting adjacent pellets. The appearance of the chain therefore resembles a ladder. This chain was a major improvement over the rubberized cotton charging belts and was installed in several new machines in Stony Brook, Orsay, Legnaro, Beijing and Ile-Ife (although that machine has never been used). This charging chain was guaranteed to carry $250\mu\text{A}$ of charge in an accelerator with at least 30% SF_6 in the insulating gas mixture. The voltage stability is greatly improved over the belt, with 1% RMS current stability reported. The construction of the Laddertron is illustrated in Fig. 6.7.

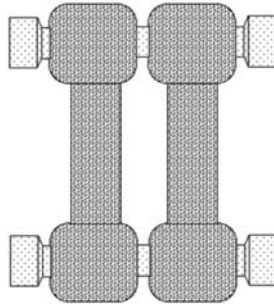


Fig. 6.7. The Laddertron charging chain

This system, however, has several significant disadvantages relative to the Pelletron system. The primary problem with this system is the weight of the chain. It is approximately 5 kg/m and requires a tension of approximately 900 N for stable running. This higher tension produces significantly higher rates of wear on the bushing in the insulating nylon links, and this results in a significantly shorter lifetime – on the order of 5000 hours (compared with perhaps $30\,000$ for a Pelletron). The chain is repairable, however, and can be disassembled to replace the worn bushings. This is a time-consuming task, as in an FN the chain consists of 5000 parts. When one is reassembling the chain, it is important to match the lengths of the parallel links to an accuracy of $\pm 0.13\text{ mm}$ to maintain the straightness of the chain.

6.3.3 Troubleshooting Chain Charging Systems

If the charging current(s) are not stable and the external power supplies have been found to be operating correctly, then a detailed inspection of the components inside the tank is necessary. Inside the tank are usually resistors mounted onto the high voltage feedthroughs that serve to limit current from sparks that would otherwise damage the external power supplies. These can be damaged, and should be measured to verify that they are within tolerance. If they are open, the charging efficiency will be reduced considerably.

The resistance of the antistatic drive sheave should be measured between the chain pellets in contact with the rim and the metal wheel. This resistance should be low, less than $20\ \Omega$ as measured with a simple hand-held multimeter. High resistance indicates that the contact bands are misadjusted or worn to the point that they should be replaced. The chain should be inspected for evidence of wear. The rivets that attach the nylon links to the chain pellets should be inspected for fret corrosion – this typically shows up as a reddish ring around the head of the rivet. If this is found, the rivet should be removed and replaced. If this is left untouched, the corrosion will continue and worsen to the point that the rivet may come loose from the pellet and the chain may break. The idler wheels, if installed, should be checked for cleanliness and bearing integrity. Dirty idlers can cause parasitic leakage from the charging chain to the column. Typically, the problem arises from grease lost from the bearings in the idlers, signaling that the bearing is at the end of its useful life and should be replaced. In the terminal, the adjustment of the gaps between the chain and the electrodes should be checked with the standard jig supplied by the manufacturer. The standard gap is 6.35 mm. For maximum charging efficiency, the gaps of the inductor and suppressor shoes at the grounded end of the column and in the terminal must be equal to match the capacitances between these electrodes and the chain. The resistance of the charge pickoff wheel to the terminal should also be checked and should be about 1 M Ω . Any evidence of lubricant leakage from the bearings in the charge pickoff wheel also indicates the need for replacement. Typically, the bearing lifetime approaches 20 000 operating hours.

When adjusting the inductor and suppressor shoes, one should observe the chain as it leaves and arrives at the drive wheels, as the path is not horizontal owing to the momentum of the chain as it leaves the wheel. After the chain path has been observed, the chain is stopped and a correction is made to the inductor electrode to tip it such that the gap between the electrode and the moving chain is constant. This may take several iterations and should result in a slight increase in the charging efficiency. Typically, the coned end of the inductor is tipped 4 to 5 mm closer to the chain to achieve the proper running gap. In the terminal, the chain approaches the charging sheave horizontally but again, when leaving the wheel, on the run back to the base end, tends to follow the curvature of the wheel, and the larger gap needs to be compensated for as described previously.

When running, the chain should be very stable, having no bounce over the length of the run. Vertical motion in the traveling chain produces excessive loads on the idler pulley bearings and in the chain itself – resulting in drastically shorter lifetimes for the chain and other mechanical components in the system. In a Pelletron, the tension in the chain is provided by a set of lead weights applied to the pendulum arm at the grounded end of the accelerator tank. The tension is adjusted by adding or removing lead weights (approximately 5.5 kg each). The typical chain tension is $\sim 34\text{ N/m}$ of chain length.

In an FN tandem, this amounts to approximately 450 N, which requires eight to eleven lead weights to achieve. Owing to the lever arms involved, this will produce a variation in the chain tension of between 450 and 650 N.

In the terminal, the Pelletron system uses a passive system to produce the down charge, whereas the Laddertron system utilizes active power supplies to achieve down charge. Charge collection is achieved by charge pickoff wheels located ahead of the terminal drive wheel. These conductive wheels are adjusted to contact the running chain but not to support it. When running, these charge pickoff wheels should remain in continuous contact with the running chain for uniform charge collection. Excessive skipping may indicate the need to adjust the tension in the chain as previously described. If the pickoff wheels are adjusted such that they are loaded by the chain, the lifetime of the bearings in these wheels will be drastically shortened, resulting in premature failure of the charging system. These wheels are tested with the chain stationary, with a 500 V megohmmeter. A good wheel will have a resistance between 1 and 500 M Ω . Resistances larger than this indicate the need for replacement. The electrodes that surround the chain at the pickoff wheel locations are also adjusted to a nominal 6.35 mm gap.

The drive sheaves are made of conductive plastic mounted onto aluminum wheels. Some installations also include metal side bands to improve the connection to the chain. Typically, for these installations the contact-band-to-chain resistance should be under 20 Ω . The resistance of the conductive plastic wheel to the metal inner wheel should be also be no more than 500 M Ω .

Drastically reduced charging efficiency can often be traced to a failure in the high-voltage feedthroughs for the inductor and suppressor electrodes at the base of the accelerator. These feedthroughs are typically protected by a large 50 kW resistor in series with the lead between the feedthrough and the electrode. These resistors are susceptible to sparks down the chain or column, and the usual failure mode is for the resistor to open up. This results in no or an unstable voltage on the inductor or suppressor electrode. These resistors should be checked at every maintenance with a simple ohmmeter and replaced if not within $\pm 10\%$ of their nominal value. It is also a good idea to install a spark gap on the electrode side of these resistors to limit the overvoltage impressed on them during a spark.

Box 2: Development of Charging Belts in Russia

V.A. Romanov

State Scientific Center of the Russian Federation, Institute for Physics and Power Engineering, Obninsk, Kaluga Region, 249033 Russia
romanov@ippe.obninsk.ru

Charging systems based on charge-carrying belts have been one of least reliable components of Russian electrostatic accelerators. The main disadvantages of existing belts are their short lifetime and the considerable wear of the working surface. Owing to the short lifetime of the belts, it has not been possible to use the experimental equipment effectively, and additional expense has been required to purchase the gas insulating mixture. The high wear of the belt increases the content of dust in the insulating gas considerably, as well as the dust deposition on the surface of the high-voltage structure, thus reducing the dielectric strength of the accelerator. Under these circumstances, it is also impossible to use a contact method for applying and removing charges. A belt made of rubberized cotton fabric has a high hygroscopicity; this prolongs the time needed to condition the belt. In addition to these disadvantages, the stiffness of the belt is insufficient, and therefore the position of the belt is unstable, resulting in high-voltage instability of the accelerator. The strength of interlayer connection is rather low (880 N/m), which very often leads to belt failures. In order to develop a reliable belt, long-term studies have been carried out at our institute in Obninsk. Many different types of synthetic, cotton and combined fabrics have been tested.

The type of belts tested at the beginning had a very high mechanical strength and wear resistance. However, their dielectric strength was rather low. When they were used, some of the fibers located in the warp of the belt burned out. The reason for the decrease of dielectric strength of the belt was finally found by tests carried out at different accelerators. Discharges occurring along the belt were found to be caused by interlayer cavities formed in some sections of the belt as a result of insufficient gluing of the different layers.

Some of the belts made with a synthetic warp showed a rather high relative elongation (more than 3%). In some accelerator designs, it is impossible to use belts with such an elongation. During the development and tests of the different belts, some other failures took place, resulting in a decrease of the charging performance.

In order to ensure high dielectric and mechanical strength as well as a permissible relative elongation value (less than 1%), a combined cotton–polyester fabric is now being used as a warp. Raw rubber for the rubber mixture was

chosen from the standpoint of ensuring wear resistance, mechanical strength and the appropriate electric characteristics.

Production of 1.2 to 3 mm thick belts with two, three or four layers has been launched. The working load of a 550 mm wide four-layer belt is about 7.9 kN for each of the two parts of the belt, while the load of a two-layer belt about 200 mm wide is about 4.9 kN. Seamless belts of up to 6 m full length can now be manufactured. Belts with a length over 6 m are manufactured using a connecting seam.

Charging belts of the latter type are used at the Obninsk accelerators and in accelerators at other Russian research centers. The methods of applying and removing charges utilized (either by charging needles or contact methods using grids and foils) vary. At all accelerators operating with these belts, a special design approach is used for applying and removing charges on the inner surface of the belt.

Table B2.1. Some results on tests on rubberized fabrics for belt charge conveyors

No.	Characteristics	Type 236	Type 1590-1
<i>1</i>	<i>Design</i>		
1.1	Number of layers	4	4
1.2	Base (fabric)	Percale	Cotton fabric with polyester
1.3	Type of (raw) rubber	Natural rubber	Methyl styrene synthetic rubber (MSSR)
1.4	Belt thickness (mm)	1.4	2.8
<i>2</i>	<i>Mechanical characteristics</i>		
2.1	Rupture force applied to 50 mm width strip (kN)	0.83	1.9
2.2	Strength of interlayer connection (kN/m)	0.88	1.67
2.3	Strain for 9.81 kN/m tension (%)	3	1
2.4	Wear resistance		
2.4.1	Duration of tests (hours)	42	1200
2.4.2	Equivalent time of accelerator operation	296	10 000
2.4.3	Lifetime of the belt in an accelerator	300–500	8000–10 000
<i>3</i>	<i>Electrical properties</i>		
3.1	Surface breakdown under atmospheric conditions (kV/cm)	14.9	14.8
3.2	Surface resistance (ohm/mm)	1.2×10^{11}	1.5×10^{12}

For instance, at the EGP-15 tandem accelerator in Obninsk, the contact method for applying the charge is based on a 100 μm thick stainless steel foil. The belt was installed in the accelerator in 1989, and after running for over 11 000 hours, its condition today is still quite good. The maximum working gradient of the potential along the belt is 1.5 MV/m. The same type of belt has been used at the EG-5 single-ended accelerator in the Neutron Physics Laboratory of the Joint Institute for Nuclear Research in Dubna since March 1997. In this accelerator, the contact method is used for charging the belt. The belt has been used more than 8500 hours and is still in good shape.

Different types of belts and their materials were tested at high-voltage in our experimental facilities. The belts were tested for rupture, stratification, dielectric strength and wear resistance. In order to test the belts for wear resistance, belt material samples were stuck on a pulley driven by the electric motor, and a contact system ensuring the required hold-down force was provided. The results of the tests are presented in Table B2.1. These results show that the new belt has high electric and mechanical characteristics, as well as high wear resistance.

It is my pleasant obligation to express sincere appreciation to my colleagues V.I. Spirin and V.V. Ecomasov for their help in carrying out the tests on the various belts.

Box 3: Cascade Generators

R. Hellborg

Department of Physics, Lund University, Sölvegatan 14, 223 62 Lund, Sweden
ragnar.hellborg@nuclear.lu.se

In the Introduction to this book, an overview of different accelerator types was given. For accelerators designed along the principle of the “direct voltage technique”, two subgroups are available: electrostatic accelerators, in which the high voltage is generated by electrostatic charging, and cascade accelerators, in which the high voltage is generated by rectifying an AC voltage. The technique for obtaining the high voltage in an electrostatic accelerator is described in detail in Chap. 6. In this box, the principal design of the power supplies for different types of cascade accelerators will be briefly described.

Asymmetrical Circuit

The original design of a cascade accelerator, first used by Cockcroft and Walton [1], can be seen schematically in Fig. B3.1. A photo of Cockcroft and Walton’s accelerator is shown in Chap. 5. The circuit is asymmetrical;

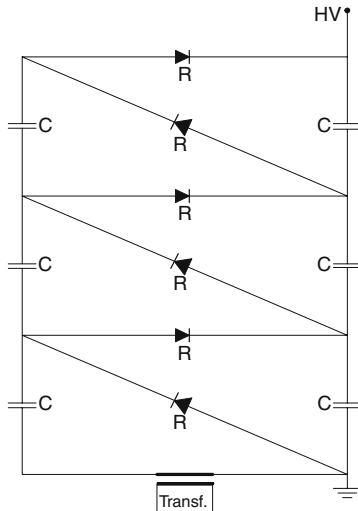


Fig. B3.1. Asymmetrical-circuit cascade generator consisting of capacitors C and rectifiers R

sometimes this word is used for this type of accelerator, and sometimes it is called the Cockcroft–Walton type. The circuit was designed to transfer AC into high-voltage DC more than ten years before Cockcroft and Walton used it in their accelerator. It was known in the electrical engineering community as the *Greinacher doubling voltage circuit* after Heinrich Greinacher – a professor of physics at the University of Bern, Switzerland – who developed this circuit around 1920 [2]. The circuit includes n identical stages (in Fig. B3.1, three stages are shown), called cascades. The circuit uses two stacks of series-connected capacitors C . The right capacitor stack in Fig. B3.1 is connected at one end to ground and at the other to the high-voltage (HV) end. The voltage across each capacitor in this stack is constant except for a ripple. One end of the left capacitor stack in Fig. B3.1 is connected to a transformer giving peak voltages of $\pm U$. The voltages at all points along this stack oscillate over a range of $2U$. Series-connected rectifiers R link the two stacks. As the voltage on the transformer oscillates, charge is transferred stepwise through the rectifiers from ground to the HV terminal. The terminal voltage will be $2nU$. The chain can be extended to higher potentials, limited only by the ability of the high-voltage terminal to hold its potential without sparking to the surroundings.

Symmetrical Circuit

In practice, the asymmetrical circuit was soon replaced by a symmetrical circuit, as seen in Fig. B3.2. This employs two transformers and two capacitor stacks (the outer two stacks in Fig. B3.2) that oscillate in voltage. Both oscillating stacks feed one fixed-voltage capacitor stack (the central stack in

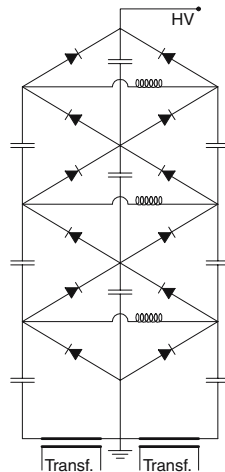


Fig. B3.2. Symmetrical-circuit cascade generator

Fig. B3.2). The advantage of the symmetrical circuit can be seen from the voltage drop ΔU and the voltage ripple δU when it is loaded with a current I . For the asymmetrical circuit, these are given by [3]

$$\Delta U_{as} = \frac{I}{fC} \frac{n}{3} \left(2n^2 + \frac{3}{2}n - \frac{1}{2} \right) \quad (\text{B3.1})$$

$$\delta U_{as} = \frac{I}{fC} \frac{n}{2} (n + 1) \quad (\text{B3.2})$$

For the symmetrical circuit, the equations are [3]

$$\Delta U_s = \frac{I}{fC} \frac{n}{3} \left(n^2 + \frac{3}{2} \right) \quad (\text{B3.3})$$

$$\delta U_s = \frac{I}{fC} \frac{n}{2} \quad (\text{B3.4})$$

Here f is the frequency of the AC supply, C is the capacitance of a given stage and n is the number of stages. For an increasing number of stages n , both the voltage drop and the ripple become considerably lower for the symmetrical circuit compared with the asymmetrical. As both the voltage drop and the ripple vary inversely with the frequency f , a high frequency of the primary sinusoidal voltage is of importance. Accelerators of up to several MV have been constructed. With currents of several hundred mA, they give a total beam power of several hundred kW. These generators have often been employed in injectors to high-energy machines, and they are commonly employed as power supplies in electron microscopy. Asymmetric and symmetric accelerators are often open and not enclosed in an accelerator tank.

Parallel-Driven Circuit

A third principle, shown in Fig. B3.3, for obtaining a high voltage for a cascade accelerator was introduced by Radiation Dynamics Inc. Their product is called the Dynamitron. Inside the accelerator tank, two large semicylindrical RF electrodes are mounted near the wall of the tank, surrounding the column. These electrodes are supplied with power from an RF oscillator at 100 kHz, and they form the tuning capacitance of an LC resonant circuit. The high-voltage column is enclosed by half rings with smooth exterior surfaces to inhibit corona and spark discharges. In these segments along the accelerator column, secondary voltages are induced by capacitive coupling. These segments are coupled to rectifiers, and the rectified voltages from each segment are added up in two rows on opposite sides to supply the terminal with a high voltage. The tank is filled in the normal way with spark-protecting gas. The DC voltage produced by each segment is 50 kV. At the end of the 1970s, Kenn Purser [4] designed a similar type of parallel-driven circuit to

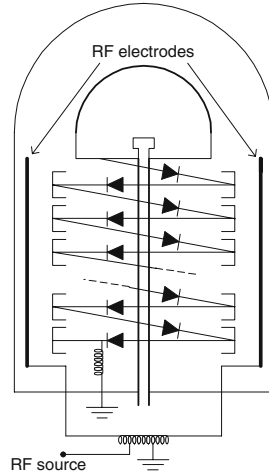


Fig. B3.3. Parallel-driven circuit

be used for the at that time new Tandetrans produced by General Ionex Corporation. The Tandetron (today produced by HVEE) is a compact tandem for material analysis, accelerator mass spectrometry, ion implantation etc. The Tandetron has a 50 kHz driver delivering several mA with very high stability. For accessibility, the high-voltage stack in Tandetrans up to 3 MV is at a right angle to the accelerator column. In the 5 MV Tandetron delivered in 2001–02 to the Centro de Micro-Analisis de Materiales in Madrid [5], the high-voltage stack is parallel to the high-energy column. A photo of the Madrid machine is shown in Fig. B3.4. A very high beam current can be accelerated in a parallel-driven accelerator, with a total power of up to 200 kW. Driving the stages in parallel instead of in series reduces the stored energy to levels comparable with electrostatic accelerators. Minimizing stored energy is important, especially in MV accelerators, because the high stored energy released in a discharge can damage capacitors, rectifiers, column components etc. Parallel-driven accelerators are ordinarily designed for higher voltages than are series accelerators and are enclosed in a tank.

Insulating-Core Transformer

The design of an insulating-core transformer is shown in Fig. B3.5, and a photo can be found in Chap. 28. The core is divided into sections separated by spacers of insulating material. The core is excited through the primary windings (using a three-phase, 400 V system at 50 or 60 Hz). Input power is magnetically coupled to secondary coils (three per deck) by a three-phase iron core electrically insulated between each deck. Each of all the secondary sections is coupled to a rectifier operating as a voltage divider and is an

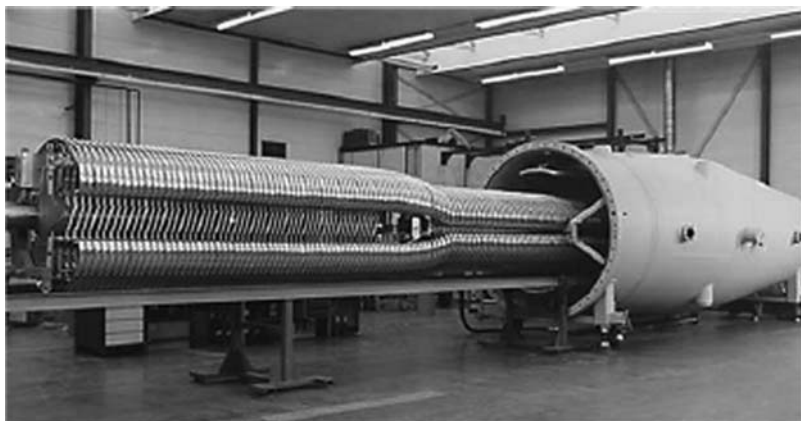


Fig. B3.4. The 5 MV Tandetron in Madrid (by courtesy of G. Garcia López)

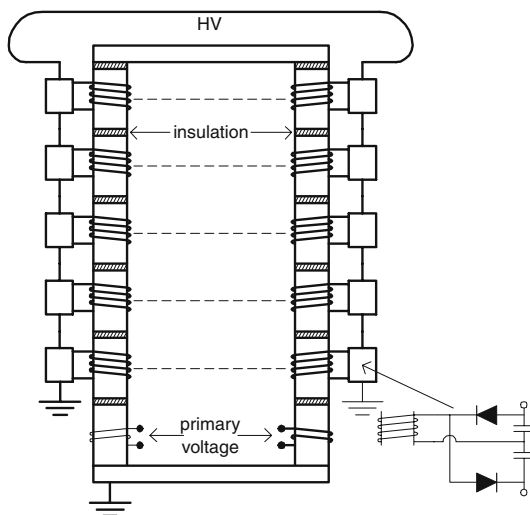


Fig. B3.5. Insulating-core transformer, two phases of the three-phase system are shown in the drawing

independent 50 kV unit. The rectifier outputs are connected in series to produce the high voltage. This type of voltage supply is housed in a tank filled with gas and can be built to be very compact. Units up to several MV and tens of mA giving a beam power up to several hundred kW are available. The accelerating column may be directly connected to the high-voltage terminal or may be physically separated from the transformer and connected to it by a high-voltage shield cable. Insulating-core transformers have long been produced by HVEE and VIVIRAD, and they are often used in industrial applications; see Chap. 28.

Acknowledgments

Ludwig Rohrer has given most valuable comments and suggestions for improvements relating to this box.

References

1. J.D. Cockcroft, E.T.S. Walton: Proc. Roy. Soc. A **129**, 477 (1930)
2. H. Greinacher: Z. Physik **4**, 195 (1921)
3. M. Minovic, P. Schulze: *Hochspannungstechnik* (VDE-Verlag, Berlin 1992)
4. K.H. Purser, R.B. Liebert, C.J. Russon: Radiocarbon **27**, 794 (1980)
5. G. Garcia Lopez et al.: The Centro de Micro-Analisis de Materiales equipped with a 5 MV Tandetron. In: *Proceedings of the Symposium of North-Eastern Accelerator Personnel, Strasbourg*, Oct. 2003

7 Voltage Distribution Systems – Resistors and Corona Points

D. Weisser

Research School of Physical Sciences and Engineering, Australian National University, Canberra, Australia
`david.weisser@anu.edu.au`

7.1 Introduction

This chapter presents the development of the two main technologies for grading the voltage of accelerators – resistors and corona systems. The voltage-holding ability of accelerators depends upon management of the distribution of electric-field stress, which relies on the voltage distribution system (Sect. 7.2). Although modern resistor systems are now the norm, corona grading provided an adequate bridging solution for large machines while resistor protection was perfected (Sect. 7.3). The large amount of energy stored in the electric field of such machines is broadcast during a spark and causes resistors to fail (Sect. 7.4). These failures motivated the improvement of resistors and techniques to protect them (Sect. 7.5). The development of resistor systems that survived sparks in large machines is largely the story of the protection strategies to reduce the coupling of spark energy to the resistors (Sect. 7.6). The crucial ingredients for such protection include spark gaps intrinsic to the accelerator, aerial effects, local shielding and the structure of the resistors themselves.

The confirmation of the success of resistor systems depends on the measurement of their resistance to a few percent, for which in-machine testing is of limited value (Sect. 7.7). It is fortunate, therefore, that modern systems are so effective in maintaining resistance value that now the main resistor failure mode is mechanical damage. Modern systems have overcome all of the historic problems, at least for machines with terminal voltages up to 25 MV.

7.2 Why Is Voltage Grading Needed?

The high-voltage limit of early electrostatic accelerators was set by discharges adjacent to the terminal along the long, uninterrupted insulating columns supporting the terminal. If the region of concentrated gradient could be shared, then the combination of several insulators and subsidiary electrodes might support a much higher voltage than would a single long insulator. The electrostatic accelerator built at Melbourne University between 1946 and 1948 had two corona rings added to the column, as seen in Fig. 7.1. These act as subsidiary electrodes to relieve the voltage stress adjacent to the terminal.

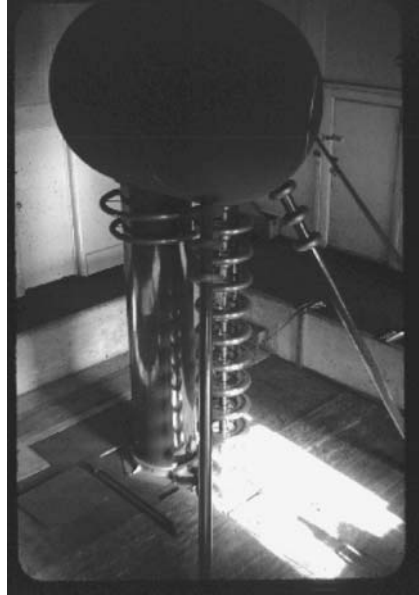


Fig. 7.1. The insulating column on the Melbourne University Van de Graaff has two corona rings near the terminal. These reduce the electric stress on the column adjacent to the terminal by spreading portions of it to the regions at the corona rings

The extension of such subdivision to the entire long column insulator stimulated the development of voltage-grading devices to control the electric field at each subsidiary electrode.

In contrast to the supporting column, the accelerator tubes had individual metal electrodes to establish the electric fields that accelerated the beam. These were only referenced to the gradient by external corona rings, which can also be seen in Fig. 7.1. Presumably, the voltage on the rings was established by fortuitous corona from the well-rounded rings. The evolution from fortuitous corona to deliberate voltage-grading devices marks the progress of electrostatic accelerators in their battle to increase the voltage at which they spark.

Machine sparks are triggered when the region of highest electric stress breaks down. This can occur either on the inside of the accelerating tube, where the breakdown is in vacuum, or in the external region of the accelerator structure, where air or high-pressure gas is the insulation environment. If the highest stress can be spread in a controlled manner among several locations then the peak stress is decreased and the machine can be pushed to higher voltages until, once again, some gap is overstressed. The voltage-grading system is the main tool to control the sharing of the high stress burden.

7.3 Corona Grading Systems

Given the early experience with corona rings, it was quite natural for machines to have evolved purposely-designed corona devices to grade the voltage from the terminal to ground. Figure 7.2 shows the corona point system mounted on the accelerator tube electrodes in Herb's high-pressure accelerator in 1937, the forerunner of modern electrostatic accelerators [1]. Connections from these corona points to the column also provided the grading for the column rings.

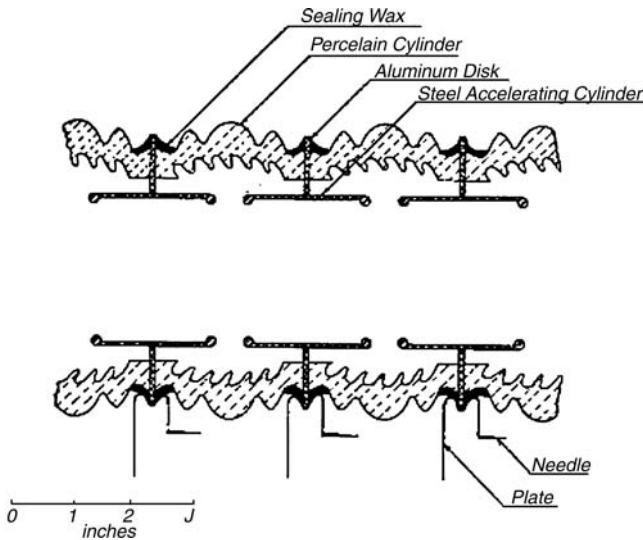


Fig. 7.2. Corona current from the needles to the opposite plates provided the voltage grading in the 1937 air-pressurized accelerator (Reprinted from [1], copyright 1937, with permission from the American Physical Society)

In spite of the head start corona systems enjoyed, high-ohmage resistors soon became the solution of choice for the machines of the day with modest terminal voltages. Resistors were convenient and commercially available, and generally were not damaged by machine sparks. However, as the terminal voltage of machines grew, so too did the failure frequency of the resistors.

The operational and financial costs of failed resistors reignited the use of corona grading systems by NEC in the 1970s as an inexpensive and spark-tolerant alternative. In Fig. 7.3, one can see the evolution from the 1937 version to the open corona point system for the NEC 14UD Pelletron. An advantage of a corona system over one based on resistors is that the voltage across a corona discharge is much less sensitive to changes in the current due to beam loading, for example. Figure 7.4 shows that a reduction by a factor of 2 in the column corona current, from 10 to 5 μA , results in only a



Fig. 7.3. Open corona point system on an NEC accelerator tube

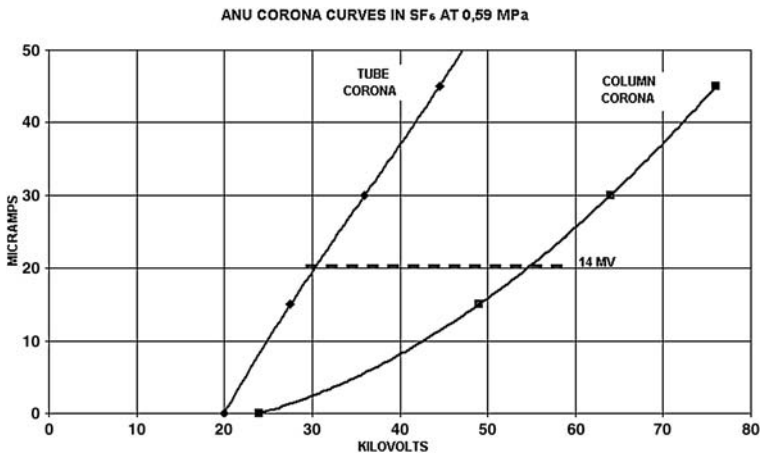


Fig. 7.4. The current carried by the corona grading system in the NEC 14UD vs. the voltage across the gap

25% decrease in voltage. This voltage stiffness allows the machine to operate stably with a smaller demand on the pellet chain charging system. Indeed, many Pelletrons operate with grading corona currents of only 1 to 5 μA .

One substantial disadvantage of the corona point system is that the corona extinguishes when the gap voltage reduces below the threshold. In the case of the column corona device illustrated in Fig. 7.4, a voltage $< 23 \text{ kV}$ across

a column corona point gap would strand the accelerator without any reliable grading at all. Shorting out some sections of the accelerator to preserve enough gradient to keep the remaining corona points lit ameliorates this problem. The process of shorting out sections of the column breaches the pressure barrier of the accelerator and so entails some risks to the expensive SF_6 inventory and to personnel. On the positive side, this solution preserves the focusing strength of the tube entrance – necessary to maintain consistent optics and therefore good beam transmission (Chaps. 8 and 13). An alternative is to reduce the pressure of the insulating gas in order to reignite the corona at the lower gradient. This option is even less palatable, since it requires operation of the complex gas-handling system and is expensive in time and in technical effort.

To obviate the need to change the gas pressure in the entire machine, NEC developed a system in which the corona point assemblies are mounted in a series array of separately pressurized insulating tubes, one of which is shown in Fig. 7.5. This tube, like an NEC accelerator tube, comprises titanium electrodes bonded to ceramic insulators protected by annular spark gaps. Each electrode, supporting a triplet of corona points in the grading tube, is attached to a corresponding electrode on either the acceleration tube or the column. The pressure in this system of tubes is easily altered to tailor the gradient for lower-voltage operation.

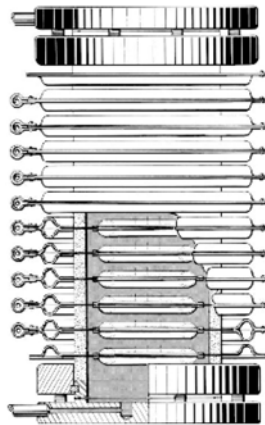


Fig. 7.5. Enclosed corona point system that allows the insulating-gas pressure to be conveniently lowered to increase the corona current for low-gradient operation

Pelletrons employ straight-field accelerator tubes rather than ones with inclined fields, so transmission degradation due to uneven voltage grading is not an issue, but gradient reliability and cost are. The choice of corona assemblies solved the cost problem but not that of gradient variations. The consequence of gradient nonuniformity is that some of the gaps operate at

20% higher voltage than the average. These overstressed gaps will be the first to break down, triggering a discharge of the entire machine. A machine operating at 12 MV, limited by a 20% gradient nonuniformity, would perform at 14.7 MV if the gradient nonuniformity were reduced to 2%. The extreme sensitivity of the voltage across a corona-graded gap to the distance from the points to the next plane, to dulling of the points with wear and to breakdown products coating the needle tip results in very nonuniform gradients [2]. Figure 7.6 shows the SF_6 breakdown products accumulated on the tip of a corona point.

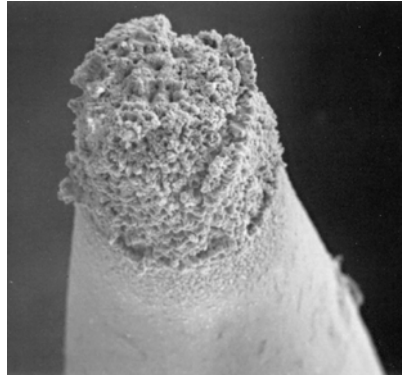


Fig. 7.6. SF_6 breakdown product deposits on a corona point tip. Magnification 100 $\times$

At best, the corona current, and so the voltage, is not constant but varies as the discharge dances around the interface between the deposit and the bare metal and occasionally from protuberances on the deposit itself. The statistical coincidences of the hundreds of corona fluctuations can instantaneously overvoltage a gap, triggering a spark.

More drastically, the corona produces corrosive SF_6 breakdown products, which attack the Pelletron charging chains, shortening their life to as little as a few hundred hours in extreme cases [3]. The corona also creates and mobilizes particulates, which, when they jump in the electric field, detonate discharges of the full machine. The enclosed corona system removes the breakdown products from the main accelerator environment, thus avoiding only some of these problems but concentrating the others in the enclosed system. However, the substantial cost of enclosed systems, their gradient nonuniformity, and their imperviousness to both inspection and easy repair, prevented their wide adoption.

The success of relatively inexpensive and robust NEC resistor systems using Welwyn [4] resistors has now supplanted both open and enclosed corona grading as the technology of choice in Pelletron accelerators.

7.4 The Need for Better Resistors

In the 1960s, HVEC developed the EN tandem accelerators, which achieved terminal voltages >5 MV. The success of the EN, which used resistors to grade the column, encouraged HVEC to produce the larger FN machines. The FNs exceeded their 7.5 MV expectation, running at and above 10 MV. These accelerators were soon joined by the HVEC MP machines, which were intended for operation at 10 MV but were pushed higher by the success of the FNs. The higher terminal voltages uncovered extra problems for the grading resistors. The resistors were exposed to a factor of 4 to 6 increase in spark energy in the higher-voltage machines, owing to the stored energy increasing with the square of the terminal voltage.

Unfortunately, the resistors then in use changed in value by $>25\%$ after a short exposure to high-voltage and sparks, causing beam transmission problems. This was because these accelerators used inclined-field accelerating tubes, in which the beam was deflected from one side of the axis to the other (Chaps. 8 and 13). Because these deflections we carefully designed to compensate each other, the tube relied on the voltage gradient being uniform. A nonuniform gradient led to the beam emerging away from the machine axis. Much research time was lost and much technical effort expended on locating open-circuit resistors and reshuffling the rest to smooth out the gradient in order to put the beam back near the axis.

An increase in resistance of a resistor by 25% will overstress its insulating gap, lowering the maximum useful voltage of the accelerator. A decrease of resistance by 25% for enough resistors exposes unchanged resistors to a higher than nominal voltage, thus also triggering sparks. Even more catastrophically, resistors frequently failed open-circuit and sparked continuously – the machine had to be opened for their replacement.

Resistor assemblies, although commercially available, failed so frequently that their cost soon became an uncomfortable burden. This drove the need to improve the reliability and to lower the cost of resistor voltage-grading systems.

7.5 Evolution of Resistors

Resistors made of metal-oxide-coated ceramic cylinders displaced carbon resistors, which required mechanical support and generally had unacceptable failure rates. Originally, metal oxide resistors used hollow ceramic-rod substrates, but these were susceptible to discharges through the center since one could not ensure that the accelerator's insulating gas would penetrate into this volume [5]. Spark damage to their protective coating was another failure mode of the standard high-voltage resistors [6]. Uncoated, solid-ceramic-core resistors are now the preferred option for use in accelerators and are available from several manufacturers.

The thick metal-oxide resistive layer is usually employed in a spiral pattern. Since a spiral is inductive and so susceptible to turn-to-turn high transient voltages and subsequent failure, low-inductance patterned resistors are an option [7]. Metal-oxide-on-ceramic resistors are sufficiently robust, however, that even ones with a spiral patterned resistive coating are entirely successful.

7.6 Spark Protection

Resistors must maintain their initial value to a few percent in spite of being subjected to severe overvoltages during sparks. Successful resistor systems minimize the voltage stress they suffer by exploiting the protection strategies discussed below.

7.6.1 Spark Gaps

In order for the minimum spark energy to be transferred to the resistors, the majority of the energy should be dissipated elsewhere in the accelerator structure. The safest medium for spark energy dissipation is in the insulating gas, since it is not permanently damaged by the discharge. The safest locations for the relieving discharges are in the robust spark gaps between the equipotential rings and in the spark gaps on the structural column.

The ring gaps are best, since they are furthest from the more delicate insulating components such as the accelerator tube and column insulators, let alone the resistors. In the NEC 14UD, the rings provide $2.6 \text{ cm}^2/\text{kV}$ of spark gap, and a similar figure would be true for the HVEC FN. In an MP, however, with fewer, larger-cross-section rings, the ring spark path is somewhat less capable.

Unfortunately, not all the spark energy travels via the ring gaps, as inspection of any other spark gap in an accelerator attests. Indeed, in the 14UD, the density of inter-ring spark marks is a maximum between the column posts and diminishes close to the posts as the post spark gaps assume more of the discharge burden.

Almost as important as the ring spark path in minimizing the exposure of the resistors to damage is the spark-energy-carrying ability of the column structure. In the 14UD, the column spark gaps provide $0.5 \text{ cm}^2/\text{kV}$ of protection. HVEC machines, with spark gap buttons, offer about $0.02 \text{ cm}^2/\text{kV}$. A similar situation obtains for grading elements mounted directly on the accelerator tube, with the NEC tube having $0.4 \text{ cm}^2/\text{kV}$ and the HVEC tube about $0.024 \text{ cm}^2/\text{kV}$. Because of these large differences, it could be argued that grading elements on an HVEC column would be subject to ~ 20 times the spark energy as those in a Pelletron. However, since the whole area of the annular spark gap does not participate in a given discharge, this is probably an overestimate. Qualitatively, the inference that “the larger the spark gap

area the better the protection” is supported by the experience of resistor lifetimes. The attrition rate for unshielded resistors in MP and FN accelerators demanded the development of local shielding for resistors [9]. This is in contrast to the success of essentially unshielded resistor elements in the largest NEC accelerator, the 25URC at the Holifield Radioactive Ion Beam Facility at Oak Ridge National Laboratory shown in Fig. 7.7 [8].

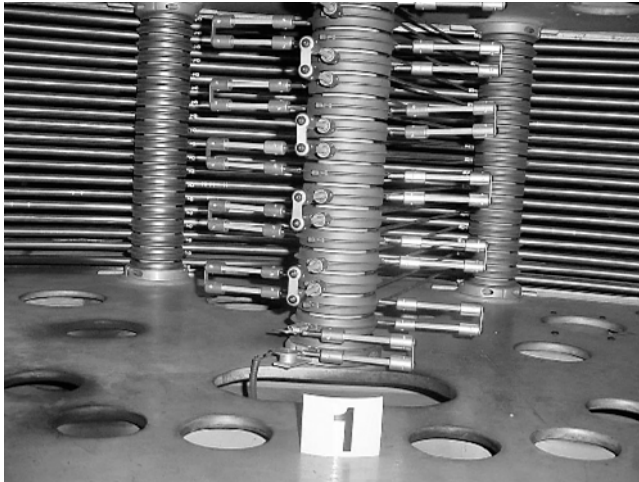


Fig. 7.7. The column resistors in the 25URC are mounted on an auxiliary post well away from the rings. The generous provision of spark gaps on the posts of NEC Pelletrons allows the resistors to survive without local spark shielding [8]

7.6.2 Coupling of Spark Energy to Resistors

The geometry of the resistors themselves and how they are attached to the column also affect their exposure to spark energy. Early resistor assemblies were series connections of individual resistors arrayed in long sticks. Figure 7.8 shows a more modern version using two metal-oxide-on-ceramic resistors based on a University of Rochester [11] innovation and developed at Brookhaven National Laboratory for their MP accelerators [9]. These assemblies, like their multiple carbon forebears, inevitably span parts of the column ~ 50 cm apart.

During a spark, sections across the diameter of the column will be at substantially different voltages because the spark gaps and/or rings on one side of the column will fire before those on the other side. A resistor string joining the two sides of the column would be subject to a megavolt RF spike rather than the 40–60 kV DC for which it was intended. In addition, the long assemblies act like dipole aerials, picking up energy from the spark’s

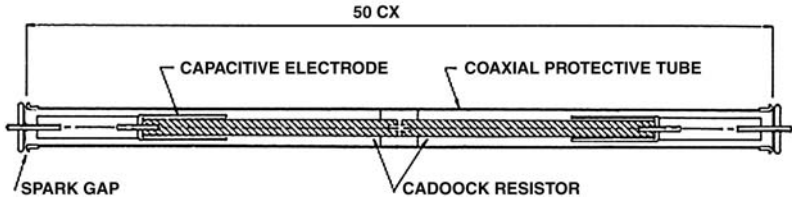


Fig. 7.8. The Brookhaven National Laboratory MP column resistor assembly (Reprinted from [10], copyright 1993, with permission from Elsevier)

electromagnetic RF field. The effective aerial of a 50 cm resistor string will absorb more RF power than will a compact assembly, typically ~ 15 cm long. As well, a long resistor string, combined with the column structure to which it is attached, forms a loop aerial that will couple inductively to the spark RF current. The loop area of $\sim 50 \times 50 \text{ cm}^2$ dwarfs the $\sim 15 \times 4 \text{ cm}^2$ of the modern compact designs [10]. The RF pickup suffered by a compact assembly could be argued to be a factor of ~ 140 less than for a long assembly.

The stresses on the resistors discussed above impose an extra gradient along the resistor axis. Since a strong pulse of electric field perpendicular to the resistor axis has been recognized as a danger too, placing them between parallel metal plates has been shown to increase their survivability [12]. Modern metal tube shielding has grown from this. If possible, even shielded resistor pairs are best dispersed to minimize the field between shielding tubes [10].

Figure 7.9 shows the compact resistor arrangement in the ANU 14UD, and Fig. 7.10 that in the FSU FN. In the 14UD, resistor pairs across adjacent voltage gaps are mounted well away from one another to reduce the



Fig. 7.9. The compact column and tube resistors in the ANU Pelletron

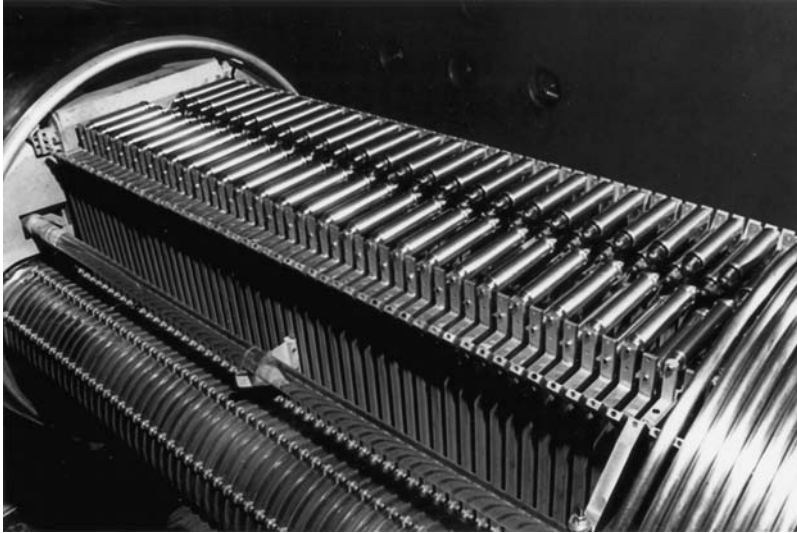


Fig. 7.10. Compact column resistors in the Florida State University FN

tube-to-tube electric field. In the FN, resistor pairs are mounted on alternate column sections for separation.

Modern resistor assemblies are compact in order to avoid spanning large distances. The compact designs also offer the operational benefit of greatly improving access to accelerator components such as the charging system and to the resistor assemblies themselves.

7.6.3 Local Shielding

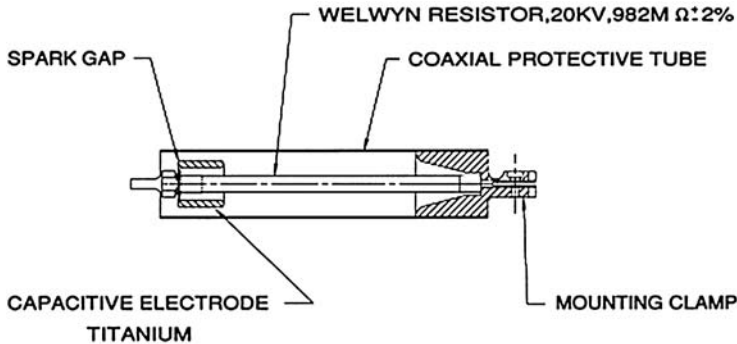
The natural response to the failure of resistors was to protect them locally with spark gaps in parallel with the resistive element. Since this did not solve the problem, more intimate spark gaps, series inductors and parallel capacitors were introduced at Rochester [11] and elaborated at Brookhaven National Laboratory [9], as shown in Fig. 7.8.

This design illustrates almost all of the important features that have evolved into present-day assemblies, with the exception that the resistors span long distances in the accelerator structure.

- Metal-oxide-coated ceramic rod resistors are used instead of wire-wound resistors or several carbon ones.
- The resistors structurally support their integral shields, spark gaps and bypass capacitors.
- A metal tube shields the resistors from direct RF spark energy.
- Metal thimbles are epoxied over the ends of the resistors to:
 - mechanically strengthen the ends of the resistors so that they can better support the structure;

- allow machining to length and squareness to ensure uniformity of the spark gaps;
- provide bypass capacitance.

The use of metal thimbles over the resistor ends is essential and universal. Figure 7.11 shows all the features of current resistor protection techniques, as used in the ANU 14UD Pelletron [13] and in many FN accelerators [14].



ANU COLUMN RESISTOR ASSEMBLY

Fig. 7.11. The column resistor for the ANU Pelletron uses a radial spark gap, in contrast to axial ones in all FN variants (Reprinted from [10], copyright 1993, with permission from Elsevier)

7.7 Resistor Performance and Testing

The criterion for satisfactory performance is necessarily imprecise, since resistance measurements done inside the accelerator are confounded by moisture on the resistors themselves and on the insulators that they span. Hygroscopic deposits, usually from SF_6 breakdown products and/or belt dust, exacerbate this problem. In almost all cases, however, in which metal oxide resistors are suspected of changing values, tests that are performed after they have been removed from the accelerator, cleaned and dried show that they are still within their original $\pm 2\%$ tolerance. In-machine testing is therefore limited to discovering gross failures – mechanical failures being the most common, if an infrequent, mode.

A variety of manufacturers have provided resistors that have performed satisfactorily in large HVEC accelerators. In most cases, resistors performed well if protected by adequate local shielding. In NEC Pelletrons, Welwyn [4] resistors have proved entirely satisfactory. Voltage grading is now successfully

done using resistors with competent and appropriate protection, thus removing resistor failure from the list of risks to efficient operation of electrostatic accelerators.

The evolution of voltage-grading systems has taught that any solution, however successful in one voltage range, may fail if applied to an accelerator with a much higher terminal voltage. It is fortunate, therefore, that it is unlikely that electrostatic accelerators will be built with voltages higher than the 15 to 25 MV now extant. However, even in more modest machines, designers will need to maintain their vigilance lest they revisit the pitfalls described in this chapter.

References

1. R.G. Herb, D.B. Parkinson, D.W. Kerst: Phys. Rev. **51**, 79 (1937)
2. D.C. Weisser, Nucl. Instr. Meth. A **268**, 419 (1988)
3. T.R. Ophel, D.C. Weisser, A. Cooper, L.K. Fifield, G.D. Putt: Nucl. Instr. Meth. **217**, 383 (1983)
4. Welwyn Components Ltd, Bedlington, Northumberland, UK: www.welwyn.com
5. L.E. Collins, F.A. Howe, R. Thorn: Proceedings of the First International Conference on the Technology of Electrostatic Accelerators, Daresbury (1973), DNPL/NSF/R5, p. 172
6. H.R.McK. Hyder: Private communication
7. Caddock Electronics, Riverside, Ca, USA: www.Caddock.com
8. M. Meigs: Private communication
9. J.W. Noé: In *Symposium of North Eastern Accelerator Personnel*, eds. E.D. Berners, U. Garg and C.P. Browne, World Scientific, Singapore, (1986) p. 168
10. D.C. Weisser: Nucl. Instr. Meth. in Phys. Res. A **328**, 138 (1993)
11. K.H. Purser, H.E. Gove, T.S. Lund, H.R.McK. Hyder: Nucl. Instr. Meth. **122**, 159 (1974)
12. J. McKibben: Panel discussion, in *Symposium of North Eastern Accelerator Personnel*, eds. J. Benson, L. Rowton, L. Tesmer, D. Darling, World Scientific, Singapore, (1991) p. 155
13. D.C. Weisser: Nucl. Instr. Meth. in Phys. Res. A **287**, 113 (1990)
14. D. Chapman: Private communication

8 Accelerator Tubes

H.R.McK. Hyder

Department of Physics, Oxford University, Denys Wilkinson Building,
Keble Road, Oxford OX1 3RH, England
`r.hyder1@physics.ox.ac.uk`

8.1 Introduction

The accelerator tube fulfils two complementary functions. As a high-vacuum insulator, it is required to support the highest possible fields. As a beam transport element, it must transmit and focus the particle beam with minimum degradation and loss.

The ideal tube, as an insulator, has the following properties:

- (a) it can sustain the same voltage gradient as the gas that surrounds the terminal and column, and the insulators which support them;
- (b) it is undamaged by sparks and transients;
- (c) it can operate at full gradient with little or no conditioning.

As a beam transport element,

- (d) it should transmit intense beams of high emittance without loss;
- (e) the focusing action must be predictable and allow for good transmission of beams with different masses, energies and charge states;
- (f) scattering should be small and aberrations unimportant;
- (g) it must generate negligible ionizing radiation, with or without beam, and
- (h) it should be unaffected by fluctuations and inhomogeneity in the column gradient.

In addition, it should have a high vacuum conductance, be rugged, compact and cheap, and last for ever.

Practical accelerator tubes fall far short of the performance of this paragon, despite seventy years of progress in understanding beam optics and vacuum insulation. Nevertheless, tube design has advanced, in spite of the difficulty of carrying out systematic studies and potentially destructive tests on operating accelerators. To these problems of cost and accessibility must be added the difficulty of interpreting data from rapid breakdowns occurring inside a closed pressure vessel.

In spite of this, tubes are now available which meet the tight specifications for voltage-holding, beam transport and reliability demanded in modern analytical and industrial accelerators.

8.2 Physical Processes Occurring in Vacuum High-Voltage Systems

An accelerator tube is made up of one or more annular insulators, bonded to metal electrodes, across which the high voltage is applied to accelerate particles through a series of central apertures. In the earliest, air-insulated accelerators, the insulators were long glass or porcelain tubes sealed to metal disks that carried cylindrical electrodes, almost as long as the insulators, through which the beam was focused and accelerated. To avoid tracking in damp air, the maximum field along the outside of the insulator was very much lower than that in the gap between the cylindrical electrodes. The importance of subdividing long insulators into short rings capable of sustaining higher fields was early recognized by Breit at the Department of Terrestrial Magnetism in Washington, even before the first nuclear experiment with an accelerator or the first use of compressed gas as an insulator. The essential features of such a multielement tube are shown in Fig. 8.1. The electrodes are thin relative to their spacing and may be flat or dished, in order to decouple the beam from the insulator. The insulators may be cylindrical or convoluted on the inner surface. The electrodes are bonded to the insulators either by a thin layer of thermoplastic adhesive, such as polyvinyl acetate, or by a diffusion bond of aluminum foil formed under sustained high pressure and temperature.

The processes that lead to breakdown and therefore limit the maximum operating field of an accelerator tube have been discussed by Hyder [2], Chatterton [3], Juttner [4], Latham [5] and Joy [6], among others. In clean,

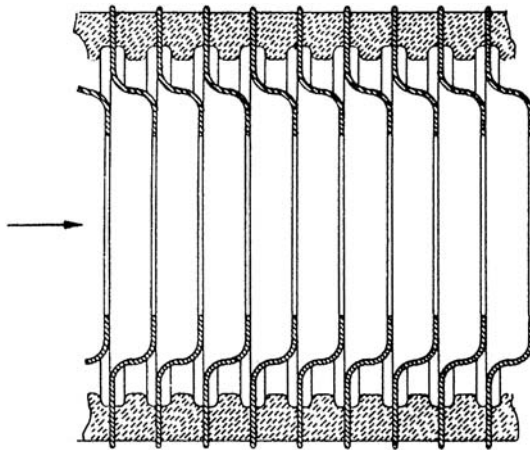


Fig. 8.1. Conventional one-inch-pitch accelerator tube, from [1]. Insulators: borosilicate glass with convoluted profile. Electrodes: dished, polished aluminum. Bond.: thermoplastic polyvinyl acetate resin

dry compressed gas, breakdown across the external surface of the insulator only occurs if the surface is defective or if a voltage surge raises the field instantaneously high above the working value. Similarly, the high dielectric strength of the insulator precludes volume breakdown unless there are defects in the bond between insulator and electrode or in the insulator itself. The areas of concern are therefore the inner surface of the insulator and the vacuum space between the electrodes.

8.2.1 Surface Effects

The tangential field across the surface of a cylindrical insulator between two plane electrodes will be uniform if the surface resistance is constant and the surface charge zero or uniform. If the surface is convoluted the field will vary. Electrons are always present to some degree, and the applied field will accelerate some of them towards the insulator. On their striking the insulator, secondary-electrons may be ejected. If the secondary-electron coefficient is less than one, the surface becomes negatively charged; if greater than one, a positive charge develops. Since the secondary emission coefficient is energy-dependent, and since electrons that strike near the cathode will tend to have lower energy than those near the anode, the surface charge, and consequently the tangential field, will not be uniform. Ion bombardment also liberates secondary electrons, and when electrical activity liberates ions from the electrodes, surface charging of the insulators will increase. Some possible paths for secondary particles are indicated in Fig. 8.2.

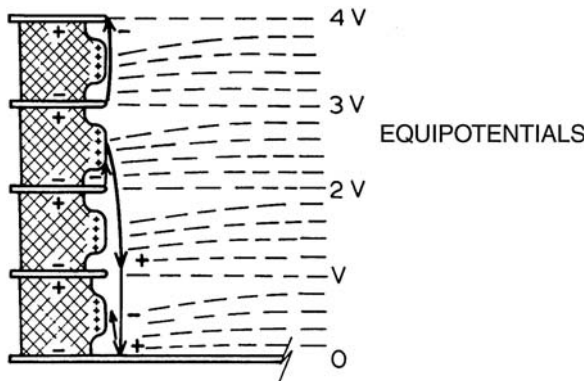


Fig. 8.2. Paths of secondary electrons and ions, showing field enhancement due to buildup of surface charge on the insulators, from [2]

An even weaker point than the insulator surface is the triple junction between vacuum, insulator and cathode. The possible situation in a tube with glass insulators and a polyvinyl acetate (PVA) bond is shown in Fig. 8.3. If

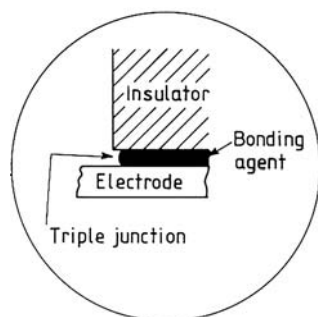


Fig. 8.3. High field near the triple junction, resulting from the glue film having a low dielectric constant and being recessed behind the insulator (Reprinted from [6], copyright (1990) with permission from Elsevier)

the glue film does not extend beyond the insulator, there is a region between insulator and cathode where the field is higher than that in the glass by a factor ϵ ($=4-5$), the dielectric constant of borosilicate glass. This leads to enhanced emission of electrons from the triple junction and thus increased surface charging of the insulator. If the glue extends beyond the insulator, as it usually does, the high field is eliminated but hydrocarbon polymer is exposed to the field and to secondary particles in the main gap. In either case this is a weak point in the system.

The alternative method of construction relies on a thin aluminum foil placed between and diffusion-bonded to the electrode and the ceramic (high-density alumina) insulator. If the foil does not reach beyond the insulator, there is likely to be a microscopic void between the insulator and electrode at the edge of the foil with a field enhancement of as much as 8-9 (the dielectric constant of alumina). The edge of the foil may further increase the field at its junction with the insulator. If the foil protrudes beyond and does not adhere to the electrode, its sharp edge may again act as a stress raiser unless chemical or other means have been successful in blending it into the electrode. Any enhanced field in this region will act as a source of electrons to charge up the insulator surface.

One of the familiar symptoms of these problems is the presence of hair-line track marks across the insulators of used tubes, caused by a discharge of the limited energy stored between adjacent electrodes. Such track marks are usually superficial and have little or no effect on the voltage-holding capability of the insulator. However, repeated discharges may eventually result in glass spalling and damage to the bond material, culminating in failure of the section to hold voltage, as seen in Fig. 8.4. Damage of this nature is often confined to sections near the terminal or the intershield, if there is one, where tank sparks induce large and rapid transients and overcome the protection offered by the spark gaps.

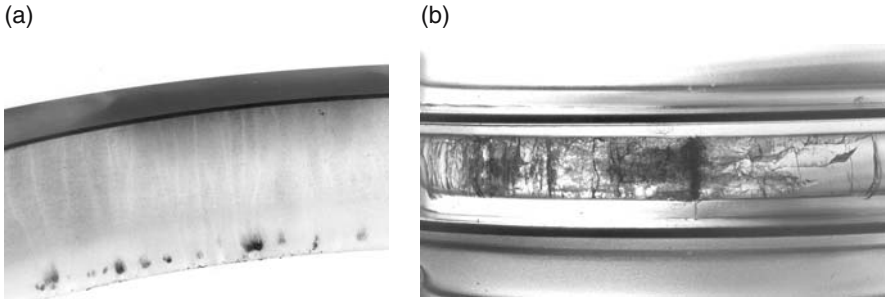


Fig. 8.4. (a) Hairline tracks and aluminum deposits sputtered from the cathode. (b) Glass damage resulting from arc discharges and tracking

One answer to these problems is to modify the shape of the insulator surface. Replacing a straight cylindrical Lucite insulator by a truncated cone was shown to improve voltage-holding between polished copper electrodes subjected to pulsed voltages, as seen in Fig. 8.5 [7,8]. The effect is similar whether the insulator tapers towards the cathode or the anode, but the conditions are very different from those in an accelerator tube. A more effective practice is to convolute the inner surface of the insulator so as to increase the tracking length and to create regions of low field. Care has to be taken to ensure that these convolutions do not actually increase the field at the triple junction and to ensure adequate protection against rapidly rising overvoltages from tank sparks, which can shatter glass if not diverted.

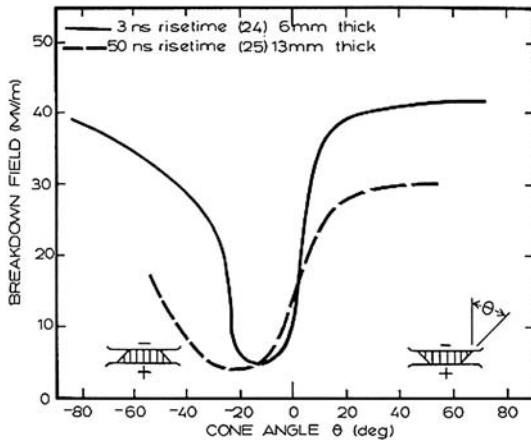


Fig. 8.5. Impulse breakdown strengths of conical Lucite insulators separating polished copper electrodes, as a function of cone angle (Reprinted from [3], copyright (1984) with permission from Elsevier)

8.2.2 Discharges in Vacuum

Juttner [4] has reviewed the different initiating processes that precede breakdown. He divides breakdown into a prebreakdown stage, an ignition stage, a current growth stage and an arc stage. There is a sharp division between the prebreakdown stage, with an upper current limit of a few mA, and the high-voltage breakdown or low-voltage arc, with a minimum current of several A. The former can be sustained solely by electron emission. The latter requires a plasma to develop, as shown in Fig. 8.6.

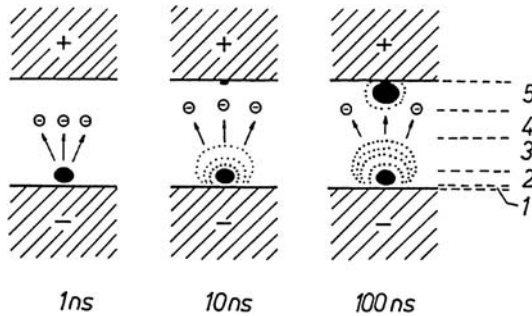


Fig. 8.6. Discharge development in different regions of the gap, from [4]: 1, space charge sheath; 2, plasma flare; 3, expanding plasma; 4, vacuum zone; 5, anode flare (Reprinted from [4], copyright (1988) with permission from Elsevier)

Electron emission from metal surfaces is described by the Fowler–Nordheim law, $i = AV^2 \exp(-B/V)$, where V is the gap voltage and A and B depend on the geometry, material and surface condition of the electrodes. At the fields of 1–3 MV/m typical of accelerator tubes, Fowler–Nordheim currents are very small; field enhancements of 10^2 to 10^4 are needed to produce currents sufficient to initiate breakdown. Sharp points or edges on the metal surface might be expected to raise the field sufficiently to emit copiously, but care is taken in manufacturing to achieve a smooth finish free from asperities. Electrodes that suffer discharges in operation, however, may suffer irreversible damage in the form of arc craters with sharp rims. However, electron-microscopic studies of high-voltage test electrodes have shown that electron emission often arises from point sources where the surface is smooth and featureless. It is assumed that these emitters are at the edges of small regions of insulating material where the potential barrier preventing emission is lowered and the energy of the electrons is increased to allow tunneling through the barrier [6], as indicated in Fig. 8.7. Adsorbed gas on the electrode surface may be one of the sources of these inclusions. Small loosely bound particles, especially conductors, may also contribute to electron emission and to the subsequent development of a full arc.

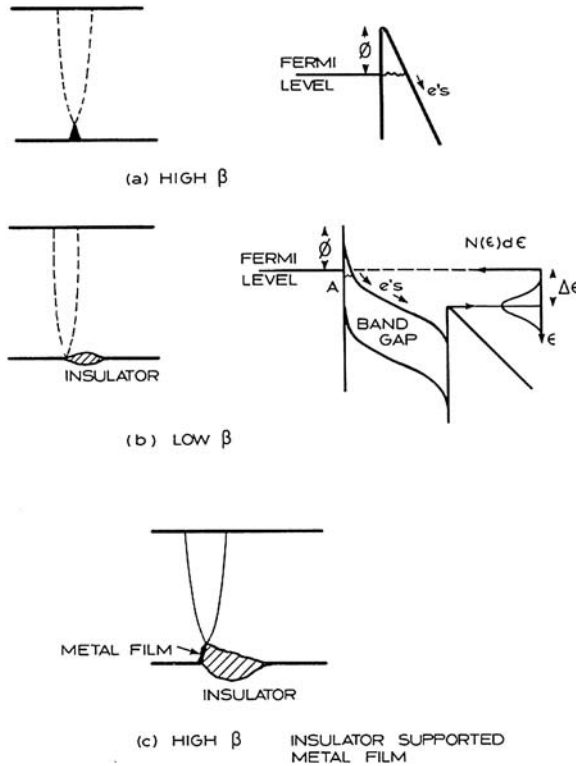


Fig. 8.7. Possible emitter structures and electron energy levels (Reprinted from [3], copyright (1984) with permission from Elsevier)

However carefully electrodes are polished and cleaned, the residual pressure in an accelerator tube, due to permeation, outgassing and particle bombardment, is likely to remain in the range 10^{-3} to 10^{-5} Pa. At these pressures monolayers will form on electrode and insulator surfaces in seconds. The surfaces will always be covered with a layer of weakly bound adsorbed gas molecules, in addition to any debris or impurities left behind from manufacture or cleaning. As the gap voltage is raised, the probability of stray ions gaining energy and releasing particles of the opposite sign on impact increases. If the number of negative ions released for each positive-ion impact is K^- and the number of positive ions from each negative-ion impact K^+ , then, when $K^-K^+ > 1$, a divergent chain reaction will take place, releasing an increasing quantity of neutral gas into the gap, as well as the ions which drive the reaction. Electrons will also take part, adding to the current, but the ions are responsible for most of the gas release. As the current grows, the voltage across the gap (fed from the resistor chain, which is a high-impedance source) decreases, slowing the process. At the same time, the surface gas density declines, multiplication ceases and the process ends.

The theory of these microdischarges is due to Gerasimenko [9], who calculated typical durations of a few hundred μs . Experiments by Schefer and Chatterton [10] have confirmed and extended this model, drawing attention to the effect of surface contaminants such as carbon on the multiplication factor of the chain reaction. At the end of the discharge, some molecular contaminants will have been dissociated and some of the desorbed gas pumped away. The electrode is subsequently able to withstand a higher field.

The relative importance of emission from asperities, from dielectric inclusions, from clumps and particles and from adsorbed gases is disputed and must, in any case, be dependent on the materials of construction, the geometry and even the operating procedure.

8.2.3 Total-Voltage Effects

The processes described above occur within a region composed of a single insulator ring bonded at each end to electrodes. In this region the voltage is limited to $\sim 50\text{ kV}$ or less, except for transient overvoltages following machine sparks. But, as ambitious designers soon found, attempts to reach higher voltages by increasing the column length and maintaining the same potential gradient were unsuccessful as soon as the tube was installed. Something was happening in the tube that depended on the total voltage, not just the field.

When a beam is present, scattering by the residual gas will give rise to ions and electrons. Some of the ions may hit electrodes in the tube, triggering energy-dependent processes leading to breakdown. Electrons stopped by the tube electrodes or beyond the end of the tube will produce bremsstrahlung, the intensity depending linearly on the atomic number of the target and more strongly on the electron energy. The resulting ionization is concentrated near the positive end of the column and consequently perturbs the column gradient. It may eventually increase enough to exceed the output of the charging system, causing the voltage to collapse. Even in the absence of a beam, stray ions and electrons may enter the tube from outside or be released near the edges of beam apertures. Some may travel long distances, gaining enough energy to initiate discharges. Measures designed to reduce or eliminate these effects are discussed in Sect. 8.3.

In 1952 Cranberg [11] reviewed the published data on breakdown voltage across gaps ranging in length from 0.1 mm to 5 m . He suggested that the observed reduction in breakdown field with increasing gap length could be due to the presence of small clumps of loosely bound material. Such clumps might, by electrostatic repulsion, be injected into the gap and gain enough energy to evaporate hundreds of atoms at the point of impact, resulting in an arc discharge.

His theory predicted that the breakdown voltage would vary as the square root of the gap length for a given pair of electrodes, in rough agreement with the existing data. Subsequent work has identified several possible processes,

depending on particle mass and terminal velocity [12]. The critical parameter is the ratio of the terminal velocity v_t to the plastic velocity v_p of the electrodes. Below v_p particles rebound elastically; above v_p the collision is inelastic, resulting in gas desorption and sometimes melting, crater formation and evaporation. The plastic velocity depends only on the yield strength and density of the material; see Table 8.2.

Considering the case where the total voltage available is, at most, that across a few pitches, four types of event can be distinguished in order of increasing radius r :

- (i) $v_t \gg v_p$, and typically $r < 0.1 \mu\text{m}$. Such particles vaporize on impact, but the number of neutrals and ions released is too small to initiate breakdown.
- (ii) $v_t \geq v_p$, $0.1 < r < 10 \mu\text{m}$. These particles can cause local melting and evaporation, liberating neutrals, ions and liquid droplets. The more energetic ones may produce enough gas and ionization to trigger breakdown. Field enhancement at crater lips and protrusions may give rise to cathode instability.
- (iii) $v_t < v_p$, $10 < r < 50 \mu\text{m}$. These particles are too slow to trigger breakdown in a single transit. Multiple bouncing impacts with charge exchange might increase their energy to bring them into category (ii).
- (iv) $v_t \ll v_p$, $r > 50 \mu\text{m}$. If such a large, slow particle approaches a cathode (or anode) protrusion, the enhanced field in the gap between particle and protrusion may result in enough current flow for melting and evaporation to take place by the Joule or the Nottingham effect.

Many observations, mostly in gaps of a few mm, have been made with the object of clarifying the importance of these processes; see, for example, Chatterton and Eastham [13]. These authors found that, after careful surface treatment and thorough cleaning, large microparticles are rare and multiple transits and bouncing unimportant. By contrast, small particles are abundant and often appear to be weakly bound [14]. These seem to give rise to more frequent breakdowns than would be expected on the basis of the above classification. Spark conditioning, however, can reduce the microparticle yield to zero except when the gap is on the verge of breakdown.

Like microdischarges, which depend on ion exchange, microparticle processes are energy-dependent and can be reduced, if not eliminated, by careful conditioning and limiting the maximum energy which can be gained in a single transit. Using modern techniques to suppress secondary particles, Cranberg's square root dependence has given way to an almost linear relation of terminal voltage to column length over the range 5–25 MV.

8.3 Beam Optics

At its simplest, the optical system of an accelerator tube consists of a strong converging entrance lens, a region of uniform longitudinal field and a weak diverging exit lens. Elkind [15] has given algebraic expressions for the first-order focusing and magnification of such a tube. More detailed treatments, using finite-element techniques to calculate field distributions and transfer matrices to handle finite-emittance beams, are due to Galejs and Rose [1] and Stenning and Trowbridge [16], among others.

The strength of the entrance lens depends on the injection energy of the beam and on the field in the tube. In many accelerators the terminal voltage, and consequently the field in the tube, may vary over a range of ten to one. To compensate for this the injection energy, or the position of the tube object, must also vary.

Because space is limited, a terminal ion source is usually close to the tube entrance. The usual practice is to preaccelerate the diverging beam from the source so as to keep the focal conditions constant as the terminal voltage changes. Within the tube, the beam may converge towards an external focus, remain parallel or even diverge slightly as long as it remains smaller than the tube apertures, whose size is limited by the need to intercept the secondary particles that trigger breakdowns. On leaving the accelerator tube, the beam may be refocused by external lenses onto a target or an analyzer magnet, reducing the high magnification that results from the proximity of ion source and tube entrance. The magnification may also be reduced by lowering the field near the tube entrance.

The beam that enters the low-energy tubes of a tandem must be brought to a focus in the stripper, close to the tube exit. Preaccelerating this beam to match it to the tandem over a range of terminal voltages means mounting the ion source on a high-voltage platform. This can be avoided, if space is available, by varying the tube object position with a zoom lens, but at the cost of changing the beam radius at the stripper. The problem can be overcome by the use of a gridded immersion lens at the tube entrance or a gridded einzel lens just before it. Another solution is the use of a “Q snout”, a method of preaccelerating the beam before it enters the main part of the tube.

The entrance lens is the most critical optical element in the tube. It is a strong lens with significant aberration; it operates on a low-energy beam; its aperture radius limits the acceptance. The temptation to inject high-emittance beams which occupy too large a fraction of this aperture is hard to resist. It is therefore important to know how much of the aperture can be filled before aberrations degrade the image or reduce the transmission.

Using modern finite-element field computation programs, Colman and Legge [17] and Trowbridge et al. [18] have calculated aberrations for simple geometries in the absence of space charge. Such programs can be used to improve the electrode geometry before and after the entrance aperture. The

exact focal power of the lens is not critical, since this can be adjusted by varying the injection energy or object position. Aberrations, however, increase the emittance of the beam, worsening the size of the beam on target and risking transmission loss.

After stripping, a tandem beam enters the high-energy tubes. At this point it is small in diameter and divergent, but the emittance will have increased owing to scattering, especially for foil-stripped heavy ions. The focusing action of the entrance lens to the high-energy tube is independent of terminal voltage, as the ratio of injection energy to field is constant. But the strength varies with the charge state of the beam, being small for singly charged ions, and more important for high-charge-state heavy ions, where it helps to compensate stripper scattering.

The assumption that the field inside the accelerator tube is uniform is only true if:

- (a) the grading resistors are all equal, and
- (b) there are no dead sections, and
- (c) the electrodes are thin.

Tolerances on high-voltage resistors are rarely as low as 1% and may rise in use to 5–10%. Surge damage can result in even larger decreases and occasionally increases big enough to cause breakdown. The effect of such random changes on beam focus are difficult to calculate but are usually small. But in inclined-field tubes, they may deflect the beam significantly.

Dead sections occur at tube joints but may also be introduced deliberately to modulate the axial field. The radial fields so generated are strong enough to suppress low-energy electrons (and ions), but too weak to have much effect on the main beam.

The focusing action of thick electrodes has been studied by Galejs and Rose [1] and Trowbridge et al. [18]. For typical geometries where $t/p < 0.1$ (t is the electrode thickness and p is the pitch), the effects are small unless low-energy beam particles are allowed to graze the edges of the electrodes.

8.4 Suppression Systems

As accelerator development led to higher terminal voltages, it became imperative to reduce or eliminate the growth of secondary-electron currents. Some success was achieved by tapering the beam apertures in such a way as to intercept at least a proportion of secondaries released within the tube [19]. Another technique, with little to recommend it, was to increase the pressure in the tube so as to scatter low-energy particles onto neighboring electrodes. The first satisfactory answer to the problem was the proposal, by Van de Graaff, to introduce transverse electrostatic fields arranged so as to sweep low-energy particles out of the beam aperture but leave the beam itself undeflected [20]. In Van de Graaff's scheme, a group of electrodes, inclined so

as to produce an upward component of the field (in a horizontal accelerator), would be followed by a slightly longer group inclined downwards. By matching the lengths of successive sections to the velocity profile of the beam, the energetic primary beam could be kept close to, and would exit on, the axis. In contrast, electrons born within these sections would be swept onto the electrodes before gaining more than a few hundred keV energy. Only electrons born in the transition regions between upward and downward fields would travel long distances. A typical electrode arrangement in such an inclined-field tube is shown in Fig. 8.8. Allowance must be made, in the alternating-field geometry, for the astigmatic focusing of the slot apertures in the electrodes and the prismatic field at the transitions. These effects have been discussed by Serbinov [21] and Koltay [22].

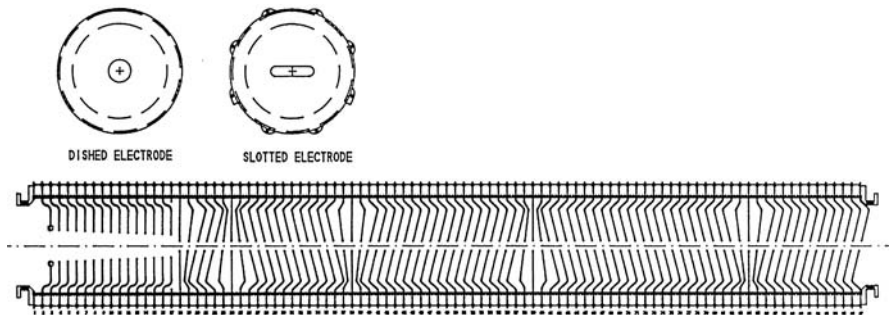


Fig. 8.8. Electrode arrangement of an HVEC inclined-field entrance tube for an MP tandem. The beam enters from the *left* and passes through a noninclined section with circular apertures, then through five inclined-field sections with slotted apertures

At about the same time, Allen [23] proposed an alternative arrangement in which the inclination direction of successive electrodes was rotated azimuthally in such a way as to suppress secondaries. By changing the rate and sense of the azimuthal rotation after each complete turn, the beam could be made to leave on axis and all secondaries could be suppressed, as indicated in Fig. 8.9. As in Van de Graaff's scheme, absolute compensation of the beam displacement is only achieved for a specific velocity profile, but for the range of velocities observed in practice the mismatch is sufficiently small for it to be corrected by external deflectors; see Fig. 8.10. Both schemes depend on a uniform field in the column and are affected by perturbations such as beam loading or damaged resistors.

A little later Howe [25] developed a third system, involving the use of permanent magnets mounted on the electrodes. Originally ring magnets, magnetized across a diameter, were mounted inside the vacuum system to produce a field on axis of ~ 0.01 T. The field direction rotated on successive

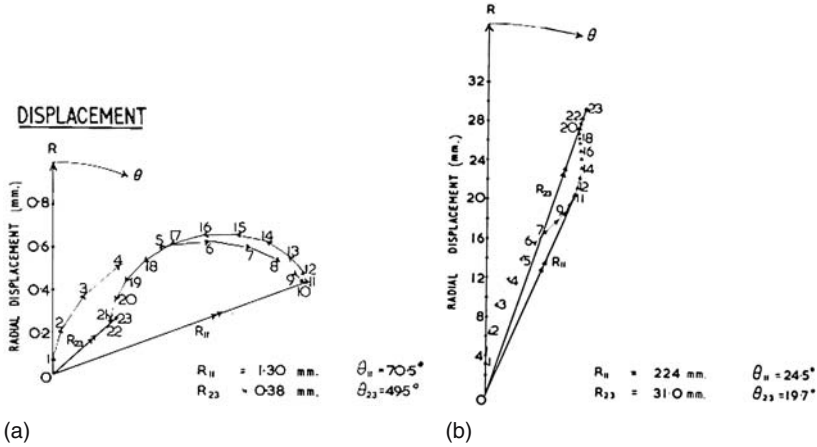


Fig. 8.9. Transverse displacements of the accelerated beam (a) and low-energy secondary electrons (b) after traveling through two sections of a spiral inclined-field tube (Reprinted from [24], copyright (1973) with permission from Daresbury Laboratory)

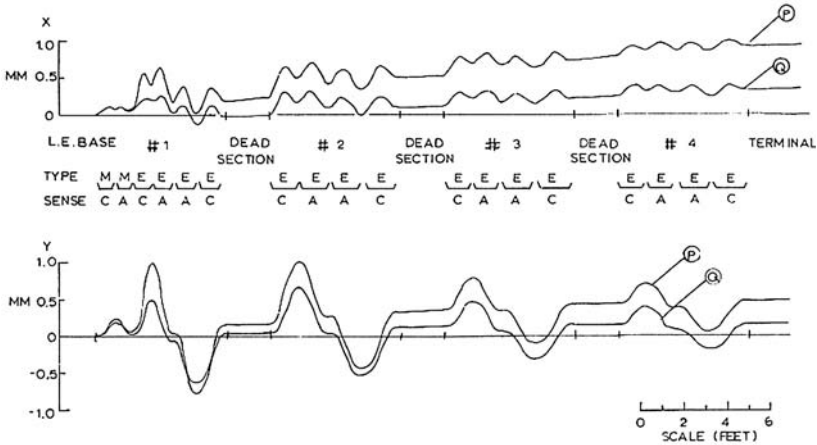


Fig. 8.10. Trajectories of axial rays through the low-energy spiral IF tubes of an MP tandem. Upper figure: horizontal plane. Lower figure: vertical plane. Key: TYPE: M, magnetic section; E, electrostatic section. SENSE: C, clockwise field rotation; A, anticlockwise field rotation. Rays: P, mass = 1, charge = $1-$, injection energy = 0.1 MeV, terminal voltage = 12 MV; Q, mass = 1, charge = $1-$, injection energy = 0.2 MeV, terminal voltage = 2 MV (Reprinted from [24], copyright (1973) with permission from Daresbury Laboratory)

electrodes so as to cancel the deflection of the beam while ensuring full electron suppression. Because the transverse impulse is independent of velocity, the final transverse momentum of the beam can be nulled for all velocity profiles. In any case, the rather weak fields needed to suppress electrons have a minimal effect on ions. A similar system, using bar magnets mounted on the edges of the electrodes outside the vacuum, has been used with Van de Graaff's inclined-field tubes to improve electron suppression at the transition points. Howe's original ring magnets have long since been replaced by compact high-coercivity Sm-Co or rare-earth bar magnets mounted inside the vacuum envelope.

A very different technique, avoiding the use of transverse fields, has been developed by Herb for the accelerators produced by NEC [26, 27]. The ceramic/titanium tubes used in these machines are made up of short sections joined together by bolted flanges. The axial field therefore varies periodically from a maximum in the middle of each section to a minimum opposite the flange joint. By suitable design of the electrodes on either side of the section joints, the field can be shaped so as to deflect any electrons released from these apertures onto nearby electrodes; see Fig. 8.11. Because the field is axially symmetric, particles on or near the axis will be transmitted, but secondary particles from the electrodes and divergent scattered particles are mostly removed, as shown in Fig. 8.12. High operating fields can be attained if the vacuum is sufficiently good.

Incorporating any of these suppression systems into an accelerator tube complicates the beam optics calculations. Computing the first-order displacement of the beam caused by the alternating or rotating inclined fields of the Van de Graaff or Allen system is straightforward. Assuming the field to be uniform across the beam aperture, it is sufficient to track the path of the axial ray as it passes through the suppression system with a known velocity profile. It is normal practice to start the inclined-field sections after a straight section of 15 to 20 electrodes that is usually magnetically suppressed and operated at reduced gradient. Recently, internal bar magnets have been incorporated in inclined-field tubes and in tubes with axial field modulation in order to reinforce electron suppression.

8.5 Design and Construction

The accelerator tube is a precision device working in a harsh environment. The designer's first task is obviously to ensure adequate mechanical strength. In horizontal machines, the tube is sometimes cantilevered from the base, having to support its own weight and sometimes an ion source and lenses. In horizontal tandems, tubes as long as 2.4 m will be simply supported solely at their ends. In vertical machines, the weight of the whole tube and the pressure of the gas will bear on the base of the tube. The insulators must also resist

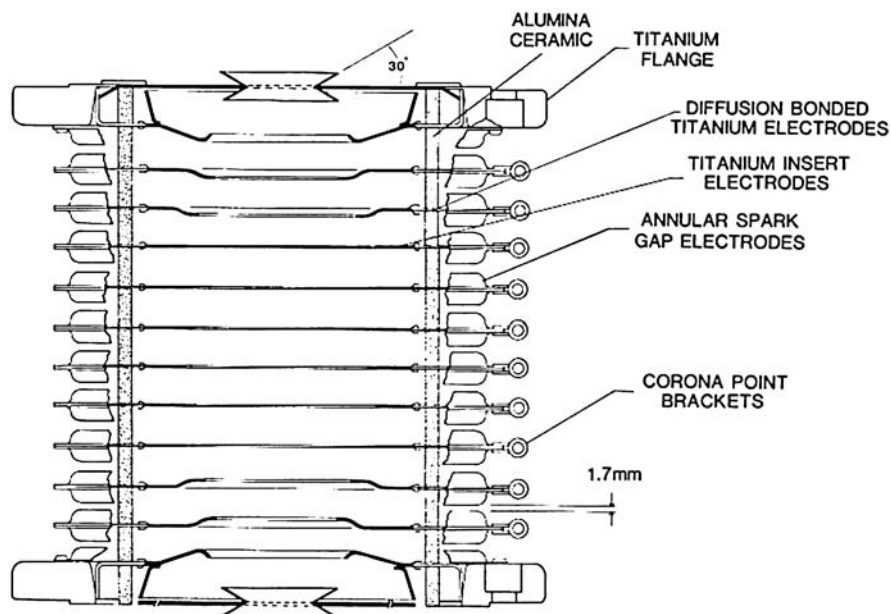


Fig. 8.11. NEC “compressed-geometry” tube section designed for the Oak Ridge 25 MV tandem, showing the shaped electrodes at the section ends that determine the electron-suppressing fields (Reprinted from [27], copyright (1988) with permission from Elsevier)

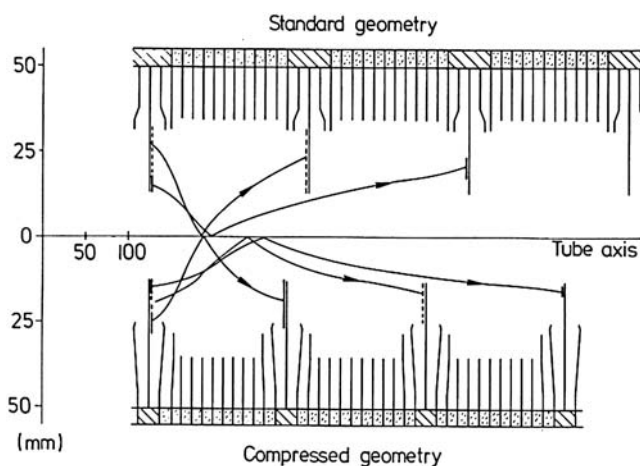


Fig. 8.12. Electron trajectories in standard and compressed-geometry NEC tubes operating at a gradient of 330 kV per tube section. Secondary electrons released from the end electrodes are captured on subsequent end electrodes before gaining excessive energy (Reprinted from [26], copyright (1984) with permission from Elsevier)

the radial force due to gas at pressures of up to 2 MPa. Deflections resulting from the cyclical variation of gas pressure must be small and reversible, consistent with the accuracy required by the beam optics. The overall length and diameter must be compatible with the layout of the column and the need for access for installation, assembly and maintenance.

The choices of materials and of the detailed design of insulators and electrodes have been made in the light of the operating conditions and physical processes already described. Vacuum conductance is clearly of major importance. The tube must tolerate variations in temperature, humidity and pressure during transport. In service, it must withstand mechanical shock during tank sparks and occasionally tremors due to earthquakes.

8.5.1 Insulators

The superior vacuum properties of glass and ceramics ensure their use in preference to plastic insulators, although these have adequate dielectric strength and resistivity.

The relevant properties of the two most widely used materials are summarized in Table 8.1. The better electrical properties of glass, coupled with adequate mechanical strength, would make it always the material of choice if it were not for the possibility of making ultra-high-vacuum, organic-free, bonds between alumina and titanium.

Table 8.1. Properties of tube insulators. Note: dielectric strengths were measured on 3 mm samples; see Aitken [28]

Insulator	Borosilicate Glass	High-density Alumina
Thermal expansion ($^{\circ}\text{C}^{-1}$)	3.8×10^{-6}	7.6×10^{-6}
Elastic modulus (GN/m^2)	68	344
Breaking strength (MPa)	35–140	360
Log volume resistivity ($\Omega \text{ cm}$)	$>10^{15}$	$>10^{14}$
Dielectric constant	4.2	9.5
Dielectric strength (MV/m)	120–180	50–70
Secondary-emission coefficient	3 (max) at 350 eV	8 (max) at 600 eV

Glass insulators are made from castings that can be inspected optically for bubbles, strings (of impurity) and freedom from stress. After annealing, they are ground flat on both end surfaces and on the interior, usually to a profile designed to increase tracking length and minimize surface charge. Flatness is carefully controlled so as to ensure uniform glue film thickness. The most usual pitch is 25 mm, but larger pitches have been specified for tubes working at modest fields. Diameters range from ~ 100 mm in some small accelerators to over 300 mm for large machines and in applications where good vacuum is critically important.

Ceramic insulators are made from pressed high-density alumina, which is then sintered at high temperature and ground flat on the ends. Control of the manufacturing process is very important in order to avoid small voids that may not be revealed by nondestructive tests. Most tubes using ceramic insulators have a pitch of 12.7 mm and a diameter of about 100 mm. The small pitch is beneficial in reducing the energy acquired by ions in interelectrode processes and in reducing the influence of the high secondary-emission coefficient on surface charging. The maximum safe diameter is limited by the differential thermal expansion between ceramic and metal and the high temperature used in the bonding process.

8.5.2 Electrodes

In the past a variety of metals have been used in accelerator tubes, including copper, which is easily cleaned and polished, and refractory metals such as molybdenum, which resist sputtering and melting in discharges. Experience gained in extensive laboratory tests, combined with the need to reduce bremsstrahlung by using materials of low atomic number, have made stainless steel and titanium the preferred choice, with aluminum a low-cost alternative for applications where heavy ion bombardment and arcing can be discounted. The relevant properties are shown in Table 8.2.

Table 8.2. Properties of tube electrodes

Electrode	Aluminum	Stainless Steel	Titanium
Atomic number	13	26.2	22
Thermal expansion ($^{\circ}\text{C}^{-1}$)	2.3×10^{-5}	9.3×10^{-6}	7.6×10^{-6}
Elastic modulus (GN/m^2)	69	193	110
Yield strength (MPa)	145	760	830
Melting point ($^{\circ}\text{C}$)	660	1530	1800
Plastic velocity (m/s)	320	530	1200

In early accelerator tubes, the long pitch of the insulators and difficulties in achieving clean vacua encouraged the designers to resort to complex re-entrant shapes for the electrodes so as to prevent scattered particles hitting the insulators and eliminate interactions between the beam and the surface charges. In today's designs, electrodes are usually thin and, if not actually flat, pressed into a dish shape or inclined at an angle to sweep secondary particles off axis. Sometimes the electrode is in two parts: a flat annulus bonded on either side to the insulator, and a removable, central insert that defines the beam aperture, intercepts unwanted secondaries and can be removed for cleaning. The aim is always to keep the peak field between adjacent electrodes as low as possible. The outer part of the electrode must be flat enough or

flexible enough to ensure a strong, vacuum-tight bond without voids. Aluminum electrodes are 2–3 mm thick, stainless steel and titanium electrodes 0.5–1.0 mm. The electrodes normally extend a few mm beyond the outer edge of the insulator and carry protective spark gaps. These may consist of six or more sphere gaps, 3–4 mm in radius, uniformly disposed around the circumference, or a pressed ring forming an annular gap close to the outside of the insulator. The spark gaps are set to break down at a slightly lower voltage than the gaps protecting the column, the actual setting depending on the rated field in the column and the expected gas composition and pressure.

8.5.3 Assembly

Tubes with glass insulators use thermoplastic resin, usually polyvinyl acetate, as a bonding material. The resin is dissolved in a suitable solvent and a controlled quantity is then deposited on both sides of the electrodes. After solvent evaporation, the end flanges, insulators and electrodes are stacked in a jig, heated in a tube oven to the softening point of the resin, compressed and then allowed to cool slowly at a controlled rate. The jig is designed to keep the tube straight and accurately aligned. The heating and cooling cycle ensures that strain, locked into the assembly because of differential thermal expansion, is kept to a minimum. Tubes as long as 2.4 m can be assembled in a single operation. A single section therefore suffices for accelerators rated at 3 MeV or even more.

Ceramic/titanium tubes rely on diffusion bonding. In this process aluminum foil, 0.1 mm thick, is cut to the same shape as the insulator ring and interposed between insulator and electrode. Assemblies, typically containing 13 insulators and electrodes, and titanium end flanges, are then jigged and placed in an oven, in which they are compressed and held at a temperature just below the melting point of aluminum. In time, the aluminum diffuses into the ceramic to form a strong, vacuum-tight bond. The temperature cycle, compressive force and ambient gas are all tightly controlled. The short length of single sections requires that all but the smallest machines will incorporate several units, introducing dead sections into the column.

8.6 Vacuum

Vacuum conditions affect tube performance in several ways. Beam losses due to scattering with residual gas are important for very intense beams, since the scattered particles may load the column and upset the gradient. Vacuum conditions are also critical where negative ions are injected, because they have large cross sections at low energy for charge exchange. The problem is especially acute in accelerator mass spectrometry, where loss-free transmission is needed for accurate measurement of isotopic intensities. Low residual pressure is also desirable because it reduces the amount of adsorbed gases

on electrodes and hence the frequency and intensity of microdischarges. Finally, the composition of the residual gas has an important influence on the production of sputtered ions and hence on the threshold for microdischarge activity.

Cockcroft and Walton were among the first users of the low-vapor-pressure hydrocarbon oils developed by Metropolitan-Vickers for diffusion pumps. However, backstreaming oil vapor was soon under suspicion as a cause of electron emission and was successively replaced as a working fluid in accelerator applications by mercury, which required trapping with liquid air, then by silicon-based fluids, and more recently by polyesters and other fluids with very low vapor pressures and improved resistance to chemical attack. These new fluids have led to a reduction in ultimate pressure of two orders of magnitude, but they pose the risk of contaminating the tube in a vacuum accident. In contrast to hydrocarbons, which form conducting surface layers, silicone oils act as insulators and can upset the functioning of high-voltage electrodes by allowing surface charges to build up and distort the field.

Complete freedom from oil contamination can be achieved at high vacuum by the use of cryopumps or sputter ion pumps. Initial evacuation requires the use of sorption traps or, more conveniently, hybrid turbomolecular/drag pumps, whose high compression ratio for heavy molecules ensures negligible backstreaming. In practice, modern turbomolecular pumps, even those with greased bearings, release negligible oil vapor and can be used as recirculators in high-voltage terminals.

It is very easy to fit a high-speed pump outside the accelerator tank and measure very low pressures above it. But gas sources inside the accelerator, such as ion source or stripper gas in the terminal, and outgassing and desorption anywhere, must be pumped through the meager conductance of the tubes. What matters is the pressure near the terminal, and this may be as much as fifty times higher than that at the pump. The need to limit the path of secondary particles means that most tubes have baffle electrodes that reduce the conductance well below that of a simple tube of the same I.D. For example, the conductance of a standard 1.8 m tube with a central aperture tapering from 63 to 38 mm radius has been measured to be 47 l/s. A more realistic design for a single EN tandem tube, having a central aperture of 12.5 mm radius and additional pumping sectors arranged as in Fig. 8.13, has a measured conductance of 21 l/s. With such tubes, a pressure rise of 10^{-4} Pa at the pump, due to stripper gas, corresponds to a pressure of at least 3×10^{-3} Pa at the terminal, making no allowance for outgassing along the length of the tube. Even higher pressures exist in single-ended machines. In large accelerators, such as MP tandems and NEC machines, the column is punctuated by substantial dead sections, and electrical power is available in some of these locations to energize sputter ion pumps. Conductances and pressure profiles for the large VIVITRON tandem at Strasbourg were reported by Heugel [29].

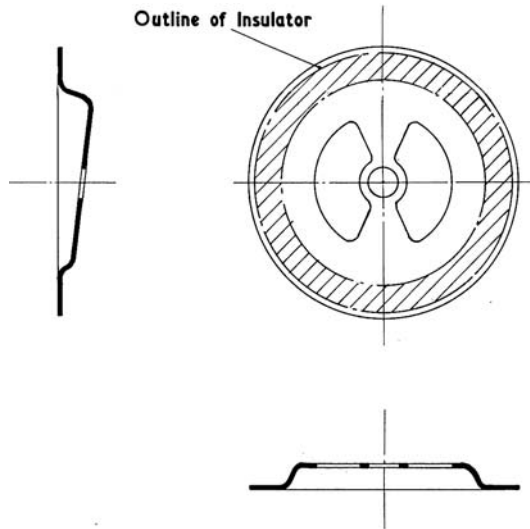


Fig. 8.13. Sections of a spiral IF tube electrode for a model EN tandem, showing pumping cutouts. The measured vacuum conductance of a single 1.8 m section is 211/s

His work emphasizes the gradual improvement in vacuum resulting from the slow decrease in outgassing rate with time.

The need for terminal pumping in large tandems has long been recognized. Sublimation pumps, cryopumps and sputter ion pumps have all been used, with varying degrees of success. Sublimation pumps have limited life and a low pumping speed for inert gases. Cryopumps are bulky and require regeneration, involving passage of all the pumped gases through the tubes. Sputter ion pumps have proved to be reliable and long-lived, provided the gas load is kept low. The use of turbomolecular pumps as stripper gas recirculators was suggested by Purser and Hyder in 1982 [30] and has subsequently been applied to reduce the gas load from terminal ion sources. In some high-current single-ended machines, attention has turned again to the provision of differential tubes, through which most of the gas generated in the terminal can be diverted away from the beam path. Satisfactory designs for such tubes, combining a high conductance with effective electron suppression, have enabled them to return to favor.

8.7 Operating Conditions

Tubes are usually shipped sealed and evacuated to ensure cleanliness and freedom from contamination and should remain so as long as possible. PVA is hygroscopic, and there are reports of glued joints deteriorating when kept

in humid conditions for long (10 years) periods. During installation, tubes are often exposed to atmosphere for long periods while assembly and alignment take place. Good practice then requires that they should be evacuated for as long as possible before voltage is applied. The use of a residual-gas analyzer, if one is available, helps to distinguish between normal outgassing and a leak. In air, the relative intensity of the mass 28 and mass 32 peaks is a useful diagnostic. A leak-tight tube in an accelerator insulated with SF_6 may still exhibit a very small mass 127 peak, but if this varies with pressure the leak must be cured before voltage is applied. It is equally important to ensure that the grading resistors are all within specification.

Initial conditioning is an important process that should only be undertaken when the necessary controls and instruments are fully operational and the electrostatic behavior of the accelerator is satisfactory. The progress of conditioning can be monitored by measuring vacuum pressure, radiation, current balance and particle emission. Typically, the voltage can be increased steadily up to about half the rated maximum before microdischarge activity results in a measurable increase in vacuum pressure. If the voltage is then held constant, the pressure should decrease, almost to the base value. Small fluctuating ion or electron currents may be observed on Faraday cups close to the tube, and radiation levels may rise above background. These effects should also decay away if the voltage is held constant. The normal procedure is next to increase the voltage in small steps, limiting the pressure rise to a few times the base pressure and pausing if the ion currents or radiation levels become erratic. Ideally this sequence should continue until the tube is operating quiescently at or above its rated voltage. In low-voltage accelerators with lead-shielded tanks, external radiation levels may be very low and conventional radiation monitors may need to be supplemented by a gamma spectrometer. In larger machines, the use of a mobile spectrometer may give valuable information about the energy and origin of abnormal sources of bremsstrahlung. Viewing ports can reveal the intensity and distribution of luminosity from microdischarges.

Conditioning to full voltage necessarily involves a higher probability of sparking than does quiescent operation. The risk of damage is much greater in large machines because of the strong dependence of stored energy on voltage. The risk of sparking is also greater because of the chance that many microdischarges will occur simultaneously in different tube sections, leading to excessive pressure rises. The use of shorting rods or cables, enabling individual sections to be taken to voltage with a small fraction of the stored energy of the whole machine, has proved effective in raising the voltage safely in multisectioned machines. Another technique, which is useful in machines of all sizes, is to apply a sawtooth waveform by computer control of the charging current. A typical amplitude for this process is 1–2% of full voltage, with a period long enough for the vacuum to return to normal between peaks.

A tube which has been conditioned to maximum voltage and then operated with beam at or very near that voltage is in a state of dynamic equilibrium. As residual gas is readsorbed on surfaces previously cleaned by microdischarges, the condition for further discharges returns. Small ion currents will flow intermittently, depleting the surface gas layers at the same rate as the readsorption. Adsorbed gas will build up when the voltage is removed and must be removed by reconditioning before returning to full voltage. Contaminants and molecules not present in the residual gas are, however, permanently removed and tube performance will improve in consequence.

Accelerator tubes operating with intense beams near the maximum voltage are at risk of damage if the beam disappears or becomes defocused. In the former case interlocks are required to sense beam loss and take effective corrective action to stop the voltage rising. If the beam becomes defocused, it must be stopped at the tube entrance and the terminal voltage frozen.

Diagnosing tube faults is a demanding part of the operator's duties. Tubes may fail because of internal defects or because of external faults such as vacuum leaks or open-circuit resistors. It is very desirable to establish the nature and location of the problem before the tank is opened, since the symptoms are likely to disappear when the voltage is removed and the gas pressure reduced. Breakdown across an individual insulator very often manifests itself as a sawtooth variation in terminal voltage and particle energy. The amplitude of this fluctuation, as seen by a capacitive pickup looking at the terminal, decreases the farther the faulty section is from the terminal. If the problem is near the baseplate, the long time constant of the terminal capacitance and the column resistors may attenuate it so much as to make it undetectable. Such failures may also generate increased bremsstrahlung with a measurable maximum energy, enabling the source to be identified. Viewing ports opposite the column can be used to reveal sparking or luminosity at the site of the problem.

Another class of faults arises when column gradients are perturbed by radiation, leakage currents or faulty resistors. Attempts to hold the terminal voltage constant then result in the unaffected part of the column being overstressed. If the original fault occurs near an inclined-field tube the gradient error will deflect the beam sideways, sometimes far enough for total beam loss, beyond the correcting power of external deflectors.

Vacuum faults are often signaled by the onset of severe conditioning at abnormally low voltage. Pressure-sensitive leaks, especially of SF_6 , degrade tube performance and may require prolonged conditioning before recovery. Tubes severely contaminated by oils or polymers must usually be removed and reconditioned or rebuilt.

The low-energy X-rays always present inside accelerator tubes are partly absorbed in the tube wall. In glass, this causes darkening due to the formation of color centers. After prolonged operation the glass may become completely

opaque, but without any deterioration of its insulating properties. Ceramic insulators also can sustain this type of radiation damage without detriment.

In the absence of vacuum accidents, contamination and external faults, well-protected tubes can continue to operate normally for ten years or more.

8.8 Conclusions

The modern accelerator tube has come a long way towards fulfilling the requirements outlined in Sect. 8.1. With proper protection it can survive undamaged in the largest electrostatic accelerators, operating at full voltage. It can be brought up to its rated voltage with modest conditioning in a well-defined way. Its focal properties can be predicted accurately; ion currents at mA levels can be transmitted with negligible loss through small machines; multiply charged heavy-ion beams of tens or hundreds of μA are available from large tandems. Small accelerators working below the neutron threshold can be operated in unshielded rooms, with a modest layer of lead surrounding the tank. Voltage stability, vacuum quality, suppression of impurity ions and beam transmission can all meet the challenges of ultrasensitive mass spectrometry and advanced ion implantation.

However, even the best tubes cannot be made to work satisfactorily at fields much in excess of 2 MV/m. At these gradients, surface charges on insulators and ion exchange between electrodes begin to affect stability and increase the probability of breakdown. Experiments have been carried out at 11 MV in an FN tandem, corresponding to a field of 2.25 MV/m over the active length of the tube. The NEC tandem at the Australian National University, Canberra, has been conditioned to 2.6 MV/m and has run experiments at 2.4 MV/m [31]. A few small accelerators have operated at slightly higher fields. In most applications, the length of the tube is not critical. A conservative tube gradient is a small price to pay for ease of operation and reliability.

References

1. A. Galejs, P.H. Rose: Optics of electrostatic accelerator tubes. In: *Focusing of Charged Particles*, vol. 2, ed. by A. Septier (Academic Press, New York and London, 1967) pp. 297–326
2. H.R. McK. Hyder: *Rev. Phys. Appl.* **12**, 1493 (1977)
3. P.A. Chatterton: *Nucl. Instr. Meth. A* **220**, 73 (1984)
4. B. Juttner: *Nucl. Instr. Meth. A* **268**, 390 (1988)
5. R.V. Latham: *Nucl. Instr. Meth. A* **287**, 40 (1990)
6. T. Joy: *Nucl. Instr. Meth. A* **287**, 48 (1990)
7. R.A. Anderson: Report on pulsed vacuum breakdown of plexiglass insulators. Sandia National Laboratory report, SAND 75 0667 (1976)
8. D. Milton: *IEEE Trans.* **EI-7**, 9 (1972)

9. V.I. Gerasimenko: Zh. Tekhn. Fiz. **38**, 155 (1968) (Sov. Phys. Tech. Phys. **13**, 107 (1968))
10. P.V. Schefer, P.A. Chatterton: IEEE Trans. **EI-11**, 12 (1976)
11. L. Cranberg: J. Appl. Phys. **23**, 518 (1952)
12. M.M. Menon, K.D. Srivastava: J. Appl. Phys. **45**, 3832 (1974)
13. D.A. Eastham, P.A. Chatterton: IEEE Trans. **EI-18**, 209 (1983)
14. G.P. Beukema: J. Phys. D **7**, 1740 (1974)
15. M.M. Elkind: Rev. Sci. Instr. **24**, 129 (1953)
16. P.J. Stenning, C.W. Trowbridge: The Pathfinder programme and its application to ion optics. Rutherford Laboratory/Reading University report, RU/RL-1 (1968), <http://www.trowbridge.org.uk/downloads.htm>
17. R.A. Colman, G.J.K. Legge: Nucl. Instr. Meth. B **73**, 561 (1993)
18. C.W. Trowbridge, K. Höffer, H.R.McK. Hyder: Fields, focusing and aberrations in electrostatic accelerator tubes. IEEE Trans. Magn. **40**, 609 (2004)
19. D.R. Chick: Nucl. Instr. Meth. **5**, 209 (1959)
20. R.J. Van de Graaff, P.H. Rose, A.B. Wittkower: Nature **195**, 1292 (1962)
21. A.N. Serbinov: Inst. Exp. Tech. (English translation) **4**, 715 (1967)
22. E. Koltay: Nucl. Instr. Meth. **66**, 253 (1968)
23. W.D. Allen: A new type of accelerating tube for electrostatic generators. Rutherford High Energy Laboratory report, NIRL/R/21 (1962)
24. H.R.McK. Hyder, G. Doucas: Experiences with suppressed accelerator tubes. In: *Proc. First Int. Conf. Tech. Electrostatic Accelerators, Daresbury* (Daresbury Laboratory, DNPL/NSF/R5, 1973) p. 352
25. F.A. Howe: IEEE Trans. **NS-16**, 98 (1969)
26. W. Assmann, G. Korschinek, H. Münzer: Nucl. Instr. Meth. **220**, 86 (1984)
27. C.M. Jones et al.: Nucl. Instr. Meth. A **268**, 361 (1988)
28. T.W. Aitken: The high voltage test programme at Daresbury for the NSF accelerator. In: *Proc. First Int. Conf. Tech. Electrostatic Accelerators, Daresbury* (Daresbury Laboratory, DNPL/NSF/R5, 1973) p. 147
29. J. Heugel: Nucl. Instr. Meth. A **287**, 109 (1990)
30. H.R.McK. Hyder: Terminal pumping using a turbomolecular pump as a recirculating gas compressor. Oxford University internal report, NPL 24/82, (1982)
31. D.C. Weissner, M.D. Malev: Nucl. Instr. Meth. A **287**, 64 (1990)

Box 4: Development of Tubes in Obninsk, Russia

V.A. Romanov

State Scientific Center of the Russian Federation, Institute for Physics and Power Engineering, 1 Bondarenko Sq., Obninsk, Kaluga Region, 249033 Russia
romanov@ippe.obninsk.ru

Work on the development and manufacture of accelerator tubes (ATs) for electrostatic accelerators has been performed for over 40 years in the Accelerator Department of the Institute for Physics and Power Engineering in Obninsk. During this time period, a large scope of work has been done, including (1) studies of some characteristics of discharge processes determining the level of the electric strength of some accelerating gaps and ATs; (2) development of methods for high-voltage tests of accelerating gaps and evaluation of ion-optic characteristics of ATs; (3) development of technology for gluing of ATs, and formulation of additional criteria that should be taken into account at the stage of manufacture of ATs designed for operational gradients over 1.2 MV/m; (4) studies of AT gaps with electrodes and insulators produced from different materials and having different shapes and vacuum surface areas; and (5) development of several designs of ATs with straight and inclined fields and their full-scale tests.

Factors to be Taken into Account when Developing Accelerator Tubes

J. McKibben [1] proposed criteria that should be taken into account in the development of long ATs. In addition to the criteria of McKibben, the following factors should also be taken into account for gradient values over 1.2 MV/m [2]:

1. In addition to the other measures aimed at an increase of the electric strength of accelerating gaps, it is expedient to reduce the active surface of the electrodes. (The electrode active surface can be determined as a part of the surface having a 20–30% higher electric field strength.) The use of corrugated electrodes is one possible approach. Treatment of *macrostressed electrode sections by dark currents* is more intensive in this case, i.e. organic impurities are removed more effectively from the work surfaces of electrodes. Besides, owing to the use of corrugated electrodes in vertical tubes, the emission of charged particles from dielectric crumbs resulting from destruction of insulating rings is significantly suppressed.

2. The so-called total-voltage effect observed in either accelerating gaps or long ATs is mainly caused by an intensification of discharge phenomena resulting from an *increase of the area* bombarded by high-energy positive and negative particles. In an ATs channel the electrodes can be considered as cathode and anode areas distributed along the tube. The role of the several-mm-wide *near-electrode discharge layer* is especially important in this case. The design of the electrode structure should, therefore reduce as much as possible the influence of this layer of dark current.
3. The maximum voltage withstood for a long time of conditioning (more than one hour) without breakdown in the gap with a dark current density within 10^{-11} A/cm² should be 2 to 3 times higher than the operational voltage of the gap.

Some Issues in Accelerator Tube Fabrication Method

A technology for glue bonding of the tube elements (i.e. electrodes and insulators) that ensured a fixed position of the metal sealing rings on the end faces of the insulators was developed. This method of electrode–insulator bonding prevents penetration of organic glue vapors into the AT volume and provides a high electric strength of the accelerating gaps with an insignificant spread of values and a rather low value of the dark current. The metal ring also provides a reliable contact between the electrode and the insulator and decreases the electric field strength in the vicinity of the triple junction (i.e. vacuum-electrode-insulator). Below, some examples of tube designs are described.

Accelerator Tube for the EG-1 Electrostatic Accelerator

This tube [3, 4], with stainless steel inclined electrodes, includes 166 units of $24 \times 180 \times 230$ mm³ insulators made of glass-ceramic and has a length of 4.125 m. In order to provide an invariable trajectory of the ion beam under conditions of change of the accelerating voltage, the tube is equipped with a section of preliminary acceleration having flat electrodes and compensative sections with inclined electrodes. An electrostatic steerer connected to the high-voltage divider of the tube is mounted at the tube exit. Another design feature is that three flat electrodes are provided at the end of the tube in order to shape and to preaccelerate an electron beam for the sake of high-voltage stabilization. This tube was manufactured in 1964 and, after several overhauls, is still in operation.

Accelerator Tube for the EG-2.5 Electrostatic Accelerator

In order to provide high spatial stability of the ion beam at the exit of the AT for the EG-2.5 electrostatic accelerator, a new AT with straight fields was

designed and manufactured. When designing the tube, the main attention was paid to the issue of the electric strength of the accelerating gaps and limitation of electron-ion exchange processes in the tube channel. Experience gained in operation of an accelerator with a large-aperture tube manufactured in 1981 was also taken into account. This AT consists of two sections, each 1090 mm long. One of these sections is shown in Fig. B4.1. The electrodes (part 1 in Fig. B4.1), made of CrNi38VTi stainless steel, have a complicated geometry. This results in firstly, a decreased active surface of the electrodes, and, secondly, protection of the tube insulators against peripheral particles of the beam. Owing to the relatively large diameter of the electrode aperture (85 mm), it is possible to maintain a vacuum in the tube at a level of 6×10^{-5} Pa. The insulating rings (part 2 in Fig. B4.1), having dimensions of $25 \times 180 \times 230 \text{ mm}^3$, are made of UF-46 ultraporcelain. The ring surface contacting the vacuum was machined in such a way that it was oriented at an acute angle with respect to the electrode planes. A vacuum-tight junction was made, in accordance with technology adopted, using a heat-resistant polymer glue with a low partial vapor pressure. The special feature of the tube is the use of grid-diaphragm units (parts 3 and 4 in Fig. B4.1) that suppress discharge processes in the accelerator channel. The design of these units makes it possible to remove gas by heating up to $400\text{--}600^\circ\text{C}$. Owing to the use of special reflectors, overheating of the glue joint is avoided. In order to increase the fraction of active gaps over the length of the AT, the sectioning pitch is half as much in the area of grid-diaphragm units.

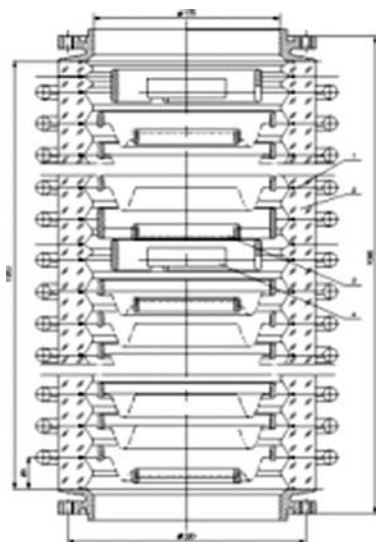


Fig. B4.1. A section of the accelerator tube of the EG-2.5 electrostatic accelerator. Numerals: 1, electrode; 2, insulating ring; 3, grid; 4, diaphragm

Accelerator Tube with Crossed Inclined Field

A tube design with crossed inclined fields was proposed in [5]. It includes several sections (Fig. B4.2, left-side part), each one consisting of axially symmetric insulating rings (1), connected by vacuum-tight sealing to the metal electrodes (2). The central sections of the electrodes, intended for formation of the accelerating field, are made as flat electrode inserts (3) with holes for the beam of charged particles. The inserts make an angle with the optical axis of the tube that is constant within one section. In between sections, a flat electrode (4) with a central hole is installed perpendicular to the axis, and each successive section (B) is turned in the same direction by an angle of 90° with respect to the previous section (A). Thus, when one goes from one section to another, a 90° turn of the lateral component of the accelerating field occurs, both the field strength and the field direction remain invariable within one section.

Calculated trajectories of the secondary electrons in both the crossed-inclined-field tube and an inclined-field tube having the same length of the inclined-field section and the same electrode aperture are given on the right side of Fig. B4.2. Comparison of the trajectories shows that the new crossed-inclined-field tube is capable of removing secondary electrons more effectively.

The author is greatly indebted to Dr. S. Bazhal for his continuous and fruitful efforts aimed at the development and manufacture of many of the accelerator tubes.

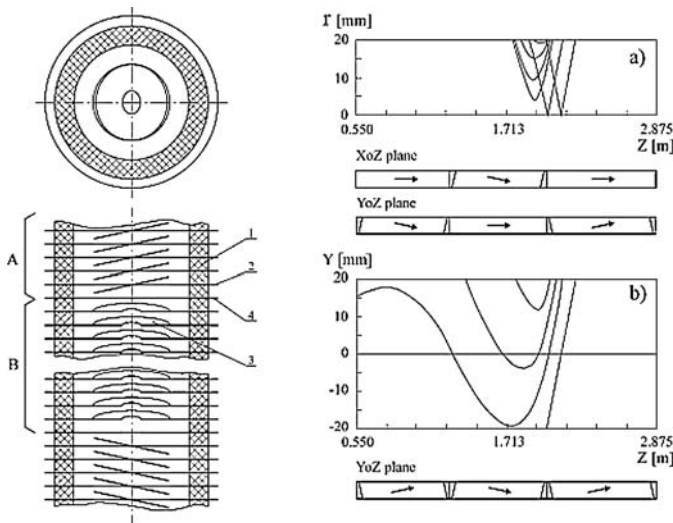


Fig. B4.2. Electrode structure of an accelerator tube with crossed inclined fields, and calculated trajectories of secondary electrons in the tube (a) with crossed inclined fields and (b) with an inclined field

References

1. J.L. McKibben: Reflection upon design criteria for a good accelerating tube. In: *Proceedings of the International Conference on the Technology of Electrostatic Accelerators, Daresbury, 1973*, ed. by T.W. Aitken and N.R.S. Tait (Daresbury 1973) pp. 295–307
2. V.A. Romanov: Features of discharge processes occurring in the accelerating tube channel and some measures for their suppression. In: *Proceedings of IX All-Union Meeting on Exchange of Experience Gained in Operation and Modification of Electrostatic Accelerators*, ed. by V. Romanov (IPPE, Obninsk 1991) pp. 62–74
3. V.A. Romanov, A.N. Serbinov: Instr. Exp. Tech. **6**, 38 (1965)
4. V.A. Romanov, A.N. Serbinov: Atomnaya Energiya **19**:2, 176 (1965)
5. S.V. Bazhal, V.A. Romanov: Accelerating tube, Russian Federation Patent No. 2089053 (1995)

9 Stabilization

L. Rohrer and H. Schnitter

Maier-Leibnitz-Laboratorium für Kern- und Teilchenphysik der
Ludwig-Maximilians-Universität München und der Technischen Universität
München Am Coulombwall 6, 85748 Garching, Germany
LA.Rohrer@t-online.de
heinz.schnitter@physik.uni-muenchen.de

9.1 Introduction

One of the outstanding features of electrostatic accelerators is the low energy spread and the high energy stability of the beam. The low energy spread results from the fact that all particles emerging from the terminal at the same time are accelerated by the same voltage to the same energy. The residual energy spread arises from processes in the ion source and, in the case of a tandem accelerator, at the stripper. The purpose of the stabilization system is to keep the energy constant. The control loop should be efficient enough that the remaining fluctuations do not further degrade the energy resolution given by the above-mentioned energy spread and by the resolution of the experimental setup.

9.2 Principles of High-Voltage Generation and Stabilization

First of all, one has to maintain the accelerating voltage constant to achieve high beam energy stability. Fortunately, the properties of electrostatic machines facilitate a terminal voltage stabilization of the order of 10^{-5} to 10^{-4} . Moreover, one can modulate the voltage of energy-defining parts of the machine to achieve the highest stability.

9.2.1 Terminal Voltage Resulting from Currents and Impedance

In a machine without any stabilization system, the terminal voltage V_t results from charging and discharge currents acting on the terminal impedance according to (9.1), as shown in Fig. 9.1:

$$\frac{V_t}{R} + C_t \frac{dV_t}{dt} = I_{ch} - (I_{p+} - I_{p-} + I_{lost}) \quad (9.1)$$

Considering the Munich MP tandem as an example, the parameters are approximately $I_{ch} = 210 \mu\text{A}$ ($\leq 600 \mu\text{A}$), $I_{p+} + I_{p-} \leq 10 \mu\text{A}$, $I_{lost} \leq 10 \mu\text{A}$, $R = 60 \text{ G}\Omega$, $C_t = 500 \text{ pF}$, and $V_t = 12.5 \text{ MV}$ ($\leq 15 \text{ MV}$).

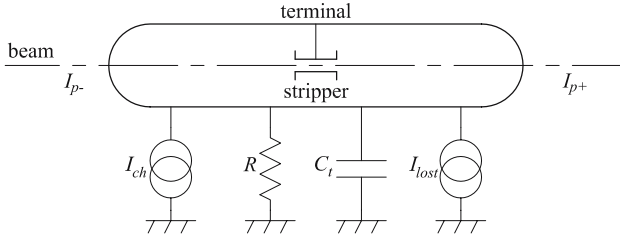


Fig. 9.1. Equivalent-circuit diagram of a tandem accelerator. I_{ch} is supplied by the charging system (the total of the up charge and, if existing, the down charge). I_{p-} is the beam current flowing from the ion source to the terminal; I_{p-} is negative. I_{p+} is the beam current flowing from the terminal to the high-energy end of the tandem. I_{lost} is the so-called lost current, mainly caused by tank gas ionization. R is the resistance of the voltage divider chains, column, and tube, if applicable. C_t is the terminal–tank capacitance

The sources of terminal voltage fluctuations are mainly the variation in the charging current, caused by inhomogeneity of the belt or the charging chains, but also fluctuations of the discharge current in the tank gas due to ionization by gamma radiation. Provided that the machine is in good shape, typical values of the voltage instability are of the order of 2 to 20 kV peak to peak. For short fluctuations, the voltage stability profits from the big time constant RC (several seconds). The terminal voltage cannot change rapidly if there are not catastrophic currents involved, as in the case of a spark. Long-term fluctuations, however, must be compensated for.

9.2.2 Basic Principle of Stabilization

Figure 9.2 shows the typical design of a stabilization system of a tandem accelerator. The main parts are a beam-energy-analyzing system, a generating voltmeter, a control amplifier, and a corona points assembly.

One cannot measure the terminal voltage with adequate accuracy by means of standard voltmeters. Therefore the beam energy is measured with a system consisting of a deflection magnet, a pair of slits, and a differential slit current amplifier. If the beam energy differs from the nominal value given by the field of the 90° magnet, the deflection angle of the magnet varies. This causes different beam currents at the slits and an error signal at the slit amplifier. Another method to measure the terminal voltage is electric-field measurement with a generating voltmeter (GVM) at the tank wall. This device works with a rotating electrode to chop the field and another electrode to pick up the current pulses, which then are rectified and amplified. The resulting signal is less accurate than the beam energy analysis described above, but in the absence of a measurable beam current it is the best choice. The best is to switch between the two modes automatically. If a beam current is present and the terminal voltage does not differ more than a given amount

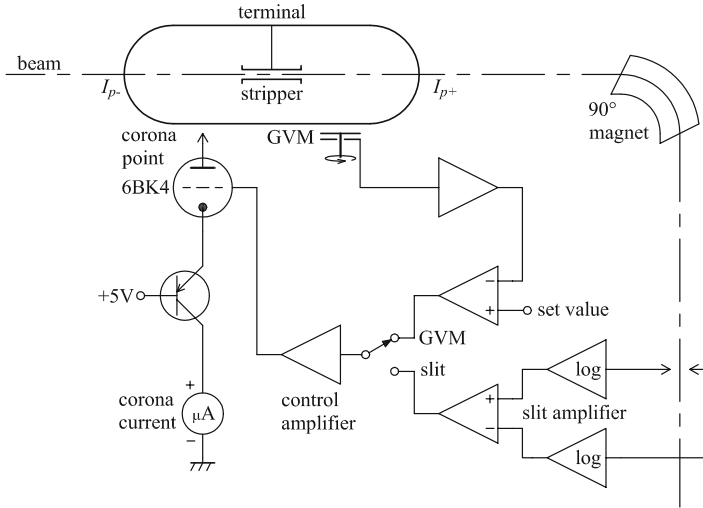


Fig. 9.2. Standard stabilization system of a tandem accelerator

(e.g. 10 kV) from the set value, slit control is used. Otherwise, if either no beam is present or the terminal voltage is too far off, so that one must suspect an unwanted beam, GVM control is applied.

The corona points produce a controlled corona discharge to compensate the fluctuations of the charging and discharging currents to and from the terminal. The assembly consists of some needles standing out of a grounded receptacle. This configuration acts as a triode, so that a swing of few kV applied to the needles is sufficient to control the corona current in the range of zero to $100\mu\text{A}$. The control voltage is supplied from an amplifier with a high-voltage vacuum tube as the output stage.

The energy-analyzing system, the amplifier, the corona points, and the terminal with its capacitance act as a feedback control system with nearly perfect integral characteristics: an energy deviation causes a proportional change of the corona current and, owing to the terminal capacitance, a corresponding ramp of the terminal voltage. Unfortunately, this ramp is delayed as a result of the slow building up of the corona discharge in the tank gas; this acts like a delay line with a propagation time of the order of 30 ms (in big machines) [1]. However, a feedback control system becomes unstable and starts to oscillate if the overall phase shift is 360° and the open-loop gain is greater than 1 at a given frequency. Owing to the intentionally negative feedback there is a phase shift of 180° , the integrator contributes 90° , and another 90° is added by the corona delay at a frequency of

$$f = \frac{1}{4T_d} \quad (9.2)$$

If the delay time T_d is 30 ms, the oscillation frequency will be 8.33 Hz.

To analyze the electrical behavior of the control loop, one can measure the response of the terminal voltage as a function of a signal applied to the grid of the corona tube. A proven method to get the frequency response is to apply sine wave voltages with variable frequencies to the grid of the corona tube and measure the amplitude and phase of the AC component of the terminal voltage using a well-calibrated capacitive pickup (described in Sect. 9.3.3). It is important that the amplitude is small enough to prevent saturation effects. On evaluating the data, one will find the corona delay time T_d and the terminal time constant $T_{RC} = R_t C_t$ (some hundreds of ms). R_t comprises the column resistance in parallel with the corona impedance. Since all other time constants in the system are comparatively small, the open-loop gain is given by

$$A_l = -A_0 \frac{e^{-pT_d}}{1 + pT_{RC}} \quad (9.3)$$

A_0 , the DC gain of the system, is set to a maximum, so that oscillation does not arise yet. The condition for that is

$$A_0 < \left| 1 + i \frac{\pi T_{RC}}{2T_d} \right| \quad (9.4)$$

The reduction of the fluctuations by the control system is described by the control factor

$$F_c = 1 + \frac{A_0 e^{-pT_d}}{1 + pT_{RC}} \quad (9.5)$$

One can compensate the delay electronically only in a small frequency range, and there is nearly no benefit from that. Therefore, if the control factor of the corona control is too low to obtain sufficient stability, one must add a faster controlling element, e.g. stripper modulation or controlled down charge, as treated in Sects. 9.4.2 and 9.4.3.

9.3 Measuring System

One can determine the terminal voltage from column current measurement or from electric-field measurement by means of a generating voltmeter. Alternatively, one can use a beam-energy-analyzing system consisting of a deflection magnet and a beam-position-measuring device.

Because of the voltage and temperature dependence of the column resistors, the measurement of the column current is the least accurate approach. Moreover, it represents only the voltage of the outer column section. The generating voltmeter is much more accurate. If it is properly positioned, its output indicates the terminal voltage without significant interference from other parts of the machine. However, provided that a macroscopic ion beam exists, the best method is beam energy analysis, as described below. It is

sensitive to the particle energy, which is the value to be stabilized. Energy deviations are indicated and corrected whatever their source may be. Examples are an instability of the injection energy, and the increasing energy loss in a stripper foil if the foil thickens during use. Energy analysis is more accurate than the other techniques and its response time is superior.

9.3.1 Energy-Analyzing System

The ion-optical components of the analyzing system are a quadrupole lens, the object slits, the magnet, and the image slits. The beam at the high-energy end of the accelerator is axially symmetric. The quadrupole doublet focuses it to an upright ellipse at the object slits. The small vertical divergence allows one to make the gap of the magnet small, and so to save energy. The magnet is usually a double-focusing 90° magnet with 26.5° shim angles, the object and image distances being twice the bending radius. The edges of the pole pieces are beveled in an approximate Rogowsky profile to avoid saturation effects, which would result in a field-dependent beam path. The field must be highly stable, since it is the measure of the energy. A closed-loop field control system provided with an NMR teslameter is capable of keeping it constant within 10^{-5} . The energy dispersion of the magnet combined with the drift space to the image waist is

$$\frac{\Delta x}{\Delta E/E} = 0.5 \left\{ R(1 - \cos \alpha) + D \left[\sin \beta_2 + \frac{\sin(\alpha - \beta_2)}{\cos \beta_2} \right] \right\} \quad (9.6)$$

where R is the bending radius, α the deflection angle, β_2 the exit shim angle, and D the image distance. For a double-focusing 90° magnet, (9.6) reduces to

$$\frac{\Delta x}{\Delta E/E} = 2R \quad (9.7)$$

An energy deviation $\Delta E/E$ of only 10^{-5} typically produces a beam offset Δx of the order of 5% of the horizontal spot size at the image slits.

The slits collect a small fraction of the beam current only, from the rim of the beam. To obtain an output signal proportional to the beam offset, independent of the beam intensity, one has to use logarithmic current-sensitive amplifiers and subtract their output voltages. The current range of the amplifiers must be very wide, and the frequency response must be fast compared with the other components of the control loop. A logarithmic converter, made as usual with a transistor in the feedback path of an operational amplifier, is rather slow at a low input current. This can be overcome by piecewise logarithmizing and subsequently adding the pieces [2,3] or by digital logarithmic conversion using a signal processor. In any case, the operational amplifier of the input stage must be excellent with respect to bias current, input offset voltage, and drift.

Assuming a normal distribution of the current density (approximately exponential in the tails), the response of the logarithmic amplifiers to the beam offset is linear to a first approximation. This is an important property of the control loop, which applies only if no secondary electrons from one slit electrode hit the other one. The electrons can be suppressed by an electric field, generated by a bias voltage from a battery. Also, the shape of the electrodes must be adequate: sharp edges are better than cylindrical electrodes. Staggered slits with the two electrodes at different positions in the beam direction can also help to avoid secondary electrons from one slit hitting the other one.

There is one more source of error in the slit signal. In some cases, for example if negative molecular ions are injected into a tandem, a mixture of various particles with different charge states is accelerated in the high-energy section. If the magnetic rigidity of unwanted particles is very close to that of the wanted particles, they can hit one of the slit electrodes and so affect the control signal. A remedy can be so-called shadow slits, a slit pair installed upstream.

Although the accuracy and the reproducibility of the analyzing system described above are excellent, the absolute value of the beam energy can be calculated only roughly from the magnet data, because the magnetic field along the beam path, including the fringing fields of the magnet, is not exactly known. A standard method for calibration uses reactions with well-known resonances, such as the excitation functions for elastic and inelastic scattering of $^{12}\text{C}(\text{p}, \text{p})^{12}\text{C}$ and $^{12}\text{C}(\text{p}, \text{p}')^{12}\text{C}$ at the resonance at 14.233 MeV [4]. If a spectrograph is available, one can use it to compare the magnetic rigidity of ions from the accelerator with that of 8.784 MeV α -particles from the decay of ^{212}Po [5]. Also, a direct measurement of the time of flight of a chopped beam over a carefully measured distance has been performed [6].

9.3.2 Generating Voltmeter

For some experiments, for example in accelerator mass spectroscopy, the terminal voltage must be stabilized even in the absence of a measurable beam. In this case a generating voltmeter (GVM) is the most accurate tool for voltage measurement and stabilization.

The generating voltmeter, called there the *Feldmühle*, was first described by Lueder and Schwenkhagen [7] and used to measure the electric field in the atmosphere at the earth's surface. In the accelerator, it is destined to measure the electric field caused by the terminal voltage. Usually the GVM is mounted very close to the tank wall in a position opposite to the terminal, where the field is determined only by the terminal voltage, and as far as possible from the corona points. The generating voltmeter (Fig. 9.3) consists of a grounded rotating disk with apertures to chop the electric field, and isolated static electrodes delivering alternating current pulses proportional to the field. The rotating plate covers and uncovers the signal plates alternately so that, owing

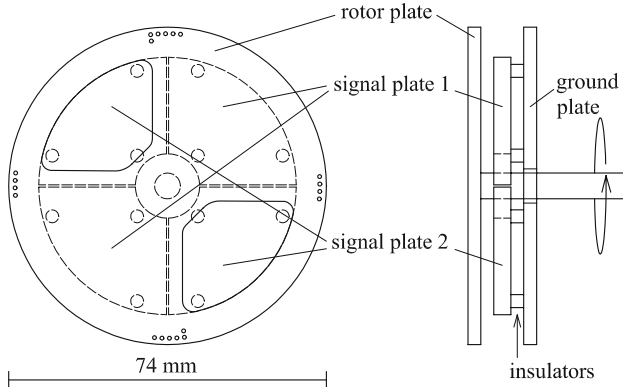


Fig. 9.3. Generating voltmeter (GVM)

to the alternating field, a current is induced in the signal plates. This current I follows the equation

$$I = V_t \frac{dC}{dt} \quad (9.8)$$

where the terminal voltage is V_t and the time-dependent capacitance between the signal plate and the terminal is C (C_1 or C_2 in Fig. 9.4). If the opening in the rotating plate has exactly the same size and form as the static signal plates, the waveform of the capacitance looks like a triangle (to a first approximation). Therefore the alternating current has a rectangular pulse form with a rather high slew rate. This causes some problems in the signal processing.

An essential improvement in the design is to use two sets of signal plates and to make the windows in the rotating plate smaller than the area of the signal plates. In this case the rotating plate covers more than the area of one signal plate, resulting in capacitances C_1 and C_2 as shown in Fig. 9.4. Therefore the readout signal (I_1 and I_2 in Fig. 9.4) is shaped in such a way that its derivative is zero at the zero-crossing (every 90°). So the readout electronics can be slow, and there is enough time to control the electronic switches.

Position holes in the rotating plate are read out by two photologic sensors to control the analog circuit shown in Fig. 9.5 via a state machine. Two index holes are used to start the phase-controlled rectifier for each 180° again. This allows one to determine the polarity of the electric field. In the first phase, from 0° to 90° in Fig. 9.4, I_1 is integrated in A_1 , and I_2 is integrated in A_2 . The results are sampled in the sample-and-hold gates A_3 and A_4 . In the next phase, I_1 is integrated in A_2 and I_2 in A_1 . This procedure is continuously repeated. The timescale for 90° is 20 ms, corresponding to the 50 Hz mains frequency, to suppress interference from the line voltage. The difference between the samples, obtained by the subtracting amplifier A_5 ,

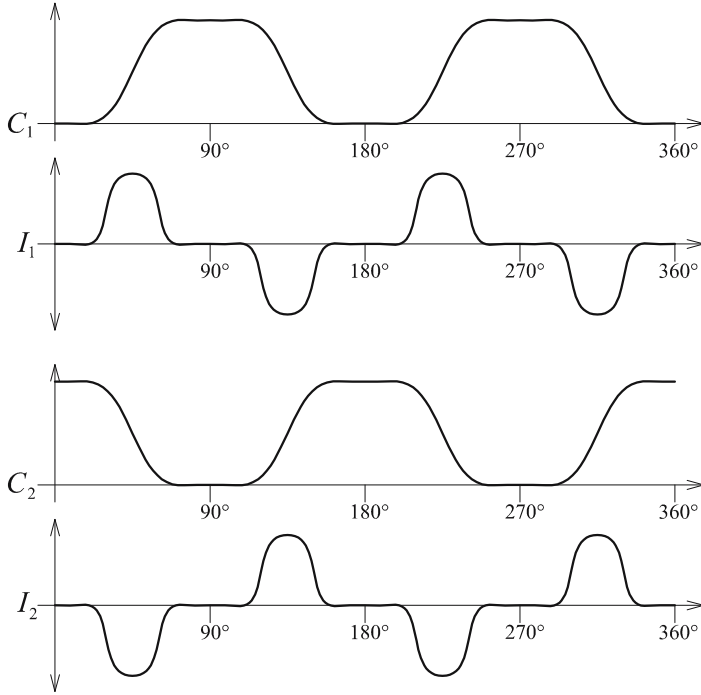


Fig. 9.4. Capacitances of the signal plates to the terminal, and currents, as a function of time

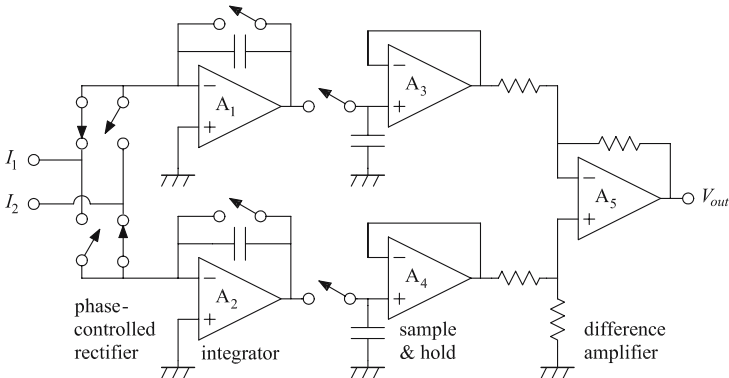


Fig. 9.5. Circuit diagram of the generating-voltmeter amplifier

represents the actual value of the terminal voltage, and common-mode noise is suppressed.

A few details are important for the accuracy of the GVM. The motor axle has to be grounded via a well-conducting collector. An approved combination is silver-graphite brushes on a highly polished stainless steel collector ring.

The insulators of the signal plates must be hidden from the surface, so that charge possibly sitting on the insulators cannot affect the signal (no printed circuit boards as signal plates!). The rotor and the stator plates should be polished, and have to be kept clean. Otherwise, an isolating film may cover the electrodes, where parasitic charge deposit may occur. The accuracy of the measurement is better than 10^{-4} .

9.3.3 Capacitive Pickup

The ripple of the terminal voltage can be measured with a capacitive pickup (CPU). This is an electrode at the tank wall connected to a current integrator. Since the electrode current is proportional to the derivative of the terminal voltage, the output of the integrator represents the AC component of the terminal voltage. An RC high-pass filter at the input of the integrator is necessary to avoid saturation of the operational amplifier by the ionization current flowing to the pickup electrode.

The CPU signal is usually displayed on an oscilloscope, acting as a monitor for the ripple. Since it is contaminated by ionization current fluctuations, it is not advisable to feed it into the control loop. This may deteriorate the voltage stability instead of improving it. An exceptional case might be the generating-voltmeter control mode if the GVM response is very slow.

9.4 Final Control Elements

The first control element for a stable terminal voltage is the charging system. Since the response of the terminal voltage to up-charge current changes is very slow, one cannot include it directly into the control system. But a uniform up charge is most important for terminal voltage stability. The most commonly used control element is a controlled corona discharge, which acts much faster. The ion transit time from the corona dominates the closed-loop behavior of most stabilizing systems. Even faster final control elements, e.g. controlled down charge [8] or stripper modulation [9], can further improve it.

9.4.1 Controlled Corona Discharge

As described in Sect. 9.2.1, in an electrostatic accelerator without any stabilizing system the terminal voltage results from an equilibrium of the charging and discharging currents in the terminal impedance. A controlled corona discharge system acts as an additional shunt load. Considering the Munich MP tandem as an example, the operating current of the shunt circuit is $50\mu\text{A}$, and therefore it has a range of zero to $100\mu\text{A}$. As shown in Fig. 9.6, the corona assembly consists of an insulated set of 12 needles standing out of a grounded mushroom. The needles are directly connected to the anode of

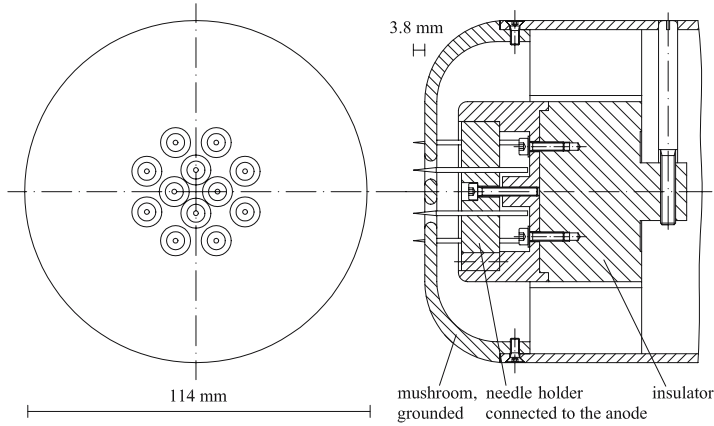


Fig. 9.6. Design of a corona needle assembly

a high-voltage vacuum tube. The grid of the high-voltage vacuum tube is driven by the output voltage of the control amplifier.

A linear positioning drive controls the distance of the corona assembly from the terminal with respect to the operating point of the stabilizing system. The higher the terminal voltage, the higher must be the distance from the terminal so that the field strength at the needles is always the same. Then all other parameters of the stabilizing system remain constant.

9.4.2 Controlled Down Charge

As mentioned above, the response of the terminal voltage to changes in the up-charge current is very slow. The calculation of the transfer function is difficult owing to the complex geometry of the accelerator structure. But the simple equivalent-circuit diagram in Fig. 9.7 shows the basics. Let us assume that a concentrated positive charge Q (on a small piece of the belt or a pellet of the chain) is moved from ground potential to the terminal. C_q is the capacitance of the charge carrier to ground, C_t the terminal capacitance to ground, and C_{qt} the capacitance of the charge carrier to the terminal. The terminal-voltage change is given by

$$\Delta V_t = Q \frac{C_{qt}}{C_q C_{qt} + C_q C_t + C_{qt} C_t} \quad (9.9)$$

When the charge carrier is on its way from ground to the terminal, C_{qt} is very small compared with C_t , and there is nearly no terminal-voltage change for many hundreds of ms. Only if the charge is very close to the terminal does C_{qt} reach the order of magnitude of C_t , and a voltage step occurs. The long delay is not tolerable in a closed loop. But if we reverse the motion and the polarity, and let the charge carrier travel from the terminal to ground,

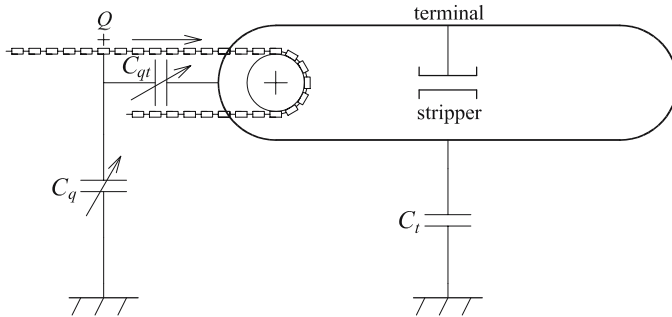


Fig. 9.7. Equivalent circuit for the charge transport

the charge leaving the terminal induces an immediate change of the voltage. Hence, a controlled down charge can be a faster control element than a corona discharge.

The actual charging device in the terminal is similar to the up-charge device at ground potential. In the case of a belt charging system, it is a screen spraying charge on the belt. The charge is supplied by a controlled current source. If the charging system is a chain or a ladder, the existing inductors in the terminal, which normally are connected to pickup wheels, must be supplied from controlled voltage sources, a positive and a negative one. The position of the charging electrodes must be as close as possible to the end of the terminal, so that the travel time of the charge inside the terminal is as short as possible. But even if the assembly is designed very well, a delay of 10 ms or more is inevitable, and this limits the control factor of the loop. A data link is necessary to transmit the control signal for the down-charge power supply from ground potential to the terminal. The transmission delay of this link should be small compared with the delay mentioned above to avoid further performance degradation.

9.4.3 Modulation of the Effective Accelerating Voltage

Since the signal delay characterizing the corona transfer function and, to a minor degree, also the response of down-charge systems, is the worst property of the control loop, a more direct control element for the effective accelerating voltage is desirable. One would like to apply directly a corrective voltage, derived from the error signal, to an energy-affecting element. The most suitable device for this purpose is a bipolar controlled current source, capable of a few kV output voltage. The voltage is determined by the current, integrated by the load's capacitance with an optional capacitor connected in parallel. In this way, a stable integral action of the control loop is achieved. It should be noted that a closed loop is possible only in the slit control mode. The GVM signal cannot be used, because it does not indicate the effect of a directly applied corrective voltage.

Points to be considered for the application of the corrective voltage are the ion injector, the high-voltage terminal or the stripper, and also the target of the beam. The latter can be neglected for two reasons. First, it is difficult to insulate the target with the complete experimental setup and to attach high voltage to it. Secondly, the effective energy cannot be measured, and no signal is available to be applied to the control loop.

To modulate the energy of the injected beam, one must apply the corrective voltage to the injector, that is, the ion source with all ion-optic elements and the associated power supplies. If the injector is already insulated and connected to a preacceleration voltage, the corrective source, controlled through a fast data link, can be inserted between the preacceleration power supply and the ion injector. A disadvantage of this method is interference to the ion optics at the low-energy end of the accelerator, due to the fluctuating injection energy. Particularly, if a buncher for a nanosecond pulsing system is installed between the injector and the accelerator, this technique is not applicable.

Modulating the terminal stripper of a tandem accelerator is much better in this respect. At the terminal, the particle energy is so high that a few kV energy modulation does not influence the optics of the high-energy tube. A further advantage of this method is that the corrective voltage acts in a way that is amplified according to the charge state of the ions. The technique is as described above: a bipolar current source controlled from the error signal via a fast data link is connected to the stripper. Only a few practical aspects must be considered. The stripper foils must be protected with a shield connected to the stripper to prevent rupture by electrostatic forces. If a gas stripper is used, one must bear in mind that the gas pressure in the pipe between the needle valve and the stripper tube is in a region where gas discharges occur even at moderate voltages. Therefore the needle valve must be connected to the stripper, and the pipe must be metallic.

Another successful method to apply the corrective voltage to the terminal has been realized at the Munich MP tandem [2]. An insulated electrode, the so-called liner, covering a part of the inner tank wall is connected to a bipolar current source controlled by the error signal. The voltage at the liner results from the current integrated by the liner capacitance. A fraction of this voltage, determined by the capacitances of the terminal to the liner and to the grounded tank wall, is induced at the terminal. Since this fraction is only about 20%, the source must be capable of ± 20 kV, so that the corrective voltage at the terminal is ± 4 kV. The most serious problem of this technique is damage by tank sparks. Since the liner is directly exposed to sparks the destructive energy is high, and it is difficult to protect the liner, the tank feedthrough, and the electronics adequately.

9.5 Functional Specification of an Ideal Stabilizing System

For many accelerator installations, a standard stabilizing system as described in Sect. 9.2.2 is sufficient. If the charging system is in good shape, the voltage fluctuations of the free-running machine are only a few kV, and this can be reduced to below 1 kV FWHM using the standard technique.

Bare corona stabilization may be insufficient if either precision experiments require an extra-stable beam energy or the stability of the free-running machine is poor and cannot be further improved. In this case stripper modulation is the best choice. One can implement it even in an existing machine, provided that the stripper can be insulated. However, the voltage span of the corrective power supply does not cover the whole control range. Therefore a corona loop or a down-charge loop is still necessary to keep the stripper modulation voltage around zero. In most cases the operator controls the up-charge current by hand. This is due to the slow response of the charging system. But, under computer control, one could mix a suitable control signal using the terminal-voltage set value and the actual value of the corona current (or the down-charge current if this is part of a control loop). In any case, it is important to avoid saturation of amplifiers and power supplies in all loops, otherwise the system tends to oscillate, even if the overall gain is set to an optimum for small signals.

Although one can reduce the terminal-voltage fluctuations to about 200 V FWHM using the techniques described above, the actual accuracy of the particle energy at the target does not correspond to this value. There are some more energy-deteriorating effects. The thermal motion of the particles in the ion source contributes 20 to 200 eV FWHM. The stripper-induced energy spread, including straggling and stripper inhomogeneity, is between 100 eV for light ions in a gas stripper and about 40 keV FWHM for heavy ions in a foil stripper. Injecting negative molecular ions and breaking the chemical bond in the stripper foil yields even higher energy straggling. Owing to the target thickness, one has to take into account similar values to those for the foil stripper. Hence, sophisticated stabilization techniques are most useful for precision experiments with light-ion beams. In the case of heavy ions, effects besides the terminal-voltage stability dominate the overall characteristics.

For the routine operation of an accelerator, the aspects of usability are very important. The operator must be able to set up the machine quickly for a new beam. The best tool for that is a computer control system, with data tables and algorithms included to determine the operating conditions and to set the values with adequate timing. So, the control system can adjust the position of the corona points, gradually set the charging current with the actual terminal voltage in check, and then hand over the control to the GVM stabilizer. With the stabilizing system in operation and the terminal voltage close to its set value, the charging current must be readjusted so that the corona current is around the nominal operating point (e.g. 50 μ A). This is

necessary also if the beam is injected, and can be done by the operator or by a control computer as well. Also, the position of the corona points has to be readjusted so that the control amplifier output is around zero.

Simultaneously, the analyzing-magnet current is set, and the NMR is switched into search mode. If the NMR signal is present, the system is switched into field-stabilizing mode. All ion-optic elements are to be preset, too.

The following steps may be taken with or without a computer algorithm. The beam is injected and optimized up to the object slits and then forwarded to the image slits. Possibly one will have to slightly modify the terminal voltage to find it there. Now, the automatic switching mode of the stabilizer described in Sect. 9.2.2 can be activated. In this mode, the terminal voltage is controlled by the slit signal if a slit current is available and the voltage is not too far from the set value; otherwise, the GVM signal is used. So the machine remains close to its operating parameters even if the beam drops for a moment, and returns to normal operation when the beam is back. Catching a different charge state or particle is prevented by the condition for slit stabilization that the voltage is close to the set value. The last step in setting up the stabilizer is to optimize the loop gain. Starting from a fairly low value, the operator increases the gain until the system oscillates. This can be easily observed at the oscilloscope showing the capacitive-pickup signal. If no oscillation occurs, even with the gain set to its maximum, the system is not in good shape, and one should try to find the reason for that (e.g. poor focusing of the beam at the slits). Otherwise, the gain should be reduced until the oscillation stops. So the system is adjusted to yield optimal beam energy stability.

References

1. E.A. Gere et al.: IEEE Trans. Nucl. Sci. **NS-14** no. 3, 161 (1967)
2. W. Assmann et al.: Nucl. Instr. Meth. **122**, 191 (1974)
3. Texas Instruments: Data sheet TL441
4. K. Sasa et al.: University of Tsukuba Tandem Accelerator Center, Annual Report (2001) p. 5
5. T. Feastermann et al.: Annual Report of the MLL, Munich (2001) p. 4
6. E. Huenges et al.: Phys. Lett. B **45**, 361 (1973)
7. W. Georgii: Feldmühle. In *Lexikon der Physik*, vol. 1, 2nd edn. by Hermann Franke (Franckh'sche Verlagshandlung, Stuttgart 1969) p. 473
8. T.W. Aitken: Nucl. Instr. Meth. **129**, 341 (1975)
9. T.A. Trainor: Proceedings of the Third International Conference on Electrostatic Accelerator Technology (Oak Ridge 1981) pp. 143–147

10 Stripper Systems

D. Weisser

Research School of Physical Sciences and Engineering, Australian National University, Canberra, Australia
`david.weisser@anu.edu.au`

10.1 Introduction

The stripping process, that is, changing the injected negative ions into positive ions, is the defining feature of acceleration in a tandem accelerator. However, it also results in loss of beam transmission efficiency and voltage-holding ability. Charge state fractionation, described in Box 5, and scattering reduce the beam intensity. How the latter also compromises the accelerator's reliable maximum voltage is explained in Sect. 10.2. Since the vacuum in the accelerator tube also affects the transmission, methods to achieve low pressure have been pursued over the years, leading to improvement of pumping systems, which is discussed in Sect. 10.3.

There are two choices of stripping media, gas and carbon foils. If stable beam transmission is the priority, then gas stripping is chosen. Gas stripper apparatus is described in Sect. 10.4. Gas stripping is also preferred for injected molecular beams and to enhance the population of low charge states. If, however, beam energy is the priority, then foil stripping is preferred at the expense of transmission efficiency and unvarying intensity. Box 6 surveys the manufacture of carbon foils, and Sect. 10.5 deals with foil stripping and failure modes. The short foil lifetimes for heavy ions limit the application of foil stripping to low- and medium-mass ions and to low-intensity applications.

In the pursuit of the highest energy from a given accelerator, further foil stripping within the accelerator, described in Sect. 10.6, provides additional substantial energy gains but with a sacrifice of intensity and machine stability. In Sect. 10.7, we canvass the need to select charge states in the terminal and how selection impacts on double-stripping operation. Methods are described to prevent undesired charge states from compromising the performance of the accelerator.

The last section deals with the use of strippers external to the tandem in laboratories with postaccelerators. The energy gain from the postaccelerator will be maximized if an external carbon stripper is used to produce higher charge states than those achievable even with a second stripper in the tandem.

The strength of heavy-ion tandem accelerator facilities is their flexibility. Being able to choose either gas or foil terminal stripping enhances flexibility, as does the option of a second stripper in the high-energy tube. These

stripping choices have demonstrated an extremely useful range of beam energies and intensities that have underpinned heavy-ion nuclear-physics research.

10.2 Interactions of the Beam with Stripper Material and Gas

By definition, tandem acceleration requires the stripping in the terminal of electrons from the injected beam. The choice of gas or a foil stripper has vacuum, beam transmission and voltage-holding consequences.

10.2.1 Interactions Between the Beam and Residual Gas

Too much gas in the accelerator tubes, from whatever source, causes trouble even for light beams. The slow-moving injected beam is readily disrupted by interaction with the gas through scattering and the production of neutrals, positive ions and electrons. This also affects positive ions in the high-energy tube, though to a lesser degree because they are moving faster. The electrons, neutrals, and positive and negative ions strike accelerator tube electrodes, producing more positive and negative ions, neutrals, electrons and even puffs of gas. These processes, at their most benign, cause dips in the voltage at the affected electrodes. If the voltage quickly recovers and the machine does not spark, this sequence is called conditioning, because the surfaces involved appear to be “cleaned”, allowing further increases in voltage. If the ion and electron currents are too large, the gradient will be upset enough to cause the machine to spark.

But it is the electrons, produced by all these processes that are the immediate enemy. A runaway situation rapidly develops if the electrons are not steered into a tube electrode before they gain enough energy to create, upon collision, more than one electron and/or ion. The X-rays from these electron strikes ionize the insulating gas outside the tube, allowing charge to be drawn from the rings and the terminal to the tank wall and so dragging down the voltage by many MV. This process is *electron loading*.

Although too much gas in the tubes causes voltage-holding problems, small amounts, up to a pressure of 10^{-4} Pa, has little voltage impact and some benefit. This by-product of gas stripper operation helps quench secondary-ion currents and consequently the electron loading. This was a necessary ingredient for the successful operation of the straight-field accelerator tubes that preceded ones with inclined fields (Chap. 8).

10.2.2 Intensity Losses and Energy Sharing

Intensity losses in stripping are caused by three main mechanisms: charge state fractionation; multiple scattering, spreading the beam entering the

high-energy tube; and the Coulomb explosion of molecular ions, which can drastically increase the angular divergence of the beam as well as increase its energy spread.

10.2.3 Multiple Scattering

The stripping process intrinsically involves multiple small-angle scatterings, which spread the beam, making its angular emittance larger than the angular acceptance of the high-energy accelerator tube [1]. For gas strippers, because they can be adjusted to have fewer atoms per unit area than foils, the increase in divergence is quite manageable. A lens between the stripper foil and the entrance to the high-energy tube mitigates this problem by capturing scattered beam and focusing it through the high-energy tube (Sect. 10.7).

10.2.4 Molecular Beams

Cesium bombardment sputter sources produce negligible atomic negative-ion beams of elements with low electron affinity such as beryllium, nitrogen, calcium and magnesium. But they do provide useful intensities of their hydrides, oxides and carbides [2]. Although the molecules are readily dissociated in the stripper, this process and associated electron stripping inevitably result in losses of intensity and spreading of the beam energy, as described below.

The energy gained by a molecule in being accelerated to the terminal is shared among the constituent atoms. For the CN^- molecule, for instance, the nitrogen leaving the stripper has 14/26 of the incident energy, thus reducing the final available energy of the nitrogen beam. The lower N energy at the terminal also results in less population in the higher charge states, further reducing the beam energy available. That is why experimenters wanting the highest energy will opt to inject NH^- instead, where the N retains 14/15 of the incident energy.

The sharing of the low-energy acceleration is not the only debilitating factor. Worse still is the divergence increase caused by Coulomb explosion [3]. As a molecule traverses a stripper foil, several electrons are almost instantaneously removed, leaving the suddenly positively charged constituent atoms dissociated but still very close together. Their mutual Coulomb repulsion produces a substantial angle between them as they leave the foil on their way to the high-energy tube. Conservation of momentum says that the heavier constituent remains closer to the axis than the lighter one. But even the N on the axis have debilities, since those emitted in the direction of the beam gain energy from the Coulomb field, while those emitted opposite to the beam direction will have less energy. The energy change is less for the heavier constituent. Thus the Coulomb explosion effect for N in CN^- is much more severe than for N in NH^- in terms of both angular and energy spreads. There is a preference, therefore, to choose molecules in which the atom of interest is bound to the lightest possible partners.

Using a gas stripper to dissociate the molecule can minimize the effect of Coulomb explosion. As a molecular ion travels through a gas, the electrons binding it are removed, allowing the first neutral atoms and then low-charge-state positive ions to move well apart before their charge states, and therefore the Coulomb field, become high. For those beams that can only be adequately produced as molecules, using gas first to dissociate the molecule and then foil to produce a higher charge state is the preferred technique.

10.3 The Vacuum in the Tubes

The transmission efficiency of the accelerator is reduced if the vacuum in the accelerator tubes is poor. The quality of the vacuum depends upon the outgassing properties of the tubes, the speed and location of the pumps, and any gas introduced by the stripping equipment.

10.3.1 Vacuum Pumps

As mentioned in Sect. 10.2, a small amount of stripper gas was useful for old-fashioned straight-field tubes to reduce electron loading. However, even a small amount of gas has negative effects on beam transmission if not on voltage-holding ability. The advent of inclined-field tubes and magnetically suppressed ones, which sharply reduce electron loading, obviated the need to quench ions with stripper gas in the tubes. Therefore, the opportunity arose to reduce the residual negative effects of gas in the tubes by improving the vacuum. The main pumps at the ends of the machines progressed from mercury diffusion pumps with cold traps to low-backstreaming oil diffusion pumps, to ion pumps, to turbopumps and to cryopumps as each type became available. The vacuum, measured just outside the accelerator over the pump, ranged from $\sim 10^{-4}$ Pa, in the early days, to $\sim 10^{-6}$ Pa, as pumps got bigger and better.

In longer machines, the low pumping conductance of the accelerator tubes themselves, ~ 25 l/s for a 1.8 m long tube, vitiates the improvement in base pressure at the pumps. To place pumps in machines that were not designed with space and electric power for them proved a continuing challenge. At first, pumps were installed in the terminal to capture the stripper gas. This will be described in Sect. 10.4. Although having pumps capable of hundreds of liters per second at the ends of the machine and in the terminal helped, it was necessary to reduce the pressure in the tubes themselves, especially in longer machines. Additional ion pumps were installed in the dead sections where accelerator tubes joined, and rotating shafts were grafted into the columns to provide power. Figure 10.1 shows the ion pumps squeezed into the dead sections in the MP accelerator at Brookhaven National Laboratory [4]. A combined titanium sublimation and ion pump was added between units 6 and 7 in the Australian National University 14UD [5] and at four locations in

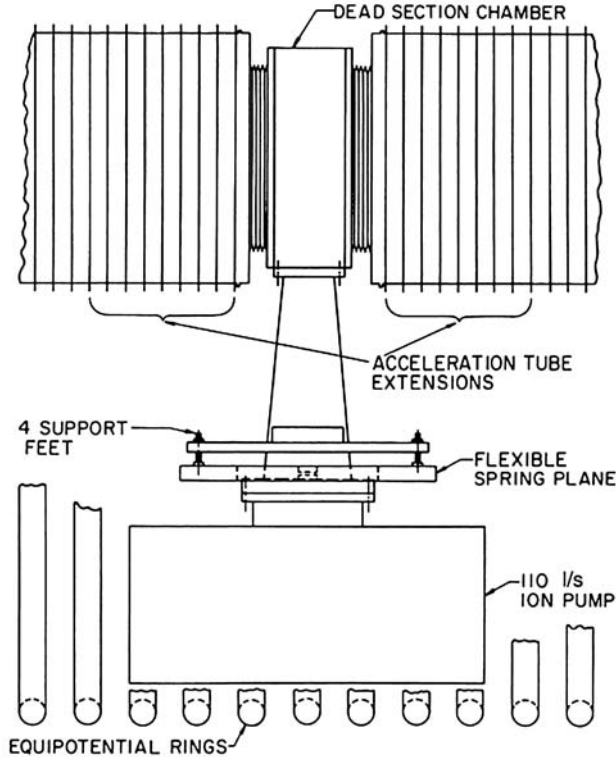


Fig. 10.1. Brookhaven National Laboratory dead-section pump (Reprinted from [4], copyright 1984, with permission from Elsevier)

the 25URC at Oak Ridge National Laboratory [6], shown in Fig. 10.2. This distributed pumping spreads the improved pressure throughout the path of the beam, greatly reducing beam-gas interactions.

The better pumps not only provide better vacuum but also inject fewer hydrocarbon contaminants into the tube. Hydrocarbons on electrode surfaces are potent sources of secondary ions, neutrals, electrons and gas, which, if present in sufficient quantity, cannot be conditioned away.

NEC led the way in reducing electrode contamination because their straight-field accelerating tubes are not very tolerant of strong electron sources. Their tubes are constructed using titanium diffusion-bonded to ceramic insulators instead of the polyvinyl acetate adhesive used to bond other tubes. To further reduce hydrocarbons, NEC also insisted on all-metal vacuum systems employing aluminum or copper gaskets, and titanium sublimation pumps coupled to ionizing pumps. However, some hydrocarbon contamination occurred while the accelerator tubes were being pumped from atmospheric pressure to ~ 0.1 Pa since oil-sealed rotary pumps were used.

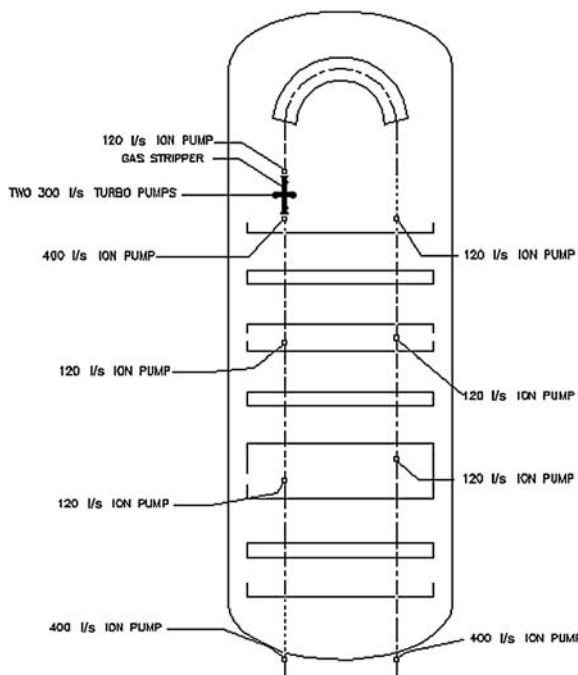


Fig. 10.2. Distributed pumping in the 25URC at Oak Ridge National Laboratory

This was in spite of employing liquid-nitrogen traps. The hydrocarbons were readily seen in residual-gas spectra. As experience with NEC machines increased, Viton O-rings were allowed, as were turbopumps, without noticeable deterioration in high-voltage performance. Whatever minor hydrocarbon contamination these introduced was readily conditioned away, allowing the tubes to perform at and above their voltage specifications.

10.4 Gas Stripping

The goal of the gas stripping apparatus in the terminal of a tandem is to present 1 to $2\mu\text{g}/\text{cm}^2$ of gas to the negative-ion beam, while minimizing the gas entering the accelerating tubes. In the simplest systems, which are adequate for light ions, gas is introduced into the middle of a canal, often $\sim 8\text{ mm}$ in diameter and $\sim 800\text{ mm}$ long. The gas exiting from the ends of the canal is restricted from going into the low-energy tube by a low-conductance section. This keeps the pressure in the low-energy tube low enough not to degrade the slow-moving injected negative ions, since the gas is preferentially pumped away via the high-energy tube. This solution is adequate for light ions and/or small accelerators, where the reasonable pumping impedance of the tubes allows the tube pressure to be adequately low.

10.4.1 Stripping Light Ions with Gas

Stripping is usually thought of as changing negative ions into positive ones, though, before the advent of the negative-helium-ion source, neutral 500 keV helium atoms were injected and then were stripped to positive using a terminal gas stripper. Early nuclear physics research using injected negative ions concentrated on light beams – hydrogen and helium isotopes. For these, gas stripping produced essentially 100% singly charged hydrogen and doubly charged helium.

There are rare situations when charge states lower than the most probable one are required. This is the case when the beam optics of the accelerator require that the terminal voltage for best beam transmission be near the design maximum. The desired beam energy is achieved by using a low charge state and the higher terminal voltage. For example, with 10 MV on the terminal and the gas stripper pressure set for equilibrium stripping, ^{12}C will have negligible intensity in the 2^+ charge state. However, if the gas stripper pressure is reduced to $\sim 20\%$ of the equilibrium value, then the 2^+ beam will have 40% of the total intensity.

10.4.2 Design Geometry

The crucial design parameter of a gas stripping system is the diameter of the stripper canal. In order for the canal not to interfere with beam transmission, its diameter must be large. Since the beam reaching the terminal generally has a low divergence, the canal length is less crucial, though the restricted terminal size limits the length of the canal. These geometry choices conflict with the goal of minimizing the gas flow to keep the pressure in the tubes low, since the conductance of the canal goes as d^3/l , where d is the canal diameter and l its length [7]. In machines primarily servicing nuclear-physics needs, the canal is often $\sim 8\text{ mm}$ in diameter and 400 to 800 mm long. In accelerators optimized for accelerator mass spectrometry, where even small changes in transmission cause problems, canals are $\sim 11\text{ mm}$ in diameter.

10.4.3 Terminal Vacuum Pumps for Gas Strippers

Various types of pumps have been used in the terminal to pump away the stripper gas exiting the canal. Ion pumps do not have the capacity to deal with the gas load, so titanium sublimation pumps were used but suffer limited lifetimes and time-varying pumping speed. Cryopumps have been used but need periodic reactivation, usually passing the gas load through the accelerating tubes. The helium compressors for the cryopumps present further challenges, since either they must be housed in the terminal or the helium gas must travel to the terminal via insulating tubing from a ground-based compressor [8]. An alternative to pumps that trap the gas is to use a turbopump to capture and then recirculate the gas back into the center of the canal [9]. Further reduction in pressure in the tubes is achievable by pumping the

gas that escapes through a low-conductance section from the turbopumped region. This differential-pumping [10, 11] solution is described in the next section, describing a typical modern gas stripper system.

Turbopumps in high-pressure, high-voltage terminals are not without problems. The most serious is the susceptibility of their power supplies to spark damage. This problem can be overcome by eliminating the power supplies by operating the pumps directly from the 400 Hz terminal alternator. This results in the pump running at about half its nominal rotational speed, with the consequent reduction in pumping efficiency. The reduction is a manageable burden on the vacuum design and has the advantage of prolonging the lifetime of the pump bearings. An electronics-free option is to use an SF_6 gas turbine to drive the pump instead of an electric motor [12]. In any case the turbopump must be modified to operate in a high-pressure environment, of 0.7 to 1.55 MPa.

Even with turbopumps employing ceramic bearings lubricated with very low-vapor-pressure grease, there is a significant hydrocarbon component in the pump backing line and thus into the stripper gas input, causing a base pressure of ~ 1.3 Pa. Since the pressure at the center of the canal may need to be less than that for molecular dissociation and for optimizing low charge states, an oil trap is required.

10.4.4 A Typical Modern Gas Stripping System

The stripper gas employed in most laboratories is either oxygen or nitrogen. Argon gas strippers are used, especially in AMS laboratories [13], to enhance high charge states where minimal noble gas gets to any ion pumps used. Typical features of a modern recirculating, differentially pumped gas stripper system are illustrated by the gas stripper in the ANU accelerator shown in Fig. 10.3 [11]. A pressure of 2.6 Pa of oxygen at the center of an 850 mm long canal will produce an integrated gas thickness, assuming a linear pressure profile, of $\sim 2 \mu\text{g}/\text{cm}^2$. Since the beam diameter at the terminal is ~ 3 mm, based on the beam spot observed on stripper foils, 8 mm was chosen for the diameter of the stripper canal itself, as well as for the low-conductance sections.

The stripper thickness is controlled with a valve that injects some gas into the turbo backing line to make up for gas that escapes through the low-conductance section. Because the quantity of gas escaping is small and the gas capture efficiency of the turbo system is high, there is a time lag of several minutes between reducing the gas input and the pressure stabilizing. A 201/s ion pump minimizes this time by pumping away gas from the volume at the exits of the canal.

After the beam traverses the stripper canal it enters another low-conductance section before passing through the region housing the foil stripper. Thus gas can be used to dissociate molecules before foil stripping or the gas can be removed and the foil stripper used on its own.

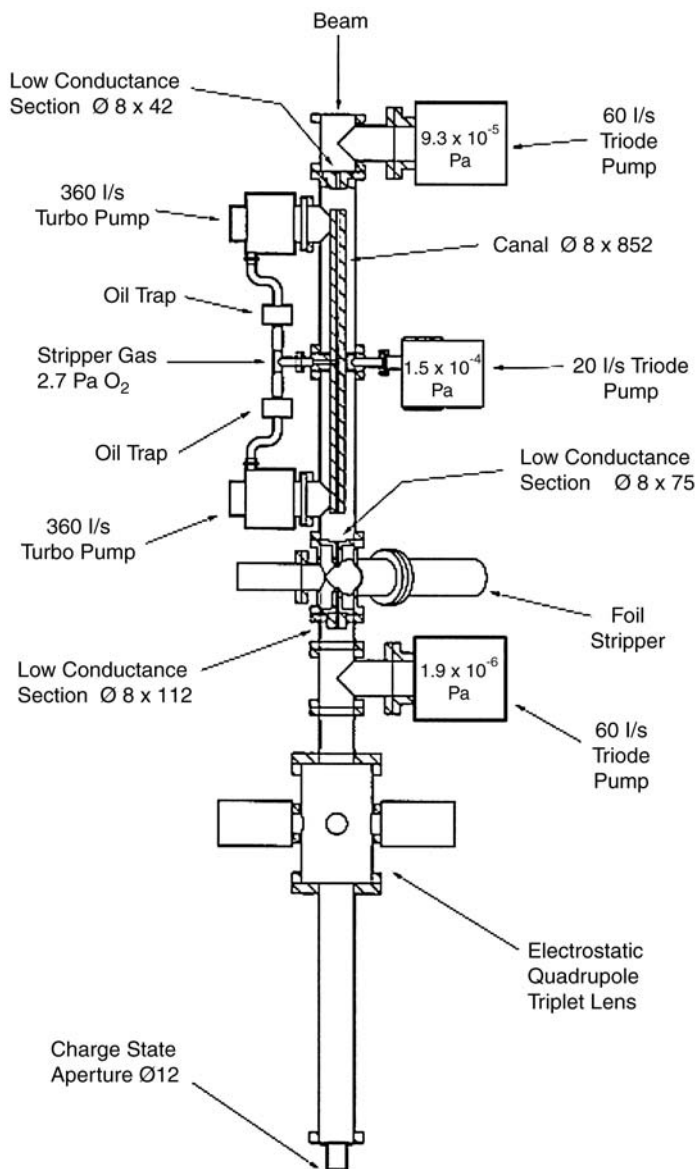


Fig. 10.3. Terminal stripper system in the ANU 14UD Pelletron. The diameters and lengths of the low-conductance sections and of the stripper canal are shown as “ $\varnothing \times \text{length}$ ” in mm

10.5 Foil Stripping

Gas stripping is perfectly adequate for the ions lighter than carbon used in the early life of tandems. As tandems started to exploit ions heavier than lithium, the need for higher energy meant that the higher charge states available from carbon foil strippers were required. This necessitated lenses to focus the beam scattered from them, as well as associated optics to stop the beam striking the high-energy tube. Electronics-free versions of foil strippers were also spark-immune, but foil thickening and breakage limited their use to ions lighter than about sulfur (Box 6).

The higher charge states from foil compared with gas stripping, a gain of two charge states for nickel beams at 15 MeV [1], comes with three costs. Firstly, the multiple scattering from the foil increases the beam emittance and thus decreases transmission. Secondly, the foil thickens under bombardment with heavy beams, resulting in a time-varying reduction in beam intensity and an increase in beam loading. Thirdly, foil lifetimes decrease as the beam mass and intensity increase. The lifetime problem has been ameliorated by the development of long-lived foils produced by various techniques (Box 6). However effective these innovations are, foil lifetime is ultimately limited by the sputtering away of the foil material [14].

Carbon stripper foils are almost exclusively used in tandem accelerators with terminal voltages up to 20 MV and for beam masses less than ~ 60 amu. Foils are successfully used for heavier-mass beams but of very low intensity for applications such as elastic recoil depth analysis using gold beams, and accelerator mass spectrometry of transuranic elements.

10.5.1 How Carbon Foils Fail

Foils almost invariably fail by splitting – well before the sputtering limit is reached. As a carbon foil is bombarded, it tightens on its frame, with stress creases radiating from the beam spot, which itself appears to remain flat, as shown in Fig. 10.4. Failure occurs as a tear, often at the edge of the beam spot or where the foil meets the frame. Once the foil is torn, the stress is relieved and if the remnant foil remains in the beam, the foil will continue to function for a very long time. This is an unpredictable but welcome event.

A confusing factor in understanding the failure of carbon foils is historic experience with carbon buildup on any foil bombarded by an ion beam in vacuum systems that contained hydrocarbons. In such vacuum systems, the buildup is generally accepted to be due to the cracking of the hydrocarbons by the beam in the beam spot. As the vacuum systems in accelerators improved, reducing the hydrocarbons in the residual gas, it was expected that foils would not get thicker and that they might last longer. This expectation proved overly optimistic. It appears that carbon migration in the foil itself continues to feed material into the beam spot, thus thickening it [14].

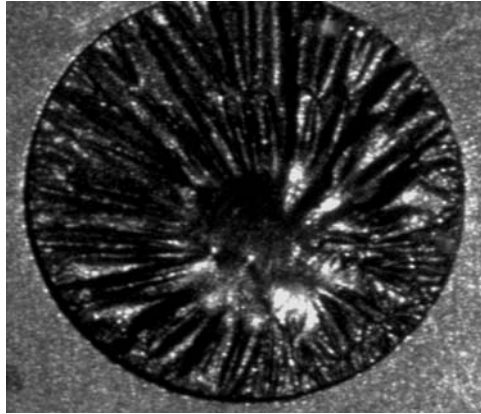


Fig. 10.4. Carbon foil stressed by beam

10.6 Second Strippers

Even higher energies are obtained for medium-mass ions by further stripping the beam a second time using a carbon foil, after the beam has gained energy from being accelerated through $\sim 1/3$ of the high-energy tube. The increase in energy and the effects of using this second stripper on beam loading are discussed next.

10.6.1 Energy Gain and Intensity Loss

Take the case of a 15 MV tandem accelerating a nickel beam with the second stripper $1/3$ of the way down the high-energy tube. After the terminal foil stripper, $\sim 27\%$ of the beam, which is in the charge state 11^+ , gains 55 MeV in traveling to the second stripper. This is a sufficiently high energy for $\sim 23\%$ of the beam to be further stripped to a charge state of 18^+ . The 18^+ component then gains an additional 180 MeV in returning to ground, resulting in a total energy of $(15 + 55 + 180)$, i.e. 250 MeV, with an overall efficiency of $\sim 6\%$. This provides enough beam intensity for most nuclear-physics experiments.

10.6.2 Beam Loading

The energy benefit of the second stripper comes with other costs in addition to the intensity penalty due to charge state fractionation. There are problems caused by both multiple scattering spreading the beam at the second stripper foil, and electrons that are emitted from it. The saving graces derive from the higher beam velocity at the second-stripper, which results in proportionally less multiple scattering than for the terminal stripper and less radiation damage. This results in second-stripper foils lasting much longer than the terminal stripper foils.

The electrons can be an immediate problem if not confined to the stripper assembly. Electrons leaving the foil that experience the electric field of the accelerating tube are accelerated toward the positive high-voltage terminal. Along the way, they strike tube electrodes and so generate copious X-ray fluxes. The X-rays ionize the insulating gas, creating severe loading of the machine and dropping its voltage by many MV. Therefore the design of the second stripper must include extremely effective electron suppression. In addition to a magnetic field and an aperture near the foil, it is essential that a suppressor electrode, held at a few kV negative, be placed between the foil and the tube going towards the terminal. This combination effectively deals with the electrons.

Although most of the electrons can be suppressed, ions cannot be. Some will strike tube electrodes in the neighborhood of the second stripper, producing electrons and their concomitant X-rays, which in turn cause loading. This connects the second-stripper region to ground in parallel with the normal resistor paths, and so reduces the voltage at the second stripper. Since the final energy of the beam depends critically on this voltage, the terminal voltage must be increased to compensate. For useful nickel beams of ~ 2 particle nA on target from the ANU 14UD, the terminal voltage will need to be increased from 15 MV without loading to 15.3 MV to compensate for the loading in the high-energy tube. This increases the gradient in the section between the terminal and the second stripper by 12%, requiring the machine to be conditioned, without beam, to at least 12% above 15 MV, viz. 16.8 MV. The loading worsens as both the terminal foil and the second-stripper foil thicken, scattering more beam onto the tube electrodes and so pushing the gradient beyond the value to which the machine has been conditioned. Beam loading limits the usable beam intensity in second-stripper operation.

It is challenging to operate a tandem accelerator near its voltage limit with double stripping. The loading causes the terminal voltage to rise, usually over several hours as the foils thicken. A rise of 0.3 MV with loading corresponds to the unloaded terminal voltage for another charge state combination at the same magnetic rigidity in the energy-analyzing magnet. Thus when either stripper foil breaks, and often it is not clear which foil has failed, the insertion of a new thin foil may allow the machine to lock onto the wrong beam at the higher terminal voltage. Worse still, the reduced loading suddenly frees the machine to rapidly increase in voltage, often until it sparks.

Despite the difficulties in using double stripping, it provides a valuable range of energies and intensities for a wide variety of nuclear-physics experiments. This is an especially important avenue to higher energies for facilities without postaccelerators but continues to fill an important niche even for facilities with them.

10.7 Charge State Selection

At tandem energies, the ions from the stripper are distributed over a range of charge states; see Box 5. The maximum intensity in any single charge state is $\sim 10\%$ for heavy beams and up to 50% for light beams [1]. Often experimenters who want the maximum possible energy choose a higher charge state and so will accept a sacrifice in intensity. A significant fraction of the other charge state beams, usually of higher intensity than the chosen one, may strike electrodes in the high-energy tube, compromising the voltage-holding ability of the machine. The several charge states from the terminal stripper can cause a confusion of beams exiting the machine when it is operated with double stripping. A charge state selector eliminates this confusion of beams by insuring only one charge state beam reaches the second stripper.

However, experience has shown that the confusion from adjacent charge states transmitted along with the desired one is not as serious as feared. Thus the complete elimination of adjacent charge states is not essential to the operation of heavy-ion tandems employing second strippers. This is just as well, since in a machine designed without a charge state selector in mind, there is usually insufficient space in the terminal to install one. Newer machines with more generous terminal space are able to employ various types of charge state selectors [10].

10.7.1 Suppression of Undesired Charge States

The selector steers the various charge state beams with an angular displacement proportional to the charge. The displacing field is adjusted so that the desired charge state beam goes through an aperture, which stops the ones that shouldn't be transmitted. The desired beam has first to be returned to the axis and then redirected along it. In a compact terminal, there is usually not enough space to accomplish this in order to uniquely select one charge state, and so some intensity of adjacent charge states gets through the selection aperture and is presented to the high-energy tube. The unwanted beam may be transmitted without striking the tube electrodes or be stopped in dead sections between the accelerator tubes [10]. In either case, the beam loading is adequately reduced.

A simpler solution to enhancing the transmission of the desired charge state from the terminal combined with the suppression of undesired ones is to focus the various charge state beams through an aperture at the entrance to the high-energy tube. This solution does require adequate space in the terminal between the lens and the charge state aperture for the lens substantially to converge the beams through it, as shown at the bottom of Fig. 10.3 [5]. The beams with substantially different charge states will be spread out across the aperture, with only on-axis beams getting through. These will be transmitted without striking the tube. The beams with adjacent charge states require about the same lens strength and so also will be transmitted without

striking the tube. This lens also serves to collect multiply scattered beam for transmission through the high-energy tube. The more elaborate charge state selectors include a lens for this purpose mounted directly at the entrance to the high-energy tube.

10.8 External Stripping

The desire for heavy-ion beams of higher energy than can be produced by tandems motivated the development of postaccelerators injected by tandems. Since the energy gain of the postaccelerator is proportional to the charge state of the injected ion, stripping of the beam after tandem acceleration is attractive.

The energy gained by an ion in a postaccelerator also depends upon how well the ion velocity is matched to the velocity acceptance of the post accelerator. A linac whose entrance section is optimized for an ion velocity of 0.1 times the speed of light, i.e. $\beta = 0.1$, is best matched to a 275 MeV nickel beam but can cope with energies down to 175 MeV. In a 15 MV tandem, the 275 MeV beam requires double stripping, with the concomitant reduction in beam intensity, while for the 175 MeV option, single terminal stripping to 12^+ suffices. External stripping to charge state 25^+ will compensate for the lower injection energy, provide similar beam intensity and avoid the difficulties of double foil stripping.

The lifetime of an external stripper foil is extremely long, since the beam intensity is small and the radiation damage from the high-energy beam is also small.

A tandem injector optimized for stable and reliable heavy-ion beams would use gas terminal stripping plus external foil stripping. The matching postaccelerator would have resonators capable of accepting the resulting lower- β beam.

References

1. J.L. Yntema: Nucl. Instr. Meth. **122**, 45 (1974)
2. R. Middleton: Nucl. Instr. Meth. **214**, 139 (1983)
3. E.P. Kanter, P.J. Conney, D.S. Gemmell, K.-O. Groeneveld, W.J. Pietch, A.J. Ratkowski, Z. Vager, B.J. Zanzbransky: Phys. Rev. A **20**, 834 (1979)
4. P. Thieberger: Nucl. Instr. Meth. **220**, 45 (1984)
5. D.C. Weissler: Rev. de Phys. Appl. **12**, 1306 (1977)
6. C.M. Jones: In: *Proc. 3rd Int. Conf. on Electrostatic Accelerator Technology*, (IEEE report 81CH163-4, 1981), ORNL, Oak Ridge, Tennessee, p. 23
7. G.L. Weissler and R.W. Carlson, *Vacuum Physics and Techniques*, New York Academic Press, 1979, p. 16
8. K.H. Purser: In: *Proceedings of the First International Conference on the Technology of Electrostatic Accelerators*, Daresbury (DNPL/NSF/R5, 1973) p. 39

9. J.B. Schroeder, C.W. Howell, G.A. Norton: Nucl. Instr. Meth. B **24**, 763 (1987)
10. H.R.McK. Hyder, J. Ashenfelter, J. Baris, C.K. Bockelman, R.O. Hamburger: Nucl. Instr. Meth. A **328**, 126 (1993)
11. D. Weisser, D. Anderson, A. Cooper, K. Fifield, G. Foote, A. Harding, N. Lobanov, A. Muirhead, H. Wallace: In: *Proc. 31st Symposium of Northeastern Accelerator Personnel*, ed. by F. Dworschak, R. Holzele (1997), Forschungszentrum, Jülich, Germany, pp. 19–30
12. R. Hellborg, K. Håkansson, M. Faarinen, M. Kiisk, P. Persson, G. Skog, K. Stenström: Pramana – J. Phys. **59**:5, 725 (2002)
13. G. Bonani, P. Eberhardt, H.J. Hofmann, T.R. Niklaus, M. Suter, H.A. Synal, W. Wölfl: Nucl. Instr. Meth. B **52**, 338 (1990)
14. R.L. Auble, J.K. Blair, D.M. Galbraith, C.M. Jones, P.H. Stelson, D.C. Weisser: Nucl. Instr. Meth. **177**, 289 (1980)

Box 5: Charge Exchange and Electron Stripping

H.J. Whitlow^{1,2} and H. Timmers³

¹ Department of Physics, P.O. Box 35 (YFL), FIN-40014, University of Jyväskylä, Finland
Harry.J.Whitlow@phys.jyu.fi

² School of Technology and Society, Malmö högskola, 206 05 Malmö, Sweden
Harry.J.whitlow@ts.mah.se

³ School of Physics, University of New South Wales at the Australian Defence Force Academy, Canberra, ACT 2600, Australia
H.Timmers@adfa.edu.au

1 Charge Exchange Processes

Charge exchange processes, where the capture or loss of electrons in a target medium changes the electrical charge state of swift ions, are of central importance for electrostatic tandem accelerators. For example, such processes are employed for:

- The formation of negative ions prior to injection into the accelerator, which is commonly used to produce He^- ions and in this case involves double electron capture by low-energy He^+ ions in an alkali-metal vapor.
- The conversion of negative to positive ions by *stripping* electrons from the injected negative ions by a target medium in the high-voltage terminal of the tandem accelerator.
- The second stripping of positive ions to create highly positive charge states, either at some point along the high-energy tube of the tandem accelerator or after acceleration. The latter may improve the suppression of background ions by the analyzing magnet or may be required to match the ion charge state with that required by a second accelerator such as a cyclotron (see Chap. 10).
- The fragmentation of molecular ions in the high-voltage terminal of the tandem accelerator. This is a crucial tool of accelerator mass spectrometry (AMS), discussed in Chap. 23. It can also be used to produce beams of elements which do not form stable negative ions. A typical example is ^{14}N , which, for example, may be accelerated by injecting the molecular ion $^{14}\text{NH}^-$ into the tandem accelerator. Electron stripping in the high-voltage terminal then renders the molecular ion unstable and releases a positive nitrogen ion.

Two fundamental charge exchange processes can be distinguished, which are *electron capture* and *electron loss*. In electron capture, the discrete charge

state q of an ion with atomic number Z_1 decreases by unity through the acceptance of an electron from the target medium, according to

$$Z_1^q + e^- \rightarrow Z_1^{q-1} \quad \text{for } q \geq 0$$

Conversely, in electron loss, an electron is stripped off and the ion charge state increases by unity:

$$Z_1^q \rightarrow Z_1^{q+1} + e^- \quad \text{for } q \geq -1$$

The conditions on q account for the fact that ions do not carry more than one negative charge. Generally, the change of the ion charge state in a target medium, such as a stripping gas or a thin, solid stripping foil, is the consequence of a multiple combination of these two fundamental processes.

Charge-State Equilibrium

As a swift ion penetrates a target medium it undergoes a large series of ion–electron collisions. The statistical probabilities for electron capture and electron loss generally differ and depend on the current charge state, the current excitation state and the velocity of the ion. This, coupled with the discrete changes of the charge state by ± 1 , implies that, as the ion traverses the medium, q changes in a stepwise manner towards a *charge-state equilibrium* q_{eq} . This is shown schematically in Fig. B5.1(a) for the case of an initially negative ion with $q = -1$ penetrating a stripping gas or foil. It is apparent that the charge-state equilibrium is a pseudoequilibrium, because the ion continues to undergo electron capture and loss processes [1]. The fluctuation of the actual charge state about the charge-state equilibrium q_{eq} results in a charge state distribution, as illustrated in Fig. B5.1(b).

The development of the charge-state distribution before equilibration can be verified experimentally. Figure B5.2 shows results for ^{12}C ions accelerated in a tandem accelerator with an N_2 gas stripper and a terminal voltage of 2.4 MV. The charge-state equilibrium q_{eq} is reached after the initially negative ^{12}C ions have traversed $\sim 0.6 \mu\text{g}/\text{cm}^2$ of the gas. In the pre-equilibrium phase, the fraction of low charge states is necessarily large, because the ions pass through the $q = 0, +1, +2$ states before equilibration with an charge-state equilibrium $q_{eq} = +2.8$. A number of useful compilations of experimentally measured charge-state distributions are available in the literature [3–8].

The charge-state equilibrium q_{eq} of a swift ion with atomic number Z_1 can be estimated from an expression based on the Thomas–Fermi effective-charge model. This expression has the form [4, 9–12]

$$q_{eq} = Z_1 \left[1 - \exp \left(-\frac{0.97v_1}{v_{TF}} \right) \right]$$

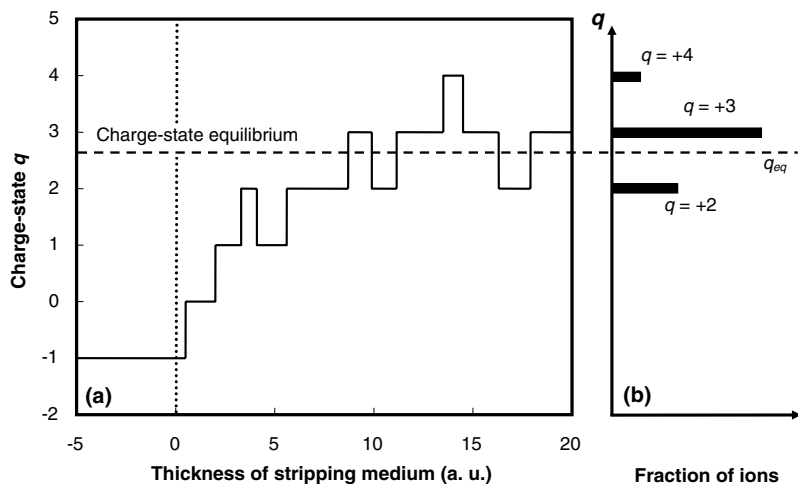


Fig. B5.1. Schematic illustration of the charge exchange processes for a negative ion penetrating a stripping gas or foil. (a) The approach to charge-state equilibrium. (b) The relation between the discrete distribution of exit charge states and the charge state equilibrium q_{eq}

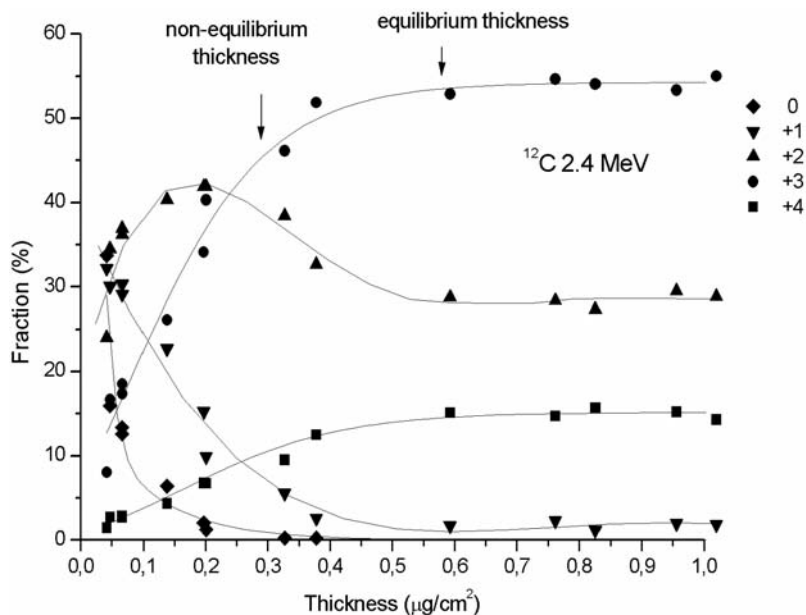


Fig. B5.2. The measured fractions of the charge states $q = 0, +1, \dots, +4$ for ^{12}C ions exiting an N_2 gas stripper in a tandem accelerator. The initial charge state and ion energy were $q = -1$ and 2.4 MeV, respectively (Reprinted from [2], copyright 2002, with permission from Elsevier)

which compares the ion velocity v_1 with the Thomas–Fermi velocity $v_{TF} = Z^{2/3}v_0$, where v_0 is the Bohr velocity. In terms of SI units, $v_0 = e^2/(4\pi\epsilon_0\hbar) = 2.188 \times 10^6 \text{ m s}^{-1}$. More sophisticated expressions for the charge-state equilibrium have been proposed [12]; however, over the energy range accessible with electrostatic accelerators (0.1–5 MeV per nucleon), this simple form is generally adequate. For a representative selection of ions, Fig. B5.3 illustrates that the charge-state equilibrium q_{eq} calculated in this way is proportional to the ion atomic number Z_1 and increases with increasing ion velocity to approach the fully stripped condition $q_{eq} = +Z_1$ asymptotically. Figure B5.3 also shows that the formation of He^- ions via double electron capture by He^+ , via $\text{He}^+ + e^- \rightarrow \text{He}^0$ and then $\text{He}^0 + e^- \rightarrow \text{He}^-$, is more easily achieved at low energies.

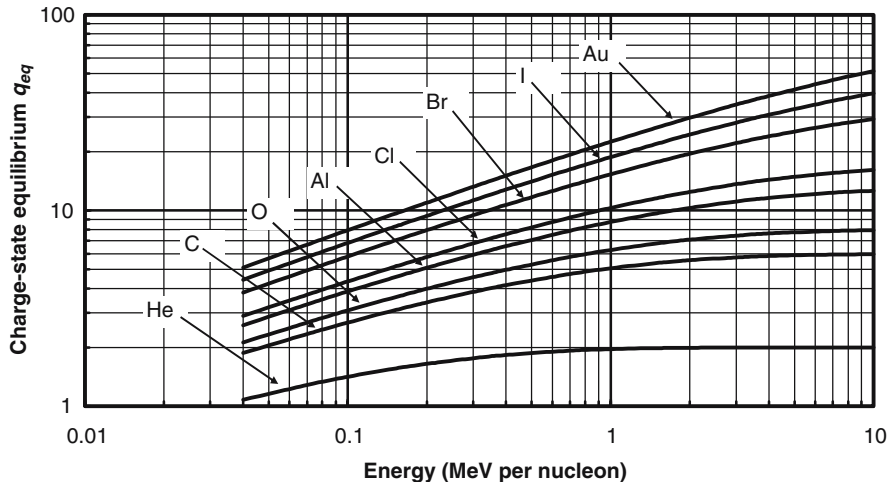


Fig. B5.3. The dependence of the charge-state equilibrium q_{eq} on energy for a range of ion species, calculated using the Thomas–Fermi effective-charge model

Gas and Foil Stripping

Foil stripping achieves a higher charge-state equilibrium than does gas stripping. As an example, Fig. B5.4 presents measured equilibrium charge-state distributions for ^{79}Br ions with an incident energy of 0.05 MeV per nucleon and an original charge state of $q = -1$ after passing through different target media, which include various gases and solid carbon. The measured charge-state equilibrium for carbon foils exceeds that for all of the gases, while also being in excess of the q_{eq} calculated using the Thomas–Fermi effective-charge model. This observation is generally attributed to the *density effect*. In dense

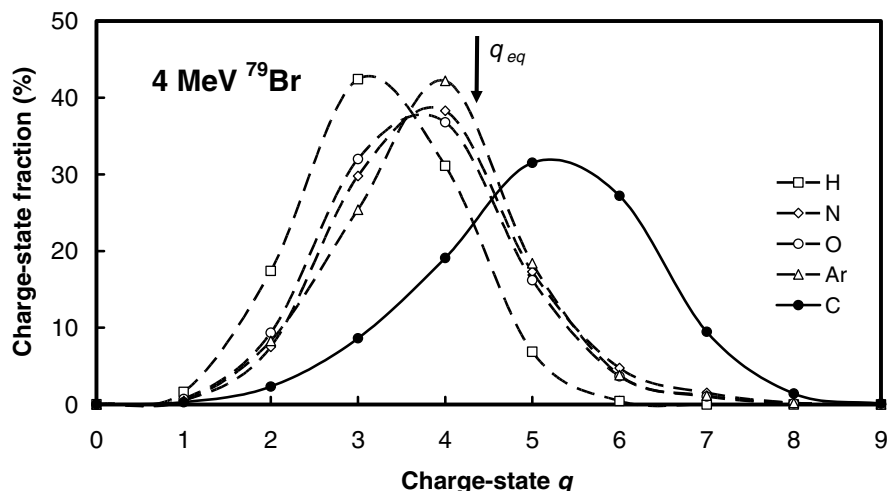


Fig. B5.4. Equilibrium charge-state distributions for 4 MeV ^{79}Br ions for various gases and for carbon-foil stripping (after Wittkower and Ryding [13]). The vertical arrow denotes the calculated q_{eq} . The curves are to guide the eye. The dashed curves denote gaseous stripping media, while the solid curve represents carbon-foil stripping

target media the electron loss of ions is enhanced, because when they are in excited states, deexcitation to the ground state may not take place before subsequent ion–electron collisions [5]. The fact that the charge-state equilibrium for the gases tends to be lower than the calculated q_{eq} has been explained, in the case of low-density gases, to arise from electron capture by doubly excited states that subsequently decay by Auger emission [5, 14]. Foil media, discussed in Box 6, are therefore, owing to their greater densities, better suited for the production of high charge states than are gases. This difference is most pronounced for heavier ions. Likewise, foils are more effective than gases for fragmenting tightly bound molecules.

References

1. P. Sigmund, L. Glasov: Nucl. Instr. Meth. B **136–138**, 47 (1998)
2. M. Kiisk, B. Erlandsson, M. Faarinen, R. Hellborg, K. Håkansson, P. Persson, G. Skog, K. Stenström: Nucl. Instr. Meth. A **481**, 1 (2002)
3. A.D. Wittkower, H.D. Betz: At. Data **5**, 113 (1972)
4. H.D. Betz: Rev. Mod. Phys. **44**, 465 (1972)
5. H.D. Betz: Appl. At. Phys. **4**, 1 (1983)
6. C.D. Moak: IEEE Trans. Nucl. Sci. **19**, 2 (1972)
7. W.N. Lennard, D. Phillips, D.A.S. Walker: Nucl. Instr. Meth. **179**, 413 (1981)
8. K. Shirma, T. Mukumo, H. Tawara: At. Data Nucl. Tables **34**, 357 (1986)

9. N. Bohr: Phys. Rev. **58**, 654 (1940)
10. J. Knipp, E. Teller: Phys. Rev. **59**, 659 (1941)
11. P. Sigmund: Phys. Rev. A **56**, 3781 (1997)
12. J.F. Ziegler, J.P. Biersack, U. Littmark: *The Stopping and Ranges of Ions in Matter*, Vol. 1, Pergamon; New York (1985)
13. A.B. Wittkower, G. Ryding: Phys. Rev. A **4**, 226 (1971)
14. G. Ryding, H.D. Betz, A.B. Wittkower: Phys. Rev. Lett. **24**, 123 (1970)

Box 6: Carbon Stripper Foils – Preparation and Quality

V. Liechtenstein

Institute of Nuclear Fusion RRC “Kurchatov Institute”, 123182 Moscow, Russia
liechten@nfi.kiae.ru

Introduction

Foil strippers of carbon are commonly employed in electrostatic accelerators for electron stripping of ion beams. Carbon foils have the advantage of being stable in vacuum at high temperatures, in combination with good electrical and thermal conductivity. Carbon has the further advantage of being the material with the lowest Z that can be fabricated into a very thin foil to minimize multiple scattering and energy straggling of the transmitted ions. In many energy ranges, lower- Z materials also can lead to higher average charge states compared with higher- Z materials [1]. However, the significant disadvantage of foil strippers is their limited lifetime due to irradiation effects, for example the thickening and shrinkage observed in carbon foils (see for example [2] and references therein. These effects both deteriorate the stripper quality, especially under heavy-ion beams, and lead eventually to the rupture of the stripper foil, with the lifetime being strongly dependent on ion mass, energy, beam density, and the vacuum environment in the terminal as well.

Preparation of Carbon Stripper Foils

Carbon Stripper Requirements

A good stripper foil should have a constant and satisfactory ion yield during the experiment. In more detail, carbon stripper foil requirements can be summarized as follows:

1. optimum thickness from the point of view of ion yield and transmission through the acceleration tube,
2. long irradiation lifetime,
3. high mechanical strength,
4. amenability to mass fabrication.

It is clear that these requirements are all related somehow to the foil thickness and preparation technique. As reported in [3, 4], the optimum thickness of a carbon stripper foil for maximal transmission at a terminal voltage of 5–10 MV is about $10\mu\text{g}/\text{cm}^2$ for light ions ($Z \leq 6$), several $\mu\text{g}/\text{cm}^2$ for

medium-heavy ions, and less than $2\mu\text{g}/\text{cm}^2$ for heavy ions ($Z > 16$). However, standard commercial foils with a thickness of several $\mu\text{g}/\text{cm}^2$ may appear impractical for terminal stripping, since they typically break in minutes when irradiated with heavy ($Z \geq 16$) ions [1,5] under normal accelerator conditions. A significant improvement in the foil stripper quality can be achieved by (a) slackening foils, and (b) preparing them by the proper technique for the optimum structure. If successful, a combination of the two methods can produce more than one order of magnitude increase in the lifetime compared with standard foils. A brief description of various preparation methods is given below for carbon strippers in the thickness range of $2\text{--}10\mu\text{g}/\text{cm}^2$, together with some comparative data.

Procedure

The majority of preparation methods for carbon stripper foils are based on deposition of the material onto a glass slide, coated with a water-soluble parting agent. The resultant film is floated off the slide and mounted on a suitable frame. The thickness of the foils (areal density) is usually measured by a light transmission method at suitable wavelengths [6]. A variety of parting agents are known from the literature [7,8]. The parting agent may influence not only the yield of usable foils but also their stripping efficiency and lifetime [9]. Most target laboratories use detergent-like parting agents, for example Teepol 610, RBS 25 Creme-Cotec, or potassium oleate ($\text{C}_{18}\text{H}_{33}\text{O}_2\text{K}$). Another large group of commonly used parting agents is the halides [9], in most cases chlorides. These have the advantage of being much more thermally stable than organic parting agents. The third group of parting agents in use are sugars, for example betaine-sucrose [10]. An important advantage of betaine-sucrose is that carbon foils produced on such a parting agent are very flexible and have a very high mechanical strength. This is due to the highly corrugated structure of the foil, which obviously replicates the significant surface roughness of the parting agent. A drawback of betaine-sucrose as a parting agent is that the covering process must be done in a humidity of near 40% to keep the proper crystallite structure of the layer [11]. Also, the nonhomogeneity of the stripper foil caused by this parting agent should not be neglected for certain experiments requiring high energy and time resolution. Free-standing carbon stripper foils of $2\text{--}5\mu\text{g}/\text{cm}^2$ are usually reinforced, prior to picking up in a suitable frame, by means of collodion (cellulose nitrate) or formvar films. The plastic coatings will evaporate in a short time when exposed to the beam.

Evaporation-Condensation Methods

The thermal evaporation of carbon in vacuum is carried out by three different techniques: (i) resistance heating of carbon filaments, (ii) arc evaporation,

and (iii) electron beam heating. Since the first application of carbon stripper foils in tandem accelerators in the 1960s [5], the arc-evaporation method has become standard, mainly because of its amenability to mass-production of inexpensive and relatively strong carbon foils in a wide thickness range [7, 12]. However, carbon stripper foils made by conventional thermal evaporation methods, especially by resistance and electron beam heating, suffer from irradiation damage that limits their use in heavy-ion tandem accelerators [9]. These foils have the shortest lifetimes. Recently, a significant improvement has been reported in the preparation procedure of carbon strippers by arc deposition called “controlled DC arc discharge”(CDAD) [13]. The major point of the CDAD method is a strictly controlled ratio of carbon clusters emitted by the anode and cathode during the DC arc discharge. The average lifetimes of very thin optimized CDAD foils were measured to be at least three times longer than those of similar standard foils.

Cracking of Ethylene Gas

Preparation of stripper foils by the DC glow-discharge cracking of ethylene gas offers improvements in lifetime by factors of 5 to 10 compared with standard carbon foils under heavy-ion bombardment, owing to the much higher resistance of ethylene-cracked foils against irradiation shrinkage [14]. The extensive development efforts since the end of the 1970s have finally resulted in the routine use of such foils for the stripping of heavy ions with $Z > 60$ in many tandem accelerators [15]. Detailed descriptions of the modifications of the ethylene-cracked foil technique are given in [5, 8, 14]. Carbon films prepared by cracking of ethylene are brittle, and the whole procedure requires experienced personnel to facilitate a reasonable foil-production yield. Also, some features of the glow discharge cracking process limit the minimal possible thickness of the foil to about $3\text{ }\mu\text{g}/\text{cm}^2$ [14]. Stripper foils prepared by cracking of ethylene have medium lifetimes [5, 8, 15].

Ion Sputtering

Ion sputtering seems attractive for the preparation of long-lived carbon stripper foils, owing to the much higher impact energy of sputtered particles compared with that in the evaporation process. This has been confirmed by using heavy-ion beam sputtering (HIBS) and ion beam sputtering with reactive-nitrogen (IBSRN) methods [13, 16]. Foils prepared using Xe^+ and Kr^+ sputtering at 3, 5, 10, and 15 keV have demonstrated significantly longer lifetimes compared with those of thermally evaporated foils under heavy-ion irradiation, and outlast even cracked ethylene foils [16]. Unfortunately, the suitability of ion-beam-sputtered foils for heavy-ion tandem accelerators seems rather problematic so far, since the minimal possible thickness of such foils is about $10\text{ }\mu\text{g}/\text{cm}^2$. In order to overcome this difficulty, magnetron sputtering can be utilized [17].

Laser Plasma Ablation

A laser plasma ablation–deposition technique to produce long-lived stripper foils has been developed at the Technical University of Munich on the basis of comprehensive investigations of the destruction mechanisms in carbon foils under heavy-ion bombardment (see for example [18] and references therein). According to the accepted theory, only carbon foils with a structure of randomly oriented nanocrystals can exhibit the longest possible lifetimes. To create such a structure, the energy of the deposited particles should be at least one order of magnitude higher than that in evaporation–condensation techniques. In order to fulfill these requirements, a high-power pulsed Nd:YAG laser (400 mJ, 10 ns) was used to shoot onto carbon targets in ultrahigh vacuum. As a result, carbon foils with the desired structure are being produced in the thickness range of 4–20 $\mu\text{g}/\text{cm}^2$. Today, stripper foils made by laser plasma deposition have been demonstrated to have the longest lifetimes of all foils. In addition, laser plasma ablation (LPA) foils have a unique mechanical strength so that, unlike any other foils, even 4 $\mu\text{g}/\text{cm}^2$ LPA foils do not need any plastic support for safe handling and mounting. However, this very successful method is extremely complicated and relatively expensive. It is also difficult to produce very thin carbon strippers by this technique. Sputter diamond-like carbon (DLC) foils, described below, have no such limitations.

Sputter Deposition of Diamond-Like Carbon Foils

The significant improvements with laser ablation stripper foils confirmed considerations that a higher energy of the deposited particles may result in an increased lifetime. This point has attracted considerable attention to diamond-like carbon films, since they are being grown using fast particles, having an energy about 30–50 times higher than that for thermally evaporated foils. Developed in the Kurchatov Institute, DLC stripper foils are produced by special-purpose DC glow discharge sputter deposition of energetic (30 eV) carbon atoms onto glass substrates cooled to liquid-nitrogen temperature. Strong and flexible DLC foils, in the thickness range from 0.6 to 20 $\mu\text{g}/\text{cm}^2$, can easily be produced by this method. The preparation and comparative testing of DLC foils under heavy-ion irradiation have been described in [19, 20] and references therein. It was observed that DLC stripper foils last more than 10 times longer than similar standard carbon foils and compare favorably with LPA foils.

Conclusions

1. A variety of advanced stripper foil preparation techniques are currently capable of fulfilling the requirements of modern tandem accelerators, although not all variables of the irradiation effects can be explained so far, and further investigations are necessary.

2. Laser plasma ablation foils and DLC foils seem to be the longest-lived stripper foils for tandem accelerators up to now.
3. In view of the necessity for thicker and long-lived strippers for higher-energy ion beams, new forms of carbon, for example fullerene (randomly oriented substitutes) and nanotubes (exceedingly high tensile strength) show promise as candidate materials.

References

1. J.L. Yntema, F. Nickel: In: *Experimental Methods in Heavy Ion Physics*, ed. by R.D. Bengthl, Lecture Notes in Physics (Springer, Berlin, Heidelberg (1987)) p. 83
2. G. Dollinger, P. Maier-Komor: Nucl. Instr. Meth. A **282**, 223 (1989)
3. G. Dollinger, P. Maier-Komor: Nucl. Instr. Meth. A **282**, 153 (1989)
4. K. Shima, S. Ishii, T. Takahashi, I. Sugai: Nucl. Instr. Meth. A **460**, 233 (2001)
5. J.L. Gallant: In: *Treatise on Heavy-Ion Science*, Vol. 7, ed. by D.A. Bromley (Plenum, New York (1985)) p. 90
6. P. Maier-Komor, G. Dollinger, E. Hammann: Nucl. Instr. Meth. A **303**, 88 (1991)
7. A.H.F. Muggleton: J. Phys. E **12**, 780 (1979)
8. P. Maier-Komor, E. Ranzinger: In: *Preparation of Nuclear Targets for Particle Accelerators*, ed. by J. Jaklovsky (Plenum, New York (1981)) p. 37
9. S. Takeuchi, C. Kobayashi, Y. Satoh, T. Yoshida, E. Takekoshi, M. Maruyama: Nucl. Instr. Meth. **158**, 333 (1979)
10. P. Maier-Komor: Nucl. Instr. Meth. **102**, 486 (1972)
11. W. Thalheimer, W. Hartmann, J. Klemm, B. Lommel: Cryst. Res. Technol. **34**, 175 (1999)
12. R. Blanc, M. Bouriant, J.P. Richauld: Nucl. Instr. Meth. A **397**, 146 (1997)
13. I. Sugai, Y. Takeda, M. Oyaizu, Y. Kawakami, Y. Hattori, K. Kawasaki, N. Hayashizaki: Nucl. Instr. Meth. A **480**, 191 (2002)
14. D.W.L. Tolfree: Nucl. Instr. Meth. **200**, 15 (1982)
15. B. Huck, E. Jaeschke, W. Kratscher, R. Repnow, H. Wirth: Nucl. Instr. Meth. **184**, 215 (1981)
16. I. Sugai, M. Oyaizu, K. Kawakami, T. Hattori, H. Tomzawa, K. Kawasaki: Nucl. Instr. Meth. A **397**, 137 (1997)
17. V.Kh. Liechtenstein, T.M. Ivkova, E.D. Olshanski, I. Feigenbaum, R. DiNardo, M. Dbeli: Nucl. Instr. Meth. A **397**, 140 (1997)
18. P. Maier-Komor, G. Dollinger, H.J. Krner: Nucl. Instr. Meth. A **438**, 73 (1999)
19. V.Kh. Liechtenstein, T.M. Ivkova, E.D. Olshanski, A.M. Baranov, R. Repnow, R. Hellborg, R.A. Weller, H.L. Wirth: Nucl. Instr. Meth. A **438**, 79 (1999)
20. V.Kh. Liechtenstein, T.M. Ivkova, E.D. Olshanski, R. Repnow, J. Levin, R. Hellborg, P. Persson, T. Schenkel: Nucl. Instr. Meth. A **480**, 185 (2002)

11 Positive-Ion Sources

L. Bartha

Institute of Nuclear Research of the Hungarian Academy of Sciences, 4001
Debrecen, P.O. Box 51, Hungary
`bartha@atomki.hu`

11.1 Introduction

This chapter is intended to center around the two main groups of ion sources of recent electrostatic ion accelerators. Sorted by placement of ion sources, the first typical group is the ion sources of single-ended machines and the terminal ion sources of tandem accelerators, and the second group is the positive-ion sources applied at the low-energy side of tandem accelerators. In all cases, so-called pressurized accelerators are considered.

In the first group, one of the following ion sources is used: an RF, a Penning or a duoplasmatron ion source. These are listed by their decreasing occurrence. Their lifetime expectancy must be over 400–1000 hours to keep maintenance down to an acceptable level. In simple solutions, the ion source is directly connected to the acceleration tube entrance, and they are matched by a single lens. In more sophisticated systems there are many lenses, a mass separator or a Wien filter between them.

In the second group, the positive-ion sources are placed in the open air, generally close to ground potential, or sometimes there is some acceleration potential between the ion source and ground. A tandem accelerator requires a negative-ion beam at its low-energy tube entrance, which can be produced directly from a negative-ion source, or indirectly using one of the former positive-ion sources, followed by a so-called charge exchanger. Either the direct or the indirect method is used; the source lifetime is less important than that of terminal sources. Tandem accelerators are often built with a double injection entrance containing an RF source or a duoplasmatron for producing light ions, and a cesium-sputtering ion source for heavier and metal ions. The advantage of this group of accelerators is obvious when a wider scale of ion species is desired.

The aim of this chapter is to survey the topic of positive-ion sources, basically in the context of single-ended accelerators. The negative-ion sources are discussed in Chap. 12. If particular positive-ion sources (e.g. sources used in low-energy, high-intensity implanters) or the more detailed theory and practice of ion sources are the main interest, readers are directed to [1–3].

11.2 The Plasma Physics of Ion Sources

The ion source is basically a plasma generator. As it is based on plasma physics, its operation can be described by basic plasma parameters and by fundamental processes.

11.2.1 Basic Plasma Parameters

Plasma Density and Degree of Ionization

The plasma state is the fourth state of matter. A plasma consists of ions, electrons and neutrals. Their number per unit volume can be given by the plasma neutral density (n_0), plasma electron density (n_e) and plasma ion density (n_i). In a plasma containing multiply charged ions, every ion component has its own ion density. In ion sources, the plasma is quasi-neutral; that is,

$$e \sum_i i n_i \cong e n_e \quad (11.1)$$

The index i goes through the different charge states ($i = 1, 2, 3, \dots Z$). This equation is not valid close to the plasma wall. Many times the plasma density simply means the density of the plasma electrons. Typical values in ion sources are 10^{10} – 10^{16} cm^{-3} .

Plasma Temperature

Generally, the concept of “temperature” is valid only for Maxwellian energy distributions, which cover many kinds of plasmas but not all. In spite of this, quite often the “plasma temperature” is used also for the plasmas of ion sources, which are not in equilibrium. The ion temperatures T_i (of i -times ionized ions) and the electron temperature T_e are not necessarily equal, and in the presence of a magnetic field the temperatures parallel and perpendicular to the field may be different, especially for the electrons. In such a case the term “plasma temperature” has no meaning. It is usual to define the plasma temperature in electron volts (eV), where 1 eV corresponds to 11 600 K. Typical plasma electron and ion temperatures are several eV. In some plasmas (e.g. ECR discharges), however, $T_{e(\text{perpendicular})}$ can be over 1 keV.

Ionization Cross Section, Mean Free Path

In the plasma, several interactions can happen in collisions: excitation, ionization, recombination and charge exchange. For every interaction, a corresponding cross section (σ) and mean free path (λ) can be defined. If one of the plasma components has a density n , then

$$\lambda = \frac{1}{n\sigma} \quad (11.2)$$

Ion Lifetime and Confinement Time

The average time between the birth and the loss of an ion with a given charge is called the lifetime. Typically τ_i is about 1 ms for highly-charged-ion sources. The lifetime in confinement sources is frequently called the confinement time. The relation between τ_i and λ is

$$\tau_i = \frac{\lambda}{v} = \frac{1}{n\sigma v} \quad (11.3)$$

where v is the (Maxwell–Boltzmann average) velocity of the ions.

Plasma Frequency

Any local deviations of the plasma particles from charge neutrality generate a force that redirects them into their original states. This process results in oscillation of the plasma particles at a frequency which is called the plasma frequency, and for the charged components of the plasma, ω_{pe} and ω_{pi} are the electron plasma frequency and ion plasma frequency, respectively. The electron plasma frequency is

$$\omega_{pe}^2 = \frac{e^2 n_e}{\varepsilon_0 m_e} \quad (11.4)$$

where ε_0 is the permittivity of free space, e and m_e are the charge and mass of the electron, and n_e is the density of electrons. A similar expression can be written for ω_{pi} by substitution of qe , m_i and n_i in place of e , m_e and n_e . Here q is the ion charge state.

Cyclotron Frequency

A charged particle with velocity v rotates in a magnetic field. The frequency of this rotation is the cyclotron frequency. For electrons, the electron cyclotron frequency is

$$\omega_{ce} = \frac{e}{m} B \quad (11.5)$$

where e is the elementary charge, B is the magnetic induction and m is the electron mass. $f_{ce} = \omega_{ce}/2\pi = 28B$ (GHz), if B is in tesla. In the case of oscillating ions, this frequency is called the ion cyclotron frequency ω_{ci} , and is given by

$$\omega_{ci} = \frac{q_i e}{m_i} B \quad (11.6)$$

where q_i is the ion charge state and m_i is the ion mass. Since the magnetic field strength applied in ion sources is usually $B = 0.1\text{--}1$ T, the cyclotron frequency for electrons is 1–30 GHz, and 100 kHz–10 MHz for ions.

Plasma Sheath and Debye Length

The charged plasma particles exert an influence on each other within a certain distance through the plasma. This distance is called the shielding distance. Particles at a greater distance than the shielding distance have no influence on the “screened” volume of the other particles. The metal wall of the plasma container or the imposed electric field similarly has no effect inside that “screened” volume, where charge neutrality is preserved. The boundary layer covering the “screened” volume is called the plasma sheath. This layer is not charge-neutral. Mainly the electrons, because of their higher mobility, follow a distribution such that it establishes an equilibrium transition layer between the plasma and its boundary. The Debye length λ_D is $1/e$ times the width of this transition layer. It is given by

$$\lambda_D = C \sqrt{\frac{T_e}{n_e}} \quad (11.7)$$

where $C = e^{-1} \sqrt{\varepsilon_0 k}$, e is the electron charge and k is the Boltzmann constant.

If a floating electrode is inserted into the plasma, it will assume a potential, called the floating potential, which is negative with respect to the plasma by about 3 to 4 times kT_e . When a high voltage is applied to this electrode, the sheath will be thicker than that of the former, unbiased case. The thickness of the high-voltage sheath is

$$d_{HVsheath} = \sqrt{\frac{V_{HV}}{kT_e}} \quad (11.8)$$

The formation of a plasma boundary has outstanding importance in the optimization of the shaping of the extracted ion beam.

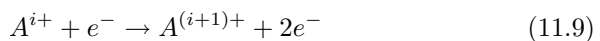
11.2.2 Fundamental Processes in the Plasma

Electron Impact Ionization

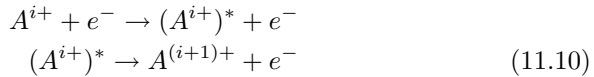
The most fundamental ionization process in ion sources is the ionization of atoms or ions carried out by energetic electrons. For maximum electron impact ionization, the electron temperature T_e (or the electron energy given by $E_e = kT_e$) should be several times larger than the ionization potential of the subshell of the atom to be ionized. In most cases the resulting ions are singly charged owing to the low energy of the colliding electrons.

Multiple ionization can take place when the electrons have an energy at least equal to the n th ionization potential. The ion charge state increases successively, starting from $i = 0$ or from some other lowest charge state, in the following ways:

- step-by-step single ionization



– single ionization by excitation in the following processes



where the asterisk means an excited state.

Ion Impact Ionization

Similarly to electron impact ionization, collision of energetic ions with neutral atoms can cause ionization. The ionization cross section is a maximum when the fast ion has a velocity equal to that of the orbital electron to be removed. This is satisfied if the ion energy is higher by the ion–electron mass ratio than the energy of the electron. The contribution of this phenomenon to the whole ionization process is rather low because of low probability of production of such energetic ions.

Surface Ionization

The phenomenon where atoms are ionized by contact with a hot metal surface is called surface ionization or contact ionization. It needs a residence time of the particles long enough (10^{-5} to 10^{-3} s) for them to come into thermal equilibrium with the hot surface. Low-work-function alkali and alkaline earth metals (Li, Na, K, Rb, Cs, Ca, Sr, Ba, etc.) in contact with refractory-metal hot plates (Ta, W, Re, Ir, Pt, etc.) are useful for ion production. For example, P_i of Cs on W at 1500 K is 0.99, and this process is used in Cs-ionizers producing Cs^+ ions for sputtering the cathodes of so-called Cs-sputtering negative-ion sources. The ionization can be increased significantly by covering the hot plate with the aluminosilicate of the alkali metal.

Field Ionization

Sharp points at high voltage can emit either electrons or ions from solids or liquids. About 10^8 V/cm electric field strength is generated at the tip of a needle or a capillary, and this very high field ionizes streaming gas, vapor or molten metal at the tip.

Ion Loss Processes

In the plasma, the opposite of the ionization process also takes place: ions can be recombined by electron capture or charge exchange. Furthermore, some of the ions and electrons exit the plasma volume and are lost at the wall of the chamber. All these processes have their own cross sections, similarly to the ionization [4].

11.3 Ion Sources

11.3.1 Low-Pressure Gas Discharges with DC Current

In the ion sources discussed in this section, ions are produced directly from a low-pressure source gas, or indirectly by bombarding ionizable substances with charged components of that gas. Depending on the ion current intensity I , this low-pressure, self-maintained gas discharge is called a Townsend discharge when $I < 10^{-6}$ A, a glow discharge when $I = 10^{-4}$ to 10^{-1} A, and an arc discharge when $I > 10^{-1}$ A, if the cross section of the discharge tube is supposed to be a few cm^2 .

Townsend Discharge

The space charge effect can be ignored because of the low current intensity. The region between the electrodes is dark, except for a slight glow appearing close to the anode.

The value of the ignition voltage U_{ign} required for the ignition of the discharge depends on the properties and pressure p of the gas, and on the electrode distance d . It has been shown that the curve of U_{ign} has a minimum in the range of $pd = 13$ to 1300 Pa cm . This minimum varies with the type of gas. The curves representing $U_{ign} = f(pd)$ are called Paschen curves. U_{ign} is also influenced by the cathode material through its secondary-electron emission coefficient, and is dramatically decreased when a heated cathode is used.

Glow Discharge

This can be experienced in the discharge volume at low gas pressures ($p = 13$ to 130 Pa) if the discharge current varies in the range $I = 10^{-4}$ to 10^{-1} A. The space charge cannot be ignored, as it causes significant distortion of the electric field.

Arc Discharge

This takes place at discharge currents $I > 10^{-1}$ A in the low-pressure discharge area.

The electron gas and the ions produced in collisions form a quasi-neutral system, which is the plasma of the arc discharge. This plasma covers almost the total volume of the discharge vessel, except for a thin layer along the vessel wall and the anode and cathode fall regions. This layer is not quasi-neutral, and its thickness along the wall is about $2\lambda_D$.

11.3.2 High-Frequency Gas Discharges

A gas discharge generated by a high-frequency field differs in many features from a DC gas discharge. Under the influence of a high-frequency field with a periodically alternating magnitude and direction, a certain amount of charged particles is unable to escape from the discharge space. As a consequence, the loss of ions and electrons is less important.

The secondary processes at the cathode surface also have smaller importance, and thus the high-frequency field can be coupled to the plasma region from both inside and outside the discharge vessel, and electrodes emerging into the plasma can be screened from ion and electron bombardment in order to avoid their sputtering and erosion. This results in a longer lifetime compared with that of DC ion sources.

The type of discharge is determined by the gas pressure p and, as a consequence, the free mean path λ and the collision frequency ν of the electrons; the frequency f of the alternating field; the electrode spacing d ; and the radius r of the discharge space.

Alternating-Electric-Field or Linear High-Frequency Discharge

The discharge tube is made of an insulating material, and a high-frequency electric field is formed by two electrodes. The electrodes are located internally or externally, one at each end of the tube. Under the influence of the electric field, the free electrons oscillate at the frequency of the field. If the electrons do not lose any energy during their oscillation (very low pressure), no energy is transferred to them from the field. If they lose energy and the energy transferred is sufficient to cause inelastic collisions, the gas molecules in the tube can dissociate. The dissociation initiates excitation and ionization of atoms and molecules.

The value of the ignition voltage U_{ign} depends on the gas species, on the frequency of the alternating field and on the gas pressure. U_{ign} is shown as a function of these parameters in Figs. 11.1 and 11.2.

Alternating-Magnetic-Field or Ring High-Frequency Discharge

The discharge tube is placed in a solenoid driven by high-frequency power. In practice, the solenoid is the inductance of a resonant circuit in a high-frequency power oscillator. The solenoid induces circular currents, and at its ends an axial electric field as well, in the discharge tube.

Let the applied alternating magnetic field be $H = H_0 \sin \omega t$. The condition for ignition of a ring discharge must satisfy the following equation:

$$rH_0 \frac{e}{m_e} = \frac{2eU_i/m_e + (\omega\lambda_e)^2}{\omega\lambda_e} \quad (11.11)$$

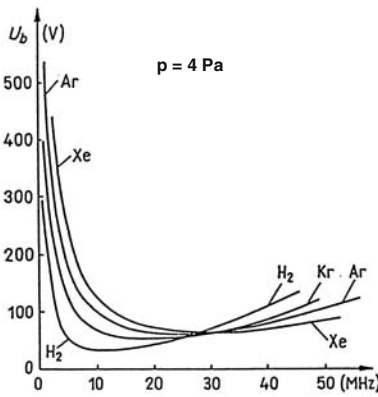


Fig. 11.1. U_{ign} as a function of high-frequency alternating field for different gases

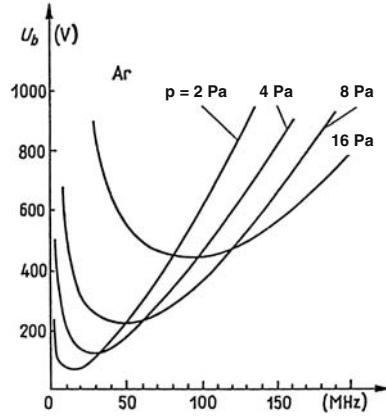


Fig. 11.2. U_{ign} as a function of high-frequency alternating field in argon for different pressures

where H_0 is the maximum value of the magnetic field; e , m_e and λ_e are the charge, mass and mean free path of an electron; and U_i is the ionization potential. The minimum value of H_0 can be calculated from (11.11), and varies with the field frequency ω and with the gas pressure p , since λ_e is proportional to $1/p$. With an axial electric field, the minimum value of H_0 can diminish.

Influence of Static Magnetic Field on Gas Discharges

In DC discharges, a magnetic field forces the electrons to circulate in a helical orbit at a Larmor frequency $f_H = eH/(2\pi mc)$, and thus the number of collisions and ionized particles increases, similarly to that what happens in the case of increased pressure. So, in the presence of a magnetic field, a given discharge intensity can be maintained at a lower pressure than without it.

In linear high-frequency discharges under the influence of a transverse magnetic field, a resonance-type increase can be observed in the electron energies, in the discharge current and in the power consumption at the frequency $\omega_H = 2\pi f_H = \omega$. Here and in the following paragraphs, ω_H is the Larmor frequency of the same arrangement in a DC electric field with an applied static magnetic field, and ω is the frequency of the alternating field.

A magnetic field parallel to the direction of the electric field applied to a linear discharge also increases the charged-particle current density, and the increase is higher when the magnetic field is inhomogeneous (see Fig. 11.3).

In a ring high-frequency discharge, its intensity increases substantially both in a transverse and in a longitudinal static magnetic field, and it is accompanied by a resonance-type power consumption and luminosity of the plasma in the $p = 0.1$ to 5 Pa pressure range. The dependence of the power

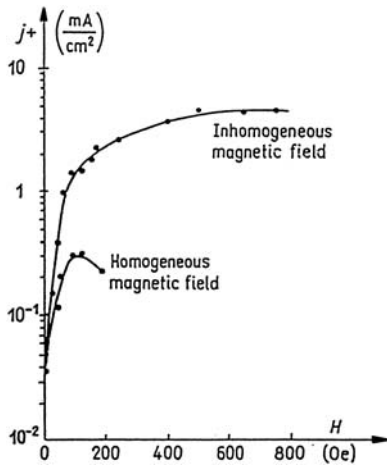


Fig. 11.3. Ion current density vs. homogeneous or inhomogeneous magnetic field

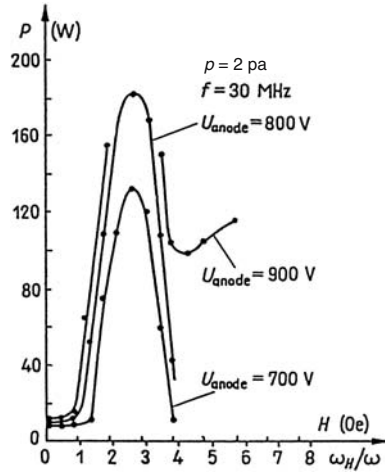


Fig. 11.4. Dependence of power consumption of ring discharge on applied transverse magnetic field and on ω_H/ω at different values of anode potential

consumption of a ring discharge on the applied transverse magnetic field and on ω_H/ω is shown (Fig. 11.4) at different values of the anode potential. The resonance in a transverse magnetic field has been observed at $\omega_H = 1.5\omega - 3\omega$, while in a longitudinal magnetic field it has been observed at $\omega_H = 3\omega - 6\omega$.

11.3.3 Traditional Ion Sources

Duoplasmatron

This is an arc discharge ion source using both a magnetic and an electric field to govern the plasma, from which the source derives its name. This source is used typically for ionizing gaseous materials. It was developed by Von Ardenne [5]. The plasma is composed of two regions inside the source: a lower-density cathode plasma maintained at a relatively high pressure (10 Pa) between the cathode and the IE, and a very high-density ($\sim 10^{14}/\text{cm}^3$) plasma at a much lower pressure (about 0.1 Pa) between the IE and the anode. The cathode plasma is compressed by a double sheath into the IE channel as in the case of the unoplasmatron and then further compressed by a more or less axial magnetic field, applied between the IE and the anode. In Fig. 11.5 a power-saving low-intensity positive duoplasmatron ion source is shown; this is advised for operation in the high-voltage terminal of a single-ended electrostatic accelerator.

Like typical duoplasmatrons, this arrangement also contains a hot cathode, and an intermediate electrode (IE) and an anode plate, both made of

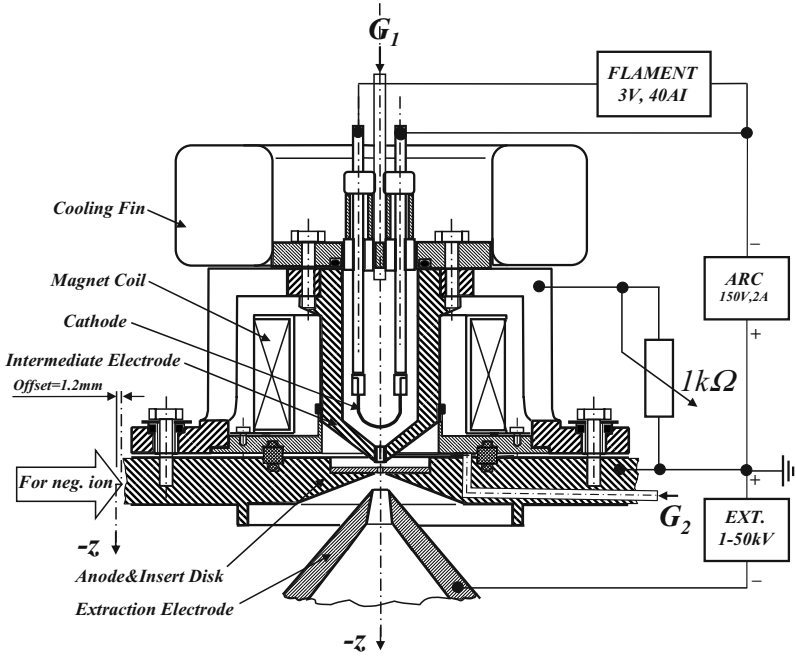


Fig. 11.5. Duoplasmatron ion source: scheme and circuitry

soft-iron in order to concentrate the magnetic flux between them along the z -axis, where the maximum magnetic induction is in the range of 0.1–0.3 T.

The cathode is heated and electrons are produced by thermionic emission. These electrons are accelerated by the electric field through the intermediate electrode toward the anode. As the IE potential is between that of the cathode and the anode, it produces some electrostatic focusing, which is accompanied by a strong magnetic focusing of the electron beam due to the magnetic flux between the IE and the anode.

When a gas (or vapor) is admitted (through G_1 in Fig. 11.5) into the cathode–IE region, the atoms or molecules enter the electron beam, and either they are ionized directly by the colliding electrons, or the molecules are dissociated and then ionized. The ion yield is dependent upon the electron beam intensity and the gas pressure. If the gas pressure is too low, the probability of collision with electrons is low. If the pressure is too high, the electron energy is insufficient and the mean free path λ is too short to cause molecular dissociation and ionization. In the case of normal operating conditions, the ions and electrons form a plasma ball in the cathode–IE region covered by an electrostatic double layer on the cathode side of the ball. The ions run toward the cathode and the electrons in the opposite direction. The number of electrons emitted from the cathode is increased by the impact of accelerated ions with the cathode.

The ion temperature in the plasma ball increases and, partly owing to compression by the double layer and partly owing to the pressure difference between the cathode-IE and IE-anode regions, the plasma material flows through the hole in the IE with these energetic ions into the IE-anode region. In this region, the strong magnetic field, regulated by the excitation of the magnet coil, causes magnetic plasma constriction. It also serves as a magnetic mirror, trapping the electrons, thus increasing their lifetime in the ionization process. For this reason, the ion temperature increases further. In the hot plasma, any neutral particle will be ionized before it escapes from the confinement area, and the plasma will go through the anode aperture. This hot plasma consists of positive ions of atomic, molecular and even triatomic particles, and, furthermore, of negative ions and electrons. Either positive or negative particles can be extracted by an extraction electrode connected to a power supply with an adequate value and polarity referred to the anode.

If this so-called extraction voltage power supply has a negative polarity with respect to the anode (as seen in Fig. 11.5), the positive ions are extracted, and the negative ones and the electrons are repelled backward. With reverse polarity, negative ions can be extracted, but the beam intensity is about two orders of magnitude lower compared with positive-ion-beam extraction. It is worth mentioning that in this case electrons can be excluded from the resulting beam by off-axis extraction (the axis of the ion source is shifted with respect to that of the extraction electrode, as shown in Fig. 11.5), or by a magnetic filter.

The resistor between the IE and the anode helps the easy ignition of the discharge, which always requires a higher voltage than does maintaining the discharge. After ignition, the IE potential gets closer to the cathode potential, to a greater or lesser extent depending on the electron current emitted from the cathode. The electron current is correctly adjusted when the potential of the IE is half or slightly above half of the arc voltage. If the IE voltage is significantly larger than this value (that is, the electron emission is below optimum), the ions (especially the heavier ones) sputter the cathode. Too high an electron emission results in high-frequency instability of the discharge.

The high-intensity arc discharge strongly erodes the IE and the anode apertures. As a consequence, the apertures will be distorted or completely closed. The lifetime of the apertures can be prolonged if sputtering-resistant liners and insert disks are applied in the IE and the anode. These are often made of graphite, tungsten, a copper-tungsten alloy, titanium or molybdenum. Basically, the lifetime of these elements and the cathode lifetime determine the source lifetime, which is around 1000 hours or above. The beam current – if everything is optimized – is roughly proportional to the area of the insert apertures, which are frequently made as an increasing series in 0.15 to 0.2 mm inner-diameter steps. Obviously, a smaller aperture results in lower gas consumption.

As the duoplasmatron is normally used only for producing light ions, the cathode lifetime is about 1000 hours. In the rare cases, when heavier ions or ions of corrosive gases are desired, these gases are admitted through the inlet G_2 of the ion source in order to avoid cathode deterioration; meanwhile, the primary discharge is maintained by the injection of a light gas. In a better solution for producing ions from heavier or corrosive gases, the ion source has a so-called expansion cup (see later), and the second gas inlet is connected to that cup.

The cathode material in H sources is typically oxide-coated tantalum, tungsten band or wire. More rarely, barium-oxide-coated platinum mesh is used. LaB_6 cathodes have, essentially, longer lifetimes, especially in high-current duoplasmatrons. If the cathode is heated directly with AC current, the filament voltage must be low (and as a consequence, the current is high), otherwise this would cause arc voltage modulation and thus excessive energy spread. In the normal case, the energy spread of the duoplasmatron ion source is about 10 to 15 eV, mainly due to the oscillation of ions in the negative anode fall of the discharge voltage distribution characteristics. A new cathode and a vented discharge chamber need a long time for outgassing, because pumping down is only possible through the small apertures of the IE and the anode. The pumping time is significantly diminished when a bypass valve is applied.

The duoplasmatron is able to produce ion currents of up to 200 mA in DC mode and a few A in pulsed mode, limited by the heat dissipation in the anode due to the high arc power, so these high-intensity ion sources are cooled by liquid coolants. In the high-voltage terminal of a single-ended machine, less than 1 mA ion current is sufficient, and the power consumption and dissipation are much lower. The magnet coil, the anode, the IE and the filament leads can be cooled by means of the high-pressure insulating gas of the machine streaming around the cooling fins. In Fig. 11.5, only the cooling fin of the IE is illustrated. A larger, ring-contoured fin is fitted to the anode, and two semicylinder-contoured fins are applied to the copper bars holding the cathode.

It is important to note that the gas pressure relations can be maintained only when the pressure on the extraction side is not higher than 10^{-3} to 10^{-4} Pa. This needs efficient pumping from the main vacuum system of single-ended machines. In the case of tandem accelerators, a directly connected high-vacuum pump is regularly applied to the ion source. The gas consumption of the source introduced is lower than $15 \text{ cm}^3/\text{h}$ under normal operation. Sometimes cm^3/h NTP, that is, cm^3/h at normal temperature and pressure, is mentioned. The standard (or normal) temperature is 273.15 K and the standard pressure is 101 325 Pa.

A further possibility for decreasing the power consumption of low-current ion sources is to replace the hot cathode by a cold, hollow cathode and the magnet coil with a permanent magnet, as announced in [6]. In this source, a permanent magnet is also used for magnetic confinement of the hollow

cathode. Initiation of the discharge requires a similar gas pressure to that which a hot-cathode source does, but the discharge voltage is 600 to 700 V, which is significantly higher than that of a hot-cathode system. After establishment, a lower value of the arc discharge can be maintained, and the gas pressure can be reduced to any desired value between 10 and 10^{-2} Pa. The power consumption remains below 62 W at 0.5 to 3.0 mA beam currents of Ar.

As the plasma density near the anode aperture is very high, many sources are equipped with a plasma expansion cup. In this cup the plasma expands and thus cools, providing an increased area over which the beam can be extracted and transported through the ion-optical system. The expansion cup is a cylindrically shaped electrode in mechanical and electrical contact with the anode at the extraction side. In some cases, the anode is made of copper and the magnetic flux immersed in the expansion cup also affects the plasma formation. The plasma sheath, the effect of the magnetic field and double-layer formation, and the extraction of the ion beam from the source plasma without and with usage of an expansion cup are discussed, and the theoretical expectations are compared with the experimental results in a brief article [7].

Duoplasmatrons are typically used for production of singly charged gaseous ions. Their ionization efficiency is about 90%, and the proton yield is 30 to 70%. Multiply charged ions can be generated by increasing the arc discharge voltage, but in DC operation this results in extremely high dissipated power. In order to avoid this, the ion source is operated in pulsed mode. Another possible way to get multiply charged ions is the optimization of the magnetic field for the maximum product $n_e \tau_i$ in DC mode. Unfortunately, the optimum magnetic field is very close to the value where instability appears. Solids, practically, are not ionized by duoplasmatrons. For this purpose, the duoPIGatron has been developed.

DuoPIGatron Ion Source

This source is a modified duoplasmatron, with a so-called reflector electrode added to it following the anode. The reflector electrode is connected to cathode potential or close to it, and the electrons coming through the anode aperture are reflected between the IE and the reflector electrode, as in PIG discharges (see later), hence the name of the source. It was constructed by Demirkhanov [8]. The oscillation of the electrons between the IE and the reflector electrode further increases their ionization efficiency, and a higher amount of gaseous ions can be produced than with the duoplasmatron. If the reflector electrode is covered by the metal of interest and a heavy gas is used as the source gas, intensive sputtering occurs on this metal surface, producing the desired metal ion. The metal ion yield increases with the magnetic field and with the arc voltage at which the gas pressure can be lowered. Singly or, in certain cases, multiply charged ions of all the elements from H to U—except

Be, K, As, Br, Rb, Ru, Rh, Te, I, Cs, Re, Os, Tl and Th—produced by the source have been reported in [9].

Penning Ion Sources

These are often called *PIG* ion sources, from F. M. Penning, the inventor of the Penning or Philips ionization vacuum gauge. Penning ion sources are arc-discharge ion sources with electrons oscillating between a hot and a cold cathode or between two cold cathodes through a hollow anode, in a magnetic field. The cathode and the anticathode are at the same negative potential with respect to the anode, of radius R , as shown in Fig. 11.6. All of these elements are placed on a common axis, which is nearly parallel to the magnetic field of a solenoid or permanent magnet. The electrons attracted by the anode and affected by the magnetic field move along an expanding helical orbit with a radius of r_e . For a given geometry, the maximum of r_e depends on the magnitudes of the electric and magnetic fields E and B , as well as on the direction of the electron velocity relative to the magnetic field. At a sufficiently high B , the maximum value of r_e becomes lower than R , and thus the electrons cross the hollow volume of the anode and proceed toward the opposite cathode. The negative potential of that cathode redirects the electrons – which have lost a certain amount of their energy in elastic and inelastic collisions – without impact with its surface. The whole process is repeated between the cathodes until the electrons have lost their energy in successive collisions with molecules to a level at which they strike the anode. The resulting ions are transported to both cathodes and release secondary electrons from them. Secondary electrons then oscillate like primary electrons and take part in the ionization process. Hence a dense plasma is formed. Owing to electron collisions with atoms and molecules, electrons diffuse through the magnetic field to the anode, where they are collected to return to the power supply, but there is another diffusion mechanism as well. The sum of the two currents gives the current of the discharge power supply.

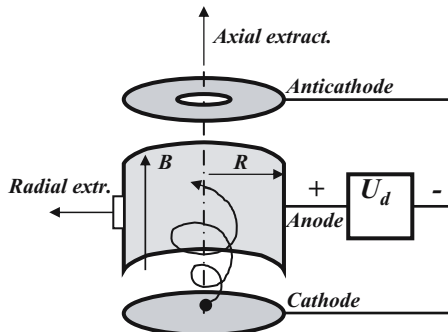


Fig. 11.6. Operation principle of the Penning ion source

The arc discharge will be self-maintained if each ion produced by a primary electron emitted from the cathode strikes the cathode and ejects at least one new electron. The secondary-electron emission coefficient δ expresses the ratio of ejected electrons n_{es} to the primary particles n_p incident on the surface: $\delta = n_{es}/n_p$. The secondary-emission coefficient of the cathode surface depends on the material of the cathode, the quality of the cathode surface and the ion energy. As a consequence, the ignition voltage U_{ign} and the formation of the self-maintained discharge are influenced by the proper choice of the cathode material and the cathode surface. By coating the cathode surface with an oxide layer, the secondary-emission coefficient can be increased. The materials used for low and higher ignition voltages U_{ign} are listed in Table 11.1. The oxide-coated materials gradually lose the oxide layer because of sputtering, and approach the U_{ign} of the pure metal. Their lifetime can be extended by 2 to 10% O₂ admixed with the ionized gas. An oxide layer deteriorated by cathode sputtering can be regenerated on the cathode surface if the discharge is operated for 10 to 30 minutes with oxygen gas. The aspects of choosing the cold-cathode material besides U_{ign} are the cathode lifetime, the value of the discharge voltage (which is lower than U_{ign}) and the discharge stability. The best material is uranium, which ensures a stable discharge at low anode voltage and with low sputtering. In most cases, titanium is also suitable. Tantalum requires a relatively low discharge voltage, and it has a very low level of sputtering. It can be used in both low- and high-current

Table 11.1. Materials used for low and higher ignition voltage U_{ign} [10]

Material	U_{ign} (V)	Sputtering Loss in H ₂ (mg/Ah)
<i>Materials suitable for low U_{ign}</i>		
Al + O ₂	350	<29
Mg + O ₂	400	
Be + O ₂	300	
Fe	400–500	68
U	500–800	<30
Ti	800–1000	<30
<i>Materials suitable for higher U_{ign}</i>		
Ni	3600	65
Zn	3600	340
Al	3500	29
Cu–Zn alloy	2800	
Monel	2800	
Cu	2300	300
C	2300	262
W	2100	57
Mo	1800	56
Ta	1700	16

sources. If the cold cathodes are intentionally heated up by the arc discharge, tantalum is a better choice.

The ions can be extracted axially along the discharge axis through one of the cathodes, or radially, where the direction of the extraction is perpendicular to the discharge axis, often through a slit aperture. Generally, PIG ion sources with axial extraction are used for higher ion currents, and radial- or transverse-extraction sources for lower ion currents, but both arrangements are suitable for both modes of operation. Radial extraction is mainly applied for cyclotrons, where the cyclotron magnet itself provides the magnetic field of the ion source.

Penning sources can be operated in two main pressure regimes: in the low-pressure regime from 10^{-6} to 1 Pa, and in the high-pressure regime from 0.1 to 100 Pa. The second regime is also divided into two types: cold-cathode PIG sources are operated with arc voltages above 1 kV at an arc power less than 1 kW, and hot-cathode PIG sources are operated with arc voltages below 1 kV. In a cold-cathode PIG operated at 1 kW or higher arc power, thermal electron emission starts. If high arc current is desired from a cold-cathode PIG (e.g. multiply charged ions are required), it must be operated in pulsed mode to limit the thermal power.

In the first heated-cathode PIG ion sources, one of the cathodes was a filament and the other cathode was floating. These sources have been used up to now in cyclotrons. In modern, high-current, heated-cathode PIG ion sources, electrons emitted from a filament, accelerated to 1 keV, are used to heat the rear side of the cathode. In this way, the cathode temperature can be controlled independently of the ion current from the plasma to the cathode, and optimum arc conditions can be adjusted. The anticathode is cold. Heated-cathode PIG ion sources deliver the highest currents of multiply charged ions of all the conventional ion sources [11]. With this type of PIG source at 1100 V and 13 A arc power, 4.9 mA Xe^{10+} and 0.01 mA Xe^{15+} ions were produced by a Ta cathode.

In single-ended machines, power consumption and vacuum load must be kept at a minimum, and the heated-cathode PIG ion source described above cannot be used. E. Heinicke, and later H. Baumann and K. Bethge, developed a PIG ion source with axial extraction suitable for these machines. The cold-cathode PIG ion source illustrated in Fig. 11.7 contains almost all of the advancements they carried out. The upper and lower cathode pole pieces, with a pot-shape and upper disk-like cores, are shown as soft-iron parts composing a soft magnetic yoke, excited by the magnet coil. The magnet coil liner, made of stainless steel, is the sidewall of the discharge chamber. The anode (molybdenum or stainless steel) is fixed in a rimmed stainless steel tube, insulated from the iron yoke by a ring-shaped insulator outside of the discharge area. The stabilized high-voltage power supplies used as discharge power supplies may have a very low internal impedance, which makes the discharge unstable. To avoid this, a resistance R is applied.

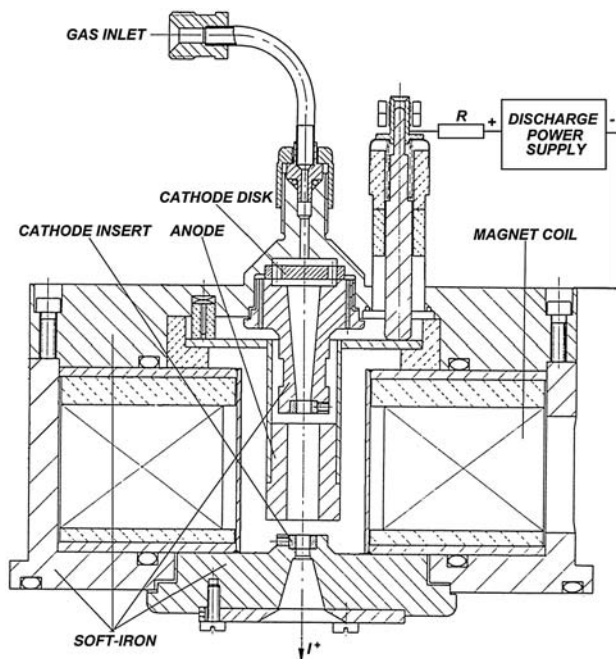


Fig. 11.7. A compact PIG ion source applicable to a single-ended electrostatic accelerator

A version of this source was published in [12]. That source – unlike the source in Fig. 11.7 – did not contain a closed magnetic yoke and thus consumed too much excitation power, and as a consequence, liquid cooling was required. The inner bore of the magnet coil bobbin was insulated from the end pieces holding the cathodes, and directly supported the anode, so discharge was possible only inside the molybdenum anode cylinder. The bore diameter of the anode was 8 mm and the length of it was 23 mm. The cathode separation was 27 mm. The volume of the discharge chamber is reduced greatly with this design, resulting in less outgassing. The cathode disks were made of tantalum. The source was operated in the high-discharge-voltage, low-discharge-current and low-pressure mode, with the following operating data: $U_d = 1$ to 6 kV; $I_d = 1$ to 20 mA; $H = 0.05$ to 0.15 T; $p \approx 0.1$ Pa. In the experiments carried out with the source, maximum total beam currents of 1 and 4 mA have been extracted with an average discharge power of about 20 W/mA. The following conclusions have been drawn about the optimal operating conditions for the production of multiply charged ions: the optimal gas flow rate and the optimal magnetic field strength have to be increased in proportion to the discharge voltage. The optimal gas flow decreases with increasing mass of the source gas, whereas the optimal magnetic field does not depend on it. This second result suggests that the magnetically guided

electron motion determines the ionization state in the discharge. This relates mainly to the primary and oscillating electrons. The ion currents for various ion species and charge states, with source operating data, are presented in Table 11.2. $I_{i,tot}$ includes the molecular-ion components as well. The energy spread of the extracted total ion beam was measured by the retarding-field method, and 60 to 80 eV was found for Ar. The gas consumption decreases with increasing mass number of the operating gas. The intensity of the ion current $I_{i,tot}$ can be expressed as $I_{i,tot} = CU_d/m_i^{1/2}$, and for the value of C , about 2.5 has been derived from the experiments.

Table 11.2. Ion currents for various ion species, and source operating data

Ion	Gas	U_d (kV)	I_d (mA)	H (T)	P (W)	$I_{i,tot}$ (mA)	$I_i^{\zeta+}$ (μA)								
							1+	2+	3+	4+	5+	6+	7+	8+	9+
He	He	3.2	15	0.07	48	3.1	3080	17							
Ne	Ne	4.5	8.4	0.076	37.8	2.7	2420	169	10	0.34					
Ar	Ar	5	6.5	0.086	32.5	2		1740	175	21	4	0.6	0.085	0.006	
Kr	Kr	4.2	4.4	0.076	18.5	1.3	1030	185	60	15	5	1.5	0.34		
Xe	Xe	5.4	3.7	0.076	20	1.25	840	230	117	40	12	4.3	1.8	1	0.42
Cl	HCl	4.2	4.8	0.097	20.1	1.6	1500	36	3.5	0.64	0.18				
Br	Br ₂	4.5	8	0.188	36	1	900	108	19	3.1					
I	HI	4	2.1	0.094	8.4	0.68	600	72	22	10	2.4	0.44			
O	O ₂	4	19	0.094	76	2.3	700	23	0.26	0.003					
S	H ₂ S	4	5.4	0.094	21.6	1.85	300	21	1.5	0.28	0.02				
Se	H ₂ Se	4.7	4.9	0.104	20.2	1.65	300	27	6	2.4					
N	N ₂	4	11.5	0.094	46	2.6	500	16	0.22	0.003					
P	PH ₃	4.1	6	0.099	24.6	1.9	150	7.7	0.82	0.12					
As	AsH ₃	4.6	5.3	0.104	24.4	1.75	200	19	5	1.9		0.1			
B	BF ₃	4	4.9	0.094	19.6	1.6	80	1.36	0.06						

This source was designed for producing gaseous ions. By replacing the Ta cathode with uranium, U ions were created by sputtering in the low-voltage, high-current discharge mode at 1.5 kV discharge voltage and 23 mA discharge current, in a 0.1 T field, with beam currents from 70 to 0.1 μA and charge states from 1+ to 8+. It is worth noting that this technique can be used in general if the desired metal, applied as a cathode, has a sufficiently high melting point for sputtering, though the metallic-ion yield is not so high.

The extraction-side cathode of the source had an opening which had not been eroded after a long operating period. The opposite cathode, originally a plain disk without a hole, showed a crater-shaped erosion profile with the diameter of the extraction-side cathode opening. This means that the plasma is concentrated along the source axis, and that the cathode lifetime, about 60 hours when Ar ions were produced, depends mainly on the wear of the plain disk cathode. Elsewhere it is found that the cathode is worn out when the diameter of the crater is equal to its depth. The other main limiting factor of PIG ion sources is the short-circuiting of the anode-cathode gap by evaporated metal layers covering the insulator surfaces, or by peeled-off ones from the anode bore surface. Both problems can be decreased by placing the

cathode disk in a field-free space in the polepiece. In the drift space formed by the cavity, the ion beam expands under the influence of its own space-charge forces from the initial 3 mm to a maximum of 7 mm before it hits the cathode disk. In this way, the sputtering current density on the cathode surface decreases by about a factor of 5, and the cathode lifetime increases in the same ratio to more than 500 hours. A further advantage of this cathode is that 90% of the sputtered cathode material is deposited inside the drift cavity. The new cathode design, published in [13], is shown as the upper cathode of the PIG source in Fig. 11.7.

The PIG ion source announced in the earlier section has been developed to produce multiply charged ions. Further developments of the source for production of ions from solids also have been published by the same authors in [14]. The basic construction of the source is very similar to that of Fig. 11.7. The power requirement of the magnet coil is only 25 W owing to the closed magnetic circuit of the soft-iron yoke, so this source runs with air cooling. Consequently, this ion source is ideal for application in single-ended machines and in high-voltage terminals of tandem accelerators as well. The three modes of operation are the following:

- *Ion production from gases.* In the basic mode, the operation data are the same as in the case of the previous source when it produced multiply charged gaseous ions. It is important to note that from the monatomic gas Ar, ions are produced with ion current intensities of 1400, 130, 17 and $2.5\text{ }\mu\text{A}$ for the 1+, 2+ 3+ and 4+ charge states, respectively, whereas from the multiatomic gas BF_3 , only ion currents of $80\text{ }\mu\text{A}$ B^+ and $1.3\text{ }\mu\text{A}$ B^{2+} are available. This means that it is better to use a solid element such as B for ion production instead of a gaseous compound.
- *Ion production from solids by evaporation.* At temperatures below 1000°C , about 30 elements have a vapor pressure in the range 10^{-2} to 10^{-1} Pa. These elements can be ionized in the high-voltage, low-current mode with the source. For this purpose the anode is replaced by a heated one, which serves as an evaporation furnace as well. The anode is fixed in the anode support by spikes to avoid thermal loss. Therefore, a power of 60 W is sufficient to heat the anode up to 700°C . In this way, the vapor of the solid can be ionized in a supporting discharge gas such as an inert gas. From P in Ne, $100\text{ }\mu\text{A}$ 1+ and $5.5\text{ }\mu\text{A}$ 2+ can be produced, from Mg in Ar, $270\text{ }\mu\text{A}$ 1+, $56\text{ }\mu\text{A}$ 2+ and $1\text{ }\mu\text{A}$ 3+ can be produced.
- *Ion production from solids by sputtering.* The evaporation method is not applicable for high-melting-point solids such as Ta and W. These elements can be applied as a cathode disk or a cathode-covering material, and sputtered by positive ions of a rare-gas discharge, and thus the required material moves to the plasma. In a PIG source such as that shown in Fig. 11.7, optimized for gaseous ions, the un-ionized part of the sputtered material is deposited on the wall of the discharge chamber and the anode. The authors have designed a modified anode and upper-cathode arrangement for

replacement of the anode and upper cathode of Fig. 11.7. In their design, the upper cathode piece is elongated and the cathode disk is placed at the tip of it, similarly to a traditional PIG cathode. In this way, the cathode separation and the discharge space between the cathodes decreased to a length of about 6 mm. The modified anode is only a thin aperture between the cathodes, supported by the old anode holder. As a result, the sputtered material moves from cathode to cathode and a small amount of it is deposited on the anode. The modified source is operated in the high-current, low-voltage discharge mode ($I_d = 40$ to 200 mA, $U_d = 0.5$ to 0.7 kV) with argon as the support gas, and Al, Ti, Fe, Cr, Ni, Cu, Mo, Ta and W singly or multiply charged ions were produced in the ion current intensity range of 20 to 200 μ A.

The value of the emittance has been given only for the high-voltage, low-current operating mode. In this mode, the emittance of the source is 0.2 to 1 cm mrad MeV^{1/2}, depending on the ion mass and beam intensity.

Investigations in [11] have pointed out that ions can be produced with sufficient efficiency and acceptable stability anywhere within a wide magnetic-field range. Owing to this recognition, many laboratories developed PIG ion sources where the magnet coils were replaced with ceramic or rare-earth permanent magnets. The PIG sources published in [15] fitted directly into the place of an HVEC RF ion source, so the ion optics were kept identical in the KN 4000 single-ended machine. One of the sources contains six ferrite ceramic magnet rings. At 0.1 T field, 1.7 kV discharge voltage, 2 to 10 mA discharge current and 4×10^{-2} Pa gas pressure, 30 nA He²⁺ current could be extracted through a cathode aperture with a diameter of 0.7 mm, which was an order of magnitude higher than the current extracted from the original RF ion source. It is impossible to pump down the residual gases through such a small aperture, and it is difficult to control the gas flow to the discharge volume, so two additional holes parallel to the extraction hole were added to increase the pumping cross section.

Both positive and negative ions can be extracted from PIG ion sources. A so-called “pocket” PIG ion source was reported recently [16] with a 3 cm diameter $\times$ 3 cm SmCo magnet and LaB₆ cathodes (with a 3 mm hole on the extraction-side cathode). The gap between the cathodes and the anode is very small compared with their diameter, similarly to the sputter mode version of the modified PIG explained previously. The source runs in the high-current, low-voltage mode at 80 mA discharge current. It can produce positive and negative ions from gaseous and solid materials with a beam current of the order of 1 mA. At 10 kV extraction voltage, about 2 mA O⁻ and F⁻ ion beams can be extracted.

RF Ion Sources

The name comes from “radio frequency”, and they are often called HF, i.e. high-frequency, ion sources, because their plasma is generated by the radio frequency field produced by a high-frequency oscillator or generator. In typical cases, an RF oscillator built with one or two transmitter electron valves provides the RF power. The frequency of the oscillator is at a constant value between 15 and 125 MHz, and the RF output power of it is 50 to 400 W.

The source is called an inductively coupled RF ion source, and the RF power is connected inductively to the discharge volume if the discharge tube is placed into a solenoid of a few turns in the resonant circuit of the oscillator. If the source is capacitively coupled, points on the previous solenoid or on a twin-lead resonant transmission line having a sufficiently high RF voltage are connected to two clips placed around the discharge volume. The discharge volume is covered by a Pyrex or quartz tube.

For the extraction of the ions, two basic methods are accepted. The first one was proposed by Thonemann et al. [17] and consists of a positive W electrode – called the probe – at the closed end of the discharge tube, and a grounded aperture made of metal at the opposite end. In the second, probeless bottle version – developed by Bayly and Ward [18] and improved by Harrison [19] and others – one of the ends of the discharge tube is closed, and both electrodes are placed at the other end concentrically to each other and to the tube. Either type of extraction can be applied with either coupling. Thus, the ion sources are called the Thonemann-type RF source with capacitive coupling, the Harrison-type RF source with capacitive coupling or with inductive coupling, etc. The different versions of RF ion sources are illustrated below by two typical examples.

RF Ion Source with Thonemann-Type Extraction and Capacitive Coupling

The RF source designed by Moak, Reese and Good [20] and shown in Fig. 11.8 is a good representative of this kind of source. The discharge tube is made of Pyrex glass with a diameter of 25 mm and a length of 220 mm. The probe or anode electrode of the source is formed by a threaded Al cylinder cemented in a vacuum-tight way in the discharge tube, cooled by an Al fin. The extracting electrode, made of Al, has a 6.3 mm long extracting canal with a diameter of 1.6 mm, and it is covered by a precisely fitted quartz tube. The oscillator, with a frequency of 100 MHz and an output power of 60 W, is connected to two clips, causing a linear high-frequency discharge in the Pyrex tube. An axial, inhomogeneous, static magnetic field is excited by the magnet coil and compresses the plasma around the axis. At 6 cm³/h gas consumption, a 1.2 mA total ion current could be extracted with a +5 kV probe voltage. The power consumption of the magnet coil is significant, and thus it is often replaced by a permanent-magnet ring.

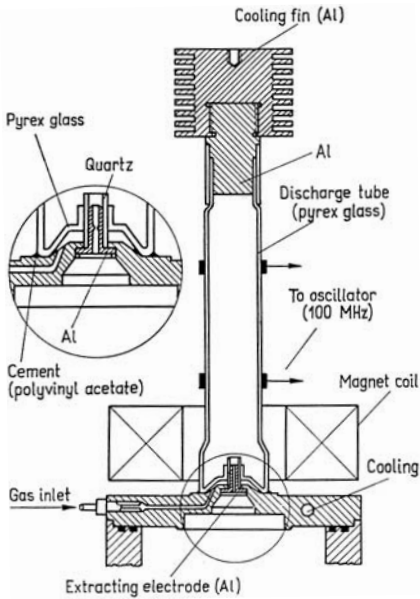


Fig. 11.8. The RF source designed by Moak, Reese and Good [20]

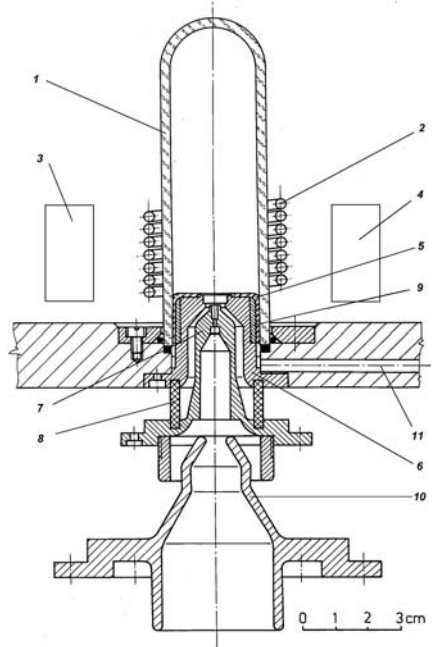


Fig. 11.9. RF ion source with a further-developed Bayly and Ward-type extraction and with inductive coupling

The RF ion source with a further-developed Bayly and Ward-type extraction and with inductive coupling shown in Fig. 11.9 represents the other group of sources [21]. In the quartz bottle (1), a high-frequency ring discharge is induced by the solenoid (2), driven with 47 MHz, 100 W RF power. A transverse magnetic field of about 5×10^{-3} T is produced between the permanent magnets (3) and (4), which greatly increases the intensity of the discharge. The potential of the anode (6) is fixed to the baseplate of the source, and the extraction electrode (7) is kept at a negative potential relative to it, which is variable between 0 and 10 kV. The quartz cup (5) shades a significant part of the anode metal surface from the plasma sheath in order to decrease recombination at it. The axial error in the manufacturing of elements (6) and (7) and in the assembly of them on the baseplate and on the accelerator tube axis is less than ± 0.03 mm, and the source is applied without any steering element at the entrance of the accelerator tube. Therefore, the deflection effect of the magnets (3) and (4) is eliminated by screening the extraction canal with a soft-iron ring (8). The source, when operated in a single-ended machine, causes a considerable gas load at the entrance of the acceleration tube, pumped only at the opposite end of the tube. The Teflon insulating ring (9) allows one to get the source gas into the tube entrance area through

the extraction bore only, and thus diminishes the vacuum load on the acceleration tube. The source gas is fed through the gas inlet (11). The proton yield of the source is about 80% if carefully optimized, and strongly depends on the clarity of the quartz discharge tube. This phenomenon can be explained by increasing recombination on the inner tube surface. Flushing of the tube with hydrofluoric acid of 4% concentration and rinsing with distilled water are required to maintain maximum proton yield.

This source is identified as ATOMKIa, with a stepped extraction canal diameter, and as ATOMKIb, with a constant diameter, in Table 11.3, where its main parameters are compared with the parameters of HVEC Model C-SO-173 (identified as SO173) and ORTEC-IONEX Model 320 (identified as IEX320) ion sources provided with Thonemann-type extraction, indicated by “Th.” “B–W” means the Bayly and Ward type of extraction. The subscript *H* means that the measurements were carried out with hydrogen. The emittances were measured at the focused beam energies and currents. The ion currents and gas consumption vs. extraction canal diameter and length are shown in Table 11.4.

The lifetime of the source is limited by erosion due to sputtering of the extraction tip, and by contamination of the discharge tube with condensed evaporated cathode materials. The condensed layer first only reduces the proton yield, and then thicker pieces of it peel off and block the extraction canal. Both symptoms increase with increasing extracted ion current and ion mass.

Table 11.3. Ion currents for various ion sources, and source operating data

Type	Ref.	Canal Dimensions		Geom. Divergence Angle deg.	Max.		Life-Time h	Emittance Measured at			ϵ	
		Length mm	Diam. mm		i_H mA	Q_H cm ³ /h		V _{Extr.} keV	V _{Focused} keV	i_H mA		
Th. SO173	[22]	12.5	2.1		2	10	700					
Th. IEX320	[23]	12.5	1.2	±5.5								
Th. IEX320	[23]	12.5	1.6	±7	0.75	4	175	2.7	35	0.5	0.35	
B–W ATOMKIa	[21]	3.0*	1.4	±10	0.25	4	600	3	10	0.1	0.43	
		4.0*	0.7									
B–W ATOMKIb	[21]	7.0	1.5	±12	3.0	11	120	3	10	0.12	0.48	

* Total length = 3.0+4.0

Table 11.4. Measured data for H gas consumption and beam intensity vs. extraction canal dimensions

Diameter (mm)		Length (mm)		Gas Consumption cm ³ /h	Ion Current mA
Canal Entrance	Canal	Canal Entrance	Canal		
1.50	0.80	4	3	4	0.22
1.50	0.90	4	3	4.7	0.24
1.50	1.00	4	3	5.5	0.45
1.50	1.20	4	3	7.8	1.60
1.50	1.40	4	3	11	2.40
	1.60		7	14	3.00
	1.80		7	28	3.90

RF sources are used typically for producing light ions such as H^+ , H_2^+ , D^+ and He^+ , but with lower intensities, heavier ions and doubly charged ions also can be extracted, as was investigated in [24] with the RF source shown in Fig. 11.9. Not only singly charged, but also molecular and doubly charged ions were produced with the application of helium, nitrogen, neon, carbon dioxide, methane and Freon-12 at moderate or low intensities. The usual frequency (47.5 MHz) of the source had been increased to 78 MHz and the gas flow minimized in order to get ions from $^4He^{2+}$ to $^{15}N^{2+}$ with usable intensities. The analysis of $^4He^{2+}$ ions is impossible with the usual magnetic and electrostatic separators since the e/m value and velocity of $^4He^{2+}$ are almost the same as those of the always present H_2^+ . To avoid this difficulty, use of the ^3He isotope has been proposed to test the efficiency of He^{2+} production [25].

11.3.4 Increased Demands

All the conventional positive-ion sources discussed up to this point are usable in single-ended machines and in terminal sources in the HV terminal of tandem accelerators, if they are designed so that their overall power and gas consumption are rather low, and they are vacuum-tight against the insulating-gas pressure. All of them can also be applied as positive or negative sources for ion injection into tandem accelerators. In these latter applications, the vacuum pumping difficulties associated with terminal sources, pressure resistance, high power consumption, the necessity for intensive cooling and shorter lifetime are not limiting factors. Still, to some extent, they are dedicated to a given purpose.

Production of ion beams with sufficient beam quality for nuclear microprobe applications is a big challenge nowadays. It is also well known that the beam quality of the ion source determines that of the accelerator. The beam quality available at some laboratories around the world was mapped by Szymanski and Jamieson [26]. Participant laboratories were asked to report their own microprobe parameters to the authors. Then they determined the beam brightness parameter B (a figure of merit of the charged-particle beam, announced elsewhere) from the formula

$$B = \frac{i}{A_o A_a E / d^2} \quad (11.12)$$

where i is the beam current on the specimen (pA), A_o is the area of the object collimator (μm^2), A_a is the area of the aperture collimator (mm^2), d is the distance between the collimators (mm) and E is the beam energy (MeV). B is obtained in the units $pA/(\mu m^2 mrad^2 MeV)$. The normal brightness and beam configuration in different laboratories are shown in Table 11.5, where the normal brightness B_n is the approximate brightness measured at a divergence of nominally 0.07 mrad for the object collimator, which has the highest brightness for the system.

Table 11.5. Normal brightness and beam configuration in different laboratories

Group	$B_n \simeq B$ at 0.07 mrad	Beam	Energy (MeV)	Ion Source	Accelerator
Debrecen	0.3	H ⁺	2	RF	Single-ended e.a. ^a
Faure	1.5	H ⁺	3	Duoplasmatron	HV KN tandem e.a.
Heidelberg	0.05	H ⁺	2.2	Penning	3 MV e.a.
JAERI	0.7	He ⁺	2	RF	NHV ^b NV3000B single-ended
Leipzig	20	H ⁺	2.25	RF	HVEE Singletron
Lund-1	2.0	H ⁺	2.7	RF	NEC 3UH
Lund-2	7.0	H ⁺	2.55	RF	NEC 3UH
Melbourne	6.0	H ₂ ⁺	3	RF	NEC 5U
Melbourne	2.0	He ⁺	3	RF	NEC 5U
Melbourne	4.5	H ⁺	3	RF	NEC 5U
Oxford	0.6	H ⁺	3	Duoplasmatron	NEC 5SDH-2 tandem
Shanghai	0.25	H ⁺	3	RF	NEC 4UH
Singapore	30	H ⁺	2	RF	HVEC AN2500
Sydney	1.0	H ⁺	3	Duoplasmatron	General Ionex Tandetron

^a e.a. = electrostatic accelerator
^b NHV = Nissin High Voltage

The following conclusions were drawn. The paraxial rays of the beam are about an order of magnitude brighter than the surrounding ones. This distribution of brightness is very similar in different accelerator systems. It seems that it does not depend on the type of machine used to provide the beam. Both single-ended and tandem machines, with RF and duoplasmatron ion sources show the same effect. The article suggests that a Penning source, which is known as a medium-quality ion beam source, is rarely used here; the duoplasmatron is used more frequently. The RF ion source is the most preferred in this field. It was also concluded that an ECR source, with 5 to 10 times higher brightness, and a gas field ionization source, with orders of magnitude higher brightness, compared with that of the best RF sources, can be promising alternatives in the future.

11.3.5 “Exotic” Ion Sources

This pallet of ion sources is much more colorful than those described in the section that dealt with traditional sources (see e.g. in [3]). Because of limited space, only two kinds of them will be briefly described below; these sources are – or can be – applied in single-ended machines.

ECR Ion Sources

The family of microwave ion sources is composed of the true ECR sources and of those operated out of the ECR regime. From (11.5) the electrons gyrate around the magnetic-field lines at the cyclotron frequency $f_{ce} = \omega_{ce}/2\pi = 28B$ (GHz), if B is in tesla. The electron gyromotion is right-handed, and

thus axially propagating right-hand-polarized waves (at a frequency of f_{ce}) can couple power into the plasma via plasma electrons. By this resonant power transfer, the electrons of the plasma can be heated to very high energies (the electron temperature can reach 10 keV). The process is called electron cyclotron resonance heating (ECRH), and the plasma is called an ECR plasma. ECRH is only possible when the electron collision frequency $\nu < f_{ce}$, i.e., the pressure of the source gas is 10^{-4} to 10^{-3} Pa. The very high-energy plasma electrons produce very highly charged ions in the ECR plasma. Microwave power can be absorbed resonantly by the plasma if the electron plasma frequency (see (11.4)) $\omega_{pe} < \omega_{ce}$, otherwise it is reflected. The microwave frequency equal to the electron plasma frequency is called the cutoff or critical frequency. Using (11.4), a critical electron density ($n_{e,crit}$) can also be derived and is given by $n_{e,crit} = 1.25 \times 10^{10} f'^2 \text{ (cm}^{-3}\text{)}$, where f' is the microwave frequency in GHz. Though, apparently, ECR cannot be created above the critical electron density, under some conditions it can be formed with a density greater than the critical density. Such a plasma is called an overdense plasma, and is advantageous when a higher ion intensity is required, for example in ion sources.

In an ECR ion source, the ECR plasma is confined by axial mirror coils or magnets, and by radial permanent magnetic multipoles. The resulting magnetic field (magnetic bucket) has a relatively low field strength in the central region, and it gradually increases in any direction from the center to the chamber wall (B -minimum geometry). This results in an effective magnetic trap for the electrons, and the space charge of the electrons keeps the ions inside this magnetic trap. The microwave power is in the range between 10 and 1000 W. Up to 5 GHz frequency and 300 W power, it is coupled with a coaxial cable; above that, it is coupled with a tube transmission line to the discharge chamber. In some cases the ECR ion source contains two discharge regions. A high-pressure (0.1 Pa) primary plasma is created in the first chamber, and the charged particles (electrons and low-charge state ions) diffuse along the magnetic-field lines into the second, low-pressure (10^{-4} Pa) chamber. The ion temperature in ECR sources is very low (lower than 1 eV), because only the electrons are heated by the microwave power in the magnetic field. The extractable ion current increases linearly with the RF power.

An elaborated compact ECR ion source has been developed and installed into the high-voltage terminal of a single-ended CN-type 6 MV electrostatic accelerator [27]. The overall power consumption of the source is 900 W. It is pressure-resistant up to 6×10^5 Pa of SF_6 . The maximum 200 W RF power is fed by eight pieces of 30 W GaAs FET amplifiers combined by a Wilkinson-combiner, and driven by a 5 GHz dielectric-resonator oscillator. High currents of lowly charged ions and reasonable currents of highly charged ions have been produced with the source, as shown in Table 11.6. The different charge states were separated with a Wien filter behind the source, and were tuned to the maximum. For lower charge states, the currents were intentionally limited to 20 μA .

Table 11.6. Analyzed currents of different charge states of the BECRIS ion source measured after acceleration and a 90° magnetic analyzer

Ions	Charge States, Currents in Electrical nA										
	1+	2+	3+	4+	5+	6+	7+	8+	9+	10+	11+
H ₂	20000										
P	20000										
D ₂	20000										
D	20000										
He	20000	10000									
N	20000	10000	2400	410	100	1.2					
O	20000	10000	400	620	130	26					
Ne	20000	10000	5000	1200	270	34	7.5				
Ar	20000	10000	3000	1450	1100	430	220	89	8	8	1.2
Xe	20000	20000	7000	4000	3000	2000	1000	1000	200	40	10

The value of the magnetic induction is 0.18 T in the ECR plane for 5 GHz. The ratio of the induction in the midplane to that in the ECR plane is 1.6, and thus the induction in the midplane is 0.11 T. For a mirror ratio of 3, a maximum magnetic induction of 0.32 T has been chosen. The radial hexapole field was realized by 36 pieces of FeNdB magnet sticks. The axial field was formed by three pairs of FeNdB rings. A schematic view of the BECRIS source is shown in Fig. 11.10.

Microwave power is often used for plasma generation at rather high pressures as well (0.1 to 1 Pa). At this pressure, the microwave-power coupling is

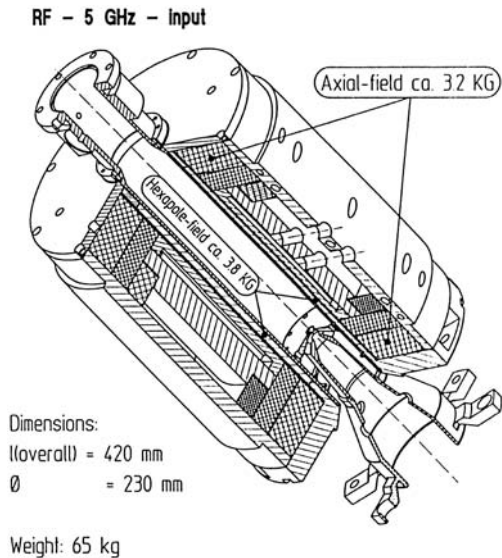


Fig. 11.10. Schematic view of the BECRIS source (Reprinted from [27], copyright 1994, with permission from Elsevier)

not resonant, since electron cyclotron motion is impossible ($\nu \gg f_{ce}$). This kind of discharge is called a high-pressure microwave discharge, and is accompanied by low-charge-state ion production. Some ion sources using 2.45 GHz microwave power produced by a magnetron high-frequency generator (made for household microwave ovens) often operate in this high-pressure discharge mode. Therefore, they are not ECR sources.

Field Ionization Sources

Field electron emission refers to the transfer of electrons from a metal surface into a vacuum under the influence of a very high electric field at the surface. Field ionization is the reverse of field emission and takes place when an atom or molecule is in an extremely high reversed-polarity electric field. The strong field polarizes the atom or molecule and lowers the potential barrier, which makes it possible for an electron to tunnel from the atom or molecule through the barrier into the metal. Using a reasonably high voltage, a very small radius of a tip is required to reach the desired 10^{-8} V/cm field strength. In practice, a needle having an apex end radius of about 100 nm and about 1 to 10 kV potential difference is applied. The small tip size results in a very high ion density, and therefore in high brightness. Examples of field ionization sources are the gas field ionization ion source (GFIS), and field-evaporation or liquid-metal ion sources (LMISs). In both ion sources, a needle or a capillary can be used. The gas pressure is 10^{-3} to 10^{-1} Pa. The needles are operated in ambient temperature or chilled to low temperatures.

Gas Field Ionization Sources (GFISs)

Figure 11.11 illustrates the two types of gas field ionization sources: (A) is the needle ionizer, and (B) is the capillary ionizer. A disadvantage of (A) is that, at high pressure, the un-ionized particles in the path of the extracted ion beam may act as scattering centers and thus degrade beam quality. An advantage is that the radius is much smaller than in the case of a capillary ionizer (a supertip). The version (B) reduces beam scattering.

A needle-type GFIS has been announced in [28]. The tip is made of tungsten or iridium. For optimum performance, the tip has to be kept at cryogenic temperatures of about 50 K. This so-called supertip has a diameter of 1 to 2 nm. The performance of the source is described in Table 11.7.

Liquid-Metal Ion Sources (LMISs)

In the needle-type version, the melted metal wets the tip. Metals with high melting points are applied in the form of low-melting-point alloys (eutectics). A Li^+ LMIS can be equipped with provision for remote wetting or rewetting of the needle. Such a source has been developed for the high-voltage terminal of an electrostatic accelerator for microprobe applications.

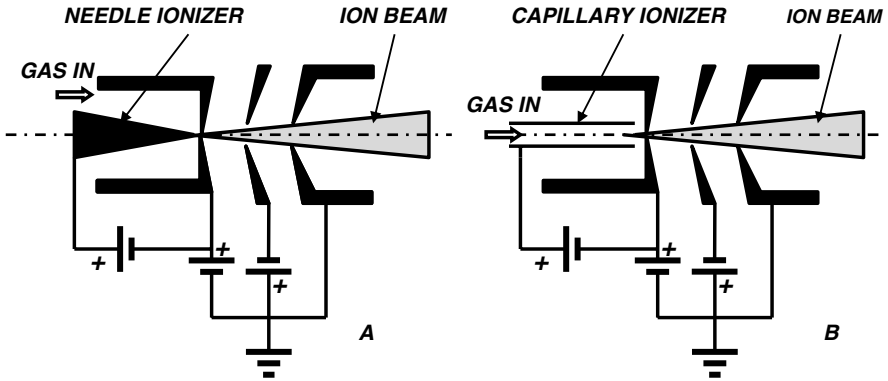


Fig. 11.11. The two types of gas field ionization sources: (A) is the needle ionizer, and (B) is the capillary ionizer

Table 11.7. The performance of the source

Particle species	e^-	H_2^+	He^+	Ne^+	O_2^+
Current (nA) into $\pm 0.5^\circ$	100	8	5	5	0.2^a
Angular current density ($\mu A/sr$)	30	35	20	20	1
Energy spread (eV)	0.3	1	1	1	1^b
Brightness ($A/cm^2 sr$)			$10^{10} - 10^{12} \text{ }^c$		
Spectral brightness ($A/cm^2 sr \cdot eV$)			$10^{10} - 10^{12} \text{ }^d$		

^a Ir supertip, maximum particle current $\approx 1 \text{ nA } O^+$ (estimate)

^b Ir supertip, extrapolated value

^c Virtual source size $d_v \approx 0.1 \text{ nm}$

^d Energy spread $\Delta E \approx 1 \text{ eV}$ (FWHM)

References

1. A. T. Forrester: *Large Ion Beams: Fundamentals of Generation and Propagation* (Wiley, New York, 1988)
2. L. Vályi: *Atom and Ion Sources* (Akadémiai Kiadó, Budapest and Wiley, London, 1977)
3. B. Wolf: *Handbook of Ion Sources* (CRC Press, Boca Raton, 1995)
4. S. Biri, E. Koltay, A. Valek: Accelerators. In: *Handbook of Nuclear Chemistry*, Vol. 5, ed. by A. Vértes, S. Nagy, Z. Klencsár (Kluwer Academic, Dordrecht, 2003) pp. 76–78
5. M. Von Ardenne; *Tabellen der Elektronenphysik, Ionenphysik und Übermikroskopie* (VEB Deutscher Verlag Wissenschaften, Berlin, 1956)
6. A. Qayyum, S. Ahmad: Nucl. Instr. Meth B **94**, 597 (1994)
7. M. E. Abdelaziz, A. M. Ghander: IEEE Trans. Nucl. Sci. **NS-14**, 46 (1967)
8. R. A. Demirkhanov et al.: Proceedings of the First International Conference on High Energy Accelerator, Brookhaven National Laboratory Report 767, p. 218 (1962)
9. B. H. Wolf: Nucl. Instr. Meth. **139**, 13 (1976)
10. L. Vályi: *Atom and Ion Sources* (Akadémiai Kiadó, Budapest and Wiley, London, 1977) p. 184

11. B. Wolf: *Handbook of Ion Sources* (CRC Press, Boca Raton, 1995) p. 75
12. H. Baumann, K. Bethge: Nucl. Instr. Meth. **122**, 517 (1974)
13. H. Baumann, K. Bethge: Nucl. Instr. Meth. **171**, 621 (1980)
14. H. Baumann, K. Bethge: Nucl. Instr. Meth. **189**, 107 (1981)
15. E. Arai et al.: Nucl. Instr. Meth. B **59** 58 (1984)
16. Y. Jinxiang et al.: Proceedings of the Second Asian Particle Accelerator Conference, Beijing, p. 179 (2001)
17. P. C. Thonemann et al.: Proc. Phys. Soc. **61**, 483 (1948)
18. A. J. Bayly, A. G. Ward: Can. J. Res. **26A**, 69 (1948)
19. E. R. Harrison: J. Appl. Phys. **29**, 909 (1958)
20. C. D. Moak, H. Reese, Jr., W. M. Good: Nucleonics **9**, 21 (1951)
21. I. Kiss et al.: Revue Phys. Appl. **12**, 1481 (1977)
22. High Voltage Engineering (Europa) N. V.: Data sheet HV-A-6021
23. General Ionex Corporation: Ion source selection data
24. L. Bartha et al.: Nucl. Instr. Meth. A **287**, 156. (1990)
25. L. Vályi: *Atom and Ion Sources* (Akadémiai Kiadó, Budapest and Wiley, London, 1977) p. 249
26. R. Szymanski, D. N. Jamieson: Nucl. Instr. Meth. Phys. Res. B **130**, 80 (1997)
27. P. Arndt et al.: Nucl. Instr. Meth. B **89**, 14 (1994)
28. S. Kalbitzer, A. Knoblauch, Invited paper at the International Conference on Ion Sources, Taormina, Sicily (7–13 September 1997)

12 Negative-Ion Formation Processes and Sources

G.D. Alton

Physics Division, Oak Ridge National Laboratory**, Oak Ridge, Tennessee
37831-6368, USA
gda@ornl.gov

12.1 Introduction

Most elements and many molecules form stable negative ions by adding an electron to the neutral atom or molecule. The existence of the negative-ion state has provided an additional means for producing charged beams. Negative ions can be formed through several processes. In this chapter, particular emphasis is placed on high-probability negative-ion formation processes of dissociative attachment, charge transfer, thermodynamic-equilibrium surface ionization and nonthermodynamic secondary-ion formation processes, and negative-ion sources. Early treatises devoted to the negative-ion state have been written by Massey [1] and Smirnov [2]. Many experimental measurements and theoretical predictions have been made of electron affinities (binding energies) of elements [3–20], as well as for electronegative molecules [21–33]. In addition, multiple electrons can be attached to certain complex-structure molecules; this subject has been reviewed [34].

Negative-ion beams have been used for many years in fundamental accelerator-based physics, as well as in applied research. As an example, H^- beams have been used as sources of neutral beams for heating plasmas by injection into controlled thermonuclear-plasma-fusion research devices. The applications have provided the impetus for the development of sources capable of producing negative ion beams containing almost every element. Sources generally are characterized by the mechanism utilized for production of negative ions (e.g. volume production in plasmas, charge exchange, thermodynamic-equilibrium surface ionization and nonthermodynamic secondary surface ionization phenomena). In certain cases, two or more mechanisms may be involved in the production process. The sources vary in operational and mechanical complexity, depending on the method required to produce the desired ion species and beam intensity for the particular application.

** Managed by UT-Battelle, LLC, for the US Department of Energy under contract DE-AC05-00OR22725.

12.2 Beam Extraction and Beam Quality Definitions

12.2.1 Beam Extraction

A number of sophisticated codes have been developed for simulating the actions of electric and magnetic fields on charged-particle beams during extraction of space-charge-dominated ion beams from plasma ion sources [35–40]. These codes have enabled the design of a variety of beam transport components, as well as the simulation of ion extraction from solid and plasma emitters. Such codes simulate the actions of the fields on charged ions moving through or accelerated by such fields. Poisson's equation

$$\nabla^2 \phi = -\rho/\varepsilon_0 \quad (12.1)$$

is solved at each point within the configuration using space-charge densities computed from the collective influence on the particle trajectories within the beam. In (12.1), ϕ is the electric potential, ρ is the charge density and ε_0 is the permittivity of free space.

The Positive-Ion Current

Sputter-type sources can be categorized according to the means of producing the positive-ion beam used to sputter the sample. Several sources utilize direct surface ionization of cesium vapor as it comes in contact with a hot, high-work-function surface to form positive ions, which are then accelerated against a negatively biased probe (cathode) containing the material of interest.

Extraction of beams at an extraction voltage $\Delta\phi_{ex}$ under conditions where ρ in (12.1) limits the ion beam current I (the value of the current that can be extracted) obeys the Child–Langmuir relation, given by $I = P \Delta\phi_{ex}^{3/2}$, where P is the perveance of the particular electrode system. For a parallel-plate electrode system, the perveance is

$$P_{pp} = \{4\varepsilon_0/9\} [2qe/M]^{1/2} \pi a^2/d^2 \quad (12.2)$$

where a is the radius of the emission aperture (circular apertures), d is the electrode spacing and M is the mass of the extracted ion species. In general, P varies from system to system because of differences in geometry and electrode design.

Optimum Perveance

The perveance for a given electrode system can be expressed in terms of the perveance P_{pp} for the parallel-plate electrode system, according to $P_{opt} \cong f P_{pp}(d_{opt})$, where f is a factor that varies from electrode system to electrode system [41].

The Negative-Ion Current

According to [41], when ions are extracted from a plasma, the plasma boundary has an optimum concave radius of curvature that minimizes the angular divergence from the source and thereby minimizes the emittances of extracted beams. Thus, the extracted current has an optimum value, given by

$$I_{opt} = P_{opt}(d_{opt}) \Delta\phi_{ex}^{3/2} = f P_{pp}(d_{opt}) \Delta\phi_{ex}^{3/2} \quad (12.3)$$

The value of the perveance at the minimum half-angular divergence ($\omega \cong 0$) is referred to the optimum perveance P_{opt} . The factor f is approximately 0.6 for extraction of space-charge-limited currents from a parallel-plate electrode system, and $f \cong 0.49$ for extraction from a two-electrode, spherical-geometry system [41].

12.2.2 Liouville's Theorem

Liouville's theorem states that the motion of a group of particles under the action of conservative fields is such that the local number density in the six-dimensional phase space volume (hypervolume) $xyzp_xp_y p_z$ everywhere remains constant. The theorem applies to an ion beam subjected to conservative fields. The quality of an ion beam is usually expressed in terms of the emittance ε and brightness B . Both are related to Liouville's theorem.

12.2.3 Emittance

For DC beam transport, the components of phase space transverse to the direction of beam motion are usually the most important. If the motions of the particles in the orthogonal planes (x, p_x) , (y, p_y) and (z, p_z) are uncoupled, the phase spaces associated with each of these planes will be separately conserved. These conserved area in the respective direction of motion is referred to as the emittance ε of the ion beam in the given direction. The conserved components of transverse phase space are taken to be elliptical in shape, the areas of which are given by $A_x = \pi\{xp_x/\pi\}$ and $A_y = \pi\{yp_y/\pi\}$. For the case where the component of momentum along the z -direction of the beam is approximately constant,

$$\begin{aligned} A_x &= \pi\{xp_z \tan \theta_{xz}/\pi\} = \pi\{xM_2 \tan \theta_{xz}/\pi\}\beta\gamma c \\ A_y &= \pi\{yp_z \tan \theta_{yz}/\pi\} = \pi\{yM_2 \tan \theta_{yz}/\pi\}\beta\gamma c \end{aligned} \quad (12.4)$$

where $\beta = v/c$ and $\gamma = 1/[1 - \{v/c\}^2]^{1/2}$. M_2 is the mass of the particle, of speed v ; c is the speed of light; and θ_{xz} , θ_{yz} are the angles that the projection of the particle's velocity in the xz and yz -planes make with the z -direction (direction of motion). In the small-angle approximation, $\tan \theta_{xz} = \theta_{xz} = dx/dz = x'$ and $\tan \theta_{yz} = \theta_{yz} = dy/dz = y'$, so that

$A_x = \pi\{xx'M_2/\pi\}\beta\gamma c$ and $A_y = \pi\{yy'M_2/\pi\}\beta\gamma c$ (relativistic case). For the nonrelativistic case, $A_x = \pi\{xx'/\pi\}[2EM]^{1/2}$ and $A_y = \pi\{yy'/\pi\}[2EM]^{1/2}$, where E is the energy of the ion beam. The emittance ε of the ion beam is proportional to the transverse phase space and thus is a conserved quantity. The following definitions related to the energy E and velocity v have been adopted historically for the normalized emittances ε_{nx} and ε_{ny} :

$$\varepsilon_{nx} \approx \pi \int \int [dx dx' / \pi]; \quad \varepsilon_{ny} \approx \pi \int \int [dy dy' / \pi] \quad (12.5)$$

or

$$\varepsilon_{nx} \approx \pi \int \int [dx dx' / \pi] \sqrt{E}; \quad \varepsilon_{ny} \approx \pi \int \int [dy dy' / \pi] \sqrt{E} \quad (12.6)$$

where the integrations are performed over the emittance contour that contains a specified fraction of the beam (e.g. 10%, 20%, ... 90%, etc.) Equation (12.6) has been historically adopted as defining the normalized emittance in the electrostatic-accelerator community.

12.2.4 Brightness

Another figure of merit, often used for evaluating the properties of ion beams, is the brightness B . Brightness is defined in terms of the ion current dI per unit area dS and per solid angle $d\Omega$, or $B = d^2I/dS d\Omega$. The normalized brightness B_n can be expressed as $B_n = 2d^2I/\varepsilon_{nx}\varepsilon_{ny}$ [42].

The normalized root-mean-square (RMS) emittance $\varepsilon_{n,RMS}$ and the normalized physical emittance ε_n are usually defined according to the following formulas:

$$\begin{aligned} \varepsilon_{n,RMS}(x, x') &= \pi[(\langle x^2 \rangle \langle x'^2 \rangle - \langle xx' \rangle^2) / \pi^2]^{1/2} \beta\gamma \\ \varepsilon_{n,RMS}(y, y') &= \pi[(\langle y^2 \rangle \langle y'^2 \rangle - \langle yy' \rangle^2) / \pi^2]^{1/2} \beta\gamma \end{aligned} \quad (12.7)$$

where the relationship between $\varepsilon_{n,RMS}$ and ε_n (90% contour value) is often defined as

$$\varepsilon_n(x, x') = 4\varepsilon_{n,RMS}(x, x'); \quad \varepsilon_n(y, y') = 4\varepsilon_{n,RMS}(y, y') \quad (12.8)$$

In this context, the normalized RMS brightness $B_{n,RMS}$ is defined by $B_{n,RMS} = 2d^2I/\{\varepsilon_{nx,RMS}(x, x')\varepsilon_{ny,RMS}(y, y')\}$, where dI is the ion beam current within a specified emittance contour.

12.3 The Negative-Ion State

12.3.1 The Electron Affinity

The processes involved in the attachment of electrons to neutral atoms or molecules to form negative ions are exothermic, in contrast to the endothermic

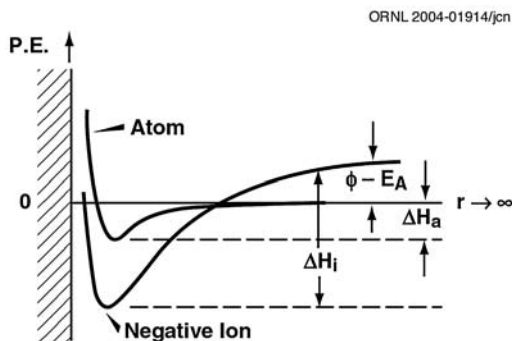


Fig. 12.1. Potential-energy curve for an atom and negative ion as a function of distance from a hot metal surface. ϕ is the work function of the metal; ΔH_a and ΔH_i are the respective enthalpies of adsorption for the atom and ion; and E_A is the electron affinity of the atom

processes required for positive-ion formation. The parameter that characterizes the negative-ion state is the binding energy of the additionally attached electron, or the electron affinity, E_A . The electron affinity of a negative ion is a measure of the stability and ease of ion formation, as well as ease of detachment, and therefore is defined as the difference between the energy of the neutral ground state E_0 and the energy of the negative ion E_- , or $E_A = E_0 - E_-$. E_A must be positive to form a stably bound negative-ion state.

Consider the potential energy of an atom with electron affinity E_A as a function of the distance from the surface of a metal of work function ϕ , as illustrated in Fig. 12.1. The energy required for removing the ion from the surface is ΔH_i , and the condition for residing on the surface as an ion is $\Delta H_i - (\phi - E_A) > \Delta H_a$, where ΔH_a is the enthalpy of adsorption. An ion supplied with an energy ΔH_i may be transferred to the continuum in either ionic or atomic form in cases where the potential-energy curves cross. At a distance far from the influence of the metal on the affinity level of the atom, the atomic and ionic potential curves will be separated by an amount $(\phi - E_A)$. The probability of arrival at a position far from the metal in a given state depends on the magnitude of $(\phi - E_A)$.

12.3.2 Electron Affinities of the Elements and of Molecules

Approximately 78% of the naturally occurring elements have positive electron affinities. Figure 12.2 displays first ionization potentials and electron affinities of the elements [3–20]. The electron affinities have values ranging from $E_A \leq 0$, to $E_A = 3.613 \text{ eV}$ for Cl. Notable examples of elements with negative electron affinities are Be [4,9] and Mg [4] of the group IIA elements, N of the group VA elements [4], Mn of the group VIIB elements [4], and all members

ORNL 2004-01835A/jcn

GROUP 1A								VIII A
1 H 13.595 0.7542	Ionization Potential Electron Affinity						2 He 24.58 0.078 *	
	II A	III A	IV A	V A	VI A	VII A		
3 Li 5.39 0.618	4 Be 9.32 0.19 *	5 B 8.30 0.277	6 C 11.26 1.263	7 N 14.54 Unstable	8 O 13.61 1.461	9 F 17.42 3.401	10 Ne 21.56 Unstable	
11 Na 5.14 0.548	12 Mg 7.64 Unstable	13 Al 5.98 0.441	14 Si 8.15 1.385	15 P 10.55 0.746	16 S 10.36 2.077	17 Cl 13.01 3.613	18 Ar 15.76 Unstable	
19 K 4.34 0.501	20 Ca 6.11 0.0245	31 Ga 6.00 0.3	32 Ge 7.88 1.233	33 As 9.81 0.81	34 Se 9.75 2.021	35 Br 11.84 3.364	36 Kr 14.00 Unstable	
37 Rb 4.18 0.486	38 Sr 5.69 0.052	49 In 5.78 0.3	50 Sn 7.34 1.112	51 Sb 8.64 1.07	52 Te 9.01 1.971	53 I 10.45 3.059	54 Xe 12.13 Unstable	
55 Cs 3.89 0.4716	56 Ba 5.21 0.144	81 Tl 6.11 0.2	82 Pb 7.41 0.364	83 Bi 7.29 0.946	84 Po 8.43 1.9	85 At — 2.8	86 Rn 10.74 Unstable	
87 Fr 4.0 0.46	88 Ra 5.3 0.100	* Meta-stable						

III B	IV B	V B	VI B	VII B	VIII B			I B	II B
21 Sc 6.56 0.188	22 Ti 6.83 0.079	23 V 6.74 0.525	24 Cr 6.76 0.666	25 Mn 7.43 Unstable	26 Fe 7.90 0.151	27 Co 7.86 0.662	28 Ni 7.63 1.156	29 Cu 7.72 1.235	30 Zn 9.39 Unstable
39 Y 6.50 0.307	40 Zr 6.95 0.426	41 Nb 6.77 0.893	42 Mo 7.18 0.746	43 Tc 7.28 0.55	44 Ru 7.36 1.05	45 Rh 7.46 1.137	46 Pd 8.33 0.557	47 Ag 7.57 1.302	48 Cd 8.99 Unstable
57 La 5.61 0.500	72 Hf 7.00 ≅ 0	73 Ta 7.88 0.322	74 W 7.98 0.815	75 Re 7.87 0.15	76 Os 8.70 1.10	77 Ir 9.00 1.565	78 Pt 8.96 2.128	79 Au 9.22 2.309	80 Hg 10.43 Unstable

Fig. 12.2. Electron affinities and ionization potentials of the elements in eV

of the group IIB elements (Zn, Cd and Hg) [4] and VIIIA elements (He, Ne, Ar, Kr, Xe and Rn) [4], while Hf is estimated to have an electron affinity $E_A \cong 0$ [4].

The group IIA elements span the complete spectrum of possible electron affinities. Be has a negative ground-state electron affinity, but is bound metastably in the $1s^2 2s 2p^2(^4P)$ state; the ion is metastable against spin-spin interactions, has a lifetime of a few μs and thus lives long enough for practical use [4, 9]. Mg, on the other hand, forms neither a bound ground state nor a metastably bound state and thus cannot be produced as an atomic negative ion [4].

In addition to negative atomic species, many molecular negative ions have been observed. In many cases, molecular negative ions containing the atom of interest have much higher electron affinities than has the atom itself and, therefore, can be formed with higher probability than the atomic species can be. In some cases, molecular ions offer the only alternative for producing beams containing elements that do not form stably bound negative-ion states. For tandem electrostatic-accelerator applications, the unwanted species can be easily rejected by collision dissociation in the positive-ion formation process (stripping) followed by magnetic (M/q) analysis for selection of the atomic species of interest. Electron affinities of a selected number of diatomic, triatomic and polyatomic molecular species are tabulated in Table 12.1 [21–33].

Table 12.1. Electron affinities (E_A) in eV of selected diatomic, triatomic and polyatomic molecules

Molecule	E_A	Molecule	E_A	Molecule	E_A
Al ₂	1.10	Al ₃	1.40	SF ₄	2.35
BO	3.12	BO ₂	3.57	SF ₆	1.05
BeH	0.7	CoH ₂	1.45	TeF ₆	3.34
C ₂	3.269	Cs ₃	0.864	UF ₅	4.4
CH	1.238	K ₃	0.956	UF ₆	5.1
CN	3.821	N ₃	2.70	WF ₅	1.25
CaH	0.93	NO ₂	2.273	WF ₆	3.36
Co ₂	1.110	Na ₃	1.019		
CrO	1.222	PH ₂	1.271		
Fe ₂	0.902				
FeH	0.934				
MgH	1.05				
MnH	0.869				
NS	1.194				
OH	1.827				
PH	1.028				
PO	1.092				
Re ₂	1.571				
ZnH	0.95				

12.3.3 Metastable-Negative-Ion Formation

There is an important class of negative ions that can be formed in excited states of the parent atom and live long enough to be useful. In order to induce electron attachment and thus form such metastable states, it is necessary to form the particular excited electronic state of the neutral atom to which

an electron can be attached. The charge transfer process (discussed later in this chapter) is particularly well suited for metastable-ion formation. The subsequently formed negative ion may live for extended periods of time if the excited compound state is forbidden to decay. Classic examples of metastable negative ions that can be formed only through attachment to excited states of the parent atoms are He^- [4–8] and Be^- [4, 9], both forming with high probability in the ^4P state through sequential charge exchange with a low-ionization-potential charge exchange vapor.

12.4 Negative-Ion Formation in a Plasma Environment

12.4.1 Radiative Electron Capture and Dielectronic Attachment

The simplest way to form negative ions is by direct capture of a free electron by a neutral atom. However for attachment to take place, the difference between the kinetic energy of the electron and electron affinity of the atom must be conserved through photon emission or momentum transfer to the motion of the atom.

Another attachment mechanism is possible, involving radiative stabilization. Attachment is possible if the incident electron has an energy such that the energy of the atom plus electron is within the level width of a doubly excited state of the atom. Thus, it is possible to capture the electron by reverting to the ground state by radiative emission; the atom may revert back to the neutral by ejecting the electron back into the continuum. Radiative electron capture and dielectronic attachment are low-probability processes and, therefore, not important in high-intensity negative-ion sources.

12.4.2 Polar Dissociation Attachment

Polar dissociation attachment may occur as a result of interactions with molecular neutrals in which sufficient energy is imparted to the molecule to excite it to an unstable state that dissociates spontaneously into positive and negative ions. Polar photodissociative attachment is a process whereby a molecule XY absorbs a photon of sufficient energy to cause spontaneous fragmentation according to the reaction



Negative-ion formation may also occur by this mechanism whenever an electron of sufficient energy for molecular excitation to an unstable state interacts with a molecule XY as follows:



We note that the electron is not itself captured, but only serves as the means by which the molecular excitation occurs.

12.4.3 Dissociation Attachment

Electrons may be stably attached to atoms during interactions with molecular neutrals according to the following process:

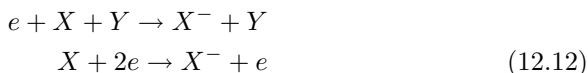


The process may be viewed as a three-body process where the excess energy released during the reaction can be absorbed by transfer to the relative motion of the atomic nuclei or fragments, and thus the state can be readily stabilized.

On the other, hand attachment to a molecule can occur for molecules with positive electron affinities without dissociation, thereby forming a molecule XY^- . Vibrational states of the molecule will be stable for frequencies $\nu < \nu'$; those with vibrational frequencies $\nu > \nu'$ will be unstable toward autodetachment. Again, this process is possible because electronic excitations take place on a timescale short with respect to the motion of nuclei (Franck–Condon principle).

12.4.4 Ternary Collisions Between Electrons and Molecules

Three-body or ternary collisions are the most efficient in producing negative ions in a dense gas or plasma at low electron energies. These occur as follows:



The first process is of much greater practical importance, because the second occurs only in systems with high electron densities.

12.4.5 Volume-Production Heavy-Negative-Ion Sources

While all of the previously discussed radiative and collisional electron capture processes may take place in certain types of negative-ion sources at some rate, radiative and dielectronic capture are known to be low-probability processes and, therefore, are not practical mechanisms for negative-ion formation. In most sources that utilize charged particles for production (electrons or ions) photopolar dissociation processes are infrequent because of the low-photon-flux environment. In plasma-type negative-ion sources, electron impact polar dissociation, dissociative attachment and three-body collisional transfer processes dominate. Sources that employ electron impact attachment methods to form negative ions directly in a plasma medium are often called direct-extraction sources. Almost any positive plasma discharge source can be used to generate negative-ion beams by simply operating the source in reverse polarity. It is now well understood that the addition of Cs or other group IA elements into plasma discharges enhances negative-ion formation. In general, plasma sources have very good beam qualities, particularly those that rely on dissociative electron attachment.

Duoplasmatron-Type Heavy Negative-Ion Sources

Moak et al. [43] first discovered that useful intensities of negative ions could be directly extracted from a duoplasmatron when operated in reverse polarity. Later it was discovered that negative ions were more abundant in the periphery of the plasma and that the yields could be enhanced by offsetting the extraction electrode with respect to the center of the plasma [44]. Other negative direct-extraction versions of the duoplasmatron source have been described in the literature, including the duodecatron [45] and triplasmatron [46]. The duodecatron has produced $>100\mu\text{A}$ of H^- and $>10\mu\text{A}$ of O^- , F^- and Cl^- , while the triplasmatron has produced $>50\mu\text{A}$ of O^- .

Diode Heavy-Negative-Ion Sources

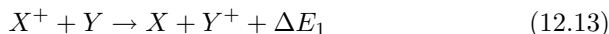
A diode source [47] has been used to generate several negative-ion species, including $600\mu\text{A}$ H^- , $20\mu\text{A}$ BO^- , $0.5\mu\text{A}$ C^- , $10\mu\text{A}$ CN^- , $4\mu\text{A}$ O^- , $50\mu\text{A}$ F^- , $4\mu\text{A}$ P^- and $4\mu\text{A}$ S^- .

Penning Discharge Heavy-Negative-Ion Sources

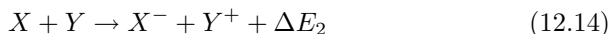
A radial-geometry cold-cathode Penning discharge source [48] has been utilized to generate a variety of DC negative-ion beams, including $60\mu\text{A}$ H^- , $1.2\mu\text{A}$ Li^- , $0.2\mu\text{A}$ BeH^- , $1.0\mu\text{A}$ MgH^- , $100\mu\text{A}$ F^- , $10\mu\text{A}$ B^- , $50\mu\text{A}$ S^- and $50\mu\text{A}$ Cl^- . The materials to be ionized are fed into the plasma discharge as gases, vaporized from an oven or sublimed from a solid rod of the material. This source type has also been used to generate $2.7\mu\text{A}$ C^- , $6.5\mu\text{A}$ Cu^- and $4\mu\text{A}$ Ni^- by sputtering solid rods made of the material of interest submerged in the plasma discharge [49].

12.5 Negative-Ion Formation Through Charge Exchange

During collisions between an ion neutral atom or molecule and another ion, neutral atom or molecule, an electron can be transferred from one of the colliding partners to the other with high probability, depending on the collision energy, ionization potentials and electron affinities of the colliding partners. A single charge transfer process is often referred to as “charge exchange” or “electron capture”. The charge transfer process can be categorized according to the energy defect ΔE involved in the transfer, as defined by the following interaction of the projectile X and target Y :



where ΔE_1 is just equal to the difference in ionization energies E_i of the colliding partners. For negative-ion formation, the energetic neutral atom X must undergo a second collision, as represented by the following reaction:

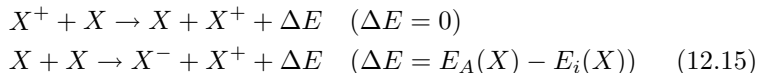


where ΔE_2 is the difference between the electron affinity $E_A(X)$ of X and the ionization energy $E_i(Y)$ of Y .

In general, the charge-transfer collision takes place at relatively large impact parameters, in which the projectile scattering angle is small and the product ion is scattered nearly perpendicular to the impact momentum vector. From the standpoint of beam quality degradation, low-momentum-transfer processes are clearly desirable. Such transfer processes can be cast into two distinct categories: (1) symmetrical (resonance) processes, and (2) asymmetrical (nonresonance) processes. In the first category, the projectile and target are the same species, while in the latter, practically more important category, the projectile and target are different.

12.5.1 Symmetrical (Resonance) Charge Exchange

In the symmetrical charge transfer process, in which the projectile and target are identical, the energy defect is equal to the difference between the ionization energies and is thus zero. Hence the cross sections can be very large. Negative-ion formation requires the following sequential projectile–target reactions:



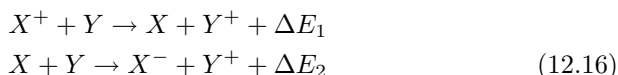
As noted, the energy defect for the second reaction is equal to the difference between the electron affinity of the first projectile atom E_A and the ionization energy E_i of the stationary target atom. Thus, a symmetrical process may not necessarily be commensurate with efficient negative-ion formation because of the possibility of a high ionization energy E_i for the donor species. However, consideration may be given to a two-step process in which the first process is symmetrical, and for which the cross section is large for neutral-energetic-atom formation, followed by an asymmetrical nonresonance process in which the energy defect is much smaller. Obviously, the energy defect can be minimized by selection of a charge transfer medium which has a low first ionization energy. Cesium has the lowest ionization energy of the naturally occurring group IA elements and is easy to volatilize. However, because of its high atomic number and mass, scattering is exacerbated relative to lower-mass members of this group such as Li. Li, however, is more difficult to use because of the relatively high temperatures required to volatilize the material.

12.5.2 Asymmetrical (Nonresonant) Charge Exchange

The formation of negative-ion beams by a two-step symmetrical/asymmetrical process would be impractical for general use, since it would involve the use of

two different charge exchange media, the first identical to that of the projectile and the second chosen because of the ease of transferring electrons to the energetic projectile X . The most general application of the charge exchange formation technique is for interactions between unlike ions. These processes occur with high probability between projectiles interacting with exchange vapors that possess a low energy defect. The probability for charge transfer from the electron donor (exchange vapor) to the projectile ion is sensitively dependent on the speed of the projectile.

The atomic charge transfer process occurs between unlike ions and differs from the symmetric resonance process in that it involves an electronic transition that requires a change in the internal energy of the system, or an energy defect ΔE . The energy defects ΔE_1 and ΔE_2 involved in the asymmetrical charge exchange formation of negative ions can be symbolically expressed through the following reactions:



ΔE_1 can be equated to the difference between the ionization energies E_i of the interacting atoms or molecules as

$$\Delta E_1 = E'(X^+) - E(X) + E(Y) - E(Y^+) = E_i(X) - E_i(Y) \quad (12.17)$$

while the energy defect E_2 is given by

$$\Delta E^2 = E(X) - E(X^-) + E(Y) - E(Y^+) = E_A(X) - E_i(Y) \quad (12.18)$$

Two-electron capture during a single collision is much less likely owing to the very high energy defect.

12.5.3 The Massey Adiabaticity Criterion

The adiabatic criterion proposed by Massey is of practical importance in asymmetrical charge transfer collisions [50]. At low projectile impact energies where the relative motion of the atoms is slow enough that the electronic motion can adiabatically adjust to small changes in the internuclear distance, the electron transfer process becomes unlikely. However, if the impact energy falls outside this “adiabatic region” and the electronic transition time is comparable to the collision time, the probability for electron transfer can be very high. The time of collision is taken as a/v , where v is the impact speed and a , the “adiabatic parameter”, is of the same order as the atomic dimensions within which the charge transfer transition becomes likely. The characteristic time for the electronic transition is given by $\hbar/\Delta E$, where ΔE is the energy defect. Thus, the condition $v \ll (a|\Delta E|)/\hbar$ characterizes the adiabatic speed region, and $v \cong (a|\Delta E|)/\hbar$ characterizes the speed region for which the maximum in the charge transfer process occurs.

The projectile energy E_{1MAX} for which the cross section reaches its maximum for a given reaction is given by

$$E_{1MAX} = F\{M_1/2\}(a \Delta E/h)^2 \quad (12.19)$$

where F is a constant, adjusted to bring the equation into better agreement with experimental measurement. Since two energy defects ΔE are involved in the negative-ion formation process involving a single exchange medium, the Massey adiabatic-maximum rule represented by (12.19) cannot simultaneously meet both optimum values. Consequently, the optimization process is ultimately a compromise between the two optimum values. However, in testing (12.19), it was found that best agreement with experiment came whenever the last reaction was taken as the dominant process. In practical units, (12.19) becomes

$$E_{1MAX}[\text{keV}] = 8.31 \times 10^{-3} M_1[\text{amu}][a(\text{\AA})\Delta E_2(\text{eV})]^2 \quad (12.20)$$

The criterion has been applied by Hasted to many charge transfer cross sections for reactions to estimate projectile energies that maximize negative-ion generation rates (maximum cross sections) [51]. Equation (12.20) has been applied by the present author to equilibrium fraction versus projectile speed data for $^{23}\text{Na}^+$, $^{27}\text{Al}^+$, $^{31}\text{P}^+$, $^{69}\text{Ga}^+$, $^{74}\text{Ge}^+$, $^{75}\text{As}^+$, $^{116}\text{Sn}^+$ and $^{208}\text{Pb}^+$ passing through ^{133}Cs vapor under the assumption that the criterion holds as an equality at the respective energy. The results of the latter calculations are given in Table 12.2. As noted, reasonable agreement is found for most of the species. The “adiabatic-maximum rule” is taken as a practical method for approximating optimum projectile energies for producing negative ions through charge exchange.

Table 12.2. Measured energy maxima and values calculated using (12.20) for charge exchange generation of selected negative-ion beams in Cs vapor

Projectile	a	$\Delta E_2(\text{eV})$	$E_{MAX}(\text{keV})$ (measured)	$E_{MAX}(\text{keV})$ (calculated)
Na^+	2.29	3.342	~ 14.5	11.2
Al^+	2.05	3.449	~ 20.2	11.2
P^+	1.97	3.144	~ 5	9.9
Ga^+	2.04	3.59	~ 41.5	30.8
Ge^+	2.02	2.657	~ 15.5	17.7
As^+	2.03	3.08	~ 15.5	24.4
Sn^+	2.15	2.778	~ 24	34.4
Pb^+	2.21	3.526	~ 35.5	105

12.5.4 Charge Exchange Cell Design

Since operational lifetime is of primary concern, special attention is given to cell design. Cell designs have increased in efficiency and lifetime in recent years, primarily because of the advent of sophisticated thermal simulation codes for accurately modeling thermal gradients in these devices. By accurately modeling thermal gradients in relation to the physical design of the cell and the materials of construction, the operational lifetimes of charge exchange cells can be very long. A cross-sectional view of a long-lifetime Li, Na, Mg or Ca cell, designed for use in producing radioactive ion beams (RIBs), is displayed in Fig. 12.3 [52]. By maintaining temperature gradients along the beam entrance and exit nozzles, the cell design ensures that the donor vapor condenses, liquefies and drains back into the reservoir along the slope of each nozzle. As noted, the cell is equipped with pneumatically actuatable Faraday cups, located at the entrance and exit of the cell to aid in characterization of the cell (measuring transmission and charge exchange efficiencies of species of interest). The cell uses heaters clamped against the body of the stainless steel cell for heating the elemental charge exchange material to the temperature that optimizes the vapor density in the path of the beam. A temperature of $\sim 450^\circ\text{C}$ produces a near-optimum target thickness ($\sim 1 \times 10^{15} \text{ cm}^{-3}$).

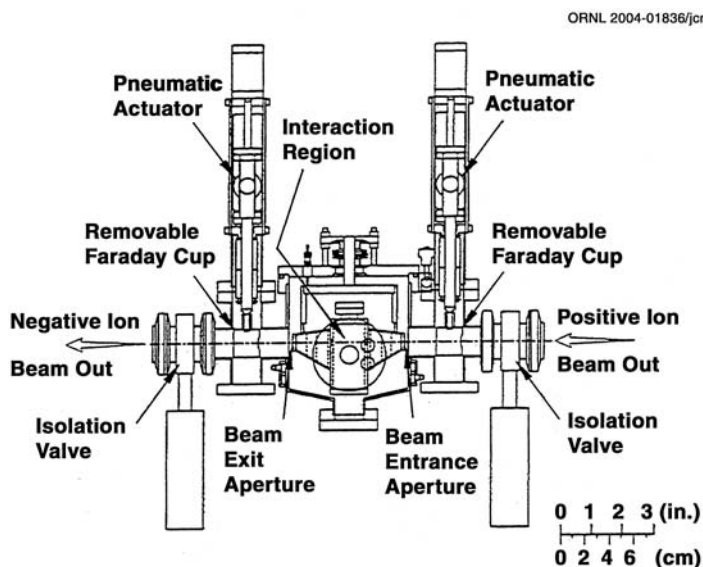


Fig. 12.3. A cross-sectional view of a long-lifetime Li, Na, Mg or Ca cell, designed for use in producing radioactive ion beams (RIBs) [52]

12.5.5 Charge Exchange Sources

The positive-ion source may be a simple gas source such as a duoplasmatron or RF source or a more universal source capable of producing positive-ion beams of volatile compounds. A schematic illustration of a duoplasmatron close-coupled to a Ca charge exchange cell is illustrated in Fig. 12.4.

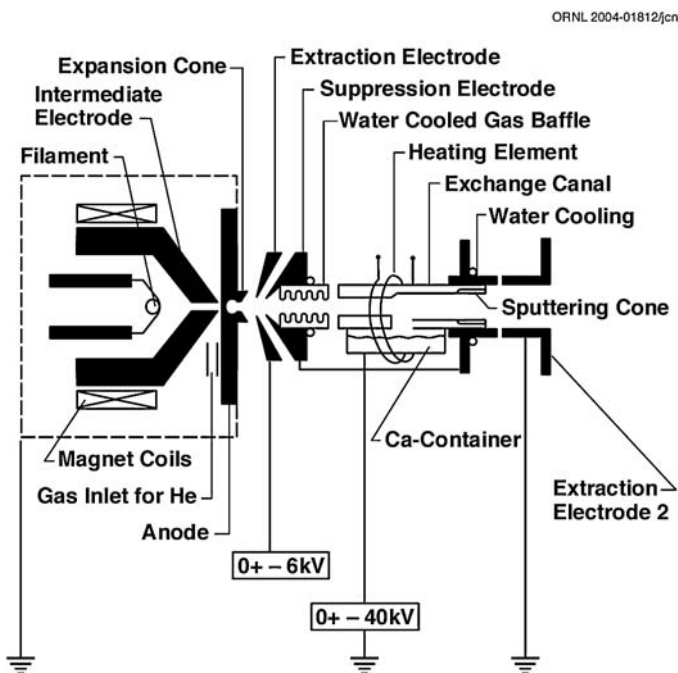


Fig. 12.4. Schematic representation of a close-coupled duoplasmatron positive-ion source and Ca vapor cell charge exchange system

12.5.6 Beam Quality Degradation Effects

In practice, consideration must be given to ease of loading, handling and ease of volatility of the donor material when designing a charge-exchange cell. In this regard, Rb or Cs is well suited for use. However, the charge exchange process can induce energy spreads of several hundred eV into beams. By its collisional nature, the charge exchange process degrades the quality of beams. For example, the mass of the primary beam in relation to the mass of the donor material is important. Ideally, a light donor material is best suited for use, since the maximum energy transfer E_{2max} is given by

$$E_{2max} = 4M_1M_2E_1/(M_1 + M_2)^2 \quad (12.21)$$

M_1 represents the mass of the primary beam of energy E_1 , and M_2 the mass of the donor material. Maximum energy transfer occurs whenever the masses of the primary beam and donor material are the same. The energy and angular spread resulting from passage of C^+ , Na^+ , Al^+ , P^+ , Ga^+ , Ge^+ , As^+ , Sn^+ and Pb^+ projectiles at the optimum energies for charge conversion in $1.1 \times 10^{15} \text{ cm}^{-3}$ Li, Na, K, Rb and Cs charge exchange vapors are displayed in Tables 12.3 and 12.4, respectively.

Table 12.3. Energy spread (in keV) induced in various projectiles during charge exchange conversion from positive to negative ions in Li, Na, K, Rb and Cs vapor. Target density $1.1 \times 10^{15} \text{ cm}^{-3}$ and length 10 cm

Projectile	Energy (keV)	Li	Na	Mg	K	Ca	Rb	Cs
C^+	20	0.015	0.093	0.095	0.18	0.18	0.45	0.68
Na^+	14.5	0.011	0.069	0.070	0.13	0.13	0.30	0.37
Al^+	20.2	0.012	0.070	0.075	0.14	0.14	0.33	0.43
P^+	5	0.014	0.068	0.068	0.10	0.11	0.23	0.28
Ga^+	41.5	0.018	0.094	0.090	0.16	0.16	0.34	0.55
Ge^+	15.5	0.029	0.10	0.12	0.19	0.18	0.34	0.45
As^+	15.5	0.027	0.11	0.12	0.19	0.18	0.33	0.43
Sn^+	24	0.044	0.17	0.18	0.29	0.27	0.42	0.49
Pb^+	35.5	0.090	0.34	0.38	0.47	0.51	0.76	0.72

Table 12.4. Angular spread (in mrad) induced in various projectiles during charge exchange conversion from positive to negative ions in Li, Na, K, Rb and Cs vapor. Target density $1.1 \times 10^{15} \text{ cm}^{-3}$ and length 10 cm

Projectile	Energy (keV)	Li	Na	Mg	K	Ca	Rb	Cs
C^+	20	5.0	19	19	27	27	59	78
Na^+	14.5	9.8	45	43	64	63	100	120
Al^+	20.2	8.5	40	35	55	54	94	110
P^+	5	39	130	130	190	200	320	410
Ga^+	41.5	8.6	28	28	45	41	73	86
Ge^+	15.5	22	71	72	110	110	160	220
As^+	15.5	22	77	73	110	110	170	210
Sn^+	24	19	58	57	85	85	150	180
Pb^+	35.5	16	50	53	76	82	130	170

12.5.7 Negative-Ion Equilibrium Fractions

The differential fraction of negative ions dF_i produced during passage through a vaporous target of thickness $d\pi$ can be expressed in terms of a set of first-order linear differential equations given by

$$dF/d\pi = \sum F_j \sigma_{ji} - \sum F_i \sigma_{ij} \quad (12.22)$$

Solution of the set of coupled equations requires the knowledge of $n(n-1)$ cross sections for a system involving n states.

Negative-ion conversion from an initially positive-ion beam interacting with a low-ionization-potential vapor can take place with high probability. The mechanism is therefore a very practical method for the production of negative-ion beams and has been utilized for this purpose for several years. As first shown experimentally by Donnally and Thoeming [53], the electron transfer process involves the transfer of a single electron in sequential collisions between the projectile and charge exchange atoms/molecules. As noted in (12.19) and (12.20), the production efficiencies depend primarily on the ion energy, the electron affinity of the element under consideration and the first ionization potential of the donor material. Several donor materials, including Li, Na, Mg, Ca, Rb and Cs, have been utilized as charge exchange media. The choice of medium affects the probability of ion formation as well as beam quality. Experimental schemes for measuring equilibrium fractions of negative ions have been described by several groups [54–56].

Equilibrium fractions for H in several group IA and group IIA vapors and for D in various group IA vapors have been measured [57]. The corresponding equilibrium fraction data for H are shown in Fig. 12.5. Many other investigations have been made of probabilities and energy dependences of charge transfer negative-ion formation, including H in Na [58] and Cs [59] vapors, and He in Li, Na and Mg [60], K [61], Rb [62], and Cs [63] vapors.

Extensive and systematic investigations of the charge exchange process for the production of many ion species have been carried out using Mg [55], Na [56] and Cs [64] vapors by researchers at the University of Aarhus. These investigations show efficiencies ranging from $\sim 0.5\%$ to $>90\%$. Mg is a more effective electron donor for high-electron-affinity elements, while Na is more effective for low-electron-affinity elements. Further evidence of the efficiency and universal character of the charge transfer process is found for group IA, IIIA, IVA and VA projectiles in cesium vapor in [64]. The corresponding equilibrium fraction dependences for these groups of elements on the projectile velocity in Cs vapor are displayed in Fig. 12.6. As noted, the maximum efficiency appears to occur at a rather well-defined velocity for each group of elements and does not vary much from group to group. Such information is very useful in selecting the most appropriate particle energy for optimizing the charge exchange efficiency when Cs is used as the donor material.

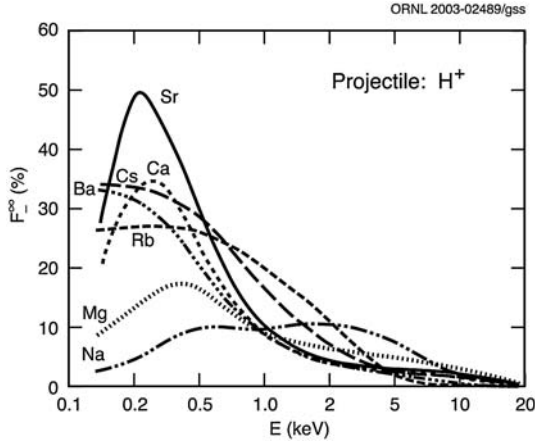


Fig. 12.5. Negative-ion equilibrium yields F_{-}^{∞} versus projectile energy E for H^{+} in several group IA (Na, Rb and Cs) and group IIA (Mg, Ca and Ba) vapors [57]

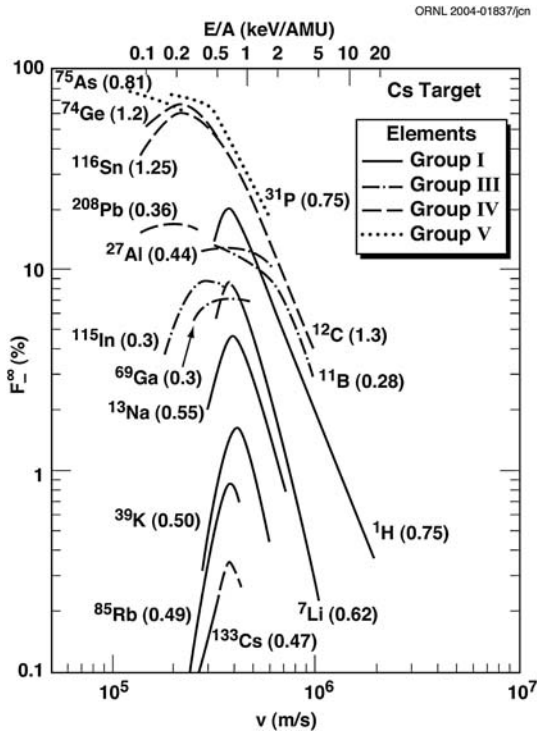


Fig. 12.6. Negative-ion equilibrium fractions F_{-}^{∞} versus projectile speed v for a few group IA, IIIA, IVA and VA elements [64]

12.6 Thermodynamic-Equilibrium Surface Ionization

Atoms or molecules impinging on a hot metal surface may be emitted as positive or negative ions in subsequent evaporation processes. The process of direct surface ionization is statistical in nature, and therefore thermodynamic arguments can be applied in deriving equations for the degree of positive- or negative-ion formation under equilibrium conditions. The subject has been reviewed in [65]. Experimental methods for negative-ion production by surface ionization have been reviewed by Kawano et al. [66–68].

12.6.1 Theory of Negative Surface Ionization

As discussed previously, the energy required for removing an ion from the surface is $\Delta H_i - (\phi - E_A) \geq \Delta H_a$. An ion supplied with energy ΔH_i may be transferred to the continuum in either ionic or atomic form where the respective potential-energy curves cross. The probability for arrival at a position far from the metal in a given state depends on the magnitude of $(\phi - E_A)$.

The probability P_i for negative-ion formation of a neutral particle of electron affinity E_A , evaporated from a hot surface with a low work function ϕ at temperature T , is given by the Langmuir–Saha relation

$$P_i = \frac{\omega_-}{\omega_0} \left(\frac{1 - r_-}{1 - r_0} \right) \exp \left(\frac{E_A - \phi}{kT} \right) \times \left[1 + \frac{\omega_-}{\omega_0} \left(\frac{1 - r_-}{1 - r_0} \right) \exp \left(\frac{E_A - \phi}{kT} \right) \right]^{-1} \quad (12.23)$$

where r_- and r_0 are the reflection coefficients of the particle at the surface and ω_- and ω_0 are statistical weights for the negative ion and neutral atom, respectively. ω_- and ω_0 are related to the total spin of the respective species and are given by $\omega = 2 \sum_i s_i + 1$, where s_i is the spin of the electron. Equation (12.23) describes an idealized situation in which there is perfect isotropy and no contamination of the ionizer surface. Moreover, the work function varies with crystalline orientation in cases where the metal is polycrystalline or the surface has uniformly or nonuniformly distributed surface contaminants. All of these effects can be taken into account by approximately summing over admixtures of existing work functions and statistical weighting factors in the respective expressions. From the relationship, it is evident that negative-ion yields could be enhanced by lowering the work function ϕ , or by increasing the surface temperature T for elements where $E_A \leq \phi$. In practice, ϕ varies with the crystalline orientation, and adsorption of highly electronegative atoms or molecules such as oxygen or the halogens raises the work function. Such contaminants can raise the work function and destabilize the formation process, and thus are the bases of the poisoning effects that will be discussed later. The computed efficiency (12.23) for negatively ionizing selected atoms and molecules striking a clean, hot LaB₆ surface is shown in Fig. 12.7.

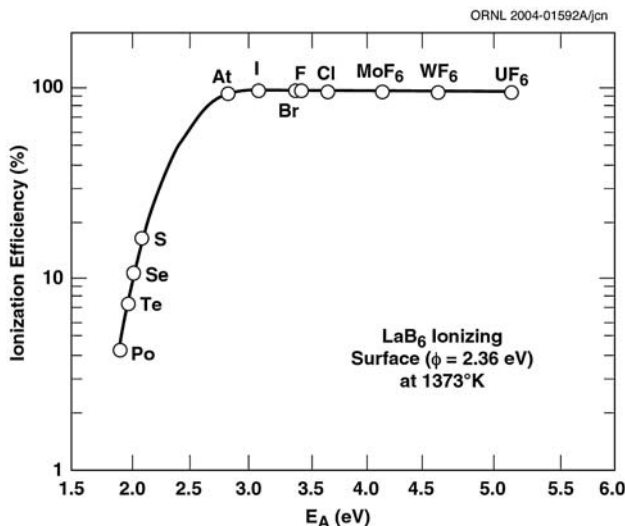


Fig. 12.7. Calculated ionization efficiency (12.23) versus electron affinity E_A of selected elemental and molecular species

12.6.2 Negative-Surface-Ionization Sources

Surface ionization can be used to great advantage for radioactive-ion-beam (RIB) applications to eliminate isobaric contaminants that may compromise experimental results with these beams, because of its highly chemically selective character. Ion sources based on the surface ionization principle are generally characterized by a high degree of ion beam purity (chemical selectivity), a limited range of species capability and excellent beam quality (low emittance). The energy spreads are typically of the order of thermal energies ($\sim 2kT \ll 1$ eV). The efficiency for negative-ion formation can be high or low, depending on the electron affinity of the species in relation to the work function of the ionizing surface. However, negative surface ionization has not been utilized frequently for generation of negative-ion beams – principally owing to the lack of chemically stable low-work-function materials for use as ionizers, in contradistinction to its positive-surface-ionization counterpart, where several high-work-function metals may be chosen for this purpose.

LaB₆ is usually used for negative surface ionization because of its relatively low work function ($\phi = 2.3$ to 3.2 eV) [69–73] and its availability, despite its widely publicized propensity for poisoning [74, 75]. The poisoning effect is attributable to the interaction of the hot LaB₆ with residual gases in the vacuum system, usually under high-flow-rate conditions or higher-than-optimum-pressure conditions. The effect raises the work function of the LaB₆ surface, thereby reducing the probability of ionizing electronegative atoms as they evaporate from the surface. Under high-flow-rate conditions, the poisoning process also affects the reliability of operation of sources equipped

with this material through-time-varying fluctuations of the ion beam intensity caused by variations in work function [75]. An increase in the work function causes an exponential diminution of the probability for negative-ion formation and, consequently, a reduction in the intensity of extracted negative-ion beams. Despite the poisoning problem, sources based on the use of LaB_6 ionizers have been described in the literature [75–78], including their use at ISOL facilities for negative-ion generation of high-electron-affinity radioactive species [77, 78]. A cross-sectional side view of the source described in [75], equipped with a spherical-sector LaB_6 ionizer, is displayed in Fig. 12.8. Ionization occurs whenever highly electronegative atoms/molecules fed into the source strike the spherical-sector ionizer. Negative ions so formed are accelerated by the converging optics of the negatively biased spherical sector extraction system and focused through the ion emission aperture.

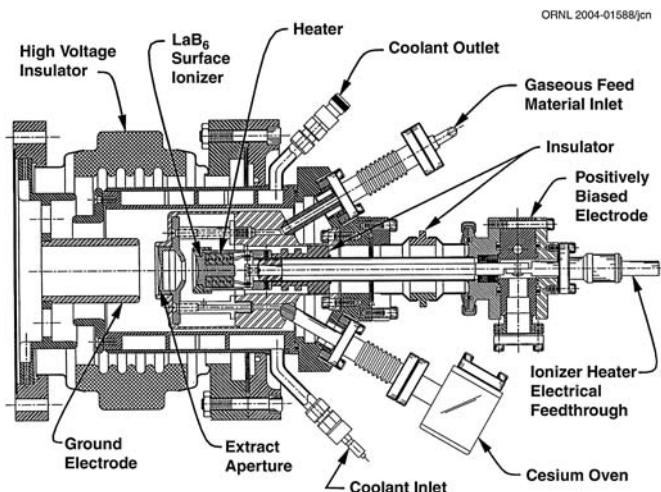


Fig. 12.8. The self-extracting negative-surface-ionization source described in [75], equipped with a spherical-sector LaB_6 ionizer

12.7 Secondary Negative-Ion Formation Processes: Nonthermodynamic-Equilibrium Surface Ionization

Since its discovery [79], the technique of sputtering a surface covered with a fractional layer of a highly electropositive adsorbate such as Cs has proved to be close to a universal method for generating atomic or molecular negative-ion beams from chemically active elements. Sputtering is generally caused by a cascade of momentum transfer processes between recoil atoms initially set in motion by an energetic projectile that leads to ejection of surface atoms.

The process is measured in terms of the sputter ratio, or yield per incident particle, S . In contradistinction to the process of conventional surface ionization, it is a nonthermodynamic equilibrium process and thus differs from thermal evaporation. For more details on the mechanisms involved in sputtering, and tabulated yield data for a variety of projectile–target interactions, early reports can be found in [80–83].

Although several independent and distinct negative-ion formation processes may coexist during sputtering, particularly from compound and alloyed samples, there is a preponderance of evidence that the mechanism of negative-ion formation during sputtering of “clean” metal surfaces with a low coverage of a highly electropositive adsorbate, such as one of the group IA elements, is a form of surface ionization. The practical implementation of the technique as a source of negative ions is quite simple, as illustrated schematically in Fig. 12.9. Positive-ion beams, formed either by direct surface ionization of a group IA element or in a Cs-rich noble-gas plasma discharge (Ar, Kr or Xe), are accelerated to between a few hundred eV and several keV, where they sputter a sample containing the element of interest. A fraction of the sputtered particles leave a negatively biased sample containing the species of interest as negative ions and are accelerated through an aperture in the source.

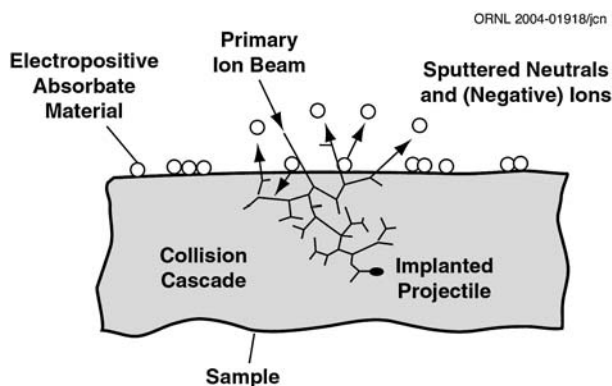


Fig. 12.9. Illustration of negative ion formation by energetic particle bombardment of a metal covered with a fractional layer of a highly electropositive adsorbate

12.7.1 Negative-Ion Sources Based on the Sputter Principle

Negative-ion source technology has steadily advanced over the years, in keeping with the continual demand for higher-intensity negative-ion beams with improved beam quality for a variety of tandem-electrostatic-accelerator-based fundamental and applied research. Heavy-negative-ion sources have been developed that utilize Cs surface ionization sources separated in space from the

sputter sample. In such sources, the energetic Cs^+ beams both sputter and lower the work function of the surface. The effect of the Cs on the work function varies from sample to sample because of differences in the saturation value for implanted Cs in the particular material surface. Thus it is difficult to achieve optimally low-work-function surfaces and consequently maximum negative-ion yields in this source type. This problem was subsequently overcome in the next-generation plasma-sputter and Cs-sputter negative-ion sources by directly feeding Cs vapor from an external oven at a controlled rate into a chamber containing a negatively biased sample probe and a means for producing positive-ion beams for bombarding samples containing the species of interest.

Plasma-sputter sources utilize either filaments or RF antennae for generating plasmas from which positive ions are extracted for sputtering the material of interest, while Cs-sputter negative-ion sources utilize hot surface ionizers for producing Cs^+ beams for this purpose. Sources based on this principle offer near-optimum conditions for generating negative-ion beams. These sources are said to be self-extracting in that negative-ion beams are accelerated from a negatively biased sputter probe within the plasma volume through an extraction aperture in the source. In addition to being versatile in terms of species, sources based on this concept are simple in design and easy to operate, and generally have long lifetimes. Because of these factors, heavy-negative-ion sources based on the sputter principle are utilized extensively in tandem electrostatic-accelerator laboratories.

Cs-Sputter Heavy-Negative-Ion Sources Equipped with Porous-W Surface Ionizers

The Müller and Hortig Negative-Ion Source

Through evolutionary processes, several types of heavy-negative-ion sources have been developed over the years since the discovery by Krohn that negative-ion yields are greatly enhanced by the presence of a thin layer of a highly electropositive alkali metal on the material being sputtered [79]. The first source to utilize the sputter principle was developed by Müller and Hortig [84]. The source utilized a continuously rotating wheel made of the material of interest, onto which was evaporated Cs metal vapor at a position diametrically opposed to the ion bombardment position. The sample wheel was bombarded with a 20 keV Ar^+ beam impinging at 20° with respect to the sample surface. Negative-ion beams were then extracted from the area of bombardment with an electrode system positioned perpendicular to the sample wheel. A simplified version of the source with improved performance characteristics was later developed [85]. A top view of the latter source is displayed in Fig. 12.10. Improvements to the source included the replacement of the Ar^+ source with a Cs surface ionization source of the porous-W type. This feature simplified source operation and eliminated the

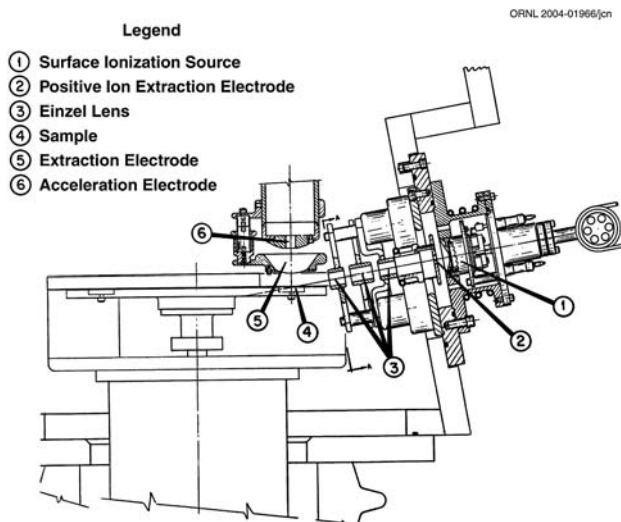


Fig. 12.10. A modified Müller and Hortig negative-ion source equipped with a Cs surface ion source and an asymmetric einzel lens for focusing Cs beams onto sputter samples [85]

necessity of using an external Cs oven. The Cs^+ beam served both to sputter the material and to lower the work function of the surface. In addition, the source incorporated an asymmetrical lens to focus 20 keV Cs^+ beams onto the rotating-wheel sample surface, thereby reducing the asymmetry of the beam spot while increasing the negative-ion beam intensity and reducing the emittance of extracted beams. Although these sources were cumbersome in design, they clearly demonstrated the viability of the technique as means of generating useful beam intensities of a wide variety of species for research, including $25\text{ }\mu\text{A C}^-$, $30\text{ }\mu\text{A O}^-$, $40\text{ }\mu\text{A F}^-$, $20\text{ }\mu\text{A C}_2^-$, $0.2\text{ }\mu\text{A Al}^-$, $44\text{ }\mu\text{A S}^-$, $100\text{ }\mu\text{A Cl}^-$, $3\text{ }\mu\text{A AlO}^-$, $1\text{ }\mu\text{A Cr}^-$, $6\text{ }\mu\text{A Cu}^-$, $0.8\text{ }\mu\text{A FeO}^-$, $14\text{ }\mu\text{A Ag}^-$, $26\text{ }\mu\text{A I}^-$, $5\text{ }\mu\text{A InO}^-$, $4.6\text{ }\mu\text{A TaO}_2^-$, $3\text{ }\mu\text{A Pt}^-$, and $0.5\text{ }\mu\text{A PbO}_2^-$.

The Middleton–Adams Source

The next generation of sources based on this technique was that of the Middleton and Adams Cs-sputter negative-ion source, illustrated schematically in Fig. 12.11 [86]. In this source, Cs^+ beams are produced by diffusing Cs vapor through a hot, porous-W ionizer (with a porosity such that $\rho = 0.8\rho_0$, where ρ_0 is the density of solid W.). Surface ionization sources generate space-charge-limited beams (several hundred μA up to a few mA) and therefore proper attention must be given to the electrode system design. The effects of space-charge on the sizes and angular distributions of 20 keV Cs^+ beams extracted from a surface ionization source with properly designed electrodes are illustrated in the simulations displayed in Fig. 12.12. The Cs^+ beams are

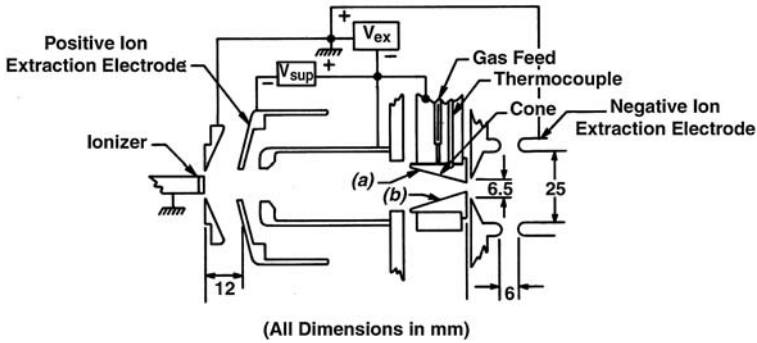


Fig. 12.11. The Middleton-Adams Cs-sputter negative-ion source [86]. The source is equipped with a Cs surface ionization source for producing a Cs^+ beam used to sputter conical-geometry samples

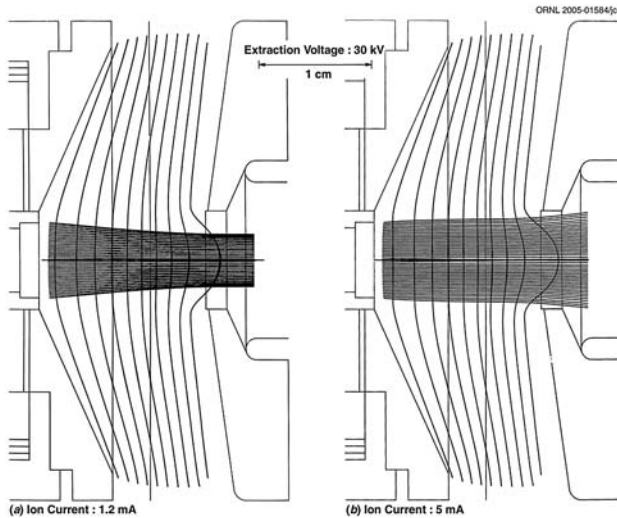


Fig. 12.12. Simulation of Cs^+ beam extraction from a Cs surface ionization source equipped with a porous-W ionizer, illustrating the influence of space-charge on beam trajectories during extraction. Surface ionization sources of this type are utilized in Middleton-Adams and refocus geometry Cs-sputter negative-ion sources. Extraction voltage $V_{ex} = 20 \text{ kV}$. Ion current (a) 1.2 mA; (b) 5 mA

accelerated to 20 keV and used to bombard the inner surface of a conical bore (half-angle 20°) in the material of interest. The Cs^+ beam serves to sputter the material of interest as well as to lower the work function of the sample surface. Negative ions are extracted at 20 keV through a 3 mm diameter hole bored into the apex of the sample. The ion optics of the negative-ion extraction region of this source are illustrated in Fig. 12.13. Because of the electrode

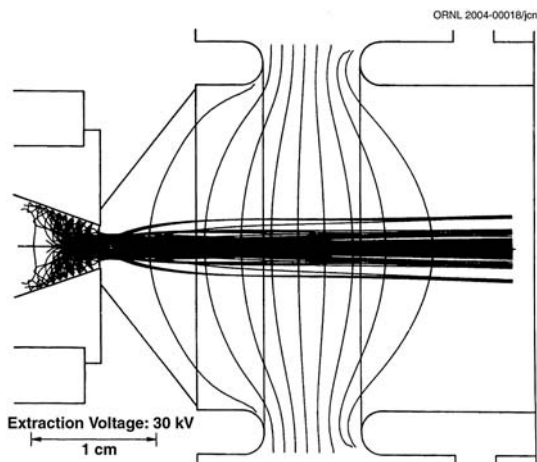


Fig. 12.13. Simulation of extraction from the Middleton–Adams or refocus-geometry Cs-sputter negative-ion source

configuration, the method of extraction and the intrinsically high energy distributions of the sputter process, negative-ion beams have high aberration coefficients and, consequently, have relatively large emittances. Because of the simplicity, reliability, versatility and long lifetime of this source type, it was quickly adopted for use in many tandem accelerator laboratories. The Middleton–Adams source was subsequently improved by introducing a lens–steerer combination between the Cs surface ionization source and the conical targets [87]. The modified source is often referred to as the refocus-geometry Cs-sputter negative-ion source.

The Inverted Cs-Sputter Negative-Ion Source

A Cs-sputter negative-ion source similar in principle to the Middleton–Adams source was developed by Chapman [88], with the exception that the sample and ionizer are reversed in position. A side view of the source is schematically displayed in Fig. 12.14. This source also utilizes samples with a conical taper in the region of ion generation that are attached to an indexable (rotatable) multisample wheel that can be loaded through an access port prior to source operation. The objective of this development was to improve the emittance and brightness of the original Middleton–Adams source while simultaneously solving the chronic ionizer erosion problems attributable to backstreaming of heavy negative ions, generated at the conical sample and accelerated back along the axis of the source to the ionizer. The Cs^+ beams are generated by diffusion through an annular, porous-W surface ionizer surrounding a central hole, through which sputter-generated negative ions pass during extraction. The species capabilities of the inverted source are similar to those of the original and refocus-geometry versions of the Middleton–Adams source. However,

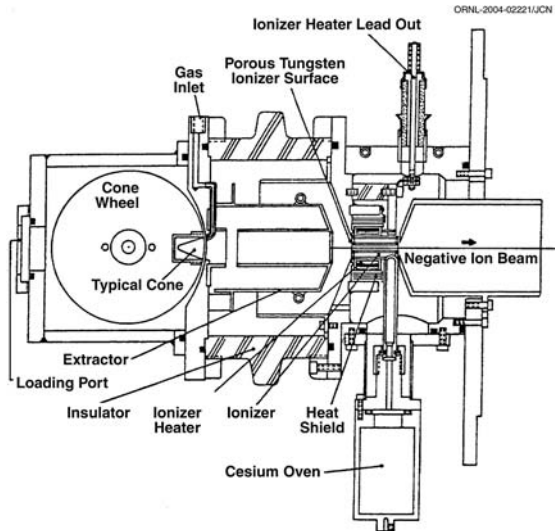


Fig. 12.14. Side view of the inverted Cs-sputter negative-ion source

because of the rather long extraction canal through the annular ionizer, extracted beam intensities from the inverted source are lower than those for the respective Middleton–Adams versions.

The ORNL KENIS Source for ISOL Radioactive-Ion-Beam Generation

Chemically active radioactive species are often released from target materials in a variety of molecular forms. For example, ^{17}F is principally released from an Al_2O_3 target material as Al^{17}F . Because of the low probability of simultaneously dissociating such molecular carriers and efficiently ionizing their atomic constituents with conventional hot-cathode, electron-impact ion sources, the species of interest are often distributed in several mass channels in the form of molecular sideband beams. Consequently, the beam intensities of the desired radioactive species are diluted. The sputter negative-ion beam generation technique is particularly effective for simultaneously dissociating molecular carriers and efficiently ionizing highly electronegative atomic constituents. Therefore, a new-concept kinetic-ejection negative-ion source (KENIS), based on this principle, was conceived to address this problem [89]. Because of geometric considerations associated with the scheme used at the Holifield Radioactive Ion Beam Facility (HRIBF) [90], the Middleton–Adams source geometry is best suited for this particular application. A three-dimensional representation of the source is displayed in Fig. 12.15, and a side view of the details of the ionization region of the source is shown in Fig. 12.16. Radioactive species, formed through fusion–evaporation nuclear reactions in a hot target irradiated with high-energy beams of either H^+ ,

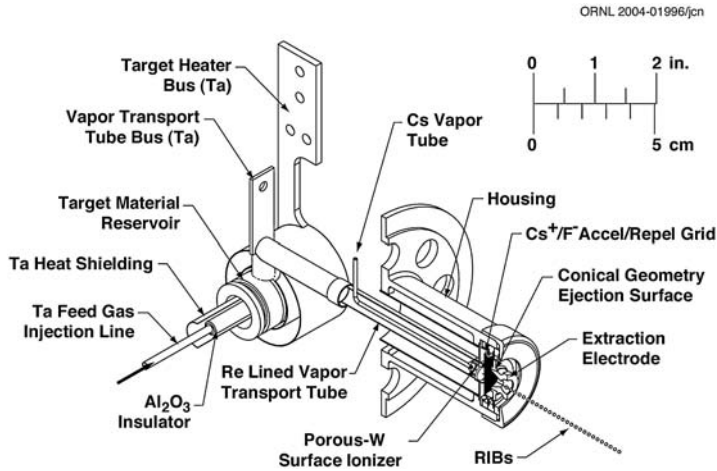


Fig. 12.15. Isometric cutaway drawing of the kinetic-ejection negative-ion source (KENIS) utilized for radioactive-ion-beam generation at the Holifield Radioactive Ion Beam Facility (HRIBF)

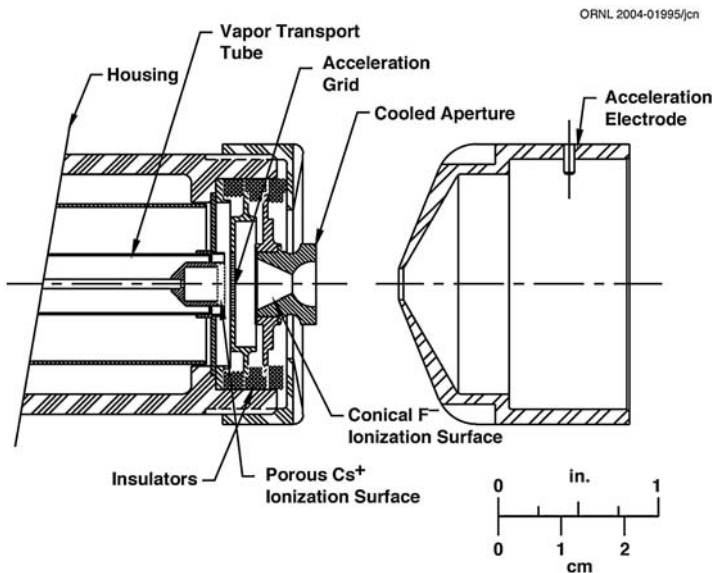


Fig. 12.16. Side view of the negative-ion generation region of the kinetic-ejection negative-ion source (KENIS) showing the porous-W Cs surface ionizer, the acceleration grid and the conical-geometry cathode where negative ions are formed

D^+ or ${}^3,4\text{He}^{++}$, are released through diffusion from thin-fiber target materials and transported to the ionization region of the source, where they are adsorbed onto a conical surface bored into a Ta cathode. Radioactive particles that collide with the cathode are bombarded with $\sim 200\text{ eV}$ Cs^+ beams produced with a porous-W ionizer. Negative ions are sputter-ejected and extracted through a 2 mm hole bored into the apex of the cone. The source is equipped with a negatively biased grid for repelling negative ions that are ejected rearward (i.e., toward the Cs surface ionization region of the source) and would be otherwise lost. The source has proven to be highly efficient for simultaneously dissociating Alf carrier molecules and forming atomic F. The source has been successfully employed on-line to generate high-intensity ${}^{17,18}\text{F}$ beams for astrophysics research at the HRIBF (using the ${}^{16}\text{O}(d, n){}^{17}\text{F}$ reaction).

Cs-Sputter Heavy-Negative-Ion Sources with External Cs Ovens

Early negative-ion source developments in the respective forms of the Müller–Hortig [84, 85], Middleton–Adams [86], refocus-geometry [87] and inverted-geometry [88] negative-ion sources clearly demonstrated their reliability, long lifetime and versatility for generating useful beam intensities of a wide variety of negative-ion species. However, the use of a surface ionization source to both generate Cs^+ beams for sputtering and simultaneously lower the work functions of sample surfaces is typically nonoptimum for the realization of maximum negative-ion beam intensities, particularly for high-sputter-ratio samples, and the extraction geometries were generally incompatible with the lowest practically achievable emittances. Following the successful development of the radial-geometry plasma-sputter source ANIS at the University of Aarhus, described below, a generation of axial-geometry sources evolved which borrowed certain features of the ANIS, including an external oven for feeding Cs vapor at a controlled rate into a chamber that houses a negatively biased probe containing the material of interest. These sources are self-extracting, and therefore the emission boundaries can be designed to ensure minimum losses during beam transport to the extraction aperture. Several sources of this type have been developed at a variety of laboratories, including those described in [91–93]. In these sources, the negatively biased sputter samples are placed in a cesium-rich environment, where they are sputtered with Cs^+ beams produced by Cs vapor in direct contact with a hot surface ionizer. The sources differ only in the geometry of the positive-ion surface ionizer, its spacing in relation to the negatively biased sample, the spacing of the sample in relation to the ion exit aperture, and the aperture size. The electrode geometry in this source type determines the optical properties of the cesium-ion beam. The perveance P of the system and the shape of the ionizer form an electrode system which determine the perveance for Cs^+ beams and therefore the magnitude of the space-charge-limited currents at a given extraction voltage. The shapes of the electrodes determine

the positive transport optics as well as the distribution of the current at impact with the sample. The positive-ion distribution, in turn, determines the shape and distribution of the negative-ion beams produced in a particular source. Since the particle speeds are lowest as they leave the sample surface, it is very important to contour the surface so that a high percentage of the sputter-ejected beams are transported back through the extraction aperture of the source. The codes described in [35–40] are extremely accurate and are valuable, cost-effective resources for quickly arriving at optimized electrode systems, as well as for simulating the beam transport optics of both positive- and negative-ion beams. Through their use, the geometry of the ionizer/negative-ion-generation electrode system can then be designed to produce the highest practically achievable negative-ion beam intensities (e.g. to optimize the perveance for Cs^+ beam generation and thereby maximize the negative-ion yields while reducing the emittance to the lowest achievable value for the source). Several single-sample axial-geometry sources, equipped with an external oven for feeding Cs vapor into a vacuum chamber housing a surface ionizer and a sputter cathode made of the material of interest, have been developed for use at a variety of laboratories, including sources equipped with cylindrical-geometry [93–98], spherical-geometry [98–104], ellipsoidal-geometry [103, 104] and conical-geometry ionizers [92, 105], among which are sources equipped with remotely indexable sputter samples [99–102, 105].

Source of Negative Ions by Cesium Sputtering (SNICs)

Following the successful radial-geometry plasma-sputter negative-ion source developments embodied in the ANIS, discussed below, a negative-ion source referred to as SNICS was developed at the University of Wisconsin [93]. This source was the first of a series of very successful axial-geometry Cs-sputter negative-ion sources that were equipped with an external oven for introducing Cs into the chamber and with means for producing Cs^+ beams for bombarding a negatively biased sputter probe containing the material of interest. A schematic representation of the source and the potential arrangement is shown in Fig. 12.17. The SNICS source utilizes a cylindrical-geometry, spiral-wound W filament for producing Cs^+ beams; these beams are accelerated to a negatively biased, concave, spherical-geometry cathode placed an optimum distance away from the ionizer. The concave spherical surface provides some focusing action on negative-ion beams during transport that helps to limit the angular excursion of extracted beams. Because of the spiral-wound filament and the cylindrical geometry of the ionization surfaces from which Cs^+ is accelerated, the wear patterns on cathodes are very complex. The observed wear pattern from this source is composed of two parts: a region of concentrated wear with a full diameter of ~ 1 mm, surrounded by a strongly graded halo region that extends out to the diameter of the cathode, depending on the position of the cathode relative to the ionizer. Thus, negative-ion beams extracted from the source have a complex distribution that affects

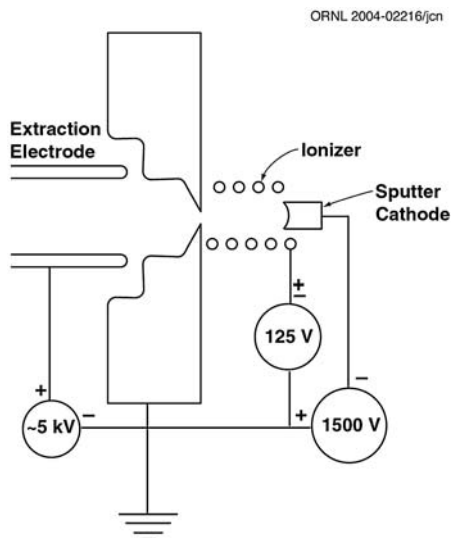


Fig. 12.17. Schematic representation of the Source of Negative Ions by Cesium Sputtering (SNICS), and potential arrangement

their emittance. During initial testing, the source was operated with Cu^- beams at intensities up to $38 \mu\text{A}$.

Source Equipped with a Conical-Geometry Ionizer

A cesium-sputter negative-ion source equipped with a conical-geometry ionizer [92,105] will serve to illustrate the principles of this source type. A cross-sectional side view of the single-sample version of the source is displayed in Fig. 12.18. The negative-ion yields are maximum at a specific Cs oven temperature, which typically is found to be independent of the species of interest. Once the oven temperature is determined, the value is usually fixed during operation of the source and never exceeded. The simplest and best method for control of beam intensity is by adjustment of the cathode sputter voltage to the desired value. Negative-ion-beam intensities versus sputter probe voltage also exhibit maxima that typically depend on the Cs oven temperature. It is also important to know the relation between Cs^+ current and sputter probe voltage, as the positive-ion current dictates the rate of sputtering. The perveance P of the electrode system for Cs^+ is $4.3 \times 10^{-9} \text{ A/V}^{3/2}$. Trajectories of Cs^+ ions flowing through the electrode structure under space-charge-limited conditions, computed with the code described in [38], are shown in Fig. 12.19. The computed ion current density resulting from positive-ion impact at the sample surface is distributed over a full diameter of 6 mm. This feature is desirable for potential use for generating radioactive ion beams, where good overlap between the condensed vapors and the cesium ion beam

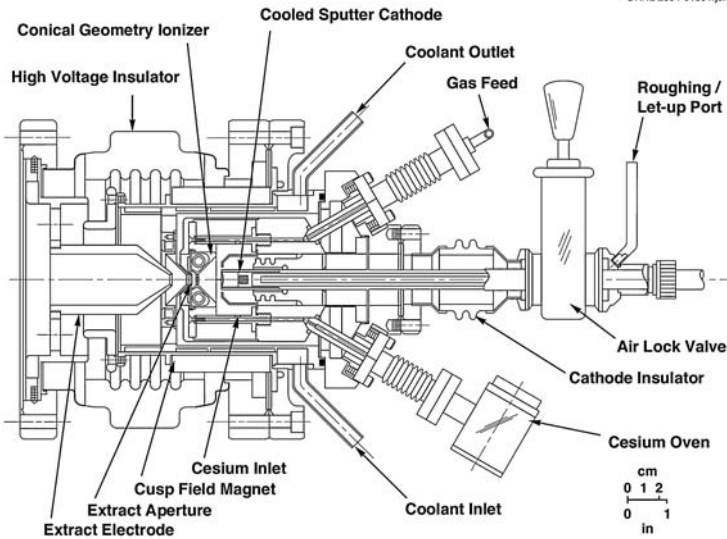


Fig. 12.18. Side view of the Cs-sputter negative-ion source equipped with a conical-geometry surface ionizer developed at the Holifield Radioactive Ion Beam Facility for stable negative-ion-beam generation

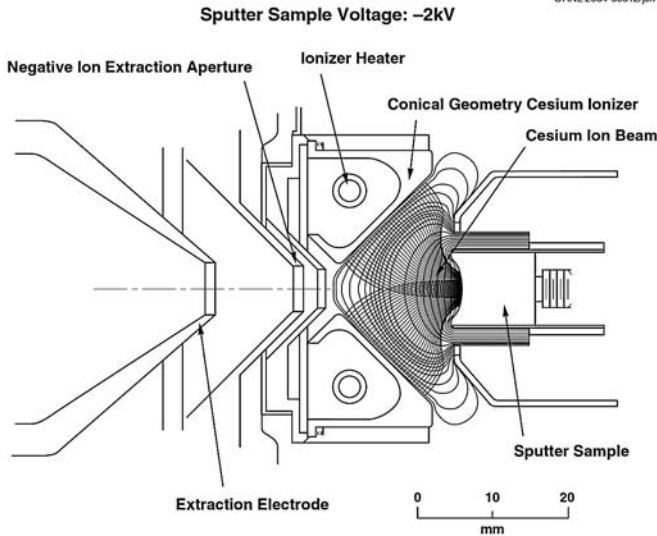


Fig. 12.19. Cs^+ beam optics of the Cs-sputter negative-ion source equipped with a conical-geometry surface ionizer

is required. An example of negative-ion trajectories flowing countercurrent to the positive-ion beam through the electrode system is shown in Fig. 12.20. These calculations were also performed with the code described in [38]. The negative-ion beam current density at the sample surface is assumed to mimic that of the positive-ion current density distribution. The influence of the positive-ion beam space charge on the negative-ion beam is included in the simulations.

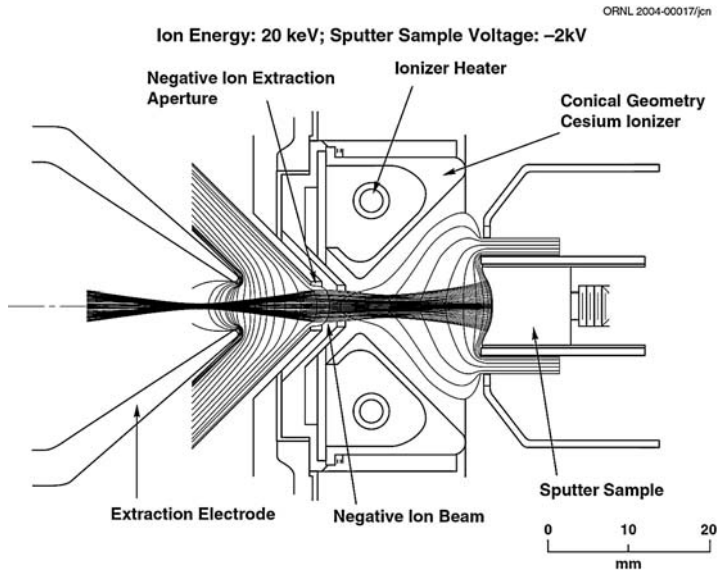


Fig. 12.20. Negative-ion beam optics of Cs-sputter negative-ion source equipped with a conical-geometry surface ionizer

Source Equipped with a Cylindrical-Geometry Ionizer

A source equipped with a solid-W cylindrical ionizer was developed at the Oak Ridge National Laboratory [95,96]. The ionization chamber and details of the ionizer and sputter probe electrode configuration for the source are shown in Fig. 12.21. The computed cesium current density distribution and observed sample wear pattern agree remarkably well for this source. The observed wear pattern is composed of two parts: a region of concentrated wear pattern with a full diameter of ~ 0.75 mm, surrounded by a low-density halo region that is uniformly worn over a diameter of ~ 4.5 mm. Thus, beams extracted from the source have two different characters: (1) a high-density beam located on axis, and (2) a uniformly distributed halo beam surrounding the high-density central region. Thus, both beams contribute to the emittance of the source with this ionizer geometry. Although the Cs^+ ion current versus

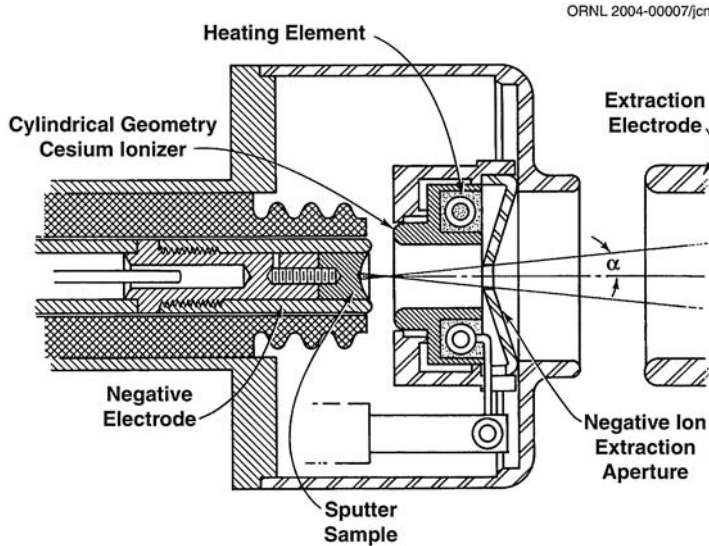


Fig. 12.21. Schematic drawing of the ion formation and extraction region of the Cs-sputter negative-ion source equipped with a smooth-surface, cylindrical-geometry W surface ionizer

sputter probe voltage for this electrode system has not been measured, according to calculations made with the code described in [35], the perveance for Cs^+ beams is very high ($P \cong 57 \times 10^{-9} \text{ A/V}^{3/2}$).

The Model 860 Cs-Sputter Negative-Ion Source Equipped with a Spiral-Wound Ta Ionizer

The type and geometry of the Cs surface ionizer and its position relative to the sputter cathode affect the positive-ion distribution at impact and, consequently, the negative-ion distribution, since the negative-ion distribution leaving the surface mimics that of the positive-ion beam. Hence, the quality of the negative-ion beam depends on the geometry of the ionizer/sputter probe electrode system. The wear pattern on the cathode determines to first order the distribution of the positive-ion beam in these sources. Codes such as those described in [35–40] predict with remarkable accuracy the observed wear patterns in these sources. A side view of the ionization region of the Model 860 source [97] is displayed in Fig. 12.22. The observed wear patterns on sputter cathodes mounted in a source equipped with a spiral-wound ionizer are found to be more complex than in those equipped with ionizers of other geometries. Typically, the central region of the cathode is found to be strongly worn, within a diameter of $\sim 1 \text{ mm}$ surrounded by a larger-diameter ($\phi \cong 8 \text{ mm}$), asymmetrically distributed halo region. The asymmetrical wear pattern within the halo region is attributable to the spiral character of the

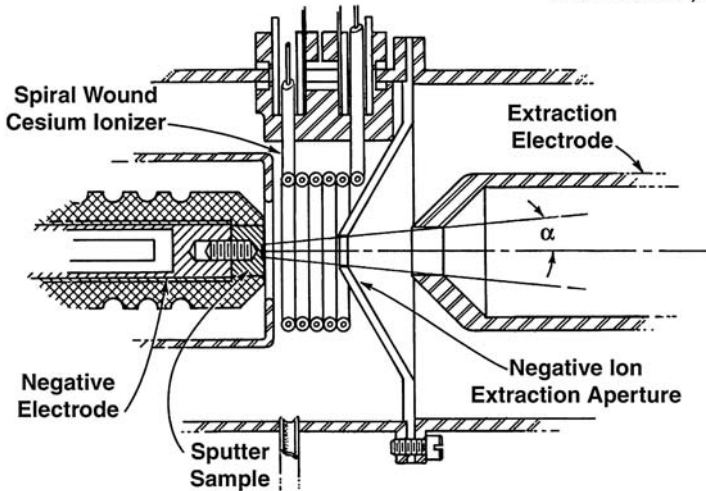


Fig. 12.22. The ion formation and extraction region of the Model 860 negative-ion source equipped with a spiral-wound Ta surface ionizer

Ta ionizer. The presence of the large halo beam, as is the case for the source equipped with a smooth-surface cylindrical-geometry ionizer [95, 96], contributes to the emittance of the source. The pervance of this electrode system has not been measured but is estimated to be $P \sim 60 \times 10^{-9} \text{ A/V}^{3/2}$.

Source Equipped with a Spherical-Geometry Ionizer

Sources have also been developed at the Oak Ridge National Laboratory that utilize spherical-geometry ionizers [98, 99]. The ionization chamber and ionizer sputter probe arrangement for a single-sample source equipped with a spherical-geometry ionizer [98] are illustrated schematically in Fig. 12.23. According to computational studies and experimental measurements, when the sputter probe is optimally positioned at the focal point of the ionizer, the positive-ion current density distribution at impact with the sample surface is $\sim 0.75 \text{ mm}$ in diameter. This particular ionizer does not exhibit a halo surrounding the central high-density distribution. Therefore the central region of the distribution serves as the sole source from which negative-ion beams are generated. According to computational analyses, with the code described in [35], the pervance for Cs^+ beams for this electrode geometry is $P \cong 2 \times 10^{-9} \text{ A/V}^{3/2}$.

Source Equipped with an Ellipsoidal-Geometry Ionizer

The ionization region of a single-sample source equipped with an ellipsoidal-geometry ionizer is shown in Fig. 12.24. The source has been briefly described in [103, 104]. When the sputter sample is placed at the focal point of the

ORNL 2004-00005/jcn

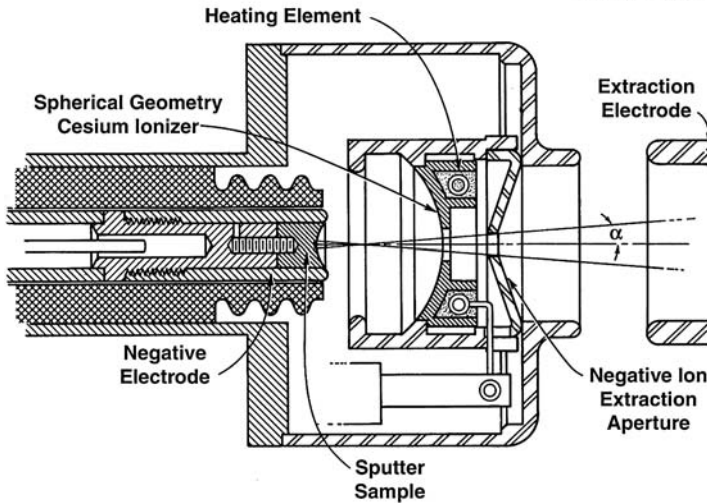


Fig. 12.23. The ion formation and extraction region of a Cs-sputter negative-ion source equipped with a spherical-sector-geometry surface ionizer

ORNL 2004-00006/jcn

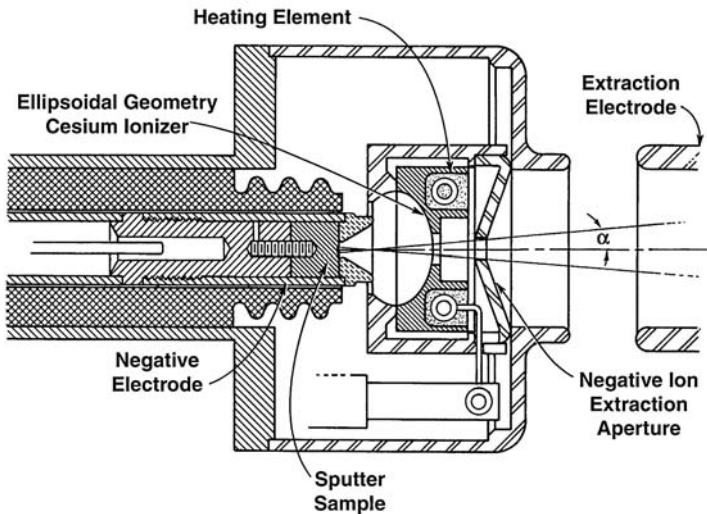


Fig. 12.24. The ion formation and extraction region of a Cs-sputter negative-ion source equipped with an ellipsoidal-geometry surface ionizer

system, the wear pattern at impact has $\phi \cong 1.25$ mm. This geometry has a high perveance in relation to other ionizer geometries with focusing attributes. According to calculations with the code described in [35], the source has a high perveance for Cs^+ beam formation of $P \sim 17 \times 10^{-9} \text{ A/V}^{3/2}$.

Multiple-Sample Sources

For applications where both high efficiency and/or high-frequency sample changes are desirable, as for accelerator mass spectrometry (AMS) applications, the ability to process multiple samples is essential. A few sources have been designed that meet this requirement, including the sources described in [99–102, 105]. A close-up of the multiple-sample source described in [105] is shown in Fig. 12.25. The source is equipped with a conical-geometry ionizer and a remotely controlled, eight-sample, wheel-type sample-indexing mechanism. The source described in [99] has a 60-sample, wheel-type computer-controlled indexing mechanism.

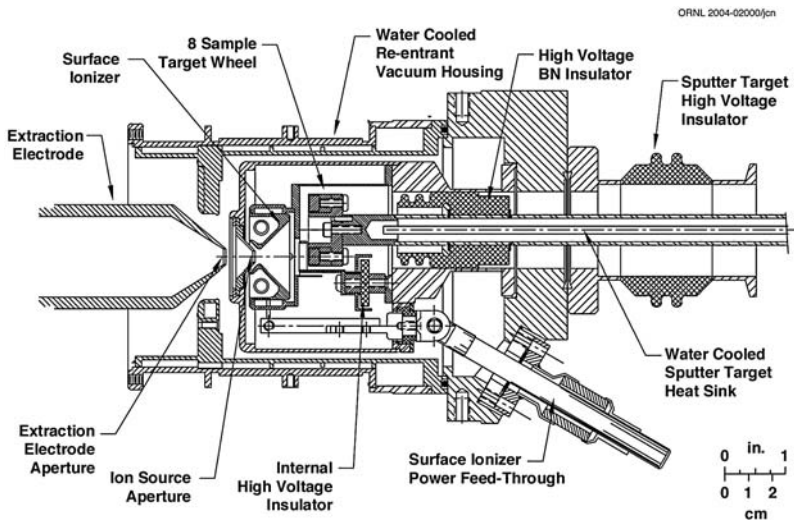


Fig. 12.25. Top-view cross section of the multisample Cs-sputter negative-ion source, equipped with a conical-geometry surface ionizer [105], that was developed for batch mode generation of radioactive ion beams at the Holifield Radioactive Ion Beam Facility

Negative-Ion-Beam Intensity Data

The negative-ion beam intensities that can be extracted from the sources previously described depend on a number of factors. The sputtering rate depends on the sample material, the magnitude of the cesium ion current and the cesium ion energy used to sputter the sample material, and these, in turn, depend on the source operational parameters, for example the cesium oven temperature and the sputter probe voltage. The space-charge-limited cesium current I^+ that can be accelerated at a given sputter probe voltage V and subsequently used for sputtering the sample depends on the perveance P of

the particular electrode configuration. The negative-ion current which can be extracted from the total current generated in the sputter process depends on the electron affinity of the species of interest, the work function of the surface, the size of the negative-ion generation region on the sample surface, the angular distribution of the negative-ion current at the ion-extraction aperture, the spacing of the sample in relation to the aperture, the aperture size and the sputter probe voltage V . Because of operational variables and the differences in the ionizer/sample electrode configuration, the negative-ion currents for a particular species will, in general, differ from source to source, and for a particular source, will vary from one operational period to another. Negative-ion yields for a particular species will depend on the chemical composition of the sample as well. The versatility of the sources described above is reflected by the wide spectrum of momentum-analyzed negative-ion beams that have been observed during their operation.

Atomic negative beams cannot be formed for elements with negative electron affinities, and elements with very low positive electron affinities often cannot be produced with intensities adequate for research. However, the chemically active elements can often be synthesized in situ by introducing appropriate chemically active gases into the ionization chamber of the source at a controlled rate. The gases react with the sputter cathode material to form molecular hydrides, carbides, oxides or halides with sufficiently high electron affinities to produce useful molecular beams containing the element of interest. The element of interest is released by accelerating the molecular beam to high energy in a tandem electrostatic accelerator, where the molecular carrier is dissociated and the atomic species are stripped to form high positive charge states in the terminal of the machine. Improved methods for producing beams of difficult elements such as the group IA elements have been reported in the literature. For example, Alton and Benjamin have developed an improved method for generating negative hydride beams of the group IIA elements [106] based on the use of composite cathodes, and Alton and Mills have developed a new method for generating beams of the group IA elements from metal carbonates [107]. Middleton and Alton have independently evaluated many solid and powder elemental, metal carbide, metal oxide, metal halide, metal carbonate, and composite-mixture cathodes, as well as techniques for in-situ syntheses of chemical compounds containing the species of interest during the sputter process. From these efforts, the best methods for forming useful beam intensities of most species have been developed. Middleton has tabulated the results of studies designed to determine the best cathode material for generating useful beams of almost every chemically active element [108].

Emittance Data

Emittance and brightness measurements for sources equipped with spherical-geometry, ellipsoidal-geometry and cylindrical-geometry ionizers, as well as

for the Model 860 source, have been reported previously [103,109,110]. The average normalized emittance versus the percentage of negative-ion-beam intensity contained within a given contour for these sources, as well as that for a source equipped with a conical-geometry ionizer [92,105], are displayed in Fig. 12.26. In these sources, the geometry of the ionizer/sputter sample system electrode determines the size and shape of the cesium ion beam at impact with the surface. For small sample sizes, a focusing-geometry electrode system may be desirable. As indicated in Fig. 12.26, the emittances of negative-ion beams extracted from sources equipped with electrode systems that focus the cesium beam on the sample surface are somewhat lower. The conical-geometry source has the lowest emittance, while that of the Model 860 is considerably larger than for those equipped with the other ionizer geometries, principally owing to the increased size of the exit aperture used in this source.

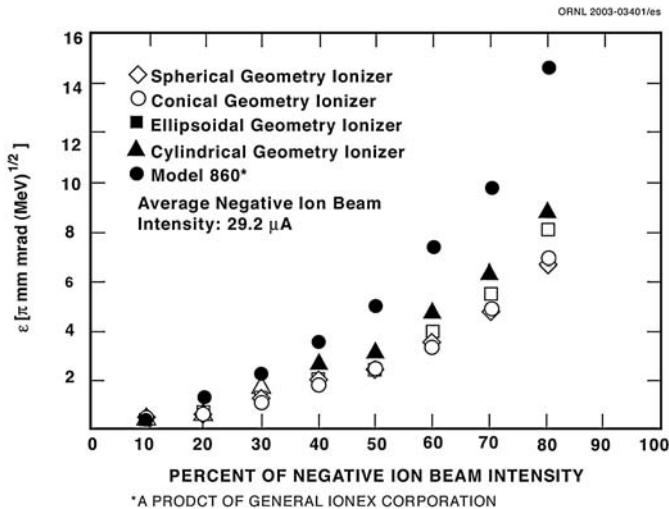


Fig. 12.26. Normalized emittance versus percentage of negative-ion-beam intensity for Cs-sputter negative-ion sources, equipped with cylindrical, spherical-sector, ellipsoidal and conical-geometry ionizers, developed at the Holifield Radioactive Ion Beam Facility, and for the Model 860 negative-ion source

Plasma-Sputter Heavy-Negative-Ion Sources

The advent of Cs-seeded plasma-sputter heavy-negative-ion sources has significantly advanced the state-of-the-art of heavy-negative-ion source technology. Several sources have been developed since the successful development of the radial-geometry University of Aarhus negative-ion source (ANIS)

[111,112], including those described in [113–115,120–126]. In this source type, negative ions, created in the sputter process, are accelerated across a double-layer plasma sheath that surrounds and exactly conforms to the shape of negatively biased spherical- or cylindrical-geometry sputter probes. In later developments, the multi-cusp-field plasma-confinement technique makes it possible to effectively confine the plasma while preserving uniform sputtering over variable-size spherical- or cylindrical-sector sputter probes, when operated in a high-density plasma mode, because of the low magnetic-field flux density in the central region of the plasma chamber. Thus, the emission area can be scaled to meet the intensity requirement of a particular experiment. The plasma sheath serves as a lens to focus the beam through the plasma to the ion exit aperture of the source. Since space-charge effects are precisely compensated during passage through the plasma, very high beam intensities can be extracted from this type of source. Thus, high beam intensities can often be realized while preserving a reasonable emittance value. The intensity levels for certain species from plasma-sputter negative-ion sources, such as those described in [120–126], are often higher by factors of 3–100 than those generated in conventional cesium-sputter negative-ion sources, and yet the emittances ε_N of the beams are reasonably small. In sources that use hot cathodes (filaments) to generate the plasma for DC operation, the source lifetime can be limited by sputter erosion of the filament. This problem can, in part, be offset by making provision for filament redundancy or by use of RF plasma generation techniques.

The tandem accelerator has also been either used or considered for use as an injector for synchrotron heavy-ion accelerators. The plasma-sputter negative-ion source is well suited for this application in that pulsed negative-ion beam intensities exceeding the practical value of $\sim 200\ \mu\text{A}$ (peak intensity) can be delivered to the synchrotron from the tandem electrostatic accelerator for a wide variety of heavy-ion species. It also offers the prospect of use for batch-mode generation of radioactive ion beams for injection into tandem electrostatic accelerators for postacceleration, because of the perfect overlap of the plasma particles that sputter the sample and the area of the sample irradiated by the production beam. In general, plasma-sputter negative-ion sources generate higher beam intensities with improved emittances than do their Cs-sputter counterparts. These sources all utilize hot cathodes for plasma ignition, with the exception of the sources described in [115,123,124], which use RF antennae for plasma ignition.

The University of Aarhus Negative-Ion Source (ANIS)

While Cs-sputter source designs that use porous-W surface ionization sources separated in space from the region of negative-ion generation are very versatile in terms of species, this method of Cs^+ beam formation rarely provides optimum cesium-layer surface coverage, critically important for generating maximum negative-ion yields. With this approach, the cesium surface

content is relegated to the saturation value for Cs in the particular material, and the residual neutral surface Cs is thought to be less than optimally distributed. (The saturation value is the amount of cesium left in the surface after the steady state has been achieved.) Since the saturation value varies inversely with sputter ratio, high-sputtering materials (Cu, Ag, Au, etc.) have a low residual cesium content and thus do not produce negative-ion yields in accordance with the magnitudes of their electron affinities. In addition, negative-ion beams have high aberration coefficients because of the shape of the negative-ion generation surface (the inner surface of a conical bore in the material of interest) and require a means for ion extraction from the surface of generation (through a hole bored into the apex of the cone). In 1975, Andersen and Tykesson of the University of Aarhus introduced a radial-geometry Cs-rich plasma-sputter negative-ion source with very high yields from elemental materials, exceeding considerably the yields of those sources that utilize Cs beams to both sputter and simultaneously lower the work functions of sample surfaces [111, 112]. A version of this source was designed and developed at the Oak Ridge National Laboratory (ORNL) [113], as displayed in Fig. 12.27. Neutral Cs is fed into the discharge chamber at a controlled rate from an external oven. Discharge support gases are metered through a leak valve into the source. A weak dipole magnetic field (~ 0.0150 T) is used to collimate a primary electron beam produced by thermionic emission from a Ta filament situated at the end of the plasma chamber. Plasmas are ignited by raising the filament to emission temperature and accelerating electrons to energies typically < 100 eV. The source is equipped with an external oven that permits precise control of the feed rate of Cs into the plasma discharge and thereby provides a means for optimizing the cesium layer thickness on samples undergoing bombardment, as required for maximizing negative-ion yields. Chemically reactive gases may also be introduced into the discharge,

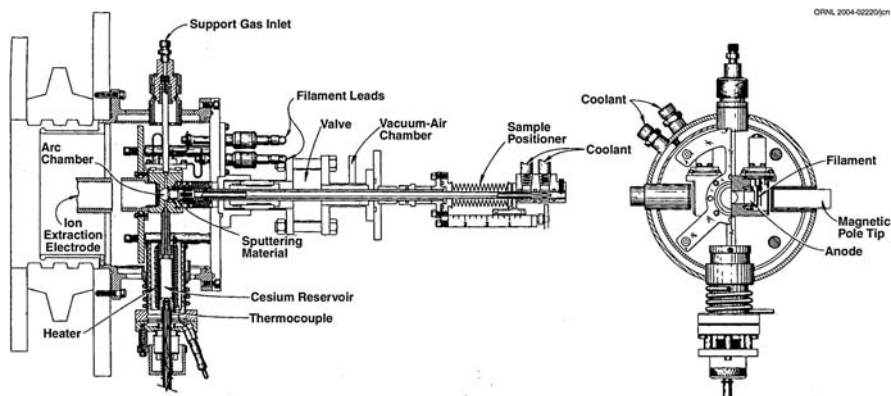


Fig. 12.27. Side and end views of the modified University of Aarhus negative-ion source (ANIS) developed at the Oak Ridge National Laboratory [113]

to react with the cathode material to form intense molecular carrier beams containing the species of interest in cases where the species has a negative electron affinity or the electron affinity is appreciably lower than that of the molecular carrier. Negatively biased samples have concave spherical emission areas that are placed away from the extraction aperture by a distance equivalent to the spherical radius of the sample so that the negative-ion beam is a minimum at the aperture. One of the great advantages of the plasma-sputter source is that the sample surfaces are uniformly bombarded by ions extracted from the plasma sheath, which precisely conforms to the shape of the emission surface. Negative ions are produced by sputtering with energetic positive ions across the plasma sheath that surrounds the spherical-geometry surface. Thus, the plasma sheath acts as a spherical lens that focuses ion beams through the extraction aperture and thus reduces the size of the extraction aperture relative to other emission geometries, thereby restricting losses of neutral Cs from the source. Since the sputter process ejects particles normal to the emission surface, beams created from a concave spherical-geometry surface have lower aberration coefficients than those for other emission geometries. This method of negative-ion generation enables variation of the sputter sample size (diameter and spherical radius) according to the intensity requirements of a particular experiment. The emittances of beams are found to depend on the sputter cathode diameter and the intensity of the species. For a cathode of diameter $\phi = 10$ mm, the normalized emittance of a $30\text{ }\mu\text{A}$ C^- beam at the 90% contour level is $\varepsilon_N \cong 1.8\pi$ mm mrad $\text{MeV}^{1/2}$ [112]. The beams reported by this group during early characterization studies of the ANIS include $1\text{ }\mu\text{A}$ Li^- , $0.8\text{ }\mu\text{A}$ BeH^- , $3\text{ }\mu\text{A}$ BeO^- , $0.04\text{ }\mu\text{A}$ B^- , $2.5\text{ }\mu\text{A}$ BO^- , $20\text{ }\mu\text{A}$ C^- , $15\text{ }\mu\text{A}$ C_2^- , $5\text{ }\mu\text{A}$ CN^- , $30\text{ }\mu\text{A}$ O^- , $15\text{ }\mu\text{A}$ F^- , $1\text{ }\mu\text{A}$ Al^- , $2\text{ }\mu\text{A}$ Al_2^- , $50\text{ }\mu\text{A}$ Cl^- , $0.9\text{ }\mu\text{A}$ TiH^- , $0.7\text{ }\mu\text{A}$ FeH^- , $6\text{ }\mu\text{A}$ Ni^- , $30\text{ }\mu\text{A}$ Cu^- , $3\text{ }\mu\text{A}$ Ta^- , and $80\text{ }\mu\text{A}$ Au^- . A modified form of the source has been extensively evaluated at the Oak Ridge National Laboratory [113,114].

The ORNL RF Plasma-Sputter Negative-Ion Source

Filament-ignited sources are limited by the lifetimes of filaments used for plasma ignition because of the erosive nature of the hot-cathode plasma discharge. The use of RF power for plasma ignition, in principle, overcomes this handicap, reduces the complexity of operation and lowers overall source maintenance. However, in practice, sources in which antennae are submerged in the plasma undergo bombardment and, consequently, their lifetimes are also limited by sputter erosion processes. This problem can be ameliorated by antennae coupling RF power through thin-walled glass or ceramic tubes that surround the plasma volume and thus protect the antenna. A compact source based on RF plasma ignition techniques, displayed schematically in Fig. 12.28, has been developed at the Oak Ridge National Laboratory [115]. The main vacuum chamber is equipped with permanent magnets positioned around the periphery of the chamber to form a longitudinal magnetic field

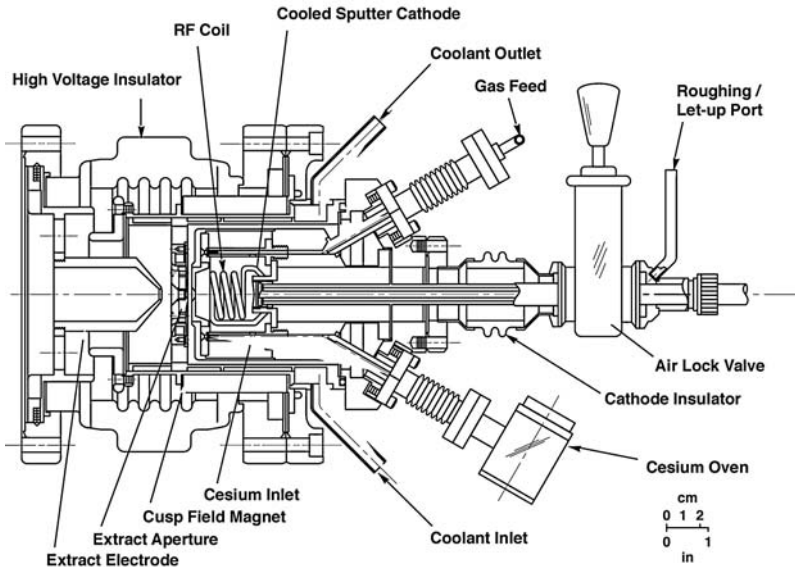


Fig. 12.28. Cross-sectional side view of an axial-geometry RF plasma-sputter negative-ion source [115]

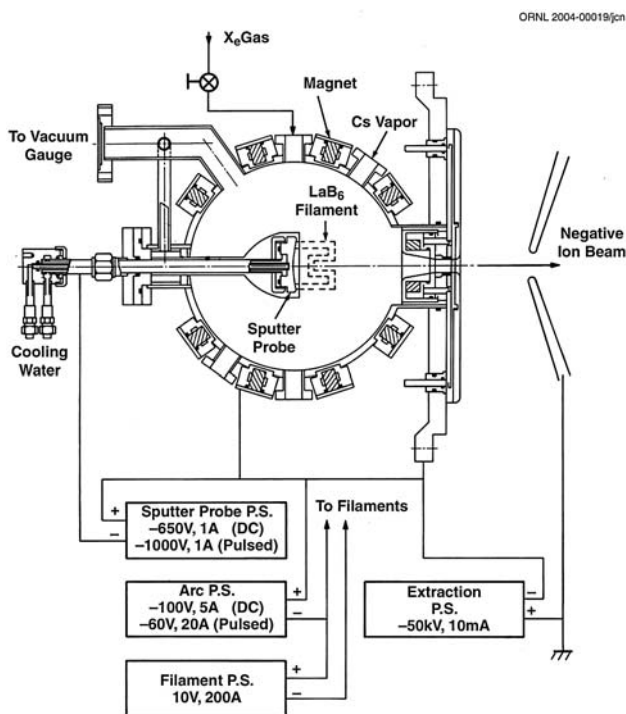
(maximum field strength 0.0380 T) for confinement of the plasma in the radial direction. The source utilizes a porcelain-coated, high- Q , self-igniting, inductively coupled antenna system, operating at 80 MHz, that has been optimized to generate Cs-seeded plasmas at low pressures (typically <100 Pa for Xe). The source can be operated in either pulsed or DC modes for the generation of negative-ion beams. Table 12.5 provides a partial list of momentum-analyzed beams produced by its use. The normalized emittance within the 80% contour for a $200\text{ }\mu\text{A Cu}^-$ beam is $\varepsilon_n \sim 7.5\pi\text{ mm mrad (MeV)}^{1/2}$. In general, the emittances of this source are lower than those for Cs-sputter sources at the same beam intensity. However, the intensities from the present source and other plasma-sputter negative-ion sources are often considerably higher for a given species.

The KEK Pulsed-Mode, High-Intensity, Plasma-Sputter Negative-Ion Source

The original multi-cusp magnetic-field plasma surface ion source, routinely employed for the production of high-intensity pulsed-mode H^- beams at the LAMPF [116,117] and at the National Laboratory for High Energy Research (KEK) [118,119], has been modified for use as a high-intensity pulsed-mode heavy-negative-ion source [120,121]. A side-view representation of the source is displayed in Fig. 12.29. The developments for heavy-ion production in the case of the KEK negative-ion source are described in [120,121]. These

Table 12.5. Negative-ion beam intensity data for the ORNL RF plasma-sputter negative-ion source

Ion	Probe Material	Beam Intensity (μA)
C^-	C	610
F^-	LiF	100
Si^-	Si	500
S^-	ZnS	500
P^-	GaP	125
Cl^-	NaCl	200
Ni^-	Ni	150
Cu^-	Cu	230
Ge^-	Ge	125
As^-	GaAs	100
Se^-	CdSe	200
Ag^-	Ag	70
Au^-	Au	250
Pt^-	Pt	125

**Fig. 12.29.** A cross-sectional view of the KEK pulsed-mode, high-intensity plasma-sputter negative-ion source [120, 121]

developments clearly demonstrated the capabilities of this source concept for generating high-intensity (several mA) pulsed negative-ion beams of a wide spectrum of atomic and molecular species. In the original source, the plasma was formed by pulsing the discharge voltage of two series-connected LaB₆ cathodes, thus creating a high-density plasma discharge from Xe feed gas, seeded with cesium vapor, during the discharge cycle. Under pulsed-mode operation at the low duty factors used (typically 2×10^{-3}), the LaB₆ cathodes exhibit very little erosion after many hours of operation. Although principally tested in a low-duty-cycle (repetition rate 1–50 Hz) macropulsed mode (pulse width 50–300 μs) suitable for heavy-ion synchrotron applications, the source can be operated in DC mode and generate heavy-negative-ion beams at mA intensity levels for many atomic and molecular species, including semiconducting-material dopants (B[−], P[−], As[−], Sb[−], etc.), as well as O[−] for isolation barrier formation. Since O[−] beams are generated from solid materials, chemical reactions between the feed gas and the hot cathodes commonly used in volume-discharge positive-ion sources, which lead to an extremely short cathode lifetime, are avoided. Table 12.6 provides a partial list of total beam intensities (peak), species and probe materials used in the KEK radial-geometry source. The normalized emittances, within the 80% contour,

Table 12.6. Negative-ion beam currents (peak) produced by the high-intensity plasma-sputter negative-ion source developed at KEK [120, 121]

Ion(%)	Probe Material	Probe Voltage (V)	Cathode Geometry	Current (mA)
C [−] (36)	C	937	Spherical	6.0
C ₂ [−] (58)	C	937	Spherical	6.0
O [−] (67)	Mo and O ₂	438	Spherical	30
Si [−] (75)	Si	937	Spherical	6.0
P [−] (44)	GaP	937	Flat	1.8
Co [−] (85)	Co	937	Spherical	6.0
Ni [−] (87)	Ni	438	Spherical	6.0
Cu [−] (77)	Cu	438	Spherical	8.2
Cu [−] (40)	CuO	438	Flat	4.5
O [−] (60)	CuO	438	Flat	4.5
As [−] (20)	GaAs	937	Flat	3.7
As ₂ [−] (52)	GaAs	937	Flat	3.7
Pd [−] (69)	Pd	937	Spherical	7.6
Ag [−] (91)	Ag	937	Spherical	6.2
Sn [−] (67)	Sn	937	Spherical	3.6
Pt [−] (71)	Pt	937	Spherical	8.1
Au [−] (73)	Au	437	Spherical	10.3
Bi [−] (6)	Bi	937	Spherical	2.7
O [−] (42)	Bi	937	Spherical	2.7

ORNL 2004-02215/jcn

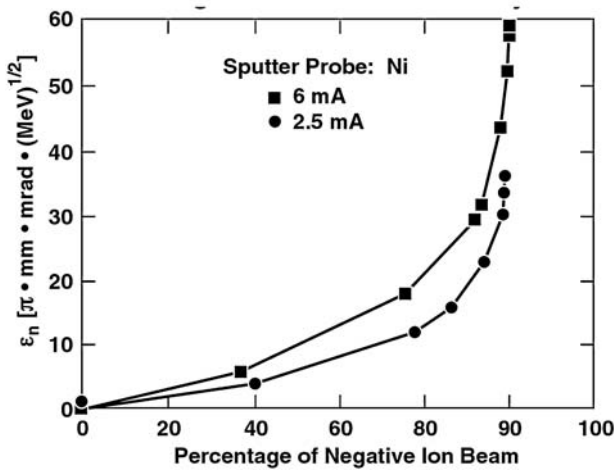


Fig. 12.30. Normalized emittances, within the 80% contour, for 2.5 mA and 6 mA Ni beams for the pulsed-mode high-intensity plasma-sputter negative-ion source [120, 121]

for 2.5 mA and 6 mA Ni beams are, respectively, $\sim 12\pi$ mm mrad (MeV)^{1/2} and 17π mm mrad (MeV)^{1/2}, as shown in Fig. 12.30. The increase in emittance is attributable to space-charge effects in the beam. An axial-geometry version of this source has also been designed for pulsed-mode operation for potential use in tandem accelerator/heavy-ion synchrotron applications at the Oak Ridge National Laboratory [122].

The Kyoto University RF Plasma-Sputter Negative-Ion Source

A DC-mode plasma-sputter negative-ion source using a 13.56 MHz RF plasma igniter has been developed at Kyoto University in Japan [123, 124]. The source, shown schematically in Fig. 12.31, differs from previously described sources of this type in that it is not equipped with provision for plasma confinement. The source generates a high-density plasma from Ar or Xe gas seeded with Cs vapor by application of RF power to a three-turn antenna. The spherical-sector cathode ($\phi = 43$ mm) is positioned 70 mm from the ion exit aperture ($\phi = 13$ mm). Beam intensities of 6.5 mA of Cu⁻, 1.6 mA of C⁻, 2.3 mA of C₂⁻, 3.8 mA of Si⁻, 0.03 mA of B⁻ and 1.0 mA of B₂⁻ have been generated in the source. The efficiencies for forming several negative ions in the source are given in Table 12.7. The emittances of beams extracted from the source have not been measured, but for the same species, intensity and extraction parameters, they are expected to be comparable to the

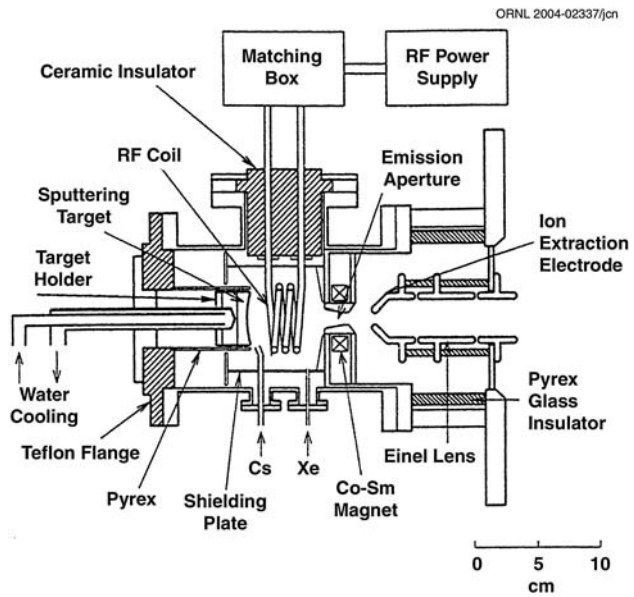


Fig. 12.31. A side view of the high-intensity RF plasma-sputter source developed at the University of Kyoto

Table 12.7. Maximum efficiencies for negative-ion production in the Kyoto University RF plasma-sputter negative-ion source

Ion	C ⁻	Si ⁻	Cu ⁻	Ge ⁻	Mo ⁻	Ta ⁻	W ⁻
Efficiency (%)	18.3	15.6	12.1	13.6	0.28	0.87	8.1

sources described in [120,121] because of the similarities in physical size of their respective sputter cathodes and emission apertures.

*The KEK/University
of Tsukuba Compact Axial-Geometry Plasma-Sputter Negative-Ion Source*

Negative-ion beams have an advantage over their positive-ion counterparts for SIMS analysis of insulating materials because they avoid the chronic charge-up problems present whenever positive-ion beams are used for this application. Mori et al. developed a compact, axial-geometry plasma-sputter negative-ion source with a 16 mm diameter cathode, operating in the DC mode with a LaB₆ cathode for plasma ignition, for such applications [125]. The source has been successfully employed by Yurimoto et al. for SIMS analyses of isotopic ratios in meteorites which intrinsically have insulating properties [126]. Beam intensities of 1.5 mA of Cu⁻ and 0.85 mA of Au⁻ have been extracted from the source and transmitted through a mass-

analysis system. The emittance of a Cu beam at the 90% contour has a value $\varepsilon_n \cong 6\pi \text{ mm mrad (MeV)}^{1/2}$.

12.8 Acknowledgments

The author is indebted to the Physics Division of the Oak Ridge National Laboratory for financial support. The author is also indebted to Dr. Yan Zhang for assisting in computing energy distributions induced in beams by the charge-exchange process, and to Dr. Yuan Liu for assisting in calculating the energy distributions of sputter-generated neutral and negative ions and for supplying the measured energy distribution for Ni beams for comparison with the Nörskov and Lundqvist theory.

References

1. H.S.W. Massey: *Negative Ions*, 3rd edn. (Cambridge University Press, London 1976)
2. B.M. Smirnov: *Negative Ions* (McGraw-Hill, New York 1982)
3. K.R. Lykke, K.K. Murray, W.C. Lineberger: *Phys. Rev. A* **43**, 6104 (1991)
4. H. Hotop, W.C. Lineberger: *J. Phys. Chem. Ref. Data* **14**, 731 (1985)
5. V. Bunge, C.F. Bunge: *Phys. Rev. A* **19**, 452 (1979)
6. G.D. Alton, R.N. Compton, D.J. Pegg: *Phys. Rev. A* **28**, 1405 (1983)
7. L.M. Blau, R. Novick, D. Weinflash: *Phys. Rev. Lett.* **24**, 1268 (1970)
8. R. Novick, D. Weinflash: *Proc. Int. Conf. Precision Measurements and Fundamental Constants*, eds. D. N. Langenberg and B. N. Taylor, Gaithersburg, MD, 1970, Natural Bureau of Standards Spec. Publication No. 343 (1971) p. 403
9. T.J. Kvale, G.D. Alton, R.N. Compton, D.J. Pegg, J.S. Thompson: *Phys. Rev. Lett.* **55**, 484 (1985)
10. D.M. Neumark, K.R. Lykke, T. Anderson, W.C. Lineberger: *Phys. Rev. A* **32**, 1890 (1985)
11. C. Blondel et al.: *Phys. Rev. A* **40**, 3698 (1989)
12. R. Trainham, G.D. Fletcher, D.J. Larson: *J. Phys. B* **20**, L777 (1987)
13. V.V. Petrunin, J.D. Voldstad, P. Balling, P. Kristensen, T. Andersen, H.K. Haugen: *Phys. Rev. Lett.* **76**, 744 (1996)
14. D.G. Leopold, W.C. Lineberger: *J. Chem. Phys.* **85**, 51 (1986)
15. J. Ho, K.M. Ervin, W.C. Lineberger: *J. Chem. Phys.* **93**, 6987 (1990)
16. T.M. Miller, A.E.S. Miller, W.C. Lineberger: *Phys. Rev. A* **33**, 3558 (1986)
17. H.H. Andersen, V.V. Petrunin, P. Kristensen, T. Andersen: *Phys. Rev. A* **42**, 3247 (1997)
18. V.V. Petrunin, J.D. Voldstad, P. Balling, P. Kristensen, T. Andersen, H.K. Haugen: *Phys. Rev. Lett.* **75**, 1911 (1995)
19. C.H. Greene: *Phys. Rev. A* **42**, 1405 (1990)
20. V.A. Dzuba, G.F. Gribakin: *Phys. Rev. A* **55**, 2443 (1997)
21. T.J. Kvale, R.N. Compton, G.D. Alton, J.S. Thompson, D.J. Pegg: *Phys. Rev. Lett.* **56**, 592 (1986)

22. R.D. Srivastava, O.M. Uy, M. Farber: *Trans. Faraday Soc.* **67**, 2941 (1971)
23. P.S. Drzaic, J. Marks, J.I. Brauman: In *Gas Phase Chemistry*, vol. 3, ed. M.T. Bowers (Academic Press, Orlando 1984) p. 167
24. K.M. Ervin, W.C. Lineberger: submitted to *J. Chem. Phys.*
25. R. Klein, R.P. McGinnis, S.R. Leone: *Chem. Phys. Lett.* **100**, 475 (1983)
26. K.J. Taylor, C.L. Pettiette, M.J. Craycraft, O. Chenovsky, R.E. Smalley: *Chem. Phys. Lett.* **152**, 347 (1988)
27. D.G. Leopold, K.K. Murray, T.M. Miller, W.C. Lineberger: unpublished data
28. A.E. Stevens, C.S. Fiegerle, W.C. Lineberger: *J. Chem. Phys.* **78**, 5420 (1983)
29. T. Anderson, K.R. Lykke, D.M. Neumark, W.C. Lineberger: *J. Chem. Phys.* **86**, 1858 (1987)
30. A.E.S. Miller, C.S. Fiegerle, W.C. Lineberger: *J. Chem. Phys.* **84**, 4127 (1986)
31. K.M. McHugh, J.G. Eaton, G.H. Lee, H.W. Sarkas, L.H. Kidder, J.T. Snodgrass, M.R. Manaa, K.H. Bowen: *J. Chem. Phys.* **91**, 3792 (1989)
32. D.G. Leopold et al.: *J. Am. Chem. Soc.* **108**, 179 (1986)
33. A.A. Viggiano, J.F. Paulson, F. Dale, M. Henchman, N.G. Adams, D. Smith: *J. Chem. Phys.* **89**, 2264 (1985)
34. M.K. Scheller, R.N. Compton, L.S. Cederbaum: *Science* **270**, 1160 (1995)
35. W.B. Herrmannsfeldt: SLAC Report No. 166 (1973)
36. J.C. Whitson, J. Smith, J.H. Whealton: *J. Comp. Phys.* **28**, 408 (1978)
37. J.H. Whealton, M.A. Bell, R.J. Raridon, K.E. Rothe, P.M. Ryan: *J. Appl. Phys.* **64**, 6210 (1988)
38. J.E. Boers: A digital computer code for the simulation of electron and ion beams on a PC, IEEE Cat. No. 93CH3334-0 (1993) 213. (PBGUNS is an electron/ion optics simulation code, developed by Thunderbird Simulations, Garland, TX, USA)
39. P. Spädtke: AXCEL-Gesellschaft für Schwerionen forschung (GSI), Interaktives Simulationsprogramm zur Berechnung von zwidimensionalen Potentialverteilungen elektrostatischer Anordnungen sowie von Ionbahnen in elektrostatischen Feldern unter Berhcksichtigung der Raumladung, Gesellschaft für Schwerionen forschung (GSI)-Report 9 (1983)
40. R. Becker, W.B. Herrmannsfeldt: *Rev. Sci. Instr.* **63**, 2756 (1992)
41. G.D. Alton, H. Bilheux: *Proc. 15th Int. Workshop on ECR Ion Sources*, Jyväskylä, Finland, JYFL Research Report No. 4/2002 (2002) p. 169
42. T.R. Walsh: *J. Nucl. Eng. C* **4**, 53 (1962)
43. C.D. Moak, H.E. Banta, J.N. Thurston, J.W. Johnson, R.F. King: *Rev. Sci. Instr.* **30**, 694 (1959)
44. P. Lawrence et al.: *Nucl. Instr. Meth.* **32**, 357 (1965)
45. H.T. Richards, G.M. Klody: *Proc. 2nd Int. Conf. on Ion Sources*, Vienna (1972) p. 804
46. J. Aubert, G. Gautherin, C. Lejeune: *Proc. 2nd Int. Conf. on Ion Sources*, Vienna (1972) p. 797
47. R.P. Bastide, N.B. Brooks, A.B. Wittkower, P.H. Rose, K.H. Purser: *IEEE Trans. Nucl. Sci.* **NS-12**, 775 (1965)
48. E. Heinicke, K. Bethge, H. Baumann: *Nucl. Instr. Meth.* **58**, 125 (1968)
49. H. Smith, H.T. Richards: *Bull. Am. Phys. Soc.* **18**, 617 (1973)
50. H.S.W. Massey: *Rep. Prog. Phys.* **12**, 248 (1949)
51. J.B. Hasted: *Proc. Roy. Soc. A* **205**, 421 (1951); **212**, 235 (1952)
52. G.D. Alton, D.L. Haynes: *Physics Division Progress Report for period ending September 30, 1994*, Oak Ridge National Laboratory, ORNL-6916, pp. 1–39

53. B.J. Donnally, G. Thoeming: Phys. Rev. **159**, 87 (1967)
54. G.D. Alton, T.J. Kvale, R.N. Compton, D.J. Pegg, J.S. Thompson: Nucl. Instr. Meth. A **244**, 142 (1986)
55. J. Heinemeier, P. Hvelplund: Nucl. Instr. Meth. **148**, 425 (1975)
56. J. Heinemeier, P. Hvelplund: Nucl. Instr. Meth. **148**, 65 (1975)
57. R.H. McFarland, A.S. Schlachter, J.W. Stearns, B. Liu, R.E. Olson: Phys. Rev. A **26**, 775 (1982)
58. G.I. Dimov, G.V. Roslyakov: Instr. Expl. Tech. **17**, 658 (1974)
59. R.J. Gurnius, C.J. Anderson, L.W. Anderson: Phys. Rev. A **16**, 2225 (1977)
60. B.A. D'Yachkov, V.I. Zinenko: Sov. Phys. – Tech. Phys. **16**, 305 (1971)
61. R.M. Ennis, Jr., D.E. Schechter, G. Thoeming, D.B. Schlafke, B.L. Donnally: IEEE Trans. Nucl. Sci. **NS-14**, 75 (1967)
62. R.J. Gurnus, L.W. Anderson: Nucl. Instr. Meth. **137**, 373 (1976)
63. A.S. Schlachter, D.H. Loyd, P.J. Bjorkholm, L.W. Anderson, W. Haeberli: Phys. Rev. **171**, 20165 (1968)
64. J. Heinemeier, P. Hvelplund: unpublished
65. M. Kaminsky: *Atomic and Ionic Impact Phenomena on Metal Surfaces* (Springer, New York 1965)
66. H. Kawano, F.M. Page: Int. J. Mass. Spectr. Ion Phys. **50**, 1 (1983)
67. H. Kawano et al.: Int. J. Mass Spectr. Ion Phys. **50**, 35 (1983)
68. H. Kawano, Y. Hidaka, M. Suga, F.M. Page: Int. J. Mass Spectr. Ion Phys. **50**, 77 (1983)
69. J.M. Lafferty: J. Appl. Phys. **22**, 299 (1951)
70. V.S. Fomenko: *Emission Properties of Materials*, JPRS-56579 (NTIS, U.S. Dept. of Commerce, Springfield, VA 1972)
71. H. Ahmed, A.M. Broers: J. Appl. Phys. **43**, 185 (1972)
72. S. Hosoki, S. Yamamoto, K. Hayakawa, H. Okano: Jpn. J. Appl. Phys. (Suppl. 2, Part 1) 285 (1974)
73. H. Yamauchi, K. Takagi, I. Yuito, U. Kawabe: Appl. Phys. Lett. **29**, 638 (1976)
74. A. Avdienko, M.D. Malev: Vacuum **27**, 583 (1977)
75. G.D. Alton, M.T. Johnson, G.D. Mills: Nucl. Instr. Meth. A **328**, 154 (1993)
76. N. Kashihiro, E. Vietske, G. Zellermann: Rev. Sci. Instr. **48**, 150 (1977)
77. B. Vosicki, T. Bornstad, L.C. Carraz, J. Heinemeier, H.L. Ravn: Nucl. Instr. Meth. **186**, 307 (1981)
78. G.D. Alton et al.: Nucl. Instr. Meth. B **211**, 425 (2003)
79. E. Krohn: J. Appl. Phys. **33**, 3523 (1962)
80. G.K. Wehner, G.S. Anderson: The nature of physical sputtering, in *Handbook of Thin Foil Technology*, Chapter 3, pp. 3-19, eds. L.I. Maissel and R. Glang (McGraw-Hill, New York 1970)
81. R. Behrish, W. Heiland, W. Poschenrieder, P. Staib, H. Verbeek (eds.): *Ion Surface Interactions and Sputtering and related Phenomena* (Gordon and Breach, New York 1973)
82. H.H. Andersen, H.L. Bay: Sputtering yield measurements, in *Sputtering by Particle Bombardment I*, chapter 4, ed. R. Behrish, Topics in Applied Physics, Vol. 47 (Springer, New York 1981)
83. R. Behrish (ed.): *Sputtering by Particle Bombardment II*, Topics in Applied Physics, Vol. 52 (Springer, New York 1983)
84. M. Müller, G. Hortig: IEEE Trans. Nucl. Sci. **NS-16**, 3523 (1962)
85. G.D. Alton: IEEE Trans. Nucl. Sci. **NS-23**, No. 2, 1113 (1976)

86. R. Middleton, C.T. Adams: Nucl. Instr. Meth. **118**, 329 (1974)
87. G. Doucas, H.R.McK. Hyder, A.B. Knox: Oxford University report (1975)
88. K.R. Chapman: IEEE Trans. Nucl. Sci. **NS-23**, 1109 (1976)
89. G.D. Alton et al.: Nucl. Instr. Meth. B **170**, 515 (2000)
90. G.D. Alton, J.R. Beene: J. Phys. G, Nucl. Part. Phys. **24**, 1347 (1998)
91. G.D. Alton: Nucl. Instr. Meth. B **73**, 221 (1993)
92. G.D. Alton: Rev. Sci. Instr. **65**, 1141 (1994)
93. G.T. Caskey, R.A. Douglas, H.T. Richards, H.V. Smithe Jr.: Nucl. Instr. Meth. **157**, 1 (1978)
94. R. Middleton: Nucl. Instr. Meth. **214**, 139 (1983)
95. G.D. Alton: Nucl. Instr. Meth. A **244**, 133 (1986)
96. G.D. Alton, C.M. Jones: Nucl. Instr. Meth. A **244**, 170 (1986)
97. The Model 860 cesium-sputter negative-ion source is a product of the former General Ionex Corporation, Newburyport, MA, USA
98. G.D. Alton, G.D. Mills: IEEE Trans. Nucl. Sci. **NS-32**, 1822 (1985)
99. G.D. Alton, M.T. Johnson: Nucl. Instr. Meth. A **328**, 154 (1993)
100. I.D. Proctor: Nucl. Instr. Meth. B **40/41**, 727 (1989)
101. T.R. Niklaus, W. Baur, G. Bonani, M. Suter, H.A. Synal, W. Wöfli: Rev. Sci. Instr. **63**, 2485 (1992)
102. G. Du: Rev. Sci. Instr. **63**, 1151 (1992)
103. G.D. Alton: Proc. 11th Symposium on Ion Sources and Ion-Assisted Technology, Tokyo (1987) p. 157
104. G.D. Alton: *Proc. of the 1989 IEEE Particle Accelerator Conference*, 89CH2669-0, **Vol. 2**, 1112 (1989)
105. G.D. Alton, B. Cui, Y. Bao, C.A. Reed, J.A. Ball, C. Williams: Proc. 8th Inter. Conf. on Heavy Ion Accelerator Technology, American Institute of Physics (AIP) **CP473**, 352 (1999)
106. G.D. Alton, J.A. Benjamin: Nucl. Instr. Meth. **211**, 1 (1983)
107. G.D. Alton, G.D. Mills: Nucl. Instr. Meth. A **276**, 388 (1989)
108. R. Middleton: *A Negative Ion Cook Book* (1989) (unpublished)
109. G.D. Alton, J.W. McConnell, S. Tajima, G.J. Nelson: Nucl. Instr. Meth. B **24/25**, 826 (1987)
110. G.D. Alton, J.W. McConnell: Nucl. Instr. Meth. A **268**, 445 (1988)
111. H.H. Andersen, P. Tykesson: IEEE Trans. Nucl. Sci. **NS-22**, 1632 (1975)
112. P. Tykesson, H.H. Andersen, J. Heinemeier: IEEE Trans. Nucl. Sci. **NS-23**, 1104 (1976)
113. G.D. Alton, G.C. Blazey: Nucl. Instr. Meth. **166**, 105 (1979)
114. G.D. Alton, R.M. Beckers, J.W. Johnson: Nucl. Instr. Meth. A **244**, 146 (1986)
115. G.D. Alton, R. Lohwasser, B. Cui, Y. Bao, T. Zhang and C.A. Reed: Proc. 8th Inter. Conf. on Heavy Ion Accelerator Technology, American Institute of Physics (AIP) **CP473**, 340 (1999)
116. R.L. York and R.R. Stevens, Proc. 3rd International Conf. On Production and Neutralization of Negative Ion Beams, ed. K. Prelec, American Institute of Physics Conf. Proc. No. 111, New York, p. 410 (1984)
117. R.L. York, R.R. Stevens, Jr., R.A. DeHaven, J.R. McConnell, E.P. Chamberlain, R. Kandarian: Nucl. Instrum. and Meth. B **10/11**, 891 (1985)
118. A. Takagi, Y. Mori, K. Ikegami, S. Fukumoto: IEEE Trans. Nucl. Sci. **NS-32** (5), 1782 (1985)

- 119. Y. Mori, A. Takagi, K. Ikegami, S. Fukumoto: Proc. 4th International Conf. On Production and Neutralization of Negative Ion Beams, ed. J.G. Alessi. (American Institute of Physics Conf. Proc. No. 158, New York, 1987) p. 378
- 120. G.D. Alton, Y. Mori, A. Takagi, A. Ueno, S. Fukumoto: Letters to the editor, Nucl. Instrum. and Meth. **A 270**, 194 (1988); *ibid* Nucl. Instrum. and Meth. **B 40/41**, 1008 (1989)
- 121. Y. Mori, G.D. Alton, A. Takagi, A. Ueno and S. Fukumoto, Nucl. Instrum. and Meth. **A 273**, 5 (1988)
- 122. G.D. Alton, Rev. Sci. Instrum. **63** 2455 (1992)
- 123. J. Ishikawa, H. Tsuji, Y. Okada, Y. Toyota, Y. Gotoh, Vacuum **44** 203 (1993)
- 124. H. Tsuji and J. Ishikawa, Rev. Sci. Instrum. **63**, 2488 (1992)
- 125. Y. Mori, Nucl. Instrum. and Meth., **A 328** 146 (1993)
- 126. H. Yurimoto, Y. Mori, and H. Yamamoto, Rev. Sci. Instr. **64** 1146 (1993)

Box 7: Tandem Terminal Ion Source

G.C. Harper

Nuclear Physics Laboratory, University of Washington, Seattle, WA, USA
gharper@u.washington.edu

Introduction

Some astrophysics experiments require ion beams at low energies to simulate the reactions that take place in stars. The accompanying low count rates from the small reaction cross sections can be improved with higher beam currents. The beam power requirements are not beyond the capabilities of an FN tandem but do require operation in an unusually low-voltage, high-current regime. A radio frequency (RF) discharge source assembly has been designed and developed for use in the terminal of the model FN tandem electrostatic accelerator at the University of Washington to produce high-intensity, low-energy beams of $^1\text{H}^+$, $^2\text{H}^+$, $^3\text{He}^+$ and $^4\text{He}^+$. A similar source arrangement was used in the FN tandem at Florida State University to accelerate $^{15}\text{N}^+$ ions.

The RF Discharge Source

The assembly, illustrated in Fig. B7.1, consists of a commercially manufactured RF discharge source [1], a porcelain insulator containing an extractor electrode and einzel lens, a double-focusing permanent dipole magnet or a spherical electrostatic deflector, vertical and horizontal electrostatic steerers, and a gas bottle manifold with associated isolation and metering valves. The RF source was specified to deliver $400\text{ }\mu\text{A}$ of $^1\text{H}^+$ from a 2 mm diameter aluminum canal while producing a gas load of 0.2–0.4 Pa l/s. Since the tandem does not have any terminal pumping, we were concerned that a large gas load would result in collisions with the low-energy particles and reduce transmission. In order to reduce the gas load, a 1 mm diameter canal was specified. The gas load from an RF discharge source varies approximately as d^3/l , where d is the canal diameter and l is the canal length [2,3]. Reduced beam current was expected using the smaller canal because source output varies as $(d/l)^2$ and as $1/\sqrt{m}$, where m is the mass in AMU [4]. In a high-intensity proton experiment, we tested the source with a high gas flow and determined that the gas load from the source operating with a 2 mm canal could be managed by our pumping system.

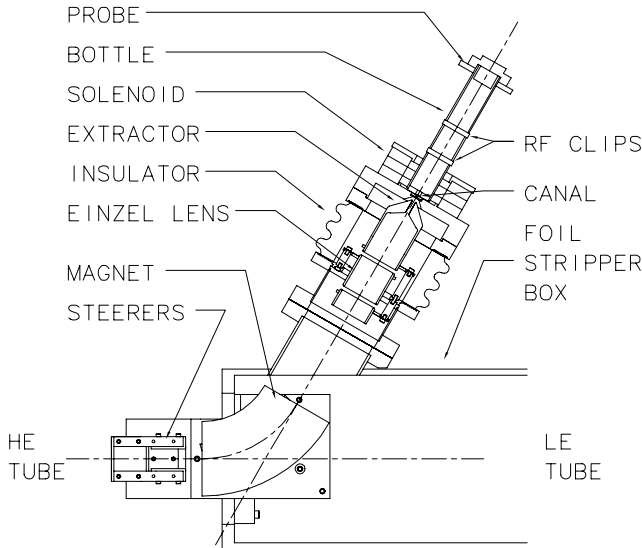


Fig. B7.1. Terminal ion source and associated optics

The Double-Focusing Magnet

Removing heat from and delivering current to an electromagnet operating inside a vacuum chamber is a difficult and complex problem. To avoid this problem, we designed and built a pair of permanent-magnet assemblies. One of the assemblies transports $^1\text{H}^+$ or $^2\text{H}^+$ and the other transports $^3\text{He}^+$ or $^4\text{He}^+$. Each dipole magnet has a 60° bending angle and a 14 cm bending radius, R . The poles are made from 8C ferrite material with the pole faces cut at 17° to the beam axes to produce double foci at $3.46R$, or about 50 cm, from the pole faces [5]. The frames are made from soft steel and bolted together. The poles are glued onto the top and bottom frame pieces. The assemblies are magnetized to produce a transverse B -field of 0.15 to 0.25 T. Thin (1.5 to 3.0 mm) steel pole caps which overlap the ceramic pole pieces by 1 to 5 mm on each side along the beam trajectory are required to tune the field to the magnitude appropriate for each individual ion species and to make the fringe fields at the pole faces more uniform. The magnet is mounted on an aluminum platform inside the foil-stripper box.

The Spherical Electrostatic Deflector

A double-focussing, spherical electrostatic deflector may alternatively be used to deflect the beam onto the accelerator axis. The advantage of such a system is that it is mass-independent and may be used with a gas manifold to rapidly switch between the ion species available. The disadvantage is that the optics

are not as well matched in the geometry available, and beam transport suffers somewhat. Two additional high-voltage supplies are required to operate the electrostatic deflector. The system was modeled using an electrostatic software package [6], and it was found that the optimum transport would occur with the deflector electrodes operated in an unbalanced mode with the inner, negative electrode at about 1.5 times the voltage of the outer, positive electrode. The bending radius R of the deflector is again 14 cm and the deflector produces foci located $2R$, or about 28 cm, from the electrode edges [7].

Beam Transport and Optics

Vertical and horizontal steering plates are mounted on the platform downstream from the deflection system. The plates are 5.7 cm wide, 4.5 cm long and spaced 2.6 cm apart, with the horizontal pair preceding the vertical pair by 3 mm. Bipolar power supplies producing ± 350 V drive the steering plates. The supplies are 4-stage multipliers running off of the 400 Hz terminal power. For simplicity, these supplies accept the same AC input lines as the normal terminal steerers.

Within the source, one of two available einzel lens assemblies is used depending on the deflection system. A 2.5 cm diameter, short lens is used with the electrostatic deflector to produce an object point in front of the deflector and an image point inside the beam tube 28 cm from the deflector. A 5 cm diameter, long lens is used with the dipole magnet to produce a virtual object 12 cm behind the actual source and an image point inside the beam tube 50 cm from the magnet.

Accelerator Tube Gradient

The entrance voltage gradient of the spiral-inclined-field tube produces lens action that is too strong for the low-energy (17 keV) beam, and the transverse electric-field components are large enough to sweep the beam into the side of the tube. The tube and the spiral inclined planes were carefully modeled using a computer program to determine how the entrance gradient could be modified to successfully transport the beam. A set of resistor values was determined by iteratively applying the code and observing the calculated variation of the displacement of the beam from the tube axis. We reduce the gradient in approximately the first quarter of our high-energy accelerator tube (the first half of accelerator tube #3) during source operation to eliminate these problems. Twenty-two resistor assemblies having 20% of the nominal column resistor value are installed at the tube entrance. These are followed by 3 assemblies with 50% value, 10 assemblies with 60% value and 14 assemblies with 70% value.

The aperture size of the spiral tube is small enough to restrict beam transport at energies below about 2 MeV. For lower-energy beams, down to 150 keV or even less, we install a KN accelerator tube in our beam tube #3 position which has noninclined planes and a much larger aperture. The noninclined-field tube planes allow secondary electrons to easily stream to the terminal. As a result this tube produces x-rays at terminal voltages above 4.5 MV that are intense enough to make the accelerator unusable.

Ion Beams and Experiments Run

The source proton current was expected to be about 100 μA with a gas load of 0.03 Pa l/s. The source produces 70 μA of protons with a measured gas load of 0.1 Pa l/s and a focused FWHM of 1.3 mm. The measured beam emittance is 1.1π mm mrad $\sqrt{\text{MeV}}$. At the 5.5 MV terminal voltage required by one of the experiments, 90% of the beam is transported through the spiral-inclined-field tubes. Over the terminal voltage range of 0.15 to 7.5 MV, over 50% of the beam can be transported using one of the configurations described. A brief summary of the experiments run and ion beams produced using this source is given in Table B7.1.

Table B7.1. Table of ion species and terminal-ion-source experiments

Ion	Energy Range (keV)	Experiment	Beam Current (μA)
$^1\text{H}^+$	160–1400	$^7\text{Be}(\text{p}, \gamma)$	10–40
$^2\text{H}^+$	770–1400	$^7\text{Li}(\text{d}, \text{p})$	15–20
$^3\text{He}^+$	5000–6000	$^6\text{Li}(^3\text{He}, \text{n})$	28
$^4\text{He}^+$	1370	$^7\text{Be}(\alpha, \gamma)$	15

References

1. National Electrostatics Corporation, Middleton, WI, USA
2. V.J. Kowalewski, C.A. Mayans, M. Hammerschlag: Nucl. Instr. Meth. **5**, 94 (1959)
3. S.K. Allison, E. Norbeck: Rev. Sci. Instr. **27**:5, 287 (1956)
4. K.R. Spangenberg: *Vacuum Tubes* (McGraw-Hill, New York 1949) p. 447
5. J.J. Livingood: *The Optics of Dipole Magnets* (Academic Press, New York 1969) p. 63
6. D.A. Dahl: *SIMION 3D* (Idaho National Engineering Laboratory, Idaho Falls, ID, USA 1995)
7. H. Wollnik: Electrostatic prisms. In: *Focusing of Charged Particles*, vol. II, ed. by A. Septier (Academic Press, New York 1967) p. 163

13 Ion Optics and Beam Transport

J.D. Larson

10011 East 35th Street Terrace South, Independence, MO 64052-1107, USA

13.1 Introduction

Effective and efficient use of ion or electron accelerators requires a knowledge of how to control and manipulate beams of charged particles using electric and magnetic fields. Because this discipline shares many concepts acquired historically through the study of light, the names “electron optics” and “ion optics” came into early usage. Subsequently, the generic expression “beam transport” was coined, in part because transporting particle beams (of any kind) over long distances both inside and outside of accelerators has come to dominate the field. In this chapter, and depending on context, sometimes reference will be made to “optics” (especially if there is close analogy to conventional light optics), and at other times to “transport”. No special distinction is intended. The word “ion” generally will mean any electrically charged particle moving (mostly) in vacuum, while “beam” will refer to collections of ions moving on average in the same direction.

At first blush, predicting particle trajectories within an electrostatic accelerator seems almost devoid of interest. In the uniform electric field of a basic accelerator tube, charged particles moving at low speeds ($v \ll c$) follow parabolic trajectories analogous to masses falling in a uniform gravitational field. The velocity component directed parallel to the tube axis undergoes constant acceleration, while transverse velocity components remain unchanged. Unfortunately, this much of the story is almost irrelevant. At each end of a tube, where the electric field strength changes, fringing fields create strong lenses [1] that dominate the overall optics. The difficulty of passing beams through the stronger of these lenses at the tube entrance without incurring substantial loss converts this to a problem of continuing interest. For the simplest accelerators it is enough to understand this much. But devices get added and machines grow in complexity. The largest electrostatic accelerators house elaborate beam transport systems that require careful study both for design and for efficient use. Outside the accelerator, the beam may have come from somewhere else (in the case of a tandem) and certainly has to go somewhere else. Both tasks require efficient coupling of the accelerator to external devices.

13.2 Methods

Throughout this discussion, the x axis points away from the reader into the page, the y axis lies within the page pointing toward the top and the z axis lies within the page pointing toward the right (longitudinal beam axis). It is common practice to select x, z to define the midplane of a dipole and, since dipoles often bend in the horizontal, x, z frequently is chosen to represent the horizontal plane. In axially symmetric systems, where the distinction between x and y disappears, an r, z frame may be substituted. The location of a particle moving from left to right along the z axis is described by (usually) small deviations from a reference particle representing the idealized center of the beam:

- x displacement from beam axis in x direction
- y displacement from beam axis in y direction
- x' trajectory slope dx/dz in x, z plane
- y' trajectory slope dy/dz in y, z plane
- δp fractional deviation from central momentum
- δE fractional deviation from central energy
- δv fractional deviation from central velocity
- δm fractional deviation from central mass
- δz deviation from reference position along z
- δt deviation from reference transit time
- $\delta \varphi$ deviation from reference r.f. phase angle

Obviously δp , δE and δv are not independent; ordinarily, one is selected as a working coordinate (typically δp for all-magnetic systems and δE for all-electric systems) and the others calculated from it as needed. Likewise, for modulated beams, only one of δz , δt or $\delta \varphi$ is chosen to describe the longitudinal position of a particle relative to the center of a beam pulse or bunch. The differential mass variation, δm , is useful primarily for spectrometry and will not be pursued here. Mass selection (e.g. for AMS) is easily calculated for each discrete mass without introducing δm .

Six deviations (e.g. x , x' , y , y' , δp and δz) are commonly used to describe the location of a particle in a six-dimensional (6D) phase space volume centered on a reference particle. For dc beams, the longitudinal deviation δz is discarded but δp is retained to describe dispersion in magnetic dipoles (alternately, δE is retained to describe dispersion in electric dipoles, or δv in velocity filters). Acceleration changes x' , y' and δp in a coherent way and once acceleration is accounted for, Liouville's theorem dictates that the total volume is conserved during beam transport even in the presence of aberration. Coherent aberrations such as sextupole (2nd order) or octupole (3rd order) in principle are reversible. In practice, although some corrections can be made, it usually proves impossible to force all of the misshapen genie back into the bottle because of the way aberrations warp the envelope of the phase space volume, making it effectively larger. Incoherent aberrations (caused, for

example, by strippers) occur at the microscopic level and are equivalent to heating the beam; these are not reversible using corrective lenses but can be overcome in cyclic machines by procedures that cool the beam.

Typically, in two dimensions (e.g. x, x'), contours of equal intensity in a beam occupy areas that are roughly elliptical to start with [2,3]. The reasons for this are varied but, in simplest form, any beam that scrapes against multiple apertures tends to lose projecting corners and smooth toward an elliptical shape in phase space. For mathematical convenience, an ellipse (or ellipsoid) is usually substituted for less tractable polygons. An ellipse remains an ellipse under linear transformations and, because most common devices operate almost linearly for small deviations and do not cause mixing of x and y coordinates (an important exception being solenoid lenses that rotate the beam around the z axis), the 2D area is usually conserved. Suppose that such an ellipse is upright with semiaxes x and x' aligned to the coordinate frame. For area to be conserved, the product xx' must remain constant. This means that reductions in spot size at this location (called a beam waist) can only come at the expense of increased angles of incidence. The largest possible angle (x'_{max}) will be limited by the distance upstream to the nearest lens and the aperture at that lens. If the focused spot is to be made any smaller, then the upstream lens must be brought closer or given a larger working aperture; lens adjustments alone cannot overcome placement. Thus, the combination of physical layout and phase space content governs the smallest spot sizes available at a given location. Collimators (e.g. successive apertures, a tube or a channel) restrict both size and angle, thus defining the maximum area of an ellipse (or other figure) that can pass through [4]. Matching phase space content to a collimator is a much more stringent requirement than simply minimizing beam size [5].

Matrix analysis begins with equations that relate the coordinates of a particle as it leaves some part of a beam transport system to the coordinates at entry. To lowest order,

$$\begin{pmatrix} x \\ x' \\ y \\ y' \\ \delta p \\ \delta z \\ 1 \end{pmatrix} = \begin{pmatrix} a_{x/x} & a_{x/x'} & a_{x/y} & a_{x/y'} & a_{x/\delta p} & a_{x/\delta z} & a_{x/1} \\ a_{x'/x} & a_{x'/x'} & a_{x'/y} & a_{x'/y'} & a_{x'/\delta p} & a_{x'/\delta z} & a_{x'/1} \\ a_{y/x} & a_{y/x'} & a_{y/y} & a_{y/y'} & a_{y/\delta p} & a_{y/\delta z} & a_{y/1} \\ a_{y'/x} & a_{y'/x'} & a_{y'/y} & a_{y'/y'} & a_{y'/\delta p} & a_{y'/\delta z} & a_{y'/1} \\ a_{\delta p/x} & a_{\delta p/x'} & a_{\delta p/y} & a_{\delta p/y'} & a_{\delta p/\delta p} & a_{\delta p/\delta z} & a_{\delta p/1} \\ a_{\delta z/x} & a_{\delta z/x'} & a_{\delta z/y} & a_{\delta z/y'} & a_{\delta z/\delta p} & a_{\delta z/\delta z} & a_{\delta z/1} \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ x' \\ y \\ y' \\ \delta p \\ \delta z \\ 1 \end{pmatrix}_0. \quad (13.1)$$

The coefficients $a_{i/1}$ accommodate contributions that do not depend on incident coordinates.

Over a distance of length l in empty space, the particle position in the x, z plane drifts as $x = x_0 + lx'_0$ (and similarly for y, z), while the divergence “angle” does not change, keeping $x' = x'_0$. A thin lens reverses this; the divergence changes as $x' = \pm x_0/f_x + x'_0$, where f_x is the focal length (which need

not be the same as f_y), while the instantaneous position remains unchanged at $x = x_0$. In matrix form, these two basic operations become the workhorses of beam transport analysis:

$$\begin{array}{ccc} \text{2D partition} & \text{drift} & \text{thin lens} \\ \left(\begin{array}{cc} a_{x/x} & a_{x/x'} \\ a_{x'/x} & a_{x'/x'} \end{array} \right), & \left(\begin{array}{cc} 1 & l \\ 0 & 1 \end{array} \right), & \left(\begin{array}{cc} 1 & 0 \\ \pm 1/f_x & 1 \end{array} \right). \end{array} \quad (13.2)$$

The thin-lens matrix may be extended to describe thick lenses (asymptotically) as a single impulse or as a product of simpler components [5,6]:

$$\begin{aligned} & \left(\begin{array}{cc} F_2/f_2 & \pm(F_1F_2 - f_1f_2)/f_2 \\ \pm 1/f_2 & F_1/f_2 \end{array} \right) \\ &= \left(\begin{array}{cc} 1 \pm \delta_2 & \\ 0 & 1 \end{array} \right) \left(\begin{array}{cc} 1 & 0 \\ \pm 1/f_2 & 1 \end{array} \right) \left(\begin{array}{cc} 1 & 0 \\ 0 & f_1/f_2 \end{array} \right) \left(\begin{array}{cc} 1 \pm \delta_1 & \\ 0 & 1 \end{array} \right). \end{aligned} \quad (13.3)$$

The $\pm$ signs select for divergent lenses (+) or convergent lenses (-). The subscripts 1 and 2 refer to the object and image sides, respectively. The focal lengths f_i and focal points (focal planes) F_i are normally positive. The differences $\delta_i = (F_i - f_i)$ are displacements of the principal planes from reference planes at points of entry and exit. The sum $\pm(\delta_1 + \delta_2)$ represents an effective (not necessarily actual) overall length. The signs in (13.3) have been altered from those of DiChio et al. [6] in order to preserve the beam transport convention that a convergent lens preceded by a drift matrix of length F_i produces a point-to-parallel focus, and followed by drift F_2 produces a parallel-to-point focus. If the determinant $f_1/f_2 < 1$, the 2D phase space area decreases (owing to acceleration); if $f_1/f_2 > 1$, the area increases (deceleration).

Continuous lensing action distributed over the length L of a component, such as occurs in quadrupole fields, may be represented for convergent (C) and divergent (D) lenses as follows [5]:

$$\left(\begin{array}{cc} \cos \theta & (L/\theta) \sin \theta \\ -(\theta/L) \sin \theta & \cos \theta \end{array} \right)_C = \left(\begin{array}{cc} 1 - \delta_C & \\ 0 & 1 \end{array} \right) \left(\begin{array}{cc} 1 & 0 \\ -1/f_C & 1 \end{array} \right) \left(\begin{array}{cc} 1 - \delta_C & \\ 0 & 1 \end{array} \right), \quad (13.4)$$

$$\left(\begin{array}{cc} \cosh \theta & (L/\theta) \sinh \theta \\ (\theta/L) \sinh \theta & \cosh \theta \end{array} \right)_D = \left(\begin{array}{cc} 1 \delta_D & \\ 0 & 1 \end{array} \right) \left(\begin{array}{cc} 1 & 0 \\ 1/f_D & 1 \end{array} \right) \left(\begin{array}{cc} 1 \delta_D & \\ 0 & 1 \end{array} \right). \quad (13.5)$$

For a particle of energy qU moving parallel to the axis in an electric quadrupole having a pole potential V_a at radius a , or in a magnetic quadrupole having a gradient $\partial B_x/\partial y = -\partial B_y/\partial x$,

$$\theta_E = (L/a)(2V_a/E\rho)^{1/2}, \quad E\rho = \beta^2\gamma mc^2/q \xrightarrow{v \ll c} 2U, \quad (13.6)$$

$$\theta_B = L[(\partial B_x/\partial y)/B\rho]^{1/2}, \quad B\rho = \beta\gamma mc/q \xrightarrow{v \ll c} (2mU/q)^{1/2}, \quad (13.7)$$

$$1/f_C = (\theta/L) \sin \theta, \quad F_C = f_C \cos \theta, \quad \delta_C = f_C(\cos \theta - 1), \quad (13.8)$$

$$1/f_D = (\theta/L) \sinh \theta, \quad F_D = f_D \cosh \theta, \quad \delta_D = f_D(\cosh \theta - 1). \quad (13.9)$$

Note that (13.4) and (13.5) reflect (13.3) in mirror-symmetric form with $f_1 = f_2 = f$, $F_1 = F_2 = F$ and unit determinant. A quadrupole singlet lens is maximally astigmatic: convergent in one plane while divergent in the other. Combinations of alternating-gradient quadrupole lenses (doublets, triplets and higher multiplet combinations) provide double-focusing capability [5,7,8]. Uniform-field (flat-pole) magnetic dipoles focus continuously like a quadrupole lens in the bending plane. In the perpendicular plane, the motion is that of a pure drift of length $\rho\theta$:

$$\begin{pmatrix} x \\ x' \\ \delta p \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \rho^{-1} \tan \beta_2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \theta & \rho \sin \theta & \rho(1 - \cos \theta) \\ -\rho^{-1} \sin \theta & \cos \theta & \sin \theta \\ 0 & 0 & 1 \end{pmatrix} \\ \times \begin{pmatrix} 1 & 0 & 0 \\ \rho^{-1} \tan \beta_1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ x' \\ \delta p \end{pmatrix}_0, \quad (13.10)$$

$$\begin{pmatrix} y \\ y' \\ \delta p \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -\rho^{-1} \tan \beta_2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \rho\theta & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ \times \begin{pmatrix} 1 & 0 & 0 \\ -\rho^{-1} \tan \beta_1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y \\ y' \\ \delta p \end{pmatrix}_0. \quad (13.11)$$

If the pole ends are rotated away from perpendicular to the beam axis by angles β_1 and β_2 then, to first order, equal and opposite thin lenses are created in the two planes. Positive rotation angles shorten the path within the dipole for particles having displacement $+x$. The third column in each matrix accounts for changes that occur when the particle momentum differs from the value that sets the bending radius ρ . Electric dipoles [9] and nonuniform magnetic dipoles [8] are only slightly more complicated. For other perspectives on dipoles and quadrupoles see [10–12].

13.2.1 Focal Constraints

Focusing requirements are satisfied by reducing one or more elements of a beam transport matrix to zero; for example, in the x, z plane,

- $a_{x/x'} = 0$, focus point-to-point (final x independent of initial x')
- $a_{x'/x} = 0$, focus parallel-to-parallel (final x' independent of initial x)
- $a_{x/x} = 0$, focus parallel-to-point (final x independent of initial x)
- $a_{x'/x'} = 0$, focus point-to-parallel (final x' independent of initial x')
- $a_{x/\delta p} = 0$, dispersionless in space (final x independent of initial δp)
- $a_{x'/\delta p} = 0$, dispersionless in angle (final x' independent of initial δp).

Note that unless axial symmetry is present, setting $a_{x/x'} = 0$ assures only a line focus in the x, y plane. A second constraint, $a_{y/y'} = 0$ in the y, z plane, is required to produce a double focus in both x and y (e.g. when using quadrupole lenses). Often two constraints are desired in one plane; for example, achieving $a_{x/\delta p} = a_{x'/\delta p} = 0$ eliminates (to first order) all dispersion introduced by upstream components [10]. Obviously, devices capable of effecting these goals must be represented in the beam transport matrix. Skills required for successful beam transport analysis include the selection, placement and adjustment of such controlling devices.

Thus far, matrices have been assumed to operate on a single ray vector representing a single particle. For most purposes a beam of particles may be described using one ray for each dimension in phase space [12–14]. In two dimensions, the perimeter of a phase space ellipse may be traced as the parameter φ ranges from 0 to 2π using the following equation:

$$\begin{pmatrix} x(\varphi) \\ x'(\varphi) \end{pmatrix} = \begin{pmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{pmatrix} \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} = \mathbf{E} \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}. \quad (13.12)$$

Columns in the beam ellipse matrix $\mathbf{E}$ comprise individual ray vectors, which must be selected for orthogonality; the most convenient starting choice is an upright ellipse (or ellipsoid) having all diagonal elements nonzero ($e_{ii} \neq 0$) and all off-diagonal elements zero. Desired conditions which apply to the beam as a whole (e.g. waist, minimum or size) may be achieved by applying constraints to the $\mathbf{E}$ matrix after it has been acted upon by a beam transport matrix. The parameter φ serves as an analytic tool for locating extrema and as a graphical tool for outlining ellipses; it is not required during beam transport calculations.

13.2.2 Focus, Waist or Minimum?

Confusion often arises from casual statements to the effect that lenses are used to bring the beam to a focus. A true focus is applicable for some purposes (notably microscopy and spectrometry), but far more commonly a beam is “focused” to a spot of minimum size or (less often) to a waist.

Focus: a location where rays that initially shared the same object position (point) or angle (parallel) arrive similarly correlated at an image; thus we have point-to-point, parallel-to-point, point-to-parallel and parallel-to-parallel foci, which are convenient for beam transport calculations because they are geometrically determined and independent of the beam (ignoring space charge). The familiar point-to-point focus ordinarily coincides with neither waist nor minimum; however, a combined point-to-point and parallel-to-parallel focus guarantees waist-to-waist transmission.

Waist: a place where the beam envelope momentarily is parallel as it passes through a local minimum in size; an upright beam ellipse. A waist is not the smallest size that can be obtained at this location; increasing the

strength of the nearest upstream lens to draw the waist upstream will produce a smaller spot (a minimum) where the waist had been. A waist is of interest because of its reflection symmetry (upstream and downstream beam envelopes are mirror images) and because small displacements (as might be caused by fluctuating lens strength) produce only second-order changes in beam size. Very small waists are always near to a point focus and to a minimum; very large waists are near to a parallel focus (an angular minimum).

Minimum: the smallest size the beam can be made (in x, y or both) at a specific location using available controls. It is always possible to reduce the beam size to a minimum even if a waist or a focus is unattainable. A calculated minimum simulates the best “focus” that operators observe on beam viewers and scanners. Minimum size or minimum angle is accessible using (from among various possibilities) the ellipse–matrix formalism (i.e. an ellipse, not necessarily upright, projecting either the smallest spatial extent or the smallest angular extent at this location).

13.3 Single-Stage Accelerators

The heart of a single-stage electrostatic accelerator is the accelerator tube, whose nonrelativistic matrix description may be approximated as follows (see Sect. 13.6 for derivation, relativistic forms and other refinements):

$$\mathbf{A} = \begin{array}{cc} \text{lens} & \text{field} & \text{lens} & \text{simple tube model} \\ \left(\begin{array}{cc} 1 & 0 \\ \frac{1}{f_2} & 1 \end{array} \right) & \left(\begin{array}{cc} 1 & \frac{2L}{R+1} \\ 0 & \frac{1}{R} \end{array} \right) & \left(\begin{array}{cc} 1 & 0 \\ -\frac{1}{f_1} & 1 \end{array} \right) & = \left(\begin{array}{cc} -\frac{R-3}{2} & \frac{2L}{R+1} \\ -\frac{3(R+1)(R-1)^2}{8LR^2} & \frac{3R-1}{2R^2} \end{array} \right) \end{array} \quad (13.13)$$

L is the length of the accelerator tube; $R = v/v_0 = (U/U_0)^{1/2}$ is the ratio of exit to entrance velocities (a velocity increase being assumed here for purposes of illustration); $f_1 = 4U_0L/(U - U_0) = 4L/(R^2 - 1)$ is the focal length of the (assumed circular and convergent) entrance lens; and $f_2 = 4UL/(U - U_0) = 4LR^2/(R^2 - 1)$ is the focal length of the corresponding exit lens.

Suppose that, starting from an object located a distance p upstream from this tube, a focus is desired at a distance q beyond the exit of the tube. The model for this is

$$\begin{aligned} \mathbf{B} &= \begin{pmatrix} b_{x/x} & b_{x/x'} \\ b_{x'/x} & b_{x'/x'} \end{pmatrix} = \begin{pmatrix} 1 & q \\ 0 & 1 \end{pmatrix} \mathbf{A} \begin{pmatrix} 1 & p \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} a_{x/x} + qa_{x'/x} & a_{x/x'} + pa_{x/x} + q(a_{x'/x'} + pa_{x'/x}) \\ a_{x'/x} & a_{x'/x'} + pa_{x'/x} \end{pmatrix} \end{aligned} \quad (13.14)$$

For a point-to-point focus, the final position of any point in the focal plane is independent of initial angle; therefore $b_{x/x'} = 0$, which leads to

$$p = -\frac{a_{x/x'} + qa_{x'/x}}{a_{x/x} + qa_{x'/x}} = \frac{4L/(R+1) + q(3R-1)/R^2}{(R-3) + (3/4)(q/L)(R+1)(R-1)^2/R^2} \quad (13.15)$$

While this ignores many details, it illustrates clearly the problem of picking an object distance p . If $q \geq 0$ (i.e. a focus downstream from the tube exit), then for $R = 1$, $p < 0$, but for $R > 3$, $p > 0$; therefore, between $R = 1$ and $R = 3$ the denominator passes through zero and p makes a transition from a virtual object infinitely far downstream to a real object infinitely far upstream. Lenses alone cannot easily correct for this. The ideal solution would be a lens coincident with the tube entrance whose strength can be varied to compensate for the changing focal length of the entrance aperture lens. The so-called “gridded lens” used in some tandem accelerators does just this by nullifying the natural aperture lens with a wire mesh grid while substituting a long-focal-length gap lens in its place [15]. Most single-stage accelerators keep R approximately constant by varying the injection energy. An acceleration gap (often the ion source extractor or a downstream gap lens) located where the beam is small permits the beam energy to be changed without significant movement of the object point as seen by the tube. A less common variant is to change L by shorting out portions of the accelerator tube as a coarse form of adjustment, narrowing the dynamic range of R and f_1 , and thereby permitting the injection energy to be kept high and more nearly constant [16].

13.4 Tandem Accelerators

A tandem electrostatic accelerator [17] consists of two acceleration stages that share a single high-voltage terminal but are separated by a charge-changing stripper. Negative ions injected into the low-energy (LE) stage accelerate to the terminal, change to positive at the stripper and accelerate a second time through the high-energy (HE) stage back to ground potential. Most tandems have a straight-through geometry; a few have been “folded” into a single-column structure by including a 180° reversing magnet in the terminal [18–21]. Except for some large machines equipped with mid-column lenses, the LE optics of tandems are characterized by (13.14). A focus (ultimately, a beam minimum) is desired at the stripper, only a short distance beyond the LE tube exit. Depending on the LE velocity gain ratio R_{LE} , the natural object location for this focus can range from far outside the accelerator to distressingly near the LE tube entrance.

13.4.1 Tandem Low-Energy Stage

Although the basics of electrostatic accelerator optics were known [22], tandems raised new challenges which had yet to be resolved. Typically, a lens outside the pressure vessel was used to get as much beam as possible into the first stage, but optical matching was alarmingly poor. A struggle ensued to improve the coupling of external sources to the tandem LE stage. The alternatives were to make efficient use of the entrance aperture lens, to alter it or to neutralize its effects. Table 13.1 offers a sampling of options;

Table 13.1. Types of beam coupling to a tandem LE stage

Type	Object	Requirements/comments
(a) Mid-column lens	Zero	Variable lens located in LE column
(b) Gap lens	Short	Used in many single-stage accelerators
(c) Constant gradient	Short	Parts of tube and column shorted
(d) Constant “Q”	Medium	Injection energy varied to match
(e) Gridded lens	Medium	Wire mesh grid on LE tube entrance
(f) Divergent lens	Medium	Requires grid in strong lens
(g) Step gradients	Medium	Gradient increased with energy
(h) Upstream lens(es)	Variable	Limited range of good matching
(i) Three-stage	Long	Ion source in injector terminal

some variants may utilize more than one entry. *Zero* object displacement is achieved by focusing the injected beam to a crossover at the LE tube entrance (effectively neutralizing the strong entrance aperture lens) and then using a mid-column lens to refocus at the stripper [23–25]. A *short* object displacement is the closest the object point approaches the LE tube entrance at the highest attainable terminal voltages; adjustments are made to the injection energy (b) or to the tube gradient (c) to keep this point fixed (or within a relatively narrow range) as terminal voltage is lowered. *Medium* displacement is achieved by injecting with energies that are kept proportional to the terminal voltage and that range into the hundreds of keV (d); by cancelling entrance fringing fields with a plane mesh grid that serves as one electrode of a weaker but still convergent gridded lens (e); by preceding the aperture lens with a strong divergent lens (for high terminal voltages), which converts to a convergent lens by reversal of polarity (for low terminal voltages) but which also requires the use of a grid (f); or by progressing in stages from lower to higher gradients (g). *Variable* object displacement (h) is the normal outcome when nothing has been done to alter the LE entrance aperture lens or stabilize it against changes in R_{LE} . By introducing more than one lens between the ion source and the tandem accelerator, a wider range of object lengths can be accommodated than with a single lens. *Long* (possibly negative) displacements are attained by injecting with another accelerator (i) that contributes substantially to the negative-ion energy and usually results in $R_{LE} < 2$. The relative merits of various of these possibilities have been explored by [16, 26–28].

Negative-ion injectors (NIIs) vary widely in design but tend to fall into two broad categories: those in which negative ions reach ground potential immediately after extraction (or after a charge-exchange process), and those mounted on an insulated high-voltage platform that delivers ions to ground potential through a “preacceleration” stage. In the first category, lenses built into the NII may be adequate to provide a focus outside; otherwise, the outgoing beam diverges. In the second category, the preacceleration stage

usually is operated much like a tandem HE stage since no penalty is incurred by crossing the beam near the preaccelerator tube entrance (often at a mass-selection aperture) and allowing it to diverge coming out. Outside, divergent beams are refocused by one or more external lenses, which may be axially symmetric electrostatic einzel lenses or electric or magnetic quadrupole lenses. Besides offering greater focal strength, quadrupoles also provide astigmatic correction for beams that emerge from the NII without axial symmetry as a consequence, for example, of passing through a mass-selection dipole. (However, from this distant upstream location, it is difficult to correct for astigmatism generated within the tandem by, for example, slot-shaped tube apertures in the LE stage.) Examples of studies of NII optics are found in [28–31].

13.4.2 Tandem High-Energy Stage

If, as is usually the case, the injection energy may be ignored, then the velocity gain ratio for the HE stage is simply

$$R_{HE} = [(V_T + qV_T)/V_T]^{1/2} = (1 + q)^{1/2}, \quad (13.16)$$

where V_T is the terminal voltage and q is a positive charge state of interest produced by stripping. The focal length for a circular aperture at the HE tube entrance is approximately

$$f_{1HE} = 4U_0L_{HE}/(U - U_0) = 4L_{HE}/(R_{HE}^2 - 1) = 4L_{HE}/q. \quad (13.17)$$

Thus, the problem of varying focal length reappears at the entrance to the HE stage as a consequence of different charge states produced by stripping rather than of changes in terminal voltage. The center of the charge state distribution (the most probable charge state) does rise with energy, initially as $E^{1/2}$ and progressively more slowly at higher energies [32]. Typically, f_{1HE} considerably exceeds the distance from stripper to HE tube; consequently, a convergent lens preceding the HE tube entrance offers improved transmission and becomes essential when a controlled focus is desired at a second stripper installed in the HE column. Additionally, if the HE tube begins with slot-shaped apertures, the astigmatism provided by a quadrupole matching lens can help compensate for aperture lens astigmatism.

Austerity prevails in small tandems, where gas strippers or a combination of gas and foil strippers may be the only components in the terminal that act on the beam. At the opposite extreme, the largest tandems have been equipped with a cornucopia of components, including devices that select one charge state from the many that emerge from a stripper. Noteworthy examples are the Daresbury vertical tandem [33], designed originally with offset LE and HE tubes linked by paired magnetic dipoles [25] and redesigned for axial geometry by changing to a displaced magnetic-quadrupole charge state

selector [34]; the Oak Ridge, folded tandem [24], in which the 180° reversing magnet performs charge selection, with electrostatic quadrupole lenses providing focusing and matching [35]; and the VIVITRON [36], in which a displaced quadrupole charge selector and matching lens were integrated into one electrostatic package that could be shoehorned into a comparatively short horizontal terminal with the charge selection aperture relocated farther downstream at a second stripper position inside the HE column [37].

13.5 Transporting the Beam to Targets

If the beam emerging from an electrostatic accelerator is not deposited directly on a target, then it is likely that it will undergo a change in angle during passage through one or more beam-bending, switching and/or analyzing dipoles. The ability to change beam direction is, by itself, a valuable utility, and multiport switching magnets facilitate directing the beam to more than one target station. In addition, dipoles perform the useful and oftentimes indispensable function of filtering out contaminant beams that originate in the ion source or arise later because of charge-changing processes. However, the first dipole encountered by a beam after acceleration is likely to be used to convert variations in beam momentum (magnetic dipole) or energy (electric dipole) into spatial displacements that are detected at slits and fed back to a system which corrects for errors in the accelerating voltage. If the angle of bend is small then the dipole acts as a relatively weak lens (see Sect. 13.2), and the beam can be focused directly through the dipole onto regulating slits. As previously mentioned, some accelerators (usually single-stage) have the capability to do this using internal lenses. More typically, however, an external lens is required. In either case, the dipole should be located close to the exit of the accelerator to minimize magnification at the slits. Conversely, if the angle of bend is relatively large, as is the case for the widely used double-focusing 90° analyzer, the beam must first be focused at the object point of the dipole; the dipole then acts as a strong lens to refocus onto the image slits. For this arrangement to work well with an external lens, the large-angle dipole must be located considerably farther away from the accelerator and a more powerful lens provided in order to achieve an intermediate object point crossover. To obtain sufficient focal strength, the external lenses usually must be quadrupole doublets or triplets. A quadrupole triplet will preserve more of the axial symmetry in an initially symmetric beam than will a doublet, but often the properties of a doublet are better matched to requirements. Either lens can produce a double focus even if the upstream object is mildly astigmatic, but a doublet magnifies more in one perpendicular plane than in the other (typically by about 2:1), and smaller magnification in the bending plane usually provides a better match to aperture constraints, as well as being highly desired to increase the resolution of the analyzing dipole.

Well-regulated electrostatic accelerators produce very little energy spread in the beam; consequently, small fluctuations in beam position caused by dispersive devices usually can be ignored. But residual dispersion may be intolerable for some applications, such as providing beam to a booster accelerator. Every change in beam direction adds or subtracts some dispersion and, when required, dispersion can be removed by balancing opposing contributions. Often this is done using mirror-symmetric configurations of dipoles and lenses such that dispersed rays converge back in mirror image onto the axis; however, neither mirror imaging nor lenses nor identical angles of bend are requisite, but at least two dipoles (including the one that initiated the dispersion) and two controlled parameters (one of which may be mirror symmetry) will be required to cancel, to first order, both the spatial ($a_x/\delta p$) and the angular ($a_{x'}/\delta p$) dispersion terms. Dispersion-control lenses typically produce point-to-point foci between the centers of dipoles (the pivot points from which dispersed rays appear to fan out), whereas the conventional beam focus may need to avoid these points (especially if a dipole is to serve as an analyzer). Satisfying all requirements is possible but challenging; see, for example, [10].

13.6 Accelerator Tube Matrices

13.6.1 Axial Accelerator Tube Model

At the core of any electrostatic accelerator is the accelerator tube, an evacuated region containing a strong longitudinal electric field. Wherever the field changes, focusing or defocusing occurs and this must be accounted for. Extracting analyses [38] show that weak modulations of the field caused by the finite thickness of tube electrodes and other internal details have appreciable effect, but often these are ignored because perturbations of comparable magnitude in the field distribution during operation remain largely unknown. Ordinarily, a variable lens system makes up for such shortcomings in the calculations. Selection and placement of such lenses constitutes an important part of optical design. Only when critical components have been positioned for good beam control and matching does it become worthwhile to shift the emphasis from idealized models to more detailed studies.

In order to evaluate velocities and times, let the particle kinetic energy be qU , where U is the potential difference required to raise the energy of a particle having charge q from zero to its present value. The total particle energy [39] is

$$E = mc^2 + qU = mc^2(1 + qU/mc^2) = \gamma mc^2, \quad (13.18)$$

where m is the (rest) mass and c the speed of light. The following conversions are convenient to work with at low kinetic energies, where qU (typically

measured in keV or MeV) is more likely to be known than the momentum or velocity:

$$\gamma = (1 - \beta^2)^{-1/2} = (\beta^2 \gamma^2 + 1)^{1/2} = 1 + qU/mc^2, \quad (13.19)$$

$$\beta = v/c = [(2qU/mc^2)(1 + (1/2)qU/mc^2)]^{1/2}/(1 + qU/mc^2), \quad (13.20)$$

$$\beta\gamma = (\gamma^2 - 1)^{1/2} = [(2qU/mc^2)(1 + (1/2)qU/mc^2)]^{1/2}. \quad (13.21)$$

In a region of electric field $\mathbf{E}$, the applied force is

$$\mathbf{F} = q\mathbf{E} = d\mathbf{p}/dt = d(\gamma m \mathbf{v})/dt = \gamma m d\mathbf{v}/dt + m \mathbf{v} \gamma^3 v c^{-2} dv/dt, \quad (13.22)$$

$$F_i = qE_i = \gamma m dv_i/dt + m v_i \gamma^3 v c^{-2} dv/dt, \quad i = x, y \text{ or } z. \quad (13.23)$$

For purely axial acceleration, $F_x = F_y = 0$; therefore, taking $i = x$ to represent x or y , (13.23) separates into functions of v_x and v , leading to

$$dv_x/v_x = -\gamma^2 v c^{-2} dv = -d\gamma/\gamma, \quad (13.24)$$

$$\int_{v_{x0}}^{v_x} v_x^{-1} dv_x = - \int_{\gamma_0}^{\gamma} \gamma^{-1} d\gamma, \quad (13.25)$$

$$\log(v_x/v_{x0}) = \log(\gamma_0/\gamma), \quad (13.26)$$

$$v_x = v_{x0} \gamma_0 / \gamma, \quad \text{or} \quad \gamma m v_x = \gamma_0 m v_{x0}. \quad (13.27)$$

In the absence of transverse forces, the transverse momenta $\gamma m v_x$ and $\gamma m v_y$ are separately conserved but transverse velocities v_x and v_y are not (except in the low-velocity limit). However, the divergence changes from x'_0 to

$$\begin{aligned} x' &= v_x/v_z = \gamma_0 v_{x0}/\gamma v_z = (\gamma_0 v_0/\gamma v)(v/v_z)(v_{z0}/v_0)(v_{x0}/v_{z0}) \\ &= x'_0/R_{\beta\gamma} + O(x'^2, y'^2), \end{aligned} \quad (13.28)$$

where

$$R_{\beta\gamma} = \beta\gamma/\beta_0\gamma_0 \xrightarrow{v \ll c} R = (U/U_0)^{1/2}. \quad (13.29)$$

Since $v_z/v = (1 - x'^2 - y'^2)^{1/2}$, and both x' and y' are presumed to be small, no significant error results from discarding higher-order terms and assuming $v_z/v \cong 1$. Note that, from (13.28), x' and y' depend only on x'_0 and y'_0 , and not on the transverse displacements x_0 and y_0 .

If separate versions of (13.23) for $i = x, y$ and z are multiplied by v_i/v and summed, then

$$\begin{aligned} q[(v_x/v)E_x + (v_y/v)E_y + (v_z/v)E_z] \\ &= \gamma m[(v_x/v) dv_x/dt + (v_y/v) dv_y/dt + (v_z/v) dv_z/dt] \\ &\quad + \gamma^3 m c^{-2} (v_x^2 + v_y^2 + v_z^2) dv/dt \\ &= \gamma m dv/dt + \gamma^3 m v^2 c^{-2} dv/dt = \gamma^3 m dv/dt. \end{aligned} \quad (13.30)$$

For accelerator tubes, E_z is the dominant term; therefore, let the left side of (13.30) be written as

$$\begin{aligned} & qE_z(v_z/v)[1 + (v_x/v_z)E_x/E_z + (v_y/v_z)E_y/E_z] \\ &= qE_z(v_z/v)(1 + x' \tan \theta_x + y' \tan \theta_y) = qkE_z, \end{aligned} \quad (13.31)$$

where $\theta_x = \arctan(E_x/E_z)$ and $\theta_y = \arctan(E_y/E_z)$ are fixed angles of inclination of fields with respect to the tube axis, $x' = v_x/v_z$ and $y' = v_y/v_z$ are the usual x and y divergences, and

$$k = (v_z/v)(1 + x' \tan \theta_x + y' \tan \theta_y) \quad (13.32)$$

adjusts for the orientation of a uniform electric field not aligned with z . Rearranging (13.30) leads to

$$dt = \gamma^3 m (qkE_z)^{-1} dv. \quad (13.33)$$

Although k is, in general, a function of velocity, the variation in k over an interval of acceleration may be small enough to allow k to be approximated as a constant; whereupon (13.33) can be integrated ([39], p. 139) from an initial state v_0 , t_0 before acceleration to a final state v , t to yield the duration of acceleration

$$\begin{aligned} t - t_0 &= \int_{t_0}^t dt = \int_{v_0}^v \gamma^3 m (qkE_z)^{-1} dv \\ &= m(qkE_z)^{-1}(\gamma v - \gamma_0 v_0) = mc(qkE_z)^{-1}(\beta\gamma - \beta_0\gamma_0). \end{aligned} \quad (13.34)$$

For axial acceleration, $\theta_x = \theta_y = 0$; consequently, from (13.32), $k = (v_z/v)$. But $v_z/v \rightarrow 1$ when $O(x'^2, y'^2)$ is neglected; therefore, $k = 1$ will be assumed from here on for the axial case.

Because prior history is not of concern, the choice of t_0 is arbitrary (v_0 is not); therefore, let $t_0 = 0$, and extract from (13.34)

$$\gamma(t) = (\beta^2\gamma^2 + 1)^{1/2} = [(qE_z m^{-1} c^{-1} t + \beta_0\gamma_0)^2 + 1]^{1/2}. \quad (13.35)$$

An integration over x may be performed by using (13.27) in the form $dx/dt = v_x = v_{x0}\gamma_0/\gamma$, assisted by $dt = mcq^{-1}E_z^{-1}d(\beta\gamma)$ obtained by differentiating (13.34) with $k = 1$:

$$\begin{aligned} x - x_0 &= \int_{x_0}^x dx = v_{x0}\gamma_0 \int_0^t \gamma^{-1}(t) dt \\ &= v_{x0}\gamma_0 mcq^{-1}E_z^{-1} \int_{\beta_0\gamma_0}^{\beta\gamma} (\beta^2\gamma^2 + 1)^{-1/2} d(\beta\gamma) \\ &= \{v_{x0}v_{z0}v_0 L\gamma_0/[v_{z0}v_0 c(\gamma - \gamma_0)]\} \\ &\quad \times \log[(\beta\gamma + \beta^2\gamma^2 + 1)^{1/2}/(\beta_0\gamma_0 + \beta_0^2\gamma_0^2 + 1)^{1/2}] \\ &= [(x'_0 L\beta_0\gamma_0)/(\gamma - \gamma_0)] \log[(\beta\gamma + \gamma)/(\beta_0\gamma_0 + \gamma_0)] \\ &\quad + O(x'^2, y'^2), \end{aligned} \quad (13.36)$$

where terms of $O(x'^2, y'^2)$ are to be discarded and E_z has been replaced, using (13.19), by

$$E_z = (U - U_0)/L = mc^2 q^{-1} L^{-1} (\gamma - \gamma_0). \quad (13.37)$$

For nearly all electrostatic-accelerator applications, a low-velocity approximation will suffice. Provided $\beta\gamma < 1$ (or, $\gamma - 1 = qU/mc^2 < 0.4$), the log function is expandable in powers of $\beta\gamma$ ([40], p. 45):

$$\begin{aligned} \log(\beta\gamma + \gamma) &= \log[\beta\gamma + (\beta^2\gamma^2 + 1)^{1/2}] \\ &= \beta\gamma - (1/2)(\beta^3\gamma^3/3) + [(1 \times 3)/(2 \times 4)](\beta^5\gamma^5/5) \\ &\quad - [(1 \times 3 \times 5)/(2 \times 4 \times 6)](\beta^7\gamma^7/7) + \dots \end{aligned} \quad (13.38)$$

$$\begin{aligned} \log[(\beta\gamma + \gamma)/(\beta_0\gamma_0 + \gamma_0)] \\ = (\beta\gamma - \beta_0\gamma_0)[1 - (1/6)(\beta^2\gamma^2 + \beta\gamma\beta_0\gamma_0 + \beta_0^2\gamma_0^2) + \dots]. \end{aligned} \quad (13.39)$$

Keeping only leading $\beta\gamma$ terms in (13.36), discarding $O(x'^2, y'^2)$, substituting for $R_{\beta\gamma}$ from (13.29) and remembering that $\beta^2\gamma^2 = \gamma^2 - 1$ leads to the following reduction:

$$\begin{aligned} x - x_0 &\cong x'_0 L \beta_0 \gamma_0 (\beta\gamma - \beta_0\gamma_0) / (\gamma - \gamma_0) \\ &= x'_0 L (\gamma + \gamma_0) / (R_{\beta\gamma} + 1) \xrightarrow{v \ll c} 2x'_0 L / (R + 1). \end{aligned} \quad (13.40)$$

To summarize, from (13.28), (13.29), (13.36) and (13.40), axial acceleration is described (for either x or y) by the first-order (linear) matrix

$$\begin{pmatrix} a_{x/x} & a_{x/x'} \\ a_{x'/x} & a_{x'/x'} \end{pmatrix} = \begin{pmatrix} 1 [(L\beta_0\gamma_0)/(\gamma - \gamma_0)] \log[(\beta\gamma + \gamma)/(\beta_0\gamma_0 + \gamma_0)] \\ 0 & \beta_0\gamma_0/\beta\gamma \end{pmatrix} \xrightarrow{v \ll c} \begin{pmatrix} 1 & 2L/(R + 1) \\ 0 & 1/R \end{pmatrix}. \quad (13.41)$$

13.6.2 Inclined-Field Accelerator Tube Model

Axially symmetric accelerator tubes are more susceptible to electrical breakdown than are tubes in which a component of the field is perpendicular to the axis. One method for generating such fields is to cant the internal electrodes at an angle of order 10° to 15° to produce a substantial transverse field component [41–43]. For convenience, assume that the transverse coordinate system has been rotated about the beam axis until an electric-field exists only in the x, z plane and that perpendicular to this plane v_y is negligibly small. After acceleration, any such rotation will have to be reversed to realign to the original axes. If $v_{\parallel}$ and $v_{\perp}$ are components of initial velocity parallel and perpendicular to the electric-field vector $\mathbf{E}$, and θ is the angle between the electric field and the tube axis then, by transformation of coordinates,

$$v_{\perp 0} = v_{x0} \cos \theta - v_{z0} \sin \theta = v_{z0} \cos \theta (x'_0 - \tan \theta), \quad (13.42)$$

$$v_{\parallel 0} = v_{x0} \sin \theta + v_{z0} \cos \theta = v_{z0} \cos \theta (1 + x'_0 \tan \theta). \quad (13.43)$$

In a coordinate system aligned to $\mathbf{E}$,

$$\begin{aligned} X'_{\perp 0} &= (v_{\perp}/v_{\parallel})_0 = (x'_0 - \tan \theta)/(1 + x'_0 \tan \theta) \\ &= x'_0/\cos^2 \theta - \tan \theta + O(x'^2, y'^2). \end{aligned} \quad (13.44)$$

In the $\perp, \parallel$ system, $E_{\parallel} = |\mathbf{E}|$ and $E_{\perp} = E_y = 0$; therefore, (13.23) transforms into

$$F_{\perp} = qE_{\perp} = \gamma m dv_{\perp}/dt + mv_{\perp} \gamma^3 v c^{-2} dv/dt, \quad (13.45)$$

which, as in (13.24) through (13.27), separates into functions of $v_{\perp}$ and v that can be integrated to yield the conservation of transverse momentum,

$$v_{\perp} = v_{\perp 0} \gamma_0 / \gamma, \quad \text{or} \quad \gamma m v_{\perp} = \gamma_0 m v_{\perp 0}. \quad (13.46)$$

After acceleration, the transformation back to v_x and v_z is

$$v_x = v_{\perp} \cos \theta + v_{\parallel} \sin \theta, \quad (13.47)$$

$$v_z = -v_{\perp} \sin \theta + v_{\parallel} \cos \theta. \quad (13.48)$$

Using (13.44) through (13.48), the ratio v_x/v_z may now be written as

$$\begin{aligned} x' &= (v_x/v_z) \\ &= [(v_{\perp}/v_{\parallel 0}) \cos \theta + (v_{\parallel}/v_{\parallel 0}) \sin \theta] / [-(v_{\perp}/v_{\parallel 0}) \sin \theta + (v_{\parallel}/v_{\parallel 0}) \cos \theta] \\ &= [X'_{\perp 0} + R_{\parallel} \tan \theta] / [R_{\parallel} - X'_{\perp 0} \tan \theta] \\ &= x'_0 R_{\parallel} (1 + \tan^2 \theta)^2 / (R_{\parallel} + \tan^2 \theta)^2 \\ &\quad + (R_{\parallel} - 1) \tan \theta / (R_{\parallel} + \tan^2 \theta) + O(x'^2, y'^2), \end{aligned} \quad (13.49)$$

where, based on the central ray trajectory, for which $v_{x0} = x'_0 = 0$ in (13.42) and (13.43), the ratio of momenta parallel to the electric vector before and after acceleration is

$$\begin{aligned} R_{\parallel} &= \gamma m v_{\parallel} / \gamma_0 m v_{\parallel 0} \\ &= (\gamma/\gamma_0) [v^2 / (v_{z0}^2 \cos^2 \theta) - (v_{z0}^2 \sin^2 \theta) / (v_{z0}^2 \cos^2 \theta)]^{1/2} \\ &= R_{\beta\gamma} [1 + (1 - \beta_0^2/\beta^2) \tan^2 \theta]^{1/2} + O(x'^2, y'^2) \\ &\xrightarrow{v \ll c} R [1 + (1 - R^{-2}) \tan^2 \theta]^{1/2}. \end{aligned} \quad (13.50)$$

Referring to (13.32) and applying (13.46), the correction factor k in the $v_{\perp}, v_{\parallel}$ plane becomes

$$k_{\parallel} = v_{\parallel}/v = (1 - v_{\perp}^2/v^2)^{1/2} = (1 - \gamma_0^2 v_{\perp 0}^2/\gamma^2 v^2)^{1/2}, \quad (13.51)$$

which is neither constant nor small, since typically $v_{\perp}/v \cong 1/4$. As a consequence, $k_{\parallel}(v)$ must be included inside the velocity integral in (13.34) when calculating the transit time, as follows:

$$\begin{aligned}
 t - t_0 &= \int_{t_0}^t dt = mq^{-1} E_{\parallel}^{-1} \int_{v_0}^v \gamma^3 k_{\parallel}^{-1} dv \\
 &= mq^{-1} E_{\parallel}^{-1} \int_{\gamma_0 v_0}^{\gamma v} (\gamma^2 v^2 - \gamma_0^2 v_{\perp 0}^2)^{-1/2} \gamma v d(\gamma v) \\
 &= mcq^{-1} E_{\parallel}^{-1} [(\beta^2 \gamma^2 - \gamma_0^2 v_{\perp 0}^2 / c^2)^{1/2} - (\beta_0^2 \gamma_0^2 - \gamma_0^2 v_{\perp 0}^2 / c^2)^{1/2}] \\
 &= mcq^{-1} E_{\parallel}^{-1} [\zeta - \zeta_0] , \tag{13.52}
 \end{aligned}$$

where

$$\begin{aligned}
 \zeta &= \beta_{\parallel} \gamma = (\beta^2 \gamma^2 - \gamma_0^2 v_{\perp 0}^2 / c^2)^{1/2} \\
 &= (\beta^2 \gamma^2 - \beta_{\perp 0}^2 \gamma_0^2)^{1/2} = (\gamma^2 - 1 - \beta_{\perp 0}^2 \gamma_0^2)^{1/2} , \tag{13.53}
 \end{aligned}$$

$$\gamma(\zeta) = (\zeta^2 + 1 + \beta_{\perp 0}^2 \gamma_0^2)^{1/2} , \tag{13.54}$$

$$dt = mcq^{-1} E_{\parallel}^{-1} d\zeta . \tag{13.55}$$

Guided by (13.36), an integration over $x_{\perp}$ may be performed using $dx_{\perp}/dt = v_{\perp} = v_{\perp 0} \gamma_0 / \gamma$ from (13.46), assisted by (13.42) as well as ζ , $\gamma(\zeta)$ and dt from above:

$$\begin{aligned}
 x_{\perp} - x_{\perp 0} &= \int_{x_{\perp 0}}^{x_{\perp}} dx_{\perp} = v_{\perp 0} \gamma_0 \int_0^t \gamma^{-1}(t) dt \\
 &= v_{\perp 0} \gamma_0 mcq^{-1} E_{\parallel}^{-1} \int_{\zeta_0}^{\zeta} (\zeta^2 + 1 + \beta_{\perp 0}^2 \gamma_0^2)^{-1/2} d\zeta \\
 &= [(x'_0 - \tan \theta) L \beta_0 \gamma_0 \cos^2 \theta / (\gamma - \gamma_0)] \\
 &\quad \times \log[(\zeta + \gamma) / (\zeta_0 + \gamma_0)] + O(x'^2, y'^2) , \tag{13.56}
 \end{aligned}$$

where, using (13.37), $E_{\parallel}$ has been replaced by

$$E_{\parallel} = E_z / \cos \theta = (U - U_0) / (L \cos \theta) = mc^2 (\gamma - \gamma_0) / (qL \cos \theta) . \tag{13.57}$$

Note that x'_0 is not cleanly separated out in (13.56), because ζ is a function of $\beta_{\perp 0}$, which, in (13.42), contains x'_0 .

The trajectory has now been tracked in the $\perp, \parallel$ plane from the equipotential U_0 , which crosses at $z = 0$, to the equipotential U , which crosses at $z = L$. Along the $x_{\parallel}$ axis, the original z axis has diverged by $-L \sin \theta$. To compensate for this, $L \sin \theta$ must be added to (13.56) to obtain the net displacement $\Delta_{\perp} = L \sin \theta + x_{\perp} - x_{\perp 0}$, along $x_{\perp}$. Because equipotentials are not orthogonal to the z axis, the endpoint will not, in general, lie in the perpendicular plane $z = L$ which bounds the accelerator tube. Since this rudimentary model does not include details of transitions from one inclination angle to

the next, the question of how a region of uniform inclined field should be terminated remains open. To continue acceleration until the equipotential plane intersecting $z = L$ is reached would change the final energy. To simply transform the $x_\perp, x_\parallel$ coordinates into x, z would leave $z_{final} \neq L$. To drift with divergence x' , until $z = L$, has some physical justification but seems unnecessarily complicated. To project parallel to $\mathbf{E}$, so that $x = \Delta_\perp / \cos \theta$, would be appropriate if the trajectories were essentially parallel to $\mathbf{E}$, but that is unlikely because trajectories are deliberately programmed to remain, on average, close to the axis and thus more nearly parallel to z . Instead, the procedure chosen here is to let

$$x - x_0 = \Delta_\perp \cos \theta = (L \sin \theta + x_\perp - x_{\perp 0}) \cos \theta, \quad (13.58)$$

implying $x_{\perp 0} = x_0 / \cos \theta$. In effect, this slides the final result parallel to z , as necessary, in order to arrive at $z = L$. Virtues of this imperfect choice are that sections having different inclination angles may be joined without passing information from section to section and that contiguous sections having the same angle join without perturbation.

By a change of variable from ζ to Z , and substituting $\gamma(\zeta)$ from (13.54), the log function in (13.56) may be expanded in powers of Z for $Z_0 \leq Z < 1$, in the same way as $\beta\gamma$ in (13.38) and (13.39):

$$Z = \zeta(1 + \beta_{\perp 0}^2 \gamma_0^2)^{-1/2}, \quad (13.59)$$

$$\begin{aligned} & \log[(\zeta + \gamma)/(\zeta_0 + \gamma_0)] \\ &= (Z - Z_0)[1 - (1/6)(Z^2 + ZZ_0 + Z_0^2) + \dots]. \end{aligned} \quad (13.60)$$

The leading term may be decomposed using (13.43), (13.50), (13.53), and (13.59) into

$$\begin{aligned} Z - Z_0 &= Z\zeta^{-1}(\zeta - \zeta_0) = Z\zeta^{-1}(\zeta^2 - \zeta_0^2)/(\zeta + \zeta_0) \\ &= Z\zeta^{-1}(\gamma^2 - \gamma_0^2)/[\beta_{\parallel 0}\gamma_0(R_\parallel + 1)] \\ &= (1 + \beta_{\perp 0}^2 \gamma_0^2)^{-1/2}(\gamma^2 - \gamma_0^2) \\ &\quad /[\beta_0\gamma_0(v_z/v_0) \cos \theta(1 + x'_0 \tan \theta)(R_\parallel + 1)]. \end{aligned} \quad (13.61)$$

Substituting (13.60) reduced to (13.61) into (13.56) and then that result into (13.58) yields the low-velocity approximation

$$\begin{aligned} x &\cong x_0 + L \sin \theta \cos \theta + [(x'_0 - \tan \theta)L\beta_0\gamma_0 \cos^3 \theta(\gamma - \gamma_0)(\gamma + \gamma_0)] \\ &\quad / \{(\gamma - \gamma_0)(1 + \beta_{\perp 0}^2 \gamma_0^2)^{1/2}[\beta_0\gamma_0 \cos \theta(1 + x'_0 \tan \theta)(R_\parallel + 1)]\} + O(x'^2, y'^2) \\ &\cong x_0 + L \sin \theta \cos \theta + [(x'_0 - \sin \theta \cos \theta)L(\gamma + \gamma_0)] \\ &\quad / [(1 + \beta_{\perp 0}^2 \gamma_0^2)^{1/2}(R_\parallel + 1)] + O(x'^2, y'^2) \\ &\xrightarrow{v \ll c} x_0 + L \sin \theta \cos \theta(R_\parallel - 1)/(R_\parallel + 1) + 2x'_0 L/(R_\parallel + 1) \end{aligned} \quad (13.62)$$

This result suggests a compromise that will extract x'_0 from within (13.56) for most applications. Let

$$H = [(\beta_0 \gamma_0 \cos \theta) / (\gamma - \gamma_0)] \log[(\zeta + \gamma) / (\zeta_0 + \gamma_0)] \xrightarrow{v \ll c} 2 / (R_{\parallel} + 1), \quad (13.63)$$

with the proviso that H is to be evaluated for a reference trajectory having $x'_0 = 0$. To balance against this, replace $x'_0 \cos^2 \theta$ in (13.56) with x'_0 , as happens in (13.62) for $\beta_{\perp 0}^2 \gamma_0^2 \ll 1$. In matrix form,

$$\begin{pmatrix} a_{x/x} & a_{x/x'} & a_{x/1} \\ a_{x'/x} & a_{x'/x'} & a_{x'/1} \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & LH & L \sin \theta \cos \theta (1 - H) \\ 0 & \frac{R_{\parallel}(1 + \tan^2 \theta)^2}{(R_{\parallel} + \tan^2 \theta)^2} & \frac{(R_{\parallel} - 1) \tan \theta}{R_{\parallel} + \tan^2 \theta} \\ 0 & 0 & 1 \end{pmatrix}. \quad (13.64)$$

The third column, containing $a_{x/1}$ and $a_{x'/1}$ elements, is introduced to accommodate particle displacements from a predefined geometric axis that are independent of initial position and angle. It also may be used to describe beam steerers [12], displaced quadrupole charge selectors [11], and steering effects caused, for example, by misalignment (deliberate or otherwise) of beam transport components. The matrix (13.64) supersedes results presented by the author in [44]. Reversals of the inclination angle within a tube module are required to keep the cumulative $a_{x/1}$ from growing too large. It is also desirable to exit each module with $a_{x/1} = a_{x'/1} = 0$. Interesting examples of this art may be found in [42–51].

In the orthogonal y, z plane, transverse momentum is unchanged by the inclined field; therefore, adapting (13.28),

$$y' = v_y / v_z = \gamma v_y / \gamma v_z = \gamma_0 v_{y0} / \gamma v_z = y'_0 / R_{\beta\gamma} + O(x'^2, y'^2). \quad (13.65)$$

During acceleration, γv_y behaves the same as $\gamma v_{\perp}$. After adjustment for the fact that coordinate transformations are not required, (13.56) through (13.62) serve as guidelines for $y - y_0$:

$$\begin{aligned} y - y_0 &= \int_{y_0}^y dy \\ &= v_{y0} \gamma_0 \int_0^t \gamma^{-1}(t) dt = v_{y0} \gamma_0 m c q^{-1} E_{\parallel}^{-1} \int_{\zeta_0}^{\zeta} (\zeta^2 + 1 + \beta_{\perp 0}^2 \gamma_0^2)^{-1/2} d\zeta \\ &= [(y'_0 L \beta_0 \gamma_0 \cos \theta) / (\gamma - \gamma_0)] \log[(\zeta + \gamma) / (\zeta_0 + \gamma_0)] + O(x'^2, y'^2) \\ &\xrightarrow{v \ll c} 2 y'_0 L / (R_{\parallel} + 1). \end{aligned} \quad (13.66)$$

Comparison of the above with (13.49) and (13.62) shows that, to within the approximations used in (13.64), the matrix for y is the same as (13.64), except that $a_{y/1} = a_{y'/1} = 0$.

References

1. C.J. Davisson, C.J. Calbick: Phys. Rev. **42**, 580 (1932)
2. A. Septier: Production of ion beams of high intensity. In: *Focusing of Charged Particles*, vol. 2, ed. by A. Septier (Academic Press, New York 1967) pp. 123–159
3. A. Nadji, F. Haas, G. Heng, C. Muller, R. Rebmeister: Nucl. Instr. Meth. A **287**, 173 (1990)
4. J.D. Larson, C.M. Jones: Nucl. Instr. Meth. **140**, 489 (1977)
5. E. Regenstreif: Focusing with quadrupoles, doublets, and triplets. In: *Focusing of Charged Particles*, vol. 1, ed. by A. Septier (Academic Press, New York 1967) pp. 353–410
6. D. DiChio, S.V. Natali, C.E. Kuyatt, A. Galejs: Rev. Sci. Instr. **45**, 566 (1974)
7. E.D. Courant, H.S. Snyder: Ann. Phys. **3**, 1 (1958)
8. S. Penner: Rev. Sci. Instr. **32**, 150 (1961)
9. H. Wollnik: Electrostatic prisms. In: *Focusing of Charged Particles*, vol. 2, ed. by A. Septier (Academic Press, New York 1967) pp. 163–202
10. H.A. Enge: Deflecting magnets. In: *Focusing of Charged Particles*, vol. 2, ed. by A. Septier (Academic Press, New York 1967) pp. 203–264
11. Z. Segalov: Nucl. Instr. Meth. **130**, 607 (1975)
12. J.D. Larson: Nucl. Instr. Meth. **189**, 71 (1981)
13. T.C. Randle: Nucl. Instr. Meth. **41**, 319 (1966)
14. B.A. Norman, W.H. Moore: Theory and design of beam transport systems. Brookhaven National Laboratory report BNL 12138 (1967)
15. N.B. Brooks, R.P. Bastide, K.H. Purser, M. Roos, P.H. Rose, A.B. Wittkower: IEEE Trans. Nucl. Sci. **NS-12**(3), 313 (1965)
16. A. Vermeer, C. Van der Leun, A.J. Veenenbos, P.E.P. Van der Vliet: Nucl. Instr. Meth. A **268**, 506 (1988)
17. L.W. Alvarez: Rev. Sci. Instr. **22**, 705 (1951)
18. H. Naylor: Nucl. Instr. Meth. **63**, 61 (1968)
19. H.R. McK. Hyder, P.J.S. Bromley-Barratt, T.R. Brock, G. Doucas, T.J.L. Greenway, A.R. Holmes, G.M. Parker, J. Takacs: Rev. Phys. Appl. **12**, 1331 (1977)
20. C.M. Jones: Rev. Phys. Appl. **12**, 1353 (1977)
21. G.A. Norton, M.L. Sundquist, R.E. Daniel, R.D. Rathmell: Nucl. Instr. Meth. **184**, 107 (1981)
22. M.M. Elkind: Rev. Sci. Instr. **24**, 129 (1953)
23. I. Ben-Zvi, M. Birk, E. Dafni, G. Hollos, R. Kaim, J.S. Sokolowski: Rehovot 14UD Pelletron – status report. In: *Proceedings of the 3rd International Conference on Electrostatic Accelerator Technology*, ed. by J.A. Martin (Oak Ridge, Tennessee 1981) pp. 59–61
24. C.M. Jones: Nucl. Instr. Meth. **184**, 145 (1981)
25. T. Joy, J.C. Lisle, W.R. Phillips: Interim report on beam optics calculations for the Nuclear Structure Facility. Daresbury Nuclear Physics Laboratory report DNPL/NSF/R3 (1972)
26. R. Hellborg, K. Håkansson, G. Skog: Nucl. Instr. Meth. A **287**, 161 (1990)
27. J.A. Van der Heide: Nucl. Instr. Meth. **95**, 87 (1971)
28. D.C. Weissler: Simple solutions to ion source matching. In: *Symposium of North Eastern Accelerator Personnel [SNEAP '90]*, ed. by T.N. Tipping, R.D. Krause (World Scientific, Singapore 1991) pp. 90–102, 171

29. F. Haas, G. Heng, J. Hoffmann, C. Muller, R. Rebmeister: Nucl. Instr. Meth. A **268**, 465 (1988)
30. E. Minehara, S. Hanashima: Nucl. Instr. Meth. A **268**, 461 (1988)
31. P. Spolaore, C. Signorini: First operation of the XTU-tandem 150 kV injector. In: *Proceedings of the 3rd International Conference on Electrostatic Accelerator Technology*, ed. by J.A. Martin (Oak Ridge, Tennessee 1981) pp. 65–67
32. K. Shima, T. Ishihara, T. Mikumo: Nucl. Instr. Meth. **200**, 605 (1982)
33. N.R.S. Tait: Nucl. Instr. Meth. **220**, 54 (1984)
34. D.A. Eastham, T. Joy, N.R.S. Tait: Nucl. Instr. Meth. **117**, 495 (1974)
35. W.T. Milner, G.D. Alton, D.C. Hensley, C.M. Jones, R.F. King, J.D. Larson, C.D. Moak, R.O. Sayer: IEEE Trans. Nucl. Sci. **NS-22**(3), 1697 (1975)
36. M. Letournel and the VIVITRON Group: Nucl. Instr. Meth. A **268**, 295 (1988)
37. E. Jegham, R. Rebmeister, J.D. Larson, A. Nadji: The Vivitron charge selector. In: *Symposium of North Eastern Accelerator Personnel [SNEAP '99]*, ed. by D.K. Hensley, N.L. Jones, R.C. Juras, M.J. Meigs, S.W. Mosko (World Scientific, Singapore 2000) pp. 11–21
38. P.H. Rose, A. Galejs: Nucl. Instr. Meth. **31**, 262 (1964)
39. T.M. Helliwell: *Introduction to Special Relativity* (Allyn and Bacon, Boston 1966)
40. R.S. Burington: *Handbook of Mathematical Tables and Formulas* (Handbook Publishers, Sandusky, Ohio 1954)
41. W.D. Allen: A new type of accelerating tube for electrostatic generators. National Institute for Research in Nuclear Science, NIRL/R/21 (1962)
42. R.J. Van de Graaff, P.H. Rose, A.B. Wittkower: Nature **195**, 1292 (1962)
43. K.H. Purser, A. Galejs, P.H. Rose, R.J. Van de Graaff, A.B. Wittkower: Rev. Sci. Instr. **36**, 453 (1965)
44. J.D. Larson: Nucl. Instr. Meth. **122**, 53 (1974)
45. J.G. Cramer: Nucl. Instr. Meth. **62**, 205 (1968)
46. B. Gyarmati, E. Koltay: Nucl. Instr. Meth. **66**, 253 (1969)
47. N.H. Merrill, S. Whineray: Nucl. Instr. Meth. **91**, 613 (1971)
48. M. Letournel, J.C. Oberlin, G. Heng: Nucl. Instr. Meth. **184**, 67 (1981)
49. J.D. Larson: Nucl. Instr. Meth. A **244**, 192 (1986)
50. X.-L. Guan: Nucl. Instr. Meth. A **268**, 376 (1988)
51. J.D. Larson: Beam optics tutorial. In: *Symposium of North Eastern Accelerator Personnel [SNEAP '99]*, ed. by D.K. Hensley, N.L. Jones, R.C. Juras, M.J. Meigs, S.W. Mosko (World Scientific, Singapore 2000) pp. 96–134

14 Beam Envelope Techniques for Ion-Optical Calculations

S. Bazhal¹ and R. Hellborg²

¹ SSC RF Institute for Physics and Power Engineering, 1 Bondarenko Sq.,
Obninsk, Kaluga Region, 249033 Russia
`bazhal@ippe.obninsk.ru`

² Department of Physics, Lund University, Sölvegatan 14, 223 62 Lund, Sweden
`ragnar.hellborg@nuclear.lu.se`

14.1 Introduction

The understanding of the analogy between propagation of light rays in a nonuniform medium and the motion of electrons in an electromagnetic field, outlined in the 1930s, served as a stimulus for the development of charged-particle optics. In the initial stage of its development, charged-particle optics mainly considered the problem of image acquisition by means of electrons. In a short time, this area of study named “electron optics” resulted in the development of the electron microscope. However, the problems of electron optics were not confined to development of a lens system for providing images with the help of electron beams. The expanding applications of electron and ion beams required the development of devices capable of providing the necessary control of beams and to deliver charged particles from their source to a distant target. The appearance and rapid growth of ion optics, for example, were mainly caused by the need for mass analysis, as well as by the development of charged-particle accelerators. To date, charged-particle optics embraces a wide range of problems connected with the application of electron and ion beams in different areas of science and technology. A huge number of publications have been devoted to this subject. Unfortunately, it is impossible to give even a brief overview of these publications in the framework of this chapter. Therefore we confine ourselves to references to the books [1–3], in which the principles of charged-particle optics are given along with an extensive bibliography.

One of the fields of application of ion optics is electrostatic accelerators. Ion-optical calculations play there (as in all other applications) a large role. The techniques for ion-optical calculations for electrostatic accelerators have undergone a very thorough change during their history of more than half a century (one of the first papers in which a detailed calculational analysis of an ion-optical system of an electrostatic accelerator was given was published in 1953 [4]). These techniques have come a long way from simple analytical calculations based on approximations of geometrical optics, towards sophisticated 3-D computer simulation. The rapid development of hardware will probably promote the creation of increasingly complicated software for

mathematical simulation of ion-optical systems. However, as was mentioned by J.D. Larson [5], application of complex numerical models without preliminary analytical calculations is fraught with the risk that the understanding of the results obtained could be lost. This observation, in spite of the impressive development of a system for data visualization, seems to remain valid in the future as well. Therefore it is essential to begin the solution of any ion-optical problems with simple analytical evaluations, then passing on to numerical calculations based on simplified mathematical models, and to use more complicated calculation techniques only at the last stage if necessary.

Electrostatic accelerators are widely used in various fields of scientific research and in industry. Quite naturally, the range of ion-optical problems arising at these accelerators is wide enough, and rather different approaches are necessary for their solution. Suffice it to mention ion microprobes and bunched beams as examples of the variety of these problems. At the same time, there is a problem that is general for the majority of applications. This is the beam transport problem.

Some examples of uncomplicated ion-optical calculations (both analytical and numerical) which can be used for electrostatic accelerators will be given in this chapter. The main attention is paid to methods for first-order calculations of beam transport taking into account the unordered spread of transverse speeds of ions, as well as the space charge forces. To describe a continuous monoenergetic ion beam, we use the concepts of four-dimensional phase space and of beam envelopes. According to Liouville's theorem, the phase volume occupied by the points representing the ion beam in the 4-D space of canonically conjugate coordinates and momenta is conserved. In the case of separation of variables in the equations of motion, projections of the phase volume on the planes of canonically conjugate variables are conserved as well. More often, however, to describe the transverse motion of ions, one considers the projection of the phase volume onto a plane of coordinates and angles, the plane XX' for example. The area of this projection divided by π , called the emittance ϵ , is an important characteristic of a charged-particle beam:

$$\epsilon = \frac{1}{\pi} \iint dx dx' \quad (14.1)$$

It is known that for a beam with finite (i.e. nonzero) emittance, one cannot point out a particle the trajectory of which could define a beam boundary. In this case the boundary is defined by an envelope of the ion trajectories, the determination of which is one of the principal problems for beams with finite emittance. To find the beam envelope in a linear approach one generally uses the matrix formalism or the solution of differential equations for the envelope. Detailed descriptions of these methods can be found in [6–9].

14.2 An Analytical Technique for Calculation of Ion Beam Envelopes

Let us consider some simple calculation techniques which can be useful in the first stage of designing the ion-optical system of an electrostatic accelerator, as well as for analysis of the optical behavior of elements of this system. The method we shall use was suggested by E.V. Shpak [10]. For an ion beam with a negligibly small current, this method allows one to find an analytical expression for the beam envelope from a family of trajectories passing through a boundary of the projection of the 4-D phase volume. In the original method [10], an analytical expression for the beam envelope is found as a function of the longitudinal coordinate z on the assumption that the beam has a crossover at the initial point z_0 (i.e. the beam in the phase plane is represented at this point by a straight ellipse).

We consider the method in more detail, having generalized it to initial conditions given by an elliptical phase contour with an arbitrarily sloped axis. We assume that at the initial point z_0 , the projection of the phase volume of the beam onto the plane XX' is bounded by an ellipse given by

$$\gamma x_0^2 + 2\alpha x_0 x'_0 + \beta (x'_0)^2 = \epsilon \quad (14.2)$$

The relations between the coefficients α , β and γ in (14.2) and the input characteristics of the beam (the envelope coordinate r_0 , the beam divergence r'_0 and the emittance ϵ) are given by the following:

$$\begin{aligned} \alpha &= -\frac{r_0 r'_0}{\epsilon} \\ \beta &= \frac{r_0^2}{\epsilon} \\ \gamma &= \frac{\epsilon^2 + (r_0 r'_0)^2}{\epsilon r_0^2} \end{aligned} \quad (14.3)$$

Let the projection of the ion trajectory on the coordinate plane XOZ be given by

$$x(z) = R_1(z)x_0 + R_2(z)x'_0 \quad (14.4)$$

where $R_1(z)$ and $R_2(z)$ are linearly independent solutions of the paraxial equation ($R_1(0) = 1$; $R'_1(0) = 0$; $R_2(0) = 0$; $R'_2(0) = 1$); and x_0 and x'_0 are the initial values of the transverse coordinate of the ion and the tangent of the angle between the ion trajectory and the longitudinal axis OZ , respectively. Equation (14.5), obtained from (14.2)–(14.4), gives a family of ion trajectories

$$x(z) - R_1(z)x_0 - R_2(z) \left(\frac{r'_0}{r_0} x_0 \pm \frac{\epsilon}{r_0^2} \sqrt{r_0^2 - x_0^2} \right) = 0 \quad (14.5)$$

The initial points of the trajectories (14.5) are located on the boundary phase contour given by (14.2). The beam envelope can be determined as the envelope of the family of the curves given by (14.5). After elimination of x_0 from

the simultaneous equations

$$\begin{aligned} F(x_0, z) &= 0 \\ \frac{\partial F(x_0, z)}{\partial x_0} &= 0 \end{aligned} \quad (14.6)$$

where $F(x_0, z)$ denotes the left-hand side of (14.5), we arrive at an expression for the calculation of the envelope,

$$r(z) = \pm \sqrt{(R_1(z)r_0 + R_2(z)r'_0)^2 + \left(\frac{R_2(z)\epsilon}{r_0}\right)^2} \quad (14.7)$$

Note that (14.7) agrees with a particular solution of the differential equation for the beam envelope given in [7]. Finding the derivative of r with respect to z , we arrive at an expression for the beam divergence r' :

$$r'(z) = \pm \frac{(R_1 r_0 + R_2 r'_0)(R'_1 r_0 + R'_2 r'_0) + R_2 R'_2 (\epsilon/r_0)^2}{r(z)} \quad (14.8)$$

(The signs “+” and “−” in (14.7) and (14.8) are related to the upper and lower branches of the envelope, respectively) And finally, equating the right-hand side of (14.8) to zero, one can find a position of the beam crossover.

We shall now consider two examples of the application of the analytical expression (14.7) to beam envelope calculation for elements of the ion-optical system of an electrostatic accelerator.

14.2.1 Focusing of an Ion Beam with Finite Emittance by an Accelerator Tube

Let us turn to Elkind’s classic work [4], in which a detailed analysis of beam focusing by an accelerator tube was given for the first time. The calculations represented in Elkind’s work were performed in the approach of geometrical optics (i.e. for a beam with zero emittance). Therefore they did not take into account the effect of an unordered spread of ion speeds on the beam focusing. We shall now solve Elkind’s problem for a beam with finite emittance.

Let the system for beam acceleration and transport consist of the following linear optical elements (Fig. 14.1): 1, a drift section between the plane of optical object and the accelerator tube; 2, a converging aperture lens at the entrance to the accelerator tube; 3, an uniform-field accelerator tube; 4 a diverging aperture lens at the exit of the accelerator tube; and 5, a drift section between the accelerator tube and the target on which the beam has to be focused [4]. The problem will be solved without taking into consideration the space charge forces. In this case the variables in the equations of motion are separated. Therefore the analysis can be confined to one of the coordinate planes (XOZ , for example).

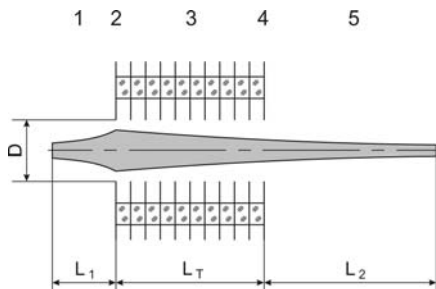


Fig. 14.1. For calculation of beam-focusing by an accelerator tube

Consider the motion of nonrelativistic ions in the drift section following the accelerator tube (Fig. 14.1, part 5). In this section the projection of the trajectory on the coordinate plane takes the form of (14.4). Using matrix formalism, we now determine $R_1(z)$ and $R_2(z)$ in (14.4). The transfer matrix of the whole system (parts 1–5 in Fig. 14.1),

$$\mathbf{R}_S = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \quad (14.9)$$

can be expressed as a matrix product of its individual ion-optical elements:

$$\mathbf{R}_S = \mathbf{R}_2 \mathbf{R}_D \mathbf{R}_T \mathbf{R}_F \mathbf{R}_1 \quad (14.10)$$

where

$$\mathbf{R}_1 = \begin{pmatrix} 1 & L_1 \\ 0 & 1 \end{pmatrix} \quad (14.11)$$

is the matrix of the drift section located before the accelerator tube (Fig. 14.1, part 1);

$$\mathbf{R}_F = \begin{pmatrix} 1 & 0 \\ 1/f_1 & 1 \end{pmatrix} \quad (14.12)$$

is the matrix of the input aperture lens (Fig. 14.1, part 2);

$$\mathbf{R}_T = \begin{pmatrix} 1 & 2L_T/(\sqrt{N} + 1) \\ 0 & 1/\sqrt{N} \end{pmatrix} \quad (14.13)$$

is the matrix of the uniform-field accelerator tube (Fig. 14.1, part 3);

$$\mathbf{R}_D = \begin{pmatrix} 1 & 0 \\ 1/f_2 & 1 \end{pmatrix} \quad (14.14)$$

is the matrix of the output aperture lens (Fig. 14.1, part 4); and, finally,

$$\mathbf{R}_2 = \begin{pmatrix} 1 & L_2 \\ 0 & 1 \end{pmatrix} \quad (14.15)$$

is the matrix of the drift section positioned after the accelerator tube (Fig. 14.1, part 5).

The following notation is used for the matrix elements: $N = \Phi_l/\Phi_0$ is the ratio of the potential of the last electrode of the accelerator tube to the potential of the first tube electrode; f_1 and f_2 are the focal lengths of the input and output aperture lenses; and L_1 , L_T and L_2 are the lengths of the first drift section, of the accelerator tube and of the second drift section, respectively (see Fig. 14.1).

The focal length of the aperture lens can be approximated by $f = 4\Phi\xi/(E_1 - E_2)$, where ξ is a function of $\Phi/(E_1 - E_2)$ and the aperture diameter D [4], and E_1 and E_2 are the fields preceding and following the aperture, respectively. For the weak output lens, this dependence can be neglected, assuming $\xi = 1$. Then, using the ratio N defined above, we arrive at

$$f_1 = -4\xi L_T/(N - 1) \quad (14.16)$$

$$f_2 = 4NL_T/(N - 1) \quad (14.17)$$

Taking into consideration (14.10)–(14.17), one can write the matrix elements R_{11} and R_{12} as

$$R_{11} = \frac{3(N - 1)(\xi - \sqrt{N})L_2}{8N\xi L_T} - \frac{\sqrt{N} - 1}{2\xi} + 1 \quad (14.18)$$

$$R_{12} = \frac{L_2}{2N} \left\{ \frac{L_1(N - 1)}{2L_T} + (3\sqrt{N} - 1) \left[1 - \frac{L_1(N - 1)}{4\xi L_T} \right] \right\} + \frac{2L_T(\sqrt{N} - 1)}{N - 1} \left[1 - \frac{L_1(N - 1)}{4\xi L_T} \right] + L_1 \quad (14.19)$$

We suppose now that the beam is focused into a waist with a radius of r_0 at a distance of L_1 from the entrance aperture lens. Let us find the distance L_2 from the exit aperture at which the output crossover is shaped. Assuming the plane of this crossover to be the end of the ion-optical system under consideration, one can write $R_1(z) = R_{11}$ and $R_2(z) = R_{12}$. Having substituted (14.18) and (14.19) into (14.7), we differentiate it with respect to z , taking into account the obvious relation between variables z and L_2 : $z = L_1 + L_T + L_2$. Solving the equation obtained in such a way with respect to L_2 , we determine the position of the output crossover

$$L_2 = \frac{4NL_T(\sqrt{N} - 1 - 2\xi)(r_0^4 + \epsilon^2 S_1 S_2)}{3(N - 1)(\xi - \sqrt{N})(r_0^4 + \epsilon^2 S_2^2)} \quad (14.20)$$

where $S_1 = 4L_T\xi/((\sqrt{N} + 1)(\sqrt{N} - 1 - 2\xi)) - L_1$ and $S_2 = 4L_T\xi(3\sqrt{N} - 1)/(3(N - 1)(\sqrt{N} - \xi)) - L_1(3\sqrt{N} - 1 - 2\xi)/(3(\sqrt{N} - \xi))$. In the extreme case of a beam of zero emittance, emerging from a point source on the optical axis (i.e. for $\epsilon \rightarrow 0$ and $r_0 \rightarrow 0$), (14.20) transforms to Elkind's formula:

$$L_2 = 4NL_T \frac{(\sqrt{N} - 1 - 2\xi) - 4\xi(L_T/L_1)/(\sqrt{N} + 1)}{4\xi(L_T/L_1)(3\sqrt{N} - 1) - (N - 1)(3\sqrt{N} - 1 - 2\xi)} \quad (14.21)$$

As a rule, $N \gg 1$ for electrostatic accelerators. Then a condition for existence of a beam crossover in the drift space following the accelerator tube is given by the inequality $S_1 S_2 < -r_0^4/\epsilon_0^2$. This inequality defines more rigid constraints for the position of the input crossover in comparison with the analogous inequality $S_1 S_2 < 0$ obtained from Elkind's formula (14.21).

One of the ion-optical problems for electrostatic accelerators lies in matching the beam to the accelerator tube. By defining a relation between the positions of the input and output crossovers, (14.20) allows us to perform estimations necessary for this problem to be solved. These estimations will be done using the low-energy accelerator tube of the 3 MV Pelletron tandem accelerator in Lund as an example. The beam-matching problem at this accelerator has already been solved earlier by a matrix method [11]. By analogy with this work, we employ in our example a simplified approximation for the input aperture lens, assuming ξ in (14.20) to be equal to unity. We shall also assume that the electrostatic field in the accelerator tube ($L_T = 1.694$ m) is uniform. Figure 14.2 gives the relationship $L_2 = f(L_1)$ calculated for the fixed ratio $N = 60$. Here the variable parameters are the radius r_0 of the beam in the plane of the input crossover and the normalized emittance ϵ_n (Fig. 14.2, curves 1–4). Curve 5, giving the relationship between the positions

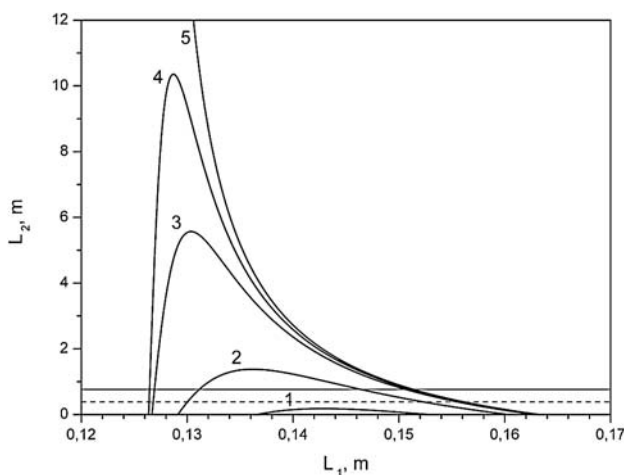


Fig. 14.2. Relationships between the positions of the input (L_1) and output (L_2) crossovers calculated for the low-energy part of the accelerator tube of the 3 MV Pelletron accelerator in Lund. 1, $\epsilon_n = 3\pi \times 10^{-6}$ m rad (MeV) $^{0.5}$, $r_0 = 0.5 \times 10^{-3}$ m; 2, $\epsilon_n = 5\pi \times 10^{-6}$ m rad (MeV) $^{0.5}$, $r_0 = 0.5 \times 10^{-3}$ m; 3, $\epsilon_n = 3\pi \times 10^{-6}$ m rad (MeV) $^{0.5}$, $r_0 = 0.25 \times 10^{-3}$ m; 4, $\epsilon_n = 5\pi \times 10^{-6}$ m rad (MeV) $^{0.5}$, $r_0 = 0.25 \times 10^{-3}$ m; 5, $\epsilon_n = 0$, $r_0 = 0$ (the calculation performed in accordance with Elkind's approximation)

of an optical object located in the drift space before the accelerator tube and its image created by the tube was calculated in Elkind's approach (14.21). For the purpose of more detailed analysis, the drift space after the accelerator tube has not been constrained to the actual drift section under the high-voltage terminal of the Pelletron accelerator (the boundary of this drift section is shown in Fig. 14.2 as a solid horizontal line). Consider the focusing of the beam onto the center of the stripper (the center of the stripper channel is marked by a dashed line). As can be seen from Fig. 14.2, the functional dependences $L_2 = f(L_1)$ plotted for a beam with finite emittance (curves 1–4 in Fig. 14.2) have a maximum. This highest possible value of L_2 decreases with increasing beam radius r_0 in the plane of the input crossover and with decreasing emittance ϵ_n . The curve 1 is completely below the dashed line, i.e. in this case the crossover cannot be obtained at the center of the stripper. The dashed line, as can be seen in Fig. 14.2, crosses the curves 2–4 twice, giving two values of L_1 for which the output crossover is in the center of the stripper. However, the lower of these two values of L_1 corresponds to an impermissible large optical magnification and therefore cannot be applied in practice. To obtain a beam crossover at the stripper, the beam needs to have a rather small radius in the plane of the input crossover and, in addition, the permissible position for this plane is constrained to quite a small part of the input drift section. The calculation and experimental results given in [11] justify this conclusion. Thus the ion-optical behavior of the system under consideration imposes a rigid limitation upon the input characteristics of the beam.

14.2.2 Application of the Beam Envelope Technique for Acceptance Calculations

Let us consider an application of the beam envelope technique to calculation of the acceptance of a part of an ion-optical system confined between two apertures. Although this problem had been already discussed by J.D. Larson and C.M. Jones in detail [12], we decided to return to the problem for the following reasons. Firstly, in the case of an elliptical phase contour, the envelope technique allows somewhat of a simplification of the derivation of the main formula for the acceptance calculation. Secondly, the method suggested gives a possibility of graphical representation of the results in the phase plane.

Then let a part of some ion-optical channel be confined between two planes z_0, z_1 . Let r_0, r_1 be, correspondingly, the radii of the input and the output apertures constraining the beam in these planes. For this part of the system, we shall derive an analytical expression for the acceptance making use of (14.7). If $R_2(z_1) = 0$, i.e. the system under consideration forms a Gaussian image in the plane z_1 , then (14.7) takes the form $r_1 = |R_1(z_1)|r_0$. In this case the transverse dimension of the beam in the image plane does not depend on the beam emittance, and it is determined by the absolute value of the

optical magnification of the system. Assuming $R_2(z_1) \neq 0$, we can express the emittance ϵ from (14.7) as

$$\epsilon = \sqrt{\left(\frac{r_0 r_1}{R_2(z_1)}\right)^2 - \left(\frac{r_0[R_1(z_1)r_0 + R_2(z_1)r'_0]}{R_2(z_1)}\right)^2} \quad (14.22)$$

Equations (14.2), (14.3) and (14.22) determine in the phase plane a family of ellipses of variable area $\pi\epsilon$. The ellipses exist under conditions given by the inequality $r_1^2 - (R_1 r_0 + R_2 r'_0)^2 > 0$. Its solution establishes the variation limits for the beam divergence r'_0 :

$$\begin{aligned} -(R_1/R_2)r_0 - (|r_1|/R_2) < r'_0 < -(R_1/R_2)r_0 + (|r_1|/R_2) ; R_2 > 0 \\ -(R_1/R_2)r_0 + (|r_1|/R_2) < r'_0 < -(R_1/R_2)r_0 - (|r_1|/R_2) ; R_2 < 0 . \end{aligned} \quad (14.23)$$

The emittance reaches its maximum value when the second term under the square root in (14.22) is equal to zero, i.e.

$$r'_0 = -\frac{R_1}{R_2}r_0 \quad (14.24)$$

The upper bound of the emittance determines the acceptance A ; it follows directly that

$$A = \left| \frac{r_0 r_1}{R_2} \right| \quad (14.25)$$

Taking into account that the emittance has been defined in the present work as the area of a phase ellipse divided by π , (14.25) coincides with the expression derived in [12]. Substituting (14.24) and (14.25) in the coefficients of (14.3), we arrive at the equation representing the acceptance in the phase plane:

$$\left(\frac{r_1^2 + r_0^2 R_1^2}{r_0^2 r_1^2}\right)x^2 + \frac{2R_1 R_2}{r_1^2}xx' + \frac{R_2^2}{r_1^2}(x')^2 = 1 \quad (14.26)$$

As an example we present acceptance calculations performed for the mass analyzer of the injector at the Pelletron accelerator in Lund [13]. The mass analyzer is a uniform-field dipole magnet with a bending angle $\varphi = 90^\circ$, a bending radius $\rho_m = 0.3735$ m and an angle of pole edge rotation $\beta = 28.2^\circ$. The cross section of the vacuum chamber of the mass analyzer has the following dimensions: 80 mm in the dispersive plane and 39 mm in the non-dispersive plane. The planes of the optical object and image are at equal distances from the boundaries of the magnetic field. A four-blade input aperture is positioned in the object plane of the magnetic analyzer. A two-blade slit device is installed in the image plane.

Since the trajectory R_2 crosses the optical axis in the image plane (i.e. $R_2 = 0$ in this plane), the analyzer acceptance does not depend on the aperture of the slit device. It is defined only by the input aperture which is used as the first diaphragm in (14.25) and by the aperture of the vacuum chamber

of the analyzer. For a wide range of initial beam conditions (beam radius, divergence and emittance), the beam envelope reaches a maximum inside the analyzer. Therefore, to use the transverse dimensions of the vacuum chamber of the analyzer in acceptance calculations performed by the “two diaphragm” formula (14.25), we have to find a cross section of the vacuum chamber that represents the chamber constraints equivalently and hence could be considered as the second diaphragm.

Consider first the dispersive (XOZ) plane of the magnetic analyzer. In this plane, the beam is defocused by the fringe fields and focused by the uniform magnetic field. Focusing is the resulting effect. Therefore, it is most natural to connect the position of the second diaphragm with a possible maximum of the beam envelope in the uniform-field region. Then, using (14.8), we can find this maximum from the equation

$$(R_1 r_0 + R_2 r'_0)(R'_1 r_0 + R'_2 r'_0) + \left(\frac{\epsilon}{r_0}\right)^2 R_2 R'_2 = 0 \quad (14.27)$$

Taking into account that $R_1 r_0 + R_2 r'_0 = 0$ for the beam of the maximum possible emittance, and that the image plane is outside the magnetic field (i.e. $R_2 \neq 0$ in the field region), we arrive at the condition for the extreme of the R_2 trajectory,

$$R'_2 = 0 \quad (14.28)$$

Thus, the position of the second diaphragm coincides with the extreme of the trajectory $R_2(z)$ determined by a linearly independent solution of the paraxial equation.

To find the trajectory $R_2(z)$ in the field region of the magnetic analyzer, we use the matrix formalism. The transfer matrix of the system, consisting of a drift section between the input aperture and the effective boundary of the magnetic field, a thin lens describing the effect of the fringe field, and the sector magnetic field can be expressed as a matrix product of these individual elements:

$$\mathbf{R} = \mathbf{R}_M \mathbf{R}_L \mathbf{R}_{Dr} \quad (14.29)$$

where

$$\mathbf{R}_M = \begin{pmatrix} \cos(z/\rho_m) & \rho_m \sin(z/\rho_m) \\ -\rho_m^{-1} \sin(z/\rho_m) & \cos(z/\rho_m) \end{pmatrix} \quad (14.30)$$

$$\mathbf{R}_L = \begin{pmatrix} 1 & 0 \\ \rho_m^{-1} \tan \beta & 1 \end{pmatrix} \quad (14.31)$$

$$\mathbf{R}_{Dr} = \begin{pmatrix} 1 & L \\ 0 & 1 \end{pmatrix} \quad (14.32)$$

are the matrices of the uniform magnetic field, of the fringing-field lens and of the drift section of length L , respectively. Multiplication of the matrices yields

$$R_2(\theta) = L \cos \theta + (L \tan \beta + \rho_m) \sin \theta \quad (14.33)$$

where $\theta = z/\rho_m$. The trajectory $R_2(\theta)$ attains an extreme at

$$\theta_e = \arctan \left(\tan \beta + \frac{\rho_m}{L} \right), \quad |\theta| \leq 90^\circ \quad (14.34)$$

After substituting θ_e in (14.33), we find this extreme of the trajectory R_2 :

$$R_2(\theta_e) = L \sqrt{1 + \left(\tan \beta + \frac{\rho_m}{L} \right)^2} \quad (14.35)$$

In the field region of the 90° magnet, the extreme is a maximum for any L , since $R_2 > 0$ and $R_2'' = -R_2$ at the point of the extreme. Taking into consideration the condition for radial focusing in a uniform sector magnetic field [3],

$$\sin \varphi + 2 \frac{L \cos(\varphi - \beta)}{\rho_m \cos \beta} - \left(\frac{L}{\rho_m \cos \beta} \right)^2 \sin(\varphi - 2\beta) = 0 \quad (14.36)$$

as well as equality of the distances from the field boundaries to the object and image planes, one can express L through the geometrical parameters of the analyzer:

$$L = \frac{\rho_m}{1 - \tan \beta} \quad (14.37)$$

And finally, substituting (14.37) in to (14.35), we arrive at an expression for the extreme of the trajectory R_2 :

$$R_2 = \frac{\rho_m \sqrt{2}}{1 - \tan \beta} \quad (14.38)$$

R_1 can also be expressed through the matrix product given by (14.29). Substituting the value of θ_e in this expression at the point of the extreme of the trajectory R_2 yields

$$R_1 = \frac{1 + \tan \beta}{\sqrt{2}} \quad (14.39)$$

Equations (14.38) and (14.39) and the dimensions of the two diaphragms determine an acceptance area in the phase plane. In accordance with (14.25), the value of the acceptance in the dispersive plane of the magnetic analyzer can be calculated from:

$$A = \frac{r_{x0} r_{x1} (1 - \tan \beta)}{\rho_m \sqrt{2}} \quad (14.40)$$

where $2r_{x0}$ and $2r_{x1}$ define the input aperture and the aperture of the vacuum chamber of the magnetic analyzer, respectively.

In the nondispersive (YOZ) plane of the magnetic analyzer, only the fringe field lenses act as focusing elements. In the first-order approach, the

absolute value of the focal length of the lenses is the same for both of the transverse planes. Therefore the transfer matrices $\mathbf{R}_M$ and $\mathbf{R}_L$ can be written as

$$\mathbf{R}_M = \begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix} \quad (14.41)$$

$$\mathbf{R}_L = \begin{pmatrix} 1 & 0 \\ -\rho_m^{-1} \tan \beta & 1 \end{pmatrix} \quad (14.42)$$

Substitution of (14.41) and (14.42) in (14.29) yields

$$R_1 = 1 - \frac{z \tan \beta}{\rho_m} \quad (14.43)$$

$$R_2 = \frac{\rho_m}{1 - \tan \beta} + z \frac{1 - 2 \tan \beta}{1 - \tan \beta} \quad (14.44)$$

In the magnetic-field region, the trajectory R_2 neither attains the extremes nor crosses the optical axes. Taking into account that $1 - 2 \tan \beta < 0$ in (14.44), we arrive at the expressions for the acceptance calculation

$$R_1 = 1; \quad R_2 = \frac{\rho_m}{1 - \tan \beta}; \quad A = \frac{r_{y0} r_{y1} (1 - \tan \beta)}{\rho_m} \quad (14.45)$$

Finally, let us consider a numerical example. Let the opening of the four-blade input aperture be $5 \times 5 \text{ mm}^2$. To reduce the unfavorable effect of aberration, we require that the beam dimensions in two transverse planes do not exceed half of the aperture of the vacuum chamber of the analyzer. This requirement results, evidently, in smaller acceptances in comparison with those which are defined by the geometrical constraints of the chamber. Taking into consideration that $2r_{x1} = 40 \text{ mm}$ and $2r_{y1} = 19.5 \text{ mm}$, we obtain the following values of these “conditional” acceptances: $A_x = 44 \text{ mm mrad}$ for the dispersive plane and $A_y = 30 \text{ mm mrad}$ for the nondispersive plane. Contours of these acceptances calculated with the help of (14.26) are given in Fig. 14.3.

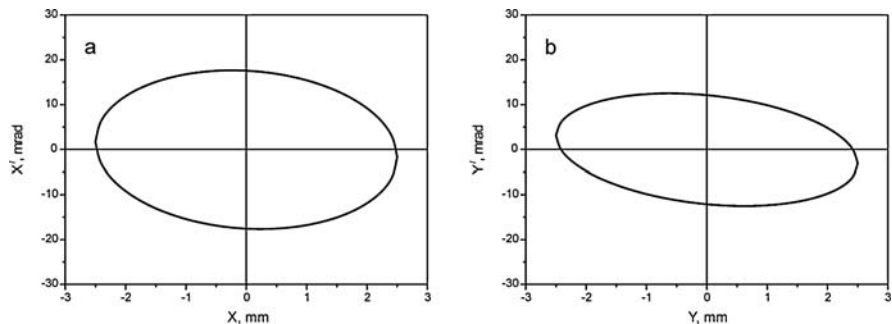


Fig. 14.3. Acceptances of the magnetic analyzer in the new injector leg of the Lund Pelletron accelerator, calculated (a) for the dispersive plane and (b) for the nondispersive plane

14.3 Differential Equations for Beam Envelopes with the Kapchinskiy–Vladimirskiy Density Distribution

Beam envelope calculations performed in a first-order approach are among the most simple and, at the same time, a rather effective way to obtain information about an ion beam with finite emittance. These calculations allow one to find the semiaxis r of the transverse cross section of the beam and the beam divergence r' as functions of the longitudinal coordinate z . Consider a method based on numerical solution of the differential equation for a beam envelope with the Kapchinskiy–Vladimirskiy density distribution [7, 9]. This method is suitable for ion-optical calculations for electrostatic accelerators as the beams for these accelerators have a moderate divergence and relatively low intensity.

Let us assume that the ion-optical channel under consideration includes the following elements: drift sections, axially symmetric electrostatic lenses, sections of the accelerator tube with a uniform field, magnetic and electrostatic quadrupole lenses, dipole analyzing magnets and spherical electrostatic analyzers. To exclude the second derivative of the axial potential $\Phi(z)$ from the envelope equations, we make use of Picht's substitution [14]

$$r_{x,y} = R_{x,y} \Phi^{-0.25} \quad (14.46)$$

which expresses projections of the beam envelope $r_{x,y}$ on the planes XOZ and YOZ by way of the auxiliary variables $R_{x,y}$. If, in (14.46), instead of the potential Φ , one uses the kinetic energy W (expressed in electron volts), the equations for the envelopes of the nonrelativistic ion beam can be written in the following general form [15]:

$$\begin{aligned} R_x'' &= kI/[W(R_x + R_y)] + (W_0\epsilon_x^2)/R_x^3 - (3/16)(W'/W)^2 R_x \pm \omega_x^2 R_x \\ R_y'' &= kI/[W(R_x + R_y)] + (W_0\epsilon_y^2)/R_y^3 - (3/16)(W'/W)^2 R_y \pm \omega_y^2 R_y \end{aligned} \quad (14.47)$$

where I is the beam current in amperes, ϵ_x and ϵ_y denote transverse emittances in mrad, and W_0 is the initial value of the kinetic energy of the ions. The coefficient k is defined by (14.48):

$$k = \frac{Z_i}{2\pi\epsilon_0} \sqrt{\frac{A}{2\eta_0}} \quad (14.48)$$

where Z_i and A stand for the ion charge state and mass number, respectively; $\epsilon_0 = 8.85 \cdot 10^{-12}$ F/m; and $\eta_0 = 0.958 \times 10^8$ C/kg. The coefficients $\omega_{x,y}^2$ describe focusing (–) and defocusing (+) effects of the linear ion-optical elements. For the different ion-optical elements mentioned above, the coefficients $\omega_{x,y}^2$ have the following values and expressions:

- **Drift space:** $\omega_{x,y}^2 = 0$.
- **Axially symmetric electrostatic lens:** $\omega_{x,y}^2 = 0$.
- **Uniform electrostatic field:** $\omega_{x,y}^2 = 0$.
- **Electrostatic quadrupole lens:** $\omega_{x,y}^2 = (Z_i U)/(a^2 W)$, where U is the voltage applied to the lens electrodes in V, and a is the aperture radius of the lens in m.
- **Magnetic quadrupole lens:** $\omega_{x,y}^2 = Z_i G_{x,y} (\eta_0/(2AW))^{0.5}$, where $G_{x,y}$, in T/m, stands for the gradients of the magnetic flux density.
- **Magnetic analyzer:** $\omega_x^2 = (1 - n)/\rho_m^2$ in the dispersive plane, and $\omega_y^2 = n/\rho_m^2$ in the nondispersive plane, where n is a field index and ρ_m is the bending radius in m.
- **Spherical electrostatic analyzer:** $\omega_x^2 = 1/\rho_e^2$ in the dispersive plane, and $\omega_y^2 = 1/\rho_e^2$ in the nondispersive plane, where ρ_e is the bending radius in m.

In the framework of this method, the electrostatic field on node points of the ion-optical axis is determined by the numerical solution of the Dirichlet problem

$$\frac{\partial^2 \Phi}{\partial r^2} + \frac{1}{r} \frac{\partial \Phi}{\partial r} + \frac{\partial^2 \Phi}{\partial z^2} = 0, \quad \Phi_S = U_i \quad (14.49)$$

where U_i is the potential of the i th electrode. The continuous distribution of the axial potential is approximated by a cubic spline.

The need for calculation of a multielectrode axially symmetric electrostatic lens is rather common in studies of the ion-optical system of an electrostatic accelerator. If the potentials of the electrodes attain only two independently variable values, then, having solved the Dirichlet problem for a pair of arbitrarily selected unequal potentials, one can easily determine potentials on the optical axis for any regime of the lens. In the case of more than two independently variable potentials the field calculation becomes complicated, inasmuch as one solution to the boundary problem is now insufficient for all regimes of the lens to be described. In principle it is possible to seek an individual solution to the Dirichlet problem for each of those regimes. However, such an approach to the problem looks rather unpractical from a computational point of view. Another way is to apply the superposition principle to some totality of solutions to the Dirichlet problem. In this case the potential on the optical axis of an n -electrodes lens can be expressed in the following way:

$$\Phi(r, z) = \sum_{i=1}^{n-1} (U_i - U_{i+1}) \Phi_i + U_n \Phi_n \quad (14.50)$$

where $\Phi_1(r, z) - \Phi_n(r, z)$ are solutions to the Dirichlet problem obtained for the linearly independent vectors of the boundary condition $\mathbf{S}_1(1, 0, 0, \dots, 0)$, $\mathbf{S}_2(1, 1, 0, \dots, 0), \dots, \mathbf{S}_n(1, 1, 1, \dots, 1)$; $U_1, U_2, \dots, U_n$ stand for the potentials of the lens electrodes. Some examples of the application of the superposition technique to calculation of multielectrode lenses, as well as estimations of the resulting error in the calculated field, are given in [16].

14.4 Examples of Beam Envelope Calculations

Consider two examples of the application of the method based on solution of the simultaneous differential equations for the beam envelopes, exploiting results obtained at the 3 MV Pelletron accelerator in Lund [17].

14.4.1 The Low Energy Part of the Lund Pelletron Accelerator

Figure 14.4 gives examples of beam envelopes calculated for the low-energy part of the accelerator between the ion source and the stripper. The calculations were performed in connection with the development and installation of a new injector [17]. Carbon ion beams with currents of 5 and 10 μA and normalized emittances of 2π mm mrad (MeV) $^{0.5}$ and 4π mm mrad (MeV) $^{0.5}$, respectively, were considered. We assumed that the following voltages were applied to the electrodes of the ion source and the ion source lens (see Fig. 14.5): $U_1 = -40$ kV, $U_2 = -35$ kV, $U_3 = -29$ kV, $U_4 = -35.4$ kV, $U_5 = -20$ kV and $U_6 = 0$ kV. All voltages except U_4 are similar to those found from experience with test running of the ion source. (U_4 in those tests was varied from -35.6 kV to -36.4 kV, depending mainly on the sputtering conditions.) These results illustrate the beam transport through the low-energy part of the accelerator. The ion source lens, in the form in which it has been provided in the design of the injector, allows ions to be focused into a waist on the beam profile monitor. The voltages of the lens electrodes used in the calculations are in agreement with their experimental values. The shape of the beam

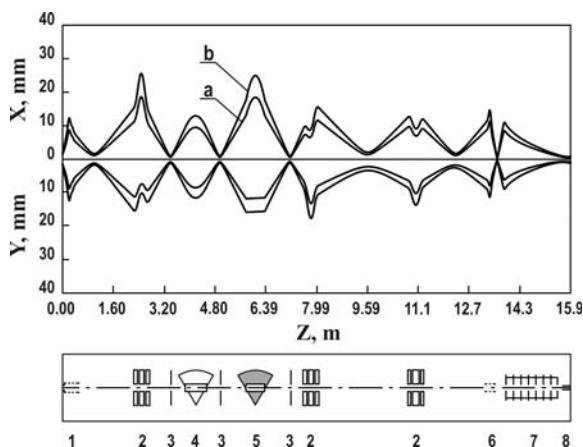


Fig. 14.4. Beam envelopes in the low-energy part of the Lund Pelletron accelerator. Calculations were carried out for a carbon ion beam under the following conditions: (a) $I = 5 \mu\text{A}$, $\epsilon_n = 2\pi$ mm mrad (MeV) $^{0.5}$, (b) $I = 10 \mu\text{A}$, $\epsilon_n = 4\pi$ mm mrad (MeV) $^{0.5}$. Numerals: 1, ion source lens; 2, electrostatic quadrupole triplet; 3, slit device; 4, spherical electrostatic analyzer; 5, magnetic analyzer; 6, einzel lens; 7, accelerator tube; 8, stripper

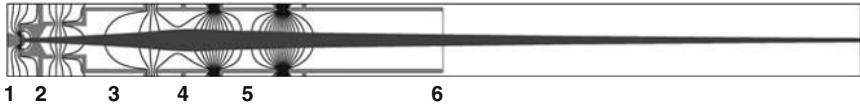


Fig. 14.5. SIMION simulation of C^- beam in the ion source section. Numerals: 1, cathode of the sputtering ion source; 2, spherical ionizer; 3, extracting electrode; 4, focusing electrode; 5, preaccelerating electrode; 6, accelerating electrode

within the electrostatic and magnetic analyzers found from these calculations indicates that the beam emittances are matched to the acceptances of the analyzer section in both transverse planes. In the dispersive plane of the analyzers, the calculated beam occupies approximately half of the aperture of the vacuum chamber. In the nondispersive plane of the magnetic analyzer, the beam transport conditions are somewhat worse. However, the transverse dimension of the beam in this case does not exceed the chamber constraints either. And finally, the ion-optical system of the new injector provides the conditions necessary to match the beam emittance and acceptance of the accelerator tube and to have a beam waist at the stripper position.

14.4.2 Ion Source Lens

To verify the calculation results for the ion source lens obtained by numerical solution of the differential equation for the beam envelopes, these calculations were repeated by use of the SIMION ion-optics simulation program [18]. Result of the SIMION simulation of the carbon ion beam in the part of the ion-optical system between the sputtered sample and the beam profile monitor, performed for the same electrode voltages as in the previous example, are given in Fig. 14.5.

The beam envelopes found from numerical solution of the differential equations (14.47) and from the SIMION simulation are given in Fig. 14.6. The calculations were carried out for two pairs of values of the beam current and emittance, namely $I = 5 \mu A$, $\epsilon_n = 2\pi \text{ mm mrad (MeV)}^{0.5}$ (Fig. 14.6, part a) and $I = 10 \mu A$, $\epsilon_n = 4\pi \text{ mm mrad (MeV)}^{0.5}$ (Fig. 14.6, part b). In both cases, the results obtained by these two different methods are in good agreement. Some differences between the beam geometries are observed. The plane of the beam waist calculated by SIMION has a small shift toward the lens. It can be explained by the influence of geometrical aberrations, which are not taken into account in the framework of the paraxial approximation used in the beam envelope method. The fact that the differences are reduced with decreasing emittance value (this value defines the highest possible angle of the ion trajectory in the beam) indicates the consistency of this assumption.

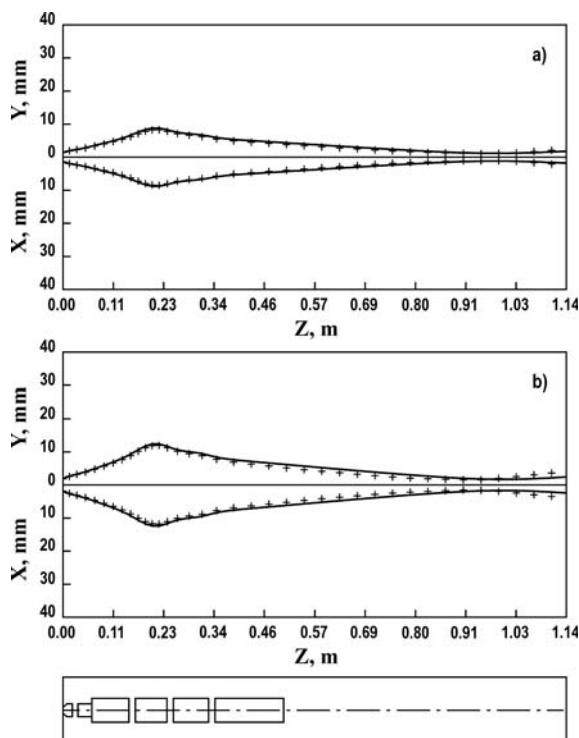


Fig. 14.6. C^- beam envelopes in the lens of the spherical-ionizer sputtering ion source, found by different calculation techniques (+++ SIMION simulation, — beam envelope method) for two values of the emittance: (a) 2π mm mrad $(\text{MeV})^{0.5}$ and (b) 4π mm mrad $(\text{MeV})^{0.5}$

References

1. H. Wollnik: *Optics of Charged Particles* (Academic Press, New York, 1987)
2. M. Szilgyi: *Electron and Ion Optics* (Plenum, New York, 1988)
3. P.W. Hawkes, E. Kasper: *Principles of Electron Optics* (Academic Press, New York, 1989)
4. M. Elkind: *Rev. Sci. Instr.* A **24**:2, 129 (1953)
5. J.D. Larson: *Nucl. Instr. Meth. A* **244**, 192 (1986)
6. K.L. Brown: *Nucl. Instr. Meth.* **187**, 51 (1981)
7. V.A. Tepljakov: *Instr. Exp. Tech.* **6**, 13 (1968) (in Russian)
8. J.D. Lawson: *The Physics of Charged-Particle Beams* (Clarendon Press, Oxford, 1977)
9. I.M. Kapchinskiy: *Particle Dynamics in Linear Resonance Accelerators* (Atomizdat, Moscow, 1966) (in Russian)
10. E.V. Shpak: *Nucl. Instr. Meth.* **213**, 171 (1983)
11. R. Hellborg, K. Håkansson, G. Skog: *Nucl. Instr. Meth. A* **287**, 161 (1990)
12. J.D. Larson, C.M. Jones: *Nucl. Instr. Meth.* **140**, 489 (1977)

13. R. Hellborg, M. Faarinen, C.E. Magnusson, S. Bazhal, V. Romanov: Nucl. Instr. Meth. A **465**, 297 (2001)
14. A.P. Banford: *The Transport of Charged Particle Beams* (Spon, London, 1966)
15. S. Bazhal, M. Faarinen, R. Hellborg, C.E. Magnusson, V. Romanov: Proceedings of the 13th International Conference on Electrostatic Accelerators, 25–28 May 1999, Obninsk, Russia, p. 160
16. S.V. Bazhal and V.A. Romanov: Proceedings of the 12th International Conference on Electrostatic Accelerators, 25–28 Nov. 1997, Obninsk, Russia, p. 229 (in Russian)
17. R. Hellborg, S. Bazhal, M. Faarinen, K. Håkansson, C.E. Magnusson, P. Persson, G. Skog, K. Stenström: Pramana – J. of Physics **59**:5, 1 (2002)
18. D.A. Dahl: SIMION 3D, Version 7.0, User's manual, Idaho National Engineering and Environmental Laboratory, INEEL-95/0403 (2000)

15 Equipment for Beam Diagnostics

M. Friedrich

Forschungszentrum Rossendorf, Institute of Ion Beam Physics and Materials
Research, P.O. Box 51 01 19, 01314 Dresden, Germany
`m.friedrich@fz-rossendorf.de`

15.1 Introduction

The ion beam from an electrostatic accelerator is described by the following parameters:

- species of ion, its charge state and energy
- beam current
- beam profile (diameter, position and intensity distribution)
- emittance and brightness
- energy spread
- stability in position and current.

During accelerator operation, not all beam characteristics can be measured with a reasonable number of instruments. Fortunately, the optimization process of the accelerator and the beam transport system requires only the measurement and control of some of the parameters.

The main parameters of an ion beam are the species, charge state and energy. Usually, it is assumed that these values are clearly determined by the operation of the ion source and the stabilized field of the deflecting magnets in connection with the acceleration voltage. Consequently, these parameters are not continuously observed during accelerator operation. Unfortunately, a mistake in this area can lead to wrong results and economic losses. Some reasons for such troubles are unexpected leaks or contaminated materials in the ion source, which result in additional oxide or hydride ions in the ion spectrum from the source. Therefore, the check of the ion spectrum from the source or in the accelerated beam should not be neglected, especially after maintenance of the ion source or a change of the ion source material.

The emittance and the brightness of an ion beam are influenced mainly by the properties of the ion source, the acceleration voltage and, in the case of tandem accelerators, the stripper density. Their values cannot be directly influenced during accelerator operation and therefore do not need to be measured continuously. Emittance measurement devices are preferably installed in ion source test stands [1]. The sources of energy instabilities and their monitoring on a display using a capacitive pickup electrode are described in Chap. 9.

In the following explanations, the experience of more than 30 years of operation and development of the electrostatic accelerators at the Forschungszentrum Rossendorf, Dresden, Germany, is summarized and focused on the necessary and appropriate equipment for beam diagnostics for electrostatic accelerators, especially for measurement and monitoring of the current, position, stability and profile of the ion beam.

15.2 Measurement of the Beam Current

The measurement of the particle current is the most important diagnostic tool. It is necessary for:

- optimizing the accelerator operation (ion source output, transmission, and beam transport system)
- calculation of the implanted fluence or as a monitoring parameter for ion beam analysis.

The beam current can be measured by destructive and nondestructive techniques. The destructive techniques can be separated into total beam-stopping devices (Faraday cups) and partially stopping devices (scanning wires and rotating-sector discs). A nondestructive technique is the measurement of the residual-gas ionization. In the case of bunched beams, the current can be measured nondestructively by capacitive pickup devices or by induction coils. Bunched beams are generated with electrostatic accelerators mainly for special applications (time-of-flight measurements) and are not considered in the following.

15.2.1 Faraday Cups

Faraday cups (FCs) are the most commonly used devices for beam current determination. The particle beam is collected in an insulated cup and measured with a conventional DC measuring technique. The design of a Faraday cup is concerned with the suppression of disturbing currents to or away from the collecting cup. Such parasitic currents are generated by:

- secondary electrons or ions from the entrance aperture or the bottom of the cup
- leakage currents from the suppression electrodes
- leakage currents from the cooling water
- charged particles from the residual gas
- thermal emission of electrons from heated surfaces.

Electrostatic and magnetic fields are applied for the suppression of secondary particles. A conventional Faraday cup with electrostatic suppression is shown schematically in Fig. 15.1. To prevent any secondary-electron emission from the suppression electrode (2), its inner diameter must be bigger

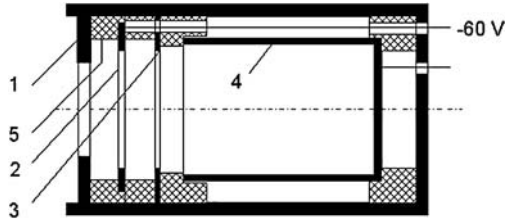


Fig. 15.1. Faraday cup with electrostatic electron suppression: 1, entrance aperture; 2, suppression electrode; 3, grounded separation electrode; 4, measuring cup

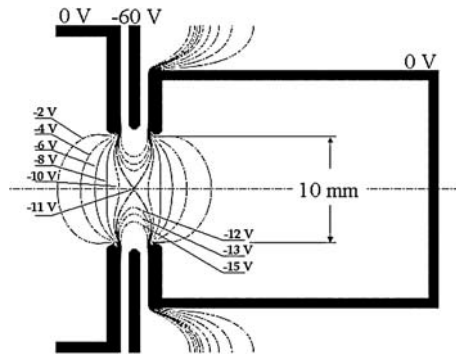


Fig. 15.2. Calculated potential distribution in an FC with electrostatic suppression

than the diameter of the entrance aperture (1). Secondary electrons cannot leave the FC (4) if the potential barrier on the optical axis exceeds their corresponding maximum kinetic energy. This potential on the axis is not identical to the potential of the suppression electrode and depends on its inner diameter and length. The calculated potential barrier on the optical axis of the FC in Fig. 15.2 is about -11 V, while the potential at the suppression electrode is -60 V. Electrons with energies of about 12 eV can pass through the entrance opening of the FC. Therefore, the beam current measured with an FC tends to a saturation current (Fig. 15.3) at suppression voltages which are higher than the maximum energy of about 10 eV of the secondary electrons. Additionally, the negative electrode collects secondary positive ions from residual-gas ionization and from the bottom of the cup, resulting in deviations of the measured current (the saturation current in Fig. 15.3) from the real beam current. In comparison with the generation of secondary electrons, this effect has a lower influence, but it must be considered for high-precision measurements.

The secondary electrons and ions can also be suppressed by magnetic fields. Under the influence of a transverse magnetic field generated by permanent magnets, the secondary particles move on bent trajectories and cannot leave the FC. A remarkable deviation of the measured current from

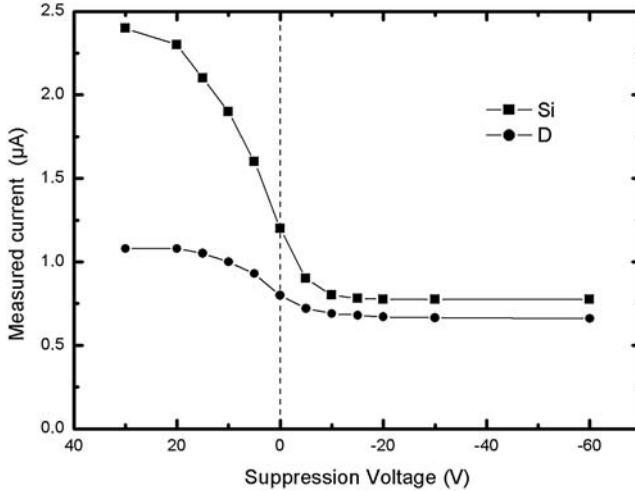


Fig. 15.3. Dependence of the measured beam current on the suppression potential for 0.7 MeV $^2\text{H}^+$ and 3 MeV Si^{2+} ions. The beam currents at positive suppression voltages indicate the different secondary-emission efficiencies

the expected value may occur if the cup is located behind an aperture (a slit device or the entrance aperture of the cup) or if the residual-gas ionization is sizable. Owing to the internal resistance of the DC measuring device, the ion beam generates a positive potential at the Faraday cup. This potential collects secondary electrons from the aperture or from the residual gas, respectively. Therefore, the measured beam current is lower than the real ion current. An additional electrostatic suppression electrode at the entrance of the FC can reduce this effect. On the other hand, the positive potential generated by the ion beam collects secondary electrons emitted from a cup or an isolated piece of quartz and improves the precision of current measurement with these devices. Using different resistors adapted to the expected range of the beam current, the current signal from an isolated screen or an FC without secondary-electron suppression can be applied in the same manner as that from an FC with additional electron suppression for optimization of the operation of the accelerator.

The arguments for and experiences with different electron suppression techniques can be summarized as follows:

- Faraday cups with electrostatic suppression of secondary electrons are versatile devices for beam current measurement.
- Faraday cups with magnetic suppression can be applied advantageously if the generation of electrons by residual-gas ionization can be neglected and no aperture is located in front of the cup.
- The influence of apertures in front of an FC with magnetic suppression can be reduced by additional electrostatic suppression.

- The potential generated by isolated beam-stopping devices can be used to improve the beam current signal from these devices.

Leakage currents can influence the measured beam current. To reduce this effect, any direct insulator connection between the suppression electrode and the measuring cup must be prevented. This can be realized by an additional grounded electrode (3 in Fig. 15.1), which collects leakage currents from the suppression electrode. This detail is neglected at some commercial FCs and the modification of these cups is nearly equivalent to a new construction.

For high beam power, the cup must be cooled. In order to reduce disturbing influences and leakage currents, distilled water or compressed air is commonly applied. For precise low-current measurements, a cooling-water line is not necessary and can be interrupted. For this purpose, the use of quick-disconnect cooling lines is advisable. If the FC is constructed using materials of high melting point (e.g. tantalum), the heat can also be removed by thermal radiation. Commercial devices based on radiation cooling are available for up to 50 W beam power [2]. A disadvantage of the radiation-cooling method is the possible gas desorption from the beam line wall, causing beam degradation.

Besides the suppression of secondary electrons, the problem of removing the FC from the beam axis has to be considered in the construction of an FC. Some commonly applied methods are used at the Rossendorf electrostatic accelerator and can be compared after several years of operation:

- In the 2 MV Van de Graaff accelerator, the FCs (or insulated quartz disks) are lifted by the magnetic field of solenoids.
- The beam diagnostic elements for the 5 MV tandem accelerator were originally equipped with electric motors.
- All retractable FCs in the 3 MV Tandetron are equipped with pneumatic cylinders.

Owing to the switching time of several seconds for motor-driven diagnostic elements, this method is not applied anymore. The switching time and reliability of the devices with magnetic solenoids and those operated with compressed air are comparable. Since pneumatic cylinders are commercially available, this method has been used in reconstruction and enlargement of the beam lines.

15.2.2 Nondestructive Beam Current Measurement by Residual-Gas Ionization

Owing to interaction of the accelerated ions with the residual gas, electron-ion pairs are generated along the beam trajectory. These electrons and ions can be separated by a transverse electric field and detected with particle detectors, Faraday cups or channel-plate amplifiers (Fig. 15.4). The ionization efficiency depends linearly on the residual-gas pressure and is also influenced

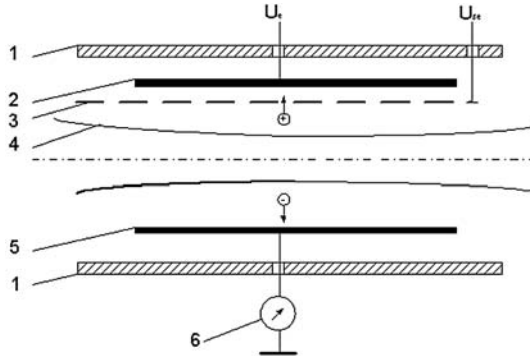


Fig. 15.4. Principle of beam current measurement using residual-gas ionization: 1, beam line wall; 2, extraction electrode; 3, secondary suppression electrode (optional); 4, ion beam; 5, collecting electrode; 6, current measurement

by the atomic number and energy of the accelerated primary ions. Owing to the calibration necessary for beam current measurements, this nondestructive method is not applied over a wide range, but it appears as a basic principle in particle detectors (gas-filled ionization chambers). A special modification for beam profiling is described in Sect. 15.4.2.

15.2.3 Partially Destructive Beam Current Measurement

The interruption of the ion beam during measurement with an FC can be reduced using an off-axis FC and short-time electrostatic deflection of the beam into the cup. This method is preferably applied when there is a constant beam current. If the beam is unstable, the average value of the beam current must be determined in fast, short sequences. This can be realized by measuring the beam particles backscattered from a rapidly rotating sector disk plated with gold (Fig. 15.5). Thereby, the ion beam is reduced by the ratio of the area of the sectors to the area of the full circle. The count rate of the backscattered particles can be converted into the real beam current by calibration using an FC. At experiments with variation of the ion energy E , the dependence of the counting rate on E^{-2} must be taken into account.

15.3 Monitoring of the Beam Diameter and Position

A simple device for observation of the beam diameter, position and stability is a screen which emits light under irradiation with the accelerated particles. This screen may consist of a metal plate coated with luminescent material (ZnS, MgO or Al_2O_3) or of a quartz disk. CsI crystals have been applied for very low currents. The coated metal plates can be produced in a simple

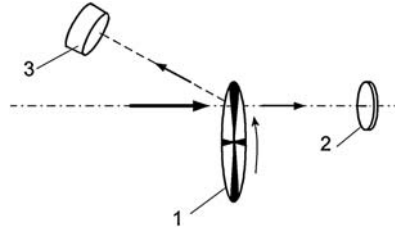


Fig. 15.5. Beam current measurement using ions backscattered from a rotating sector disk: 1, sector disk; 2, target; 3, particle detector

way, for example by moving a metal plate in the vapor of a burning piece of magnesium. An additional advantage is the electric conductivity of a metal plate. In contrast to insulating screens, no discharge effects appear on its surface that may be falsely interpreted as instabilities in the beam position. On the other hand, the lifetime of coated screens is limited, especially for MeV ions. Therefore, the most commonly applied material for beam observation screens is quartz. For suppression of discharge effects, the irradiated surface can be covered with a metallic net. Quartz emits a blue light under irradiation with ions and electrons. This can be observed directly through a glass window or in a remote mode by use of a TV camera. For higher beam power, the light emission changes to glow colors of red, yellow and white. Ion beams with such high beam power can be observed with quartz disks only for a short time to prevent damage to the material. At high beam power, the infrared radiation from a metallic plate in conjunction with a dedicated infrared-sensitive camera can be used for beam monitoring [3].

The main disadvantage of the observation screens mentioned above is the interruption of the beam. Therefore, in the FZ Rossendorf, a control method without complete beam interruption has been developed, and has been applied over a wide range. In accordance with the requirements of the experiments and the stabilizing circuits and to define the optical axis of the system, some slit devices are installed in the beam line of the electrostatic accelerator, especially at the entrance and exit of deflection magnets, in front of focusing lenses and at the places of beam crossovers. A small part of the beam intensity hits the slit plates. The electric signals from a 4-sector slit are amplified by a 4-channel preamplifier and 4-channel amplifier, both with adjustable gains, and are visually displayed in a 4-channel 10-element LED bar graph array (Fig. 15.6). Owing to the identical arrangements of the LED display and the slit devices, the accelerator staff get immediate information about the beam position and stability. The combination of these LED displays with retractable Faraday cups behind the slits has proved to be an effective piece of equipment for beam transport optimization in electrostatic accelerators.



Fig. 15.6. Beam diagnostic equipment at the Rossendorf 3 MV Tandetron: beam position monitoring by slit devices and LED display (*center*), NEC beam profile monitor (*left*), current measurement using retractable FC (*top*), and computer-controlled parameter variation (*foreground*)

15.4 Beam Profile Monitors

15.4.1 Beam Profiling by Sensing Wires

The position and intensity distribution inside an ion beam can be measured by a net of insulated thin wires [4]. The displayed current signals from each wire give information about the intensity distribution in the horizontal and vertical directions. This basic principle is modified in commercial beam profile monitors (BPMs) using a wire moving in two perpendicular directions through the cross section of the beam. Inside the BPM from NEC, Middleton, USA, a helical grounded wire is moved by a motor (Fig. 15.7, [2]). The rotation axis is arranged at 45° to the horizontal and vertical directions. The wire sweeps across the beam twice during each revolution to give a y -profile in one half-revolution and an x -profile during the next half-revolution. The secondary electrons generated on the wire are collected by a cylindrical electrode and give information on the beam intensity at the wire position.

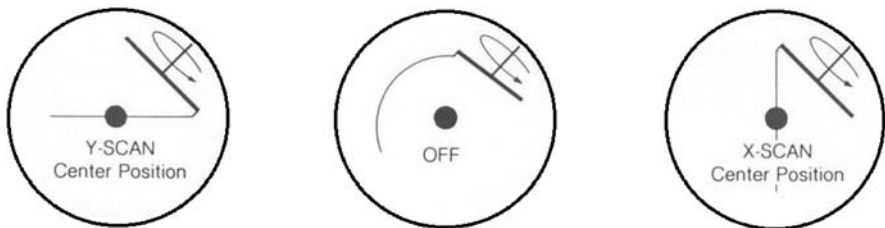


Fig. 15.7. Principle of NEC beam profile measurement using rotating helical wire (Reprinted from [2], with permission from NEC)

The signal is displayed for both directions on an oscilloscope (see the oscilloscope on the left side of Fig. 15.6). A modified version of this BPM can be applied for very low currents (particles/s) [5]. Here, a solid-state detector collects counts from several turns of the wire. The signal is fed into an MCA card in a standard PC. The display is similar to the high-current mode. The emission of secondary electrons depends on the atomic mass number and energy of the accelerated ions. Using calibration measurements for different ions and energies, the signal of the BPM can also be used for beam current measurements.

Inside the BPM from HVEE, Amersfoort, Netherlands, a Y-shaped sensing wire sweeps through the beam. The scanner head is mounted at 45° to the horizontal beam axis [6]. The collected beam current from the scanning wire is displayed for the horizontal and vertical directions on an oscilloscope (Fig. 15.8), whose time base moves synchronously with the sweep of the scanner. The sweeping of the scanner is controlled by the drive electronics together with so-called power and reference coils interacting with permanent magnets at the base of the sensing wires. By installation of additional electrodes or by adding a positive bias voltage to the scanning wire to suppress secondary ions, the BPM can also be applied for beam current monitoring without dependence on the energy and atomic number. A potential-separated preamplifier with a 30 V bias potential has been applied in the BPM of the Rossendorf 5 MV tandem accelerator. This modification has proved to be helpful, especially in beam profile measurements for negative ions, where the secondary electrons can reduce or completely compensate the primary signal from the ion beam.

The displayed signals from both types of BPM are nearly identical; for the handling conditions also, no remarkable differences exist.

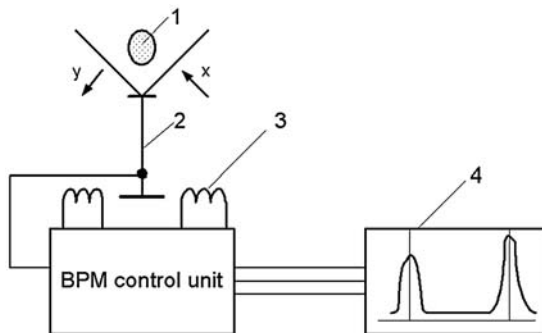


Fig. 15.8. Principle of HVEE beam profile measurement using scanning Y-shaped wire: 1, ion beam; 2, sensing wire; 3, power and reference coils; 4, oscilloscope

15.4.2 Beam Profiling Using Gas Ionization

At the cyclotron U-120 at the FZ Rossendorf, a BPM using residual-gas ionization was tested. It was manufactured in the RRC Kurchatov Institute, Moscow [7]. The electrons (or ions) generated in the residual gas by the ion beam are extracted by a transverse electric field (x -direction in Fig. 15.9). According to their transverse coordinate of generation in the electric field, the extracted electrons have different energies. After passing through a resolution slit and deflection by an electrostatic mirror, the electrons hit an electron-optical converter equipped with a channel-plate amplifier. The signals from the converter are displayed on a TV or PC monitor. In the y -direction, the electrons hit the converter according to the x -coordinate of ionization inside the extracting electric field; the perpendicular transverse coordinate y is not influenced by the electrostatic mirror. Consequently, both transverse coordinates of the detected electrons on the observation screen are a definite function of their generation points inside the beam cross section. Together with the measured vacuum in the beam line, the detected electron current also allows ion beam current monitoring. The resolution of this BPM would also allow its application to electrostatic accelerators.

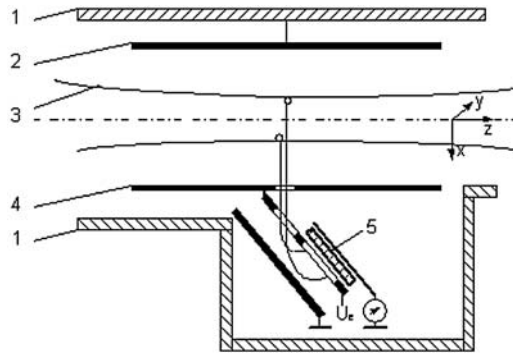


Fig. 15.9. Beam profile monitor using residual-gas ionization: 1, beam line wall; 2, grounded condenser plate; 3, ion beam; 4, extraction condenser plate connected to deflecting plate of electrostatic mirror; 5, channel-plate amplifier of electron-optical converter

15.5 Beam Stoppers and Safety Equipment

The Faraday cups installed in the beam line of an accelerator can be applied for beam current measurements and also to stop the beam during breaks in the experiment, as in the case of radiation hazards. In electrostatic tandem accelerators, the ion beam is preferably stopped with an FC in the

injector region. Special considerations are necessary for accelerator operation near the maximum terminal voltage. A fast beam stop in the injector region causes an increased terminal voltage and may create a voltage breakdown with consequent damage to the accelerator. At the high-energy side of a tandem or in a single-stage accelerator some additional consideration, must be taken into account in the application of an FC for beam stopping owing to the higher beam power and possible nuclear reactions. The sputtering and heating effects of high-energy ion beams can damage the FC, and this results in a reduced accuracy of current measurements. Because beam stopping on the high-energy side does not interrupt the energy-stabilizing circuit, this method has been preferred by accelerator staff. Therefore, a dedicated cooled FC should be installed, which is able to stop all kinds of high-energy ion beams. At the Rossendorf 5 MV tandem accelerator, a radiation-cooled FC has been applied for this purpose. The stopper must consist of material with a high atomic number (e.g. tantalum), and the irradiated bottom should be exchangeable to prevent nuclear reactions. After long-time acceleration of Li ions, a beam stopper can create unexpectedly high neutron radiation during proton acceleration owing to the nuclear reaction ${}^7\text{Li}(p, n){}^7\text{Be}$.

For high-accuracy implantation experiments, a combination of fast electrostatic beam deflection and conventional beam stopping using a retractable FC has been advantageously applied. After the required implanted fluence has been reached, the ion beam is deflected outside the implantation chamber on to the wall of the beam line, and an FC before the beam-scanning system prevents the implantation of ions backscattered from the wall.

In the case of a radiation hazard, the safety equipment must stop the ion beam fast, reliably and completely. In tandem accelerators, this can be realized by a fast, pneumatically moved FC with an electronic interlock to prevent any unwanted retraction of the FC. Owing to the possible generation of radiation in a beam stopper, a retractable FC used for radiation safety purposes in the beam line of a single-stage accelerator must be combined with switching off the accelerating voltage.

References

1. P. Strehl: Ion beam diagnostics, in: B. Wolf (editor), *Handbook of Ion Sources*, CRC Press, Boca Raton, 1995, p. 385
2. NEC: Pelletrons, vacuum components and beam-handling catalog
3. H. Büttig: Nucl. Instr. Meth. **203** (1982) 69
4. F. Loyer: Proc. 11th Int. Conf. on Cyclotrons and their Applications, 1987, Tokyo, p. 449
5. G. Norton: Pramana – J. Phys. **95** (2002) 745
6. HVEE: Beam-profiling system manual A-4-35-253
7. V.Y. Mikhailov, V.V. Leonov, V.A. Rezvov, A.A. Roschin, V.I. Sklyarenko, L.I. Yudin, A.I. Artemev, T.Y. Rakhimbabaev: Instr. Exp. Techn. **38** (1995) 717

16 Computer Control

M.L. Roberts

Woods Hole Oceanographic Institution, Mail Stop No. 8, Woods Hole,
MA 02543, USA
`mroberts@whoi.edu`

16.1 Introduction

In the past two decades, the greatest improvement to the performance and ease of operation of electrostatic accelerators has undoubtedly been the development and implementation of computerized control systems. In the “old days”, control of an electrostatic accelerator and associated components was accomplished through a large central console containing a massive collection of knobs, switches, meters, dials, and indicator lights. This central console was often located some distance from the accelerator, which necessitated long (and hence expensive) control cable runs that made the system susceptible to ground loops and electromagnetic interference. Each element in the accelerator system was typically controlled by a custom-fabricated chassis in the control console and this individuality increased costs, caused difficulties in repair and maintenance, and created a system not readily amenable to change. These major difficulties aside, however, perhaps the biggest disadvantage of these “knob-based” consoles was the fact that start-up of a typical accelerator system (or even changing from one set of parameters to another) could require the assistance of one or more skilled operators and require hours of “tuning” and/or “retuning”. Modern, properly designed and implemented computer control systems have alleviated many of these issues.

An old “knob-based” console for accelerator control is shown in Fig. 16.1. Figure 16.2 shows the operator interface of a modern, computer-controlled accelerator system. The difference is striking, given that the complexity and number of elements controlled in each accelerator system are similar. Comparing the two photographs, one can begin to understand that a computerized control system is less expensive to implement, more reliable, more precise, less expensive to operate, easier to modify, more flexible in the rapid shift from one set of parameters to another, and more capable than a “knob-based” system. The “knob-based” system is only more impressive in scale.

16.2 Software and Hardware

A large array of software and hardware is available and appropriate for use in an accelerator control system. Because new technology, new products,



Fig. 16.1. A “knob-based” central control console used to control the High Voltage Engineering Corporation model FN tandem accelerator and associated components at the Triangle Universities Nuclear Laboratory, Duke University, Durham, NC. The photo is circa 1970 and courtesy of Chris Westerfeldt, TUNL, Duke University



Fig. 16.2. A satellite computer used to control the High Voltage Engineering Corporation model FN tandem accelerator and associated components at the Center for Accelerator Mass Spectrometry, Lawrence Livermore National Laboratory. Essentially the entire accelerator system can be controlled from a single computer screen. Photo courtesy of William Fields, CAMS, LLNL

and new software are continuously being developed, any detailed discussion of computer control hardware and software is almost instantly out of date. Nevertheless, a generalized description of some of the software and hardware typically found in an accelerator control system is worthwhile.

Figure 16.3 shows a generalized computer control system. The operator interacts with software on a computer that in turn communicates to a device interface. The device interface contains all the analog and/or digital inputs and outputs (I/Os) needed to control the particular device(s). A device is any of the multitude of power supplies, solenoid valves, beam profile monitor(s), oscilloscopes, etc. needed to operate the accelerator. In small accelerator systems, with only a few devices to control, the control system may have only

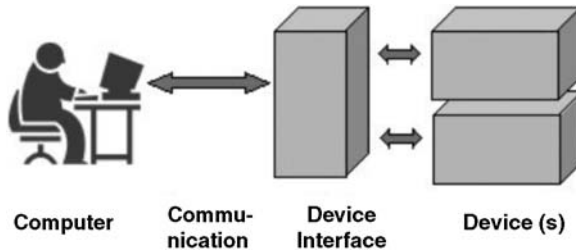


Fig. 16.3. A generalized computer control system

one device interface, which may in fact be directly embedded in the computer. Large accelerator control systems, with hundreds of devices, may have multiple computers and multiple device interfaces distributed throughout the accelerator facility.

Several software packages are suitable for use in an accelerator control system, and only a few brief comments can be made about these very complex software programs. The majority of accelerator laboratories that have upgraded their infrastructure from a “knob-based” to a “computer-control” system seem to have used either LabVIEW [1] or EPICS [2]. LabVIEW is suitable for small to medium-sized accelerator systems, is cross-platform compatible, and uses a graphical programming language that is relatively easy to learn. EPICS, or Experimental Physics and Industrial Control System, is primarily used on large accelerator systems. EPICS requires considerable computer expertise to implement, and is specially designed for high-bandwidth, real-time networking applications in which tens or even hundreds of computers are linked together. Two other software packages that have been used in accelerator control applications are InTouch [3] and Vsystem [4]. The two largest commercial manufacturers of electrostatic accelerators, National Electrostatics Corporation and High Voltage Engineering Europe, use control software developed in-house [5, 6].

Various communication schemes between the computer and the device interface are frequently found in accelerator control systems. These include copper cable (i.e., the General Purpose Interface Bus (GPIB), RS-232, RS-485, etc.), fiber optics (glass or plastic), and networks (usually a local area network but occasionally the Internet). Copper cable, especially GPIB, can offer high data transfer rates. Distances are limited and electromagnetic interference can be a problem. A network can communicate over long distances but can be limited to low data transfer rates. Accordingly, many networked computer control systems have computers embedded within the device interface. The embedded computer takes care of local, speed-critical tasks and only system changes are transmitted back to the main control computer. Fiber-optic communication offers good data transfer rates, works over moderate distances, and is relatively immune to electromagnetic interference. Furthermore, many accelerator laboratories wish to control various devices at the terminal of the

accelerator or at ion source potential. Fiber-optic communication is ideal in situations requiring high-voltage isolation.

The device interface is essentially the interconnect between the computer and the particular device(s) that need to be controlled. One of the more popular device interfaces is Computer Automated Measurement and Control (CAMAC). CAMAC devices were first developed in 1969 and were designed for use by the high-energy physics data acquisition community. Since many accelerators were (and are still) being used for high-energy physics research, it is only natural that these devices would find their way into accelerator computer control systems. Many manufacturers [7] make various CAMAC modules. Typical modules include analog outputs (to control a power supply or device), analog inputs (to read back voltages or currents), digital outputs (to control a solenoid valve or switch), digital inputs (to read back the status of a solenoid valve or switch), timing generators, counters, and waveform recorders. Increasingly, many feel that CAMAC is becoming obsolete technology. Some CAMAC users are switching to VME (Versa Module Europa) or VXI (VME eXtensions for Instrumentation). Most manufacturers of CAMAC modules also make VXI modules. Compared with CAMAC, VXI offers better immunity from electromagnetic interference. Unfortunately, VXI devices tend to be more expensive and less densely packed than comparable CAMAC modules.

Device interfaces designed specifically for computer control of accelerator systems are also available. Group3 [8] has a line of products in which fiber optics are used to link a series of small, distributed modules. A module may contain one or more analog outputs, analog inputs, digital outputs, digital inputs, stepper motor controllers, communication ports, etc. One Group3 module even has provision for an embedded PID (proportional–integral–derivative) control algorithm that can be useful for closed-loop control of various devices (e.g., a momentum-analyzing magnet). A Group3-based control system is easily expandable, and the fiber-optic communication provides high-voltage isolation and good noise immunity. Overall, Group3 control products have found wide acceptance in many accelerator laboratories.

Other manufacturers also make instrumentation useful in accelerator control systems. Besides LabVIEW, National Instruments makes a diverse array of device interfaces, including digital oscilloscopes and motion controllers that have found use in control systems. Industrial control system hardware such as programmable logic control (PLC) has also been used in some accelerator control systems. Two PLC brand names are MODICON [9] and Allen-Bradley [10]. PLC technology is simple, inexpensive, and robust but can lack the control precision demanded in most accelerator operations.

16.3 Operator Interface

More important than the choice of software and hardware is the manner in which accelerator personnel interact with the control software. A poorly implemented interface can outweigh all possible positive features of a computer control system. A good interface can greatly enhance the usefulness of a computer control system. What makes a good computer control system, however, is somewhat dependent upon the eye of the beholder. Accelerator operators want a control system with a quick response, and need tools to analyze ongoing operations and make correlations between parameters and measured values. Maintenance personnel want to monitor magnet currents and the voltages of the power supplies, and to have tools that provide information for analyzing and investigating problems. Computer support personnel have their own requirements to monitor system performance and error logging. Accelerator users typically want an on/off button. The end result is that the control system must contain hardware and software components that allow the users of the accelerator to control the accelerator system in the most efficient and effective manner possible.

The best accelerator computer control systems have a minimum of display windows and are graphically based (i.e., the use of tables of parameters is avoided). In small accelerator systems, the entire system can often be displayed on a single computer window. Nonessential information such as setup parameters, maintenance diagnostics, and nonroutine procedures are not continuously displayed, and are made accessible from separate (and usually hidden) computer windows. It is often helpful to have a flowchart or basic outline of the accelerator and beam transport elements. This outline helps the infrequent or novice user understand the flow of the beam and the spatial relationship of the various devices. Faraday cups and vacuum valves can be inserted or retracted at the push of a mouse button. Power-supply settings can be changed by clicking on a device and entering a new value or by assigning the device to a control knob. Error conditions (such as an out-of-range power supply) can be indicated by having the device icon change color or shape. A brightly colored error indication will draw the eye much faster than scanning a list of parameters looking for differences. In addition, provisions should be made so that previous set points can be retrieved and current set points logged and saved for future retrieval. If anything can be sequenced or automated, it should be.

The response time of the accelerator computer control system should also be considered. Early computerized control systems often displayed a noticeable lag between when a computer button or knob was pushed or turned and when the physical device actually responded. This slow response was annoying and made beam tuning difficult. The increased speed of modern computers has largely solved this problem. However, consideration should still be given to leaving high-frequency devices such as beam profile monitors

and Faraday cup current measurements outside of the computer control system, with only the control of such devices in the control system.

Finally, control of items involving either personnel safety or instrument protection should be independent of the computer control system. Such items include but are not limited to radiation interlocks, vacuum interlocks, and high-voltage interlocks. Primary control of such items should be through hardwired systems. It is perfectly reasonable to monitor or back up such systems with the computer control system, but a computer must never be the primary system when safety is involved.

16.4 Special Algorithms

Various special routines or algorithms have been developed that allow the users of an accelerator to control the accelerator system in the most efficient and effective manner possible. Although the exact details of these algorithms will vary with the details of the individual control systems, the general principles described should be useful in many accelerator computer control systems. These routines include “flat-topping”, “scaling”, “conditioning”, closed-loop control, and auto-tuning.

Output from a typical so-called “flat-topping” routine is shown in Fig. 16.4. “Flat-topping” involves slewing a selected optical element over

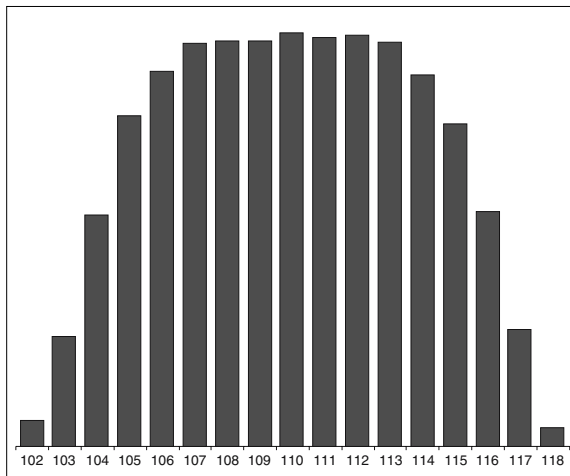


Fig. 16.4. Output from a “flat-topping” routine. The x -axis is the device set point varied over some user-defined range. The y -axis is an arbitrary measured parameter (in this case a Faraday cup current). “Flat-topping” allows the operator to set a device in the middle of the “flat-top” region of the tuning response curve. With a “knob-based” control system the operator might inadvertently tune the device near one of the “edges”

some user-defined range and displaying the value of that element against a measurable parameter (e.g., current from a Faraday cup or counts from a detector). “Flat-topping” allows the operator to precisely set the value of a selected element to the optimum value. “Scaling” involves using basic physical formulas to “scale” the accelerator from one operating point to another. The change in operating point could be either to a new energy setting or to a new mass or both. These algorithms can be surprisingly precise and are of great use in laboratories that utilize many different types of ions and/or a broad range of energies.

Routines can be designed to aid in the “conditioning” of the accelerator to high voltages. A “conditioning” routine might involve ramping the terminal potential in a sawtooth fashion in which the accelerator terminal potential is raised by a user-defined value for a user-defined time. The terminal potential is then dropped (again by a defined value for a defined time) and the process repeated as often, and as long, as necessary to reach the desired terminal voltage. Many laboratories have found this method of conditioning more effective than a slow incremental increase in terminal potential. The computer control system can relieve the operator of this tedious and boring procedure.

The computer control system can also be used to stabilize, or closed-loop control, a device such as a bending magnet. Using a Hall probe, algorithms can be developed to adjust the output of a power supply to maintain a precise magnetic field. Since Hall probe readings are typically more precise and stable than power supply current readings, these techniques provide a more stable ion beam than what could be obtained if one were to rely only on the internal stability of the power supply. In any closed-loop system, however, care should be taken to avoid control offsets and oscillations. Various texts on control loops are available [11, 12].

Finally, some accelerator laboratories have implemented routines to automatically tune beams. The accelerator mass spectrometry (AMS) group at the Vienna Environmental Research Accelerator, University of Vienna, Austria has developed a tool that maximizes a measurable parameter (i.e., a Faraday cup current) by adjusting accelerator parameters (e.g., steerer voltages, magnet currents, and slit positions) [13]. Such routines are valuable in maximizing ion transmission, especially in cases where apertures are narrow and “flat-top” transmission is difficult to obtain. High and reproducible ion-optical transmission is essential in AMS measurements since beam losses can directly influence measured isotope ratios.

16.5 Summary

Given changing technology, the large array of available software and hardware, and the personal preference of the individuals involved, it is almost certain that no two computer control systems for electrostatic accelerators

are exactly the same. Nevertheless, most accelerator control systems use similar hardware and have similar design philosophies. Compared with a “knob-based” system, computer control systems are less expensive to implement, more reliable, more precise, less expensive to operate, easier to modify, and more flexible in the rapid shift from one set of parameters to another. Furthermore, a computer control system has expanded capabilities that cannot be readily achieved by a “knob-based” system. More information about computerized control systems is available from the above-referenced manufacturers or a variety of reports from specific accelerator laboratories [14–19].

References

1. LabVIEW, National Instruments Corporation, Austin, TX, USA
2. EPICS, or Experimental Physics and Industrial Control System, is also the name of the collaboration of organizations that were and are involved in the software’s development and use. Los Alamos National Laboratory, Los Alamos, NM, USA, and Argonne National Laboratory, Argonne, IL, USA originally wrote EPICS jointly
3. InTouch[®], Wonderware/Invensys Systems, Lake Forest, CA, USA
4. Vsystem[®], Vista Control Systems, Inc., Los Alamos, NM, USA
5. National Electrostatics Corporation, Middleton, WI, USA has developed a computer control software package called AccelNET. Details can be found in “Automated accelerator controls for a 3 MV tandem Pelletron”, R.D. Rathmell, R.L. Kitchen, T.R. Luck, M.L. Sundquist, Nucl. Instr. Meth. Phys. Res. B **56/57** (1991) 1072
6. High Voltage Engineering Europa B.V, Amersfoort, The Netherlands has developed in-house a dedicated Microsoft Windows-based computer control software program
7. There are many manufacturers of CAMAC devices, including Kinetic Systems Company, LLC, Lockport, IL, USA; Joerger Enterprise, Inc., East Northport, NY, USA; BiRa Systems, Inc., Albuquerque, NM, USA; and CAEN S.p.A., Viareggio, Italy
8. Group3 Technology, Auckland, New Zealand
9. MODICON, Schneider Electric, Groupe Schneider, North American Division, Palatine, IL, USA
10. Allen-Bradley[®], Rockwell Automation Control Systems, Milwaukee, Wisconsin, USA
11. K.J. Åström, T. Hägglund, *PID Controllers: Theory, Design, and Tuning*, 2nd edn. (January 1, 1995), Published by The Instrumentation, Systems, and Automation Society, Research Triangle Park, NC 27709 USA
12. A. Datta, M.-T. Ho, S.P. Bhattacharyya, *Structure and Synthesis of PID Controllers*, Springer Berlin, Heidelberg, 2000
13. P. Steier, S. Puchegger, R. Golser, W. Kutschera, A. Priller, W. Rom, A. Wallner, E. Wild, Developments towards a fully automated AMS system. Nucl. Instr. and Meth. B **161–163** (2000) 250–254

14. R.D. Rathmell, R.L. Kitchen, T.R. Luck, M.L. Sundquist, Nucl. Instr. Meth. Phys. Res. B **56/57** (1991) 1072
15. M.L. Roberts, T.L. Moore, Nucl. Instr. Meth. Phys. Res. **56/57** (1991) 1080
16. J.R. Lutz, J.C. Marsaudon, Nucl. Instr. Meth. Phys. Res. A **328** (1993) 113
17. T.M. DeTurck, D.J. Treacy Jr., D.L. Knies, K.S. Grabowski, C. Knoll, C.A. Kennedy, G.K. Hubler, in: J.L. Duggan, I.L. Morgan (Eds.), Conference Proceedings 475, American Institute of Physics Press, New York, 1999, p. 668
18. N. Akasaka, A. Akiyama, S. Araki, K. Furukawa, T. Katoh, T. Kawamoto, I. Komada, K. Kudo, T. Naito, T. Nakamura, J. Odagiri, Y. Ohnishi, M. Sato, M. Suetake, S. Takeda, Y. Takeuchi, N. Yamamoto, M. Yoshioka, E. Kikutani, Nucl. Instr. Meth. Phys. Res. A **499** (2003) 138
19. D.S. Barton, S. Binello, W. Buxton, T. Clifford, T. D'Ottavio, H. Hartmann, L.T. Hoff, R. Katz, S. Kennell, T. Kerner, J. Laster, R.C. Lee, A. Marusic, R. Michnoff, J. Morris, B.R. Oerter, R. Olsen, J. Piacentino Jr., J.F. Skelly, Nucl. Instr. Meth. Phys. Res. A **499** (2-3) (2003) pp. 356-371

17 Radiation Protection at an Accelerator Laboratory

R. Hellborg¹ and C. Samuelsson²

¹ Department of Physics, Lund University, Sölvegatan 14, 223 62 Lund, Sweden
`ragnar.hellborg@nuclear.lu.se`

² Department of Medical Radiation Physics, Lund University, University
Hospital, 221 85 Lund, Sweden
`christer.samuelsson@radfys.lu.se`

17.1 Introduction

The phenomenon called “radiation” is the transport of energy in the form of a stream of atomic particles or electromagnetic quanta (photons). No supporting medium is required. Radiation can be divided into ionizing and non-ionizing. Ionizing radiation has a higher energy than nonionizing. (As a rule of thumb, ionizing radiation has an energy of the order of atomic or molecular binding energies, that is, 10 eV or higher, while nonionizing radiation has an energy below 10 eV.) Ionizing radiation can – as the name suggests – ionize material when interacting. Ionizing means that electrons are removed from the atoms/molecules in the material by the radiation. In this way, charged particles, i.e. ions, are produced. If this happens in a human body, radiation injuries can result. The interaction of radiation with matter is discussed in Sect. 17.2 and its consequences in a human body are discussed in Sect. 17.4.

Ionizing radiation is not a new phenomenon connected with human activity. It has always been available and is present throughout the environment. However, it is only during the last century that man has learned to detect ionizing radiation and to produce artificial ionizing radiation. In Sect. 17.3, the quantities and units used within the field of radiation protection are discussed, and in Sect. 17.5, detectors for ionizing radiation are briefly presented. In Sect. 17.6, dose measurements are outlined.

Many accelerators are used at low intensities and have therefore normally a low radiation level. However, it should be remembered that in general any accelerator can produce hazardous levels. Even if the ion (or electron) source is switched off, stray electrons can be accelerated over the high-voltage gap, producing bremsstrahlung when hitting material. The radiation hazards for different types of accelerators and for different kinds of radiation are discussed in Sect. 17.7.

For standard use with low intensities, only minimal shielding is normally required. This could be a risk, as personnel may become careless if they consider the radiation hazard as negligible. It is therefore important to measure the radiation level whenever an uncertainty exists. Safety considerations in an accelerator laboratory is discussed in Sect. 17.8. Finally, in the Appendices

of this chapter, the regulatory standards for radiation protection, attenuation of photons, and conversion between fluence and dose are given.

In Box 8, the nonradiation hazards (i.e. typical industrial hazards such as high-pressure gas hazards, electrical hazards, and fire and explosive hazards) in an accelerator laboratory are discussed.

17.2 Radiation and Its Interaction with Matter

Ionizing radiation may stem from three different origins: radioactive decay, accelerated particles or extraterrestrial (cosmic) sources. The properties of the more common type of ionizing radiation are summarized in Table 17.1. In the decay of a radioactive nucleus, its surplus energy is transferred to photons or to ionizing particles in a complex manner. The radiation emitted directly from the nucleus can be predominantly electrons (β^- -particles), positrons (β^+ -particles) or photons (γ -quanta), and for heavier nuclei also ^4He ions (α -particles). Part of the surplus energy may support processes causing vacancies in the electron shells outside the nucleus. When these vacancies are refilled, X-ray photons or so-called Auger electrons with discrete energies are emitted from the atom. The yield per decay of individual X-ray photons and Auger electrons, as well as for the primary emitted particle/photon, is fixed for a specific decay, but is only seldom 100%. Thus, the number of photons of a specified energy generated in a radioactive source is only occasionally identical to the number of decays.

The *radioactive decay* of the nucleus is statistical in nature. Therefore, it is impossible to predict when any given nucleus will disintegrate. Extensive experiments on radioactive materials have shown that the decay for a given initial mass of material is accurately exponential:

Table 17.1. Examples of different types of ionizing radiation. Charge is given in units of the elementary charge (1.602×10^{-19} C)

Type	Origin	Process	Charge	Rest Mass (kg)	Energy Spectrum
α -particle	Nucleus	Nuclear decay or reaction	+2	6.664×10^{-27}	Discrete
β^- -particle	Nucleus	Nuclear decay	-1	9.110×10^{-31}	Continuous
β^+ -particle	Nucleus	Nuclear decay	+1	9.110×10^{-31}	Continuous
γ -ray	Nucleus	Nuclear de-excitation	0	0	Discrete
X-ray	Electron cloud	Atomic de-excitation	0	0	Discrete
Neutron	Nucleus	Nuclear reaction or spontaneous fission	0	1.678×10^{-27}	Continuous or discrete
Fission fragment	Nucleus	Fission	10-30	$1.4 - 2.8 \times 10^{-25}$	Continuous

$$N_t = N_0 e^{-\lambda t} \quad (17.1)$$

which is in accordance with its stochastic nature. Here N_t is the number of independent radioactive nuclei at time t in a sample, N_0 is the number of radioactive atoms present at the beginning of the observation ($t = 0$) and λ is a constant called the disintegration or decay constant. The time interval $t_{1/2}$ during which half of the atoms disappear by decay, denoted the half-life, is given by $t_{1/2} = (\ln 2)/\lambda$. The *activity* of a radioactive material is the number of decays per unit time, and the number of decays per second is a convenient unit of measurement. In the SI system, this unit is called the becquerel (Bq). However, it should be observed that the activity says nothing about the kind of radiation emitted, nor about its energy.

The kinetic energy available from radioactive transformation is at most a few MeV, while electrostatic accelerators may generate electrons and singly charged ions of higher energies, but rarely above 5 and 10 MeV, respectively. The interaction with matter for ionizing particles in the MeV energy range will be very briefly outlined in the following paragraphs. The range of these particles is schematically illustrated in Fig. 17.1 (see also Sect. 17.7.1 and Table 17.6).

Photons lose their kinetic energy to atomic electrons, either partly (a Compton collision) or totally (the photoelectric effect). Photons above 1.02 MeV passing near a nucleus may, additionally, create an electron-positron pair. The large and few energy losses experienced by photons before they are annihilated mean firstly that the attenuation can be described by an exponential expression, $\Phi(x) = \Phi(0)e^{-\mu x}$, and secondly that the number of ion pairs created by a photon itself is vanishingly small compared with the number of ionizations caused by the generated photoelectrons, Compton-scattered electrons and pair production electrons. Photons are accordingly denoted as *indirectly ionizing*. (In the expression, Φ is the photon fluence, x is the material thickness and μ is the attenuation coefficient).

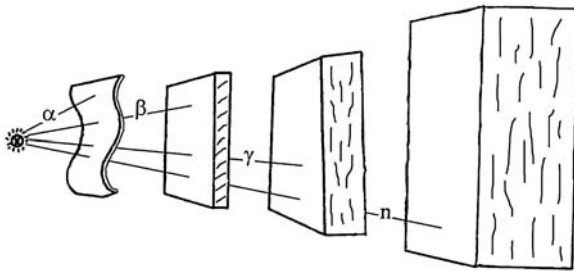


Fig. 17.1. A simple illustration of what is needed to stop energetic α, β, γ and neutron rays: a 0.2 mm sheet of paper, a 100 mm thick piece of wood, half a meter of concrete and a few meters of concrete, respectively

The two major energy loss processes for *electrons* (β -particles) are collision with atomic electrons and bremsstrahlung emission (see Sects. 17.3.2 and 17.7.1). In collisions with atomic electrons, the incident electron may lose up to half of its kinetic energy in a single encounter. Collision losses clearly dominate over bremsstrahlung losses in the electron energy range discussed here, particularly in materials of low atomic number. The fairly long range and irregular path of energetic electrons slowing down in matter means, on a microscale, that the ionization events are separated by distances of the order of a micrometer. Electrons and photons belong to the category of *sparsely ionizing radiation*.

Protons, deuterons, ^4He ions (α -particles) and other heavy charged particles are more than three orders heavier than an electron and can therefore lose only a minute fraction of their kinetic energy even in a head-on collision event with an atomic electron. On the other hand, the probability (cross section) for the event is very large, which means that heavy charged particles are *densely ionizing*. The slowing-down path is short and straight, and losses due to bremsstrahlung are insignificant.

The interaction of *neutrons* with matter is different from that of charged particles, as neutrons readily collide and interact with any nuclei encountered. A neutron is not itself ionizing, but if it hits a nucleus, it may activate it or cause emission of a γ -ray or a charged particle, indirectly giving rise to ionizing radiation. In hydrogenous material, fast neutrons slow down rapidly owing to collisions with protons. In a collision with a heavy nucleus, such as uranium for instance, the neutron loses very little of its kinetic energy. The probability of a neutron capture process is large for some light nuclei, including hydrogen, and in practice neutron radiation is always accompanied by a more or less significant amount of photons. The optimum strategy to eliminate fast neutrons is to use a proton-rich material such as water or concrete to slow them down to thermal energies and then to capture them with a material with a high capture cross section (see Sect. 17.7.2). The penetration of fast neutrons through concrete is higher or comparable to that of 1 MeV photons, and the order of half a meter or more of ordinary concrete may be needed in order to reduce the fast fluence rate by a factor of one hundred. Fundamental details about ionizing radiation and its interaction with matter can be found in [1] and [2], respectively.

17.3 Quantities and Units

The quantities and units used within the field of radiation protection are somewhat impenetrable to the layman owing to the vast range of options of measurable and nonmeasurable, mean-value and stochastic, plain physical and risk-weighted, source-related and target-related, and rate and time-integrated quantities. In this section, only the most essential quantities are mentioned and commented on. Quantities and units used within the radiation

protection field are defined by the International Commission on Radiation Units and Measurements (ICRU), and the interested reader should consult its reports on this matter for further guidance [3, 4].

17.3.1 Radiation Field Quantities

Despite the fact that ionizing radiation is quantized and its interaction is stochastic in nature, the radiation field and the energy transfer from the field to matter are mostly described in mean-value, nonstochastic terms. The particle fluence rate $\dot{\Phi}$, for instance, is thus the *mean number* of particles incident on a sphere of unit cross-sectional area per unit time. In textbooks on physics $\dot{\Phi}$ is sometimes denoted as the “particle flux”, but the recommendation by the ICRU [3] is to reserve this term for the number of particles per unit time. The particle fluence Φ (m^{-2}) is numerically identical to the total particle path length traveled per unit volume, a relation useful in dosimetric calculations. Staying with dosimetry issues, it is the kinetic energy available in a radiation field that is of prime interest, not the number of particles carrying the energy. Consequently, such quantities as the energy fluence Ψ (J m^{-2}) and the energy fluence rate $\dot{\Psi}$ ($\text{J m}^{-2} \text{s}^{-1}$) are defined.

When one is performing detailed calculations of how radiation energy is transferred to matter, the basic field quantity is the particle fluence distribution with respect to direction and kinetic energy, $\dot{\Phi}_{\Omega,T}$ ($\text{m}^{-2} \text{s}^{-1} \text{J}^{-1} \text{steradian}^{-1}$). The physical meaning of $\dot{\Phi}_{\Omega,T} d\Omega dT(\mathbf{r})$ is the number of particles per second and unit area at a point $\mathbf{r}$ in the room that fulfill the criteria of having kinetic energies between T and $T+dT$ and of being confined to the solid angle $d\Omega$ in the direction defined by the unit vector $\boldsymbol{\Omega}$. The quantity $\dot{\Phi}_{\Omega,T}$ is also known as the angular flux in the field of radiation transport theory.

17.3.2 Interaction Quantities

The probability that an ionizing particle or photon will interact with atoms (or electrons) in its path is described by the (microscopic) *cross section*, σ (m^2). If a particle can undergo different and independent kinds of interactions, the total cross section σ equals the sum $\sum \sigma_i$, where σ_i is the cross section for the interaction of type i . The photon cross section, for instance, can be split into five components, $\sigma = \tau + \sigma_c + \sigma_{coh} + \kappa + \nu$, where τ denotes the photoelectric effect, σ_c Compton scattering, σ_{coh} coherent scattering, κ pair production and ν interactions with the nucleus.

The *attenuation coefficient* μ (m^{-1}) is the *macroscopic* cross section, i.e. the number of target entities (atoms or electrons) per unit volume, n_t , times the (microscopic) cross section σ . The *mean free path* of an uncharged particle equals μ^{-1} (m), i.e. the reciprocal of the attenuation coefficient μ .

In radiation dosimetry, the focus is not on the interaction as such, but instead on how much of the particle energy is transferred to the material by the

interaction. For photon interactions, for instance, the attenuation coefficient μ is weighted by the mean fraction $\bar{f}$ that is transferred to charged particles, taking into consideration the partial probabilities of the photoelectric effect, Compton scattering etc. The resulting coefficient, $\mu_{tr} = \bar{f}\mu$, is called the *energy transfer coefficient*. A certain fraction g of the kinetic energy transferred to atomic electrons in photon interactions will be converted to photon energy (bremsstrahlung etc.) during the retardation of these electrons. The *energy absorption coefficient* $\mu_{en} = \mu_{tr}(1 - g)$ takes this loss into consideration.

The *stopping power* S (J m^{-1}) of a material, for charged particles, equals dT/dl , where dT is the energy lost by the particle in traversing a distance dl in the material. The stopping power S is the sum of at least three independent components, $dT/dl = (dT/dl)_{el} + (dT/dl)_{rad} + (dT/dl)_{nuc}$, where the index *el* denotes energy losses due to collisions with electrons, the index *rad* denotes radiative energy losses in bremsstrahlung processes, and the index *nuc* denotes energy losses in which the transferred recoil energy is imparted to atoms. The first two components are usually referred to as the *collision stopping power* and the *radiative stopping power*, respectively.

17.3.3 Dose Quantities

The basic quantity for estimating radiation risk is the *absorbed dose* D (J kg^{-1}), the specific energy imparted. The SI unit J kg^{-1} in this case has been given the special name “gray” (Gy). D is a *mean-value* quantity and does not take into account the stochastic character of the absorption process. In the high-dose range, say for D larger than 100 mGy, there are many energy deposition events per human cell, and the absorbed dose D becomes a good descriptor of the energy imparted and can be expected to correlate well with the severity of acute radiation effects.

The higher biological effect of densely ionizing radiation per unit dose has led to the attachment of a weight factor, the *radiation quality factor* ω_R , to D . The weighted quantity $H = \omega_R D$ is called the *equivalent dose* and the SI unit in this case, J kg^{-1} , has been given the special name “sievert” (Sv). The value of ω_R picked for different types of radiation is a compromise, an adaptation of the variable RBE (relative biological to effectiveness) values obtained in irradiation experiments. For sparsely ionizing radiation, such as γ -rays, X-rays and electrons (β), the radiation quality factor ω_R equals unity.

The equivalent dose H_T in tissue T is the quantity used for dose limits for individual organs. As H_T is a weighted quantity, it is normally not experimentally measurable. This remark is also valid for the *effective dose* $E = \sum \omega_T H_T$ (Sv), where the *tissue weighting factor* ω_T weights the body organs with respect to carcinogenic and hereditary effects. The sum of all ω_T factors is normalized to unity. Dose limits in working life are expressed in units of E , as this is considered to be the best quantity available for estimating the probabilities of cancer and hereditary effects. As the numerical

value of E depends on the carcinogenicity and hereditary damage of the radiation, it should not be used outside this area, for instance in forecasting acute radiation effects.

In operational radiation protection work, the three measurable quantities *ambient dose equivalent* H^* , *directional dose equivalent* H' and *personal dose equivalent* $H_p(d)$ have been introduced as substitutes for the effective dose and equivalent dose. The depth d is mainly limited to values of 10 mm, in popular terms the “deep dose”, and 0.07 mm, the skin dose or “shallow dose”. If the numerical value of H^* , H' or H_p , obtained with a properly calibrated instrument, is below the relevant dose limit, it is considered that conformity with the legal demands of the radiation protection system has been demonstrated. The definitions of H^* , H' and H_p are intricate and are aimed at radiation standardization laboratories. In operational protection work, knowledge of the precise definition of these quantities is not necessary.

The *kerma* K (Gy) is defined by the relation $K = dT_{tr}/dm$ ($= \Psi\mu_{tr}/\rho$), where dT_{tr} is the sum of the kinetic energies of all the charged particles liberated by uncharged particles in a mass dm of the material. If the energy inflow and outflow of secondary charged particles cancel out (charged-particle equilibrium, CPE) in dm , the kerma K numerically equals the absorbed dose D .

The exposure X is an old quantity emanating from the use of open-air ionization chambers as primary-standard instruments for calibration purposes. The old special unit röntgen (R) for exposure corresponds to $2.58 \times 10^{-4} \text{ C kg}^{-1}$ (or approximately 0.0087 Gy expressed as kerma in air). Dosimetry standard laboratories are phasing out both the kerma and the exposure quantities in favor of the operational quantities H^* , H' and H_p .

17.4 Radiation and Living Material

Ionizing radiation interacts on the atomic level as outlined in Sect. 17.2. Any charged secondary particle created by the primary beam ionizes the material along its track while slowing down. The mean energy absorbed to create one ionization (one ion pair) is 34 eV for electrons, and about the same value also for heavier charged particles, when stopped in living tissue. If the ionization takes place within a biomolecule such as DNA (deoxyribonucleic acid), the lesion is denoted as “direct”. If an ion pair created in any other type of molecule causes damage to a biomolecule, this is an indirect effect caused by attack from chemical radicals. Water plays an important role here, as the water content of the human body is so high.

17.4.1 Early Radiation Effects in Humans

Whole-Body Response

High irradiation doses, of order of 1 Gy and higher within less than a few hours, cause cell killing to such an extent that the function of organs will be significantly impaired or destroyed all together. Below a threshold dose of about 0.25 Gy, the cell-killing effect, even in sensitive organs is small enough to be compensated for and is not clinically detectable unless very sophisticated types of chromosomal or physiological analysis are applied.

The response to a whole-body dose in excess of 1 Gy is almost immediate, within hours, owing to damage to sensitive cells in the gastrointestinal tract. The symptom of nausea that follows, and at higher doses vomiting and diarrhea, increases in severity with dose.

As the variation in individual sensitivity to early radiation effects is small, all individuals irradiated above a certain threshold will show symptoms, and early effects are synonymously denoted as “deterministic”. Another, perhaps more common, designation is “acute”, emphasizing the short-termness (days, weeks or months) and distinguishing it from late effects (years), such as cancer and hereditary disorders.

The approximate dose thresholds for different acute radiation syndromes following brief and protracted exposures are listed in Table 17.2. Modern medicine offers treatments that relieve the symptoms following accidental overexposure, but these are rarely curative when the dose exceeds 6 Gy. On the other hand, single whole-body irradiations below 2 Gy are considered nonlethal.

The Skin Response

Soon after Röntgen’s discovery of X-rays, skin redness (erythema) and, after massive and extended exposures, loss of skin and ulceration were observed from this new type of radiation. The soft (i.e. low-energy) X-rays of the early 1900s made the skin, or to be more precise, the basal cells just below the skin surface, a critical organ for acute radiation damage. A dose of about 6 Gy is the threshold for a so-called main erythema reaction about 1–2 weeks after irradiation. Before that, within hours, the skin reacts with a mild and transient redness if the absorbed dose exceeds about 2 to 3 Gy.

Another indication of a high dose to the skin is loss of hair (epilation). The threshold dose for temporary epilation is about 3 Gy, while a dose in excess of about 7 Gy is necessary to make the hair loss permanent.

Effects During Pregnancy

Animal studies and the epidemiological results from the bombing of Hiroshima and Nagasaki (H–N) indicate that a growing embryo is prone

Table 17.2. Estimates of the thresholds for deterministic effects in the adult human testes, ovaries, lens and bone marrow ([5], p. 103)

Tissue and Effect	Threshold		
	Equivalent Dose, Single Brief Exposure (Sv)	Equivalent Dose, Highly Fractionated or Protracted Exposure (Sv)	Annual Equivalent Dose Rate for Many Years (Sv y^{-1})
<i>Testes</i>			
Temporary sterility	0.15	NA*	0.4
Permanent sterility	3.5–6.0	NA	2.0
<i>Ovaries</i>			
Sterility	2.5–6.0	6.0	>0.2
<i>Lens</i>			
Detectable opacities	0.5–2.0	5	>0.1
Visual impairment (cataract)	5.0	>8	>0.15
<i>Bone marrow</i>			
Depression of hematopoiesis	0.5	NA	>0.4

* NA denotes “not applicable”, since the threshold is dependent on dose rate rather than on total dose

to radiation damage. When the embryo is irradiated during 8–15 weeks after conception in humans, the probability of severe mental retardation is believed to be 40% per Sv and 10% per Sv during weeks 16–25 ([6], p. 231). A downward shift in IQ score is interrelated with this risk, about 30 IQ units per sievert during the most sensitive period of 8–15 weeks [5]. These risk figures are mainly based on the H–N statistics, which also indicate that there is a threshold of about 0.2 Sv for these interferences with the developing human brain.

Taking into account the risks of both induction of malformations and childhood cancer (0–19 y) after in utero exposure, the ICRP concludes that only for fetal doses in excess of 100 mGy may there be medical reasons for terminating a pregnancy [7].

17.4.2 Late Effects in Humans

Radiation-Induced Cancer

That ionizing radiation is a carcinogen has been proven beyond doubt for brief or extended exposures to high doses. In contrast to the acute effects discussed above, radiation-induced cancers are characterized by the following:

1. They are stochastic.
2. There is a long delay (years or decades), the latency period, between the exposure event and the outbreak of the disease.
3. In an irradiated population, the radiation-attributable cancer risk to still-unaffected individuals is in force for a long period after the minimum latency period.
4. The damage to an affected person will not increase with dose.

The first point means randomness in the sense that it is not possible to predict the individual persons in an irradiated population that will be affected, just an expected value of how many. The minimum latency periods for solid tumors and leukemia are assumed to be about 10 and 2 years, respectively [8].

The proof of a connection between ionizing radiation and cancer stems from many different areas, for example medical irradiation procedures of the past and uranium miners. The dominating source of statistics, however, is from the follow-up of the Japanese bomb victims [9].

Cancer Risks After High-Dose Exposures

Natural and radiation-induced cancers cannot be distinguished from each other. In order to isolate the influence of radiation, the number of radiation-induced cancers must outnumber the expected variation in natural cancer incidence rates. Only extensive irradiations to many people can achieve this as, fortunately, the gathered experience from accidents and old medical procedures reveals that ionizing radiation is a fairly weak carcinogen. The excess lifetime morbidity risk is of the order of a few percent following a single dose of 100 mSv.

Compared with adults, small children and the human embryo are more prone to develop radiation-attributed cancers later in life. The rarity of natural childhood cancers also favors the prospect of identifying any such radiation effects. Cancer induction after in utero exposure to diagnostic X-rays has been studied extensively in the past and indicates an increase in childhood cancer risk by about 40% for doses of about 10–20 mGy ([10], Appendix G, §245). Today's investigation of the fetus with ultrasound has essentially removed this childhood cancer risk from the scene.

The most extensive radiation detriment data on an adult population stem from the follow-up of the Hiroshima and Nagasaki populations. The fatal-cancer incidence in a cohort of 50 000 persons who were significantly exposed, i.e. received a whole-body dose >5 mGy, has been compared with a control group of 36 500 persons exposed to <5 mGy. As seen in Tables 17.3 and 17.4, the lifetime risk attributable to the bombs exhibits a strong dependence on dose and age at exposure. As an average over gender and age, the lifetime cancer mortality probability in the H–N statistics is 10–12% and about 1% per Sv for solid cancers and leukemia, respectively.

Table 17.3. The lifetime excess probability due to radiation in the significantly exposed group in Hiroshima and Nagasaki with an average effective dose of 0.2 Sv. The natural lifetime risk for an unexposed person is shown for comparison [9]

Age at Exposure	Probability			
	Excess Women	Natural Women	Excess Men	Natural Men
10 y	0.05	0.19	0.03	0.26
30 y	0.03	0.20	0.02	0.28
50 y	0.01	0.15	0.01	0.18

Table 17.4. The percentage of fatal cancers attributable to the bombing of Hiroshima and Nagasaki as a function of dose [9]

Effective Dose Range	Attributable Fraction	
	Leukemia	Solid Cancer
5–200 mSv	14%	2%
200–500 mSv	45%	12%
500–1000 mSv	74%	23%
>1000 mSv	84%	39%

Cancer Risks After Low-Dose Exposures

Considering that a single or at least very few ionizations may alter cell genetics, radiation protection expert organizations such as the ICRP consider it pertinent to base radiation hygiene recommendations on a linear–no-threshold (LNT) hypothesis of cancer induction. This hypothesis simply states that the probability of dying from a radiation-caused tumor (including leukemia) is directly proportional to the effective dose E and that no threshold exists below which the probability is zero. Applying the epidemiological data given above to working-life conditions, the slope of the LNT curve is reduced by a factor of 2. This reduction is motivated by the fact that working-life exposures are typically protracted and of low dose rate, leaving room for cellular repair. This reasoning has led the ICRP to recommend so-called nominal risk coefficients for fatal cancer of 4.0 and 5.0% Sv^{-1} applicable to a working population and a population including also children, respectively, and to low-dose and low-dose-rate exposures.

The LNT hypothesis, which has given birth to concepts such as ALARA (as low as readily achievable) and collective dose (man-sievert), is a cornerstone of modern radiation protection recommendations. In recent years, however, the uncritical use of LNT has been questioned. The criticism is based on mainly two things:

- 1. Predicting low-dose (<50 mSv) consequences, for example the number of cancer mortalities, using the LNT hypothesis gives a false impression of knowledge that is not there.
- 2. The LNT hypothesis leads to a regulation of risks trivial to the individual, which causes mistrust in the regulatory authority.

Radiation epidemiology is not sensitive enough to resolve the low-dose issue, and one has to await future research in fields such as cellular genetics and signaling in order to quantify cancer probabilities after mSv (or lower) irradiations. Meanwhile, it is wise to extract, at most, a precautionary principle from the LNT hypothesis. The cancer risk at a dose below say 50 mSv presumably resides in the interval from zero up to at most 10 times the predictions of LNT. With such an uncertainty, it is sound to be very restrictive with risk quantifications. The complexity of cellular response to external agents such as ionizing radiation, revealed by modern cell physiology, strongly support such a cautionary posture.

Hereditary Effects

The detailed mapping of the human genome in recent years and the improved understanding of human hereditary processes indicate that the probability of hereditary effects following exposure to ionizing radiation is negligible compared with the natural risks in the dose range below 100 mGy (see Table 17.5).

Hereditary diseases are broadly classified as either single-gene (Mendelian), chromosomal and multifactorial types. Multifactorial diseases result from a

Table 17.5. Estimates of hereditary risks to offspring from exposure to a single parent (1st) generation and two (1st and 2nd) parent generations. The assumed doubling dose is 1 Gy. From [11]

Disease Class	f_s	r_D (Gy ⁻¹ per 10 ⁶ progeny in generation)		
	Per 10 ⁶ live births	one 1st generation exposed	two	two both gen. exp.
<i>Mendelian</i>				
Dominant + X	16 500	750–1500	500–1000	1300–2500
Recessive	7 500	0	0	0
<i>Chromosomal</i>	4 000	a	a	a
<i>Multifactorial</i>				
Chronic	650 000 ^b	250–1 200	250–1 200	250–1 200
Congenital ^a	60 000	2 000	400–1 000	2 400–3 000

^a The chromosomal damage is included partly in the class of dominant and X-linked diseases, and partly in the class of congenital abnormalities.
^b Frequency in population.

large number of interacting genetic and environmental factors, and the natural incidence is high. The multifactorial class includes both abnormalities seen at birth (congenital abnormalities), and adulthood diseases such as high blood pressure and diabetes.

The so-called doubling dose is a basic concept for hereditary-risk estimations. The doubling dose D_{dbl} is the absorbed dose that provokes as many mutations as those which occur spontaneously in one generation.

For dominant single-gene disorders, a mutation will ultimately, if not lethal, give a hereditary ailment in the living offspring. For the recessive type of disorders, and for polygenic and multifactorial diseases, the probability of manifested hereditary damage in the first-generation offspring may be close to zero. In order to take this into consideration, the probability of mutation is multiplied by a disease incidence factor Q . The hereditary risk per unit dose, r_D , can then be written in the form $r_D = f_s D_{dbl}^{-1} Q$ (Gy^{-1}), where f_s is the baseline spontaneous-incidence rate. For the incidence factor, we have $0 \leq Q < 1$, and on the basis of mouse data, it is believed that the doubling dose for most human gene locations is at least 1 Gy.

In Table 17.5, an estimation of the hereditary risks is summarized. It must be remembered that hereditary damage may affect quality of life in a way that ranges from the almost unnoticeable to the most severely disabling. This implies that it is not very meaningful to add the incidence rates of Table 17.5 together without weighting for the severity of the hereditary disorder in question. It is also obvious from the figures in Table 17.5 that the assumed doubling dose of 1 Gy far from doubles the incidence rates in the first few generations of offspring, i.e. the Q factor is of the order of one percent or lower. This statement is also valid in the case where the irradiation is duplicated in the first few generations.

A permanent and significant increase in gonad dose to a population will, however, in the long run, change the baseline frequency f_s , and in this new hereditary equilibrium, the increased mutation rate will result in an identical increase in disease frequency.

17.5 Detecting Ionizing Radiation

All detectors for ionizing radiation are based on the same fundamental principle – the transfer of energy to the detector. In the detector, the energy is converted into some other form that can be registered. Modern detectors are essentially electrical – at some point, the information from the detector is transformed into electrical pulses that can be treated by electronic means.

It should be recognized that all instruments are limited to measuring certain types of radiation within a fixed range of energy. Outside these limits, instrument readings are not to be trusted.

17.5.1 X-Ray, β -Ray and γ -Ray Detection

Ionization detectors were the first electrical devices developed for radiation detection. These instruments are based on the direct collection of the ionization electrons and ions produced in a gas by passing radiation. Three basic types of detectors have been developed – the ionization chamber, the proportional counter and the Geiger–Müller counter. These types of detectors are today mostly used as radiation monitors as they are cheap, simple to operate and easy to maintain.

The *scintillation detector* makes use of the fact that certain materials when irradiated emit small flashes of light, i.e. they scintillate. When coupled to an amplifying device such as a photomultiplier, this scintillation light is transmitted through a shaped light pipe to the photocathode of a photomultiplier. There, photons release electrons, which are accelerated and focused onto the first dynode. For each primary electron hitting a dynode, between two and five secondary electrons are released. Up to 15 multiplying stages can be used, reaching overall multiplying factors of up to 10^9 . A few incident photons therefore produce a measurable electrical pulse, which can then be analyzed and counted electronically.

Semiconductor detectors are based on crystalline semiconductor materials, most notably silicon and germanium. The basic operating principle of semiconductor detectors is analogous to that of gas ionization devices. However, the medium is now a solid material. The passage of radiation through the solid material creates electron–hole pairs (instead of electron–ion pairs). The advantage of a semiconductor is that the average energy required per electron–hole pair is some 10 times smaller than that required for gas ionization. Thus, the amount of ionization produced for a given particle energy is an order of magnitude greater, resulting in superior energy resolution. Moreover, semiconductor detectors have a greater density and therefore a greater stopping power than gas detectors. As a result, they are more compact in size and can have a very fast response time.

17.5.2 Tritium Detection

It is important that proper ventilation techniques are used to ensure that any release of tritium is adequately exhausted from the laboratory. It is also necessary to provide air monitoring for releases of tritium gas. A common device for tritium monitoring is an ionization detector, through which air is drawn at a fixed rate. The ionization chamber may be preceded by an ion collector and/or a filter in order to remove other ions or radioactive material. Tritium contamination can be detected and analyzed through the use of samples collected with filter paper or other type of traps and then transferred to a liquid scintillation solution. Liquid scintillation counters are also the instrument of choice for monitoring tritium in urine samples.

17.5.3 Neutron Detection

Neutron radiation may be accompanied by relatively high levels of γ -radiation. Consequently, in order to measure the neutron level adequately, it is necessary to have an instrument insensitive to γ -radiation. Ionization instruments, in general, are usually not satisfactory for measuring neutrons, since they are also sensitive to γ -radiation.

The boron trifluoride (BF_3) proportional counter provides a sensitive detector for a neutron survey instrument and can be relatively insensitive to γ -radiation. This instrument uses BF_3 gas enriched to more than 95% in the isotope ^{10}B , and typically the detector has about a 100 mm active length. The BF_3 counter is sensitive to thermal neutrons. It is also possible to detect intermediate and fast neutrons by enclosing the detector in polyethylene or paraffin wax, which brings down the neutron energy before the neutrons enter the BF_3 gas. With a suitable moderator configuration, it is possible to achieve a count rate which approximates the dose-equivalent rate $\dot{H}$ for neutrons in the energy range below 10 MeV.

The ^3He neutron detector is more sensitive than the BF_3 counter. The detector is based on the reaction $^3\text{He}(n, p)\text{T}$. The proton and the tritium will ionize the gas and an electrical pulse will be obtained for every absorbed neutron. A disadvantage is that the ^3He gas is very expensive, and to have a sensitive detector, the detector needs to be large. The sensitivity is strongly dependent on the ^3He gas pressure, volume and enrichment and can be up to several hundred pulses per incoming neutron per cm^2 .

Cerium-doped lithium silicate glasses are widely used as neutron detectors. Recent developments incorporate this scintillator into plastic fibers acting as light waveguides [12]. Both flat and bent large-area neutron detectors can be built by this technique. The light output is monitored at the end of the glass fibers.

17.6 Radiation Dose Measurements

Generally, suitable dosimeters tend to mimic human tissue in atomic composition in order to avoid the problems of converting the measured signal to another material that is too different. This is of special importance when dealing with the sometimes very strongly atomically dependent cross sections of neutrons, and in gamma radiation field measurements for less well-known spectral distributions. Now, this choice of a suitable atomic composition may not be a problem for the user at an accelerator laboratory, since the decision on composition has already been taken by the instrument manufacturer. The important thing is that the manufacturer can describe the application areas for the instrument and that the instrument is calibrated in quantities acceptable to the radiation protection authority surveying safety issues. The user's

responsibility is to use a dosimeter suitable for the radiation to be monitored, to check the functionality and to calibrate the dosimeter on a regular basis.

In accelerators emitting pulsed radiation, the capability of a dosimeter to cope with the high dose rates in the pulse peak needs to be considered. As mentioned in Sect. 17.6.1, some solid-state dosimeters can reproduce very high dose rates accurately. Many type of gas-filled detectors may be duty-cycle-dependent in the high-dose-rate domain. Further guidance on this subject is given in the ICRU Report 20 [13].

An issue that accelerator dosimetry measurement has in common with all high-energy photon applications involving small radiation field sizes is the dose buildup within surfaces facing air or gaseous media in the direction of the incoming photons. When photons irradiate a solid material in air, the maximum dose in the solid is reached at a depth corresponding to the range of the secondary electrons generated in the solid by the photons. Thin-walled ionization chambers, open dosimeters etc. must be covered with an appropriate amount of material to register this maximum properly. A set of buildup covers of different thicknesses may be necessary in order to handle the diverse photon geometries and energies encountered.

In order to perform dose measurements that can be used for legal purposes, the dosimeter used must be calibrated with respect to or traceable to established dose quantities, such as the environmental dose equivalent $H^*(10)$. With modern dosimeters, this is usually no problem.

17.6.1 Personal Monitoring

Individual dosimetry is normally mandatory in all controlled areas of an accelerator. The exposure of people can be measured by using pocket pen dosimeters, film badges, pocket ionization chambers, thermoluminescent dosimeters and diode detectors. There is, however, a trend today allowing only legally authorized dosimeter types from likewise authorized dosimetry laboratories, when the dosimeter is intended for personal record-keeping.

Film badges can be made sensitive to γ - and X-rays, electrons, and neutrons. The use of filters makes possible the separation of different kinds of radiation and different energies. The radiation produces blackening of the emulsion, which can be determined photometrically after development. Integrating dosimeters based on the thermoluminescence (TL) principle are gradually replacing the old film badge as the dominating legal dosimeter. A thermoluminescent material stores the ionization energy in the crystalline lattice in such a way that after exposure it can be released in the form of light when the material is heated. The amount of light is directly proportional to the absorbed dose accumulated by the dosimeter material. Following special routines for how the material is temperature-treated and stored when not in use, a TL dosimeter can be reused hundreds of times. Compared with the film badge, the TL dosimeter is superior in sensitivity, dose and dose-rate

linearity, and energy dependence for photons. The most common TL phosphor, LiF, is linear in dose rate up to 10^9 Gy s^{-1} [14], a useful capacity in pulsed radiation fields. A TL badge can hold several small TL tablets, for instance one with high sensitivity, e.g. ^6LiF , to neutrons and one with low sensitivity, e.g. ^7LiF . Film and TL badges should be read out on a regular basis. A common integration time is 1 month.

The pocket pen dosimeter is based on the old electroscope-discharging principle and is directly readable by the wearer. Modern direct-reading electronic personal dosimeters (EPDs), mostly based on multidiode designs, have made the pocket pen dosimeter obsolete. Today EPDs come in lightweight designs ($<100 \text{ g}$) offering separate shallow- and deep-dose readings, dose and dose rate histories, and visual and audible alarms for dose and dose rate, and with computer software for easy changing of dosimeter parameters, visualization of dose history etc. EPDs have high dose sensitivity and are usually calibrated in both $H_p(10)$ and $H_p(0.07)$ (see Sect. 17.3.3).

17.7 Radiation Hazards

The primary beam creates secondary radiation by collision with atoms in its path, and induced activity by nuclear reactions. The secondary radiation is prompt and can be a problem only while the beam is on, while the induced activity constitutes a source that is also alive after the beam has been stopped. The small aperture and high intensity of the primary beam makes it a dangerous source of radiation that can cause local radiation injuries to the human body. Fortunately, in most accelerators the primary beam is enclosed in such way that it is not accessible. As a rule, bremsstrahlung and neutrons, owing to their high penetrating abilities, dominate the radiation hazard and determine the degree of protection necessary in an accelerator environment.

17.7.1 Electron Accelerators

Electrons

The primary beam of electrons may be accessible if the beam is brought out into the room air. The acceleration of electrons by an electrostatic accelerator means a limit on the kinetic energy T of the electrons of about 5 MeV. The maximum range of electrons is approximately proportional to the inverse density of the stopping material, with the consequence that electron ranges are usually expressed in surface weight units, e.g. g cm^{-2} . In the energy interval 2–5 MeV, the electron range in any material in cm can in practice be set to $0.6 \times \rho^{-1}T$, if T is in MeV and the density ρ is in g cm^{-3} . This means that 5 MeV electrons are stopped by a few cm of unit-density material, while the range in air is about one thousand times longer. It is a good practice to stop the electrons in a low-atomic-number material as this will minimize bremsstrahlung formation.

Gamma Radiation

Any charged ionizing particle may lose energy in a bremsstrahlung (X-ray) process. The energy distribution of the generated bremsstrahlung is continuous, with a maximum energy equal to the initial electron energy. The cross section for the process in a material is approximately proportional to Z^2/m^2 , where Z and m are the atomic number of the material and the mass of the charged particle, respectively. Thus, bremsstrahlung is a potential safety problem only for electrons.

The angular dependence of the X-rays generated in thick targets is to a large extent governed by the angular distribution of the electrons that are slowing down. Increasing the electron energy from, say, 100 keV and upwards causes the X-ray beam to tilt more and more in the electron direction. The photon energy fluence at large angles, say larger than 60° , is dominated by low-energy photons created by energy-degraded electrons in the target. For increasing primary electron energies, this behavior will be more pronounced.

Low-energy X-rays are generated wherever low-energy electrons are decelerated, for example in cavities and klystrons. For electrons in the energy interval 0.5–5 MeV, the forward-scattered photons create a dose-equivalent rate of roughly $14T^3$ Sv h⁻¹ at 1 m per mA if T is expressed in MeV [15].

Every point in an accelerator in which high-energy electrons strike matter is a bremsstrahlung source. Wherever the electron beam is deflected, or passes through slits, windows, collimators and so forth, enhanced emission of bremsstrahlung photons should be suspected. Poor vacuum is another source of bremsstrahlung, owing to collisions of electrons with residual-gas molecules. A partial loss of vacuum can be almost instantaneous, and the subsequent rapid and pronounced increase in dose rate along the beamline makes this kind of failure extra insidious.

Neutron Radiation

Photonuclear processes, such as (γ , n), show a broad resonance cross section (the giant resonance) centered around 15 to 25 MeV in most materials. The threshold is not lower than 6 MeV, with exceptions for deuterium (2.2 MeV) and ^9Be (1.67 MeV). Thus, neutrons are not produced by electrostatic electron accelerators unless deuterium or beryllium targets are deliberately exposed to bremsstrahlung.

17.7.2 Proton and Light-Ion Accelerators

Compared with electrostatic electron accelerators, the radiation environments around accelerators used for light ions up to and including helium ions are more complex. The much lower intensity of bremsstrahlung and the short range of the primary radiation facilitate the radiation protection work.

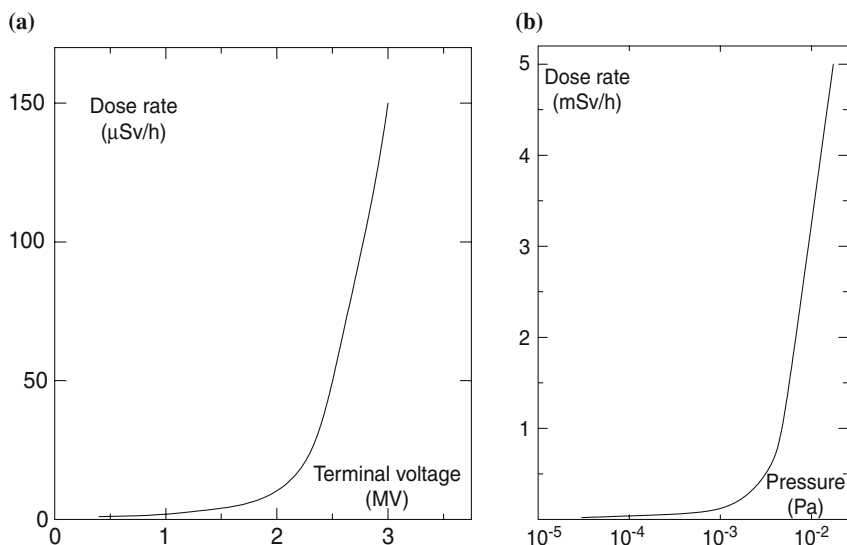


Fig. 17.2. Bremsstrahlung dose rate 1 m from the accelerator tank just opposite the terminal for $\sim 3\mu\text{A}$ protons injected into a fairly new 3 MV tandem Pelletron. (a) As a function of the terminal voltage up to the maximum design voltage. The pressure at the low-energy side just outside the tank was 1×10^{-5} Pa and the pressure in the terminal was 5×10^{-4} Pa; (b) As a function of the mean pressure in the low-energy accelerator tube. The curve was obtained at a terminal voltage of 2.0 MV

Bremsstrahlung cannot be avoided, however, as electrons are unintentionally accelerated in all evacuated high-voltage parts of the accelerator.

Examples of the bremsstrahlung around an electrostatic accelerator are shown in Fig. 17.2. The curves were obtained with a 3 MV tandem Pelletron using a beam of protons. Similar curves have been obtained with other ion beams. The detector was an organic scintillator designed for monitoring low background levels. Figure 17.2a demonstrates the strong increase of the dose rate when the terminal voltage approaches the maximum design voltage. The strong dependence of the dose rate on the pressure in the accelerator tube is demonstrated in Fig. 17.2b. A similar strong increase of the dose rate can also be obtained by changing the optics of the beam entering the tube so that they become worse. Both figures were obtained for a constant beam current. If the beam current is changed (keeping the terminal voltage and tube pressure constant), a linear relation between dose rate and beam current is obtained (at least for moderate currents, 1–10 μA).

Many light-ion and proton reactions create neutrons, and it is usually these neutrons that determine the radiation shield necessary. The most common shielding material is concrete. The high content of hydrogen in the concrete efficiently slows down the neutrons to thermal energies and the neutrons

are then, with a high probability, lost by a capture process in hydrogen, i.e. ${}^1\text{H}(n_{th}, \gamma){}^2\text{H}$. In the capture a 2.2 MeV photon is created, which, in turn, because of its high penetration, may cause a protection problem. The mean free path in ordinary concrete is about 0.1 m for 2.2 MeV photons. To avoid the production of 2.2 MeV gamma radiation, boron can be mixed into the concrete. The capture cross section of boron due to the reaction ${}^{10}\text{B}(n_{nt}, \alpha){}^7\text{Li}$ is about 10 000 times bigger than the corresponding cross section of hydrogen capture ditto, and the emitted gamma ray has an energy of only 0.48 MeV.

Protons

The neutron yield from protons impinging on thick targets increases rapidly with energy, from the order of per mill per proton at 10 MeV to tens of percent at 100 MeV. Generally, the neutron yield for most materials is similar to that for copper and iron, or lower; it may be even a factor of ten lower for proton energies well above the threshold for neutron production in the material [15, 16]. (Materials with low thresholds are discussed in the next section.)

17.7.3 Low-Voltage Neutron Generators

High neutron dose rates can also be created in low-voltage accelerators with proper projectile and target materials. Low-voltage neutron generators function by accelerating positive ions across a potential drop of a few hundred kV or less. Because of the high yield of energetic neutrons, the reaction of deuterium with tritium as the target is the method of choice. The reaction ${}^3\text{H}({}^2\text{H}, n){}^4\text{He}$ (also denoted as $\text{T}(\text{d}, n)$ or D-T) is exoergic with a large energy release, 17.6 MeV. For thin targets, the maximum yield occurs at a deuteron energy of 107 keV. For thick tritium targets, the yield increases rapidly up to 600 keV and then more slowly [17]. The energy of the neutrons created is somewhere in the interval 13–16 MeV for accelerating potentials less than 400 kV, with the maximum and minimum energies in the forward and backward directions, respectively. By convention, these neutrons are referred to as 14 MeV neutrons, from the kinetic energy in the center-of-mass coordinate system.

The typical yield of 14 MeV neutrons per mA for 150 keV deuterons gives a fluence rate of about $8 \times 10^5 \text{ cm}^{-2} \text{ s}^{-1}$ at one meter from the target. This fluence rate is equivalent to an effective dose rate of 1.4 in the AP and 1.1 Sv h^{-1} in the ROT geometry (see Sect. 17.B.5).

Another exoergic reaction used for low-voltage neutron accelerators is the D-D (deuteron-deuteron) reaction. The neutron yield at 150 kV is less than one percent of that for the D-T reaction and the neutron energy is much lower, in the range 2–3 MeV [17]. The effective neutron AP dose per mA at 1 m from the deuteron target is approximately 0.6 mSv h^{-1} . D-D accelerators

do not generally present any important radiation hazards compared with D-T generators. It should be noticed, however, that if deuterium is accumulated somewhere in the beamline before the actual target, extra neutron sources outside the expected one are created.

17.7.4 Heavy-Ion Accelerators

Typically, accelerating lithium and heavier ions generates only moderate radiation problems if the projectile ions are kept within the vacuum part of the accelerator. Hazards to pay attention to are neutron-producing reactions and the source of characteristic X-rays created in the material in which the ions are stopped, and, as in all “evacuated” electrostatic environments, bremsstrahlung from inadvertently accelerated electrons.

17.7.5 Induced Radiation

In contrast to the particle beam itself, the induced activity in accelerator materials and components constitutes a delayed source of radiation that exists also after the accelerating voltage has been switched off. The thresholds for nuclear reactions in ordinary metals and materials are such that induction of radioactivity is an almost nonexistent problem for electrostatic electron accelerators. Proton and deuteron beams and the accompanying neutrons may produce measurable quantities of activity, but the specific activity concentrations in accessible areas are usually too low to be hazardous. In addition, the majority of radionuclides are short-lived, and by delaying access to target areas after shutdown and delaying maintenance programs by one or a few days, the exposure of personnel can be minimized. Evidently, the highest activity concentrations should be expected in targets, beam dumps, slits and collimators.

The unavoidable induction of activity in components and materials is complex, and the reader is advised to consult the rich literature (e.g. [15, 18, 19]) on the subject to solve specific problems.

For light-ion electrostatic accelerators, thermal-neutron reactions induce activity in air, copper, steel, concrete etc. The resulting specific activity is usually less than 1 Bq g^{-1} and with negligible radiation risks to the personnel. Long-lived products, such as ^{60}Co , however, may cause administrative problems at decommissioning, owing to the very strict regulations for disposal of weakly radioactive materials.

17.7.6 Tritium Targets

Tritium, or ^3H , is the only radioactive isotope of hydrogen. It is a pure beta emitter with a half-life of about 12 years. The maximum energy of the beta particles is only 18.6 keV, corresponding to a range of less than about

0.6 mg cm^{-2} . Thus, tritium is only a potential hazard when taken into the human body, for example by inhalation. Even so, the dose factor is low, owing to a short biological half-life, roughly 2 weeks for most compounds, and the low ionizing energy emitted per decay. The low dose factor, about 30 mSv for an intake of 1 GBq tritium, must, however, be seen in the perspective of the extremely high tritium target activities that exist and the natural tendency of tritium to become airborne, particularly at elevated temperatures.

The typical target construction is tritium adsorbed in a thin metal layer. Tritium desorbs spontaneously from the metal at room temperature and is, furthermore, driven out by the bombardment with projectile particles. Under normal conditions, the tritium leaving the target is exhausted via the vacuum pumps and the ventilation. The tendency of tritium to become airborne implies that tritium targets not in use should be stored in airtight containers. Such containers must be opened in well-ventilated places.

The bulk of the tritium desorbing from a target will pass straight through the vacuum pumps used and be ventilated out. Small amounts of tritium can be found in pump oils and other components. One exception is pumps of the ion-getter type. These retain tritium and may contain TBq of tritium if used for many targets [17].

17.8 Safety Considerations

17.8.1 Administrative Safety Program

The work at an accelerator facility must be organized in such a way that risks from radiation and other hazards are minimized. The details of an accelerator safety program depend on local conditions, but the basic safety strategies to follow have been set out by various national and international bodies (see for instance [20–23]). Evidently, any accelerator safety program must conform with the current national legislation. For many accelerators, the normal situation is a low radiation intensity. This is a risk, as staff easily become careless, considering the radiation hazard as negligible. This may lead to insufficient notice being taken of changed conditions (increased pressure in the vacuum system, new components of unsuitable material along the beam path, use of a target material with a low (p, n) threshold etc.). It is necessary to avoid careless habits in an accelerator laboratory. An accelerator must be respected for its maximum radiation capability rather than its customary conditions. In this section, we give some general comments on the administrative part of a radiation safety program, leaving more technical aspects to Sect. 17.8.2.

To be effective, a radiation safety program must have full support, both mental and budgetary, from the management, but also be proportionate and well adapted to all interests in the organization. A radiation safety program should, in a radiation safety manual, clearly define the duties and responsibilities of all employee categories, including the management, and state the

qualification requirements at different levels in the organization. The duties and the possible delegation in the safety organization should preferably refer to named persons, and the delegation scheme should be known and easily accessible to all employees. The safety program should cover topics such as categorization of workers and workplaces, education and training of personnel, routine controls and inspections, and rules for visitors. The safety procedures may be quite extensive, and it is hence important to state in the manual how safety work should be documented and reported. The local radiation expert in charge is a key person concerning radiation safety issues and it is good practice to let him/her report directly to the management of the facility in order to avoid on-the-floor conflicts between research and safety interests. At large-scale accelerator facilities, the expert normally reports to some kind of radiation safety steering committee.

An accelerator workplace can be very dynamic, with shifting beam rooms and intensities, mobile protective shields, etc. The safety manual must define who is authorized to make changes to beamline setups and how changes with significance for safety shall be reported in the organization. Pre-prepared plans for how to handle emergencies and accidents are also an important part of an operational safety program. Conventional accidents should not be forgotten, as they are normally much more probable than irradiation mishaps involving radiation doses to employees.

17.8.2 Technical Safety

The radiation safety at an accelerator facility depends on numerous factors. In addition to administrative procedures and the shielding the building itself provides, technical systems aimed at minimizing radiation and other hazards to the staff are of great significance. The ideal technical safety system is, if necessary, redundant; works passively in the background; is not sensitive to human errors; and does not interfere unduly with the primary accelerator assignments.

Shielding the staff from external radiation, and the use of distance, stationary (wall) shields and locks are preferred as compared with warning lights, audio alarms and other techniques more dependent on human behavior. The amount of shielding material must be dimensioned in accordance with the maximum practicable workload of the accelerator and the most “unfavorable” particle accelerated. It must be ensured that radiation attenuation at doors, windows, ventilation ducts and all other penetrations through a shielding wall, is sufficient or is remediated by additional shielding blocks. The effect of sky-shine through the roof must be considered, as this can provide a significant contribution to the total radiation levels observed outside thick shielding walls.

Concrete is a natural choice for a shielding material. Both as a part of the building construction and as additional mobile shields, concrete is inexpensive, reliable, structurally useful and easily installed. Owing to these

practical aspects and its content of hydrogen, concrete is especially suitable for shielding against neutrons (see Sect. 17.7.2).

Access control and radiation warning systems are essential parts of a radiation safety program, particularly at facilities provided with several target rooms. A common way to prevent entry to high-radiation areas is by interlocks. If the interlock is broken or loses power, the beam should be stopped at a safe position and a restart must commence at the position of the interlock. There is a great deal of variation in the interlocking arrangements to be found in different laboratories. This depends on differences in accelerator-specific space or component limitations and on differences in design philosophy, cost factors and users' requirements.

In high-level radiation areas accessible to staff, the control circuits of the accelerator should produce a warning by light or an easily identified sound prior to start-up, to allow anyone trapped in the laboratory to leave the area. The duration of the warning must be long enough for a person to safely leave the area or reach the nearest scram switch.

Detectors to monitor the X- and γ -ray and neutron radiation levels in various areas have been installed in some laboratories. Such area monitoring is not necessary if the radiation level is unlikely to change. Area monitors are sometimes used to switch off the beam or give a warning if the permissible level has been exceeded. One has to remember that even if the charging supply is switched off, a high voltage – and therefore an increased radiation level – can be obtained because of self-charging, whenever the belt (or chain) is moving.

It is important that survey and monitoring instruments are properly maintained and calibrated. The use of a check source for verifying the fitness of the detector prior to each measurement is highly recommended. At least every second year, the absolute sensitivity of the detector should be controlled against secondary-standard instruments or calibrated in a standards laboratory.

17.A Appendix: Recommendations by ICRP

Most national regulatory standards in the field of ionizing-radiation protection are based on recommendations issued by the International Commission on Radiological Protection (ICRP). Two principles are recommended for practices involving exposure to a workforce: firstly, the exposure must be justified, and secondly, all justified exposures must be optimized [5]. Justification means that no practice should be adopted unless the benefits to exposed individuals or to society balance out the harm caused by the radiation exposure. The aim of the optimization is to keep all exposures as low as readily achievable (ALARA), taking also economic and social factors into account.

In order to guarantee that justified and optimized exposures do not lead to unacceptable risks to any individual, the ICRP also recommends dose limits. The basic recommendation for workers is that the effective dose E should be kept below 100 mSv average over any consecutive period of 5 years and below 50 mSv in any single year. The annual equivalent dose to skin, hands and feet should be limited to 500 mSv, and for the lens of the eye, the corresponding value is 150 mSv.

Additional restrictions apply to occupational exposure of pregnant women. The ICRP recommends that the working conditions of a pregnant worker, after the declaration of pregnancy, should be such that it is unlikely that the additional equivalent dose to the conceptus will exceed about 1 mSv [21].

The recommendations by the ICRP are extensive, but generic in character, and for operational radiation protection work, applicable national legislation and local radiation safety rules need to be consulted for more practical guidance.

17.B Appendix: Half-Value Layer (HVL) and Fluence–Dose Conversions

17.B.1 HVL for Photons

The thicknesses needed to attenuate the fluence rate of primary photons to half of its value for the photon energy range between 0.1 keV and 3 MeV are given in Table 17.6.

Table 17.6. The thickness needed to attenuate the fluence rate of primary photons to half of its value (the HVL) for different materials. Note that in broad geometries, the contribution from secondary photons can be significant. In narrow-beam geometries, the HVL is related to the attenuation coefficient μ by $\text{HVL} = (\ln 2)/\mu$

Photon Energy keV	Air m	Water cm	Al mm	Fe mm	Pb mm	Glass mm	Concrete mm
0.1	1.1×10^{-4}	1.2×10^{-5}	—	—	—	—	—
1	1.5×10^{-3}	1.7×10^{-4}	2.4×10^{-3}	1.0×10^{-3}	1.2×10^{-4}	8.9×10^{-4}	8.7×10^{-4}
3	3.3×10^{-2}	3.6×10^{-3}	3.3×10^{-3}	1.7×10^{-3}	3.1×10^{-4}	5.5×10^{-3}	6×10^{-3}
10	1.1	0.13	0.10	5.2×10^{-3}	4.7×10^{-3}	0.17	0.11
30	15	1.9	2.3	2.3	2.0×10^{-2}	3.5	2.4
100	35	4.1	15	2.4	0.11	17	16
300	50	5.8	25	8.0	1.5	26	27
1000	84	9.8	42	15	8.6	44	46
3000	150	17	73	24	14	77	80

17.B.2 Photon Effective-Dose Factors

The effective dose (in pSv) per unit photon fluence (in cm^{-2}) for a standard adult person, as a function of photon energy, is illustrated in Fig. 17.3a. Curves for three different irradiation geometries have been included.

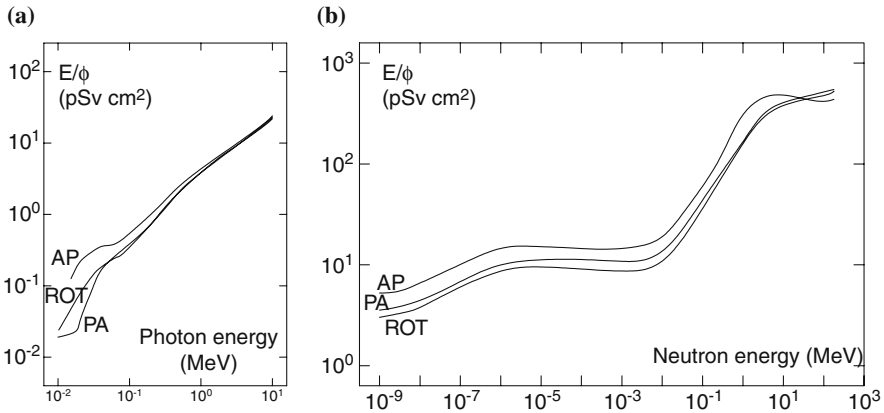


Fig. 17.3. The effective dose E per unit photon (a) and neutron (b) fluence in units of pSv cm^2 for a standardized adult person. AP = irradiation in the direction front to back, PA in the direction back to front, and ROT in all horizontal directions [24]

17.B.3 Neutron Cross Sections

Neutron interaction cross sections vary widely with the neutron energy and target material. This complexity of the interaction makes accurate neutron-shielding calculations demanding and extensive, and calculations utilizing intricate computer programs are the norm. Databases covering neutron cross sections, thick-target yields etc. can be found on the Internet, for example at <http://www-nds.iaea.org/> and <http://www.nndc.bnl.gov/nndc/exfor/>.

17.B.4 Neutron Attenuation

Several textbooks and other publications have gathered data for neutron transmission through concrete and other materials (e.g. [2, 25]). The transmission factor depends on the initial neutron energy distribution, the angle of incidence, the isotopic content of the shielding material, the quantity of interest etc. The allotted space does not, unfortunately, permit us to cover the complex field of neutron attenuation in this book.

17.B.5 Neutron Effective-Dose Factors

The effective dose (in pSv) per unit neutron fluence (in cm^{-2}) for a standard adult person, as a function of the neutron energy, is illustrated in Fig. 17.3b. Curves for three different irradiation geometries have been included.

References

1. D. Brune, R. Hellborg, B.R.R. Persson, R. Pääkkönen: *Radiation – at Home, Outdoors and in the Workplace* (Scandinavian Science Publishers, Oslo 2001)
2. E. Profio: *Radiation Shielding and Dosimetry* (Wiley, New York, 1979)
3. International Commission on Radiation Units and Measurements: *Quantities and Units in Radiation Protection Dosimetry*, ICRU Report 51 (International Commission on Radiation Units and Measurements, Bethesda, MD, 1993)
4. International Commission on Radiation Units and Measurements: *Fundamental Quantities and Units for Ionizing Radiation*, ICRU Report 60 (International Commission on Radiation Units and Measurements, Bethesda, MD, 1998)
5. International Commission on Radiological Protection: *1990 Recommendations of the International Commission on Radiological Protection*, ICRP Publication 60 (Annals of the ICRP, Vol. 21, No. 1–3, Pergamon, Oxford, 1991)
6. E. Hall: *Radiobiology for the Radiologist*, 5th edn. (Lippincott Williams & Wilkins, Philadelphia, 2000)
7. International Commission on Radiological Protection, *1990 Pregnancy and Medical Radiation*, ICRP Publication 84 (Annals of the ICRP, Vol. 30, No. 1, Pergamon, Oxford, 2000)
8. United Nations Scientific Committee on the Effects of Atomic Radiation: *Sources, Effects and Risks of Ionizing Radiation*, Report to the General Assembly, with Scientific Annexes (United Nations, New York, 1988)
9. Radiation Effects Research Foundation: <http://www.rerf.jp>
10. United Nations Scientific Committee on the Effects of Atomic Radiation: *Sources and Effects of Ionizing Radiation*, Report to the General Assembly, with Scientific Annexes (United Nations, New York, 2000)
11. United Nations Scientific Committee on the Effects of Atomic Radiation: *Hereditary Effects of Radiation*, UNSCEAR 2001 Report to the General Assembly, with Scientific Annexes (United Nations, New York, 2001)
12. M.J. Weber, M. Bliss, R.A. Craig, D.S. Sunberg: *Radiat. Eff. Defects Solids* **134**, 23 (1995)
13. International Commission on Radiation Units and Measurements: *Radiation Protection Instrumentation and Its Application*, ICRU Report 20 (International Commission on Radiation Units and Measurements, Bethesda, MD, 1971)
14. F. Attix: *Introduction to Radiological Physics and Radiation Dosimetry* (Wiley, New York, 1986)
15. A.H. Sullivan: *A Guide to Radiation and Radioactivity Levels Near High Energy Particle Accelerators* (Nuclear Technology Publications, Ashford, England, 1992)
16. K. Tesch: *Radiat. Prot. Dosim.* **11**:3, 165 (1985)

17. National Council on Radiation Protection and Measurements: *Radiation Protection and Measurement for Low-Voltage Neutron Generators*, NCRP Report 72 (National Council on Radiation Protection and Measurements, Washington, DC, 1983)
18. M. Barbier: *Induced Radioactivity*, (North-Holland, Amsterdam, 1969)
19. C. Biratti, M.C. Cantone, A. Ferrari, M. Silari: Nucl. Instr. Meth. B **43**, 119 (1989)
20. National Council on Radiation Protection and Measurements: *Operational Radiation Safety Program*, NCRP Report 127 (National Council on Radiation Protection and Measurements, Washington, DC, 1998)
21. International Commission on Radiological Protection, *General Principles for the Radiation Protection of Workers*, ICRP Publication 75 (Annals of the ICRP, Vol. 27, No. 1, Pergamon, Oxford, 1997)
22. American National Standards Institute: *Radiological Safety in the Design and Operation of Particle Accelerators*, ANSI N43.1 (American National Standards Institute, New York, 1979)
23. L.E. Moritz: Radiat. Prot. Dosim. **96**:4, 297 (2001)
24. International Commission on Radiation Units and Measurements: *Conversion Coefficients for Use in Radiological Protection Against External Radiation*, ICRU Report 57 (International Commission on Radiation Units and Measurements, Bethesda, MD, 1998); also published by International Commission on Radiological Protection, as ICRP Publication 74 (Annals of the ICRP, Vol. 26, No. 3–4, Pergamon, Oxford, 1996)
25. National Council on Radiation Protection and Measurements: *Protection Against Neutron Radiation*, NCRP Report 38 (National Council on Radiation Protection and Measurements, Washington, DC, 1971)

Box 8: Nonradiation Hazards and Safety Considerations

R. Hellborg¹ and C. Samuelsson²

¹ Department of Physics, Lund University, Sölvegatan 14, 223 62 Lund, Sweden
`ragnar.hellborg@nuclear.lu.se`

² Department of Medical Radiation Physics, Lund University, University Hospital, 221 85 Lund, Sweden
`christer.samuelsson@radfys.lu.se`

In Chap. 17, the radiation hazards at an accelerator laboratory were emphasized. In this box, the nonradiation hazards are discussed. The most conventional of these – electrical hazards, fire and explosion hazards, mechanical hazards, and toxic hazards – are no less important in the operation of accelerators and should not be overlooked. In some countries, a risk assessment of the planned activity must be undertaken before a laboratory experiment.

Electrical hazards of a broad range are present in an accelerator laboratory. The most serious danger is that of fatal electric shock. However, even nonfatal electric shocks can cause severe injury, for example as a result of falls or impact against nearby equipment. Electrical burns, caused by the passage of a large current through the skin or tissue, are often deep and painful and heal slowly. Electrical accidents frequently cause serious damage to equipment, including damage by fire and smoke. No set of rules can completely insure safety, but they should include the following:

1. Remove fuses or switch off breakers to prevent accidental application of power when working on the equipment.
2. Provide and use grounding hooks.
3. Never work alone on electrical equipment.
4. Never take chances.
5. Avoid forming a circuit to ground through your body.

Not to be overlooked are interlock systems dealing with extreme electrical hazards, particularly high-voltage DC power supplies. Switches on the access doors to high-voltage compartments should remove power and discharge capacitor banks when the doors are opened. In general, any electrical hazard which has a protective enclosure should have interlock switches to open the power circuit if the enclosure is open. For RF systems, two categories of hazard exist: firstly, the electrical hazard associated with high voltage, and secondly, the hazard associated with the RF power, which can produce intense local heating in nearby objects. Intense RF fields should be confined by screen enclosures. Devices which present extreme hazards include injectors, electron guns and ion sources. Injectors can be biased at up to several hundred kV, which is high enough to create ionization of the surrounding air, and

can cause sparks in air over long paths. Personnel should therefore be kept at a safe distance away from equipment operating at these voltages. Many laboratories have interlocked enclosures equipped with doors that automatically ground the source deck when open.

The principal *fire and explosion hazards* are associated with that of electrical fires. The hazard can be high, as complex electrical apparatus is often used close to flammable or explosive materials. Those can be combustible liquids, solvents and gases. The quantity of flammable materials in the accelerator laboratory should be kept to a minimum. When large quantities of flammable gases are used, an emergency remote shutoff valve should be incorporated in each such piping system, if possible. All tubing and connections used in systems for flammable fluids should be examined periodically to maintain integrity during extended periods under high-temperature conditions such as might be encountered in a fire. In sputtering sources, cesium metal is used for the sputtering process. Cesium reacts violently with water, and owing to its reaction, it must be stored and handled in the absence of air. Cesium is normally delivered in glass ampoules of 1–5 g. The ampoule must be opened and the cesium transferred to the oven of the source in, for example, a glove box containing a dry protection gas. Only well-instructed and trained persons should handle cesium in connection with the charging and cleaning of a sputtering source. If possible, the experimental area should be isolated from the accelerator area by a fire- and blast-resistant wall to prevent or minimize damage to the other area. Emergency lighting should be provided. When flammable/explosive materials are handled in open reservoirs, the electrical installation in the laboratory must, in some countries, be safe against initiating explosions, for example not sparking when the light is switched on.

Mechanical hazards could be falling accidents because of unsafe conditions, which sometimes are tolerated in accelerator laboratories because of the temporary nature of, for example, experimental setups. Objects which are not bolted to the floor should be braced or arranged so that they can resist side forces without overturning. Mechanical forepumps are usually oil-sealed. The area around the pump tends to become slippery and can present a hazard. By careful maintenance of the pump, the oil leakage can be minimized, and by the mounting of the pump in a metal tray, the oil can be confined to areas close to the pump. The lower part of a diffusion pump becomes hot enough to cause burns. This part of the pump should therefore be protected. The pressure difference (0.1 MPa) between the inside of a vacuum enclosure and outside may initiate an accidental implosion of an evacuated brittle container, resulting in flying fragments. Glass and perspex viewports represent an implosion hazard. In the vicinity of powerful magnets, there can be a leakage of magnetic flux. Iron objects near a large operating magnet can cause serious injuries. High-pressure gases, either in cylinders (with pressures up to 20 MPa or even higher) or in the accelerator tank itself

(with pressures up to 1–1.5 MPa), are reservoirs of potential energy which, if released in an uncontrolled manner, can cause serious injury to personnel. The failure of a cylinder or its fittings may occur explosively, or may result in the cylinder becoming a high-velocity projectile from the jet reaction of exhausting gas. Cylinders must be stored upright and secured against overturning. They may not be exposed to heat and should never be dropped.

Toxic hazards of various kinds are often encountered in an accelerator laboratory. Ozone may be produced by corona discharge around high-voltage terminals. Ozone is a strongly oxidizing gas, attacking metal and rubber rapidly. It also has a serious effect on the respiratory system. Concentrations in excess of a few tenths of a part per million cause discomfort to exposed individuals in the form of headache, and dryness of the throat and of the mucous membranes of the nose and eyes. To reduce the hazard, ventilation should therefore be provided. Nontoxic but suffocating gases are extensively used in accelerator laboratories. Such gases may accumulate or be confined in an area, and those that are heavier than air will accumulate near the floor. The most common example of these gases is SF_6 , used in most modern accelerator tanks. Many metals and compounds have toxic properties. Some of the most serious hazards include beryllium, cadmium, mercury and lead. In the case of *beryllium* metal and its oxides, all handling of this material should be done under controlled conditions and confined to a ventilated hood or dry-box. *Cadmium* is another metal which should be handled with care and in a ventilated hood. *Mercury* was, earlier, commonly used in laboratories, for example in diffusion pumps, manometers, thermometers, relays, switches and batteries. Mercury slowly vaporizes at room temperature (vapor pressure at room temperature 0.160 Pa) and thus contaminates the air whenever it is exposed. The permitted level (in Sweden, $30\text{ }\mu\text{g}/\text{m}^3$, and in the USA, $100\text{ }\mu\text{g}/\text{m}^3$) will quickly be passed. Repeated inhalation of mercury vapor may result in severe cumulative poisoning. Liquid mercury can be absorbed through the skin. The great danger of mercury has forced several countries to take steps in the direction of completely eliminating its use. As an example, in Sweden, all use of mercury has been forbidden for some years, and all mercury available at industrial sites, laboratories and private homes is collected for safe disposal. Old laboratories in which mercury has been used are investigated by specially trained dogs to find and collect traces of mercury, for example in plumbing pipes. *Lead* is commonly used in accelerator laboratories for radiation shielding. The most usual manner in which lead absorption occurs is by inhalation of the fumes from molten lead. Therefore, when one is working with molten lead, adequate ventilation should be provided. However, handling of old lead blocks could also be dangerous, as the dust from lead oxide could be inhaled. To minimize this risk, lead blocks used for radiation shielding are often painted. When a toxic material is handled, this should be done in well-vented boxes connected to a safety ventilation system. The latter may, of course, not be connected to the ordinary ventilation system.

In conclusion, it is important to use good common sense when evaluating safety practices. A constant awareness of possible hazards is most important for all personnel in the laboratory, and especially for the person responsible for safety at the accelerator facility.

Box 9: Confined-Space Maintenance

G.A. Norton

National Electrostatics Corp., Middleton, WI 53562-0310, USA
nec@pelletron.com

A confined space is typically defined as a volume which is large enough to enter and has restricted means of entry and exit. It is not designed for continuous human occupation and has the potential, if ventilation is insufficient, for containing a hazardous atmosphere. In addition, in the United States, the Occupational Safety and Health Administration (OSHA) has an additional clause describing a confined space within an internal configuration having inwardly converging walls or a floor which slopes downward and tapers to a smaller cross section. The Swedish Work Environment Authority, in its definition of a confined space, also includes a requirement that it is a temporary workplace. All of this certainly describes the pressure vessels for electrostatic accelerators.

Some of the hazards found inside the accelerator tank are immediately obvious. In the case of the large vertical tandems, maintenance is often performed at extreme heights. This can also be true for the large horizontal tandems, where the distance from the column to the tank is sufficient that falling from the column level would cause significant injury. There are also the obvious hazards from moving mechanical parts such as the charging belt or charging chain, and rotating power shafts. Work inside the high-voltage terminal also involves working with high-current or high-voltage power supplies. The use of various types of volatile solutions for degreasing or general cleaning can produce hazardous vapor concentrations. However, the potentially greatest hazard is the least obvious.

In the early days of accelerator development, the insulating gas was high-pressure air. It was found that the increased concentration of oxygen led to the occurrence of fires during a spark. The column damage was often substantial. Therefore, the insulating gas was changed to a nitrogen/carbon dioxide mixture. While this mixture is still used, a large number of modern electrostatic accelerators are now using 100% sulfur hexafluoride (SF_6). In both cases, these gases are often benignly described as biologically inert. It is this property that makes these insulating gases very dangerous.

The human breathing reflex is triggered by the buildup of carbon dioxide in the bloodstream. Therefore, walking into an oxygen-free volume, breathing normally, unconsciousness often results in about a minute with no warning. If some oxygen is present, hypoxia results. The major early symptoms of

hypoxia include a feeling of euphoria combined with poor judgment or, more bluntly, stupidity.

Although this hazard can be avoided with adequate ventilation, extra care must be taken with sulfur hexafluoride. While the nitrogen/carbon dioxide mixture has about the same density as ambient air, sulfur hexafluoride is 5 times more dense and will tend to pool in low areas. There have been cases where even though the tank has two open hatches at each end with a fan in one of them, the SF₆ has remained undisturbed enough that personnel near the bottom of the tank have been in an SF₆/air mixture. This is immediately made apparent by the deepening of the person's voice. Just as breathing helium will cause a voice to increase in pitch, breathing SF₆ will cause the voice to lower in pitch. As soon as this is noticed, it is extremely important the personnel leave the area.

Many countries have government agencies that issue strict guidelines for working in confined spaces. In the United States, these rules are issued by the Occupational Safety and Health Administration, as mentioned above. Their rules focus on the proper preparation before entering a confined space, with adequate communication and ventilation while inside a confined space. These rules require adequate training of personnel with access to the confined space, as well as the availability of emergency equipment and crew. A brief look at rules and regulations for other countries indicates similar priorities. For example, the Swedish Work Environmental Authority has recommendations but no specific rules about work in confined spaces. These recommendations (AFS 1993:3) point out the importance of having good communication between employer and employee, so that the planning and realization of the work can be done in agreement.

The details for the preparation to enter a confined space vary from country to country, and from laboratory to laboratory depending upon the size and configuration of the accelerator tank. In general, before entry can be allowed into the pressure vessel, the gas transfer system must be secured or "locked out" to prevent insulating gas from reaching the pressure vessel while open. In addition, the charging system, power inside the tank and, in some cases, the complete control system are also locked out. This may change during maintenance as the testing of various column components is done. There may be other requirements, such as grounding the accelerator column to prevent charge buildup.

In all cases, it is necessary that the pressure vessel be thoroughly vented with ambient air. For many large machines, the venting process continues for 12 hours or more before personnel are allowed access. Then, before any entry, the oxygen level is monitored. In the case of the large, vertical tandems the oxygen monitor is lowered to the interior base of the tank.

At this point, some organizations require the safety department to issue a Confined Space Entry Permit. Oftentimes, this permit takes the form of a checklist to assure that all safety precautions have been taken and the

necessary safety equipment is present. Then, only trained staff are allowed entry to prepare the interior of the accelerator tank for maintenance. Oftentimes, this requires at least one trained person outside the tank in constant communication with trained personnel inside the tank. This preparation consists of assembling floorboards in the case of horizontal accelerators or securing the service platform in the case of vertical accelerators. Then the necessary safety equipment, such as communication devices and oxygen monitors, is installed in the tank. At this time, the necessary tools and lights are also brought in.

Once the interior of the accelerator tank is ready for accelerator maintenance, often the requirements are relaxed to where constant communication is no longer required. However, most organizations will require at least two people in close proximity in case an emergency develops. In some cases, there is a sign-in/sign-out procedure whenever entering or exiting the pressure vessel. In all cases, it is expected that only persons thoroughly familiar with the accelerator be allowed entry. All guests must be carefully escorted.

Many accelerator laboratories have other confined spaces, such as the insulating-gas storage tanks. Often, the procedures for working inside these tanks are more rigorous. In the case of the very large liquid storage vessels at the Holifield Radioactive Ion Beam Facility, there is always one person outside who is in constant communication with the person inside the storage vessel. In addition, the person working inside wears a safety harness to allow remote retrieval by the person outside unless the harness interferes with the safe performance of the job at hand.

In order to maintain a competent safety staff, many laboratories have annual refresher courses to train the safety personnel in the use of respirators and other safety equipment and procedures. This assures that all proper care is taken at each accelerator opening.

Box 10: Earthquake Protection – for Pelletrons

G.A. Norton

National Electrostatics Corp., Middleton, WI 53562-0310, USA
nec@pelletron.com

Pelletron accelerators are located throughout the world, including the seismically active area of the Pacific known as the “Ring of Fire”. Of the more than 30 Pelletron accelerators in Japan, only two are equipped with earthquake protection systems. These are the largest Pelletrons in the area, the 12 MV vertical tandem at the University of Tsukuba and the 20 MV vertical, folded tandem at the Japan Atomic Energy Research Institute (JAERI). Most of the Pelletrons in the Pacific Rim are smaller systems rated at 3 MV and below. While they have no earthquake protection systems, only one of these has seen significant damage.

On 17th of January 1995, a magnitude 7.2 earthquake struck the region of Kobe and Osaka in south-central Japan. The accelerator laboratory of the Kobe University of Mercantile Marine (KUMM) is located very close to the epicenter, which was just a few kilometers off shore. The KUMM accelerator was a 1.7 MV tandem Pelletron model 5SDH. It had a relatively short column structure, only about 3 m long. The earthquake caused the beamline bellows at each end of the pressure vessel to separate and the high-energy magnetic quadrupole lens and switching magnet to be thrown from their support stands. However, the 5SDH column structure showed no outward signs of damage.

In this case, the damage could have been substantially reduced if the pressure vessel support and external beamline supports were designed with earthquake considerations in mind. For example, this could include securely attaching support stands to the floor. However, the protection of the accelerator column itself is highly dependent upon the spectrum of the local expected seismic activity, the column design and overall column size. In the case of the KUMM accelerator, the accelerator column was relatively small and very rigid in both the horizontal and the vertical direction. That is not the case for the very large, vertical 12 MV conventional tandem and the vertical, folded 20 MV tandem, both also in Japan. Both of these column structures are about 18 m high. However, the 12 MV column is a conventional tandem, attached at the pressure vessel at each end. The 20 MV folded tandem is a single-ended, vertical-column structure.

There are two different types of earthquake protection in use on the large Pelletron accelerators, active and passive. The vertical tandems at the

University of Tsukuba in Japan and at the Nuclear Science Centre in New Delhi, India are protected by active coupling systems. The vertical, single-ended column structures of the tandem Pelletrons at the Oak Ridge National Laboratory (ORNL) and at JAERI are protected by passive isolation systems at the column base.

The active systems used on the 12 MV tandem and the 15 MV tandem consist of four pneumatic, radial ram assemblies located at the high-voltage terminal level on the pressure vessel. When an earthquake is detected, the rams rapidly (within a few seconds) move into the tank to prevent excess radial motion of the terminal. In this way, the vertical tandem column is supported at each end and the terminal is directly coupled to the tank and the building.

This system has proven successful during a “Miyagi” earthquake that affected the Tsukuba area in 1978. The magnitude was reported to be 7.4, with an epicenter about 100 km from the laboratory. The 90° inflection magnet, located about 40 m above ground with a weight of 3 tons, was displaced by 2 cm. All the shield doors, which are heavier than 2 tons, moved about 20 cm. However, the earthquake rams acted as designed and there was no damage to the tandem column structure. It must be noted that even without the earthquake rams, it is not clear that the column would have sustained any significant damage (communication from Professor Kohei Furuno).

Unlike the traditional, straight-through, vertical tandem configuration of the machines above, the folded tandems at JAERI and ORNL are vertical, single-ended, freestanding column structures with a massive 180° magnet located in the high-voltage terminal. These are tall, top-heavy freestanding structures of titanium and ceramic. Calculations have shown that the best method of protection is isolation from the pressure vessel and building. To accomplish this, the columns rest on large thrust bearings at the tank base. Approximately 2 m above the bearings, still at ground potential, the column is radially attached to the tank walls via a viscous damping system. In the case of the JAERI 20 MV tandem, the column is allowed to move ± 5 cm relative to the tank. In all cases, the acceleration tube is rigidly attached to the column structure. The tube and column beamlines are isolated from the rest of the beamline and vacuum system by bellows assemblies.

The earthquake protection systems at JAERI and ORNL have been in use since their original manufacture in the late 1970s. The New Delhi 15 MV tandem Pelletron went on-line in the mid 1980s and the Tsukuba 12 MV tandem went on-line in the mid 1970s. All have experienced minor to major earthquakes with no column damage.

Box 11: Earthquake Protection of the Bucharest FN Tandem Accelerator

S. Dobrescu and L. Marinescu**

National Institute for Physics and Nuclear Engineering, POB MG-6, Bucharest,
Romania
`sdob@ifin.nipne.ro`

The HVEC FN tandem accelerator, commissioned in Bucharest in March 1973, is operated at a terminal voltage of up to 8.5 MV and delivers a wide range of ion species from protons to gold ions, used for atomic- and nuclear-physics research as well as for applications (AMS, PIXE, ERDA and RBS analysis). The operation of the Bucharest tandem accelerator was interrupted for long periods of time owing to the disastrous damage caused to the column structure in the tank [1, 2] by major earthquakes in 1977 (magnitude 7.2 on the Richter scale) and 1986 (magnitude 6.9). After the second event and the subsequent second very expensive repair of the tandem column, a concept for earthquake protection of the accelerator was developed.

The concept of an earthquake protection system was based on a study aimed at providing a better understanding of earthquake effects upon the column structure as a result of the dynamic interaction between the seismic ground motion, the building, the accelerator tank and the column. In order to experimentally characterize the dynamic properties of this complex system, its response to the microseismic movements of the ground and to oscillations induced by 20 m deep underground TNT explosions close (150–215 m) to the building were recorded. The most critical component of the horizontal tandem accelerator is the column, since it presents a capability to resist accelerations of only $0.1g$ for transverse horizontal forces, while the peak acceleration during the 1977 earthquake was $0.4g$ ($g = 9.81 \text{ m/s}^2$). More than that, the column has a natural mechanical frequency of 2.1 Hz in the horizontal direction, coinciding with the dominant frequency at the maximum acceleration specific to the earthquakes produced in the Vrancea region of Romania, as well as with the eigenfrequency of the rocking motion of the building. The origin of the column damage resided therefore in the coincidence between the natural frequencies of the column, the building and the ground motion during the earthquakes, resulting in a mechanical resonance centered around 2 Hz. Several technical solutions for an earthquake protection system for the Bucharest tandem accelerator were taken into consideration:

- improvement of the radial stiffness of the column with a portico system, such as that used in the Vivitron tandem accelerator, or with four rods automatically pushed into the tank in order to rigidly support the terminal;

** Deceased.

- isolation of the tank from ground oscillations by a pendulum-type system or by efficient oscillation damping through a soft elastic suspension with springs and viscodampers.

The last solution, implying support of the accelerator tank (14 m long, 3.6 m diameter and 60 t weight) on spring and damper units, was preferred, being considered technically and economically the most adequate for the Bucharest tandem accelerator. The four accelerator tank legs have each been equipped with special elastic units, designed and made by GERB GmbH, Essen, Germany. Each unit consists of four soft springs and one viscous damper. When the accelerator tank is freely swinging on these elastic suspensions, its natural frequencies are lowered to 0.5 Hz in the horizontal direction and 0.9 Hz in the vertical direction. Except in the vicinity of the resonance, the force transmitted by an undamped isolator is practically proportional to the square of the ratio between its natural frequency and the excitation frequency. Therefore the elastic supports reduce the horizontal force induced by an earthquake in the column by a factor of over 10, sufficient to provide efficient column protection for earthquakes as strong as the one in 1977. The maximum allowed displacements of the tank are ± 180 mm horizontally and ± 50 mm vertically.

A special problem arises from the fact that the soft springs do not ensure the stability of the tank position required by the accelerator tubes and the alignment of the stripper with the general ion-optical axis of the tandem. In order to solve this problem, special devices, which might be called “mechanical fuses”, have been designed [3] and installed. Each tank leg is provided with four mechanical fuses, and there are two more at the two ends of the tank. The mechanical fuses continuously lock the tank in the desired position, dictated by the alignment requirements, and release the tank on to the elastic modules at the beginning of an earthquake, but only if the horizontal acceleration of the floor exceeds a preset value. The mechanical fuses also allow vertical adjustment of the tank position. The general layout of the system in Fig. B11.1, and a view of a spring and damper unit equipped with mechanical fuses is shown in Fig. B11.2.

The simultaneous release of all mechanical fuses is automatically performed by a specially designed triggering system consisting of an earthquake sensor and a nitrogen-pressurized network of tubes connected to the mechanical fuses. The mechanical fuses are locked as long as the pressure in the gas network is in the range $(8 - 10) \times 10^5$ Pa. If the pressure drops below a preset value, the mechanical fuses are unlocked and the accelerator tank is released on to the springs.

The earthquake sensor, which is completely free from conventional sources of energy that may fail during an earthquake, consists of a 76 mm diameter steel ball seated in an unstable balance at the top of a 3 m long vertical tube. When the ground acceleration exceeds $0.05g$, the ball falls down and activates a release system consisting of switches and a pressurized glass tube placed at

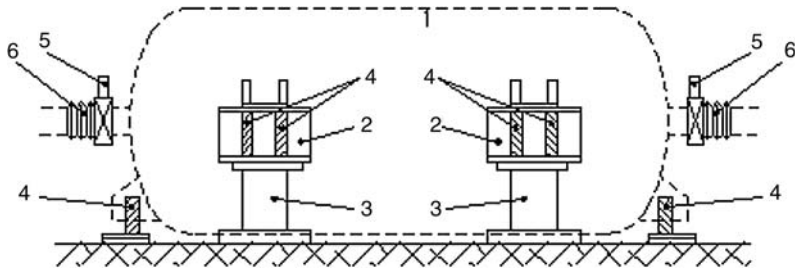


Fig. B11.1. General layout of the earthquake protection system: 1, tank; 2, spring and viscodamper units; 3, tank legs; 4, mechanical fuses; 5, protection gate valves; 6, fragile aluminum tubes



Fig. B11.2. View of an elastic unit with springs, damper and mechanical fuses

the bottom of the vertical tube. The switches, activated by the falling ball, switch off the accelerator belt motor and close the gate valves installed on both ends of the accelerator tubes in order to prevent damage to the tubes caused by breaking of the vacuum (these actions occur at $t = 0.5 - 1$ s after the moment $t = 0$ when the ball starts to fall). After that, the falling ball breaks the glass tube, the pressure drops, unlocking the mechanical fuses, and the tank is completely released on to the springs, slow vertical damped oscillations of the tank starting at $t = 2.5$ s. In order to protect the LE and HE beam components, fragile short aluminum tubes were inserted at both ends of the tank, between the beam components and the accelerator tubes. When the tank is released on to the springs, these two tubes are broken after the closing of the gate valves.

The system has been tested several times by forced excitation of the earthquake sensor and, so far, only once in real conditions (May 1990, earthquake of magnitude 6.4), when the system behaved properly [4, 5].

References

1. S. Dobrescu: Nucl. Instr. Meth. **184**, 103 (1981)
2. L. Marinescu, V. Zoran, S. Dobrescu, G. Pascovici, G. Semenescu, G. Marmureanu, H. Sandi, T. Sireteanu: Nucl. Instr. Meth. A **287**, 127 (1990)
3. Romanian Patent No. 102594 (1990)
4. L. Marinescu, G. Pascovici, V. Zoran, R. Dumitrescu, T. Sireteanu, E. Iordachescu, I. Filimon, A. Winkler: Nucl. Instr. Meth. A **328**, 90 (1993)
5. L. Marinescu, M. Petrascu, R. Dima, D. Serban: Nucl. Instr. Meth. A **328**, 82 (1996)

18 Electrostatic-Accelerator Free-Electron Lasers

A. Gover¹ and Y. Pinhasi²

¹ Tel Aviv University, School of Engineering, Physical Electronics Dept.,
Tel Aviv, Israel

`gover@eng.tau.ac.il`

² The College of Judea and Samaria, Dept. of Electrical and Electronic
Engineering, Ariel, Israel

`yosip@eng.tau.ac.il`

18.1 Introduction

Free-electron lasers (FELs) are high-power sources of electromagnetic radiation that utilize accelerated electrons which are oscillating transverse to their propagation axis. The name “free-electron laser” was chosen to distinguish this device from conventional quantum lasers (light amplification by stimulated emission of radiation), where the radiation is a consequence of transitions of bound electrons between atomic energy states.

The foundations of FELs go back to the early investigations of stimulated Thomson and Compton scattering carried out by Kapitza and Dirac in 1933 [1]. Motz, at Stanford, in 1951, studied theoretically [2] and experimentally [3] the radiation emitted when an accelerated electron beam passes through a succession of magnetic fields of alternating polarity. This “synchrotron undulator radiation” is a common source of X-ray and UV radiation in synchrotron facilities. It is, of course, a spontaneous-emission radiation source. The FEL is, in fact, the stimulated-emission version of synchrotron undulator radiation.

In 1957, Phillips of the General Electric Microwave Laboratory invented the Ubitron (undulating-beam interaction) [4, 5]. This mm-wavelength electron tube may be regarded as the first version of an FEL or, rather an FEM (free-electron maser). Both an amplifier and an oscillator version were built. However, the full potential of this device was not recognized, and the research work was terminated in 1964.

The interest in interaction between an undulating e-beam and electromagnetic field as related to generation of coherent radiation, arose again when Madey proposed in 1971 his idea [6] of the free-electron laser [7]. Two successive experiments were performed at Stanford in the infrared (IR) regime. First, an amplifier at $10.6\,\mu\text{m}$ was announced in 1976 [8], and laser oscillation at $3.4\,\mu\text{m}$ was obtained the following year [9]. These experiments utilized a relativistic electron beam from a linear accelerator (Linac).

18.2 Fundamentals of Free-Electron Laser Operation

In principle, an FEL consists of an accelerated electron beam traveling along a periodic undulator structure, which forces the electrons to oscillate in the transverse direction (see Fig. 18.1). The undulator is usually composed of alternating magnets, equally spaced, with a period λ_W .

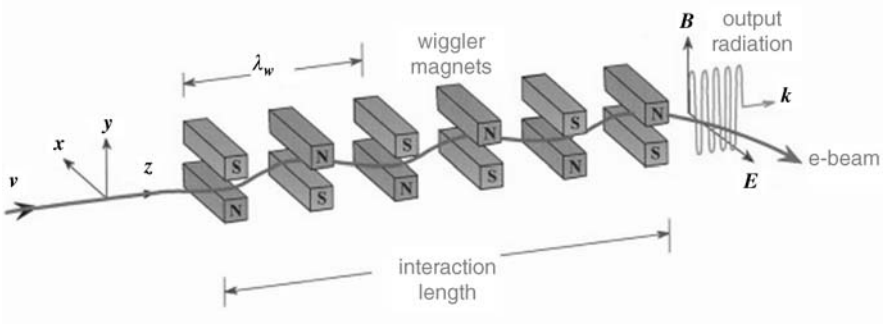


Fig. 18.1. Electron passing through a wiggler, emitting undulator synchrotron radiation

Each electron is a moving dipole radiator, emitting a Doppler-shifted wave packet of undulator synchrotron radiation [3], which propagates in free space or in a waveguide. In free space and in an overmoded waveguide, the wavelength of the radiation is given approximately by

$$\lambda_S = \frac{\lambda_W}{\beta_z(1 + \beta_z)\gamma_z^2} \quad (18.1)$$

where $\beta_z = v_z/c$ (v_z is the axial electron velocity and c is the speed of light), and $\gamma_z = 1/\sqrt{1 - \beta_z^2}$ is the axial relativistic Lorentz factor, related to the beam energy E_k and the wiggler parameter $a_W = eB_W/k_W mc$ through the relations

$$\gamma_z = \frac{\gamma}{\sqrt{1 + a_W^2/2}} \quad (18.2)$$

$$\gamma = 1 + \frac{E_k}{m_e c^2} \quad (18.3)$$

where $k_W = 2\pi/\lambda_W$, B_W is the magnetic-field amplitude of the periodic planar wiggler (undulator) and m_e is the electron mass. In the relativistic limit $\beta_z \rightarrow 1$, the radiation wavelength is approximated by

$$\lambda_S \approx \frac{\lambda_W}{2\gamma_z^2} = \frac{\lambda_W}{2\gamma^2} (1 + a_W^2/2) \quad (18.4)$$

If a waveguide is used inside the wiggler, (18.1) and (18.4) should be modified, and a longer wavelength is expected [10].

In stimulated-emission devices (an FEL amplifier or oscillator), a radiation wave enters the undulator collinearly with the e-beam. The beating of the transverse electromagnetic-field components of the radiation wave with the wiggler-induced periodic transverse velocity of the electrons creates a moving “ponderomotive force wave”, directed in the axial (z) direction. This propagates synchronously with the electrons and modulates their axial velocity, forming density bunching in the beam at the radiation frequency. If the ponderomotive wave is slightly slower than the electron velocity, the electrons lose energy to the wave. As a result, the electrons generate radiation by stimulated emission, and amplification occurs.

18.3 Electrostatic-Accelerator FELs

Free-electron lasers utilizing electrostatic accelerators (EAs) of a few MeV normally operate in the mm and IR wavelength regime. Figure 18.2 shows a curve of radiation frequency as a function of acceleration voltage for an EA-FEL, assuming free-space propagation (18.1) and a wiggler with $\lambda_W = 4.44$ cm (Table 18.1).

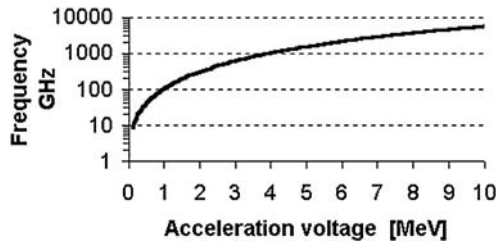


Fig. 18.2. Frequency vs. acceleration voltage for an EA-FEL

While most FEL facilities are based upon linacs, which produce ultra-short (ps range) electron-beam pulses, a few facilities in the world utilize EAs that enable continuous-wave (CW) or quasi-CW (long-pulse) operation. EA-FELs are also characterized by high average power generation, high energy-conversion efficiency and high spectral purity. The high quality (small emittance and low energy spread) of the electron beam in the electrostatic accelerator is crucial for attaining FEL operation at short wavelengths. The unique features of EA-FELs make them naturally fitted for a variety of applications in the present and in the near future [11].

The first demonstration of an EA-FEL, operating at the far-IR band, was made at the University of California Santa Barbara (UCSB) in 1984 [12–14]. The experiment employed a 6 MV Pelletron accelerator, and produced 10 kW over the range of 390–1000 μm . Subsequently, the FEL was upgraded to cover

Table 18.1. Parameters of an EA-FEL

Accelerator	
Electron beam energy	$E_k = 1\text{--}3\text{ MeV}$
Beam current	$I_0 = 1\text{--}2\text{ A}$
Undulator	
Type	Magnetostatic planar wiggler
Magnetic induction	$B_W = 0.2\text{ T}$
Period length	$\lambda_W = 4.444\text{ cm}$
Number of periods	$N_W = 20$
Resonator	
Waveguide	Curved parallel
Transverse mode	TE_{01}
Round-trip length	$L_C = 2.62\text{ m}$
Out-coupling coefficient	$T = 7\%$
Total round-trip reflectivity	$R = 65\%$

the entire wavelength range from 2.5 mm to 30 μm by adding two more wigglers of shorter period. Figure 18.3 shows a photograph of the UCSB FEL laboratory. The electron beam originates from a 2A Pierce-type electron gun located in the negatively charged high-voltage (HV) terminal inside the accelerator tank (top right part of Fig. 18.3). The high-current beam is transported down the acceleration tube into the beam-switchyard room (Fig. 18.4), where

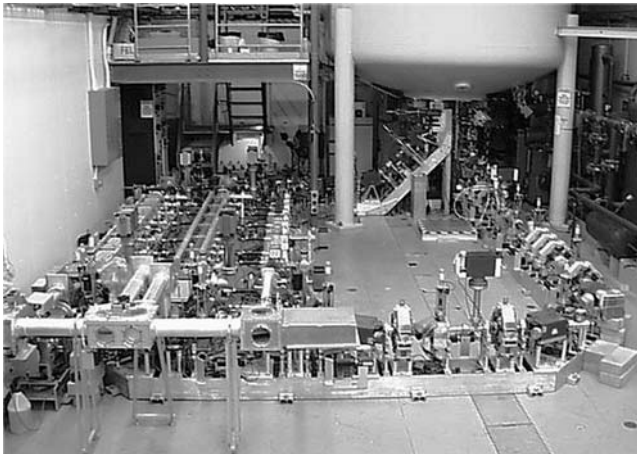


Fig. 18.3. The UCSB 6 MV FEL Laboratory (Reprinted from http://sbfel3.ucsb.edu/fel_lab.html, with permission from UCSB)

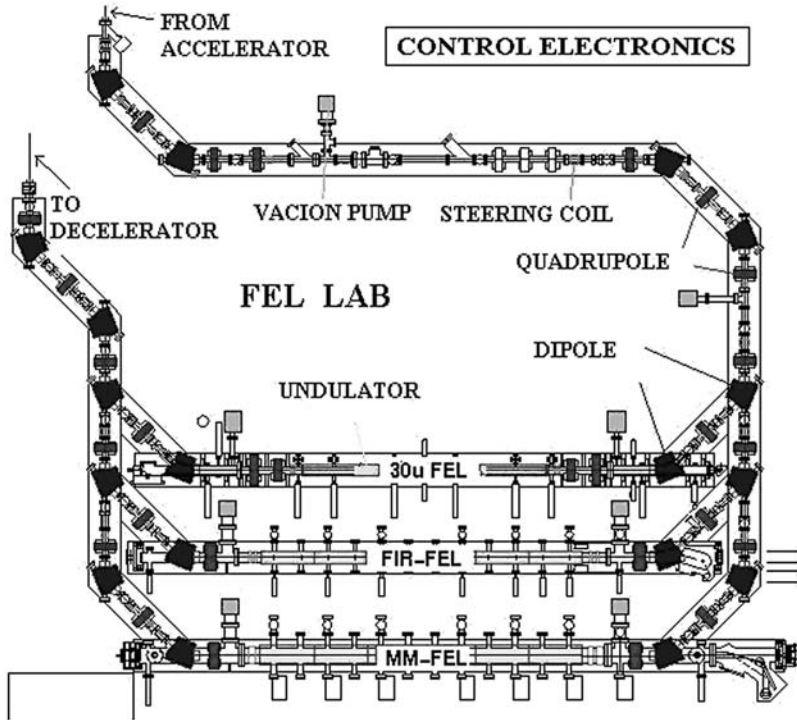


Fig. 18.4. The UCSB 6 MV beam switchyard (Reprinted from http://sbfel3.ucsb.edu/6mv/6mv_swtychd.html, with permission from UCSB)

it can be directed at choice to pass through any of the three wigglers (mm-wave, FIR and $30\mu\text{m}$). The used e-beam is then bent and directed back into the accelerator terminal by means of low-dispersion bending magnets and a vertical deceleration tube, which goes parallel to the acceleration tube inside the tank. The beam is then collected by a multistage collector, closing the high-current-beam electrical circuit at the HV terminal. The UCSB FEL has been operating for almost two decades as a Radiation User Center.

An alternative configuration for an EA-FEL was developed in Israel by a consortium headed by Tel Aviv University, based on the 6 MV EN tandem accelerator at the Weizmann Institute (shown in Fig. 18.5) [15, 16]. In this configuration, the wiggler (and resonator) is placed in the positively charged HV terminal, and the electron gun and collector are external to the tank at ground potential (see Fig. 18.6). The accelerator, originally an ion accelerator, was converted into an electron accelerator and its internal transport system was modified (two out of four acceleration tube sections were removed) to accommodate a magnetostatic wiggler and electron-optical focusing, steering and diagnostic elements (see Fig. 18.6). First lasing and high-coherence single-mode operation were demonstrated in 1997 [16, 17]. In 2001, the FEL

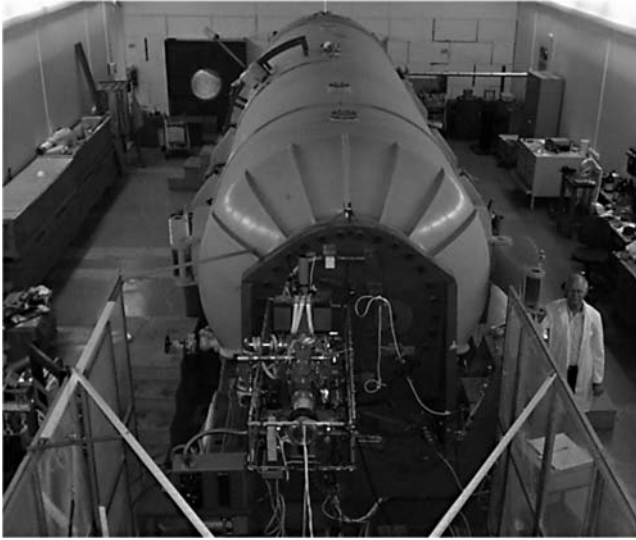


Fig. 18.5. The Israeli electrostatic-accelerator FEL

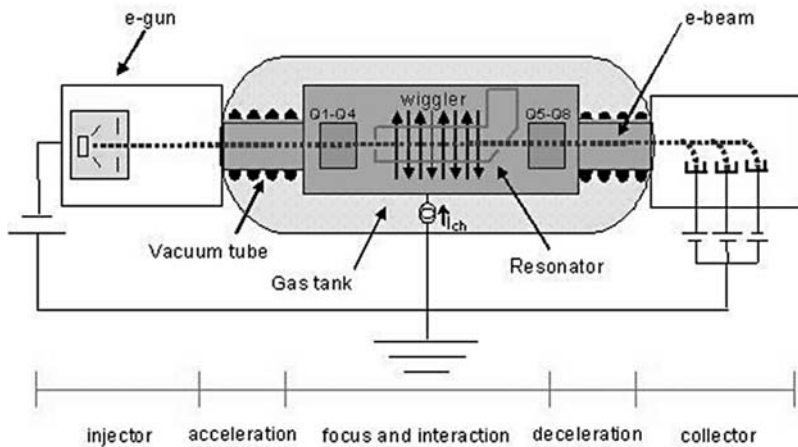


Fig. 18.6. The internal-cavity straight-transport-line EA-FEL scheme

was relocated to Ariel to a dedicated radiation-users' building, and has returned just recently to operation in a new configuration. In the present FEL, the mm-wave radiation, generated in the resonator, is separated from the electron beam by means of a perforated Talbot-effect bent reflector [18]. A quasi-optical delivery system, composed of overmoded corrugated waveguides and dielectric lenses, transmits the out-coupled power through a window in the pressurized-gas accelerator tank and into the radiation-users' room,

across a 1.3 m radiation protection concrete wall (see the horizontal aluminum waveguide tube at the top right corner of Fig. 18.5). The Israeli FEL was built to serve as a radiation-user facility. In recent experiments [19], 0.5 kW RF power at 100 GHz was delivered to the users' rooms. It is now being further developed to operate at high average power (1 kW) in the range of 70–130 GHz.

The capability of the EA-FEL to produce almost continuous high average power at mm wavelengths with wide-range tunability (by varying the beam energy) has aroused interest in applying it for electron cyclotron resonance heating (ECRH) of magnetically confined plasmas for controlled thermonuclear fusion. Such an FEM was suggested and constructed by the Dutch FOM Institute for Plasma Physics [20,21], and attained a record high power of 730 kW in a few ms at 206 GHz, and pulses of 80 kW tens of ms long in the 200–300 GHz range. Single-mode operation and real-time, wide-range step frequency tuning, which is needed for the fusion application, were also demonstrated. The accelerator used in this case was charged by a 2 MV 20 mA insulated-core transformer HV power supply.

Two other noteworthy experiments have been carried out in Korea and at CREOL (University of Central Florida). In 2001, the Korean FEM demonstrated first lasing in a straight-line-transport, positively-charged-terminal configuration. No accelerator was used; rather, an open-air-insulation transport system, based on a 500 kV power supply and an e-gun, were used to demonstrate lasing at 35 GHz at an RF power of 1 kW [22]. At CREOL, a compact FEL based on a mini-wiggler ($\lambda_W = 8$ mm) and a Pelletron accelerator has been developed [23]. In 2000, lasing in the FIR region ($\lambda = 400 \mu\text{m}$, $P = 100$ W) was demonstrated with it for the first time. This FEL was transported recently to the University of Hawaii and is being reassembled there.

18.4 EA-FEL Design Considerations

The EA-FEL is inherently an energy retrieval (current recirculation) device. This means that the entire high current (I_b) of the accelerated electron beam (in the order of amperes), except for a small interception current I_{int} , must be transported after passing through the wiggler (where it transfers part of its kinetic energy to the radiation field), and decelerated down to the moderate (tens of kV) voltages of the collector high-current power supplies. Through the collector power supplies, the e-beam electric circuit closes back to the electron gun. Because the electron-beam energy distribution spreads out during the interaction with the radiation wave inside the wiggler, it is desirable to collect the decelerated electrons with a multiple-electrode collector. This means that each electron-beam energy class hits the corresponding collector electrode with minimal kinetic energy, and consequently minimal heat waste and secondary-electron emission. This technology of a “multistage depressed collector” was developed before in the field of microwave tubes.

The beam current recirculation scheme makes the EA-FEL a potentially high-efficiency, high-average-power radiation source. However, its primary benefit is making it possible to transport a high beam current (a current of the order of amperes is needed to surpass the FEL lasing threshold) in an electrostatic accelerator that can supply only a very low charging current I_{ch} (of the order of mA in electrostatic accelerators and tens of mA in high-voltage electrodynamic accelerators) to the HV terminal. The current recirculation scheme makes it possible to maintain the passage of a high-current beam through the wiggler, because this current is not supplied through the terminal-charging circuit. However, in order to maintain the steady-state voltage of the terminal, the transport efficiency from the e-gun to the collector $\eta_{trans} = 1 - I_{int}/I_b$ (I_{int} is the e-beam current intercepted along the acceleration-deceleration transport line) must be high enough to keep the current balance

$$I_{ch} = I_L + I_{int} \quad (18.5)$$

where I_L is the constant leakage current (the corona discharge current and the current through the series bleed resistors). Typically, a transport efficiency in excess of $\eta_{trans} = 99.9\%$ ($I_{int}/I_b < 10^{-3}$) is required to maintain condition (18.5) for steady-state lasing (CW) operation.

Attaining high transport efficiency is difficult, and therefore all present EA-FELs are operated in a (long-) pulse operating mode. In this mode, the pulsed e-beam current intercepted $I_{int} = (1 - \eta_{trans})I_b$ exceeds the net terminal charging current $I_{ch} - I_L$, and therefore the terminal voltage drops during the pulse duration τ_p according to

$$\Delta V_T = I_{int}\tau_p/C_T \quad (18.6)$$

where C_T is the capacitance of the HV terminal relative to the tank enclosure. The terminal voltage drop causes a decrease in the e-beam kinetic energy $\Delta E_k = mc^2 \Delta\gamma = e \Delta V_T$, which downshifts the synchronism wavelength (18.4). If the e-beam pulse is long enough to enable oscillation buildup in the FEL oscillator up to saturation, and consequently establishment of a monochromatic stored radiation field in the resonator, then as the terminal voltage continues to drop, the stored radiation field of the oscillator gets out of synchronism with the slowing-down beam, and the FEL gain drops. Because of the finite bandwidth of the FEL gain curve, only a relative energy drop of

$$\frac{\Delta\gamma}{\gamma} = \frac{1}{2N_W} \quad (18.7)$$

(N_W is the number of wiggler periods) shifts the gain curve out of the frequency of the stored radiation, and causes the FEL to stop lasing and decay. Thus, consideration of (18.6) and (18.7), using (18.3), determines the maximum duration of a single-frequency lasing pulse τ_p permissible. Typical pulse durations in existing EA-FEL are less than ms.

The possibility of obtaining long-pulse operation (and potentially CW operation) demonstrates an important feature of the EA-FEL – high temporal coherence. Because the FEL is essentially a homogeneously broadened laser, the mode competition process in the resonator during the oscillation buildup period winds up with single-mode operation [14, 17, 24]. Figure 18.7 shows the spectrum of the single-mode radiation measured for the Israeli FEL [17]. It shows an FWHM emission bandwidth of $\Delta f = 1$ MHz that corresponds to a relative linewidth of $\Delta f/f \cong 10^{-5}$, still Fourier-transform-limited by the finite pulse duration ($\tau_p = 20 \mu\text{s}$).

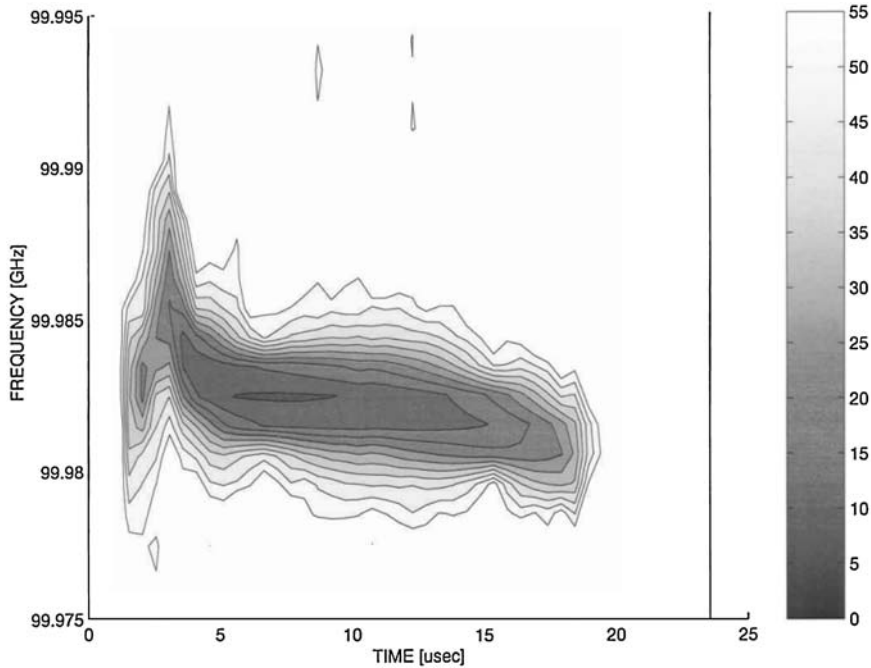


Fig. 18.7. Spectrum of the coherent single-mode radiation emitted from the tandem EA-FEL [17]. The FWHM bandwidth is $\Delta f = 1$ MHz (the intensity scale is logarithmic)

The saturation radiative extraction power for the FEL is given by

$$\Delta P_{ext} = \eta_{ext} E_k I_b / e \quad (18.8)$$

where the extraction efficiency $\eta_{extr} \cong 1/2N_W$ can be understood as a result of the electrons getting out of synchronism owing to their energy loss in the resonator (compare (18.7)). The power coupled out of the resonator at saturation depends on the power transmission T of the out-coupling mirror and the total round-trip loss factor $1 - R$ (R is the resonator round-trip

reflectivity):

$$P_{out} = \frac{T}{1-R} \Delta P_{ext} \quad (18.9)$$

As in any laser, there is always an optimal transmission factor T , for which P_{out} is maximized. Its determination requires numerical solution of the FEL equations in the nonlinear (saturation) regime [25].

The design of the resonator is a difficult part of EA-FEL design. Because of the long wavelength of operation, diffraction is significant, and usually an over-moded waveguide resonator is used. If the electron beam can be bent into and out of the resonator axis (as in [13,23]) then slit out-coupling can be used in the output mirror. In [17,19,21] Talbot-effect reflector configurations [18] were used, enabling out-coupling of the radiation to the side of the resonator axis. The e-beam is injected, and then enters and leaves the resonator through holes in the mirrors, keeping a straight trajectory along the coaxial accelerator and resonator.

An example of EA-FEL construction, some parameters based on the Israeli EA-FEL are given in Table 18.1. For these parameters, an extraction efficiency $\eta_{extr} = 2.5\%$ and an output power $P_{out}/I_b = 7 \text{ kW/A}$ are expected at saturation, according to (18.8) and (18.9). So far, $P_{out} = 5 \text{ kW}$ (0.5 kW at the users' room) has been measured in this device at $I_b = 1.8 \text{ A}$, which is before saturation.

In conclusion in relation to the design considerations, an important note should be made concerning the choice of accelerator polarity and configuration. A configuration with an external resonator and wiggler is advantageous for application in a radiation-users' facility, mostly because it leaves room for the design of extended and multiple wigglers and resonators, and the radiation can be coupled easily out of the resonator directly into the user rooms (see the horizontal aluminum pipe in the bottom left corner of Fig. 18.3). However, for future specific short-wavelength and high-power applications, an internal-cavity configuration is advantageous. Because of voltage breakdown considerations, it is possible to build high-voltage electrostatic accelerators (which are needed for short-wavelength operation) only with a positive-polarity terminal. Furthermore, in this configuration, the e-gun and collector power supplies are placed at ground potential. At the high average power levels (of the order of MW) at which EA-FELs can potentially operate, it would be inconceivable to place the large power supplies in the terminal and transfer the grid power to them mechanically, as is done in the negative-polarity terminal configuration. In addition, the simple straight geometry of the electron beam transport in the internal-cavity configuration (Fig. 18.6) promises better transport efficiency and consequently higher-average-power operation.

User Center For FEL Radiation

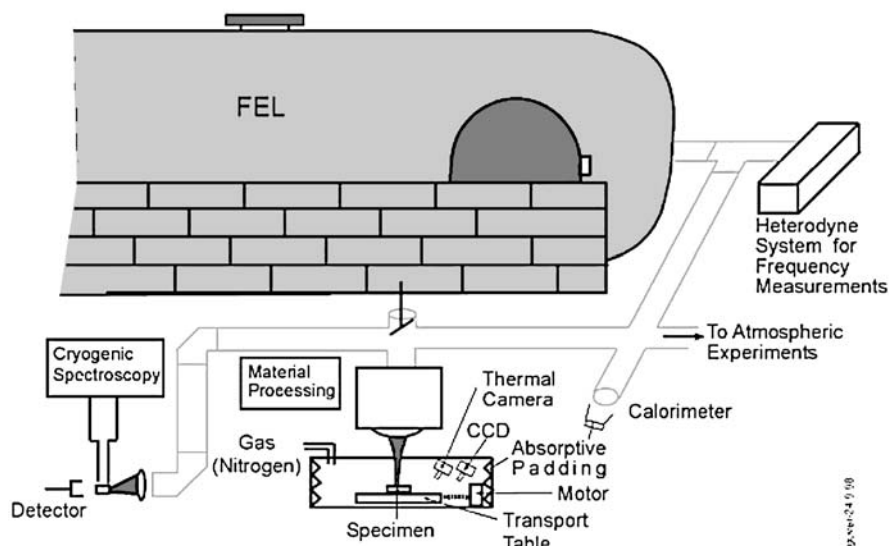


Fig. 18.8. Conceptual scheme of a radiation user facility based on an EA-FEL

18.5 EA-FEL Applications

The main application of the EA-FEL at present (as with most other kinds of FELs) is as a radiation user facility for scientific studies. In this connection, the main feature taken advantage of is the wide-range tunability in a wavelength region (from mm-wave to IR) which is deficient in tunable sources. For some studies, the high power and narrow spectrum of the radiation source are also important characteristics. Figure 18.8 depicts a conceptual design of some radiation exposure stations in a radiation user facility based on an EA-FEL [26].

The EA-FEL, as an intense tunable FIR source, is a useful study tool in condensed-matter physics research. It has been used in linear and non-linear spectroscopy of elementary excitations in materials: phonons, magnons, excitons and shallow impurity levels in semiconductors, small-energy-gap materials such as high-temperature superconductors (HTSCs), quantum wells, superlattices, and various mesoscopic semiconductor structures and devices [27]. In biological research, such a source can be useful in spectroscopic investigations of vibrational levels in proteins and DNA molecules. It can be used also for studies of damaging or therapeutic effects of mm-wave or FIR radiation interacting with biological tissues.

The potential of the EA-FEL for high average power and high efficiency holds promise for the development of energy-related industrial and

commercial applications. The MW-level design and the partial demonstration of operation of the FOM FEL indicate that such high-power applications are feasible with further investment in technological development. The fusion application (ECRH heating of magnetically confined plasmas in tokamaks) requires a few MW of power at short mm wavelengths, step-tunable by 35% on a timescale of seconds. More moderate average power levels (kW to tens of kW) are required for applications in thermal material processing (ceramic sintering, surface treatments, cutting, curing, drying etc.) [28]. Interesting applications and schematic designs have been proposed for compact EA-FEL systems to be used in free-space atmospheric beam propagation. These include the concept of radiative energy transmission to atmospheric or ionospheric unmanned airborne vehicles through atmospheric transmission windows (35 GHz and 94 GHz) [26, 29]. Other such applications include extremely high frequency (30–300 GHz) mm-wave radars and high-resolution imaging systems, communication (to satellite) systems, and remote sensing of gases and aerosols in the atmosphere [26].

Acknowledgments

This work was performed in the Israeli FEL National Knowledge Center supported by the Ministry of Science, Israel, and was supported in part by the Ministry of Infrastructure.

References

1. P. L. Kapitza, P. A. M. Dirac: Proc. Camb. Phil. Soc. **29**, 247 (1933)
2. H. Motz: J. Appl. Phys. **2**, 527 (1951)
3. H. Motz, W. Thon, R. N. Whitehurst: J. Appl. Phys. **24**, 826 (1953)
4. R. M. Phillips: IRE Trans. Electron. Dev. **ED-7**, 231 (1960)
5. R. M. Phillips: Nucl. Instr. Meth. A **272**, 272 (1988)
6. J. M. J. Madey: J. Appl. Phys. **42**, 1906 (1971)
7. J. M. J. Madey, H. A. Schwettman, W. M. Fairbank: IEEE Trans. Nucl. Sci. **NS-20**, 980 (1973)
8. L. R. Elias, W. M. Fairbank, J. M. J. Madey, H. A. Schwettman, T. L. Smith: Phys. Rev. Lett. **36**, 717 (1976)
9. D. A. G. Deacon, L. R. Elias, J. M. J. Madey, G. J. Ramian, H. A. Schwettman, T. L. Smith: Phys. Rev. Lett. **38**, 892 (1977)
10. E. Jerby, A. Gover: J. Quantum Electron. **QE-21**, 1041 (1984)
11. A. Gover, A. Friedman, A. Drobot: Laser Focus World, pp. 95–104 (October 1990)
12. L. R. Elias et al.: Nucl. Instr. Meth. A **237**, 203 (1985)
13. L. R. Elias: IEEE J. Quantum Electron. **QE-23**, 1470 (1987)
14. L. R. Elias, G. Ramian, R. J. Hu, A. Amir: Phys. Rev. Lett. **57**, 424 (1986)

15. A. Gover, E. Jerby, H. Kleinman, I. Ben-Zvi, B. V. Elkonin, A. Fruchtman, J. S. Sokolowski, B. Mandelbaum, A. Rosenberg, J. Shiloh, G. Hazak, O. Shahal: Nucl. Instr. Meth. A **296**, 720 (1990)
16. A. Abramovich et al.: Appl. Phys. Lett. **71**, 3776 (1997)
17. A. Abramovich, M. Canter, A. Gover, J. S. Sokolowski, Y. M. Yakover, Y. Pinhasi, I. Schnitzer, J. Shiloh: Phys. Rev. Lett. **82**, 6774 (1999)
18. I. M. Yakover, Y. Pinhasi, A. Gover: Nucl. Instr. Meth. A **358**, 323 (1995)
19. A. Gover, A. Faingersh, A. Eliran, M. Volshonok, H. Kleinman, S. Wolowelsky, B. Kapilevich, Y. Lesser, Z. Seidov, M. Kanter, A. Zinigrad, M. Einat, Y. Lurie, A. Abramovich, A. Yahalom, Y. Pinhasi, E. Weisman, J. Shiloh: "Radiation Measurements in the New Tandem Accelerator FEL", Nucl. Instr. Meth. Phys. Res. A **528**, 23 (2004)
20. P. W. van Amersfoort, W. H. Urbanus, A. G. A. Verhoven, A. Verheul, A. B. Sterk, A. M. van Ingen, M. J. van der Wiel: Nucl. Instr. Meth. A **304**, 168 (1991)
21. W. H. Urbanus et al.: Phys. Rev. Lett. **89**, 214801-1 (2002)
22. B. C. Lee, S. K. Kim, Y. U. Jeong, S. O. Cho, B. H. Cha, J. Lee: Nucl. Instr. Meth. **375**, 28 (1996)
23. L. R. Elias, I. Kimel, D. Larson, D. Anderson, M. Tecimer, Z. Zhesu: Nucl. Instr. Meth. A **304**, 219 (1991)
24. B. G. Danly et al.: Phys. Rev. Lett. **65**, 2251 (1990)
25. A. Abramovich, Y. Pinhasi, A. Yahalom, D. Bar-Lev, S. Efimov, A. Gover: Nucl. Instr. Meth. Phys. Res. A **475**, 579 (2001)
26. Y. Pinhasi, I. M. Yakover, A. Eichenbaum, A. Gover: IEEE Trans. Plasma Sci. **24**, 1050 (1996)
27. J. Allen et al.: Nucl. Instr. Meth. Phys. Res. A **358**, 536 (1995)
28. A. Gover, A. L. Eichenbaum, S. Efimov, J. Sokolowski, M. Tecimer, Y. Yakover, A. Abramovich, M. Canter, Y. Pinhasi, A. Yahalom, I. Schnitzer, Y. Shiloh: "User Application for the TAU Tandem Electrostatic Accelerator FEL at Long Wavelengths", FEL User Workshop at 22nd Int. FEL Conf., Durham, N.C., Aug. 2000
29. I. Kimel, L. R. Elias: Nucl. Instr. Meth. A **358**, 186 (1995)

Part III

Research Fields and Their Technical Requirements

Introduction to Part III – Research Fields and Their Technical Requirements

R. Hellborg¹ and J. McKay²

¹ Department of Physics, Lund University, Sölvegatan 14, 223 62 Lund, Sweden
`ragnar.hellborg@nuclear.lu.se`

² Box 463, Deep River, Ontario, Canada K0J 1P0
`mckay@magma.ca`

It is an ambitious project to summarize all the fields of research which use electrostatic accelerators. The accelerator, as a research instrument, was a product of nuclear physics. Now, only a small fraction of machines are concerned with nuclear physics or basic research. The imagination with which people have applied the techniques of nuclear physics has led to uses never imagined by the originators of electrostatic machines. The topics covered in this Part describe the range of research areas utilizing such accelerators but inevitably omit many ingenious and important applications.

Basic research in nuclear and atomic physics continues to push the limits of accelerator technology. The search for isotopes near the proton and neutron drip lines calls for beams of exotic ions, including radioactive ions. Complex and sensitive detector arrays are needed to observe the products of the resulting reactions. Atomic physicists continue the examination of highly excited states, rare electron configurations, and the interaction of beams and matter in various states. Such studies refine our basic understanding of matter.

While electrostatic accelerators are often thought to be primarily instruments for research, today the vast majority of machines are used in applications. The great virtues of ion beams produced by electrostatic machines are their precision and versatility. This leads to great utility in the analysis or modification of materials. Atomic and nuclear interactions allow atom-by-atom measurements of or interactions with the target. Implantation of ions can be carefully controlled or measured by depth and location.

Materials analysis examines the distribution and density of introduced, contaminant or trace elements in many materials. These materials can be substrates for solid-state devices, biological or geological samples, or archeological or other materials. Accelerator mass spectrometry in particular has pushed the limits of sensitivity. Microprobe beams are used to explore distributions. Accelerators are also being employed in systems that survey bulk samples (such as cargo) for explosives or other contraband. Alteration of materials by precision implantation or bulk modification by deposition of energy or ions can be done in some applications, in spite of the current limitations inherent in electrostatic machines. Even isotope production is practical in some instances, where controlled reactions can be used to produce proton-rich or very pure products.

It is interesting to note that while most applications machines are of relatively low voltage, the latest generation of tandem accelerators for applications are rated at 5 MV. This is the same voltage as the first research tandem built almost fifty years ago. The new machines feature reliability and stability unimaginable in the early days. The larger tandems still in operation are ideal machines to serve as prototypes to examine and test new techniques and applications. The future may well see an increasing range of energies and ion species from new generations of modern electrostatic accelerators.

19 Isotope Production for Medical Applications

A.D. Roberts^{1,2,3}, T.E. Barnhart¹, and R.J. Nickles¹

¹ University of Wisconsin, Department of Medical Physics, Madison, WI, USA

² University of Wisconsin, Department of Psychiatry, Madison, WI, USA

³ University of Manchester Molecular Imaging Centre, Christie Hospital, UK

`arobert5@facstaff.wisc.edu`

`tebarnhart@wisc.edu`

`rnickles@facstaff.wisc.edu`

19.1 Introduction

The use of radiation for medical imaging and therapy has a long history, originating almost immediately in the earliest days of the X-rays discovered by W.C. Röntgen in 1895, and the subsequent discovery of natural radioactivity by Becquerel in 1896 and the separate isolation of radium by Pierre and Marie Curie in 1898. The use of radioactive materials was limited to natural radioisotopes until the demonstration of induced nuclear transformations using an accelerated beam by Cockcroft and Walton [1]. Despite the relatively widespread use of accelerators in the following decade for radioisotope production, these early machines were quite limited in the amount of useful radioactive material they could produce.

The field quickly evolved with the advent of nuclear reactors and improved charged-particle accelerators as part of the weapons programs in the Second World War. By far the most prolific source of man-made radioisotopes was from the energetic neutrons from nuclear reactors. While medical applications of these isotopes quickly evolved, the scarcity of reactor sources and the nature of neutron-induced reactions generally limited the field to the use of long-lived, neutron-rich nuclei. These isotopes and subsequent procedures are by far the most widely used in nuclear medicine today. There are numerous biological applications using reactor-produced radionuclides such as ^{14}C and ^{32}P . Medical applications include in vivo photon imaging (SPECT, single-photon-emission computed tomography) of source compounds labeled with ^{125}I , ^{99}Tc and many other isotopes. Other reactor-produced isotopes have found uses in radiotherapy, taking advantage of the local cell-killing capabilities of heavy-particle disintegration, particularly α -particles.

While the application of reactor-produced isotopes for medical purposes continues to be important, it was quickly realized that there are advantages to using accelerator-produced isotopes. First, the flexibility of using beams of well-defined energy allows for controlled selection of nuclear reactions. Next, the use of positively charged beams, typically of protons, deuterons or helium nuclei, allows the selective production of proton-rich radioisotopes, greatly

expanding the choice of isotopes available for medical use. More importantly, for imaging applications, radioisotopes can be produced in this manner with a change in proton number, facilitating easier chemical separation of the reaction products from the starting material. This can be crucial in the synthesis of compounds suitable for medical use, allowing complex radiochemistry to be performed, and efficient removal of the often toxic bulk material.

A further need for charged-particle-produced radioisotopes is in the imaging application of PET (positron emission tomography). Isotopes that decay by positron emission can be located using pairs of coincidence scintillators tuned to the characteristic 511 keV gamma rays emitted at near 180° upon positron-electron annihilation. By using rings of detector pairs, three-dimensional distributions of PET radioisotopes can be quantitatively imaged. In addition to the true three-dimensional distribution quantitation, there are significant advantages of using PET radioisotopes in medical studies, including the availability of chemical compounds labeled with naturally occurring light radioisotopes such as ^{15}O , ^{13}N and ^{11}C , and lifetimes well matched to the imaging requirements of physiological pathways of interest, seconds to hours rather than years.

While much of the early development work on isotope production originated out of nuclear-physics laboratories using electrostatic accelerators, most recent installations use commercially available small cyclotrons. The main beam requirements are energy and current, and the precise energy definition required for nuclear-physics experiments is less important. Nonetheless, isotope production with tandems still has a role in the field. The accelerator technology is well established, robust and comparatively inexpensive, and in some cases may be a matter of necessity, for example through collaboration with established physics or engineering tandem facilities. Furthermore, few commercial production cyclotrons allow much tuning of the beam shape and energy, hampering some basic development work on isotope production. Finally, most commercial cyclotrons require costly technological additions to provide different particle beams (protons vs. deuterons, typically), a feature readily available with the electrostatic tandem.

19.2 Historical Perspective

Since its initial development in the 1930s [2], the electrostatic accelerator, insulated under high pressure, has dominated the field of nuclear investigations. This was the direct result of the exquisite control of such beam variables as:

- energy, with an achievable resolution of $\delta E/E \approx 10^{-6}$ [3]
- geometry, with microbeams of μm dimensions [4]
- polarization [5, 6]
- heavy-ion capability, with the tandem today acting as the second accelerator in radioactive-ion-beam experiments [7].

Yet curiously, the electrostatic accelerator has seldom played a major role in radionuclide production for radiochemical or biological applications. This is due to a mismatch between the technical finesse of the machine and the brute needs of high currents on thick targets, analogous to pulling a plow with a thoroughbred. Cyclotrons arose as the more logical and reliable source of proton beams of 50 μA , 10–30 MeV, and have now evolved into near turnkey “black boxes” that almost disappear into the infrastructure that underlies a modern medical facility. This evolutionary path was shaped by the perceived beam current limitations of the tandem, a limitation that has been removed by today’s multicusp negative-ion sources [8].

In the past, occasionally situations arose that exploited the tandem’s advantages. A number of important radionuclides can be best produced by the irradiation of extremely costly, isotopically enriched targets. Examples of this include $^{78}\text{Kr}(\text{p}, \alpha)^{75}\text{Br}$ [9], $^{80}\text{Kr}(\text{d}, \text{n})^{81}\text{Br}$ [10] and today’s interest in making ^{124}I from enriched ^{124}Te . In cases where target material costs can reach many tens of US dollars/mg, a thick target ($\approx 100 \text{ mg/cm}^2$) demands shrinking the diameter of both the beam and the target to the order of mm. In this application, the tandem excels with its tight beam emittance matched to miniature-target configurations [11].

A second application flexing the electrostatic tandem accelerator was an unsuccessful effort to make useful activities of radionuclides farther removed from beta-stability through heavy-ion reactions. As an instructive example, the $^{40}\text{Ca}(^{16}\text{O}, \alpha)^{52}\text{Fe}$ reaction invites the production of an important biological tracer without recourse to an isotopically enriched target or beam stock. In fact, the measured production rate was of the order of hundreds of Bq/($\mu\text{A h}$) (RJN, unpublished data), disappointing even though it was performed parasitically in the Faraday cup of a “dues-paying” reaction study upstream. Heavy-ion reactions are poorly suited for the production of radiotracers simply because of the minuscule utilization of the beam. This utilization reaches nearly one ^{18}F nucleus created per thousand protons for the (p, n) reaction on ^{18}O at 11 MeV, but is drastically reduced by the Z^2 dependence of the stopping power of the ^{16}O beam, and the coupled problems of making beams of tens of μA of negative ions and stripping them to high charge states. Finally, the large number of competing exit channels saps the total reaction cross section above the Coulomb barrier, routing most of the flux into few-particle transfer reactions, resulting in products that could have been reached much more simply by a light-ion approach.

What has changed the landscape, in the seven decades since its birth, is the refinement of the tandem electrostatic accelerator into a reliable, high-current resource that rivals today’s commercial cyclotron in almost every respect, save for compactness and familiarity within the radiochemistry community. The former limitation is site-specific, the latter amenable to education.

19.3 Isotope Production Physics

Regardless of the accelerator used, the first key to determining its applicability to production of a given radioisotope is an evaluation of the total reaction yield. Most common PET radioisotopes have had their yields measured over a variety of energies, but it can be useful to revisit these data, particularly in poorly documented cases or where novel radioisotopes are proposed. These calculations are particularly important when evaluating the use of linear accelerators, which most commonly have lower energies than comparably sized cyclotrons available today.

If the total reaction cross section is known, as is often the case for most isotopes used in medicine, the total reaction yield for a thin target can be calculated in the usual way. The nomenclature for the yield Y at a given energy is typically given in MBq/ μ A at saturation bombardment for short-lived isotopes (less than a few hours), and MBq /(μ A h) for longer-lived isotopes.

Thin targets are rarely used in isotope production, particularly with lower-energy beams. The reaction of interest is typically the first channel available, for example (p, n) or (d, n), so to maximize yield the target material is thick enough to completely stop the beam. The details of the target thickness are sometimes difficult to determine, particularly in cases of gas targets that may see localized density reduction depending on the beam current and energy deposition. However, provided the target is truly thick and stops the beam at some unspecified point, the yield can be calculated using the well-known linear stopping power of the target material dE/dx . Published stopping-power data are quite sufficient for these calculations [12]. Equation (19.1) gives the calculated differential yield for a thick target. Integration over the appropriate energy range gives the total yield. The energy loss (dE/dx), the energy-dependent cross section (σ), the target density used for the energy loss (ρ) and the atomic weight (AW) are required for each calculation. The conversion factors give the reaction rate in MBq and the beam current in μ A.

$$\frac{dY}{dE} = \left(\frac{dx}{dE} \right) \sigma \rho \left(\frac{6.02 \times 10^{23}}{AW} \right) \left(\frac{\text{MBq}}{10^{-6} \text{ s}^{-1}} \right) \left(\frac{1}{1.6 \times 10^{-19} \mu\text{A}} \right) \left(\frac{1}{10^6 \text{ s}} \right) \quad (19.1)$$

As an example, several “standard” PET radioisotope yields have been calculated up to 10 MeV, as shown in Fig. 19.1. This is by no means a complete listing, but can serve in determining the applicability of a given accelerator to this task. Further discussion of these target systems is given in Sect. 19.5.2.

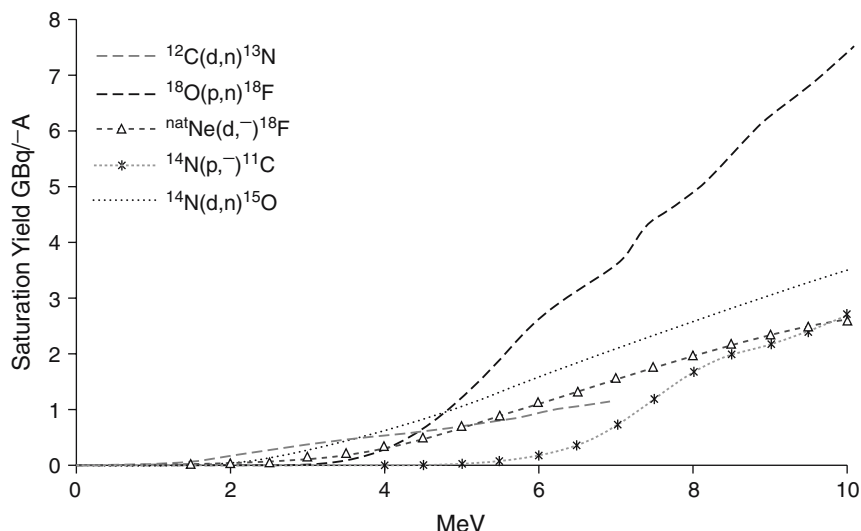


Fig. 19.1. Calculated thick-target saturation yields for some common PET radioisotopes [13]

19.4 Targetry Considerations

While total reaction yield is clearly important for radioisotope applications, there are several other considerations in developing production systems, particularly for medical applications.

By definition, a tracer imaging study requires a total mass of the material being tracked that does not itself perturb the physiological system studied. Therefore, it is crucial to maximize the specific radioactivity (SA) of the product, or the ratio of radioactive to nonradioactive tracer material. This is typically given in terms of GBq/ μmol . Some studies require higher SA than others, clearly, and the targetry requirements can be adjusted accordingly. For example, an imaging study using a ^{11}C -labeled neuroreceptor ligand may require a starting source of ^{11}C with an SA > 370 GBq/ μmol , while a simple ^{15}O -labeled-water study for flow may have no SA requirement to speak of.

Another consideration is radionuclidic purity. Impurities can interfere with the identification of the desired isotope activity and concentration, and, in the case of medical imaging, can cause problems with radiation dosimetry. In most cases, for PET applications, radioisotopic purity is not difficult to achieve. The low-energy beams used with tandem accelerators and the simplicity of the target materials rarely allow reactions other than the desired ones. In the case where other reactions are possible, or where the target composition presents alternate target nuclei to the beam, the radioactive impurities can usually be separated chemically, or in some cases of short

impurity half-life simply allowed to decay. Nonetheless, new systems must be monitored for impurities, and validated for their adequate removal.

Isotopes used in medicine are typically incorporated into some chemical form of physiological interest after production on the accelerator, so care must be taken to produce the isotope in an appropriate form to match the next step of chemical synthesis. For example, ^{13}N is readily made by low-energy deuteron irradiation of graphite, but the isotope is chemically trapped in the carbon matrix and not readily available for synthesis into a useful tracer form (e.g. ^{13}N -ammonia).

In addition to the correct chemical form of the radiotracer, care must be taken to minimize the amount of any stable chemical in the target material that could interfere with the subsequent chemistry. Using PET as an example again, one of the primary contaminants possible with aqueous ^{18}F -fluoride production is hydrocarbons. Trace amounts of ethanol, for example, can severely harm the downstream nucleophilic-substitution chemistry.

19.5 Examples of Isotope Production Systems

Every installation of an accelerator facility for medical radioisotope production will have its own unique requirements. While covering all possible contingencies is beyond the scope of this chapter, it can be useful to consider the specific example of a PET radioisotope production facility at the University of Wisconsin (UW), Madison. Much of the basic facility and targetry requirements met in this installation can serve as an example for what is needed elsewhere.

The Keck Laboratory for Functional Brain Imaging was created at the UW Waisman Center to provide a multidisciplinary resource for brain imaging and development studies. The lab, opened in early 2001, incorporates high-field MRI, high-resolution EEG, transcranial magnetic stimulation, and a full clinical and research PET facility.

Radioisotopes for the PET program are provided on site by an NEC 9SDH-2 tandem accelerator, purchased in 1996. The 9SDH-2 Pelletron was designed to provide $100\text{ }\mu\text{A}$ of 6 MeV protons or deuterons within a maximum 10 mm diameter beam spot. The actual performance regularly exceeds these specifications. The Torvis multicusp ion source typically achieves more than $150\text{ }\mu\text{A}$ [8]. The two chains, rated at $150\text{ }\mu\text{A}$ each, charge the high-voltage terminal, with demonstrated accelerated-beam currents in excess of $115\text{ }\mu\text{A}$. The dome voltage of 2.97 MV required for 6 MeV single-charge beams (with a 50 keV ion source voltage) is conservative, and the accelerator has been run up to 3.48 MV (for a 7.0 MeV beam). The beam optics components include low-energy steering, and high-energy quadrupole focusing and steering magnets. The tuning capabilities, coupled with an in-line rotating-wire beam profile monitor, allow fine, continuous control of the beam shape and position. The

independent adjustment of beam width in two dimensions is typically from 6 to 10 mm, with a full practical range of 2 to 20 mm FWHM.

Although this machine is intended primarily for the continuous production of short-lived tracers labeled with ^{15}O ($t_{1/2} = 122\text{ s}$) or ^{17}F ($t_{1/2} = 65\text{ s}$), high-power target systems have been developed to provide $[^{15}\text{O}] \text{H}_2\text{O}$ (yield at saturation 329 MBq/ μA), $[^{17}\text{F}] \text{F}_2$ (936 MBq/ μA), $[^{18}\text{F}]$ fluoride (411 MBq/ μA), $[^{18}\text{F}] \text{F}_2$ (370 MBq/ μA), in-target production of $[^{13}\text{N}] \text{NH}_3$ (78 MBq/ μA), $[^{11}\text{C}] \text{CO}_2$ (311 MBq/ μA) and $[^{11}\text{C}] \text{CH}_4$ (303 MBq/ μA).

Figure 19.2 shows the general layout of the tandem lab and the radiochemistry support facilities. Several points should be made that distinguish this layout from a typical nuclear-physics or engineering lab. First, this facility is intended for short-lived-radiotracer production. Placement of the accelerator close to the intended scanner is crucial. While some PET radioisotopes can be effectively transported from the production facility to the end-use point, many installations must also account for the fact that the staff responsible for the accelerator may also be involved in the radiochemistry and the actual performance of PET scan protocols. In the UW example, the decision was made to combine all the PET support activities in one location, thus minimizing staff requirements as well as the losses due to long transit times. The tandem vault is located immediately adjacent to both the radiochemistry labs and the PET scanner suite.

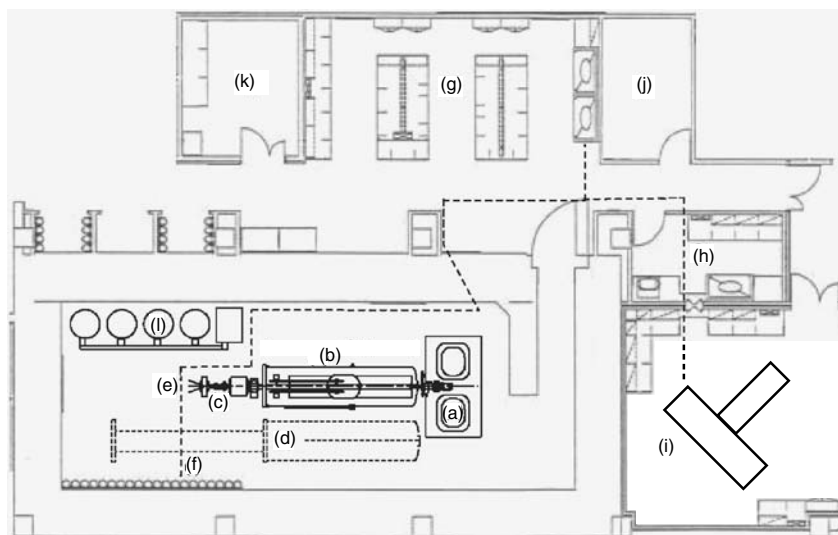


Fig. 19.2. Keck Laboratory radioisotope production and imaging area, including (a) Torvis ion source and control cabinets, (b) NEC 9SDH-2 accelerator, (c) focusing and switching elements, (d) accelerator position for maintenance, (e) target area, (f) gas supply lines, (g) lab area, (h) radiopharmacy, (i) PET scanner, (j) animal scanner and (k) machine shop

This lab arrangement puts rather stringent requirements on the radiation shielding, primarily for the neutrons resulting from (p, n) and (d, n) reactions during irradiation. While much can be achieved with localized shielding around the target areas, the UW installation went with a complete accelerator vault. The walls are standard concrete, >1.8 m thick all around. Care was taken to prevent direct neutron shine through any of the service access ports, with significant bends placed in the largest penetrations for air handling to ensure no leakage. The door is shielded from the target area by a dry-stacked wall of cement blocks, and is constructed of boron-impregnated high-density polyethylene backed by lead for stopping the neutron-capture gamma rays.

In total, the shielding and safety requirements for a medical installation are a significant cost, comparable to the accelerator cost itself. The source radiation can be significantly higher than that found in most nuclear-physics installations, and there is the added complication of shielding for the general public (e.g. PET scan subjects), sensitive local equipment (e.g. the PET scanner and other nuclear detection equipment) and the radiation workers themselves.

Shielding of personnel from the product radionuclide dose is obviously essential. In the case of PET radioisotopes, it is not uncommon to start with >100 GBq of activity produced. The primary shielding consideration is the 511 keV gamma ray dose, since the direct positron energies are typically < 2 MeV and easily shielded. Often the material is transported around a laboratory, from the accelerator target to chemistry stations to scanners, and must be handled appropriately at all points. The radioactive product is delivered via small-bore stainless steel or Teflon tubing from the target end of the accelerator, through a conduit in the shield walls, to a lead-lined trenching system leading to shielded chemistry stations or directly to the PET scanner. Table 19.1 lists the shielding characteristics for various materials, which can serve as a guide for analyzing transport lines, trenches etc. The gamma transmission drops exponentially with distance. Distance is also an effective means of minimizing dose rates. Equation (19.2) gives the dose rates measured for ^{18}F in air, at a distance d from the source:

$$\text{Dose rate} = 0.163 \frac{\mu\text{Sv m}^2}{\text{MBq h}} d^{-2} \quad (19.2)$$

Table 19.1. Shielding characteristics for various materials for 511 keV gamma rays. The fraction penetrating drops exponentially with thickness as $e^{-\mu d}$

Material	μ (cm $^{-1}$)
Lead	1.54
Steel	0.41
Cement	11
Glass	12

19.5.1 Target Entrance Windows

The UW installation was primarily designed for the shorter-lived PET radioisotopes, e.g. ^{15}O , although several standard and nonstandard medical isotopes have been produced with the NEC 9SDH-2 tandem. Most of the systems were based on existing PET targetry developments, and several excellent summary texts have been published on the field, e.g. [14, 15].

One of the primary concerns with radioisotope production common to most targetry systems is the proper handling of the beam power. The total beam power is linear with current and energy, typically on the order of 1 kW or more. This amount is not typically problematic, and simple water cooling of the target body material is generally sufficient, with examples shown for the targets discussed below. When one is using lower-energy beams, typical of most electrostatic-accelerator installations, special consideration is necessary to take to account of the increase in linear energy loss in the target windows. Liquid, gas and some volatile solid targets require thin windows to separate the target material from the beamline vacuum. Given the typical low yields for lower-energy beams, these target windows must be thin to preserve the available energy. Figure 19.3 shows the energy dependence of the beam energy loss in two common entrance-window foils, Havar and aluminum. These windows must be thick enough to withstand the pressure without rupture.

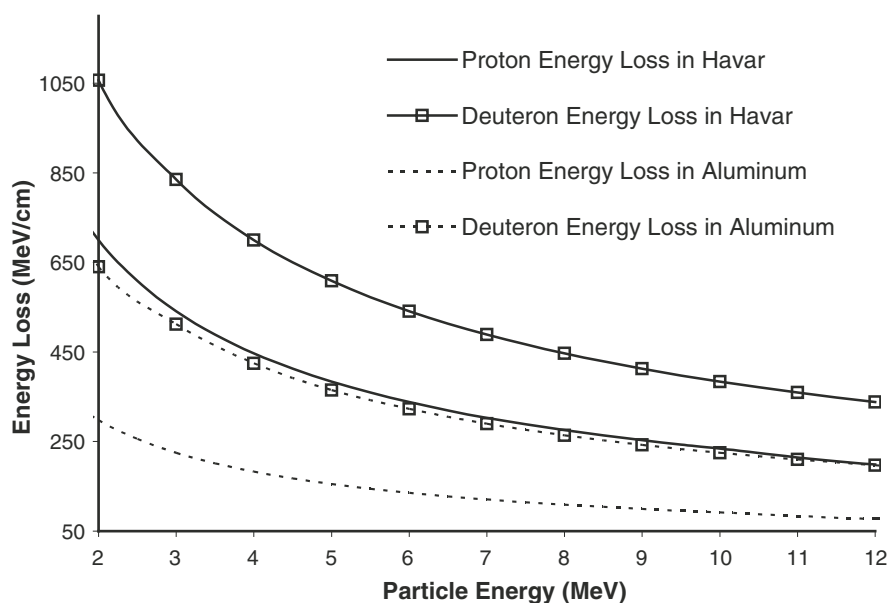


Fig. 19.3. Energy loss of protons and deuterons in aluminum and Havar

One solution to the problem of sealing a pressurized target volume to a vacuum system was provided by the method of the double-foil helium-cooled window [16]. This is a method that has been successfully used for radioisotope production with electrostatic accelerators [17]. An alternative and in some cases improved method has been reported [18, 19], employing a single thin window supported by a high-transmission grid. Gridded windows are now the preferred choice for radioisotope production targets for the UW 3 MV electrostatic tandem accelerator. The removal of one target window foil reduces the energy loss, and the low beam emittance allows the use of deep grids and efficient water cooling [20].

The support grid pattern consists of circular holes arranged in a hexagonal pattern. Trials were done with differing-size holes to maximize the allowable beam current. Figure 19.4 shows the basic design of this grid, with 80 holes of 1.7 mm diameter. The grid is constructed from a single aluminum unit, incorporating water cooling for maximum heat transfer with no material discontinuities. The deep grid holes provide increased material for heat transfer to the water cooling with negligible loss of the near-parallel beam. The aluminum was machined to a minimum wall thickness of 0.18 mm between the holes, with a grid depth of approximately 12.5 mm along the beam path. Deeper grid holes up to 25 mm deep have been used with no change in performance with the tandem accelerator. The targets for the tandem accelerator had the grid holes arranged to cover a 2.85 cm^2 area to minimize the heating of the grid. The grids are water-cooled through two straight channels on opposite sides of the support grid. Chilled water at 18°C flows at 2.3 l/min through the cooling channels of the grid. The support grid mounts onto the beamline with a KF-40 quick-connect flange. Single entrance-window foils have been tested with thicknesses ranging from 12.7 to $25.4 \mu\text{m}$ aluminum and 2.5 to $12.7 \mu\text{m}$ Havar.

Limitations of the water-cooled support grid have been found from foil failures occurring at high beam currents with narrower beam profiles. Typically the beam profile of 6 MeV deuterons is run at 8 mm FWHM in both directions, but when the beam profile is changed to a narrower 5.5 mm FWHM in both directions, $25.4 \mu\text{m}$ aluminum foils fail at above 70 μA (by developing

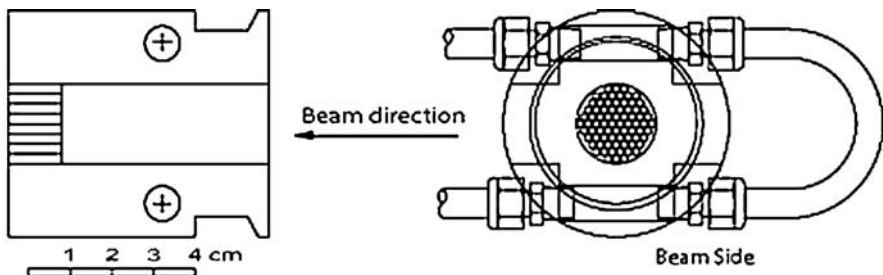


Fig. 19.4. Basic schematic of standard water-cooled support grid

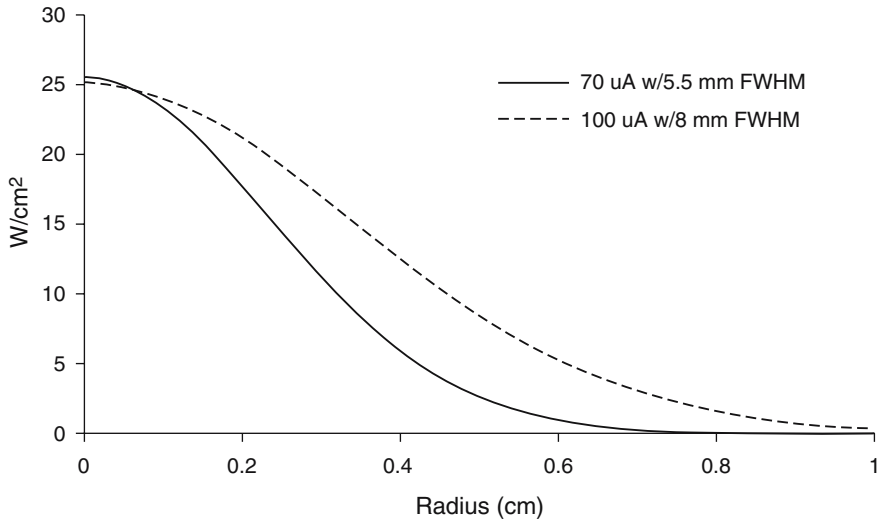


Fig. 19.5. Maximum power on a 1.6 mm diameter hole grid vs. beam profile on a 25 μm aluminum entrance window [20]

pinholes). Figure 19.5 shows the energy deposition from 6 MeV deuterons in 25.4 μm aluminum foil for both 70 μA 5.5 mm FWHM and 100 μA with 8 mm FWHM, assuming Gaussian profiles. From the plot, the energy per unit area is higher for the 70 μA with beam at the center; above this energy, the foils fail. The 100 μA beam is lower at the center, and without the resulting failures. This demonstrates the upper limit on the energy per unit area for these foils on grids with 1.7 mm diameter holes. This could also be used to estimate the performance with higher-energy protons and deuterons of the water-cooled support grid.

Optimization of the hole size was performed by calculation of the maximum current density on the target material. The peak current of the beam profile at the maximum allowable beam current on a particular grid gives the maximum current density on the grid. Estimation of the maximum current density on the grid for holes smaller than those constructed was performed by a least-squares fit of the known maximum current densities for the various grid sizes, including the maximum current density for an unsupported single-foil target window. The transmission for all grids was estimated under the constraint of a consistent minimum wall thickness of 0.18 mm, which is reasonable for fabrication by standard or wire-electron-discharge machining, with 0% transmission for infinitely small grid holes and 100% transmission for an unsupported window. The product of the transmission and the maximum beam current density on the grid is the maximum current density on the target material. Figure 19.6 shows the maximum current density on the target material vs. the grid hole size, using 25.4 μm aluminum target windows. The

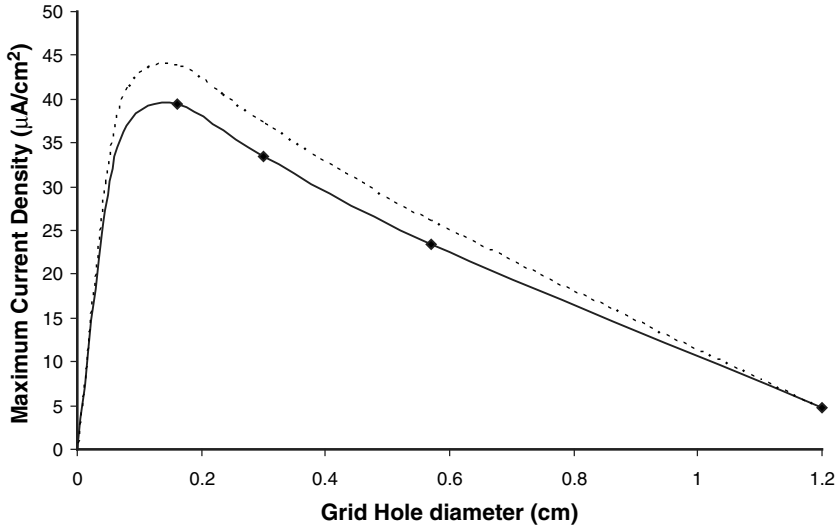


Fig. 19.6. Maximum current density ($\mu\text{A}/\text{cm}^2$) on the target material vs. grid hole diameter. In each case a minimum amount of material between holes of thickness 0.18 mm is assumed. The maximum current density is the product of the grid transmission fraction and the measured maximum allowable peak current density ($\mu\text{A}/\text{cm}^2$). Points (black dots) correspond to the measured maximum beam current with a $25.4\mu\text{m}$ aluminum window at ≤ 350 kPa on a flow-through gas target. The dashed curve is the estimated improvement in maximum current density if it is corrected for the improved transmission with hexagonal holes [20]

points correspond to the measured maximum beam current density and the grid hole size. The solid curve represents the estimations obtained from the measurements down to smaller grid hole diameter. The optimum point for maximum current density on the target material lies near the smallest grid hole diameter tested, below which the grid transmission loss dominates the maximum current density at smaller hole sizes.

Hexagonal grid holes have been proposed as a means of improving the in grid transmission [21, 22]. The improvements in grid performance obtained by using hexagonal holes instead of circular holes can also be calculated from the maximum current density on the target material. The improvement in transmission by the reduction of material is approximately 12%. Figure 19.6 shows the improvement in the maximum current density on the target material as a dashed curve above the calculation for circular holes.

19.5.2 Examples of Production Targets

Oxygen-15 ($t_{1/2} \sim 2$ min)

Oxygen-15 is one of the earliest radioisotopes used for PET studies, and continues to be used for studies of fast processes such as blood flow, using primarily $[^{15}\text{O}]$ water but in some cases $[^{15}\text{O}]$ butanol or other freely diffusible compounds. It has also been used directly in the form of $[^{15}\text{O}]$ O_2 gas for tissue oxygen utilization measurements, as well as $[^{15}\text{O}]$ CO and CO_2 for blood volume and flow measurements, respectively. The simplest reaction to use is $^{14}\text{N}(\text{d}, \text{n})^{15}\text{O}$, taking advantage of both the high yields at low energy and the economical use of the natural isotopic abundance of the target material. Gas targets are used, typically with 99% N_2 gas and an admixture of an appropriate gas to form the required chemical product. In the case of $[^{15}\text{O}]$ water, hydrogen is used as the mix gas. Similarly, replacing the hydrogen with oxygen produces $[^{15}\text{O}]$ O_2 .

Figure 19.7 shows the typical gas target used for ^{15}O systems. The target body is aluminum, a preferred material for many systems owing to ease of fabrication, high thermal conductivity, and low residual radiation activation from stray-beam impact. The gas chamber of the target body is 19 mm in diameter and 127 mm in length, with the outside diameter of the target being ~ 50 mm. The 25 μm aluminum target window is supported with the cooled grid as discussed.

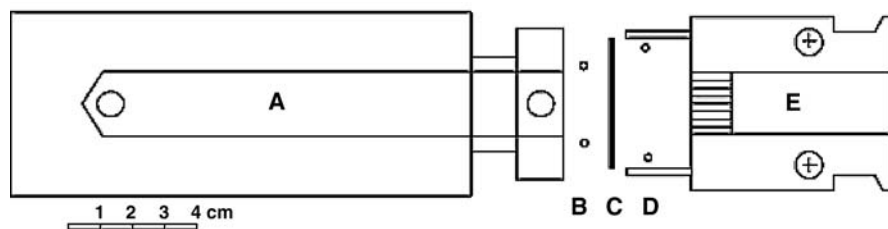


Fig. 19.7. Typical ^{15}O gas target assembly. (A) is the T-6061 aluminum gas target body, with a 19 mm i.d. and 127 mm length bore; (B) and (D) are Viton O-rings; (C) is the aluminum entrance foil; and (E) is the water-cooled support grid

Typical operation parameters for in-target $[^{15}\text{O}]$ water production use a flow-through technique. 350–700 kPa of premixed $\text{N}_2/1\%\text{H}_2$ fills the gas target body and flows to the chemistry area at a flow rate of 200–400 cm^3/min . A lower target pressure allows reasonable gas flow rates in the collection vials without the use of metering or needle valves in the gas stream. As all fittings and valves in the gas stream collect and condense the $[^{15}\text{O}]$ water, reducing the yield collected from the target, a minimum should be used between the target and the collection area. The beam current for $[^{15}\text{O}]$ water production

is usually between 10 and 20 μA . Irradiation and collection last for about 2–4 min, minimizing the amount of NH_3 produced. Collection of the activity for the study results in >4 GBq of $[^{15}\text{O}]$ water, of which 350 to 2500 MBq is injected.

Unlike $[^{15}\text{O}]$ water, $[^{15}\text{O}]$ O_2 does not stick to the target and the lines coming from the target to the chemistry hood. This target is generally shot as a static target and dumped in a reservoir with a minimum (reasonable) volume for inhalation. The target pressures are higher, ~ 650 kPa, and take advantage of the full thick-target yield. The beam current for $[^{15}\text{O}]$ O_2 production is similar to that for $[^{15}\text{O}]$ water, at 10–20 μA for oxygen utilization studies. Irradiation times for O_2 production are usually about 4–5 min. These preparations generally result in >10 GBq of $[^{15}\text{O}]$ O_2 .

Carbon-11 ($t_{1/2} = 20$ min)

Carbon-11 has been widely used for the labeling of novel research tracers for PET. The half-life is long enough to probe more complex physiological parameters, such as specific binding of radioligands to neuroreceptors. The fact that it is carbon means complex molecules can be turned into PET tracers with no chemical differentiation from the cold compound. This is particularly important when one wants to use PET to follow the tracer kinetics of labeled drugs at subpharmacological levels without affecting the behavior of the drug.

Several avenues have been explored for ^{11}C production; these include $^{11}\text{B}(\text{p}, \text{n})^{11}\text{C}$ and $^{14}\text{N}(\text{p}, \alpha)^{11}\text{C}$. The first provides reasonably high yields at low energy, but requires the use of solid targets, typically boron oxide, and subsequent extraction of ^{11}C for chemical use. The gas phase production with $^{14}\text{N}(\text{p}, \alpha)^{11}\text{C}$ suffers from a lower yield; however, it can greatly simplify the subsequent chemistry.

^{11}C is typically produced in the form of CO_2 or methane (CH_4). The basic system is identical to that for ^{15}O production, in that the target gas is primarily natural nitrogen, with a small admixture of an appropriate balance gas to produce the desired product. In the case of $[^{11}\text{C}]$ CO_2 , the mix is $<1\%$ oxygen, while for methane, a higher mix of hydrogen is optimal, on the order of 10%.

The UW $[^{11}\text{C}]$ methane target is machined from stainless steel rather than aluminum. The cylindrical gas chamber has a 19 mm diameter and a 135 mm length. Unlike the aluminum gas targets, which have good thermal conductivity, the stainless steel target requires a water jacket that is TIG welded to the target body at the front and rear. The gas stream enters at the rear of the target and leaves at the front on the mounting flange. A water-cooled support grid secures and cools the entrance window (Havar).

In-target $[^{11}\text{C}]$ methane is produced using a static shot technique. The target pressure inside the stainless steel gas target body is increased to 1 MPa for this reaction, as the protons have a longer range than deuterons. The

increased pressure also compensates for the reduced nitrogen and the reduced target thickness arising from the premixed $\text{N}_2/10\%\text{H}_2$. The beam current for $[^{11}\text{C}]$ methane production is usually between 50 and $100\mu\text{A}$ at 6.8 MeV, with a yield of 303 MBq/ μA .

Fluorine-18 ($t_{1/2} = 110\text{ min}$)

While ^{11}C has some clear advantages for producing PET tracers identical to many known drugs or natural compounds, the short half-life does limit its potential for widespread distribution. Also, some physiological processes may require longer times to reach equilibrium, requiring scans of 1–2 h to interpret fully.

The main reaction channels available with low-energy beams are $^{18}\text{O}(\text{p}, \text{n})^{18}\text{F}$ and $^{20}\text{Ne}(\text{d}, \alpha)^{18}\text{F}$. With higher-energy machines ($>10\text{ MeV}$), the yield advantage for the $^{18}\text{O}(\text{p}, \text{n})^{18}\text{F}$ reaction is quite substantial (see Fig. 19.1) and most target systems use the expensive isotopically enriched ^{18}O , in the form of either water or oxygen gas. While these systems have been used successfully with tandem accelerators [17, 23], the yield advantage compared with the neon target is diminished, and allows the use of natural neon, with no costly isotopic enrichment.

$^{20}\text{Ne}(\text{d}, \alpha)^{18}\text{F}$, for $[^{18}\text{F}]\text{F}_2$ Production

Two systems for ^{18}F production using $^{20}\text{Ne}(\text{d}, \alpha)^{18}\text{F}$ are in use at UW. The first system produces low-specific-activity $[^{18}\text{F}]$ fluorine gas (F_2). The basic methodology has been widely used in PET [24–27], producing ^{18}F -labeled compounds such as $[^{18}\text{F}]$ fluoroDOPA for Parkinson's disease research.

$[^{18}\text{F}]\text{F}_2$ is produced in an aluminum-body gas target chamber similar to the ^{15}O system, although some institutions use nickel or stainless steel targets to minimize reactions of fluorine with the walls. The target gas is natural neon, with a nominal 0.5% cold fluorine. Using a 6 MeV deuteron beam with the tandem accelerator, the UW target produces $[^{18}\text{F}]\text{F}_2$ at a saturation yield of 370 MBq/ μA at $100\mu\text{A}$. The target can be run in a continuous-flow or a static mode, although in the flow mode care must be taken not to exceed the maximum fluorine load capacity of the downstream chemistry (typically $<100\mu\text{mol}$ of F_2 for most preparations).

$^{20}\text{Ne}(\text{d}, \alpha)^{18}\text{F}$, for Aqueous $[^{18}\text{F}]$ Fluoride Production

The low specific activity from the gas system above limits its use for medical applications. Where high specific activity is required, the ^{18}F can be extracted in the form of aqueous fluoride, with no addition of cold fluorine. This is the approach most commonly used in PET, and is the direct result of proton irradiation of $[^{18}\text{O}]\text{H}_2\text{O}$ targets. However, one can still explore the production of no-carrier-added fluoride using the neon gas targetry, at a significant saving

in target materials compared with enriched ^{18}O . In the pure neon gas system, the ^{18}F ions drift to the walls of the target chamber and stick. The ^{18}F can then be washed off the walls with water, resulting in aqueous ^{18}F fluoride. In this system, the wall material and treatment can have a dramatic effect on the performance of the system. Several groups have studied empirically the use of different materials for similar wash-off systems [23, 28–30]. While there is still some debate as to the exact mechanisms involved, some common results are beginning to emerge. Niobium seems to be well suited to use with fluoride systems, both for direct fluoride from ^{18}O water and for our indirect wash-off target using the neon reaction.

The gas target for the production of ^{18}F fluoride via the $^{20}\text{Ne}(\text{d}, \alpha)^{18}\text{F}$ reaction varies from the other gas targets. The aluminum target body has been nickel-plated, and a niobium tube (15.7 mm I.D.) is inserted in the bore of the target. Holes drilled in the niobium tube align with the gas ports at front and rear, and with a third port centered on the bottom of the target, to drain the water wash. Water is preheated to 85°C , and kept hot through the washing process with heat tape wrapped around the target. The total water volume is approximately 20 ml per wash. Water washes of the interior surface of the gas target have given yields of 407 MBq/ μA at 85 μA .

$^{18}\text{O}(\text{p}, \text{n})^{18}\text{F}$, for Aqueous ^{18}F Fluoride Production

Liquid target systems have been developed for production of ^{18}F fluoride from $^{18}\text{O}(\text{p}, \text{n})^{18}\text{F}$ using enriched water targets with electrostatic tandem accelerators [17, 30, 31], as well as with most medical production cyclotrons (see e.g. [32]). The volume of the liquid target is generally kept low to minimize the use of enriched isotopes. The target cooling is more critical as hundreds of watts are usually imparted into less than 1 ml of liquid. The target pressure in a sealed system increases with the temperature of the liquid as it turns to vapor. Poisons from the target chamber can also detrimentally affect the chemistry post-irradiation. In the case of ^{18}F , aluminum cannot be used as a target body, since it not only binds the ^{18}F fluoride, but also poisons radiosyntheses. For this reason, silver, niobium or titanium is a better alternative.

Figure 19.8 shows a schematic drawing of the simple high-pressure, silver-body target used on the UW tandem. The water delivery lines are Teflon or HPLC (high-pressure liquid chromatography)-grade stainless tubing (0.8 mm I.D.). The switch valves at the target are manual stainless steel ball valves with Teflon packing (Whitey, 40 series). The target body is silver, with stainless steel tubing top and bottom for water loading. The tubing is close-fit to the silver, and then silver-soldered in place. The cylindrical beamstrike volume is 12.7 mm O.D. by 3 mm deep, sufficient to stop 6 MeV protons in water. The beamstrike volume is 380 μl , with a total of 500 μl required to fill between the switch valves. The target mounts onto the water-cooled support grid, similar in design to the one used for the gas target, with the window

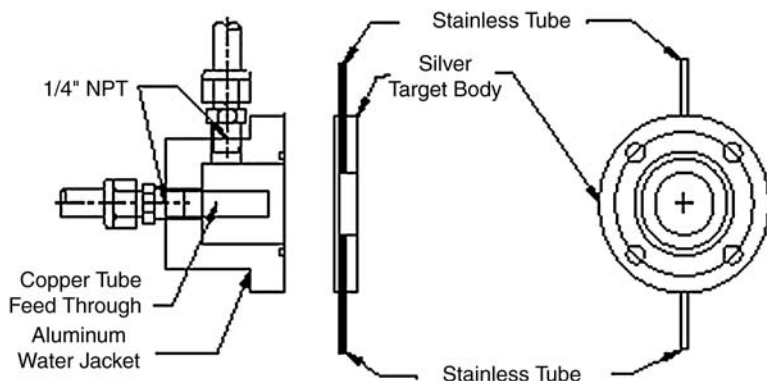


Fig. 19.8. Design of the UW ^{18}O water target

foil compressed against a Teflon O-ring. A water jet at the back of the body cools the target. 6.0 MeV protons irradiate the water through $12.7\text{ }\mu\text{m}$ Havar. The beam profile is kept to a maximum of 6 mm FWHM on the water target. While the reaction yield at ~ 6 MeV is significantly less than that for higher-energy machines, several GBq of useful ^{18}F can be produced, sufficient for research PET purposes.

References

1. J.D. Cockroft, E.T.S. Walton: Proceedings of the Royal Society A, **136**, 619 (1932)
2. R.G. Herb: IEEE Transactions, **NS-30**(2), 1359 (1995)
3. S.E. Arnell: Nuclear Physics, **24**, 500 (1961)
4. G.M. Klody, J.B. Schroeder, J.A. Ferry, T.J. Pollock, E.D. Adams, R.G. Herb: Nuclear Instruments and Methods in Physics Research, **56-57**, 704 (1991)
5. W. Haeberli, W. Gruebler, P. Extermann, P. Schwandt: Physical Review Letters, **15**(6), 267 (1965)
6. W. Haeberli, M.D. Barker, C.A. Gossett, D.G. Mavis, P.A. Quin, J. Sowinski, T. Wise, H.F. Glavish: Nuclear Instruments and Methods in Physics Research, **196**, 319 (1982)
7. K.E. Rehm, M. Paul, A.D. Roberts: Nuclear Physics A **616**, C115 (1997)
8. M.L. Sundquist, J.R. Adney, R.C. Schmidt: Nuclear Instruments and Methods B **99**, 684 (1999)
9. A.M. Friedman, O.J. DeJesus, P. Harper, et al.: Journal of Labeled Compounds and Radiopharmaceuticals, **19**(11-1), 1427 (1982)
10. P.V. Harper, B. Rich, K.A. Lathrop, B. Mock: Early Electrostatic Accelerators and Some Later Developments, IEEE Transactions, Report for the ID Atomic Energy Commission, IAEA-SM-185, P. 14 (1974)
11. R.J. Nickles, M.E. Daube, G.D. Hutchins: Journal of Labeled Compounds and Radiopharmaceuticals, **11**, 1365 (1983)
12. J.F. Janni: Atomic Data and Nuclear Data Tables, **27**, (1982)

13. S.M. Qaim, F.T. Tarkanyi, P. Oblozinsky, K. Gul, A. Hermanne, M.G. Mustafa, F.M. Nortier, B. Scholten, Y. Shubin, S. Takacs, Y. Zhuang, IAEA-TECDOC-1211, May 2001
14. J.C. Clark, P.D. Buckingham: *Short-Lived Radioactive Gases for Clinical Use*. Butterworths, London, 1975
15. G. Stöcklin, V.W. Pike: *Radiopharmaceuticals for Positron Emission Tomography*. Kluwer Academic, Dordrecht, 1993
16. B.W. Wieland: Proceedings of the First Workshop on Targetry and Target Chemistry, Heidelberg, pp. 14–16 (1985)
17. T. Ohlsson, A. Sandell, R. Hellborg, K. Håkansson, C. Nilsson, S.E. Strand: Nuclear Instruments and Methods in Physics Research A, **379**, 341 (1996)
18. R.J. Nickles: Nuclear Instruments and Methods in Physics Research A, **177**, 593 (1980)
19. D.J. Schlyer, M.L. Firouzbakht, I. Garcia, R.A. Ferrieri: Application of accelerators in research and industry, AIP Conference Proceedings 392, p. 1363 (1997)
20. T.E. Barnhart, A.K. Converse, K.A. Dabbs, R.J. Nickles, K.R. Buckley, S. Jivan, T.J. Ruth, A.D. Roberts: Water-cooled grid support system for high power irradiation with thin target windows, submitted to Applied Radiation and Isotopes **58** (1) pp. 21–26, 2003
21. G. Bida, R.E. Ehrenkaufer, A.P. Wolf, J.S. Fowler, R.R. MacGregor, T.J. Ruth: Journal of Nuclear Medicine, **21**, 758 (1980)
22. J.A. Nye, D.W. Dick, R.J. Nickles: 17th International Conference on the Application of Accelerators in Research and Industry, Denton, TX, AIP Conference Proceedings 680, P. 1098 (2002)
23. T.E. Barnhart, R.J. Nickles, A.D. Roberts: 17th International Conference on the Application of Accelerators in Research and Industry, Denton, TX, AIP Conference Proceedings 680, P. 1086 (2002)
24. G. Bida, B.W. Wieland, J. Lenz, C. Alvord: Proceedings of the Ninth International Workshop on Targetry and Target (Chemistry, Turku, Finland, pp. 24–29, May 23–25 2002)
25. G. Blessing, H.H. Coenen, K. Franken, S.M. Qaim: Applied Radiation and Isotopes, **37**, 1135 (1986)
26. V. Casella, T. Ido, A.P. Wolf, J.S. Fowler, R.R. MacGregor, T. Ruth: Journal of Nuclear Medicine, **21**(8), 750 (1980)
27. B.W. Wieland, D.J. Schlyer, A.P. Wolf: International Journal of Applied Radiation and Isotopes, **35**, 387 (1984)
28. F. Helus, V. Uhler, G. Wolber, H. Gasper, W. MaierBorst: Journal of Radioanalytical and Nuclear Chemistry, **182**, 445 (1994)
29. T.J. Ruth, K.R. Buckley, K.S. Chun, E.T. Hurtado, S. Jivan, S. Zeisler: Applied Radiation and Isotopes, **55**, 457 (2001)
30. R.J. Nickles, R.D. Hichwa, M.E. Daube, G.D. Hutchins, D.D. Congdon: International Journal of Applied Radiation and Isotopes, **34**, 625 (1983)
31. A.D. Roberts, R.J. Davidson, R.J. Nickles: 15th International Conference on the Application of Accelerators in Research and Industry, Denton, TX, AIP Conference Proceedings 475, p. 106 (1998)
32. M.R. Kilbourn, P.A. Jerabek, M.J. Welch: International Journal of Applied Radiation and Isotopes, **36**, 327 (1985)

20 Nuclear Structure

C. Fahlander and D. Rudolph

Department of Physics, Lund University, Box 118, 221 00 Lund, Sweden
`claes.fahlander@nuclear.lu.se`, `dirk.rudolph@nuclear.lu.se`

20.1 Introduction

The atomic nucleus is a finite many-body quantal system consisting of a well-defined number of protons, Z , and neutrons, N , commonly called the nucleons, which are bound together by the strong nuclear force. Under its influence, the nucleus displays a remarkable diversity of phenomena and symmetries involving many different degrees of freedom, and still, after decades of studies, when pushed to the very limits of its stability, it continues to surprise physicists as completely unexpected properties are revealed by increasingly sensitive instruments. The improved sophistication of experimental techniques in the past two decades has allowed us to study nuclei under extreme conditions, and to follow the evolution of nuclear structure to the limits in several directions. At the limit of angular momentum, we study extremely rapid nuclear rotations in individual nuclei. At the limit of excitation energy, we probe nuclear properties at high temperature. In nuclei very far away from the valley of stability, we seek the limits of existence of nuclear matter: Where are the proton and neutron drip lines, and where does the Table of Elements end?

The principal goal is to come to an understanding of the strong nucleon–nucleon interaction. It is affected by the nuclear medium, which responds differently when pushed to extremes in rotational frequency, temperature, isospin, or mass. The information we seek is revealed by studies of the excitation modes and decay properties of the nucleus. Many exciting new discoveries have been uncovered in nuclear-structure physics in recent years, only a few of which can be discussed in this chapter. It will focus on a selection of results under extreme conditions obtained with state-of-the-art γ -ray multidetector arrays.

20.2 Instrumentation

20.2.1 Electrostatic Accelerators

Electrostatic accelerators are ideal for the study of nuclear reactions (see Chap. 21) and nuclear structure. They can deliver a large variety of stable

beams, from protons up to very heavy ions, with large duty cycles and high intensities. The type of nuclear reaction to employ in a specific experiment is determined by the nucleus of interest and the particular quantum states one wants to investigate. Nuclei on the neutron-deficient side of stability are most efficiently reached by heavy-ion fusion-evaporation reactions, whereas moderately neutron-rich nuclei can be populated in, for example, quasi-elastic or multinucleon transfer reactions. However, the production rate of nuclei with extreme values of isospin is small in almost any kind of nuclear reaction involving stable beams and targets. This must be compensated for by high efficiency and selectivity of the detector systems.

The first electrostatic generator accelerating a nuclear beam was reported by Van de Graaff in 1931 [1], but it took until 1951 before the first real electrostatic accelerator was produced by HVEC. During the 1950s and 1960s, HVEC delivered more than forty tandem accelerators from their very successful EN and FN series of accelerators to laboratories and universities all over the world. These two decades saw a very rapid development of the basic nuclear models of the atomic nucleus, which was the result of a very close interplay between theory and experiment, and it is clear that the early tandem accelerators played an enormously important role in this development. Today it is possible to accelerate heavy ions to beam energies beyond the Coulomb barrier for many projectile–target combinations using tandem accelerators of the FN, MP, XTU, and ESTU types from HVEC, Pelletron accelerators from NEC, or high-voltage machines built by physicists at their home institutes. An example of the latter has been the VIVITRON tandem accelerator (terminated in December 2003) at the Institut de Recherches Subatomique (IReS) in Strasbourg. To extend our present knowledge of nuclear structure with stable-beam experiments requires high-intensity accelerators. However, for a full exploration of the nuclear landscape we must finally employ radioactive-ion-beam (RIB) facilities. Electrostatic accelerators will, in this context, continue to play a role as demonstrated at, for example, the Holifield Radioactive Ion Beam Facility at Oak Ridge National Laboratory, USA, where the produced radioactive ions are injected into and accelerated by a 25 MV tandem accelerator.

20.2.2 Detectors

The most powerful tool to study the structure of nuclei is high-precision γ -ray spectroscopy of the decay of their excited states. During the last two decades, developments in the techniques of γ -ray spectroscopy have led to the construction of large 4π multidetector systems comprising several hundreds of escape-suppressed detector elements. These developments have set new limits on γ detector technology and have provided a significant step forward in basic nuclear-structure research.

One such state-of-the-art instrument has been EUROBALL [2–4], which is shown in Fig. 20.1. It comprises 239 Ge crystals covering 45% of the solid

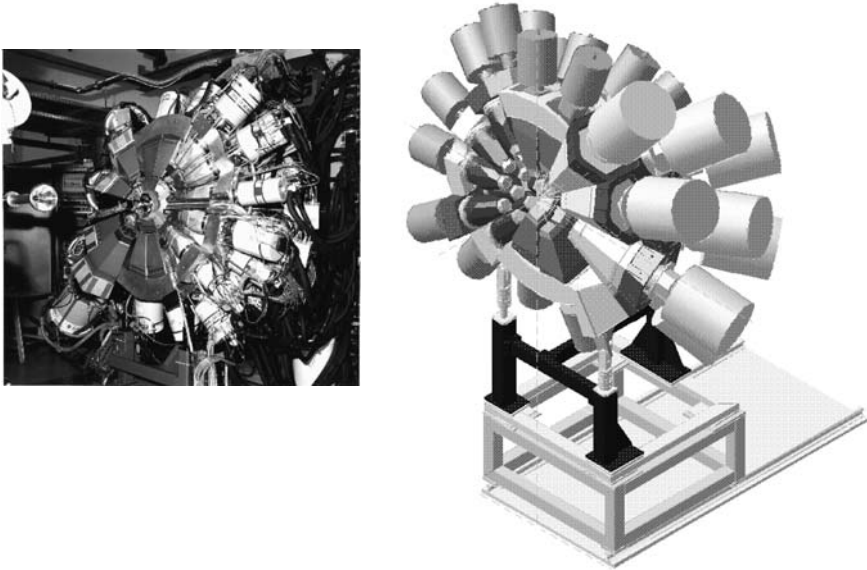


Fig. 20.1. The *left-hand* side shows a picture of the open EUROBALL array at Strasbourg, France. The *right-hand* side sketches the Ge detectors of EUROBALL. Courtesy of John Simpson

angle, and was installed at two of the main nuclear-structure facilities at Europe: the XTU-ALPI accelerator facilities of the Legnaro National Laboratory (LNL), Italy, and the VIVITRON tandem accelerator facility at IReS, Strasbourg, France. GAMMASPHERE is a similar instrument developed in the USA [5–7].

Both EUROBALL and GAMMASPHERE combine high efficiencies for detecting γ -rays with high energy resolution. In particular, they are very powerful for the measurement of high-multiplicity γ -ray cascades. Still, their performance is limited to an efficiency of about 10% for γ -ray energies of 1 MeV, and to a response function (signal-to-noise-ratio) with a maximum of 60% of the total intensity in the full energy peak. The latter is achieved by rejecting the signal from the Ge detector when the surrounding scintillator shield detects Compton-scattered γ -rays from the Ge crystal. These suppression shields limit the coverage with Ge detectors to less than one half of the solid angle, reducing the high-resolution efficiency considerably. The next major advance requires the replacement of the passive suppression shields by active Ge detectors. A solution to the problem of Compton scattering is the tracking of the γ -ray interaction path. Recent advances in crystal segmentation technology and digital signal processing have opened up the possibility of operating the Ge detectors in a position-sensitive mode with high count rates. This enables the construction of a compact 4π array made solely of

Ge detectors. It is expected that segmented Ge detectors combined with the tracking technique will improve the sensitivity for coincidence γ -ray detection by two orders of magnitude over the current generation of Ge arrays. Two such arrays are presently being developed, AGATA [8] in Europe and GRETA [9] in the USA.

To study exotic nuclei, the γ -ray detection has to be combined with efficient and selective ancillary detectors, such as devices for the detection of light charged particles, neutrons, evaporation residues, fission fragments, and binary-reaction products.

20.3 Limits of Angular Momentum

Nuclei generate angular momentum in basically two different ways, by means of the single-particle motion of the individual nucleons and by the collective vibrational or rotational motion of the nucleus as a whole.

The total angular momentum of a nucleon, $\mathbf{j}$, is the sum of its orbital angular momentum, $\mathbf{l}$, and intrinsic spin, $\mathbf{s}$. Owing to the specific properties of the so-called pairing interaction, nucleons prefer to couple into pairs with angular momentum zero, similar to the Cooper pairs of electrons. Therefore, in the ground state of an even-even nucleus, the total angular momentum is also zero, and nuclear matter in its ground state is superfluid. To generate angular momentum from single-particle motion requires the successive breaking of nucleon pairs combined with the promotion of individual nucleons to higher-lying orbitals, which generally leads to an irregular sequence of states and γ -decay patterns. In contrast, regular sequences of states are due to collective rotational motion. The states form rotational bands for which the excitation energy is proportional to the square of the angular momentum, i.e. $E_x \sim I(I+1)$.

Single-particle and collective modes of excitation can coexist in one single nucleus. Figure 20.2 shows the decay scheme of the doubly magic nucleus ^{56}Ni , which is spherical in its ground state and develops the expected irregular pattern on top of it (left-hand side of Fig. 20.2). At about spin $J = 10\hbar$ and an excitation energy of ~ 10 MeV, however, two regular rotational bands compete with the spherical states (middle of Fig. 20.2). While most of the decay intensity proceeds via normal γ -decay paths into the ground state of ^{56}Ni , one state in the second rotational band emits, in about 50% of cases, a monoenergetic proton, leading to the spherical ground state of ^{55}Co [10].

In quantum mechanics, a rotation can arise only for systems having a deformed shape. In an even-even deformed axially symmetric nucleus, the rotation takes place around an axis R perpendicular to the symmetry axis z (see Fig. 20.2). The deformed nucleus is brought into rotational motion by bombardment with a nuclear projectile. The quantum mechanical description of the rotor gives rise to rotational quantum states, and information

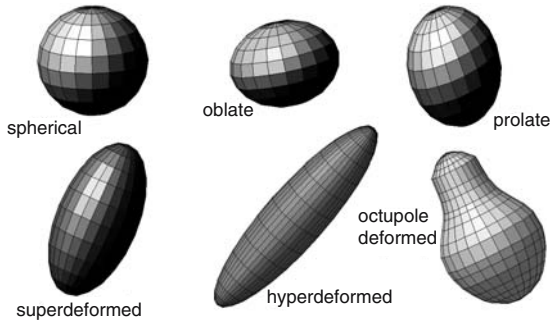


Fig. 20.3. Shapes of nuclei

interest to probe the properties of such highly excited single-particle states in deformed axially symmetric nuclei. These so-called *K*-isomers can have lifetimes of several years [12].

The advances in nuclear instrumentation have allowed new insights into nuclear behavior at low angular momentum also. Vibrational multiphonon states have been observed in several nuclei, which provide a more profound understanding of the mechanisms of nuclear vibrations. Important concepts in the description of low-lying collective states in nuclei are the dynamical symmetries, which have gained enormous success with the interacting-boson model [13]. Recently, a related concept, that of the so-called critical point symmetries, has been introduced to elucidate the nature of quantum phase transitions in nuclei [14]. Experimental programs around these symmetries are ongoing at the tandem accelerator laboratories at Yale University, USA, and the University of Cologne, Germany.

20.3.1 Superdeformation and Chaos in Nuclei

The most common deformed shapes of nuclei are axially symmetric prolate shapes. They are found in the ground states of nuclei in, for example, the rare-earth and actinide regions of the nuclidic chart. Some of the actinide nuclei have been found to be superdeformed in a second minimum of the potential-energy surface [15], in which the ratio between the long and short axes of the nucleus is 2:1 (see Fig. 20.3). At high angular momentum, superdeformation has been inferred from long cascades of equally spaced γ -rays, first discovered at the tandem accelerator at Daresbury Laboratory, UK, in 1986 in the nucleus ^{152}Dy [16]. With the EUROBALL and GAMMASPHERE arrays, designed to pick out high-multiplicity cascades of γ -rays, discrete nuclear quantum states of superdeformed rotational bands have since then been identified in many nuclei in several mass regions throughout the nuclidic chart [17].

The rotational motion in the superdeformed band is ordered, i.e., it is well characterized by a set of quantum numbers. The heads of the superdeformed

bands lie about 3–5 MeV above the normally deformed states with similar angular momentum. Since the level density increases exponentially with excitation energy, the superdeformed states are embedded in a sea of numerous normal deformed states, which may be described in terms of chaotic motion. The superdeformed states represent order in chaos, and the γ decay-out from the superdeformed bands serves as a probe to study chaos in nuclei [18].

20.3.2 Hyperdeformation and Clustering Phenomena

A prolate axially symmetric structure with a long-to-short axis ratio of 3:1 is called hyperdeformed. It has been predicted in different mass regions at very high angular momenta and excitation energies, and possible evidence for such a highly deformed structure has been sought in many experiments. The entrance channel into the hyperdeformed structure is very weak, and much emphasis has been devoted to studying the interplay between reaction dynamics, binding energies, and fission barriers to optimize the population of the hyperdeformed structures. It has been possible to observe, through the so-called ridge structure in two-dimensional γ -ray spectra, rotational structures from very weak decay paths of the nucleus. The width between the ridges is inversely proportional to the moment of inertia of the deformed nucleus, and in some nuclei the deduced moments of inertia correspond to very large deformations. Ongoing experimental work is trying to find evidence of discrete γ -ray transitions deexciting the hyperdeformed rotational states. An extensive experimental search for these transitions was recently undertaken in ^{126}Ba [19] at EUROBALL.

Very elongated shapes are also predicted in light nuclei. They are related to nuclear clustering phenomena, i.e., to strongly bound substructures such as alpha particles. Three alpha particles in a row, for example, would constitute a hyperdeformed structure in the ^{12}C nucleus. Their existence is mainly deduced from the observation of resonances in binary reaction channels, but recently such states have been searched for in γ -ray spectroscopy using EUROBALL coupled to dedicated ancillary detectors.

20.3.3 Triaxial Rotational Motion and Wobbling

At high excitation energy, nuclei may also develop shapes in which the axial symmetry is broken. This has to do with a subtle balance between the effects of the Coriolis and centrifugal forces and the influence of specific nucleon orbitals coming close to the Fermi surface of the nucleus. The rotation of an even–even nucleus in its ground-state configuration gives rise to a regular rotational band. However, the regularity may be influenced by the excitation of individual nucleons to higher-lying orbitals, which will tend to align their angular momenta with the axis of collective rotation. A classical image of such a situation is known from studies of the rigid top, which under specific initial

conditions will produce a characteristic precession phenomenon. Generally, one associates the word “wobbling” with this type of motion. A prerequisite for a wobbling mode in a nucleus is the breaking of the axial symmetry of the nucleus. The precession motion is represented by small-amplitude fluctuations of the rotational axis away from the principal axis, which destroys the regularity of the rotational band. In addition, the wobbling mode of the nucleus is associated with a wobbling-phonon quantum number, giving rise to one- and two-phonon states, on top of which the rotational motion is superimposed. Excitations of the wobbling mode were predicted long ago [20, 21] and have only recently been identified in Lu isotopes [22, 23] owing to the high sensitivity of EUROBALL.

20.3.4 Magnetic Rotation and Chiral Symmetry

Low-lying excited states in near-spherical nuclei are based on single-particle excitations, which implies an irregular pattern of emitted γ -rays. However, regular patterns of energies in γ -ray cascades have been detected in nuclei that were known to be almost perfect spheres. The observed angular momenta in these nuclei are generated by a few of the valence protons and valence neutrons. The behavior has been termed “magnetic rotation” because the sequences of transitions arise not from the collective rotation of a charged, prolate deformed nucleus, which emits electric quadrupole radiation, but from the anisotropy of the currents in the nucleus. These currents produce a rotating magnetic dipole moment, giving rise to a regular sequence of magnetic-dipole radiation. This has been investigated at both the EUROBALL and the GAMMASPHERE detector array [24].

A related topic is spontaneous chiral symmetry breaking, giving rise to left- and right-handed systems as in molecules. This has been discovered in odd-odd nuclei with triaxial shapes. The angular momenta of a valence proton, a valence neutron, and the core rotation are mutually perpendicular. The angular-momentum vectors can thus form a left- and a right-handed system, related by the chiral operator, which combines time reversal and rotation by 180 degrees. Spontaneous chiral symmetry breaking is manifested as degenerate doublets of γ transitions in two rotational bands [25].

20.4 Nuclear Structure Far from Stability

Most nuclear models, both phenomenological and microscopic, give a relatively good description of the properties of stable and near-stable nuclei, but it remains to be seen to what extent these models are directly applicable to nuclei with extreme values of the ratio N/Z . One of the central issues of contemporary nuclear-structure physics is to explore these nuclei, because they are also of immediate interest in astrophysical processes. They are loosely

bound quantum systems, and their structures, their dynamics, and their spectral responses are different from those of stable nuclei. The proton drip line lies relatively close to the valley of stability. However, since neutrons have no charge, many of them can be added to stable nuclei before the next one added becomes unbound, and neutron skins and halos may develop. Neutron-rich nuclei are much more difficult to reach experimentally than neutron-deficient nuclei.

What are the properties of the effective nucleon–nucleon interaction in a nuclear medium which are different from those in stable nuclei? The average mean field experienced by a single nucleon will be modified. The spin–orbit splitting of the mean-field orbitals will change owing to a more diffuse nuclear surface of the weakly bound nucleons. This will influence the magic numbers for nuclei far from the valley of stability. Superconducting correlations will play an enhanced role. One expects to encounter novel types of shell structures, new collective modes, new pairing phases, or regions of nuclei with special deformations and symmetries. There is a great need for experimental information on the proton-rich as well as neutron-rich side of the valley of stability to obtain a complete picture of the internal structure of weakly bound nuclei.

20.4.1 Neutron–Proton Pairing

Isospin, or isotopic spin, $\mathbf{t}$, is simply a mathematical form to express the fact that the proton and the neutron are two different quantum states of the same particle, the nucleon. They have the same isospin quantum number $t = 1/2$ but different isospin projection quantum numbers, $t_z = -1/2$ for the proton and $t_z = +1/2$ for the neutron. A pair of nucleons thus has total isospin $T = 1$, with projections $T_z = -1, 0$, and $+1$, for the proton–proton (p–p), neutron–proton (n–p), and neutron–neutron (n–n) pairs, respectively, or $T = 0$ with projection $T_z = 0$ for an n–p pair. Nucleons are fermions, and thus their intrinsic spin is $s = 1/2$. Quantum mechanically, there are only two possible spin directions for each nucleon, $s_z = +1/2$ (spin up) and $s_z = -1/2$ (spin down), which sum to $S_z = 0$ or 1 for a nucleon pair (see Fig. 20.4). In the $T = 1$ isospin triplet state, the three different pairs all must have $S_z = 0$ owing to the Pauli principle, but the n–p pair may also couple to $S_z = 1$ in the $T = 0$, $T_z = 0$ isospin singlet state. The total isospin projection of the nucleus is $T_z = (N - Z)/2$, which is obtained by summing over all the nucleons in the nucleus. For an isospin-symmetric nucleus, this gives $T_z = 0$.

The importance of n–p pairing is well known from the binding properties of the deuteron, but its effects on the structure of heavier nuclei are normally hidden because of the preponderance of like-nucleon pairs. In $N = Z$ nuclei, however, the correlation from n–p Cooper pairs may coexist with, or replace, the p–p and n–n Cooper pairs, finally giving rise to a possible n–p pair condensate.

The mass of any nucleus is less than the sum of its proton and neutron masses. In creating the nucleus, some of the original mass of the Z protons and N neutrons has been transformed into binding energy according to the famous law of Einstein, $E = mc^2$. The energy gain can be calculated according to the mass formula of Weizsäcker from the 1930s. This formula contains a term which takes into account the pairing interaction and the fact that even–even nuclei are more strongly bound than their odd–even or even–odd neighbors. In the 1960s, however, it was discovered that the $N = Z$ nuclei were lighter than the mass formula predicted. To account for this extra binding, a new term was introduced, the Wigner term, which takes into consideration the very special situation when the numbers of protons and neutrons are exactly equal. It was not, however, understood why this term was needed. One possibility may have to do with the fact that, differently from electrons, nucleons have four different possibilities for forming pairs (Fig. 20.4). The $N = Z$ nuclei gain extra energy and become lighter because of the $T_z = 0$ neutron–proton pair correlations. The effect is only observed in $N = Z$ nuclei. If there are two more neutrons than protons, or vice versa, the effect is only half as big, and if there is an imbalance of four nucleons the effect is completely gone [26]. To form this very special neutron–proton pair condensate, it seems as if the first prerequisite is isospin symmetry of the nuclear matter.

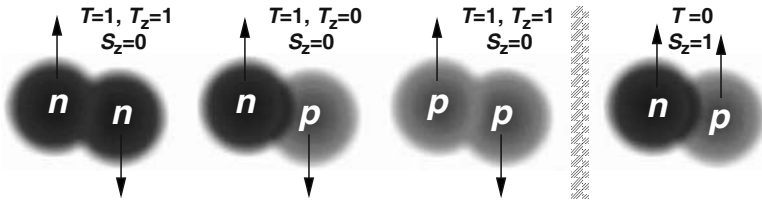


Fig. 20.4. Coupling schemes of dinucleons in the isospin formalism

20.4.2 Doubly Magic $N = Z$ Nuclei

The $N = Z$ nuclei are quantum systems where the protons and neutrons occupy orbitals with all quantum numbers equal except charge. There are 99 such nuclei which are bound with respect to particle emission. The lightest is the deuteron, which has one proton and one neutron. It is bound in its $T = 0$ state but not in its $T = 1$ state, i.e., the $T = 0$ n–p pair interaction is a few hundred keV more attractive than the corresponding $T = 1$ coupling. However, this preference changes at larger values of Z . The heaviest $N = Z$ nucleus in which it has been possible to study excited states is ^{88}Ru [27]. Large variations are found in the shapes of the $N = Z$ nuclei, ranging from spherical via oblate to prolate deformation [28, 29]. They are thus sensitive probes of how small changes in the properties of the last few valence nucleons

can result in dramatic changes in the shape of the nucleus and, hence, in the average nuclear mean field.

At the proton and neutron numbers 2, 8, 20, 28, and 50, which correspond to the elements helium, oxygen, calcium, nickel, and tin, the nuclei are nearly spherical in their ground states. These are the known magic numbers, which are related to the shell structure of the nucleus. In the spherical-shell model, each nucleon is assumed to move in an average potential generated by its interactions with all the other nucleons in the nucleus. Such a potential leads to the prediction that the quantum levels in a nucleus form shells within which several nucleons can reside. A magic number is the maximum number of nucleons which the Pauli principle allows for a particular shell. A critical ingredient of the average potential to account for the observed magic numbers is the spin-orbit interaction, which is a property of the nuclear force reflecting the fact that it is dependent on the direction of the nucleon's intrinsic spin vector $\mathbf{s}$ relative to its orbital angular-momentum vector $\mathbf{l}$. The shell model also takes into account the residual interaction between the valence nucleons that is not absorbed into the average potential. It provides a beautiful framework for predicting and understanding the lowest-lying states in near-spherical nuclei.

The magic $N = Z$ nuclei are in fact doubly magic. Their protons and neutrons occupy the same quantum states, and the proton and neutron wave functions have a large overlap. The wave functions reinforce each other's structure and amplify shell effects, and these nuclei are ideal testing grounds for the spherical-shell model (see Fig. 20.2). The nucleus ^{100}Sn is a special case, since it is the heaviest doubly magic isospin-symmetric nucleus that is stable against particle emission. The bare fact of its existence was proven a few years ago at the accelerator facilities of the Gesellschaft für Schwerionenforschung (GSI) in Darmstadt, Germany [30] and at the Grand Accélérateur National d'Ions Lourds (GANIL) in Caen, France [31]. To gain insight into the structure of ^{100}Sn requires the study of γ -rays emitted from its excited states, which is not feasible with present-day instruments owing to the very low production rate of ^{100}Sn . However, it is possible to study nuclei close to ^{100}Sn . The closest neighbors with known excited states are presently ^{98}Cd [32] and ^{102}Sn [33], which are only two proton holes and two neutrons, respectively away from the ^{100}Sn core. They were observed for the first time at the tandem accelerator laboratory of the Niels Bohr Institute in Denmark, and the lifetimes of a low-lying isomeric 8^+ state in ^{98}Cd and a 6^+ state in ^{102}Sn were measured. These lifetimes provide information on the polarizability, or rigidity, of the ^{100}Sn core. A doubly magic nucleus is expected to be stiff against such polarization effects. The measured lifetimes in ^{98}Cd and ^{102}Sn indicate that the ^{100}Sn core is indeed stiff against polarization attempts from protons, but rather soft against the interaction with neutrons. This may be an effect of the closeness of ^{100}Sn to the proton drip line, or of the importance

of the $T = 0$ n-p pairing, or it may indicate a breakdown in isospin symmetry between neutrons and protons.

20.4.3 Mirror Nuclei and Isospin Mixing

Isospin-breaking effects can preferably be studied in pairs of mirror nuclei, in which the numbers of protons and neutrons are interchanged, i.e., they lie on opposite sides of the $N = Z$ line. Isospin-breaking effects lead to shifts of typically 10–100 keV between the excitation energies of corresponding states in a mirror pair, the so-called mirror energy differences (MEDs), which have proven to be precise and challenging probes of nuclear structure. The experimental studies so far have concentrated on the $1f_{7/2}$ shell, i.e., nuclei between ^{40}Ca and ^{56}Ni [34]. This region is one of the best understood theoretically, and high-quality wave functions are available from modern shell-model calculations, which greatly simplifies the task of explaining quantitatively the MEDs. The most striking outcome of these studies is the importance of nuclear-charge-symmetry-breaking effects that may be as important as the Coulomb force [35].

The result of a recent study of the lighter $A = 35$ mirror pair [36] is displayed in Fig. 20.5. A comparison of the excitation schemes of the weakly populated $T_z = -1/2$ isotope ^{35}Ar with the well-known $T_z = +1/2$ mirror partner ^{35}Cl reveals two remarkable features: (i) a dramatic difference in the decay patterns of the $7/2^-$ states, and (ii) a very large MED value for the $13/2^-$ states.

The explanation of the latter requires consideration of the hitherto neglected relativistic electromagnetic spin-orbit coupling between nucleons, which for electrons gives rise to the fine structure in atomic physics. Shell-model calculations show that the wave functions of the $13/2^-$ states are rather pure – a single proton or neutron for ^{35}Ar or ^{35}Cl , respectively, in the

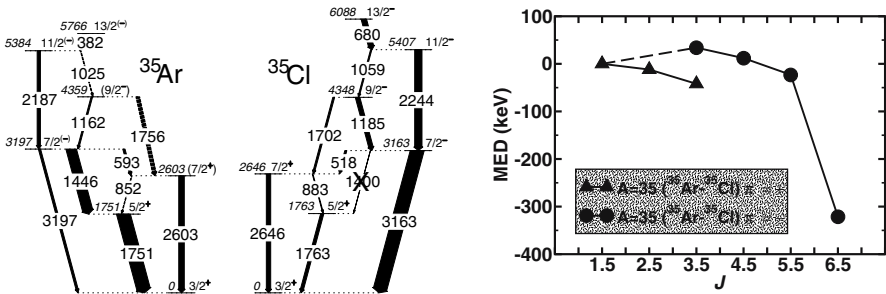


Fig. 20.5. The *left-hand* side provides relevant parts of the decay schemes of the $A = 35$ mirror nuclei ^{35}Ar and ^{35}Cl [36]. The *right-hand* side shows the mirror energy differences of the $A = 35$ pair, $\text{MED} = E_x(^{35}\text{Ar}) - E_x(^{35}\text{Cl})$, as a function of the angular momentum J

$1f_{7/2}$ orbital above the $N = Z = 20$ shell gap coupled to a $J^\pi = 3^+$, $T = 0$ neutron–proton pair in the $1d_{3/2}$ orbital below this gap. The excitation of a particle from a $j = l - s$ orbital ($1d_{3/2}$) into a $j = l + s$ orbital ($1f_{7/2}$) combined with the single-particle character of the states leads to a contribution of some 200 keV from the electromagnetic spin–orbit coupling to the MED of the $13/2^-$ states. Note that the opposite signs of the gyromagnetic factors of protons and neutrons lead to a coherent contribution when one is looking at the *difference* between the energies of the excited states. The remaining part of the MED of the $13/2^-$ states is thought to be due to established isospin-breaking sources. The wave functions of all other states in Fig. 20.5 involve mixtures of proton and neutron excitations and are thus similar for both members of the mirror pair, which results in a cancellation of the MED [36].

Breaking of the isospin symmetry is also reflected in a mixing of different T states. It is expected to occur most strongly in the heaviest $N = Z$ nuclei. An understanding of the mechanism of isospin mixing is crucial to performing reliable corrections when deriving the coupling constant from, for example, the $\log(ft)$ values of superallowed Fermi beta decay. Such measurements can have a direct impact on the unitarity test of the Cabibbo–Kobayashi–Maskawa (CKM) matrix for fundamental interactions. One way to study the violation of isospin symmetry is through the observation of isospin-forbidden $E1$ transitions in even–even $N = Z$ nuclei. Another way is by comparing $E1$ transitions in mirror nuclei. In ^{35}Ar (see Fig. 20.5) the 1446 keV $E1$ branch from the 3197 keV $7/2^-$ state clearly dominates the 3197 keV $M2$ transition, while the corresponding 1400 keV $E1$ decay is essentially absent in ^{35}Cl . This state decays directly to the ground state through the strong 3163 keV $M2$ transition [36].

20.4.4 Exotic Particle Decay of Nuclei

Proton decay from ground states or low-lying isomeric states limits the existence of nuclear matter on the neutron-deficient side of the nuclidic chart. In fact, proton-rich nuclei offer a unique opportunity to study the nuclear system beyond the drip line. The protons are no longer bound, but because of the Coulomb and centrifugal barriers of the proton, even excited states may live long enough such that they can be studied by γ -ray or proton emission. The proton escapes the nucleus by quantum mechanical tunneling, which is a widespread phenomenon in the natural sciences. Examples range from biophysical transport via chemical bonding to electron tunneling in quantum transistors.

Ground-state proton decay competes with β^+ radiation and the (partial) half-life decreases rapidly with increasing decay energy, i.e., the more exotic the isotope the larger its proton decay branch will be, until the last proton cannot be held inside the barrier anymore. Typically, proton-unstable nuclei are produced in fusion–evaporation reactions, are then in-flight mass separated, and finally implanted in pixelized silicon detectors. A space–time

correlation between the implantation signal and the subsequent proton decay is applied to identify a proton-emitting nucleus or study its decay in detail [37, 38]. Most of the experimental studies are being performed at the national laboratories in Legnaro, Italy, and Argonne and Oak Ridge, USA. In parallel, recent years have seen significant steps in the theoretical description: decay rates of spherical and deformed proton emitters can be predicted reliably, and fine structures in the decay have proven to be sensitive probes of the wave function compositions of the nuclear states involved [37, 38]. The first evidence for two-proton radioactivity has been found in ^{45}Fe following projectile fragmentation of 75 and 650 MeV/nucleon ^{58}Ni beams at GANIL and GSI, respectively, in 2002 [39, 40].

During recent years, the new nuclear decay mode of discrete-energy prompt proton and α -particle emission from excited states has been established in nuclei near ^{56}Ni [41, 42]. An example of prompt proton decay is shown in Fig. 20.2. Unlike the ground-state proton emitters, the prompt particle emission competes with γ -rays instead of β^+ radiation. This places the timescale of the decays into the $10^{-12} - 10^{-15}$ s regime, and allows their study in “prompt” coincidence with preceding and subsequent γ -rays emitted from the parent and daughter nuclei, respectively. The prompt particle decays proceed from highly deformed or superdeformed initial states into spherical daughter states. This implies a drastic rearrangement of the nuclear mean field in the course of the decay. Hence, the decay mode may be viewed as a self-regulated two-dimensional quantum tunneling process, which is unique in Nature.

20.5 Outlook

An enormously rich variety of nuclear-structure phenomena has been uncovered in recent studies of atomic nuclei using stable beams and stable targets. Electrostatic accelerators have played a very important role in these studies. The field of study concerning nuclei at the limit of angular momentum has been enriched enormously. Owing to the development of very sensitive instrumentation and refined experimental techniques, it has also been possible to extend the field of studies to nuclei far from stability, still using stable beams and targets. Large parts of the EUROBALL and GAMMASPHERE experimental programs have been dedicated to the investigation of the structure of neutron-deficient nuclei through a large variety of dedicated ancillary detectors for reaction products.

In the case of nuclei not too far away from the valley of stability, we can continue to investigate them using high-intensity beams of stable ions in combination with very efficient detectors, but γ -ray detector arrays are now also starting operation at the existing radioactive-ion-beam facilities. RISING [43] is the next phase of the EUROBALL collaboration. It allows efficient γ -ray spectroscopy of exotic nuclei produced in fragmentation or induced fission of

relativistic heavy-ion beams, which are selected by the fragment separator (FRS) at GSI. Other highly efficient γ -ray arrays are EXOGAM [44] at Spiral I at GANIL, and MINIBALL [45] at REX-Isolde at CERN, Switzerland. These existing facilities already provide exciting nuclear-structure data, and they offer tantalizing glimpses of scientific opportunities in the years to come.

However, it is the extremes in the ratio N/Z of neutrons to protons which define the limits of existence for nuclear matter. This field of research in nuclear physics will be opened up by the second-generation radioactive-beam accelerators. There is presently an increased scientific interest in the physics that can be probed with such beams, and enormous worldwide activity is going on in the construction of several large radioactive-nuclear-beam facilities throughout Europe, the USA, and Japan (see, for example, [46–48]). These facilities will make possible a considerable enlargement of the field of nuclear spectroscopic studies. When the drip lines are approached, detailed information on the rp- and r-process nuclei will become available. Exploring these limits will enhance new information, not only about the fundamental properties of the nuclear many-body system, but also about astrophysical processes and the origin of the elements, and about fundamental symmetries.

References

1. R. J. Van de Graaff: Phys. Rev. **38**, 1919 (1931)
2. EUROBALL III, a European γ -ray Facility, eds. J. Gerl and R. M. Lieder, GSI Laboratory Report, Darmstadt (1992)
3. J. Simpson: Z. Phys. A **358**, 139 (1997)
4. Achievements with the EUROBALL Spectrometer, Scientific and Technical Activity Report, LNL-INFN (REP) 201/2004, 1997–2003, eds. W. Korten and S. Lunardi (2003)
5. GAMMASPHERE Proposal, a National γ Ray Facility, eds. M. A. Deleplanque and R. M. Diamond, Lawrence Berkeley National Laboratory Report (1988)
6. I.-Y. Lee: Nucl. Phys. A **520**, 641c (1990)
7. GAMMASPHERE, The Beginning...1993–1997, Science Highlights booklet, ed. M. A. Riley (1998); <http://www-gam-lbl.gov/>
8. AGATA, Technical Proposal for an Advanced Gamma Tracking Array for the European Gamma Spectroscopy Community, eds. J. Gerl and W. Korten, GSI, Darmstadt (2001)
9. M. A. Deleplanque et al.: Nucl. Instr. Meth. A **430**, 292 (1999)
10. D. Rudolph et al.: Phys. Rev. Lett. **82**, 3763 (1999)
11. A. V. Afanasjev, D. B. Fossan, G. J. Lane, I. Ragnarsson: Phys. Rep. **322**, 1 (1999)
12. P. Walker, G. Dracoulis: Nature **399**, 35 (1999)
13. F. Iachello, A. Arima: *The Interacting Boson Model* (Cambridge University Press, Cambridge 1987)
14. F. Iachello: Nucl. Phys. News **12**:3, 17 (2002)
15. S.M. Polikanov et al.: Sov. Phys. JETP **15**, 1016 (1962)
16. P. J. Twin et al.: Phys. Rev. Lett. **57**, 811 (1986)

17. B. P. Singh, R. B. Firestone, S. Y. F. Chu: *Table of Superdeformed Nuclear Bands and Fission Isomers*, second edition, Nucl. Data Sheets **97**, 241 (2002)
18. S. Åberg: Nature **417**, 499 (2002)
19. B. Herskind et al.: Proceedings of the International Conference on the Labyrinth in Nuclear Structure, July 2003, Crete, in press
20. A. Bohr, B. Mottelson: *Nuclear Structure*, Vol. II (Benjamin, Reading, MA 1975)
21. I. Hamamoto: Phys. Lett. B **193**, 399 (1987)
22. S. W. Odegård et al.: Phys. Rev. Lett. **86**, 5866 (2001)
23. G. B. Hagemann, I. Hamamoto: Nucl. Phys. News **13**:3, 20 (2003)
24. R. M. Clark, A. O. Machiavelli: Annu. Rev. Nucl. Part. Sci. **50**, 1 (2000)
25. S. Frauendorf: Rev. Mod. Phys. **680**, 463 (2001)
26. W. Satula et al.: Phys. Lett. B **407**, 10 (1997)
27. N. Märginean et al.: Phys. Rev. C **65**, 051303(R) (2002)
28. C. J. Lister et al.: Phys. Rev. C **42**, R1191 (1990)
29. *Proceedings of Pingst 2000 – Selected Topics on $N = Z$ Nuclei*, June 2000, Lund, Sweden, eds. D. Rudolph and M. Hellström (Bloms i Lund AB, Lund, 2000); <http://pingst2000.kosufy.lu.se/proceedings.asp>.
30. R. Schneider et al.: Z. Phys. A **348**, 241 (1994)
31. M. Lewitowicz et al.: Phys. Lett. B **332**, 20 (1994)
32. M. Górska et al.: Phys. Rev. Lett. **79**, 2415 (1997)
33. M. Lipoglavšek et al.: Phys. Lett. B **440**, 246 (1998)
34. M. A. Bentley et al.: In [29], pp. 222–231
35. A. Zuker et al.: Phys. Rev. Lett. **89**, 142502 (2002)
36. J. Ekman et al.: Phys. Rev. Lett. **92**, 132502 (2004)
37. *Proceedings of the First International Symposium on Proton-Emitting Nuclei*, Oak Ridge, TN, U.S.A., October 1999, ed. J.C. Batchelder, AIP Conference Proceedings **518** (2000)
38. *Proceedings of the Second International Symposium on Proton-Emitting Nuclei*, Legnaro, Italy, February 2003, eds. E. Maglione and F. Soramel, AIP Conference Proceedings **681** (2003)
39. J. Giovinnazzo et al.: Phys. Rev. Lett. **89**, 102501 (2002)
40. M. Pfützner et al.: Eur. Phys. J. A **14**, 279 (2002)
41. D. Rudolph et al.: Phys. Rev. Lett. **80**, 3018 (1998)
42. D. Rudolph et al.: Phys. Rev. Lett. **86**, 1450 (2001)
43. RISING letter of intent, “Gamma Spectroscopy with RISING at the FRS”, eds. G. de Angelis et al., GSI, Darmstadt (2001)
44. J. Simpson et al.: Heavy Ion Phys. **11**, 159 (2000)
45. P. Reiter et al.: Nucl. Phys. A **701**, 209c (2002)
46. EURISOL: Web page, <http://www.ganil.fr/eurisol/>
47. H.H. Gutbrod, K-D. Gross, W.F. Henning, V. Metag, An International Accelerator Facility for Beams of Ions and Antiprotons, Conceptual Design Report, GSI, Darmstadt (2001)
48. RIA: Web page, <http://www.phy.anl.gov/ria/>

21 Nuclear Reactions

L. Corradi

INFN, Laboratori Nazionali di Legnaro, Viale dell'Università 2, I-35020
Legnaro (Padova), Italy
`corradi@lnl.infn.it`

21.1 Introduction

This chapter deals with the field of reaction mechanisms studied at tandem accelerators and is divided into four sections. The first two (elastic and transfer reactions) concern peripheral-reaction studies, those in which nuclei enter into contact through the tails of the Coulomb + nuclear potential. The other two (fusion and fission) concern compound-nucleus reactions, in which nuclei fuse together, forming an evaporation residue or decaying through fission. Some representative results obtained during last few years from different laboratories are reported with the main idea of showing items having potentialities for future interesting studies at energies available at tandem accelerators. The chosen examples are often a question of the author's taste, and the omission of results from other laboratories is only due to limited space.

21.2 Elastic Scattering

Elastic scattering provides important information on the interaction potential between colliding nuclei, and a knowledge of its size and shape is the basis of the description of all scattering processes [1–3]. A wealth of effects show up in reactions between complex nuclei, since the nucleus can both refract and absorb the incident waves. This is reflected in the angular and energy dependence of the observables related to the elastic channel [4]. Examples are given below of recent results obtained with relatively light ions and using radioactive beams with tandem accelerators.

21.2.1 Rainbow Scattering

In heavy-ion collisions, the absorption of flux from the initial (elastic) channel and the corresponding feeding of other reaction channels are usually very strong. In other words, if the impact parameter is smaller than a critical value, the corresponding partial-wave contribution to the elastic scattering amplitude is strongly damped. In relatively light systems, especially with

closed- or semiclosed-shell nuclei, absorption is weaker, and both diffractive and refractive effects in elastic scattering can be observed [5, 6]. Following studies performed at energies ≥ 20 MeV/A [7, 8], new measurements at 5–10 MeV/A have recently been performed in a broad angular range, leading to observation of specific oscillating structures interpreted as a nuclear rainbow [9, 10]. Figure 21.1 shows representative cross sections for $^{12}\text{C} + ^{12}\text{C}$, $^{16}\text{O} + ^{16}\text{O}$ and $^{16}\text{O} + ^{12}\text{C}$ systems measured at the VIVITRON tandem in Strasbourg with a high-resolution magnetic spectrometer (Q3D type) and position-sensitive silicon detectors in kinematic coincidence.

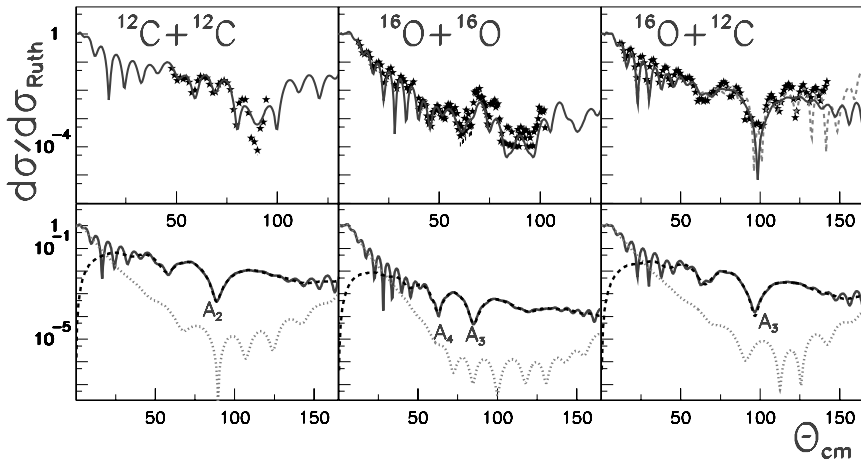


Fig. 21.1. Elastic angular distributions for $^{12}\text{C} + ^{12}\text{C}$ scattering at $E_L = 102$ MeV (left), $^{16}\text{O} + ^{16}\text{O}$ at $E_L = 124$ MeV (middle) and $^{16}\text{O} + ^{12}\text{C}$ at $E_L = 116$ MeV (right). Stars are experimental data and lines are calculations. The *full* curves are optical-model fits, while the *dotted* and *dashed* curves correspond to the near-side and far-side contributions, respectively (see text). The *dashed* curve for the $^{16}\text{O} + ^{12}\text{C}$ case includes the exchange of an α particle between the cores (data adapted from [9, 10])

The cross sections extend over six orders of magnitude in yield. The oscillations at forward angles can be interpreted as being due to diffractive effects [5]. The new feature is the observation of broad structures at large angles which are attributed to an Airy interference pattern, namely a nuclear rainbow phenomenon. In semiclassical language, different trajectories contribute to the total scattering amplitude at the same physical angle. In the near-side–far-side decomposition model [4], one considers trajectories originating from different sides of the scattering potential, and the interference between them leads to Fraunhofer diffraction at forward angles. The interference between two far-side components (one more peripheral, the other more internal) leads instead to rainbow scattering. These refractory effects

are visible as Airy minima in the angular distributions (components A_2 , A_3 and A_4 in Fig. 21.1). The appearance of rainbow scattering in experimental data is possible when the nuclear potential is strong enough (deep real part) to deflect particles into negative angles and when absorption is incomplete (shallow imaginary part). These studies then allow one to investigate the optical potential at small interaction radii.

21.2.2 Elastic Scattering Studied with Radioactive Beams

The study of properties of nuclei far from stability is a research area of highest priority nowadays [11–13]. In general, reaction mechanism studies with radioactive beams are feasible at complex accelerator facilities, where tandem accelerators are useful in different ways. The appropriate material can be directly provided at the ion source. Alternatively, nuclei of interest can be produced at the ion source via nuclear reactions with ions from another accelerator used as an injector, or one can use secondary reactions in appropriate targets.

An example of ongoing experiments is in the $N = Z$ region [12] (proton-rich nuclei), where the neutron–proton pairing interaction [14, 15] or even α -cluster properties are expected to play a significant role [16]. In the elastic scattering of α particles on light nuclei up to ^{40}Ca , structures and/or enhanced cross sections are observed at backward angles. Since the effect is particularly pronounced for $N = Z$ nuclei, it is of interest to investigate how these properties develop with radioactive proton-rich nuclei heavier than ^{40}Ca . Figure 21.2 shows the results of a measurement of elastic (+ inelastic) scattering for $\alpha + ^{44}\text{Ti}$ performed at Argonne, where a radioactive ^{44}Ti beam was accelerated on to a ^4He gas target [17]. Elastic scattering and inelastic scattering to the first 2^+ level of ^{44}Ti were measured in a wide angular range, by detecting α particles in solid-state strip detectors. The data have been compared with calculations performed within the framework of the distorted-wave Born approximation [1] using standard optical-potential parameters. It is presently not understood why the cross section enhancements at backward angles are not as strong as for $A \leq 40$ nuclei.

Another interesting area is that of weakly bound nuclei [11]. In these nuclei the last nucleons have very small binding energies (i.e. even less than 1 MeV), and the way in which scattering processes occur may differ strongly from that for stable nuclei. For instance, the nucleus can easily break up, affecting both the elastic and the other reaction channels. At Oak Ridge [18], the 25 MV tandem accelerator was used to deliver a ^{17}F beam with intensities similar to the Argonne experiment. The beam was produced in $^{16}\text{O}(\text{d}, \text{n})^{17}\text{F}$ reactions with 44 MeV deuterons provided by a cyclotron, extracted from the ion source, mass analyzed and injected into the tandem. Elastic scattering was measured in the $^{17}\text{F} + ^{208}\text{Pb}$ reaction, with projectile-like nuclei detected in silicon detector arrays. In Fig. 21.3, the experimental elastic angular distribution is shown together with coupled-channel calculations. The data show

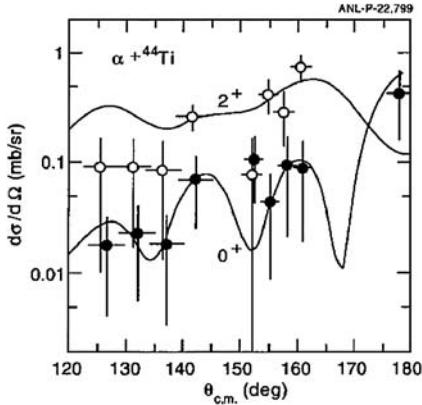


Fig. 21.2. Differential cross sections for elastic scattering of $\alpha + {}^{44}\text{Ti}$ (*solid points*) and inelastic excitation (*open points*) of the first 2^+ state in ${}^{44}\text{Ti}$. The lines are optical-model calculations (Reprinted from [17], copyright 2002, with permission from APS)

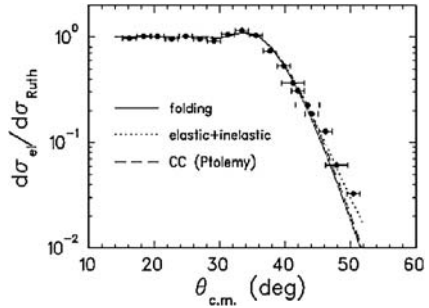


Fig. 21.3. Elastic-over-Rutherford angular distribution for the reaction ${}^{17}\text{F} + {}^{208}\text{Pb}$ at $E_L = 170$ MeV. The points are the data; the lines are coupled-channel calculations (Reprinted from [18], copyright 2002, with permission from APS)

some features of elastic scattering with heavy ions, namely the elastic scattering is close to the Rutherford scattering up to some angular range where interference between different partial waves produces an oscillating structure, and then it decreases as absorption into open reaction channels takes place. Studies of this kind allow one to investigate how the properties of unstable nuclei influence the interaction potential, for instance its absorption and transparency characteristics. This is even more interesting for neutron-rich nuclei, with their extended radius [19].

21.3 Transfer Reactions

Transfer reactions between heavy ions at energies close to the Coulomb barrier provide invaluable information for both nuclear-structure [2, 14] and reaction-dynamical studies [20, 21]. In single-particle transfer, the populations of specific excited states of the final nuclei give the possibility of deducing shell properties through the determination of the energy levels and their strength. In the transfer of two nucleons, one can investigate effects related to nucleon correlations, pairing in particular [15]. In multinucleon transfer, one can study how the various degrees of freedom, namely single-nucleon, pair or even cluster transfer modes, evolve as a function of the number of transferred nucleons [22, 23].

21.3.1 Studies Performed with Magnetic Spectrometers

In measurements using magnetic [24] and time-of-flight [25] spectrometers one could unambiguously identify by mass and charge the nuclei produced in transfer reactions up to the pickup of six neutrons and the stripping of six protons. Figure 21.4 shows a mass spectrum obtained in the reaction $^{58}\text{Ni} + ^{124}\text{Sn}$ studied at Argonne [24]. High-quality data of this kind, with good mass, charge and Q -value resolution, give possibilities of properly investigating different aspects of transfer reactions, such as the interplay between single-particle and pair transfer modes. The spectrum displays a characteristic behavior of the various inclusive data obtained so far, namely quite a regular drop of the yield for pure neutron transfer channels, indicating that neutron transfer is likely dominated by a sequential process in which nucleons are transferred independently. Cross section enhancements for channels corresponding to an even number of transferred nucleons may show up owing to the onset of pair modes [15]. Their evidence requires a careful analysis; for instance it is not clear if the pair transfer proceeds in a cluster- or sequential-like picture.

In data collected with the tandem and booster system in Legnaro (INFN) with a time-of-flight spectrometer [26], we measured for the $^{58}\text{Ni} + ^{208}\text{Pb}$ system the final mass and charge yields of light reaction products, differential

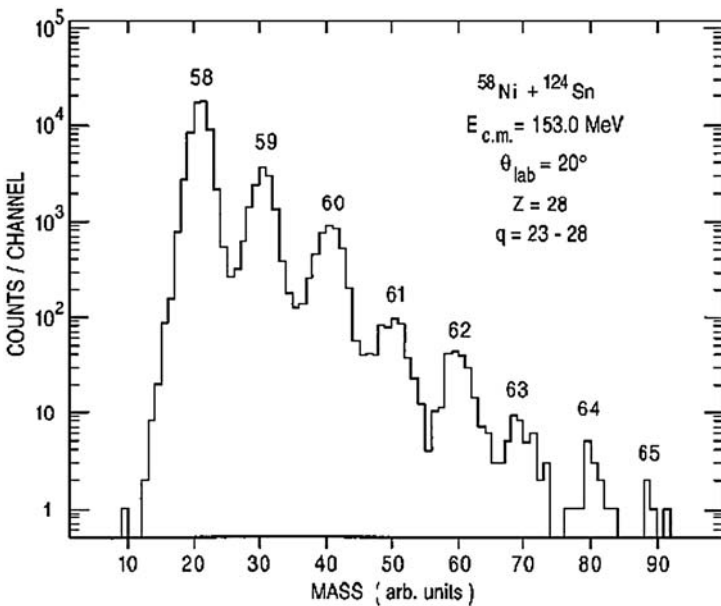


Fig. 21.4. Mass spectrum for transfer channels obtained in the reaction $^{58}\text{Ni} + ^{208}\text{Pb}$ with a high-resolution magnetic spectrometer (Reprinted from [24], copyright 1998, with permission from APS)

and total cross sections, and total kinetic-energy losses. The experimental total cross sections are reported in Fig. 21.5 and are compared with calculations performed within a semiclassical model [27], which has been successfully used in the comparison with experimental data for various systems (see [26] and references therein). The model calculates all the transitions from an initial to a final single-particle state (belonging to different nuclei), summing over all the allowed angular-momentum transfer and considering the occupation probability of the orbitals.

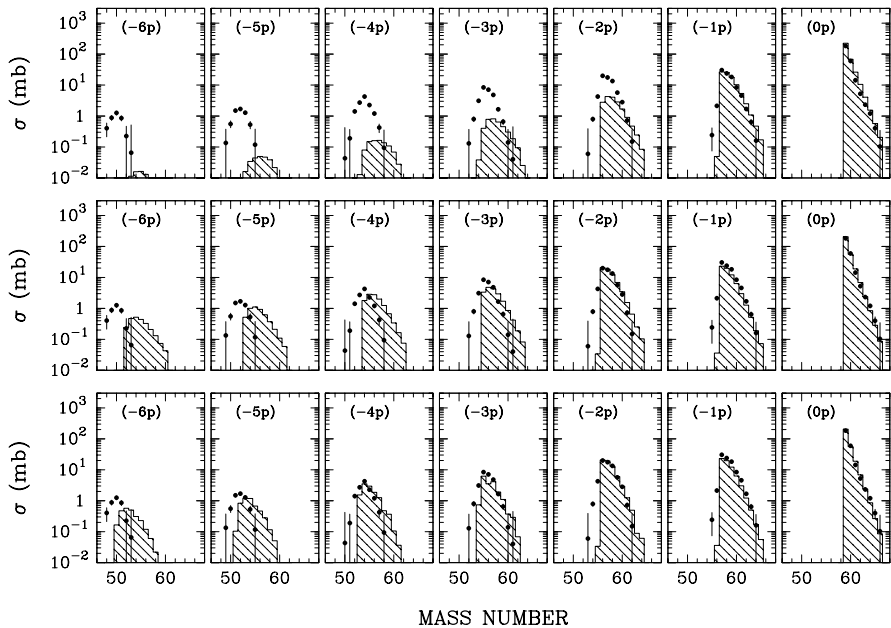


Fig. 21.5. Total cross sections integrated over angle and Q -value for the transfer channels observed in the reaction $^{58}\text{Ni} + ^{208}\text{Pb}$ at $E_L = 328.4$ MeV. The points and histograms are the experimental and theoretical values, respectively. The calculations shown in the *top* row include only single-nucleon transfer modes; those in the *middle* row have in addition a proton pair mode; and in the *bottom* row, also evaporation effects are taken into account (Reprinted from [26], copyright 2002, with permission from APS)

In the top row of the figure, the total cross sections are compared with calculations where nucleon transfer is treated in a successive approximation considering all the transitions as independent. Good agreement between data and theory is obtained for the transfer of pure neutrons (0p) and for channels involving the stripping of one proton (−1p). As more protons are transferred, the calculations are not able to follow the trend of the data, which tend to develop a population on the light-isotope side. The discrepancies indicate that

degrees of freedom beyond single-particle transfer modes have to be incorporated into the theory, or that more complex processes, such as deep inelastic components, play an important role. A new degree of freedom has then been added, namely the transfer of a pair of protons, and this is incorporated into the model by using a macroscopic form factor [28]. With a strength adjusted to reproduce the pure $-2p$ channel, the predictions for the other charge transfer channels are also much better (middle row). The treatment of the proton pair mode is, at present, only at a phenomenological level and it is still a challenge to relate it microscopically to the pair correlations in nuclei.

In the bottom row of the figure, the further effect of nucleon evaporation from the primary transfer products generated by large energy loss is taken into account, bringing the theoretical mass distributions even closer to the data. This is an important fact since a smooth transition from the quasi-elastic to the deep inelastic regime takes place after several nucleons are transferred, and this transition regime, where shell effects still play a significant role in the dynamics, is poorly understood in detail.

21.3.2 Studies with New-Generation Instruments

New spectrometers have been very recently constructed with solid angles more than an order of magnitude larger than the conventional ones. With such devices, one can greatly extend the studies of transfer reactions. First, one can investigate in detail the transition regime mentioned above. At least in some selected cases, one can distinguish the populations of specific excited states of transfer products and study how nucleon correlation behaves in reactions with heavy ions. The pairing interaction acts in such a way as to redistribute the flux among specific final states, for example 0^+ states connected with zero transferred angular momentum. This was observed years ago in studies performed with very light ions, namely of (p, t) or (t, p) type [29]. With heavy ions, one has the possibility to study the behavior of these modes (called pair-rotational or pair-vibrational [15], depending on the shell properties of the nuclei involved) as the number of nucleons increases.

Multinucleon transfer reactions may also constitute a valuable tool, at least in specific mass regions (i.e. actinides and transactinides), to produce neutron-rich isotopes [13]. Spectroscopic studies of neutron-rich nuclei produced via multinucleon transfer are an area of increasing interest thanks to the development of large γ arrays [30]. The study of the lowest excited levels of these neutron-rich nuclei gives an indication of changes in shell structure and of the modification of nucleon correlation. The coupling of a large-solid-angle magnetic spectrometer with a large γ array is therefore an ideal instrument to make spectroscopic measurements, and this is presently being pursued at INFN-Legnaro with the PRISMA setup [31]. In Fig. 21.6, an example is given of a mass spectrum obtained in the reaction $^{54}\text{Cr} + ^{124}\text{Sn}$ as a projection of a bidimensional matrix of time of flight vs. position at the focal plane of the spectrometer. One observes peaks corresponding to the population of

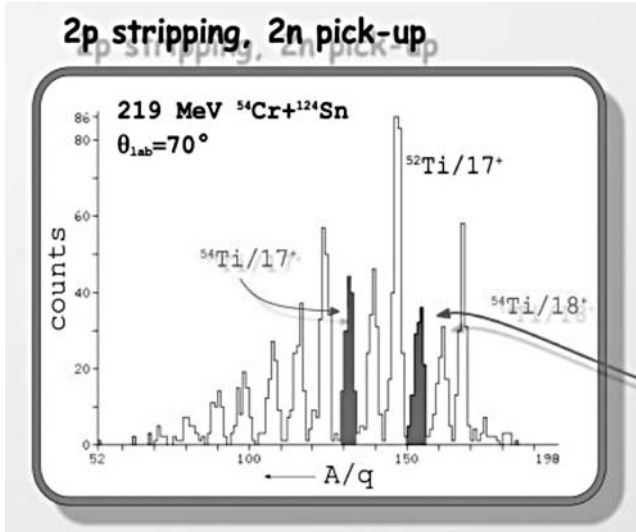


Fig. 21.6. Mass spectrum of transfer products obtained in $^{54}\text{Cr} + ^{124}\text{Sn}$ at $E_L = 219$ MeV with the spectrometer PRISMA. The *shaded* peaks correspond to the same nucleus but with different atomic charge states

neutron-rich ^{54}Ti ($-2p + 2n$ transfer channel). The nuclei are produced with different atomic charge states, whose distribution is seen in the spectrum.

With stable beams and at the energies and intensities typical of tandem accelerators, one can presently reach regions moderately far from stability (on average $\simeq 3\text{--}5$ nucleons from the last stable isotope), but one can investigate nuclei through the entire nuclear chart, provided suitable projectiles and targets are chosen. Optimum- Q -value arguments show how the population of nuclei evolves via multinucleon transfer from proton stripping and neutron pickup with neutron-poor projectiles to proton pickup and neutron stripping with neutron-rich projectiles [32]. In this way, a significant cross section of neutron-rich heavy targets is predicted, which makes this reaction mechanism a promising one, especially with radioactive beams.

21.4 Near-Barrier Fusion

Sub-barrier fusion reactions between heavy ions are characterized by large enhancements in the cross sections with respect to calculations based on one-dimensional barrier penetration models. These enhancements are due to the coupling of different degrees of freedom acting in the tunneling process, mainly static deformations and surface vibrations of nuclei [33–35], making fusion reactions a nice example of interplay between nuclear structure and reaction mechanisms [36, 37]. Recently, detailed studies have started to be

performed with weakly bound nuclei [38–40] and with very heavy ions [41] to investigate the competition of fusion with other reaction channels at energies both below and above the Coulomb barrier [42, 43].

21.4.1 Precision Measurements of Fusion Excitation Functions

Within certain approximations in solving the coupled equations of the scattering process, it can be demonstrated [33] that the tunneling process can be seen as acting through a series of different barriers, each of which is related to the coupling to a particular channel with some weight. Developing this concept of “barrier distributions”, it was suggested [44] that a structure in the fusion excitation function would become visible if measurements of very high precision were made. The theoretical ideas were soon taken up successfully by an Australian group [45], who first showed how one could get the characteristic distributions expected from coupling to deformations. Information on the effects of the different barriers can be derived through the second derivative of the fusion excitation functions. Precision studies of fusion cross sections on systems having different nuclear-structure properties may reveal peculiar distributions associated with coupling to specific reaction channels [37]. A representative example of the extracted barrier distribution for the system $^{16}\text{O} + ^{144}\text{Sm}$ [46], compared with coupled-channel calculations, is shown in Fig. 21.7. The double-peak structure in the second derivative of the excitation

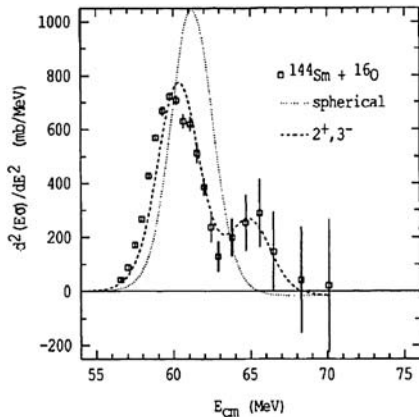


Fig. 21.7. Fusion barrier distributions for $^{16}\text{O} + ^{144}\text{Sm}$. The points are data, and the lines are calculations. The *dotted* line corresponds to a single barrier (no coupling); the *dashed* line includes the coupling to the 2^+ and 3^- states in ^{144}Sm (from [46])

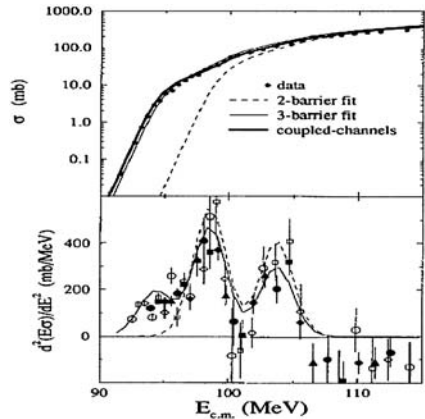


Fig. 21.8. Fusion excitation function (*top*) and barrier distributions (*bottom*) for $^{58}\text{Ni} + ^{60}\text{Ni}$. The points are experimental data, and the lines are calculations (Reprinted from [47], copyright 1995, with permission from APS)

function can be interpreted as an effect of barrier splitting due to coupling of the ground state to the lowest 2^+ and 3^- states of the spherical nucleus ^{144}Sm , where the 3^- contribution dominates. In the no-coupling limit (dotted line), the barrier distribution shows only a single peak.

Coupling to states belonging to rotational bands gives rise to an almost continuous distribution of barriers. Distinguished multipeak structures in vibrational nuclei, on the other hand, can be associated with the effect of single-phonon and multiphonon excitations. The fusion cross sections and the barrier distribution for the $^{58}\text{Ni} + ^{60}\text{Ni}$ system [47], where the nuclei have a similar vibrational structure, are displayed in Fig. 21.8. The data can be reproduced by invoking the effect of three main barriers. A coupled-channel calculation shows that these barriers can be generated by including multiphonon excitations, namely multiple quadrupole (2^+) phonons in each nucleus and mutual excitations.

21.4.2 Fusion and Complex Reaction Channels

A presently highly debated item is whether and how transfer channels affect the fusion probability. This is an important subject, since in transfer channels nucleons move from one nucleus to the other and therefore these channels can be considered as first steps in the path eventually leading to fusion. It was suggested that transfer channels with positive ground-state Q -values (Q_{gs}) increase the fusion probability below the Coulomb barrier owing to coupling effects [33]. Different shapes for the barrier distributions were predicted [48] depending on the simultaneous or sequential character of the transfer channels and on the Q -values involved. Disentangling the influence of transfer on fusion from that of the other competing channels requires a careful “control” of nuclear-structure effects. A proper choice of projectile and target is therefore needed. An example is given by a study of $^{40}\text{Ca} + ^{90,96}\text{Zr}$ systems, whose results are shown in Fig. 21.9 [49, 50]. In this case all neutron transfer channels for ^{90}Zr have negative Q_{gs} , while those for ^{96}Zr are positive. One observes a dramatic difference between the fusion cross sections and barrier distributions for the two cases. The ^{90}Zr case has a multipeak structure, while the ^{96}Zr case develops a broad distribution. Although this different behavior can be qualitatively attributed to the effect of transfer channels in ^{96}Zr , coupled-channel calculations reproduce both cases by including phonon excitations only [50]. It is important to say, that the same calculations reproduce the measured transfer cross sections for both systems.

Another highly disputed subject, both experimentally [38, 39] and theoretically [40], is whether fusion cross sections are enhanced or suppressed when weakly bound nuclei interact. As mentioned in Sect. 2.2, peculiar properties of scattering processes are expected owing to the larger radii and lower binding energies of valence nucleons. On one side, fusion can be enhanced owing to the extended matter radii, which lower the barrier, and owing to the coupling

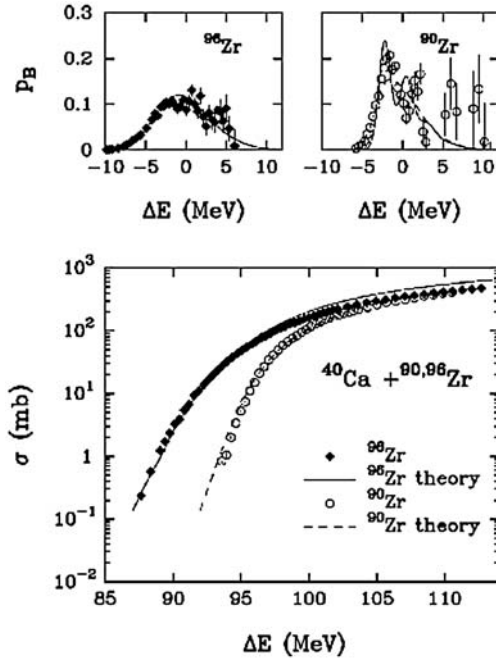


Fig. 21.9. Fusion cross sections (*bottom*) and barrier distributions (*top*) for $^{40}\text{Ca} + ^{90,96}\text{Zr}$. The points are experimental data [49], and the lines are calculations [50] including the first 2^+ and 3^- levels of both nuclei (Reprinted from [49], copyright 2000, with permission from APS)

to the continuum states, which gives rise, for example, to a low-lying soft dipole mode. On the other side, breakup channels, namely the separation of a nucleus into two (or more) pieces owing to the Coulomb and nuclear fields, may even suppress fusion owing to the depletion of the incoming flux. Various experiments are being performed at radioactive-beam facilities, but very significant results have been obtained also with tandem accelerators. The $^9\text{Be} + ^{208}\text{Pb}$ system was studied at Canberra [39] by measuring with high precision both evaporation residue cross sections (detecting α decays from the compounds) and fission products (via a couple of large-area position-sensitive parallel-plate detectors). Figure 21.10 shows the results. The complete fusion cross sections (filled circles) are the sum of the evaporation residue and fission, this last channel being the dominant one. Coupled-channel calculations reproduce the complete fusion cross sections and the barrier distributions only by applying a scaling factor, supporting the idea that fusion is suppressed. Experimentally, α -decay channels belonging to nuclei generated in incomplete fusion have been observed, consistent with the opening of breakup channels during the approaching phase of the interacting nuclei.

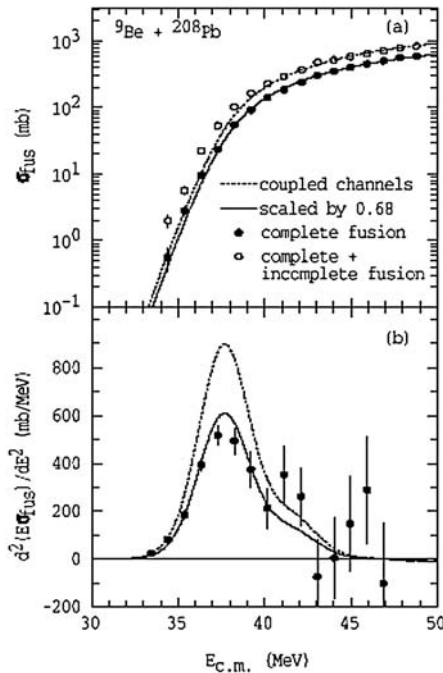


Fig. 21.10. Experimental cross sections for complete fusion (*filled circles*) and the sum of complete + incomplete fusion (*hollow circles*), and barrier distributions for ${}^9\text{Be} + {}^{208}\text{Pb}$. The lines are calculations (Reprinted from [39], copyright 1999, with permission from APS)

21.5 Fission Phenomena

Fission is one of the most investigated areas in nuclear physics [51,52]. In recent studies one could identify, using neutron-induced fission or by detecting spontaneously fissioning nuclei, ternary and quaternary fission modes [16]. At relativistic energies, people have started to investigate [53] the mass and charge yields of nuclei produced via photofission. At tandem accelerators, some examples of present studies concern the multimodal decay of fissioning nuclei [16] and the competition between fusion–fission and quasi-fission processes [43].

21.5.1 Entrance Channel Effects in the Fission Process

In heavy-ion reactions, passing over the potential barrier does not guarantee the formation of a compound nucleus and eventually an evaporation residue. Both excitation energy and angular momentum force the system to reseparate into two fission fragments. Before reseparation, the system may or may not have gone through the equilibration of all the degrees of freedom, leading

to what are named fusion–fission (FF) and quasi-fission (QF) processes, respectively [54]. Detailed investigations of the interplay between FF and QF are presently being performed to study the influence of shell properties on the decay pattern, and the mechanism of fusion hindrance at or above the Coulomb barrier, which has important consequences also for the formation of superheavy elements [55, 56]. The results of an experiment [57] in which fission fragments have been studied for the systems $^{12}\text{C} + ^{204}\text{Pb}$ and $^{48}\text{Ca} + ^{168}\text{Er}$ are shown in Fig. 21.11. The two reactions were chosen to produce the same compound nucleus (^{216}Ra) at about the same excitation energy. In this way, i.e. by making a direct comparison for cases leading to the same compound starting from different mass asymmetries in the entrance channel, one can study how QF and FF processes compete. The fission fragments were detected in kinematic coincidence with a two-arm time-of-flight spectrometer. From the coincidence one can extract distributions of the mass (M) and total kinetic energy (TKE). In the $^{12}\text{C} + ^{204}\text{Pb}$ case the M – TKE matrix (panel a) has a near-triangular shape and a near-Gaussian mass projection (panel b), as expected from symmetric fission centered at masses $A \simeq (A_p + A_t)/2$. This is what is predicted by liquid-drop and diffusion models, without the influence of shell properties [51]. These models also predict well the shape of the TKE distribution (full lines). The small peaks in the variances of TKE (panel d) are expected from the asymmetric decay mode of FF. The behavior of the same quantities for $^{48}\text{Ca} + ^{168}\text{Er}$ is very different. In panel (e) one sees, besides the events whose masses are close to those of the projectile and targets (coming from deep inelastic collisions), a wider mass and TKE distribution. Wings in the mass distribution are clearly visible. By subtracting from the total mass the part corresponding to the central Gaussian component (symmetric fission), one obtains the full circles in panel (f). This and other characteristics of TKE are attributed to the QF process, which competes with FF with a strong dependence on the mass asymmetry.

The data are qualitatively described in calculations using the concept of a dinuclear system, in which the potential-energy surfaces (PESs) generated by including all terms of interaction between nuclei are modulated by specific shell properties. The mass asymmetry dependence of the PES should also play a very important role. In asymmetric collisions, the PES favors absorption of the light nucleus into the heavy one, leading to a compact evaporation residue. In more symmetric collisions, the PES favors a mass drift from the heavy to the light partner, leading to a more elongated configuration and a higher probability for QF. These concepts have been applied in recent measurements of fusion cross sections for different systems, where fusion inhibition has been shown to take place owing to the effect of QF [42].

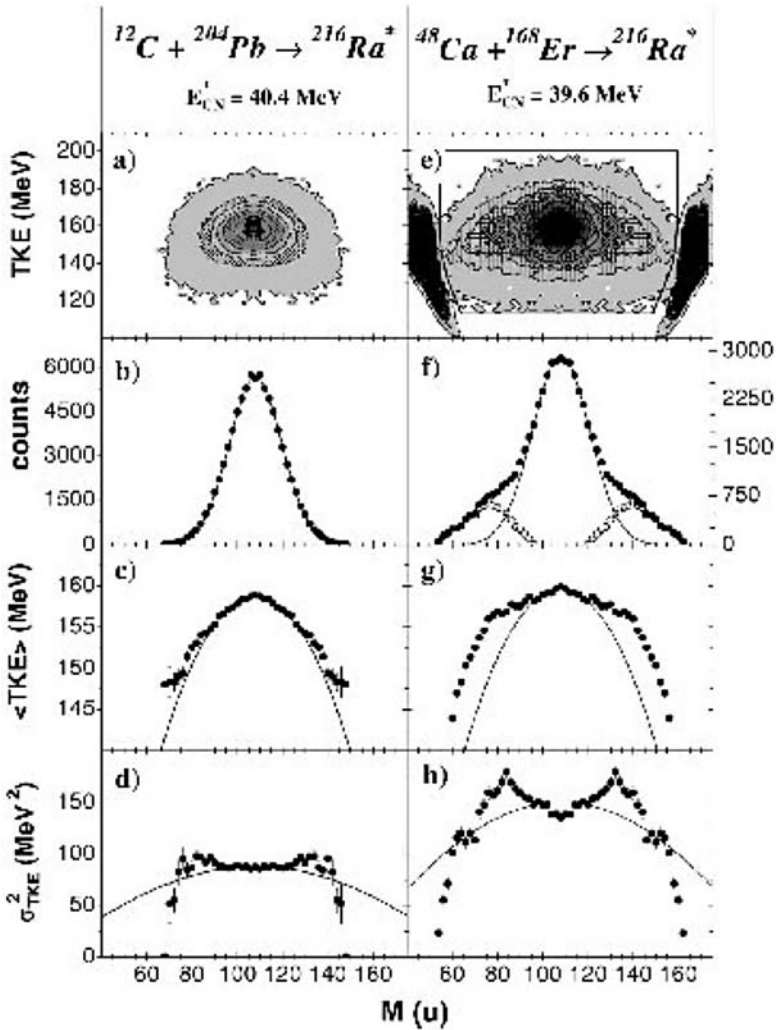


Fig. 21.11. Mass-energy distributions of correlated fission fragments in $^{12}\text{C} + ^{204}\text{Pb}$ and $^{48}\text{Ca} + ^{168}\text{Er}$ as a function of fragment mass M . Panels (a) and (e) are two-dimensional plots of the total kinetic energy (TKE) vs. M ; (b), (f) and (c), (g) are the projections on the mass and TKE axes, respectively; (d) and (h) are the variances. Panels (f), (g) and (h) include only events inside the closed line in (e). The lines and open circles are described in the text (Reprinted from [57], copyright 2003, with permission from APS)

References

1. G.R. Satchler: *Direct Nuclear Reactions* (Clarendon, Oxford 1983)
2. R.A. Broglia, A. Winther: *Heavy Ion Reactions*, Lecture Note Series Vol. 84, (Addison-Wesley, Redwood City, California (USA) 1991)
3. P. Ring, P. Schuck: *The Nuclear Many-Body Problem* (Springer, New York 1980)
4. G.R. Satchler: Phys. Rep. **199**, 148 (1991)
5. M.E. Brandan, G.R. Satchler: Phys. Rep. **285**, 143 (1997)
6. R.R. Betts, A.H. Wuosmaa: Rep. Prog. Phys. **60**, 819 (1997)
7. D. T. Khoa, W. von Oertzen, H.G. Bohlen, F. Nuoffer: Nucl. Phys. A **672**, 387 (2000)
8. A.A. Ogloblin: Phys. Rev. C **62**, 044601 (2000)
9. M.P. Nicoli et al.: Phys. Rev. C **61**, 034609 (2000)
10. S. Szilner et al.: Phys. Rev. C **64**, 064614 (2001)
11. P.G. Hansen, A.S. Jensen, B. Jonson: Annu. Rev. Nucl. Part. Sci. **45**, 591 (1995)
12. W. Nazarewicz: Nucl. Phys. A **654**, 195c (1999)
13. H.L. Ravn, J. Lettry, T. Nilsson (eds.): *Proceedings of the Fifth Int. Conf. on Radioactive Nuclear Beams*, Divonne, France, 3–8 April, 2000, Nucl. Phys. A **701** (2002)
14. A. Bohr, B. R. Mottelson: *Nuclear Structure* (World Scientific, Singapore 1998), Vols.I and II
15. R.A. Broglia, O. Hansen, C. Riedel: In *Advances in Nuclear Physics*, ed. by M. Baranger and E. Vogt (Plenum, New York 1973), Vol. 6, p. 297
16. R. Jolos, W. Scheid (eds.): *Proceedings of the Int. Symposium on Nuclear Clusters: From Light Exotic to Superheavy Nuclei*, Rauischolzhausen, Germany, 5–9 August 2002 (EP Systema, Debrecen, Hungary 2003)
17. K.E. Rehm et al.: Phys. Rev. Lett. **89**, 132501 (2002)
18. J.F. Liang et al.: Phys. Rev. C **65**, 051603(R) (2002)
19. I. Tanihata et al.: Phys. Lett. B **206**, 592 (1988)
20. P.E. Hodgson: *Nuclear Reactions and Nuclear Structure* (Clarendon, Oxford 1971)
21. L.J.B. Goldfarb, W. von Oertzen: in *Heavy Ion Collisions*, ed. by R. Bock (North-Holland, Amsterdam 1979), p. 215
22. K.E. Rehm: Annu. Rev. Nucl. Part. Sci. **41**, 429 (1991)
23. W. von Oertzen, A. Vitturi: Rep. Prog. Phys. **64**, 1246 (2001)
24. C.L. Jiang et al.: Phys. Lett. B **337**, 59 (1994); Phys. Rev. C **57**, 2393 (1998)
25. L. Corradi et al.: Phys. Rev. C **59**, 261 (1999); Phys. Rev. C **63**, 021601(R) (2001)
26. L. Corradi et al.: Phys. Rev. C **66**, 024606 (2002)
27. E. Vigezzi, A. Winther: Ann. Phys. (N.Y.) **192**, 432 (1989)
28. C.H. Dasso, G. Pollaro: Phys. Lett. B **155**, 223 (1985)
29. M. Igarashi, K. Kubo, K. Yagi: Phys. Rep. **199**, 1 (1991)
30. C.W. Beausang, J. Simpson: J. Phys. G **22** 527 (1996)
31. A.M. Stefanini et al.: Nucl. Phys. A **701**, 217c (2002)
32. C.H. Dasso, G. Pollaro, A. Winther: Phys. Rev. Lett. **73**, 1907 (1994)
33. C.H. Dasso, S. Landowne, A. Winther: Nucl. Phys. A **405**, 381 (1983); Nucl. Phys. A **407**, 221 (1983)

34. S.G. Steadman, M.J. Rhoades Brown: *Annu. Rev. Nucl. Part. Sci.* **36**, 649 (1986)
35. M. Beckerman: *Rep. Prog. Phys.* **51**, 1047 (1988)
36. A.B. Balantekin, N. Takigawa: *Rev. Mod. Phys.* **70**, 77 (1998)
37. M. Dasgupta et al.: *Annu. Rev. Nucl. Part. Sci.* **48**, 401 (1998)
38. K.E. Rehm et al.: *Phys. Rev. Lett.* **81**, 3341 (1988)
39. M. Dasgupta et al.: *Phys. Rev. Lett.* **82**, 1395 (1999)
40. K. Hagino et al.: *Phys. Rev. C* **61**, 037602 (2000)
41. K. Nishio et al.: *Phys. Rev. C* **62**, 014602 (2000)
42. A.C. Berriman et al.: *Nature* **413**, 144 (2001)
43. D.J. Hinde, M. Dasgupta, A. Mukherjee: *Phys. Rev. Lett.* **89**, 282701 (2002)
44. N. Rowley, G.R. Satchler, P.H. Stelson: *Phys. Lett. B* **254**, 25 (1991)
45. J.X. Wei et al.: *Phys. Rev. Lett.* **67**, 3368 (1991)
46. J.R. Leigh et al.: *Phys. Rev. C* **52**, 3151 (1995)
47. A.M. Stefanini et al.: *Phys. Rev. Lett.* **74**, 864 (1995)
48. N. Rowley, I.J. Thompson, M.A. Nagarajan: *Phys. Lett. B* **282**, 276 (1992)
49. H. Timmers et al.: *Phys. Lett. B* **399**, 35 (1997); *Nucl. Phys. A* **633**, 421 (1998)
50. G. Pollaro, A. Winther: *Phys. Rev. C* **62**, 054611 (2000)
51. R. Vandenbosh, J.R. Huizenga: *Nuclear Fission* (Academic Press, New York 1973)
52. Yu.Ts. Oganessian, Yu.A. Lazarov, G.T. Seaborg, W.D. Loveland: In *Treatise on Heavy-Ion Science*, Vol. 4, ed. by D.A. Bromley (Plenum, New York 1985), p. 3, p. 255
53. K.-H. Schmidt et al.: *Nucl. Phys. A* **665**, 221 (2000); *Nucl. Phys. A* **693**, 169 (2001)
54. M. Lefort: In *Heavy Ion Collisions*, ed. by R. Bock (North-Holland, Amsterdam 1979), p. 45
55. P. Armbruster et al.: *Phys. Rev. Lett.* **54**, 406 (1985)
56. Yu.Ts. Oganessian et al.: *Nature* **400**, 242 (1999)
57. A.Yu Chizhov et al.: *Phys. Rev. C* **67**, 011603(R) (2003)

22 Detection of Explosives and Other Threats Using Accelerator-Based Neutron Techniques

T. Gozani

Ancore Corp, an OSI Systems Company, 2950 Patrick Henry Dr., Santa Clara,
CA 95054, USA
tsahi@ancore.com

22.1 Introduction

Neutral particles, in particular neutrons and photons, are very suitable for probing deeply and nonintrusively into sealed objects. The main reasons are their good penetration, which makes available two- (and sometimes three-) dimensional radiographic images, and, very importantly in the case of neutrons, the elemental/material information they provide.

Nonintrusive inspection of objects of all sizes, from luggage to shipping containers and from postal parcels to trucks, is a vital component of any national security, from aviation to land and sea ports of entry. The paramount importance of these inspections is more obvious in the era after the 11th of September 2001, as the spectrum of threats is wider and the probability of occurrence more real.

To be effective, inspection technologies need to be sensitive, material-specific, rapid, flexible and automatic. Currently used inspection technologies such as x-ray radiography and vapor/trace detectors are severely deficient in this respect.

Neutron-based technologies, on the other hand, provide accurate, rapid and automatic detection of a wide array of threats: explosives, chemical agents, nuclear materials and devices, other hazardous materials, drugs, etc. The nuclear-based techniques achieve these feats through the production of characteristic elemental gamma rays from nuclear reactions, primarily thermal-neutron capture and/or inelastic scattering of fast neutrons. Small charged-particle accelerators, generating copious numbers of neutrons, are essential sources for these techniques.

22.2 Characterization of Threats

Each threat, for example explosives, chemical agents or fissile nuclear material, differs from benign materials in its chemical and hence elemental composition. This difference makes nuclear, and especially neutron-based, techniques a unique tool for detecting the presence of concealed threats amongst benign cargo.

Many of the threats are made of the basic four “organic” elements: hydrogen, carbon, nitrogen and oxygen. Additional elements found in significant amounts are, for example, chlorine in some explosives, drugs and blister chemical agents, while phosphorus, sulfur and fluorine are present in certain nerve agents. Common fissile nuclear materials are ^{235}U and ^{239}Pu . Examples of elemental compositions of various threats and common benign materials are provided in Table 22.1. The table shows that explosives are typically rich in nitrogen and oxygen and “poor” in hydrogen and carbon. In general, elemental features can be expressed by their absolute values, by ratios of elements or by other functional relationships that allow better material discrimination.

Table 22.1. Examples of elemental compositions of some threats and common benign materials (concentrations of the elements are given in weight %)

Threat	C	H	N	O	P	F	Cl	S	N/H	N/C
<i>Explosives</i>										
C4	21.9	3.6	34.4	40.1	0	0	0	0	10	2
TNT	37	2.2	18.5	42.3	0	0	0	0	8	1
PETN	19	2.4	17.7	60.8	0	0	0	0	7	1
AN (ammonium nitrate)	0	5	35	60	0	0	0	0	7	∞
<i>Chemical agents</i>										
Sarin	34.3	7.1	0	22.9	22.1	13.6	0	0	0	0
VX	49.5	9.7	5.2	12	11.6	0	0	12	1	0
CA (H-Cyanide)	44.5	3.7	51.8	0	0	0	0	0	14	1
HD (mustard gas)	30.2	5	0	0	0	0	44.6	0	0	0
Phosgene	12.1	0	0	16.2	0	0	71.7	0	NA	0
<i>Common benign</i>										
Water	0	11.1	0	88.9	0	0	0	0	0	0
Paper	44	6	0	50	0	0	0	0	0	0
Plastic	86	14	0	0	0	0	0	0	0	0
Salt	0	0	0	0	0	0	60	0	NA	NA

22.3 Nuclear Reactions and Signatures

Detection of elements relies on the ability of the neutrons to efficiently interact with them, that is, to have a high enough interaction cross section, and the fact that the interaction results in the emission of detectable gamma rays, and in a few cases neutrons.

The nuclear interactions most employed for inspection and real-time assay are neutron capture and inelastic scattering. The signature emission in the former is the gamma rays from deexcitation of the $A + 1$ nucleus created

in the reaction (A is the atomic mass). Inelastic neutron scattering leaves the scattering nucleus in an excited state, and its deexcitation leads to the emission of characteristic gamma rays.

The neutron capture cross section (for which the common symbol is $\sigma(n, \gamma)$) is in general (except for strong resonances in some elements) inversely proportional to the velocity (and hence the square root of the energy) of the neutron. Thus the lower the energy, or the more “thermalized” the neutrons are, the higher the cross section. In principle, all types of nuclei undergo thermal capture, though with vastly different cross sections. Oxygen is practically undetectable through capture (cross section 0.0002 barn; 1 barn = $1 \times 10^{-28} \text{ m}^2$), and carbon also has a very low cross section (0.0035 barn), while hydrogen, nitrogen and most metals are readily detectable through this process. The neutron cross sections for hydrogen and nitrogen are 0.33 and 0.07 barn, respectively. The cross sections for metals are in general higher. The branching ratios (i.e., the percentage of deexcitations leading to the detected gamma rays) are 100% and 14% for H (2.2 MeV) and N (10.8 MeV), respectively. A typical thermal-capture spectrum is shown in Fig. 22.1 (lower curves) and discussed there as part of the conventional

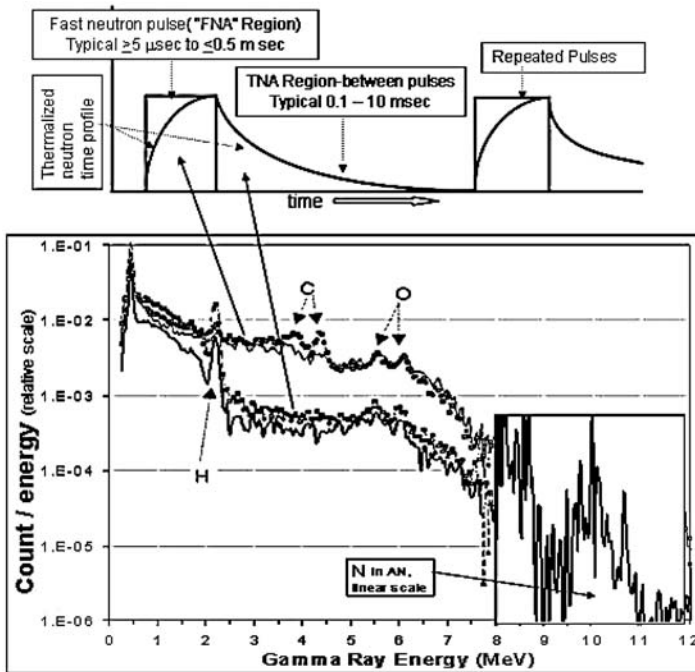


Fig. 22.1. Principles of conventional pulsed-neutron inspection (PNI), utilizing a combination of FNA and TNA techniques. The lower two curves represent the TNA spectra

pulsed-neutron technique. The main features in these spectra are the hydrogen line at 2.22 MeV and nitrogen line at 10.8 MeV (along with their single-escape peaks, which are 0.511 MeV lower).

The inelastic processes have thresholds below which the specific reaction cannot take place. These are approximately 2.6 MeV for nitrogen, 4.8 MeV for carbon and 6.4 MeV for oxygen. The cross sections for these interactions, symbolized by $\sigma(n, n'\gamma)$, are generally much lower than $\sigma(n, \gamma)$. They range from less than 10^{-3} to a few barns. Fortunately, the values of $\sigma(n, n'\gamma)$ for the organic elements are among the highest, between 0.05 and 0.5 barn. The behavior of these cross sections as a function of energy and the resulting gamma ray signatures (as detected by NaI scintillation detectors) are shown in Fig. 22.2.

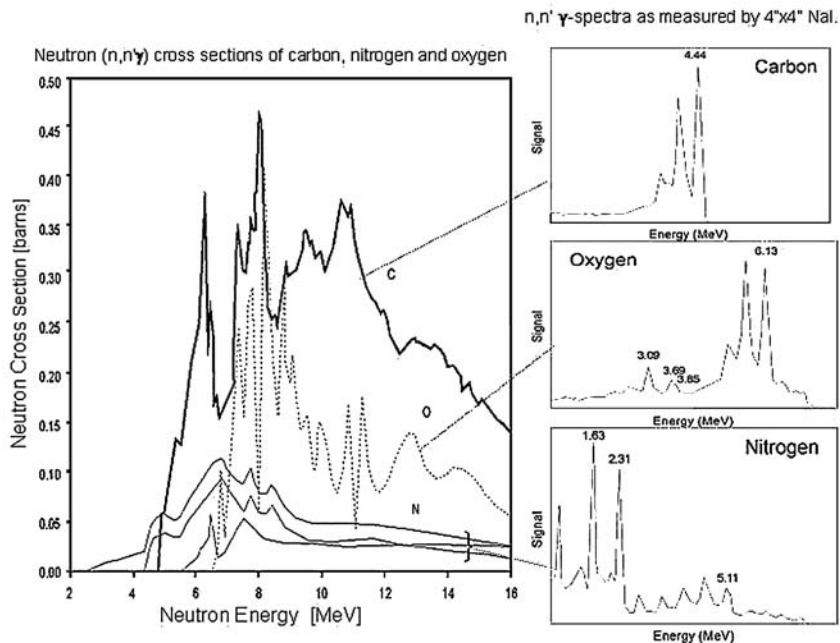


Fig. 22.2. Neutron inelastic cross sections in organic elements (carbon, nitrogen and oxygen) and the resulting gamma ray signatures, as detected by a $10 \times 10 \times 10 \text{ cm}^3$ sodium iodide detector

22.4 Accelerator-Based Neutron Sources

A wide variety of neutron sources can be used as part of nonintrusive material-specific interrogations. When the main nuclear process is neutron capture (n, γ), all accelerator-based sources are usable, as the neutrons are slowed down

in the system, and most of the reactions take place when they are fully thermalized. The selection of the source is then determined by the available intensity, total cost (and cost per neutron), energy of the neutrons (in general, the lower the better), etc. When fast-neutron reactions are the whole or part of the interrogation process, the energy, energy resolution and time structure, in addition to intensity, cost and size, are very critical in the selection process.

Table 22.2 lists possible accelerator neutron sources for interrogation systems. Four categories of sources are included: accelerator-based low-energy (<150 kV, which includes the important class of sealed-tube (d, T) neutron generators), accelerator-based medium-energy (0.5–2 MV), accelerator-based “high-energy” (>2 MV) and others (for example based on photoneutron production employing an electron linac).

Table 22.2. Accelerator-based sources for neutron-based inspection techniques

Nuclear Reaction	Strength (typical) (n/s)	Neutron Energy (MeV)	Neutron Spectrum	Pulsing Capabilities (pulse width)	Typical Repetition Rate	Possible Applications
<i>Accelerator-based: low-energy (<150 kV)</i>						
<i>Sealed tube</i>						
(d, D)	10^6 – 10^8	3.2	Narrow	$>\mu\text{s}$	DC–10 kHz	TNA/FNA, fissile detection
(d, T)	10^8 – 10^9	14	Narrow	$>\mu\text{s}$	DC–10 kHz	TNA, FNA, fissile detection
<i>Accelerator-based: medium energy (>500 kV–2 MV)</i>						
<i>Electrostatic, RFQ</i>						
(p, Li)	10^8 – 10^{10}	0.15–1	Narrow	DC–ns; μs	RFQ, 180 Hz	TNA/FNA fissile
(d, Be)	10^9 – 10^{11}	Variable	Broad	DC–ns; μs	RFQ, 180 Hz	PFNA, fissile
<i>Accelerator-based: “high”-energy (>2 MV)</i>						
<i>Electrostatic</i>						
(d, D)	10^9 – 10^{11}	Variable	Mono-energetic	ns	1–10 MHz	PFNA, fissile
<i>Electrostatic, cyclotron</i>						
(d, Be)	10^9 – 10^{12}	Variable	Broad	ns	1 MHz	NRA
<i>Others</i>						
<i>Medium-energy ($E < 10$ MeV) electron linac, using bremsstrahlung photonuclear reaction with a converter</i>						
(γ , Be), (γ , D)	10^9 – 10^{11}	Broad, “Fission-like”		μs	180 Hz	TNA, fissile

22.5 Generic Neutron Inspection Technologies

Three generic neutron-based inspection techniques can be readily identified: TNA, FNA and NRA. They and their main “derivatives,” such as PNI, PFNA and API, are briefly described in textual form in the appendix of this chapter and in numerous articles (see e.g. [1, 2], which include a comprehensive bibliography). Table 22.3 takes a different look at the same three techniques, however, from the point of view of the element detected. It elucidates, the fact that hydrogen can be directly detected by TNA through the (n, γ) process and by NRA through the energy dependence of the neutron elastic-scattering process. Oxygen, carbon and nitrogen can be determined by the FNA and NRA techniques. Nitrogen, however, can also be detected by the TNA process. The ability of the probing neutrons to create a certain fingerprint and its strength depend greatly on the neutron interaction cross section with the chemical element of interest, as discussed in Sect. 22.3.

22.6 Time-Dependent Effects

The TNA technique is the only one that can use either low-cost radioisotopic neutron sources, such as ^{252}Cf , or an accelerator-based electronic neutron generator (ENG). All other techniques require a charged-particle accelerator to generate neutrons. The ENG offers the flexibility of being switched on and off or repetitively pulsed. The latter brings forth the possibility of separating the fast-neutron-induced signals and background from those induced by thermal neutrons. These sources are typically pulsed at 100 to 5000 pulses per second with a corresponding time between pulses of 10 to 0.2 ms and with a typical pulse width of 5 to 500 μs . All the fast-neutron interactions are confined within the pulse. Following the pulse, only the thermal neutrons are present (decaying with a characteristic time constant of the system called the neutron thermal “die-away” time), resulting in a better signal-to-background for neutron capture gamma ray spectroscopy.

ENGs are usually low-voltage (50–150 kV) deuteron accelerators based on the cascade principle (see Box 3). In the most common form, the accelerator tube, target and ion source are all in a small sealed tube. The targets are typically made of tritiated titanium (or scandium) coated on a copper substrate. The most prolific neutron-generating nuclear reaction at low accelerating voltages is (d, T) , which generates 14 MeV neutrons. A vehicular explosive detection system employing such a neutron generator is described in Sect. 22.7.

The NRA technique requires either the use of variable-energy (0.5–5 MeV), high-resolution, monoenergetic neutron beams and any sensitive neutron detectors, or a broad-energy (e.g., white-spectrum) neutron beam with a high-resolution neutron spectrometer using the neutron time-of-flight (ToF) technique. The implementation of a variable-energy neutron beam is very

Table 22.3. Basic nuclear-physics characteristics of neutron-based techniques

Element	Generic Technique	Nuclear Reaction	Neutron Energy Range	Effective Cross Section (barn)	Practical Fingerprint (γ -rays or neutrons)
H	TNA	(n, γ)	Thermal	0.33	2.2 MeV γ -ray
H	FNA	$(n, n'\gamma)$	≤ 14 MeV	0	None
H	NRA	$(n, n)^*$	0.1–5 MeV	10–2	Change in total neutron σ vs. energy
C	TNA	(n, γ)	Thermal	3.5×10^{-3}	Very weak 4.93 MeV γ -ray
C	FNA	$(n, n'\gamma)$	≤ 14 MeV	~ 0.2 – ~ 0.45	4.43 MeV γ -ray
C	NRA	$(n, n)^*$	0.1–5 MeV	A few barns, with a few resonances	A few resonances and other structure in σ
N	TNA	(n, γ)	Thermal	0.011	10.8 MeV γ -ray
N	FNA	$(n, n'\gamma)$	≤ 14 MeV	0.02–0.1	5.11, 2.31, 1.64 MeV γ -rays
N	NRA	$(n, n)^*$	0.1–5 MeV	A few barns; very few and very narrow resonances	Weak signature
O	TNA	(n, γ)	Thermal	2×10^{-4}	None
O	FNA	$(n, n'\gamma)$	≤ 14 MeV	0.1–0.4	6.13 MeV γ -ray
O	NRA	$(n, n)^*$	0.1–5 MeV	In excess of 10 barns, many strong resonances	Very strong signature
Cl	TNA	(n, γ)	Thermal	33	6.11 MeV and other strong γ -rays
Cl	FNA	$(n, n'\gamma)$	≤ 14 MeV	0.15–0.3	1.76, 1.22, 3.16 MeV and other γ -rays
Cl	NRA	$(n, n)^*$	0.1–5 MeV	Weak, low- E_n resonances	Weak signature
“Metals” (Al, Si, Fe etc.)	TNA	(n, γ)	Thermal	0.1–10	Numerous strong γ -rays with generally medium to high energy
“Metals” (Al, Si, Fe etc.)	FNA	$(n, n'\gamma)$	≤ 14 MeV	0.1–2	Generally very narrow and dense resonances
“Metals” (Al, Si, Fe etc.)	NRA	$(n, n)^*$	0.1–5 MeV	2–5	Si strong; others have weak signatures

* (n, n) signifies an elastic scattering interaction of a neutron.

difficult, and thus the second approach is the one studied. The neutron ToF technique requires that the neutron generation be in the form of very narrow pulses, typically of the order of 1 ns, and with the time between pulses long enough that all neutrons belonging to a pulse vanish before the next pulse is generated. For a flight path of 5 m and the neutron energy range of interest, this time is of the order of 500 ns, entailing a repetition rate of 2 MHz or less.

Similar temporal behavior of the neutron production is required for the PFNA technique, though for completely different reasons. The ns pulsing of the monoenergetic (in the range of 7.5 to 9 MeV) neutron beam in the PFNA technique both results in a superior signal-to-background ratio (by separating the two) and allows the determination of all the local elemental densities. More detailed discussion of the PFNA system is given in Sect. 22.8.

22.7 Conventional Pulsed-Neutron Inspection (PNI) – Example of an Accelerator-Based Nonintrusive Inspection System

This technique was developed in the early 1960s for oil logging, nuclear-reactor start-up and the detection of nuclear materials. With this technique one can separate the (n, γ) reactions (TNA), which continue to occur after the fast neutrons have stopped being injected into the system, from the mixture of $(n, n'\gamma)$ (i.e. FNA) and TNA reactions taking place during the pulse. This is shown in Fig. 22.1. A sequence of neutron pulses, roughly $5\mu\text{s}$ wide or wider (depending on the allowable duty factor, the available intensity and the neutron “die-away” time constant of the system), is injected into the inspected item. The spectrum of the gamma rays produced during the pulse is measured, typically using high-efficiency NaI (sodium iodide), BGO (bismuth germanate oxide) or, sometimes, high-resolution Ge (germanium) solid-state detectors. Figure 22.1 shows spectra collected over many 14 MeV neutron pulses (upper curves in the spectra box), which are dominated by the fast-neutron interactions, as manifested by the presence of the oxygen and carbon inelastic lines. The materials inspected here were paper (indicated by a square marker), as a benign material, and ammonium nitrate (AN), as an explosive (smooth lines). Note that the carbon lines in AN (which does not have carbon) are very weak, representing background from shielding materials.

The spectra after the pulse are solely due to (n, γ) reactions and the weak, delayed oxygen activation (from $\text{O}(n, p)\text{N}$, which β -decays, with a half-life of 7.5 s, to ^{16}O , giving the same gamma rays as from inelastic scattering on ^{16}O). The spectra are characterized by the strong capture line of hydrogen and, in case of the explosive, the unique high-energy gamma ray (10.8 MeV) from capture in nitrogen. The observation of this line is the first and the primary indication of the presence of explosives (or fertilizers – which can be

used as explosives). The other elemental lines and spectral features in this technique are used to eliminate or confirm the presence of explosives. The technique is effective for detection of the bulk amounts of explosives that are typical for vehicle bombs.

Vehicle explosives detection systems (VEDS) are being used in various parts of the world.

22.8 PFNA – Example of a High-Performance Accelerator-Based Nonintrusive Inspection System

22.8.1 Introduction

Pulsed fast-neutron analysis (PFNA [3]) is a technique which uses a collimated pulsed beam of fast neutrons to excite the nuclei of common elements in bulk materials. The primary interactions of interest for contraband detection are gamma ray emissions following inelastic scattering of the fast neutrons with carbon, nitrogen and oxygen (and any other element present). Direct imaging of the elemental content of the material is accomplished by using a time-of-flight (ToF) technique to identify the position of the interactions, and gamma ray spectroscopy to identify the elemental gamma rays (see Figs. 22.3 and 22.4). The ratios of elemental signatures, or other combinations are used to identify contraband.

22.8.2 PFNA Nanosecond Pulsed-Neutron Production System

The cargo inspection system application of PFNA requires a beam of monoenergetic (with resolution $\Delta E/E < 1\%$) neutrons in an energy range of about 7 to 9 MeV, which can be raster scanned across the face of an inspected object such as a truck or cargo container. The required neutron beam is produced using a deuteron (^2H) beam in excess of $75\mu\text{A}$, with higher current desirable, from an NEC Pelletron tandem (models 9SDH-2 and 10.5SDH-4 are being used). An injector/buncher system was also developed for PFNA by NEC [4]. The ion source for this system is the Toroidal Volume Ion Source (TORVIS), which was an adaptation of a Brookhaven National Laboratory design by Prelec and Alessi [5]. The negative ions are accelerated to 80 keV and analyzed in a 90° dipole magnet before entering the bunching system. The 2.5 or 5 MHz double-drift buncher is designed to achieve a bunching efficiency of around 40% and a pulse width of 1 ns FWHM and 2 ns FWTM when the beam achieves its full energy (about 6 MeV). The fully accelerated deuteron beam enters a self-cooled deuterium (D_2 gas) target and produces neutrons via the $\text{d}(\text{D}, \text{n})^3\text{He}$ reaction.

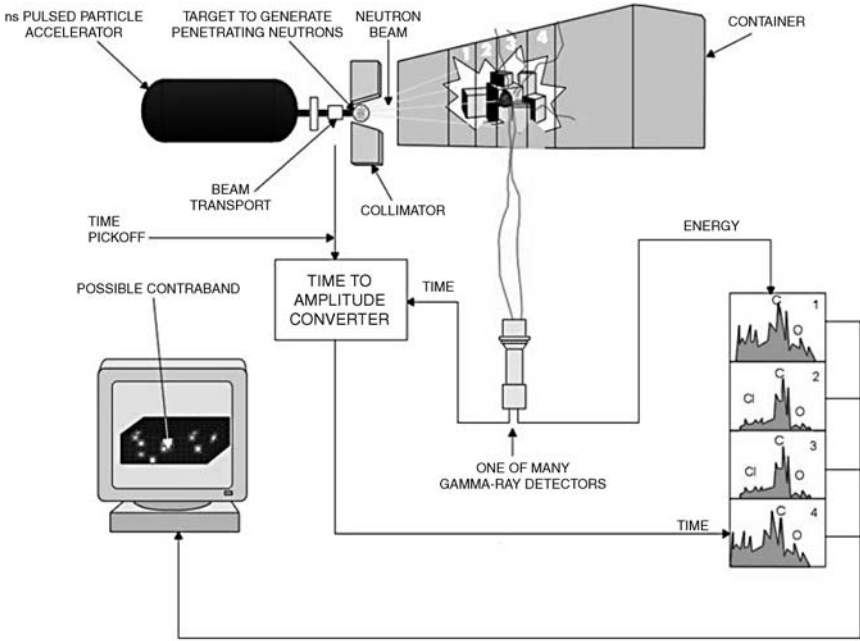


Fig. 22.3. Principle of PFNA

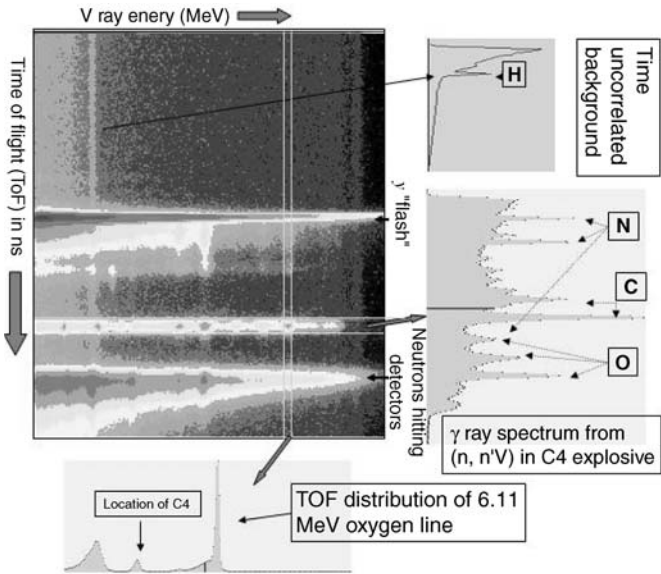


Fig. 22.4. Basic PFNA three-dimensional information – time of flight (ToF), γ -ray spectra and intensity

22.8.3 Scanning System

The oscillating neutron beam is achieved by mechanically rastering the deuteron beam and the beam line. After acceleration, the deuteron beam is taken through 90° using a dipole magnet. One unique requirement of the neutron production system is that it must not degrade the nanosecond time structure of the beam. This is done by focusing the beam to a spot, a few mm wide, near the center of the bend of the dipole. The dipole magnet, 90° beam line and gas cell are all rotated about the original beam axis to produce a vertical mechanical rastering of the beam. The 90° beam line is about 2 m long from the rotation axis to the gas cell. In operation the beam line can be rotated over a variable and wide angular range about the horizontal position. This motion is accomplished by mounting the dipole, 90° beam line components and deuterium gas cell on a platform which rotates on two bearings concentric with the beam axis. The platform and beam line hardware are shown in Fig. 22.5. The platform, magnet, vacuum hardware and counterweight weigh roughly 4500 kg. The scanning speed of the platform is adjustable. The platform is driven by a hydraulic actuator or an electric motor. The rotating and fixed beam lines are one continuous high-vacuum system employing a special high-vacuum rotary union.

The deuteron beam current deposits several hundred W of heat in the gas cell. The pressurized deuterium cell is separated from the evacuated beam drift tube by a $9\text{ }\mu\text{m}$ Havar foil and is self-supporting. The window is cooled by jets blowing deuterium gas on the interior face of the window. The deuterium is pumped through the cell in a closed-loop system and cooled with a heat exchanger. In order to not damage the window, the beam spot must be kept

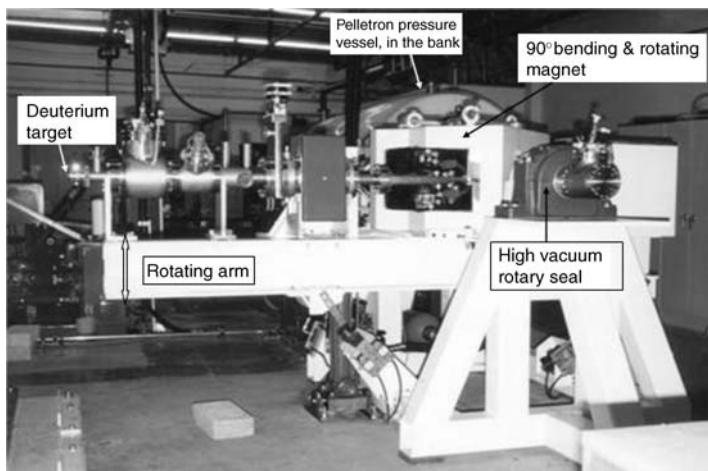


Fig. 22.5. The original PFNA scan arm, which rotates the beam line and neutron production target, before installation at Ancore's facility in Santa Clara, CA, USA

large on the window. The beam optics of the system were designed so that a proper-size spot is achieved at the gas cell when the beam is focused in the center of the bend of the dipole.

Horizontal collimation is accomplished using a fixed slot collimator. Vertical collimation is achieved by a variable collimator located at the end of the scan arm. This collimator can be adjusted to change the neutron beam cross section (the “pixel”).

The uniqueness of PFNA is its ability to determine the elemental (and from that the material) content in each volume element (called a voxel) of the inspected object. The spatial information is achieved, first, by collimating the relatively forward-peaked neutron beam from the $^2\text{H}(\text{d}, \text{n})^3\text{He}$ reaction. The point of interaction is provided by measuring the time of flight of a packet (or pulse) of neutrons (flying at a velocity of $\approx 4\text{ cm/ns}$) from its point of creation (the target) to the time when the gamma ray detector detects the appropriate rays from the $(\text{n}, \text{n}'\gamma)$ reaction. By correcting for the photon flight time, knowing the raster neutron beam angle and the position of the detector (there are more than 100 detectors distributed around the inspection tunnel), the system determines the location of the event. The spectrum of the detected gamma rays provides information on the material type at that location.

The three pieces of information: ToF, energy and intensity, collected every 10 to 100 ms (depending on the scanning speed and the size of the object), are shown in Fig. 22.4 (gray scale code: white to black indicates very high to very low intensity). This measurement was longer than the measurement done in a routine scan to show all the key physics aspects of PFNA. Figure 22.4 shows the time (abscissa) and energy (ordinate) obtained from inspection of a sample of C4 explosive placed in the middle of a marine container. It also shows a slice along the time corresponding to the location of the C4, providing the γ ray spectrum of the explosive. Also shown is an energy slice, corresponding to the oxygen signal from the explosive, showing its precise location. Capture γ rays that are the result of neutrons thermalized in the container and its surroundings, appear as a constant (time-“uncorrelated”) background and can easily be subtracted. The hydrogen capture line is shown in the top right corner. The other two features of the plot are the “ γ flash” created by the photons accompanying the neutron production hitting the detectors, and the second, later, broad time peak, representing the “n flash” i.e. scattered neutrons hitting the detectors. The time between these two time markers is when the PFNA information is collected. Through distributed computing and via a large number of very intensive computations, readily done in today’s high-end computers, the PFNA system makes an automatic decision as to whether a threat exists and where it is, and presents it to the operator.

22.8.4 Examples of PFNA Inspection Results

The overall PFNA system configuration is shown in Fig. 22.6. All the major components are identified there. The PFNA operator screen (item 11 in Fig. 22.6) contains, in a compact way, the inspection decision.

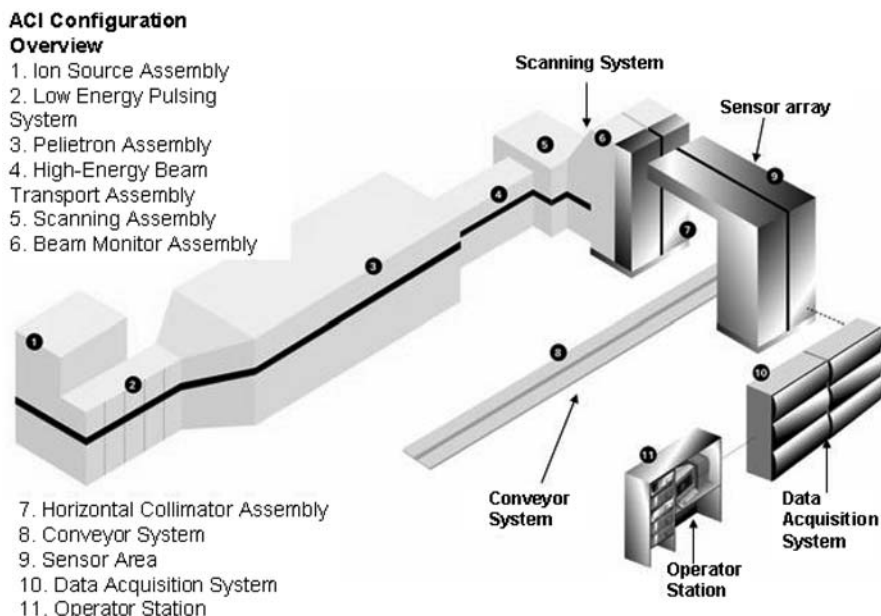
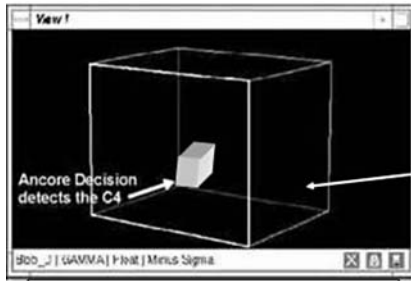


Fig. 22.6. Generic PFNA inspection system configuration with its main components

Examples of actual results are provided in Figs. 22.7 and 22.8 as an illustration of the power of PFNA inspection. Each figure shows the physical object in which the threat was concealed, the radiography image (obtained via the γ flash) of the inspected object and, finally, the PFNA decision image. Figure 22.7 shows this for a small piece of explosive in an air cargo container, whereas Fig. 22.8 does so for a large marine shipping container with concealed explosives and drugs.

22.9 Conclusions

Neutron-based inspection techniques are unique because of their great elemental (and hence material) specificity and sensitivity, and the ability to localize the threat. Particle accelerators from below 100 kV to 6 MV (or 3 MV tandems) have already played a very important role in the evolution of these



Impossible to discern explosive in the cluttered x-ray image



PFNA detects the explosives ("the needle"), and determines its locations. The benign materials ("the hay stack") become transparent

C4 hidden inside computer monitor



Fig. 22.7. Test results: PFNA inspection of an air cargo container loaded with computer monitors with a small concealed piece of explosive

PFNA detects explosives and other contraband hidden in scrap metal/machine goods inside trucks or other containers

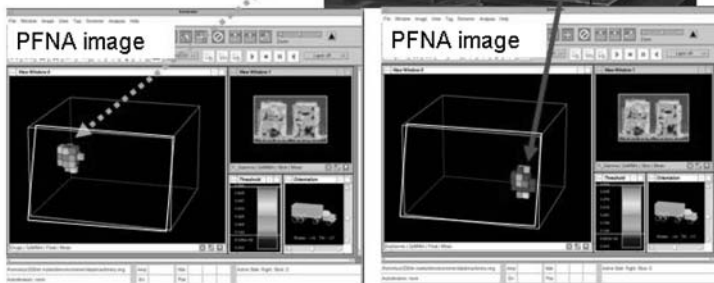
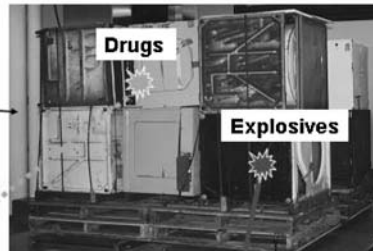


Fig. 22.8. Test results: PFNA inspection of a scrap metal shipment with hidden explosives and drugs

techniques. These accelerators will continue to increase their contribution to the inspection arena with the availability of higher-current, more compact and lower-cost machines.

22.A Appendix: Brief Description of Neutron-Based Inspection Techniques

TNA – thermal-neutron analysis. A method based on the capture of thermal neutrons by nuclei, creating high-energy gamma rays which are characteristic of the specific nuclei. Thermal neutrons are produced by slowing down fast neutrons from isotopic sources or accelerators in a specially designed moderator and in the interrogated object itself. The gamma rays are detected by an array of detectors surrounding the object. Spatial information on the interacting nuclei of interest (nitrogen, chlorine etc.) is also obtained from the detector array.

FNA – fast-neutron analysis. A method based on the interactions of fast neutrons, mostly by inelastic neutron scattering, with the nuclei of interest. Characteristic gamma rays from nitrogen, carbon, oxygen, chlorine and other elements can be measured. An array of gamma ray detectors surrounds the object to yield the spatial distribution of the elements of interest similarly to TNA.

PFNA – (nanosecond) pulsed fast-neutron analysis. A method based on fast-neutron interactions that yield characteristic gamma rays, similarly to FNA. The technique employs the time-of-flight (ToF) technique to obtain the spatial distribution of the signal. The ToF is obtained by using very narrow, about 1 ns, pulses of monoenergetic neutrons with a frequency of 1 to 10 MHz.

PNI – pulsed-neutron interrogation (an FNA/TNA combination). The interrogating neutrons are repetitively pulsed (typically 5–500 μ s wide pulses, 100 to 1000 times a second). During the pulse both fast- (i.e., FNA) and thermal-neutron (TNA) reactions take place. Between pulses only thermal neutrons exist, decaying with the system “die-away” time constant (typically 0.5 to 5 ms), and pure TNA signals are obtained. PNI is uniquely suitable for detection of fissile materials. Between the repetitive pulses of fast neutrons the thermal neutrons create fissions that can be detected via the presence of prompt epithermal neutrons, delayed neutrons, and prompt and delayed gamma rays.

API – associated-particle imaging. A technique to tag source neutrons in time and direction, by the associated charged particle emitted simultaneously in the nuclear reaction that generates the neutron, for example ^4He and ^3He in the (d, T) and (d, D) neutron-producing reactions, respectively.

NRA – neutron resonance attenuation. A neutron radiography technique measuring the areal density (density times thickness) of elements present in the interrogated object. The technique takes advantage of sharp resonances

and other features in the neutron total cross section in the energy range of 0.5 to 5 MeV. Two-dimensional projections of the areal density of N, C, O and H can be obtained using fast neutrons with a broad energy spectrum, generated in narrow pulses, to perform time-of-flight neutron spectrometry.

References

1. S.M. Khan (ed.): Proceedings of the 1st International Symposium on Explosive Detection Technology, FAA Technical Center, Atlantic City International Airport, November 1991
2. T. Gozani; "Neutron based non intrusive inspection techniques", Proceedings of the International Conference on Neutrons in Research and Industry, June 1996, Crete, G. Vourvopoulos ed., Proceedings of SPIE, Vol. 2867, p. 174
3. <http://www.ancore.com>
4. R.D. Rathmell et al.: "Pelletron accelerators for PFNA and MeV X-ray applications", Proceedings of the Contraband and Cargo Inspection Technology International Symposium, ONDCP and NIJ, 28–30 Oct. 1992, Washington, DC, p. 455
5. K. Prelec, J.G. Alessi: *Rev. Sci. Instr.*, **61**:1, 415 (1990)

23 Accelerator Mass Spectrometry

L.K. Fifield

Department of Nuclear Physics, Research School of Physical Sciences and Engineering, Australian National University, ACT 0200, Australia
`keith.fifield@anu.edu.au`

23.1 Introduction

Accelerator mass spectrometry (AMS) is a mass-spectrometric technique which incorporates an accelerator to achieve much higher sensitivities than are attainable with conventional mass spectrometers.

Usually, the atoms to be counted are radioactive with a long half-life, and are rare. The archetypal example is ^{14}C , which has a half-life of 5730 years and an abundance in living organisms of 10^{-12} relative to stable ^{12}C . Using AMS, the radiocarbon age of a sample less than 10 000 years old can be determined to a precision of 0.5% in a few minutes using a milligram or less of carbon. The sensitivity, which is more relevant than precision for very old or very small samples, is $\sim 10^4$ atoms of ^{14}C .

The utility of AMS is not limited to ^{14}C alone, however, and many other isotopes are amenable to the technique, the most important of which are ^{10}Be ($T_{1/2} = 1.5$ Ma), ^{26}Al (720 ka), ^{36}Cl (301 ka) and ^{129}I (15 Ma).

AMS emerged from nuclear-physics laboratories with electrostatic tandem accelerators in the late 1970s. The first AMS observations of ^{14}C from natural materials were reported simultaneously by two groups, one working at McMaster and the other at Rochester [1, 2].

Following these first successful demonstrations, dedicated AMS systems based on the Mark I Tandetron [3] were soon developed, and the first three were installed at Arizona, Oxford and Toronto in 1981–82. At the same time, several nuclear-physics laboratories added an AMS component to their research programs, and a number of these have evolved into dedicated AMS facilities. Techniques for other isotopes, in particular ^{36}Cl [4], ^{129}I [5], ^{10}Be [6] and ^{26}Al [7], were soon developed.

Beginning in about 1990, a series of second-generation purpose-built systems operating at 3 MV have had a major impact on the field. Subsequently, the development of systems based on small accelerators operating at 0.5 MV or less is bringing high-precision, high-throughput AMS for radiocarbon within reach of more laboratories. There are presently (2003) about 50 laboratories worldwide which are actively involved in AMS.

In parallel with the growth in the number of facilities, applications of the technique have burgeoned. Traditional ^{14}C dating is no longer the dominant application. Carbon-14 remains the most important isotope, but much of the

research is now directed towards an understanding of global climate change via studies of oceanic circulation, atmospheric processes and past climates. Biomedical applications of ^{14}C are also becoming increasingly important. Cosmic-ray exposure dating, which exploits the buildup of ^{10}Be , ^{26}Al or ^{36}Cl in surface rocks, is making an increasingly significant contribution to studies of landscape evolution and paleoclimates. Beryllium-10 has also been used in studies of soil processes, the deposition of ocean sediments, and the sources of volcanic rocks. Both ^{36}Cl and ^{129}I have been used in hydrology and in tracing the migration of nuclear waste, while ^{26}Al and ^{41}Ca have found application in biomedicine. New applications continue to surface, often spawning new developments in technique.

In this review, I shall concentrate on the current state of the art in AMS technique, and on recent innovations. The principal areas of application of AMS will be briefly surveyed to give a flavor of the science. More detail can be found in earlier reviews [8–11].

23.2 Principles of Accelerator Mass Spectrometry

23.2.1 The Need for AMS

For rare, long-lived radioisotopes, AMS overcomes certain fundamental limitations of both decay-counting and conventional mass spectrometry. These limitations are best illustrated using the familiar example of ^{14}C .

There are 6×10^{10} atoms of ^{14}C in 1 g of modern carbon, but only 14 decay per minute. In order to achieve 0.5% statistical precision with decay counting, it would be necessary to count for at least 48 hours.

A mass-spectrometric method could in principle count a much higher fraction of the ^{14}C atoms in a sample than decay-counting. Conventional mass spectrometers are not practicable, however, because the ^{14}C ions are masked by relatively intense fluxes of ions of the ^{14}N isobar, by background from tails of the stable isotopes ^{12}C and ^{13}C , and also by molecular ions such as ^{13}CH , $^{12}\text{CH}_2$, ^{12}CD and $^7\text{Li}_2$. As outlined below, all of these limitations are overcome in an accelerator mass spectrometer.

In marked contrast to decay-counting, typical ^{14}C counting rates in an AMS system are $\sim 100 \text{ s}^{-1}$, and only 1 mg of carbon is sufficient for a measurement.

23.2.2 Principles of AMS

Accelerator mass spectrometry combines the high efficiency of mass spectrometry with excellent discrimination against isobaric, isotopic and molecular interferences. It achieves this by:

- using negative ions
- dissociating molecular ions after a first stage of acceleration
- identifying individual ions after a second stage of acceleration and analysis.

Figure 23.1 shows a schematic representation of an accelerator mass spectrometer. Each of its essential features is described in sequence below. Details specific to individual isotopes will be taken up later.

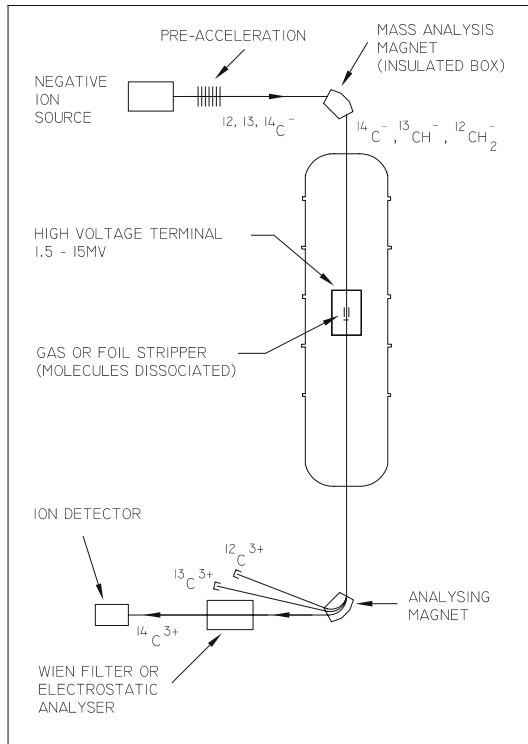


Fig. 23.1. The essential features of an AMS system. Mass-14 ions are shown being transmitted around the mass-analysis magnet. Periodically or simultaneously, mass-12 and mass-13 ions are transmitted, and measured in the off-axis cups after the analyzing magnet

- Negative ions are generated from the sample in a cesium sputter source, preaccelerated to 30–200 keV and mass-analyzed by a magnet. In the cases of ^{14}C , ^{26}Al and ^{129}I , the choice of negative ions eliminates the isobaric interferences because ^{14}N , ^{26}Mg and ^{129}Xe do not form stable negative ions.
- The mass-analyzed negative ions are accelerated to the positive high-voltage terminal of the tandem accelerator, where they pass through a gas

- or foil stripper. A crucial role of the stripper is to dissociate any interfering molecular ions.
- The now positively charged ions are further accelerated back to ground potential in the second stage of the tandem accelerator. Subsequent magnetic and electric analyzers select the ions of interest with a well-defined combination of charge state and energy.
 - Individual ions are identified and counted by a detector. Identification is necessary because ions other than those of the AMS isotope may also reach the detector.
 - Generally, AMS determines the abundance ratio of the rare isotope to a stable abundant isotope of the same element, $^{14}\text{C}/^{12}\text{C}$ for example, by also measuring the flux of the abundant species as an electrical current in a Faraday cup after acceleration.

The above gives the bare essentials of an AMS system. In the following, the various components are considered in somewhat more detail.

The Ion Source

AMS ion sources are almost exclusively cesium sputter sources (see Chap. 12). Most are “high-intensity” sources, in which the hot tantalum ionizer and the sample are within the same volume containing Cs vapor, although some Cs-gun-type sources are still employed. The majority are either the 846 model source from HVEE, which has a 60-sample carousel, or the MC-SNICS manufactured by NEC, which is available with either a 40- or a 134-sample wheel. In both sources, the carousel or wheel can be replaced through a vacuum lock in order to minimize downtime. In the HVEE source, it is possible to move the sample relative to the Cs beam to avoid cratering and to utilize the sample material efficiently. Driven by the need for both higher throughput and higher precision, significant progress has been made in understanding and improving these sources, and $^{12}\text{C}^-$ currents as high as 200 μA have been achieved from graphite samples. Typical negative-ion beam currents for the species commonly used in AMS are given in Table 23.1.

Table 23.1. Typical beam currents from high-intensity cesium sputter sources used for AMS

AMS Isotope	Material	Selected Ion of Stable Isotope	Negative-ion Current (μA)
^{10}Be	BeO	$^9\text{Be}^{16}\text{O}^-$	1–20
^{14}C	Graphite	$^{12}\text{C}^-$	20–100
^{26}Al	Al_2O_3	$^{27}\text{Al}^-$	0.1–1
^{36}Cl	AgCl	$^{35}\text{Cl}^-$	5–25
^{129}I	AgI	$^{127}\text{I}^-$	2–10

At Oxford, a source which uses CO_2 as well as graphite has been in operation for several years [12], and similar sources are beginning to appear in other laboratories. In this case, the disadvantage of lower beam currents ($\sim 10\text{ }\mu\text{A}$) is compensated by the substantial simplification in sample preparation and the capability to run samples as small as a few μg .

The Injection System

Specialized injection systems have been developed in order to inject both the rare and the abundant isotopes into the accelerator.

In a sequential system, switching between isotopes is accomplished by changing the energies of the negative ions in the mass-analysis magnet after the ion source. This is achieved by applying different voltages to the electrically isolated vacuum box of the magnet. In order to minimize the effect of fluctuations in the ion source output, switching times should be short and repetition rates high. Switching times down to μs are employed, and typical repetition rates are 10 s^{-1} , with $>90\%$ of the time spent on the rare isotope. Typically, the very intense stable beams are pulsed into the accelerator for $100\text{ }\mu\text{s}$ or less in order to avoid beam-loading effects. Electrostatic lenses at the beginning and end of the insulated section are required to ensure identical trajectories of the different isotopes.

Alternatively, the stable and radioactive isotopes can be injected simultaneously. Simultaneous-injection systems employ a sequence of dipole magnets and lenses which allow the different isotopes to follow different trajectories after leaving the ion source before being recombined at the entrance to the accelerator. This approach was pioneered at McMaster [13] and has been adopted for the Mark II Tandetron systems [14]. The recombinator for the latter machines, which are used almost exclusively for ^{14}C , is depicted in Fig. 23.2. Attenuation of the intense ^{12}C beam is accomplished by interposing a rotating slotted wheel in its path where the separation of the isotopes is a maximum.

In either case, the beam currents of the stable beams are measured in off-axis Faraday cups after the postacceleration analyzing magnet (Fig. 23.1).

The Accelerator

The appropriate size of accelerator is determined by the following two factors.

- *The charge state of the positive ions.* Until recently, it was widely accepted that the minimum charge state for AMS was 3^+ . This was based upon the fact that some singly or doubly charged molecular ions are stable, $^{13}\text{CH}^+$ and $^{12}\text{CH}_2^{2+}$ for example. If these were to survive the stripping process, and at the gas stripper pressures typically employed, some certainly do, they would be essentially indistinguishable from ^{14}C with the same charge

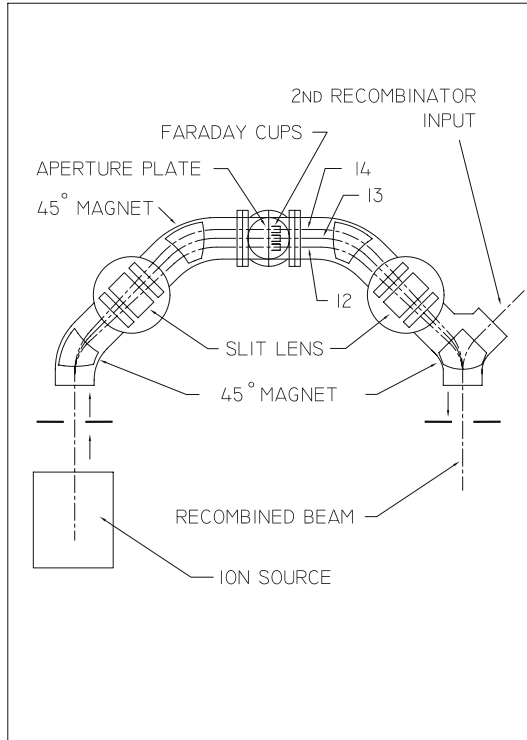


Fig. 23.2. The recombimator injection system used by AMS systems based on Mk II Tandetrans

state. Hence, they were thought to preclude the use of 1^+ or 2^+ ions, not only for ^{14}C , but for all AMS isotopes. Since the 3^+ stripping yield for carbon is appreciable only above 2 MeV, the requirement was for accelerators operating at 2 MV or more.

In one of the major advances in AMS of the past few years, the Zürich group has demonstrated [15] that it is possible to dissociate essentially all of the ^{13}CH and $^{12}\text{CH}_2$ molecules at gas stripper thicknesses of $\sim 2 \mu\text{g}/\text{cm}^2$. The original demonstration employed a 500 kV accelerator and the 1^+ charge state, but subsequently the Zürich group has taken miniaturization a step further by showing the feasibility of a system with a footprint of only $3 \times 2.3 \text{ m}^2$ based on a 200 kV vacuum-insulated accelerator. In parallel, NEC have developed a system employing only a single stage of acceleration to 300 kV across a standard air-insulated acceleration tube. The gas stripper, final analysis stages and detector are all at high voltage. Both systems utilize commercial high-voltage power supplies.

- *Ion identification.* For some AMS isotopes, the choice of negative ions does not exclude the isobar. The best-known example is ^{36}Cl , where negative

ions of the ^{36}S isobar are produced equally readily in the ion source, and inevitably accompany the ^{36}Cl all the way to the final detector. Discrimination of a few ^{36}Cl ions in the presence of a much greater flux of ^{36}S ions is only possible at energies above ~ 50 MeV, and hence AMS of ^{36}Cl requires larger accelerators.

The tandem electrostatic accelerators employed in AMS may therefore be conveniently grouped into four categories on the basis of their maximum terminal voltage V_T .

1. $V_T < 1$ MV. Following the successful demonstration that radiocarbon AMS was possible using 1^+ ions, several systems based on a 500 kV accelerator have now been delivered by NEC or are on order. Although they are presently being used almost exclusively for high-precision radiocarbon measurement, their potential for other isotopes is being actively explored by the Zürich group, and promising results for ^{10}Be , ^{26}Al , ^{41}Ca , ^{129}I and plutonium have been reported [16, 17].
2. $1 < V_T < 3$ MV. This category includes the Tandetrans, now manufactured by HVEE, and 3 MV accelerators from NEC, and presently constitutes the majority of dedicated AMS systems. They are predominantly used for high-precision ^{14}C measurements, but also for ^{10}Be , ^{26}Al and ^{129}I . The Tandetrans are charged by a radio-frequency voltage-doubling power supply.
3. $4.5 < V_T < 9$ MV. Machines in this category are principally extensively modified ex-nuclear-physics accelerators, predominantly EN and FN accelerators manufactured by HVEC in the 1960s and 70s. More recently, NEC have entered this category with a purpose-built 5 MV system, of which four have been delivered. HVEE also now offer a 5 MV Tandetrans system. Essentially the full range of AMS isotopes can be covered by accelerators of this size, although the separation of ^{36}Cl from ^{36}S is challenging at less than 6 MV.
4. $V_T > 10$ MV. These are operational nuclear-physics accelerators on which AMS typically takes 20% of the beam time. Modifications to the injection system or accelerator specifically for AMS are usually minimal. The higher energies available from these larger machines are particularly advantageous for ^{36}Cl and for isotopes requiring a gas-filled magnet. In addition, much of the development of new isotopes has been carried out on them. They include the NEC 14UD Pelletron and the HVEC MP and XTU accelerators.

The desirable criteria for an AMS accelerator are that transmission of ions through the accelerator should be high and reproducible, that this transmission should be insensitive to small changes in the injection or accelerator parameters, and that the terminal voltage should be very stable. High and flat-topped transmission is achieved by using large-diameter acceleration tubes, spacious vacuum chambers within magnets, and apertures which are

as large as possible. Ion-optical transmission is typically $>80\%$ in dedicated AMS systems. Consistency of transmission is best achieved using gas stripping in the high-voltage terminal, and most of the facilities which aim at high-precision ^{14}C use an argon gas stripper. Gas stripper canals are typically 8 mm in diameter and may constitute the smallest restriction in the system. Careful attention to ion optics is therefore required to ensure that transmission losses are minimal.

Stabilization of the accelerator's terminal voltage presents a particular challenge. The usual slit-stabilization method, which depends on nA currents striking slits after the final analyzing magnet, is not applicable, since there are typically only a few ions per second of the AMS isotope. Two solutions have been adopted. Either the signal from a generating voltmeter (GVM) is used, or the off-axis position of one of the stable beams is monitored by a Faraday cup which is split in two to give separate left and right signals [18]. The former is the more widely used and is capable of excellent stability. For example, the fluctuations in the terminal voltage of the 14UD accelerator at the Australian National University are ~ 1 kV in 14 MV under GVM stabilization. Of course, it helps if the accelerator is intrinsically very stable, and Pelletron charging systems and the solid-state power supplies of the Tandetrans have distinct advantages over the older rubberized belts in this regard.

Postacceleration Analysis

Following acceleration, the charge state and energy of interest are selected by magnetic analysis. The stable isotopes have different radii of curvature and are collected in off-axis Faraday cups at the focal plane of the magnet. At the exit from a typical 1 m radius analyzing magnet, the separation between ^{12}C and ^{14}C is 8 cm, and hence purpose-built magnets with wide pole pieces at the exit are required to accommodate the different ion species.

An additional analysis stage, either an electrostatic analyzer or a velocity filter, follows the analyzing magnet to remove that small fraction of molecular fragments which has acquired the correct energy to follow the same trajectory as the AMS isotope through the magnet. These can otherwise cause unacceptably high counting rates in the ion detector. Rates depend on the vacuum in the high-energy acceleration tube, and because some unwanted ions inevitably leak through even the additional analysis stage, the higher the vacuum the better. To this end, gas strippers which employ turbomolecular pumps in the high-voltage terminal to recirculate the gas [19] are now widely employed, and high-pumping-speed cryopumps or turbomolecular pumps are installed as close as possible to the end of the high-energy tube. Acceleration tubes of the NEC design, in which ceramic insulators are bonded without adhesives to titanium electrodes, offer advantages in terms of higher vacuum over the more widely used tubes in which glass insulators are bonded to metal electrodes with organic adhesives.

Detectors

A range of ion detectors have been variously employed to count the AMS isotope.

- *Silicon detectors.* These measure only the energy of the ion. In some cases, this is sufficient for unique identification of the AMS isotope, since any background ions have a different mass and therefore energy. A drawback of silicon detectors is that they are susceptible to radiation damage.
- *Ionization chambers.* Ionization chambers are more robust, and can measure not only the total energy of an ion but also its rate of energy loss as it slows down in the detector gas. A typical multielement ionization chamber is depicted in Fig. 23.3. Until recently, gas-confining windows were almost exclusively Mylar, typically $1.5\mu\text{m}$ thick, but the advent of silicon nitride windows with thicknesses as low as $0.03\mu\text{m}$ seems set to confer substantial gains in energy resolution, especially for low-energy heavy ions. Propane, isobutane and P-10 (90% argon, 10% methane) are the usual counter gases employed. Isobutane has the advantage in some applications of a higher stopping power but is comparatively expensive, while propane is cheap and readily available. Electrons produced by the passage of the energetic ion drift towards the anode in the transverse electric field. The anode is subdivided into sections, each of which collects those electrons produced beneath it, thereby measuring the energy lost by the ion along that portion of its track. Since ions of different Z lose energy at different rates, this

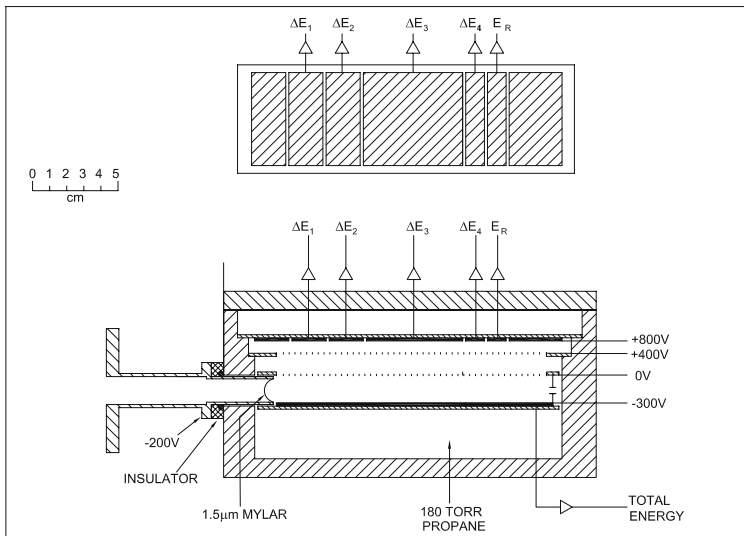


Fig. 23.3. A cross section through the multielement ionization chamber used at the Australian National University. The *upper* panel shows a plan view of the anode electrode

energy-loss information permits the separation of isobars, ^{36}Cl and ^{36}S for example, which have identical energies when they arrive at the detector. Issues specific to certain isotopes will be taken up in Sect. 23.3.

Hybrid detectors, which employ a gas ionization chamber for energy-loss measurement and a silicon detector to measure the residual energy, have also been used [20].

- *Time-of-flight systems.* For heavier ions, the energy resolution of an ionization chamber or silicon detector is insufficient to resolve neighboring masses. Better separation is possible by combining the energy measurement with a determination of the ion's velocity via a time-of-flight measurement. Such systems have been used for ^{129}I , and for very heavy isotopes such as ^{236}U .

A time-of-flight system consists of “start” and “stop” detectors separated typically by 2 m. Start detectors invariably consist of a thin carbon foil and a microchannel plate, which multiplies the electrons liberated from the foil by the passage of the ion. The stop detector may be similar to the start detector, in which case it is backed by a silicon detector or ionization chamber for energy measurements. Alternatively, a silicon detector (which may be part of a hybrid system) is often used to provide both timing and energy signals. Typical time resolutions achieved with such devices are 300–500 ps, which is more than adequate to separate ^{129}I from ^{128}Te and ^{127}I .

A drawback of time-of-flight systems is the loss of efficiency associated with scattering from grids and from the start foil, and efficiencies typically range between 50 and 80%. Scattering from the foil is crucially dependent on the foil thickness, and diamond-like carbon foils as thin as $0.5\text{ }\mu\text{g}/\text{cm}^2$ are now available.

- *Gas-filled magnets.* Isobars, which have the same mass and energy after acceleration, may be separated by passage through a gas-filled region within a magnetic field. Owing to the difference in Z , their average charge states are different in the gas-filled region, and hence they follow different trajectories in the magnetic field. The unwanted, high-counting-rate isobar can then be intercepted by an appropriately positioned baffle. The AMS isotope continues to an ionization chamber which distinguishes between the AMS isotope and events due to tails of the isobar distribution. An excellent review of gas-filled magnets as applied to AMS has been given by Paul [21]. Purpose-built systems have been installed at Zürich [22] and Munich [23]. Existing magnetic spectrometers have also been pressed into service [24].
- *X-ray detectors.* Fast ions may be identified by the characteristic X-rays they emit following excitation in a foil, thereby allowing isobar separation at energies where ionization chambers give rather poor discrimination. This technique has been applied to the detection of several isotopes, including ^{36}Cl , $^{59,63}\text{Ni}$, ^{60}Fe , ^{79}Se and ^{126}Sn [25–27]. The optimal foil material has $Z_{\text{foil}} \approx Z_{\text{ion}} + 2$, and X-ray yields are strongly dependent on energy. At the

energies obtainable from a 9 MV accelerator, yields vary from one X-ray per incident ion for light ions such as Si to one X-ray per 300 incident ions for ^{106}Pd [27]. Only a fraction of these X-rays are actually detected, however, owing to the limited solid angle subtended by the detector and an intrinsic detector efficiency of less than unity.

23.3 Techniques for Individual Isotopes

Each AMS isotope is a little different, and has its own special considerations and refinements of technique. These are discussed below.

23.3.1 Carbon-14 ($T_{1/2} = 5730 \text{ a}$)

Applications of ^{14}C place more stringent demands on precision than any other isotope, and a precision of 0.3%, corresponding to 25 years in radiocarbon age, is presently the benchmark for dedicated AMS laboratories. In order to attain this level of precision it is necessary, first, to obtain sufficient ^{14}C counts that the statistical uncertainty is at the desired level, and, second, to achieve a high degree of reproducibility.

In order to achieve a statistical uncertainty of 0.3%, 150 000 counts are required, both for the sample of unknown age and for the standard relative to which it is measured. Since the ^{14}C counting rate from a 5000-year-old sample is typically 70 s^{-1} , a sample of this age must be run for at least 30 minutes to obtain the requisite number of counts. This count rate assumes a source output of $50 \mu\text{A}$ of $^{12}\text{C}^-$, a charge state fraction of 50% and ion-optical transmission of 80%. Generally, the requisite number of counts will be accumulated over a number of runs interleaved with standards and other samples, allowing an estimate of the external error and hence a check on the reproducibility of the system.

Good reproducibility requires that transmission be insensitive to small changes in any of the system parameters. Considerable care is therefore taken during setting up to ensure flat-topped transmission in all parameters.

Modern systems incorporate a high level of automation to control not only the fast-cycling sequence, but also the sequencing of samples and, increasingly, automated tuning. Unattended operation for the duration of a complete wheel or carousel of samples is becoming commonplace. In addition, the computer control program maintains a watch on the integrity of the data by continuously monitoring the ratios of the different isotopes as a measurement proceeds. Possible problems can then be identified and flagged for later consideration, or an operator alerted.

For high-precision dating, it is necessary to correct for the natural fractionation inherent in biological processes. Carbon from C3 plants (e.g. forest trees) is depleted by 1.5% in ^{13}C relative to C4 plants (most grasses), which

are in turn depleted by 1% relative to marine organisms. Hence, in addition to the AMS measurement of the $^{14}\text{C}/^{12}\text{C}$ ratio, a measurement of the $^{13}\text{C}/^{12}\text{C}$ ratio is also required. In the past, the $^{13}\text{C}/^{12}\text{C}$ ratio was usually measured off-line on a subsample of CO_2 with a conventional mass spectrometer because the measurement of this ratio by the AMS system was insufficiently precise. Modern systems, however, now achieve precisions of 2 per mil or better in the AMS measurement of the $^{13}\text{C}/^{12}\text{C}$ ratio, thereby avoiding the extra step. An additional benefit is that any fractionation introduced during conversion of the sample to graphite or in the ion source is automatically accounted for.

23.3.2 Beryllium-10 ($T_{1/2} = 1.50 \text{ Ma}$)

Beryllium does not form a stable atomic negative ion. Samples are therefore prepared as beryllium oxide and the BeO^- molecular ion is selected for analysis. Currents of several μA are obtained, and can be enhanced by mixing the BeO with niobium metal powder [28].

At the stripper, the ^{10}Be atom carries only 10/26 of the energy of the molecular ion and thus strips to a lower average charge state than if it had the full energy. Hence, the 2^+ charge state is employed on the smaller machines operating at 2–3 MV, and the 3^+ charge state on larger accelerators operating at $>5 \text{ MV}$. In the latter case, a foil stripper following the gas stripper leads to a higher yield of the 3^+ charge state. A foil stripper alone is not suitable, because “Coulomb explosion” of the molecule as it breaks up in the foil results in substantial losses due to increased divergence and energy spread.

Boron-10 is the stable isobar of ^{10}Be . It readily forms BO^- ions, and hence ^{10}B ions inevitably accompany the ^{10}Be ions after acceleration. Despite the best efforts of the chemist, typical counting rates of these unwanted ^{10}B ions are greater than 1 MHz. Two solutions to this problem have been adopted, depending upon whether the accelerator is large ($\geq 5 \text{ MV}$) or small ($\sim 2 \text{ MV}$).

At the final energies of 20 MeV or more achieved with the larger accelerators, the difference in stopping range of ^{10}B and ^{10}Be ions can be exploited. Boron-10 ions may be stopped in a gas cell or a foil before the detector, allowing the lower- Z ^{10}Be ions, which retain $\sim 40\%$ of their initial energy, to be detected in an ionization chamber immediately behind the stopping region. Discrimination between ^{10}Be ions and other species such as ^9Be and ^7Be may be improved by taking two or more energy-loss signals from this detector.

Beryllium-7 ions from the $^1\text{H}(^{10}\text{B}, ^7\text{Be})^4\text{He}$ reaction are a potential source of background when ^{10}B fluxes are high. It is necessary, therefore, to avoid any hydrogenous component in the ^{10}B absorber. Argon is used in gas-absorber cells, and Havar (a Co/Cr/Fe/Ni/W alloy) for stopper foils or windows. An advantage of using a gas cell as the ^{10}B absorber is that it may be configured as an ion chamber. A ^{10}B flux of 10^6 s^{-1} produces an ion current of $\sim 100 \text{ nA}$, which not only provides an indication of the boron flux, but may also be used to tune the AMS system for optimal ^{10}Be transmission.

At the lower energies of ~ 5 MeV available from accelerators operating at ~ 2 MV, straggling in the range of the ^{10}B ions is too large to permit the use of a passive absorber. Instead, a $200\text{ }\mu\text{g}/\text{cm}^2$ carbon foil is interposed in the path of the ions before the final magnetic analysis [29]. Boron-10 ions lose 300 keV more energy in this foil than do the ^{10}Be ions. Hence all but 0.2% of the ^{10}B ions are rejected in the magnet, whereas 20% of the ^{10}Be ions reach the detector. An ionization counter can then cope with the much reduced rate of ^{10}B ions, which is typically in the range of a few kHz. In this case, the full energy of the ions is deposited in the detector, allowing ready discrimination between ^{10}Be and ^{10}B ions.

The $^{10}\text{Be}/^9\text{Be}$ ratio can be determined either by injecting $^{10}\text{BeO}^-$ and $^9\text{BeO}^-$ sequentially, or by a novel simultaneous-injection method [30]. The latter exploits the fact that $^{10}\text{Be}^{16}\text{O}^-$ and $^9\text{Be}^{17}\text{O}^-$ are injected into the accelerator together and that, after acceleration, $^{17}\text{O}^{5+}$ ions differ by only 1% in magnetic rigidity from $^{10}\text{Be}^{3+}$ ions. Hence, the $^{17}\text{O}^{5+}$ ion current, which is a surrogate for the ^9Be current, can be collected continuously in an off-axis Faraday cup after the postacceleration analyzing magnet.

23.3.3 Chlorine-36 ($T_{1/2} = 301\text{ ka}$)

The ^{36}S isobar is the principal challenge confronting AMS measurement of ^{36}Cl . Although ^{36}S constitutes only 0.02% of natural sulfur, a mere 1 ppm of sulfur results in a ^{36}S counting rate in the detector of 1000 s^{-1} . In contrast, a typical environmental sample will have a $^{36}\text{Cl}/\text{Cl}$ ratio of 10^{-13} and a ^{36}Cl counting rate of only $\sim 0.5\text{ s}^{-1}$. Since a sensitivity of 10^{-15} in the ratio is desirable, the detector must provide a discrimination factor of at least 10^6 between ^{36}Cl and ^{36}S ions.

Such discrimination requires the higher energies available from the larger accelerators. Effective separation is best achieved if ^{36}S penetrates significantly further into the detector than ^{36}Cl . At energies above the Bragg peak at about 24 MeV, there is a $\sim 13\%$ difference in energy loss between ^{36}Cl and ^{36}S . If accumulated over a sufficient distance in the detector gas, this leads to a substantial spatial separation of the Bragg peaks, as illustrated in Fig. 23.4. In a detector such as that shown in Fig. 23.3, this spatial separation results in large differences in the energy-loss signals near the end of the range. The other energy-loss signals preceding the Bragg peak also provide useful discrimination. At energies above 100 MeV, rejection ratios of better than $10^6:1$ are achieved at counting rates up to 10^4 s^{-1} . Although discrimination against ^{36}S is optimal for accelerators operating in excess of 10 MV, ^{36}Cl is nevertheless measured routinely at several laboratories using accelerators operating as low as 6 MV. Facilities at Purdue and Livermore carry out measurements with ~ 60 MeV ions from FN accelerators [31,32], and measurements are conducted with 48 MeV $^{36}\text{Cl}^{7+}$ ions from an EN accelerator at Zürich [33]. There is also a push to perform ^{36}Cl measurements on the new generation of 5 MV

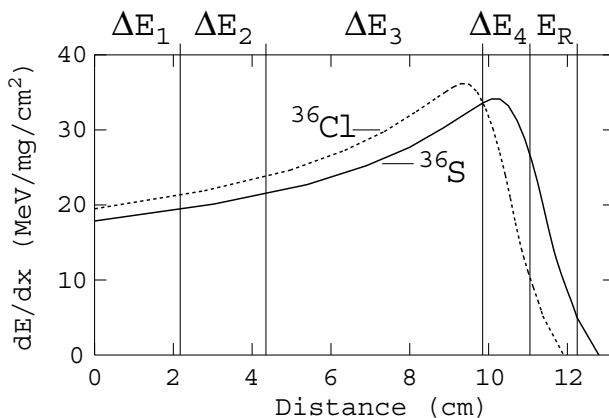


Fig. 23.4. Energy loss as a function of distance into a gas-ionization counter for 154 MeV ^{36}Cl and ^{36}S ions. The positions of the anode electrodes of the detector shown in Fig. 23.3 are indicated. Large differences between the two isotopes in the areas under the curves for individual electrodes are evident, particularly for ΔE_4 and E_R

systems, and acceptance tests for the new 5 MV system at East Kilbride in Scotland included a successful demonstration of a ^{36}Cl capability.

Clearly, it is advantageous to reduce the sulfur content of the sample as far as practicable, and this is crucial for systems operating at 5–8 MV. The silver chloride samples themselves are purified by precipitating barium sulfate from alkaline solution [34]. In addition, it is crucial to ensure that any parts of the sample holder that can be sputtered by the cesium beam are also very low in sulfur. This is achieved by masking any exposed surfaces with either silver bromide or the sample itself. Commercial silver bromide is fortuitously low in sulfur.

23.3.4 Aluminum-26 ($T_{1/2} = 720 \text{ ka}$)

Magnesium does not form a stable negative ion. Hence, smaller as well as larger accelerators are suitable for ^{26}Al measurements. The principal limitation arises from the reluctance of aluminum to form negative ions owing to its low electron affinity (0.44 eV). Beam currents of Al^- ions from Al_2O_3 samples as high as $2 \mu\text{A}$ have been reported, but 0.2–0.5 μA seems to be more typical. Consequently, and especially for geological samples which have $^{26}\text{Al}/\text{Al}$ ratios of less than 10^{-13} , running times per sample tend to be long, and measurement precision low. Generally, the insulating Al_2O_3 is mixed with an approximately equal weight of silver powder to ensure good electrical and thermal conduction.

An odd charge state is always employed. An even charge state, $^{26}\text{Al}^{6+}$ for example, would be plagued by intense counting rates of $^{13}\text{C}^{3+}$ from the

injection and subsequent breakup of the $^{13}\text{C}_2^-$ molecular ion. These ions have the same E/q and mE/q^2 as $^{26}\text{Al}^{6+}$ and therefore pass all magnetic and electric analyzers.

The AlO^- ion is produced more than an order of magnitude more prolifically than the Al^- ion. Unfortunately, MgO^- ions are also formed in the ion source, and count rates of ^{26}Mg ions at the detector are prohibitively high if the AlO^- ion is selected for injection. A possible solution would be to use a gas-filled magnet to greatly reduce the ^{26}Mg counting rate, but apart from a few measurements at Munich, this has not been seriously pursued to date, despite its obvious attractions.

23.3.5 Iodine-129 ($T_{1/2} = 15 \text{ Ma}$)

Xenon-129 does not form a stable negative ion, and hence AMS of ^{129}I is not troubled by the stable isobar. Consequently, ^{129}I can be measured as effectively with a small accelerator as with a large one [35]. Several μA of iodine beam can be obtained from AgI samples, and sensitivities of 10^{-15} in the $^{129}\text{I}/^{127}\text{I}$ ratio are readily achievable.

Background from ^{127}I ions, which are difficult to resolve from ^{129}I in the detector, is the principal challenge to be surmounted. A very small fraction of the $^{127}\text{I}^-$ ions can acquire sufficient additional energy from the sputtering Cs^+ beam to be injected along with the ^{129}I ions. Subsequent charge-changing processes during acceleration and analysis can result in a very small, but significant, proportion of these arriving at the detector. An electrostatic analyzer (ESA) between ion source and mass-analyzing magnet eliminates this “sputter tail”, and is essential on small accelerators. In the absence of such a preinjection energy analysis, the ^{127}I contribution can be minimized by sputtering with low-energy, typically 2 keV, Cs^+ ions. Systems which employ preacceleration to $\sim 100 \text{ keV}$ before the magnetic mass analysis enjoy an additional advantage. At this energy, an ^{127}I ion must acquire 0.8 keV from the Cs^+ beam in order to be injected, whereas at the $\sim 20 \text{ keV}$ characteristic of the smaller accelerators, only an additional 0.16 keV is required. Since the probability of an energy transfer δE decreases as $\sim 1/(\delta E)^2$, the number of injected $^{127}\text{I}^-$ ions is reduced by a factor of ~ 25 at the higher preacceleration energy.

For those systems employing a preinjection ESA, the only ions arriving at the detector with the correct charge state are ^{129}I . Lower-mass ions of the same m/q have significantly different energies and hence a detector with only modest energy resolution is sufficient.

Larger systems without a preinjection ESA generally operate at higher injection energy, and the postacceleration ESA or velocity filter is then usually sufficient to eliminate any remaining ^{127}I ions. Where this is not the case, a time-of-flight system may be employed to separate ^{129}I from those ^{127}I ions which elude the high-energy electrostatic analysis.

23.3.6 Calcium-41 ($T_{1/2} = 103$ ka)

A molecular negative ion is required for AMS of ^{41}Ca , owing to the very low binding energy of the atomic Ca^- ion. CaH_3^- has generally been the ion of choice, because the KH_3^- ion is not stable and hence the ^{41}K isobar is eliminated. Beam currents of several μA can be obtained from solid calcium hydride [36]. Production of calcium hydride is, however, a labor-intensive process, requiring in vacuo distillation of CaO to calcium metal.

Alternatively, the CaF_3^- ion may be used [37], since the KF_3^- ion is also unstable. The starting material is CaF_2 which is very much simpler to produce than CaH_2 . Simplicity of sample preparation is, however, offset by the disadvantages of lower beam currents, a more extensive suite of molecular fragments, and lower final energy, and most ^{41}Ca measurements to date have employed the trihydride ion.

23.3.7 Heavy Elements ($A > 180$)

Measurement of plutonium isotopes at very low concentrations in environmental samples has provided the principal impetus to the development of AMS techniques for very heavy elements. Other isotopes such as ^{236}U , ^{237}Np , $^{226,228}\text{Ra}$ and ^{182}Hf have also been explored.

In the case of plutonium, there are no stable isotopes and hence no macroscopic beam to which to measure a ratio. Instead, a “spike” of a few pg of ^{242}Pu is added to the sample prior to chemical processing, and the ratios of ^{239}Pu and ^{240}Pu to ^{242}Pu are determined by ion-counting all three isotopes.

Molecular PuO^- ions are selected for injection into the accelerator. In order to minimize scattering losses, the gas stripper is operated at a thickness of $\sim 0.2 \mu\text{g}/\text{cm}^2$. The operating voltage of the accelerator is generally limited by the bending power of the postacceleration analysis system. At the ANU for example, 24 MeV Pu^{5+} ions are at the limit of the $ME/q^2 = 210 \text{ MeV amu}$ analyzing magnet [38]. At Livermore, a 30° analyzing magnet permits a higher energy of 39 MeV [39]. Ionization chambers with an energy resolution of $\sim 3\%$ for these plutonium ions provide excellent discrimination against lower-mass ions of the same M/q arriving at the detector. Sensitivities of fewer than 10^6 atoms have been achieved for $^{239,240,242,244}\text{Pu}$ and ^{237}Np . For ^{239}Pu , this is two orders of magnitude better than α -particle counting.

With minor variations, the method is applicable to ^{237}Np and $^{226,228}\text{Ra}$. In the former case, ^{242}Pu is again used for normalization, although ^{236}Np would be preferable if it could be obtained. In the radium case, the yield of RaO^- ions is poor, and the RaC_2 negative ion is used instead.

Recently, it has been demonstrated that plutonium can be measured almost equally well with a small accelerator operating at only 300 kV. The stripping yield of 3^+ ions is surprisingly high, and a gas ionization chamber with a 50 nm silicon nitride window provides good discrimination between 1.2 MeV Pu^{3+} ions and any 2^+ ions with the same mass-to-charge ratio [17].

The presence of ^{236}U is a characteristic signature of uranium that has been through a nuclear reactor, and hence this isotope is of interest in environmental and nuclear-safeguards monitoring. The much more abundant ^{238}U and ^{235}U constitute more potent sources of background than for the isotopes above, and it is necessary to supplement the total-energy measurement with a time-of-flight measurement in order to eliminate them. Sensitivities in the $^{236}\text{U}/^{238}\text{U}$ ratio of $\sim 10^{-12}$ have been achieved, which are sufficient for determination of ^{236}U concentrations in natural uranium ores [40].

None of the above isotopes has a stable isobar. In contrast, AMS measurement of ^{182}Hf must contend with interference from the ^{182}W isobar. At the energies available from tandem accelerators, it is not possible to separate the two, and one must rely on a background subtraction based on measured counting rates of other tungsten isotopes. Clearly, the lower the ^{182}W contribution the better, and it has been shown that use of the HfF_5^- ion suppresses the tungsten by nearly four orders of magnitude [41].

23.3.8 Other Isotopes

Techniques have been developed for the measurement of several other isotopes, but none of these has yet found wide application. For completeness, each is considered briefly below.

^3H ($T_{\frac{1}{2}} = 12 \text{ a}$)

Decay counting is routinely used for determining tritium in water, and has high sensitivity. Nevertheless, AMS offers the advantage of smaller sample size and simpler sample preparation. At Rossendorf, a 3 MV Tandetron is used for depth profiling of tritium in carbon tiles from fusion reactor walls. Recently, this has been complemented by a small, SF_6 -insulated 100 kV accelerator [42]. In addition, a tritium capability is under development on the dedicated biomedical AMS system at the Lawrence Livermore laboratory.

^{32}Si ($T_{\frac{1}{2}} = \sim 140 \text{ a}$)

Silicon-32 is produced in the atmosphere by spallation of argon. Fallout is about 2 atoms/ m^2/s , i.e. about 10% of ^{36}Cl . Despite the comparatively short half-life of ^{32}Si , AMS offers the advantages over conventional decay counting of smaller sample size and simpler sample preparation. It has potential for dating ice in the 50–1000 year range in temperate-zone glaciers [43], and in biomedicine [24].

A gas-filled magnet is required to separate ^{32}Si from ^{32}S ions. Fluxes of the latter are typically $>10^7 \text{ s}^{-1}$. A discrimination factor of 10^{12} can be achieved with a combination of a gas-filled magnet and a suitable detector [24].

**^{53}Mn ($T_{\frac{1}{2}} = 3.8 \text{ Ma}$), ^{59}Ni ($T_{\frac{1}{2}} = 60 \text{ ka}$), ^{60}Fe ($T_{\frac{1}{2}} = 1.5 \text{ Ma}$)
and ^{63}Ni ($T_{\frac{1}{2}} = 100 \text{ a}$)**

A gas-filled magnet has been used successfully [44] to separate these four isotopes from their stable isobars. Sensitivities of $\sim 2 \times 10^{-14}$ in the $^{53}\text{Mn}/\text{Mn}$, $^{59}\text{Ni}/\text{Ni}$ and $^{63}\text{Ni}/\text{Ni}$ ratios have been achieved. For the more favorable case of ^{60}Fe , where the ^{60}Ni stable isobar differs by two in Z , sensitivity is significantly better at $\sim 2 \times 10^{-16}$.

Finally, where $^{59}\text{Ni}/\text{Ni}$ or $^{63}\text{Ni}/\text{Ni}$ ratios are $\sim 10^{-10}$ or higher as in nuclear waste, the technique of projectile X-ray emission may be employed (Sect. 23.2.2) [27, 45].

Studies of meteorites have been the principal applications of both ^{53}Mn and ^{59}Ni . Measurements of ^{60}Fe in ferromanganese crusts have provided evidence for a nearby supernova within the last 5 Ma [46]. Nickel-63, created by the $^{63}\text{Cu}(n, p)^{63}\text{Ni}$ reaction, is being used to check the fast-neutron dosimetry of the Hiroshima atomic bomb [47].

^{90}Sr ($T_{\frac{1}{2}} = 28.5 \text{ a}$)

Strontium-90 can be measured with high sensitivity using conventional decay counting, but AMS offers the advantage of a faster response in the event of a nuclear accident. Higher energies are required in order to be able to discriminate between ^{90}Sr and its ^{90}Zr isobar. Paul et al. [48], using a ^{90}Sr energy of 131 MeV, have reported the best sensitivity to date, corresponding to $^{90}\text{Sr}/\text{Sr} \sim 3 \times 10^{-13}$.

23.4 Applications

AMS has found application in many areas of science. In the following, a brief overview of its contribution to the most significant areas is presented.

23.4.1 Archaeology

AMS has largely supplanted liquid scintillation counting for radiocarbon dating. It offers higher throughput and smaller sample size with little or no compromise in precision. Together, these enhance the reliability of the dating by permitting more dates per site and by allowing dating of individual seeds or pieces of charcoal, for example, that are truly representative of the archaeological context. In addition, a more rigorous chemical precleaning of the sample is possible when only a milligram of carbon is sufficient.

Until recently, the expectation that the superior efficiency of the AMS technique would allow ^{14}C dating to be pushed back beyond the $\sim 50\,000$ year limit of conventional decay counting was frustrated by backgrounds due

to younger carbon that was added to a sample in the course of its history and during preparation. As the mechanisms of these processes have become better understood, however, this expectation is finally being realized [49,50].

A few illustrative examples of the application of AMS radiocarbon dating to archaeological finds are given below.

Cave and rock art is one area where the small-sample capability of AMS allows the ^{14}C dating of the paintings themselves via milligram amounts of charcoal, or other organic matter incorporated in the pigments, without significant damage to the art. Recent spectacular cave-art finds in France at Grotte Cosquer near Marseilles and Grotte Chauvet in the Ardeche have been dated to 19 to 27 ka and 31 ka, respectively [51,52], which are the earliest dates ever obtained for prehistoric paintings. Similar techniques are being applied to rock art in the Americas and Australia in order to shed new light on the antiquity and development of human occupation of these continents.

Another high-profile find was Ötzi, the Ice Man, whose well-preserved body was found in 1991 in the Ötztal Alps, South Tyrol, Italy. This remarkable find included clothes and shoes, a bow, a quiver of arrows, and a hand axe. AMS measurements on several artifacts, as well as on the body itself via bone and tissue specimens, place the date of his death at 4546 ± 17 radiocarbon years before present [53,54]. This radiocarbon age translates into a calendar age between 3100 and 3350 BC.

The famous Shroud of Turin was also dated by AMS. Three laboratories each received about 2 cm^2 of linen from the Shroud, and their concordant results [55] pointed to a medieval date (1290–1360 AD at 90% confidence) for the Shroud. This date is close to the year 1353 when the Shroud entered the historical record.

The raw datum from a radiocarbon measurement is the $^{14}\text{C}/^{12}\text{C}$ ratio. Since this ratio in atmospheric CO_2 has not been constant in time but has fluctuated in response to a number of factors, including solar activity and the geomagnetic field, the relationship between the measured ratio and the calendar age of the sample is not a simple one. Natural archives which can be precisely dated by other means are required in order to calibrate the radiocarbon timescale in terms of calendar time.

Tree rings, for which a continuous annual record is now available back to 12 400 years before present, have allowed the construction of a high-precision calibration curve from the late-glacial period to the present [56]. A section of the curve is shown in Fig. 23.5. Beyond the tree ring record, other archives are required. These include corals, which can be U/Th dated, annually laminated lake and marine sequences, and marine sediments with chronologies tied to “well-dated” Greenland ice cores via fluctuations in the ^{18}O to ^{16}O isotope ratio. A reasonable consensus exists back to ~ 26 ka, but the different archives diverge significantly between 26 and 40 ka.

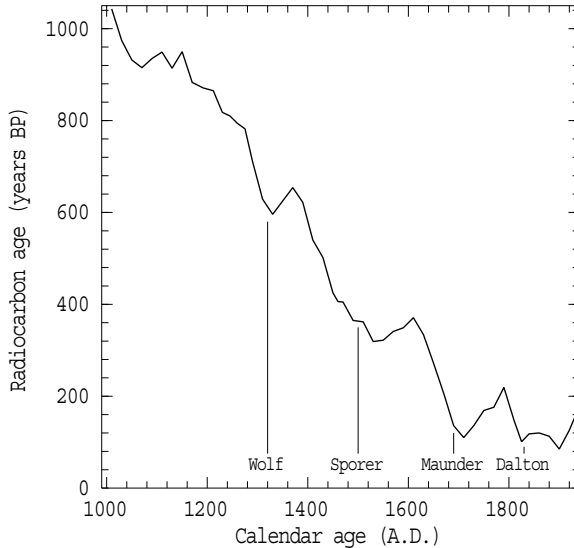


Fig. 23.5. The portion of the ^{14}C calibration curve from 1000 to 1920 AD. Minima in this curve, which correspond to periods of elevated ^{14}C production, line up with the well-known minima in sunspot activity. (After [56])

23.4.2 Exposure-Age Dating

Secondary cosmic rays, principally fast neutrons and muons, produce the long-lived isotopes ^{10}Be , ^{26}Al and ^{36}Cl in situ by interactions with suitable target nuclei in surface rocks. Provided that the production rates are known, the buildup of one or more of these isotopes may be used to determine the time of first exposure of the rock at the earth's surface. Direct dating of the advance and retreat of glaciers using either transported boulders or polished bedrock surfaces is providing essential paleoclimatic data for testing models of the earth's climate. Other geological processes or features such as meteorite impacts, fault movements, lava flows and landslides, and wave-cut platforms can also be dated. In addition, erosion rates can be determined for surfaces that have been exposed for long enough to have attained saturation to provide information on landscape evolution on 100 ka to 10 Ma year timescales. For a recent review of this rapidly expanding field, see Gosse and Phillips [11].

23.4.3 Ice Cores

Polar ice preserves a continuous record of the ^{10}Be and ^{36}Cl which are produced by cosmic-ray interactions in the atmosphere. Detailed studies of these isotopes are providing valuable information about past variations in solar activity, the strengths of the terrestrial and solar magnetic fields, and their links to climate [57].

Ice cores are also unique archives of anthropogenic ^{36}Cl . Nuclear-weapons testing in the late 1950s injected large quantities of ^{36}Cl into the stratosphere, where it was well mixed before falling out. Ice cores from high-accumulation sites such as Dye three in Greenland preserve an annual record of this fallout, which at its peak was three orders of magnitude above the normal cosmogenic rate [58] (see Fig. 23.6).

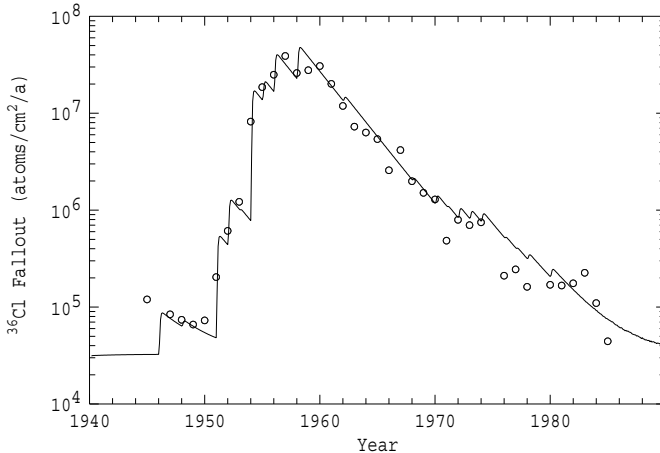


Fig. 23.6. The ^{36}Cl bomb pulse in an ice core from Dye 3, Greenland. The solid line is the result of a box-model calculation which incorporates the various test explosions (after [58])

23.4.4 Deep-Sea Cores

Cores from ocean sediments are providing a wealth of information about past climates, particularly over the 120 ka span of the most recent glacial cycle. Establishing an absolute chronology for these stages has, however, proved to be difficult. Fortunately, the most recent part of the record, which embraces the abrupt end about 15 ka ago of the last major glaciation, and the brief return of cold conditions about 11 ka ago during the Younger Dryas period, falls within the range of radiocarbon dating. AMS, because it permits the dating of individual species of foraminifera hand-picked from cores, has been instrumental in the construction of accurate chronologies spanning these changes [59].

23.4.5 Oceanography

A very large number of ^{14}C measurements have been performed by the National Ocean Sciences AMS facility at the Woods Hole Oceanographic Institute as part of the World Ocean Circulation Experiment (WOCE) [60].

Questions being addressed are the turnover and mean residence times of deep ocean water, mixing between basins, and the transfer of heat from low to high latitudes. Bomb-produced ^{14}C has provided valuable information about surface ocean circulation and atmosphere-ocean exchange processes, and here the amplitude of the signal is more than 20%. The bomb pulse has not yet penetrated into the deep basins of the world's oceans, and lateral gradients there are due to mixing of waters which were last exposed to atmospheric exchange at different times. Gradients are typically only 2 to 3% across an entire basin, and 0.3–0.4% precision is required. An AMS measurement needs only 0.5 l of water, which greatly simplifies sample collection and storage compared with the 200 l that were formerly required for decay counting.

Knowledge of present-day circulation in the Arctic Ocean is crucial to understanding heat transfer from low to high latitudes. Serendipitously, nuclear-fuel reprocessing plants at Sellafield in Cumbria (UK) and La Hague on the Cherbourg Peninsula (France) have, since 1954 and 1970 respectively, been potent point sources of ^{129}I and ^{99}Tc . Currents running through the English Channel and up the west coast of the UK have carried these isotopes into the Arctic Ocean. Surveys of ^{129}I levels, both in surface waters over an extended area and as a function of depth, are providing valuable insights into circulation patterns and vertical mixing in this important area [61], and work is under way to complement these with ^{99}Tc .

23.4.6 Biomedicine

Liquid scintillation counting of ^{14}C is very widely employed in biomedical research and in drug testing. AMS offers the major advantage of requiring very much smaller doses, thereby reducing both the expense of labeled compounds and the problems of disposal, and allowing studies in humans. The high throughput of AMS is also highly advantageous, since meaningful conclusions can only be drawn from studies which incorporate large numbers of subjects and controls. Pioneering work in this area has been directed towards measuring, at environmentally realistic doses, the rates at which known mutagens bind to DNA [62]. AMS facilities dedicated to biomedical applications have been constructed at the University of York (UK), Lawrence Livermore Laboratory and MIT. Recently, considerable progress has been made in coupling gas and liquid chromatographs directly to an AMS system to allow on-line compound-specific ^{14}C analyses [63].

23.4.7 Hydrology

Chlorine-36 is the principal AMS isotope used in hydrology. A useful review has been given by Fontes and Andrews [64]. It has found application in determining the age of groundwater, in measuring recharge rates, in studying past climates, and in investigating the hydrology of nuclear-waste analogues and potential waste disposal sites.

Extensive use of ^{129}I , often accompanied by ^{36}Cl , has also been made to investigate the origins and residence times of hydrothermal fluids and oil-field brines [65].

23.4.8 Extraterrestrial Material

AMS studies of long-lived isotopes such as ^{10}Be , ^{14}C , ^{26}Al , ^{36}Cl , ^{41}Ca and ^{59}Ni in meteorites have been directed towards establishing the nonterrestrial origin of a putative meteorite, determining the irradiation history of the parent body in space, and establishing the time at which a given meteorite fell to earth. A useful review has been given by [66].

23.5 Conclusions and Prospects

In the 26 years since the first demonstration that ^{14}C could be detected at natural levels using an electrostatic tandem accelerator, the field of accelerator mass spectrometry has expanded into many areas of science. Despite its maturity, growth continues at a rapid pace. There is a strong push towards smaller, cheaper systems; techniques for new isotopes are continually being developed; and new facilities are proliferating. In parallel, there is increasing sophistication in the area of applications, often driving improvements in technique.

References

1. D E Nelson, R G Korteling, W R Stott: *Science* **198**, 507 (1977)
2. C L Bennett, R P Beukens, M R Clover, H E Gove, R B Liebert, A E Litherland, K H Purser, W E Sondheim: *Science* **198**, 508 (1977)
3. K H Purser, R B Liebert, C J Russo: *Radiocarbon* **27**, 794 (1980)
4. D Elmore, B R Fulton, M R Clover, J R Marsden, H E Gove, K H Purser, L R Kilus, R Beukens, A E Litherland: *Nature* **277**, 22 (1979)
5. D Elmore, H E Gove, R Ferraro, L R Kilus, H W Lee, K H Chang, R P Beukens, A E Litherland, C J Russo, K H Purser, M T Murrell, R C Finkel: *Nature* **286**, 138 (1980)
6. J Klein, R Middleton, H Tang: *Nucl. Instr. Meth. Phys. Res.* **193**, 601 (1982)
7. R Middleton, J Klein, G M Raisbeck, F Yiou: *Nucl. Instr. Meth. Phys. Res.* **218**, 430 (1983)
8. R C Finkel, M Suter: *Adv. Anal. Chem.* **1**, 1 (1993)
9. C Tuniz, J R Bird, D Fink, G F Herzog: *Accelerator Mass Spectrometry: Ultrasensitive Analysis for Global Science* (Boca Raton, FL: CRC Press) (1998)
10. L K Fifield: *Rep. Prog. Phys.* **62**, 1223 (1999)
11. J Gosse, F M Phillips: *Quaternary Sci. Rev.* **20**, 1475 (2001)
12. C Bronk-Ramsey, R E M Hedges: *Nucl. Instr. Meth. Phys. Res. B* **123**, 539 (1997)

13. J R Southon, D E Nelson, J S Vogel: Nucl. Instr. Meth. Phys. Res. B **52**, 370 (1990)
14. K H Purser, T H Smick, R K Purser: Nucl. Instr. Meth. Phys. Res. B **52**, 263 (1990)
15. M Suter, S W A Jacob, H A Synal: Nucl. Instr. Meth. Phys. Res. B **172**, 144 (2000)
16. H-A Synal, M Dobeli, S Jacob, M Stocker, M Suter: Nucl. Instr. Meth. Phys. Res. B **223–224**, 339–345 (2004)
17. L K Fifield, H A Synal, M Suter: Nucl. Instr. Meth. Phys. Res. B **223–224**, 802–806 (2004)
18. M Suter, G Bonani, C Stoller, W Wölfl: Nucl. Instr. Meth. B **5**, 251 (1984)
19. G Bonani, P Eberhardt, H J Hofmann, T H Niklaus, M Suter, H A Synal, W Wölfl: Nucl. Instr. Meth. Phys. Res. B **52**, 338 (1990)
20. E Boaretto, D Berkovits, G Hollos, M Paul: Nucl. Instr. Meth. Phys. Res. B **50**, 280 (1990)
21. M Paul: Nucl. Instr. Meth. Phys. Res. B **52**, 315 (1990)
22. U Zoppi, P W Kubik, M Suter, H A Synal, H R von Gunten, D Zimmerman: Nucl. Instr. Meth. Phys. Res. B **92**, 142 (1994)
23. K Knie, T Faestermann, G Korschinek: Nucl. Instr. Meth. Phys. Res. B **123**, 128 (1997)
24. J Popplewell, S J King, J P Day, P Ackrill, L K Fifield, R G Cresswell, M L di Tada, K Liu: J. Inorg. Biochem. **69**, 177 (1998)
25. H Artigas, J L Debrun, L Kilius, X L Zhao, A E Litherland, J L Pinault, C Fouillac, C J Maggiore: Nucl. Instr. Meth. Phys. Res. B **92**, 227 (1994)
26. M J M Wagner, H A Synal, M Suter, J L Debrun: Nucl. Instr. Meth. Phys. Res. B **99**, 519 (1995)
27. J E McAninch, L J Hainsworth, A A Marchetti, M R Leivers, P R Jones, A E Dunlop, R Mauthe, J S Vogel, I D Proctor, T Straume: Nucl. Instr. Meth. Phys. Res. B **123**, 137 (1997)
28. D Fink, B McKelvey, D Hannan, D Newsome: Nucl. Instr. Meth. Phys. Res. B **172**, 838 (2000)
29. G M Raisbeck, F Yiou, D Bourles, J Lestringuez, D Deboffe: Nucl. Instr. Meth. Phys. Res. B **5**, 175 (1984)
30. R Middleton, J Klein: Phil. Trans. Roy. Soc. Lond. A **323**, 121 (1987)
31. D L Knies, D Elmore: Nucl. Instr. Meth. Phys. Res. B **92**, 134 (1994)
32. M L Roberts et al.: Nucl. Instr. Meth. Phys. Res. B **123**, 57 (1997)
33. H A Synal, J Beer, G Bonani, C Lukaszczuk, M Suter: Nucl. Instr. Meth. Phys. Res. B **92**, 79 (1994)
34. N J Conard, D Elmore, P W Kubik, H E Gove, L E Tubbs, B A Chrnyk, M Wahlen: Radiocarbon **28**, 556 (1986)
35. L R Kilius, N Baba, M A Garwan, A E Litherland, M-J Nadeau, J C Rucklidge, G C Wilson, X L Zhao: Nucl. Instr. Meth. Phys. Res. B **52**, 357 (1990)
36. D Fink, R Middleton, J Klein: Nucl. Instr. Meth. Phys. Res. B **47**, 79 (1990)
37. S P H T Freeman, J S Vogel: Mass Spectrom. Ion Proc. **143**, 247 (1995)
38. L K Fifield, A P Clacher, K Morris, S J King, R G Cresswell, J P Day, F R Livens: Nucl. Instr. Meth. Phys. Res. B **123**, 400 (1997)
39. J E McAninch et al.: Nucl. Instr. Meth. Phys. Res. B **172**, 711 (2000)
40. P Steier, R Golser, W Kutschera, V Liechtenstein, A Priller, A Valenta, C Vockenhuber: Nucl. Instr. Meth. Phys. Res. B **188**, 283 (2002)

41. C Vockenhuber, M Bichler, R Golser, W Kutschera, A Priller, P. Steier, S. Winkler: Nucl. Instr. Meth. Phys. Res. B **223–224**, 823–828 (2004)
42. M Friedrich, W Pilz, G Sun, R Behrisch, C Garcia-Rosales, N Bekris, R D Penzhorn: Nucl. Instr. Meth. Phys. Res. B **172**, 655 (2000)
43. U Morgenstern, L K Fifield, A Zondervan: Nucl. Instr. Meth. Phys. Res. B **172**, 605 (2000)
44. K Knie, T Faestermann, G Korschinek, G Rugel, W Ruhm, C Wallner: Nucl. Instr. Meth. Phys. Res. B **172**, 717 (2000)
45. P Persson, M Kiisk, B Erlandsson, M Faarinen, R Hellborg, G Skog, K Stenström: Nucl. Instr. Meth. Phys. Res. B **172**, 188 (2000)
46. K Knie, G Korschinek, T Faestermann, C Wallner, J Scholten, W Hillebrandt: Phys. Rev. Lett. **83**, 18 (1999)
47. T Straume et al. : Nature **424**, 539 (2003)
48. M Paul, D Berkovits, L D Cecil, H Feldstein, A Hershkowitz, Y Kashiv, S Vogt: Nucl. Instr. Meth. Phys. Res. B **123**, 394 (1997)
49. M I Bird, L K Ayliffe, L K Fifield, C S M Turney, R G Cresswell, T T Barrows, B David: Radiocarbon **41**, 127 (1999)
50. M J Nadeau, P M Grootes, A Voelker, F Bruhn, A Duhr, A Oriwall: Radiocarbon **43**, 169 (2001)
51. J Clottes, J Courtin, H Valladas, H Cachier, N Mercier, M Arnold: Bull. Soc. Préhistorique Française **89**, 270 (1992)
52. J Clottes, J M Chauvet, E Brunel-Deschamps, C Hillaire, J P Daugas, M Arnold, H Cachier, J Evin, P Fortin, C Oberlin, N Tisnerat, H Valladas: C.R. Acad. Sci. Paris **320**, 1133 (1995)
53. R E M Hedges, R A Housley, C R Bronk, G J van Klinken: Archaeometry **34**, 337 (1992)
54. G Bonani, S D Ivy, I Hajdas, T R Niklaus, M Suter: Radiocarbon **36**, 247 (1994)
55. P E Damon, D J Donahue, B H Gore, A L Hatheway, A J T Jull, T W Linick, P J Sercel, L J Toolin, C R Bronk, E T Hall, R E M Hedges, R Housley, I A Law, C Perry, G Bonani, S Trumbore, W Woelfli, F C Ambers, S G E Bowman, M N Leese, M S Tite: Nature **337**, 611 (1989)
56. M Stuiver, P J Reimer, J W Beck, E Bard, G S Burr, K A Hughen, B Kromer, F G McCormack, J van der Plicht, M Spurk: Radiocarbon **40**, 1041 (1998)
57. J Beer: Space Sci. Rev. **11**, 53 (2000)
58. H A Synal, J Beer, G Bonani, M Suter, W Wölfli: Nucl. Instr. Meth. Phys. Res. B **52**, 483 (1990)
59. K A Hughen, J R Southon, S J Lehman, J T Overpeck: Science **290**, 1951 (2000)
60. K F Von Reden, A P McNichol, J C Peden, K L Elder, A R Gagnon, R J Schneider: Nucl. Instr. Meth. Phys. Res. B **123**, 438 (1997)
61. G M Raisbeck, F Yiou: Sci. Total Environ. **237/238**, 31 (1999)
62. J S Vogel: Nucl. Instr. Meth. Phys. Res. B **172**, 884 (2000)
63. P L Skipper, B J Hughey, R G Liberman, M H Choi, J S Wishnok, R E Klinkowstein, R E Shefer, S R Tannenbaum: Nucl. Instr. Meth. Phys. Res. B **223–224**, 740–744 (2004)
64. J C Fontes, J N Andrews: Nucl. Instr. Meth. Phys. Res. B **92**, 367 (1994)
65. J E Moran, U Fehn, J S Hanor: Geochim. Cosmochim. Acta **59**, 5055 (1995)
66. G F Herzog: Nucl. Instr. Meth. Phys. Res. B **92**, 478 (1994)

24 Atomic Collisions in Matter

J. Keinonen

Department of Physical Sciences, Accelerator Laboratory, P.O. Box 43, 00014
University of Helsinki, Finland

24.1 Introduction

Energetic ions (in the energy range from the eV region to the MeV region considered in this chapter) collide with electrons and nuclei when penetrating into materials. In solid matter, the energy loss of ions is connected to the first-order effects on the atoms of the material, particularly the electronic excitation and displacement of lattice atoms, and the production of plasmons and phonons. The trajectory of an ion is determined by successive inelastic and elastic binary encounters with the lattice atoms. The implantation of ions results in a distribution of stopped ions and damage. In addition to the ion irradiation and implantation of materials, atomic collisions are utilized in methods of ion beam analysis of materials, such as Rutherford backscattering spectrometry (RBS) and elastic recoil detection analysis (ERDA).

In this chapter, a basic formalism is used to illustrate collisions of ions with atoms, i.e. their nuclei and electrons. Then some phenomena in atomic collisions for ions penetrating into materials are described. Finally, effects in materials caused by the impinging ions are discussed. Computer methods used to calculate the slowing down of ions, and interatomic potentials defining the force between nuclei of the ion and target atom are described in two appendixes.

24.2 Slowing Down of Energetic Ions

The stopping of energetic heavy ions in matter has received much theoretical and experimental interest for decades. Bohr's theory [1] of the stopping of charged particles in matter was extended by him [2] to point out the importance of screening due to projectile electrons in the slowing down of fission fragments. Since the appearance of the theory by Lindhard, Scharff, and Schiøtt (LSS theory) [3], heavy-ion stopping has commonly been divided into three velocity regions. At low velocities, the stopping force is taken to be proportional to the ion velocity [4] and was described originally by the theories of Firsov [5] and Lindhard and Scharff [6]. At high velocities, the basis for stopping calculations is the Bethe formula [7]. At intermediate velocities around the stopping maximum, the stopping force is characterized

by a Bethe-type formula in conjunction with an effective ion charge. Brandt and Kitagawa [8] established an explicit connection between the ion charge and stopping force. The book by Ziegler, Biersack, and Littmark [9] covers the physical phenomena and history associated with the penetration of energetic ions into solids. More recently, the book edited by Smith et al. [10] has covered the theory, simulation, and applications of atom and ion collisions in solids and at surfaces.

Comparisons with experimental values have indicated that the predictions of the analytical models are not sufficiently accurate. Some recent models improving the accuracy of stopping calculations with numerical methods are the binary theory [11], the generalization of the Friedel sum rule [12] for non-crystalline targets, and the convolution approximation [13] for crystalline targets.

Analytical models of particle penetration rely on the use of statistical-physics methods and transport equations [14] for the slowing down of an ion in a homogeneous material. Successive collisions are assumed to be statistically independent. Since these equations can be solved only for a limited set of environments, computer simulations based on the binary-collision approximation (BCA) (for example in [9]) and molecular-dynamics (MD) techniques (see [10]) have become widely used; for the techniques, see Appendix 24.A.

In the basic description of the slowing down of energetic ions produced in an accelerator and forming a beam with an incident flux density n , the ions impinging on solid matter are scattered from atoms. For each energy E of the ion and energy T transferred to the atom, the differential of the scattering cross section $\sigma(E, T)$ gives the fraction of ions n_Ω scattered into a solid angle $d\Omega$ from the incident flux, i.e. $d\sigma(E, T) = \sigma(E, T) d\Omega = n_\Omega/n$.

In a medium where atoms are distributed randomly with a density N , the cross section $d\sigma(E, T)$ defines the probability $dP(E, T)$ for a collision, i.e. $dP(E, T) = N\Delta z d\sigma(E, T)$, where Δz is the distance traversed by the ion. In a number of collisions, the average energy loss $\langle\Delta E\rangle$ is given by

$$\langle\Delta E\rangle = \int T dP = -N\Delta z \int_{T_{\min}}^{T_{\max}} T d\sigma, \quad (24.1)$$

where $T_{\min}$ and $T_{\max}$ are the minimum and maximum transferred energy, respectively. This equation defines the nuclear stopping power $(dE/dx)_n$ and the stopping cross section $S_n(E)$:

$$\left(\frac{dE}{dx}\right)_n = -N \int T d\sigma = -NS_n(E). \quad (24.2)$$

In addition to the energy loss in elastic collisions with atoms, the ions lose energy also in collision with electrons, in inelastic scattering of ions. This electronic stopping power is written in the form

$$\left(\frac{dE}{dx}\right)_e = -NS_e(E). \quad (24.3)$$

If the nuclear and electronic stopping powers are independent, the total stopping power is

$$\frac{dE}{dx} = -N[S_n(E) + S_e(E)] . \quad (24.4)$$

The two main problems in the stopping theory are the calculation of $\sigma(E, T)$, and of the energy loss spectrum from that.

The nuclear and electronic stopping powers are illustrated in Fig. 24.1, where the stopping power is shown as a function of ion velocity for silicon ions slowing down in silicon. The crystalline structure is not taken into account in the calculations.

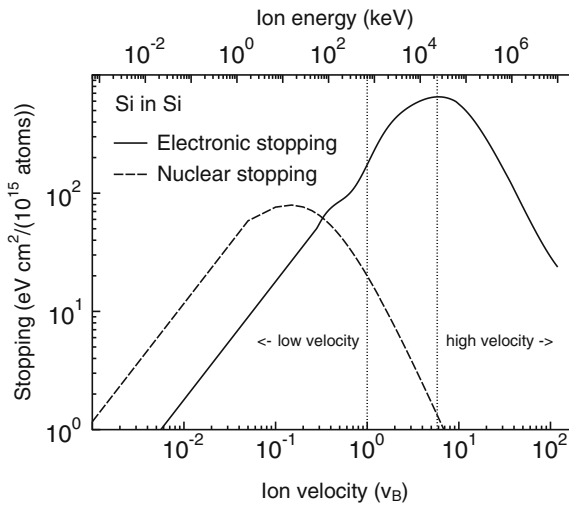


Fig. 24.1. Stopping powers for silicon ions slowing down in silicon as a function of the ion velocity. The velocity is given in the units of Bohr velocity v_B , namely the fine-structure constant of the atom times the velocity of light in vacuum. The stopping powers were calculated according to the empirical parametrization by Ziegler, Biersack, and Littmark by using the program SRIM [9]. The maximum of the electronic stopping power occurs at much higher velocities than that of the nuclear stopping. With decreasing velocity, the electronic stopping decreases before the nuclear stopping. This leads to the fact that for heavy ions, the nuclear stopping power dominates at low velocities

24.3 Collision of an Ion with an Atom

The basis for the description of collisions between impinging ions (atomic number Z_1 , mass m_1) and electrons or nuclei of matter (atomic number Z_2 , mass m_2) is the binary-collision approximation [9, 10]. A collision of an ion with a target atom and electron is described as a scattering in a central force field $F(r) = -\partial V(r)/\partial r$, where $V(r)$ is the central potential and r is the

distance between the ion and the target atom or electron. The basis of the interaction is the Coulomb potential $V(r) = Z_1 Z_2 e^2 / r$. According to classical mechanics, the center-of-mass scattering angle θ is given by

$$\tan \frac{\theta}{2} = \frac{b}{2p}, \quad (24.5)$$

where p is the impact parameter (the distance of the extrapolated initial ion trajectory from the atom) and b is the collision diameter, defined by $Z_1 Z_2 e^2 / b = MV^2 / 2$ for the reduced mass $m = m_1 m_2 / (m_1 + m_2)$. If the target atom is assumed to be at rest initially, the energy loss deduced from conservation of energy and momentum is

$$T = T_{\max} \sin^2 \frac{\theta}{2}, \quad (24.6)$$

where $T_{\max} = [4m_1 m_2 / (m_1 + m_2)^2] E \equiv \eta E$. The differential scattering cross section is now $d\sigma = \pi (b/2)^2 (T_{\max} / T^2) dT$, and the average energy loss according to (24.1) is

$$\langle \Delta E \rangle = N \Delta z 2\pi \frac{(Z_1 e)^2 (Z_2 e)^2}{(m_2 v)^2} \ln \frac{T_{\max}}{T_{\min}}, \quad (24.7)$$

where $m_2 = m_n$, the charge is $Z_2 e$, and $N = N_n$ for the collisions with nuclei, and $m_2 = m_e$, the charge is $-e$, and $N = Z_2 N_n$ for the collisions with electrons. Equation (24.7) is based on the free-body scattering for ion velocities higher than the Bohr velocity v_B . It implies that for the same ion velocity, the electronic energy loss is much higher than the nuclear energy loss, i.e. $\langle \Delta E \rangle_e / \langle \Delta E \rangle_n \approx m_n / Z_2 m_e \gg 1$, and that the energy loss is proportional to v^{-2} , i.e. E^{-1} .

24.3.1 Elastic Collisions

The calculation of the energy loss in elastic collisions is complicated owing to the electron clouds screening the ion and target nuclei. The interaction potential is the product of a spherically symmetric screening function and the Coulomb potential Q_1 / r . There are many different formulations for the screening function [9]. See Appendix 24.B for interatomic potentials. Using an appropriate screening function, the nuclear stopping of any ion in any target material can be evaluated analytically [15].

For a potential V , center-of-mass energy E_c , and impact parameter p , the final angle of scatter is obtained from the scattering integral:

$$\theta = \pi - 2p \int_{r_{\min}}^{\infty} r^{-2} \left[1 - \frac{V(r)}{E_c} - \frac{p^2}{r^2} \right]^{-1/2} dr. \quad (24.8)$$

For an ion at energy E_0 , the average energy transferred, namely the nuclear stopping cross section, is obtained by summing over all impact parameter values see (24.1) and (24.2):

$$S_n(E_0) = 2\pi\eta E_0 \int_0^{p_{\max}} \sin^2 \frac{\theta}{2} p \, dp. \quad (24.9)$$

In computer simulations using the BCA and MD techniques, an explicit analytical form for S_n is not needed. The nuclear energy loss is obtained directly from the interaction potential between the ion and the atoms.

The impinging ions produce recoil atoms, which produce new recoils and so forth. The energy lost in the collisions goes into heat and deformation of the material in which the slowing down occurs, as the atoms spread the kinetic energy in a series of collisions, namely in collision cascades.

24.3.2 Inelastic Collisions

Electronic stopping is the main source of energy loss for ions moving faster than v_B (see Fig. 24.1). There are several mechanisms contributing to the electronic stopping at ion velocities where relativistic effects can be neglected: (s1) momentum exchange in a collision between the ion and a free electron in the target material, (s2) ionization of the ion, (s3) capture of an electron by the ion, (s4) excitation of the ion, (s5) excitation of a target atom, (s6) ionization of a target atom, and (s7) collective effects such as polarization and plasmon excitation.

The electronic stopping takes place both during the collision of the ion and atom and between these collisions. Between the collisions, a constant slowing force acts on the ion owing to momentum exchange with the electrons in the material (s1). The stopping during a close collision of the ion and atom is connected to electron exchange between them (s4 and s5). In addition to these frozen-charge energy losses, there are charge-exchange events (s2, s3, and s6). The relative importance of the different contributions to the total electronic stopping power depends on the ion velocity [3]. There are three velocity regions: the high-velocity region for $v > v_B Z_1^{2/3}$; the low-velocity region for $v < v_B$, where the upper limit is the Bohr velocity of the target electrons or, in the electron gas theory, the Fermi velocity v_F ; and the intermediate-velocity region for $v_B < v < v_B Z_1^{2/3}$. The velocity $v_B Z_1^{2/3}$ is the mean velocity of the electrons filling the levels of a neutral atom with nuclear charge Z_1 , as obtained from the Thomas–Fermi statistical theory [15].

For high ion velocities, Bohr calculated the stopping cross section for electron scattering from a moving ion (see [16]). A quantum-mechanical model by Bethe [7, 17] is used for a fully stripped ion. In collisions with single electrons, the energy transferred to an electron $\Delta E(p)$, corresponding to electron excitation, results in the electronic stopping cross section

$$S_e(E_0) = 2\pi \int_{p_{\min}}^{p_{\max}} \Delta E(p) p \, dp = 4\pi \frac{Z_1^2 e^4}{m_e v_0^2} \int_{p_{\min}}^{p_{\max}} \frac{dp}{p}, \quad (24.10)$$

according to (24.1) and (24.3). A calculation based on the Born approximation gives the stopping power [7]

$$-\left(\frac{dE}{dx}\right)_e = \frac{4\pi Z_1^2 e^4 N_n Z_2}{m_e v^2} \ln \frac{2m_e v^2}{I}, \quad (24.11)$$

where I is the mean excitation energy for the atom in its ground state.

Bethe's model [7] and correction terms arising from the shell corrections [18,19], the Bloch correction [20], the polarization effect [21] (or Barkas effect or Z_1^3 contribution), and relativistic effects [22] give an extensively used analytical description of the stopping. Owing to the needs of ion beam analysis, the electronic stopping power has been much studied at high ion velocities. The studies have resulted in accurate stopping tables for many ions in different materials.

At low velocities, the electronic stopping power involves mainly the mechanisms s1, s4, and s5. The most important one is s1 [15]. The theoretical description has been improved during the last two decades by nonlinear electron gas models. The main reason for the interest in the stopping power in this velocity region is the extensive use of ion implantation in the semiconductor industry.

The early model by Firsov [5] described the electronic stopping by the local transfer of energy from electrons of the ion to electrons of the atom. The retarding force acting on the ion leads in a change of the momentum. The electrons return when the ion moves away, but there is no back transfer of momentum because the electrons fall into higher energy levels. The electronic stopping cross section for media where atoms are located randomly is obtained from the energy transferred in one interaction between two atoms by integrating over all possible impact parameters, as in (24.10).

In binary collisions, the scattering angles are affected by the inelastic energy loss. The scattering angle is an average of the scattering angles before and after the collision. The scattering angle before the collision is given by (24.8), and after the collision by a similar integral but with a reduced energy and impact parameter [10].

In modern electronic-stopping calculations, the medium in which slowing down occurs is described as an electron gas or plasma with a constant density, and the ion is a perturbation in the gas [9]. The semiclassical analysis of the perturbation model uses Poisson's equation for a charge interacting with a polarizable medium characterized by a dielectric function [15]. The stopping integrals in the dielectric formalism [23] are calculated using Lindhard's linear approximations for the dielectric function both for slow and fast ions.

In the model by Lindhard and Scharff [6], the energy loss of an ion is proportional to the velocity:

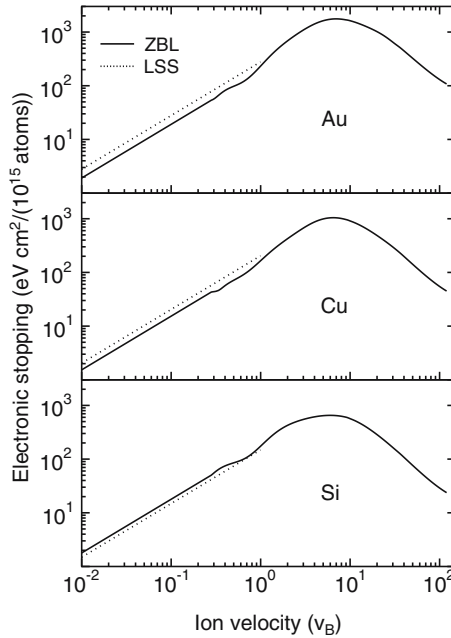


Fig. 24.2. Electronic stopping powers for silicon ions slowing down in silicon, copper, and gold. The LSS values have been calculated only up to v_B . For the ZBL values, see [9]

$$S_e(E) = \xi_e 8\pi e^2 a_B \frac{Z_1 Z_2}{Z} \frac{v}{v_B}, \quad (24.12)$$

where $Z = (Z_1^{2/3} + Z_2^{2/3})^{3/2}$ and $\xi_e \approx Z_1^{1/6}$ (see Fig. 24.2). The first general model for the nuclear and electronic stopping of an ion was presented in 1963 by Lindhard et al. [3], who based their treatment of atoms on the statistical Thomas–Fermi atom model. The LSS model gives analytical formulas for the nuclear and the electronic stopping of any ion in any target material in the low-velocity region.

Most of the theories for electronic stopping at low energies use the local-density approximation. It is assumed that each volume element in the solid is an independent, free plasma with an electron density. The electronic stopping power is position-dependent and proportional to the velocity of the ion; $S_e = \int I(v, \rho) Z_1^*(v)^2 \rho dV$, where $I(v, \rho)$ is the stopping interaction function of an ion of unit charge, $\rho(r)$ is the electron gas density or the electron density of the material in which slowing down occurs, and $Z_1^*(v)$ indicates that the charge of the ion differs from the atomic number and depends on the velocity. Various expressions for the electronic stopping power depend on the description of the plasma in the solid (e.g. [24]).

In the intermediate-velocity region, ions lose and capture electrons. This velocity region is the most difficult one to describe theoretically, because all the mechanisms s1–s7 have to be evaluated. The heavier the ion is, the more electrons participate in the exchange processes. The stopping power depends strongly not only on the ion velocity but also on the atomic numbers of the ion and target atoms [25]. The stopping power is affected by the fact that the quantum states of the electrons are different for different atoms. The electron capture and loss processes (s2 and s3) can have a notable contribution to the stopping.

By using an empirical scaling rule, the stopping power for a heavy ion (HI) is obtained from the stopping for a proton (H) at the same velocity. The stopping power is factorized into the electronic stopping for the proton and the effective charge of the heavy ion:

$$S_{\text{HI}} = S_{\text{H}}(Z_{\text{HI}}^*)^2 = S_{\text{H}}(Z_{\text{HI}}\gamma)^2, \quad (24.13)$$

where Z_{HI} is the atomic number of the heavy ion and γ is its fractional effective charge. The fractional effective charge is defined such that the effective charge of the ion $Z_1^*(v, Z_2)$ at a velocity v in the medium (Z_2) is $\gamma = Z_1^*(v, Z_2)/Z_1(v, Z_2)$. The fractional charge of a proton is equal to one. Many different Z_1 - and Z_2 - dependent formulations have been proposed in the literature for the heavy-ion fractional effective charge [9, 26–30]. Sigmund and Schinner emphasize [11] that the significance of static projectile screening is heavily overestimated in conventional effective-charge theory.

The analytical models have typically predicted the stopping power by interpolating between the Bethe formula above and linear models below the stopping maximum. The model by Brandt and Kitagawa [8] provides a description frequently used for the stopping power of a partially ionized heavy ion. According to the BK model, the energy loss increases for small impact parameters owing to reduced screening.

For the energy loss of intermediate-velocity heavy ions in electronic plasma, Ziegler, Biersack, and Littmark [9] constructed a model (ZBL) based on the ideas of the BK model [8]. The stripping of electrons in a heavy ion is calculated by comparing the ion's electron velocities with the relative velocity v_r between the ion and the electronic velocity of the medium. For a known charge state, the electron distribution of the ion $\rho(r) = (N/4\pi\Lambda^3)(\Lambda/r)\exp(-r/\Lambda)$ is defined, with a screening length $\Lambda = 2a_{\text{B}}(1-q)^{2/3}/[Z_1^{1/3}(1-(1-q)/7)]$. N is the number of electrons remaining in the ion. The fractional ionization $q = (Z_1 - N)/Z_1$ was deduced for heavy ions in an extensive data analysis to be $q = 1 - \exp \sum [a_i(v_r/v_{\text{B}}Z_1^{2/3})]^{b_i}$, where the constants a_i and b_i were obtained by fitting experimental data. The effective charge to be used in the calculation of the stopping power according to (24.13) is obtained from the fractional ionization

$$\gamma = q + 0.5(1-q) \left(\frac{v_{\text{B}}}{v_{\text{F}}} \right)^2 \ln \left[1 + \left(2 \frac{\Lambda v_{\text{F}}}{a_{\text{B}} v_{\text{B}}} \right)^2 \right]. \quad (24.14)$$

The ZBL parametrization predicts empirically the stopping of any ion in any target material (see Fig. 24.2). Since the Fermi velocity has a constant value for each target material, the ion velocity and atomic number are the only parameters in the stopping function. Owing to several fitted parameters, the predicted electronic stopping powers are reasonably accurate, but the nonlocality limits the validity and accuracy for stopping calculations in crystalline structures in cases where channeling of ions (see Sect. 24.4.5) takes place.

24.4 Atomic Collisions in Matter

24.4.1 Energy Straggling

Ions are deflected and slowed down in matter by scattering from electrons and nuclei of substrate atoms. In implantation and irradiation experiments, there are many ion trajectories involved. The transport equations are the basis for analytical descriptions of ion penetration in solids. If an ion bombardment has a distribution of ions $F(\mathbf{r}, \mathbf{v}, t)$, where $\mathbf{r}$ and $\mathbf{v}$ are the position and velocity vectors, respectively, of the ions at time t , the forward form of the transport equations, namely the Boltzmann equation, is [10]

$$\begin{aligned} \frac{\partial F(\mathbf{r}, \mathbf{v}, t)}{\partial t} = & -\mathbf{v} \cdot \nabla F(\mathbf{r}, \mathbf{v}, t) - N\mathbf{v}F(\mathbf{r}, \mathbf{v}, t) \int d\mathbf{v}' \sigma(\mathbf{v}, \mathbf{v}') \\ & + N \int \mathbf{v}' d\mathbf{v}' F(\mathbf{r}, \mathbf{v}', t) \sigma(\mathbf{v}', \mathbf{v}) + \mathcal{S}(\mathbf{r}, \mathbf{v}, t). \end{aligned} \quad (24.15)$$

$F(\mathbf{r}, \mathbf{v}, t) d\mathbf{r} d\mathbf{v}$ is the probability to find an ion in the volume $(\mathbf{r}, d\mathbf{r})$ moving with velocity $(\mathbf{v}, d\mathbf{v})$ at time t , $\sigma(\mathbf{v}, \mathbf{v}')$ is the cross section for scattering from $\mathbf{v}$ to $(\mathbf{v}', d\mathbf{v}')$, and $\mathcal{S}(\mathbf{r}, \mathbf{v}, t)$ represents the distribution of ions arriving in the sample.

For the case where the ions pass through a thin foil with small energy losses and negligibly small deflections, the one-dimensional function $\phi(z, E)$ representing the number of ions traversing the foil, of thickness z , with a constant velocity E_0 is

$$\begin{aligned} \frac{\partial \phi(z, E)}{\partial z} = & -\phi(z, E)N \int dT \sigma(E, T) + N \int dT \sigma(E + T, T)\phi(z, E + T) \\ & + \phi_0 \delta(z)\delta(E - E_0), \end{aligned} \quad (24.16)$$

where the first integral refers to the energy before the collision, the second integral refers to the final energy, and the last term is the energy distribution of the incoming ions.

For a strongly peaked scattering cross section at a small energy loss, i.e. $\sigma(E + T, T) = \sigma(E, T)$, the integral can be given to first order, where the stopping is $S(E) = \int dT T \sigma(E, T)$ and the straggling is

$$\Omega^2(E) = \int dT T^2 \sigma(E, T) , \quad (24.17)$$

by the equation

$$\frac{\partial \phi(z, E)}{\partial z} = NS(E_0) \frac{\partial \phi(z, E)}{\partial E} + \frac{N \Omega^2(E_0)}{2} \frac{\partial^2 \phi(z, E)}{\partial E^2} . \quad (24.18)$$

The solution with the boundary condition $\phi(0, E_0) = \phi_0 \delta(E - E_0)$ results in the Gaussian approximation describing the spectra of ions traversing a thin film and losing an energy $\Delta E = E_0 - E$,

$$\phi(z, \Delta E) = \frac{\phi_0}{[2\pi z N \Omega(E_0)]^{1/2}} \exp \left(-\frac{[\Delta E - z N S(E_0)]^2}{2 z N \Omega(E_0)} \right) . \quad (24.19)$$

After traversing a distance Δz in matter and losing energy as given by (24.7), the ions have a straggling

$$\Omega_{n/e}^2 = N \Delta z 4\pi (Z_1 e)^2 (Z_2 e)^2 \frac{m_{\text{ion}}^2}{(m_{\text{ion}} + m_{n/e})^2} \left(1 - \frac{T_{\text{min}}}{T_{\text{max}}} \right) \quad (24.20)$$

owing to elastic collisions with nuclei (n) and inelastic collisions with electrons (e). For heavy ions, the nuclear contribution is sizable.

In the region of high ion velocity, the straggling is almost independent of projectile velocity and is given by the formula derived by Bohr [31, 32],

$$\Omega_B^2 [\text{keV}^2] = 0.26 Z_1^2 Z_2 N \Delta z [10^{18} \text{at./cm}^2] . \quad (24.21)$$

This straggling formula has been improved by many authors to extend its applicability to lower velocities of light ions. Lindhard and Scharff [33] proposed a correction for low ion velocities. Yang et al. [34] summarized the development of straggling calculations [35–44], undertook a survey of H, He, and heavy-ion straggling data, and developed a fitting function for the Chu model [36]. By use of the effective charge and a scaling approach for energy straggling and considering correlation and charge-exchange effects, functions were obtained for heavy-ion straggling.

24.4.2 Multiple Scattering

In collisions with nuclei and electrons, the ions undergo scattering, as demonstrated in Fig. 24.3. The average deflection is zero because of symmetry,

but the average square deflection is finite and results in an angular divergence of the ion beam. The angle of scattering θ given in center-of-mass coordinates in (24.5) transforms, in laboratory coordinates, to $\phi = \tan^{-1}[m_2 \sin \theta / (m_1 + m_2 \cos \theta)]$. The average square deflection for small scattering angles, analogously to (24.17), is

$$\langle \Phi^2 \rangle = N \Delta z \int \phi^2 d\sigma, \quad (24.22)$$

where the integral goes from zero to a maximum value of the angle. If the probability of wide-angle scattering is assumed to be negligible, the angular spread of the ions is related directly to the stopping power, i.e. $\langle \Phi^2 s \rangle \approx N \Delta z (m_2/m_1 E) S(E)$. The angular distribution of the ion beam can be approximated by a Gaussian form,

$$F(\Phi) = \frac{1}{(2\pi\langle\Phi^2\rangle)^{1/2}} \exp\left(-\frac{\Phi^2}{2\langle\Phi^2\rangle}\right). \quad (24.23)$$

These equations have to be modified to take into account the large-angle scattering due to the nuclear stopping power. Early theoretical work on small-angle scattering of ions has been summarized in [45]. To meet the needs of experiments, the theory was further developed by Sigmund et al. [46–48] and, very recently, in [49].

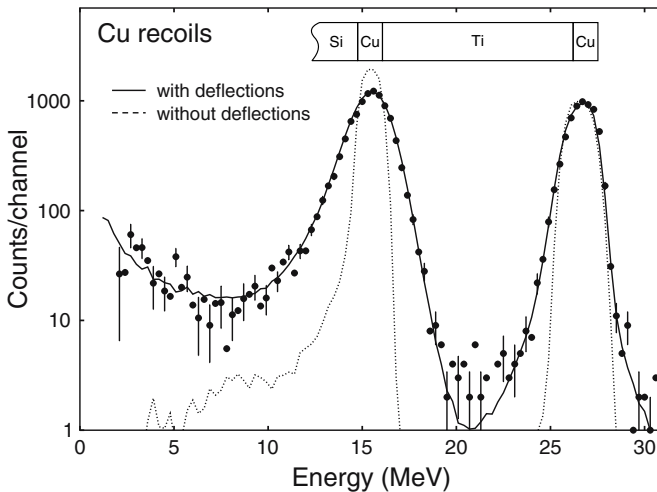


Fig. 24.3. Multiple scattering of Cu recoils in titanium as observed in TOF-ERDA measurements [60]. The Cu recoils were produced by bombardment of 23 nm thick Cu marker layers on the surface of a 200 nm thick Ti foil and at the interface between the Ti foil and Si substrate by 53 MeV $^{127}\text{I}^{10+}$ ions. The dots are experimental values. The simulated spectrum without scattering but with energy straggling is indicated by “without deflections”, and the simulated spectrum with the scattering by “with deflections”

24.4.3 Z_1 and Z_2 Oscillations

In comparisons of experimental stopping data with the predictions of the theories of Firsov [5] and Lindhard and Scharff [6], the experimental data showed an oscillatory structure in the stopping power vs. Z_1 [50–54]. The Z_1 dependence is enhanced under channeling conditions, under which the moving particle experiences glancing collisions with the lattice atoms at an almost constant impact parameter. A number of investigations have dealt with models with the aim of explaining the oscillatory structure, including [50–56].

The electron shells were explained to be the reason for the fluctuation of the stopping of an ion (Z_1) as a function of the material in which slowing down occurs (Z_2) and for the fluctuation of the stopping in a given material for different ions [56–59]; these functions are called Z_2 and Z_1 oscillations, respectively (see Fig. 24.4).

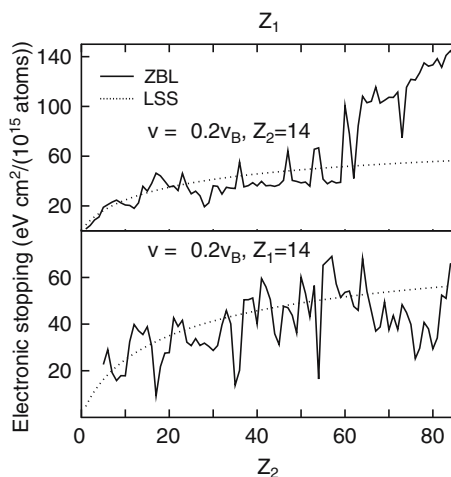


Fig. 24.4. Electronic stopping powers in different materials illustrating the Z_1 and Z_2 oscillations; for the ZBL values, see [9]

According to the model by Echenique et al. [56], the Z_1 oscillations can be taken into account by including a momentum-transfer cross section σ_{tr} in the equation for the low-energy stopping power of a free-electron gas $S_e(v) = nvv_F\sigma_{tr}(v_F)$, where nv is the uniform current of electrons scattered by a screened potential and $v_F\sigma_{tr}$ is the integrated scattering rate. The final expression for the stopping power depends on the approximation used for σ_{tr} . For a spherically symmetric scattering potential the stopping cross section is [56]

$$S_e(v) = \frac{3v}{k_F r_s^3} \sum_{l=0}^{\infty} (l+1) \sin^2(\delta_l(E_F) - \delta_{l+1}(E_F)) , \quad (24.24)$$

where $r_s = [3/(4\pi n)]^{1/3}$ is a measure of the electron density defined as the radius of a sphere whose volume is equal to the volume per conduction electron, and $\delta_l(E_F)$ represents the scattering phase shifts of electrons scattered by the effective potential of the atom for the l th partial wave in the expansion of the electron wave function. Echenique and coworkers defined, in an operational manner, an effective charge $Z_1^* = [(dE/dx)_{Z>1}/(dE/dx)_{\text{proton}}]^{1/2}$. Using the density-functional approach and the effective charge, they succeeded in scaling the electronic stopping powers for $Z_1 \geq 2$ particles to that of protons.

When the electronic stopping for channeled ions is plotted versus the atomic number of the target material Z_2 , it shows oscillations similar to those for Z_1 . On the basis of experimental data, Brandt and Kitagawa [8] concluded that a comprehensive description of low-velocity electronic stopping powers could be given if reference is made not to Z_2 but to the valence electron density.

24.4.4 Compound Materials

Bragg's rule [61] states that the stopping power of a compound target is the weighted average of the atomic stopping powers of the constituents $S = \sum a_i S_i$, where the sum of the weights a_i describing the fractions of the atoms in the compound is equal to one. Deviations from this rule occur owing to differences in the electronic structure between a free atom and an atom bound in a molecule or an alloy. Consequently, deviations from the rule should be expected at low energies, where the relative contribution from valence electrons to the stopping power is large, and for very light elements, where the valence electrons constitute a major fraction of the total number of electrons. The deviations are most pronounced, about 10–20%, around the stopping maximum for light organic gases and for solid compounds containing heavier constituents, such as oxides and nitrides. In an empirical model [62], two contributions are assumed, the effect of the closed electron shells of atoms and the effect of the chemical bonds. The average accuracy of the calculations obtained was better than 2% for compounds with known bond structures.

24.4.5 Effects of the Crystalline Structure of Matter

The crystalline structure of the target material has a large effect on the stopping [63, 64]. The periodicity in the structures of crystalline materials forms channels for the impinging ions. The atom and electron densities have lower values in the channels than in the bulk. The electron density can be several decades lower in the middle of a channel than near a nucleus. Thus the contribution to the stopping from the bound electrons is less important

in channels. The low atom and electron densities lead to lower nuclear and electronic stopping powers, respectively, in the channels than elsewhere in the bulk.

There are two different approaches to calculating the electronic stopping in channels, namely methods based on the binary-collision approximation (BCA), with the impact parameter as a variable, and MD calculations, using the local electron density as the variable. The benefit of BCA calculation is that the different stopping contributions are well separated by the use of cross sections and are straightforward to calculate. MD calculations have the advantage of using an accurate local electron density inside the material and the possibility of including many-body collisions.

Stopping powers in channels have been calculated by MD simulations using models based on the BK theory [65,66] and spherical charge densities for the target atoms. The effective charge of an ion was calculated using a fitted Fermi velocity for the material as the only free parameter. Owing to the fitted values, the model gives very good results for channeled low-energy ions. The results were improved in [67] by the use of a more accurate three-dimensional electron density distribution for crystalline silicon to obtain the local electron density and local Fermi velocity inside the crystal.

24.5 Collision Cascades and Implantation-Induced Damage

In noninsulating materials, the damage in a collision cascade is produced mainly by the nuclear energy deposition. In insulators, the electronic energy loss can also contribute to the production of atomic damage, although the detailed mechanisms for this are not fully understood.

In materials consisting of light atoms and having a low atomic density, such as silicon, the cascades are roughly linear and no large liquid zones form within the cascade. The cascade development and damage distribution can be represented rather well by BCA simulations, especially if the parameters in the simulation are calibrated with MD simulations [68]. In semiconductor materials, regions of the crystal can become strongly disordered by the irradiation, forming amorphous zones upon cooling down owing to the low recrystallization rate of these materials [69]. In heavy, dense metals, large liquid zones formed during irradiation are recrystallized almost perfectly during the cooling-down phase of the cascade and only a few isolated point defects remain [70,71]. The total damage production at keV energies is about a factor of five less than expected from a linear cascade model [72,73].

Surface effects in collision cascades induced by a single ion are divided into four categories. The first effect, sputtering of single recoil atoms by ballistic collisions (Fig. 24.5a), is well understood from classical theory [74,75] and simulations (see e.g. [76–81]). For light materials and ions penetrating deep into the sample, sputtering can be the only surface effect of a cascade. In

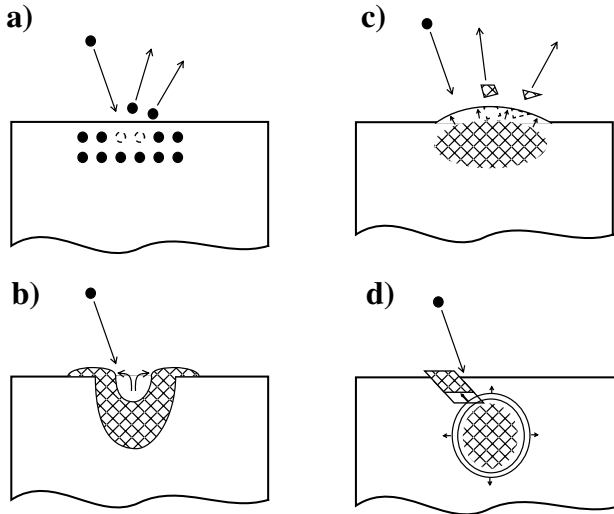


Fig. 24.5. Different mechanisms for ion-irradiation-induced damage close to the surface. (a) sputtering of single recoil atoms by ballistic collisions, (b) plastic flow of hot liquid onto the surface, (c) microexplosions, and (d) coherent displacement of atoms

dense materials, the collisions produced by a heavy ion, typically at an energy of a few keV, can be so well localized that they can produce a liquid-like zone inside the material or close to the surface [70]. The other three surface effects all require that such a zone is formed somewhere in the cascade. The second effect, plastic flow of hot liquid onto the surface, can result when a cascade is centered inside the sample, but is bounded by the surface so that liquid atoms can flow onto the surface (Fig. 24.5b) [82]. The third effect, “microexplosions”, occurs when the liquid zone is so close to a surface that the pressure wave from the cascade essentially ruptures the surface (Fig. 24.5c). In this case, pockets of hot liquid can explode out from the surface as a direct result of the collision cascade [80,83]. The fourth kind of surface damage effect is a coherent displacement of atoms, leading to the formation of an adatom island (Fig. 24.5d) [84].

The microexplosion effect corresponds to the nonlinear sputtering regime, and has thus been indirectly observed in numerous experimental sputtering studies. It is expected to produce craters on the sample surface, which have been observed in several experiments [85,86]. A recent comparison of simulations and experiments has shown that the observed crater formation can be explained by the liquid flow and microexplosion mechanisms [83,87].

24.A Appendix: Computer Methods to Calculate the Slowing Down of Ions

24.A.1 Binary-Collision Approximation Method

Analytical calculations of binary collisions have been used to deduce the stopping power as a function of the ion velocity. Computer simulations which treat the successive collisions as binary collisions are called binary-collision approximation (BCA) [88] methods. These methods are based on Monte Carlo (MC) techniques. One of the main tools in the field of statistical particle penetration has been the use of transport equations and MC techniques to solve them numerically [15]. Particle penetration was one of the first major applications of computer simulations in physics [15].

In the MC technique, a random number is used to determine the free-flight path l of an ion from an exponential distribution $F(l) = (1/\lambda)e^{-l/\lambda}$, where $\lambda = 1/(N\sigma(E))$ is the mean free path, N is the atom density of the target, and $\sigma(E)$ is the cross section for all possible collisions under consideration. The type of a collision is defined by a random number, resulting in a changed path for the ion, namely a new direction, charge state, energy, etc. Each collision is treated as binary, neglecting the rest of the environment. This procedure is then iterated until the ion has lost all of its energy. The implantation of ions is simulated by following the ions until they stop, and a histogram of the penetration depths from the target surface is used to give the range profile of the ions.

An MC calculation is very fast with modern computers, requiring typically seconds or minutes of total simulation time. There are several BCA codes [77, 89–91] available, of which the code most often used is TRIM [91] and later parametrizations of it [9].

24.A.2 Molecular-Dynamics Method

In the molecular-dynamics (MD) method, the movement of the ion is simulated more accurately than in the BCA methods. There are two factors behind the better accuracy: (i) the properties of the target structure can be described more realistically, and (ii) the ion can interact with several atoms at a time.

In MD simulations, the atoms in the system are given initial spatial and momentum coordinates. By solving the Newtonian equations of motion in a small time step for all the atoms, the atomic coordinates are changed. This process is then iterated until the given criterion (for example the maximum time) for the termination of the simulation is fulfilled. The accuracy of the ion movement depends on the length of the time step. The movement of the atoms is determined by the forces acting between them. The forces depend on the interaction potentials of the atoms, which are usually divided into repulsive and attractive parts.

The MD method makes it possible to include all the many-body collisions neglected in the BCA simulations. Thus it is the only way to calculate ion movement for low energies or in the case of atom-cluster implants. The electronic structure of the target material can also be constructed accurately, because the spatial coordinates of all atoms in the simulation are determined during every time step.

A full MD simulation is very computer-time-consuming, and this limits its use to very low energies (<1 keV/amu). Different methods have been developed to limit the calculations only to the vicinity of the moving ion (e.g. [66,92]). Thus the simulation time with normal desktop computers needs only to be from minutes to hours for statistically significant results. The drawback is that the cascades are not taken into account, which means that the change in the target structure cannot be calculated directly.

24.B Appendix: Interatomic Potentials

The forces between atoms define not only the scattering of an incoming ion by an atom of matter but also almost all physical and chemical phenomena in matter. The forces are derived from an interatomic-potential model function V depending on the positions of the atoms. For a system of N atoms whose position vectors are $\mathbf{r}_i$, the force $\mathbf{F}_i$ on the i th atom is given by $\mathbf{F}_i = -\partial V(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)/\partial \mathbf{r}_i$.

For the case of two atoms, $\mathbf{F} = -\partial V(\mathbf{r})/\partial \mathbf{r}$, where $\mathbf{F}$ is the force acting on either of the particles, whose mutual separation is $\mathbf{r}$. For two-body interactions it can be assumed that $V \rightarrow 0$ as $r \rightarrow \infty$ and that $V(r) \sim 1/r$, and, owing to the strong nuclear repulsion, $V \rightarrow \infty$ as $r \rightarrow 0$.

The interatomic force is defined by a screened Coulomb potential

$$V = \frac{Z_1 Z_2 e^2}{r} \chi(r/a_0), \quad (24.25)$$

where the screening function is $\chi(r/a_0)$. It is given as a function of distance in units of $a_0 = 0.8853a_B Z^{-1/3}$, where a_B is the Bohr radius. Several screening functions have been reported in the literature (see Fig. 24.6). These include the function given by Bohr, $\chi(x) = e^{-x}$; the Thomas–Fermi function, $\chi(x) = [1 + (x^3/144)^{0.8034/3}]^{-3/0.8034}$; the Molière function, $\chi(x) = 0.35e^{-0.3x} + 0.55e^{-1.2x} + 0.10e^{-6.0x}$; and the Lenz–Jensen function, $\chi(x) = e^{-q}(1 + \sum a_i q_i)$, where $q = 3.11126x^{1/2}$.

In computer simulation codes, the standard semiempirical interatomic-potential function (ZBL) [9] $\chi(x) = \sum a_i e^{-b_i x}$ is conventionally used. The values of the coefficients have been obtained by fitting to quantum mechanically calculated data.

At distances of around 0.1–0.4 nm, the interatomic potential is attractive. The properties of diatomic molecules, binding energies, and equilibrium

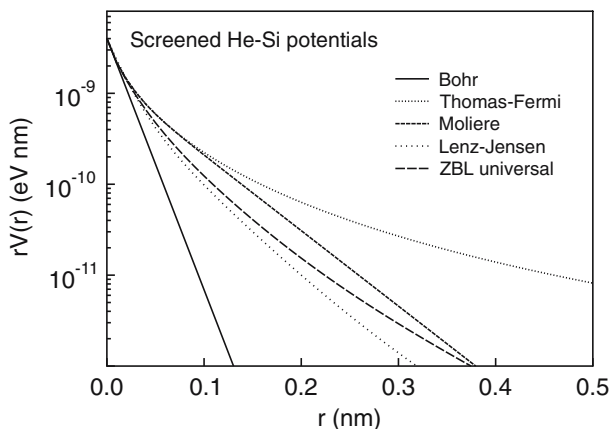


Fig. 24.6. Screening function for different potentials [9]

separations, as well as the properties of matter at and close to thermal equilibrium, are determined by a potential containing a repulsive and an attractive part.

References

1. N. Bohr: *Phil. Mag.* **25** (1913) 16
2. N. Bohr: *Phys. Rev.* **58** (1940) 654
3. J. Lindhard, M. Scharff, H. E. Schiøtt: *Mat. Fys. Medd. Dan. Vid. Selsk.* **33** (1963) 1
4. E. Fermi, E. Teller: *Phys. Rev.* **72** (1947) 399
5. O. Firsov: *Sov. Phys. JETP* **9** (1959) 1076
6. J. Lindhard, M. Scharff: *Phys. Rev.* **124** (1961) 128
7. H. Bethe: *Ann. Phys.* **5** (1930) 324
8. W. Brandt, M. Kitagawa: *Phys. Rev. B* **25** (1982) 5631
9. J. F. Ziegler, J. P. Biersack, U. Littmark: *The Stopping and Range of Ions in Matter*, Pergamon, New York, 1985. The computer program SRIM is based on this ZBL parametrization of the electronic stopping power. The version SRIM2003.24 has been used in this chapter
10. R. Smith, M. Jakas, D. Ashworth, B. Owen, M. Bowyer, I. Chakarov, R. Webb (eds.): *Atomic and Ion Collisions in Solids and at Surfaces*, Cambridge University Press, Cambridge, 1997
11. P. Sigmund, A. Schinner: *Nucl. Instr. Meth. Phys. Res. B* **195** (2002) 64
12. N. R. Arista: *Nucl. Instr. Meth. Phys. Res. B* **195** (2002) 91; see also erratum: *Nucl. Instr. Meth. Phys. Res.* **207** (2003) 232.
13. P. L. Grande, G. Schiwietz: *Nucl. Instr. Meth. Phys. Res. B* **195** (2002) 55
14. P. Sigmund: *Phys. Scr.* **28** (1983) 257
15. A. Gras-Marti, H. M. Urbassek, N. Arista, F. Flores (eds.): *Interaction of Charged Particles with Solids and Surfaces*, Plenum; New York, 1991

16. P. Sigmund, A. Schinner: Eur. Phys. J. D **12** (2000) 425
17. P. Sigmund, A. Schinner: Eur. Phys. J. D **15** (2001) 165
18. J. R. Sabin, J. Oddershede: Phys. Rev. A **26** (1982) 3209
19. J. R. Sabin, J. Oddershede: At. Data Nucl. Data Tables **31** (1984) 275
20. H. Bichsel: Phys. Rev. A **65** (2002) 52709
21. J. M. Pitarke, R. H. Ritchie, P.M. Echenique: Phys. Rev. B **52** (1995) 13883
22. E. Uggerhøj: Phys. Scr. **28** (1983) 331
23. D.G. Arbo, J.E. Miraglia: Phys. Rev. A **58** (1998) 2970
24. P. M. Echenique, I. Nagy, A. Arnau: Int. J. Quantum Chem. **23** (1989) 521
25. P. Sigmund, A. Fettouhi, A. Schinner: Nucl. Instr. Meth. Phys. Res. B **209** (2003) 19
26. L. Nortcliffe, R. Schilling: Nucl. Data Tables **7** (1970) 233
27. J. Forster, D. Ward, H. Andrews, G. Ball, G. Costa, W. Davies, I. Mitchell: Nucl. Instr. Meth. **136** (1976) 349
28. J. Anthony, W. Lanford: Phys. Rev. A **25** (1982) 1868
29. F. Hubert, R. Bimbot, H. Gauvin: At. Data Nucl. Data Tables **46** (1990) 1
30. J. F. Ziegler: *Handbook of Stopping Cross-Sections for Energetic Ions in All Elements*, Pergamon, New York, 1980
31. N. Bohr: Mat. Fys. Medd. Dan. Vid. Selsk. **18** (1948) 8
32. N. Bohr: Mat. Fys. Medd. Dan. Vid. Selsk. **24** (1948) 19
33. J. Lindhard, M. Scharff: Mat. Fys. Medd. Dan. Vid. Selsk. **27** (1953) 15
34. Q. Yang, D. O'Connor, Z. Wang: Nucl. Instr. Meth. Phys. Res. B **61** (1991) 149
35. E. Bonderup, P. Hvelplund: Phys. Rev. A **4** (1971) 562
36. W. Chu: Phys. Rev. A **13** (1976) 2057
37. F. Besenbacher, J. Andersen, E. Bonderup: Nucl. Instr. Meth. **168** (1980) 1
38. M. Livingston, H. Bethe: Rev. Mod. Phys. **9** (1937) 245
39. L. Landau: J. Phys. USSR **8** (1944) 201
40. P. Vavilov: Sov. Phys. JETP **5** (1957) 749
41. P. Schulek, B. Golovin, L. Kulyukina, S. Medved, P. Pavlovich: Sov. J. Nucl. Phys. **4** (1966) 996
42. C. Tschalaer: Nucl. Instr. Meth. **64** (1968) 237
43. H. Bichsel, R. Saxon: Phys. Rev. A **11** (1975) 1286
44. C. Sofield: Nucl. Instr. Meth. B **45** (1990) 648
45. W. Scott: Rev. Mod. Phys. **35** (1963) 231
46. P. Sigmund, K. Winterbon: Nucl. Instr. Meth. **119** (1974) 541
47. A. Marwick, P. Sigmund: Nucl. Instr. Meth. **126** (1975) 317
48. P. Sigmund, J. Heinemeier, F. Besenbacher, P. Hvelplund, H. Knudsen: Nucl. Instr. Meth. **150** (1978) 221
49. G. Amsel, G. Battistig, A. L'Hoir: Nucl. Instr. Meth. Phys. Res. B **201** (2003) 325
50. J. Ormrod, H. Duckworth: Can. J. Phys. **41** (1963) 1424
51. J. Ormrod, J. MacDonald, H. Duckworth: Can. J. Phys. **43** (1965) 275
52. J. Ormrod: Can. J. Phys. **46** (1968) 497
53. B. Fastrup, A. Borup, P. Hvelplund: Can. J. Phys. **46** (1968) 489
54. P. Hvelplund, B. Fastrup: Phys. Rev. **165** (1968) 408
55. K. Winterbon: Can. J. Phys. **46** (1968) 2429
56. P. M. Echenique, R. M. Nieminen, J. C. Ashley, R. H. Ritchie: Phys. Rev. A **33** (1986) 897

57. I. Nagy, A. Arnau, P. M. Echenique: Phys. Rev. A **40** (1989) 987
58. K. Shima, N. Kuno, M. Yamanouchi: Chin. J. Phys. **26** (1988) 292
59. G. de M. Azevedo, M. Behar, P. L. Grande: Phys. Rev. B **63** (2001) 64101
60. K. Arstila, T. Sajavaara, J. Keinonen: Nucl. Instr. Meth. Phys. Res. B **174** (2001) 163
61. W. Bragg, R. Kleeman: Phil. Mag. **10** (1905) 318
62. J. Ziegler, J. Manoyan: Nucl. Instr. Meth. Phys. Res. B **35** (1988) 215
63. I. Ruge, J. Graul (eds.): *Ion Implantation in Semiconductors*, Springer, Berlin, 1971
64. S. Tian: J. Appl. Phys. **93** (2003) 5893
65. D. Cai, N. Grønbech-Jensen, C. M. Snell, K. M. Beardmore: Phys. Rev. B **54** (1996) 17147
66. K. M. Beardmore, N. Grønbech-Jensen: Phys. Rev. E **57** (1998) 7278
67. J. Sillanpää, K. Nordlund, J. Keinonen: Phys. Rev. B **62** (2000) 3109
68. H. L. Heinisch, B. N. Singh, T. Diaz de la Rubia: J. Nucl. Mat. **212–215** (1994) 127
69. M.-J. Caturla, L. A. M. T. Diaz de la Rubia, G. H. Gilmer: Phys. Rev. B **54** (1996) 16683
70. T. Diaz de la Rubia, R. S. Averback, R. Benedek, W. E. King: Phys. Rev. Lett. **59** (1987) 1930; see also erratum: Phys. Rev. Lett. **60** (1988) 76
71. K. Nordlund, M. Ghaly, R. S. Averback, M. Caturla, T. Diaz de la Rubia, J. Tarus: Phys. Rev. B **57** (1998) 7556
72. D. J. Bacon, T. Diaz de la Rubia: J. Nucl. Mat. **216** (1994) 275
73. R. S. Averback, T. Diaz de la Rubia: Displacement damage in irradiated metals and semiconductors, in: *Solid State Physics*, ed. H. Ehrenfest and F. Spaepen, volume 51, pp. 281–402, Academic Press, New York, 1998
74. R. Behrisch (ed.): *Sputtering by Particle Bombardment I*, Springer, Berlin, 1981.
75. P. Sigmund: Nucl. Instr. Meth. Phys. Res. B **27** (1987) 1
76. H. H. Andersen: Appl. Phys. **18** (1979) 131
77. W. Möller, W. Eckstein: Nucl. Instr. Meth. Phys. Res. B **2** (1984) 814
78. R. Smith, R. P. Webb: Nucl. Instr. Meth. Phys. Res. B **59/60** (1991) 1378
79. J. Likonen, M. Hautala: J. Phys. Condens. Matter **1** (1989) 4697
80. H. M. Urbassek, K. T. Waldeer: Phys. Rev. Lett. **67** (1991) 105
81. W. Eckstein, S. Hackel, D. Heinemann, B. Fricke: Z. Phys. D: At. Mol. and Clusters **24** (1992) 171
82. M. Ghaly, R. S. Averback: Phys. Rev. Lett. **72** (1994) 364
83. M. Ghaly, K. Nordlund, R. S. Averback: Phil. Mag. A **79** (1999) 79
84. K. Nordlund, J. Keinonen, M. Ghaly, R. S. Averback: Nature **398** (1999) 49
85. K. L. Merkle, W. Jäger: Phil. Mag. A **44** (1980) 741
86. S. E. Donnelly, R. C. Birtcher: Phys. Rev. B **56** (1997) 13599
87. M. Morgenstern, T. Michely, G. Comsa: Phil. Mag. A (1999) **79**
88. M. T. Robinson, I. M. Torrens: Phys. Rev. B **9** (1974) 5008
89. M. T. Robinson: Nucl. Instr. Meth. Phys. Res. B **67** (1992) 396
90. M. Hautala, I. Koponen: Defect Diffus. Forum **57–58** (1988) 61
91. J. P. Biersack, L. G. Haggmark: Nucl. Instr. Meth. **174** (1980) 257
92. K. Nordlund: Comput. Mater. Sci. **3** (1995) 448

25 Modification of Materials by MeV Ion Beams

Y. Zhang¹ and H.J. Whitlow^{2,3}

¹ Pacific Northwest National Laboratory, P O Box 999, Richland, WA 99352, USA
Yanwen.Zhang@pnl.gov

² Department of Physics, P.O. Box 35 (YFL), FIN-40014 University
of Jyväskylä, Finland
Harry.J.Whitlow@phys.jyu.fi

³ School of Technology and Society, Malmö högskola, 206 05 Malmö, Sweden
Harry.J.whitlow@ts.mah.se

25.1 Introduction

Today's fast-developing technologically based society places ever accelerating demands on new materials and materials-processing methods. Leading-edge fields as diverse as biomedical tissue engineering, quantum devices, optical and magnetic information storage technology, and immobilization of actinides all require nanoscale engineering through controlled materials modification. The evolution of these advances from the research science stage to the industrial-applications phase is a particularly challenging task.

Amongst the beam-processing methods (electron, X-ray, laser etc.) for materials modification, MeV ions occupy a unique place. They interact strongly with both the atomic and the electronic structures of the target material to produce a broad plethora of modifications with well-defined penetration characteristics. These modifications can be precisely controlled because MeV ions can be steered and focused using electrostatic and magnetic fields, and also close process control is possible because their delivery can be monitored electrically.

A comprehensive review of materials engineering with MeV ions will not be attempted here. Instead, the scope of this chapter is to introduce the basic concepts with some practical examples of their use in advanced materials engineering.

25.2 Characteristics of MeV Ion Bombardment for Materials Engineering

Ions penetrating matter can modify the material by direct implantation of a foreign atom into the target material, or as a result of the ion stopping, which deposits kinetic energy from the projectile ion into the material. This deposition of energy takes place, as discussed in Chap. 24, by scattering of

the ion by the electrons and nuclei in the target [1]. The coupling of the deposited energy to the electronic and structural system of the material results in a wide plethora of primary effects, illustrated schematically in Fig. 25.1. From the viewpoint of materials modification, we are generally interested in persistent changes such as amorphization of the crystalline structure. These modifications may be *direct*, where no further processing is needed, or *latent*, where an additional stage of processing is required to achieve the desired modification, such as lithography, where the latent image is brought forth by development. The relative importance of the different processes depends on the nature of the material, for example, for dielectric materials the electronic processes are of great importance, whilst for metals, where the electronic excitation relaxes almost instantaneously via electron-phonon interactions, modifications induced by nuclear displacement dominate.

The energy transferred in an individual collision with an MeV ion may be large, extending to keV energies for electrons and MeV for recoiling nuclei. It follows that these recoiling particles can themselves scatter, creating a cascade of energetic secondary electrons and recoil nuclei. As the energies of the primary ion and secondary particles decrease, a very wide plethora of processes take place, as shown in Fig. 25.1. The secondary particles in turn will cause tertiary processes, such as defect agglomeration, forming extended defects such as dislocations and voids; chemical-bond scission; and other forms of

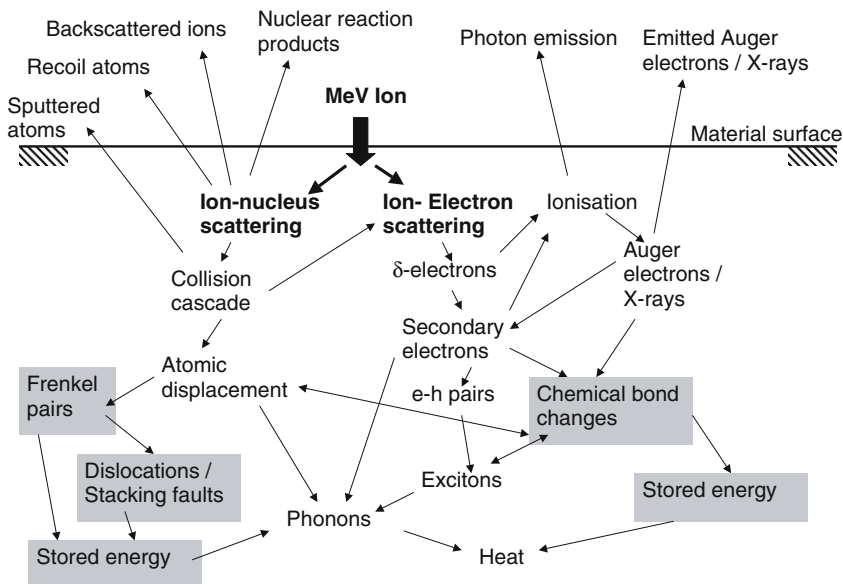


Fig. 25.1. Schematic illustration of the evolution of the processes taking place in a solid material irradiated with MeV ions. The processes shaded gray correspond to long-term effects

radiolysis. In addition, various forms of prompt processes, such as phonon creation, secondary electron emission, sputtering, X-ray and visible photon emission, and backscattering, can take place. Prompt emissions are generally not important for modification, but can be utilized as a monitor signal for process control. It is useful to bear in mind that at any time, the energy balance requires that the energy deposited in the material per ion is the sum of the energy of emitted particles, moving particles, stored energy associated with induced defects, chemical-bond changes and heat. Moreover, the sum of the momentum of all moving particles in the cascade must sum to the primary-ion momentum.

25.2.1 Fundamental Interactions Between MeV Ions and Materials

As an energetic ion penetrates the surface of matter, the probability of scattering with electrons and the atomic nuclei of the material rises from zero outside of the material to some finite value determined by the appropriate cross section (Chap. 24). The successive scattering from atomic nuclei and electrons along the ion trajectory gives rise to a stopping force¹ on the projectile, transferring energy to the target material. This slowing down gives rise to the two basic forms of ion beam modification of materials:

- i. If the material is thick enough, the ion, after traversing a distance R , where

$$R = \int_E^0 \frac{dE}{(-dE/dx)} \quad (25.1)$$

termed the range, will lose all its energy because of the stopping force and come to rest. This process, termed ion implantation, is widely used to modify materials by introducing foreign atoms.

- ii. The material is modified through deposition of energy by interaction with the ion, which induces chemical, structural and electronic changes.

25.2.2 Characteristics of MeV Ion Penetration

The salient features of MeV ion bombardment are illustrated in Fig. 25.2, which presents the results from SRIM binary collision approximation (BCA) model calculations [2] of 7 MeV ^{14}N ions (0.5 MeV per nucleon) incident on a Si target. This combination was chosen because it represents an intermediate-mass target and an intermediate-mass ion with an energy that can easily be obtained from both single-ended and tandem MeV ion accelerators. It is evident from this figure that:

¹ The stopping force $(-dE/dx)$ is, strictly incorrectly, often termed the stopping power. See Appendix 25.A.

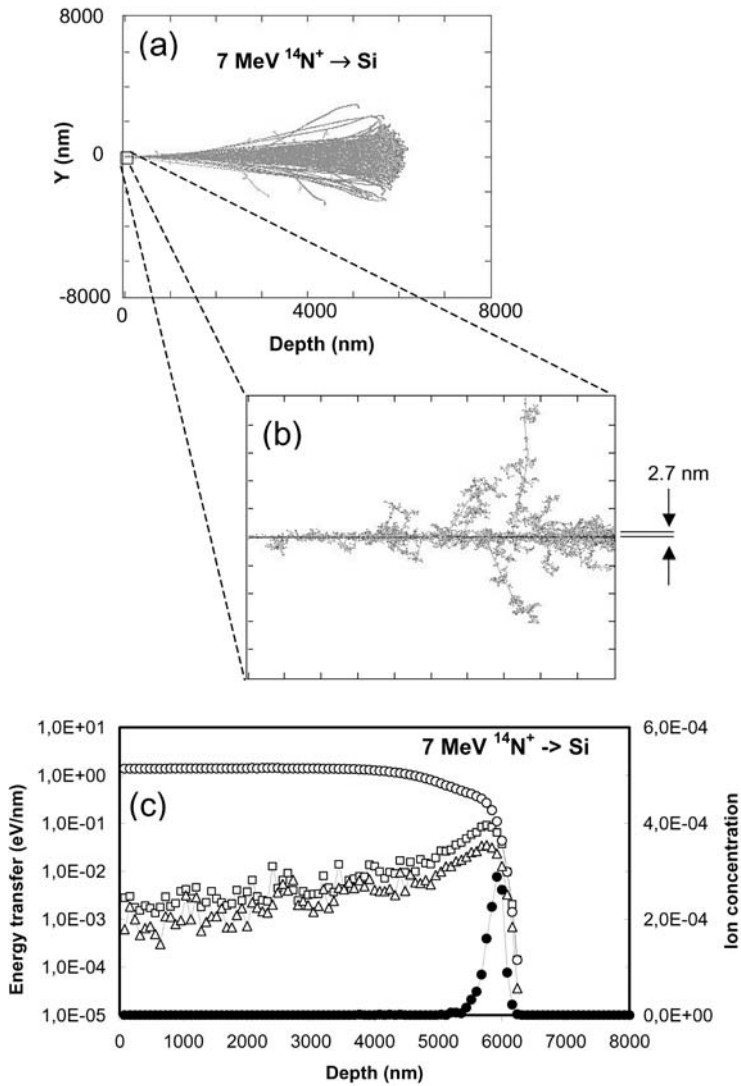


Fig. 25.2. SRIM simulation [2] of 7 MeV ^{14}N ions impinging on an amorphous Si target. The ion–target combination represents an intermediate-mass ion on an intermediate-mass target material. (a) Whole depth region, showing the deep penetration. (b) 100 nm thick, 200 nm wide surface region, showing the small spreading of the primary ions. (c) Depth distribution of ion projected range (*closed circles*), energy deposited as primary ionization (*open circles*), energy deposited in creating nuclear recoils (*open squares*) and energy deposited in ionization by nuclear recoils (*open triangles*)

- i. Almost all of the ions come to rest at a depth of $\sim 5.8 \mu\text{m}$ with a straggling of $\sim 1 \mu\text{m}$. The ratio of the straggling to the projected range is much less than for low-energy (keV) ions.
- ii. Energy deposition in electronic excitations dominates the stopping process.
- iii. Energy deposition in nuclear scattering processes takes place along the entire ion track but is an order of magnitude smaller than the electronic energy deposition. These processes are largest close to the end of the ion range. The majority of the energy deposited in nuclear scattering ends up as electronic excitation through electronic stopping of the energetic recoils.
- iv. At the end of the range, the lateral radial spreading of the trajectories is large ($\sim 400 \text{ nm}$). However, in the near-surface region, where only little nuclear scattering takes place, the radial spreading of the trajectories is correspondingly small ($\sim 3 \text{ nm}$). Although in this near-surface region there is radial transport of energy by nuclear recoils, the overwhelming contribution to the electronic energy deposition comes from the primary ions.

25.2.3 Nuclear Collision Cascades

The dynamics of nuclear collision cascades are described in Chap. 24; therefore, only a cursory treatment will be given here. In contrast to the case for keV ion irradiation, the mean free path for nucleus–nucleus scattering is large. This results in small, isolated subcascades along the primary-ion trajectory. These can be seen extending sideways in Fig. 25.2(b). This is particularly the case close to the surface, where the cross section for nuclear scattering is small and the mean free path between nuclear scattering events is correspondingly large. For nanometer materials, the structures are often just a few hundred nm thick. Thus MeV ions will pass through these structures creating very few cascades compared with the case for keV ions.

Nuclear collisions lead to the formation of Frenkel pairs. A *Frenkel pair* is composed of two point defects: a *vacancy*, which is a site where an atom would normally be located, and an *interstitial*, which is a site where an atom is located that would not normally be there. The situation for MeV ions is very closely similar to that for low-energy ion bombardment. The process is essentially the same for other forms of irradiation, and the physics is discussed in detail in the book by Thompson [3] and the compendium by Andersen [4]. If the interstitial and its associated vacancy are close together, they can undergo correlated recombination, and the defect energies are then converted to phonons (heat). Alternatively, the interstitial may be annihilated at another vacancy (uncorrelated recombination) or a defect sink, again releasing heat. The effect of the buildup of primary defects may lead to their agglomeration to form so-called extended defects such as voids, dislocations stacking and faults [1, 3]. If the fluence is sufficiently large, these amorphous subdomains

overlap and a layer that is totally amorphous will form. The defects produced during the slowing processes are not thermalized. This can be utilized in defect-stimulated regrowth of amorphous layers and domains [1].

25.2.4 Electronic Processes

Electron Cascades

In addition to the nuclear collision cascade processes discussed above (Sect. 25.2.3 and Chap. 24), the electronic stopping of the primary ion and secondary recoil nuclei can induce materials modification. The electronic stopping for MeV ions is much greater than the nuclear stopping, and consequently electron-induced materials modification by MeV ions may be more pronounced than for keV ions, where the contribution from nuclear stopping is more prominent. The electronic stopping creates excited electrons by ion–electron scattering. This scattering can excite electrons in both the projectile and the target atoms [5], from both core and valence electron levels, to unfilled bound states (so-called *resonant excitation*) or to free states (*quasi-classical scattering*) [6]. For quasi-classical scattering, the energy of the scattered free electron is $E_f = T - E_b$, where $T \approx E (4m_e/M_1) \cos^2 \varphi$ is the energy transferred to a free electron in an ion–electron collision with an ion of mass M_1 , where E_b is the electron binding energy. For valence electrons E_b is of the order of a few eV and we may, to a good approximation, consider them to be free classical electrons ($E_b = 0$). The primary electrons (δ -electrons) will then have energies extending from zero ($\varphi = \pi/2$) to a maximum T_{max} at $\varphi = 0$ (see Fig. 25.3), which corresponds to 2.2 keV for 1 MeV protons. The δ -electrons undergo elastic and inelastic scattering with the target electrons as they move through the target material. In the scattering process, the δ -electrons transfer kinetic energy to the target electrons, resulting in a cascade of secondary electrons, shown schematically in Fig. 25.3. The energy deposited in the scattering process can lead to modification of the material, for example by breaking and forming chemical bonds in dielectric materials around the ion track. The extent of the secondary-electron-induced modification will then be governed by the dose at that point in the material. In general, the ion track dose model [7] is a good starting point. The differential cross section $d\sigma$ for an energy transfer T to $T + dT$ between an ion of charge Z_1e and an electron is

$$d\sigma = \frac{Z_1^2 e^4 dT}{8\pi \varepsilon_0^2 m_e v_1 T^2} \quad (25.2)$$

Note that this expression (25.2), in SI units, differs from the cgs counterpart usually quoted (with $e^2 \Rightarrow e^2/4\pi\varepsilon_0$) and that Z_1 can be an effective charge. The energy transfer T depends on the recoil angle (Fig. 25.3). The number n of δ -electrons scattered through φ to $\varphi + d\varphi$ becomes, by substitution for dT/T^2 and $M_1 = 2E/v_1^2$,

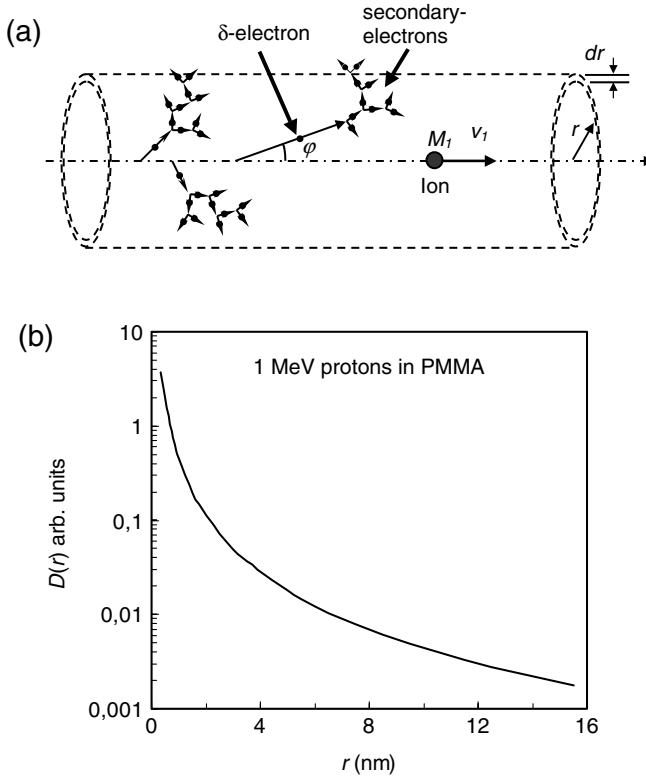


Fig. 25.3. (a) Schematic illustration of secondary-electron cascades for an MeV ion moving through matter. (b) Radial electron energy dose distribution $D(r)$ for 1 MeV protons in poly(methylmethacrylate) (PMMA). Note that the $D(r)$ scale is logarithmic

$$dn = N_e \frac{Z_1^2 e^4 \sin \varphi}{8\pi \epsilon_0^2 m_e v_1 \cos^3 \varphi} d\varphi \quad (25.3)$$

The cross section dn approaches infinity asymptotically for $\varphi \rightarrow \pi/2$, and the majority of δ -electrons are directed perpendicular to the beam, with energies T close to zero. δ -electrons scattered with smaller φ will have greater energy T . The range of electrons in the material can be conveniently approximated by the empirical relation

$$R(T) = (a/\rho) T^\beta \quad (25.4)$$

where $a = 5.2 \times 10^{-4}$ and $\beta = 5/3$ for ρ in g cm^{-3} and T in eV. After some manipulation detailed elsewhere [7] and making the approximation that $R(T)$ represents the maximum radial spread about the centerline of the beam for secondary electrons with energy T , the dose (eV/unit volume) at a distance r from the ion track can be obtained:

$$D(r) = N_e \frac{Z_1^2 e^4}{8\pi\epsilon_0^2 m_e v_1 \beta r^2} \left(1 - \frac{r}{R_{max}} \right), \quad \text{for } r \leq R_{max} = R(T_{max}) \quad (25.5)$$

The dependence of $D(r)$ on r , shown in Fig. 25.3(b) for 1 MeV protons in poly(methylmethacrylate) (PMMA), reveals that the electron dose is very sharply localized within a few nm from the ion track. Electron-mediated ion beam modification will thus be characterized by the superposition of intense localized electron doses extending along the ion tracks. Often one is interested in the net dose $D(p)$ at some point p in a material irradiated with a parallel flux of ions. This is the superposition of the dose $D(r_i)$ contributed by each of the ions impinging at a radial distance r_i from p within a radius R_{max} :

$$D(p) = \sum_{r_i \leq R_{max}} D(r_i) \quad (25.6)$$

R_{max} corresponds to the range of electrons with the maximum classical energy (25.5).

There exist a number of notable shortcomings in the simple track model outlined above.

- i. The model of the electron cascade ignores the lateral spreading of secondary electrons.
- ii. Only free classical excitation is considered, and resonant excitation is neglected.
- iii. The empirical assumption about electron ranges is in reasonable agreement with experiment in the keV region. However, because of the $1/T^2$ dependence of the differential scattering cross section in (25.2), most electrons will have lower energies, where the validity of the energy-range relation (25.4) is questionable and little experimental data exist.
- iv. The asymptotic behavior of $D(r)$ at small r – see (25.5) and Fig. 25.3 – is unphysical because the dose becomes infinite.
- v. Penetrating ions undergo lateral spreading as a result of nuclear and electronic scattering.

A complete treatment of the dose $D(r)$ about an ion track is analytically and computationally not straightforward [8] and requires a realistic 3-D calculation of electron slowing down, for example using Monte Carlo codes, such as CASINO [9–11], where the low-energy electronic stopping is determined from measurements of the complex optical refractive index and electron energy loss spectra [11]. In order to include the nuclear scattering contribution to the average dose $D(r)$ at a radius from the ion impingement axis, it is necessary to convolute $D(r)$ with the radial spreading from nuclear scattering events.

Plasma Column

The intense ionization along the ion track results from the rapid falloff with r of the radial dependence of the dose about an ion track (Fig. 25.3(b)). In the ion track, secondary electrons move radially away from the center, leaving behind less mobile positively charged holes. These exert an electrostatic restraining force on the electrons, which confines them to form a plasma column of electron-hole pairs some tens of nm broad that extends along the ion track. The plasma column may decay by spontaneous electron-hole recombination, releasing phonons (heat), or by ambipolar diffusion of electrons and holes out of the plasma column. The dominant mode of decay is determined by the electronic structure of the material. In metals, the electron-hole recombination dominates, whereas in semiconductors and dielectrics, the plasma column may be long-lasting and the column decays primarily by ambipolar diffusion rather than recombination. In dielectric materials, excitation may be so great that it introduces structural changes, for example by forming and breaking chemical bonds in polymers, and creating color center defects in alkali halides.

25.3 Applications of Materials Modification by MeV Ion Beams

25.3.1 MeV Implantation into Silicon Carbide (SiC)

The electrical, chemical, thermal and mechanical properties of silicon carbide (SiC) make electronic devices based on SiC superior to those based on Si for high-power and high-frequency applications, as well as for operation in harsh environments (e.g. high temperature and high radiation). Success in fabricating high-quality SiC has promoted worldwide activity in establishing technologies that make full use of this unique semiconductor.

For production of SiC-based devices, ion implantation is the only low-temperature doping technique because thermal diffusion of dopants requires extremely high temperatures. There is, however, a great challenge with ion implantation because it inevitably produces defects and lattice disorder, which not only deteriorate the transport properties of electrons and holes, but also inhibit electrical activation of the implanted dopants. Point defects or the growth of extended defect structures often leads to high leakage currents, poor and uneven injection during forward bias, and the premature breakdown under reverse bias observed in manufactured SiC diodes. To investigate the origin of these difficulties in SiC device fabrication and assess performance in high-radiation environments, MeV ions have been used to introduce damage into various SiC polytypes [12–19]. The disorder level of the atomic displacements or lattice disorder produced depends on the ion fluence, ion flux and implantation temperature. As an example, recent results [12, 19]

using 1.1 MeV Al_2^{2+} ions to investigate damage accumulation during ion bombardment are presented here.

The relative disorder on both the Si and the C sublattices at the damage peak for samples implanted at 150 K is shown in Fig. 25.4(a) [12]. The data indicate a predominantly sigmoidal dependence of the damage on increasing dose at the damage peak. At low ion fluence, the greater rate of C disordering is consistent with lower threshold displacement energy and a greater production rate of C defects relative to Si. These results are consistent with MD simulations of the displacement cascades [14]. MeV ion beams have also been used to study the annealing behavior. The example in Fig. 25.4(b) shows the recovery of relative Si disorder resulting from isochronal annealing [12]. Similar recovery behavior is observed for the C sublattice. In this process three distinct recovery stages are observed, which can be associated with mobile interstitials and vacancies in different defect configurations. Dynamic recovery of interstitials and vacancies is conspicuous during implantation at elevated temperatures [19]. This is illustrated in Fig. 25.5, which compares the disorder profiles for 1.1 MeV Al_2^{2+} implantations at 150 and 450 K. During

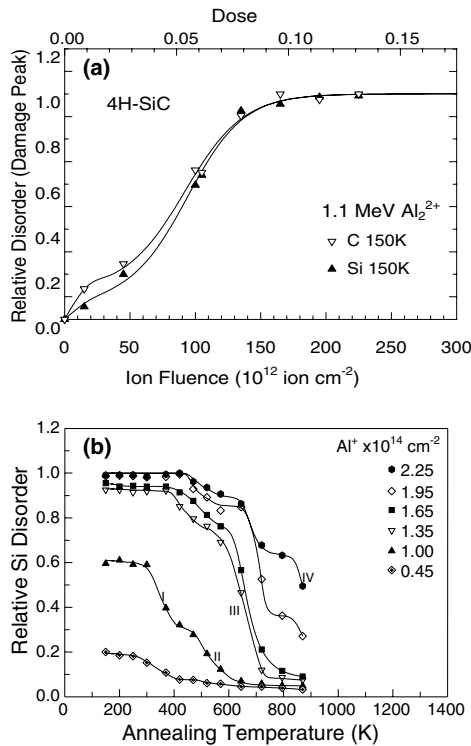


Fig. 25.4. (a) Relative disorder as a function of ion fluence for Al-implanted 4H-SiC. (b) Isochronal recovery of relative Si disorder at the damage peak

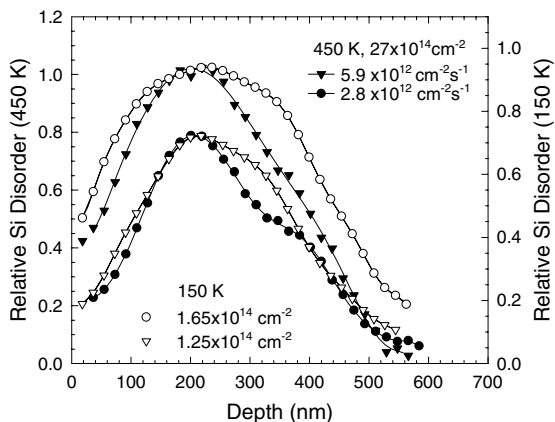


Fig. 25.5. The relative Si disorder profiles of samples that we implanted at different temperature, ion fluence and flux

implantation at 450 K, dynamic recovery of interstitials and vacancies occurs at a much higher rate, which suppresses the damage accumulation. The dynamic recovery implies that about 20 times higher fluence is needed at 450 K to produce the same damage level as at 150 K. The energy deposition rate through nuclear scattering can be controlled during MeV ion bombardment via the ion flux. Figure 25.6 shows high-resolution transmission electron microscope (HRTEM) images from the damage peak region. For the low-flux sample, as shown in Fig. 25.6(a), the basal-plane structure is maintained, while localized strain contrast is visible. Only a few occurrences of plane bending or termination are perceived. For the high-flux sample, a high concentration of planar defects and larger linked amorphous domains are observed, as shown in Fig. 25.6(b).

25.3.2 MeV Ion Irradiation Studies in Pyrochlore Materials

Pyrochlore materials ($A_2B_2O_7$) and perovskite-type oxides (ABO_3) are attracting great interest because of the capability to incorporate different elements in the A and B sites of their chemical structure. This capability suggests a wide range of applications, such as fuel cells [20, 21], catalysts [22, 23] and the immobilization of actinide-containing nuclear waste [24–29]. In actinide-bearing phases, considerable radiation damage due to alpha decay results in amorphization, macroscopic swelling and order-of-magnitude increases in dissolution rates [28–31]. This manifests itself as macroscopic changes in structure and chemical durability, which affect the long-term performance of the actinide waste forms [27–33].

Studies [29, 30, 34, 35] of actinide-doped/natural pyrochlores and related structures indicate that pyrochlores with Ti, Nb and Ta as the major B-site

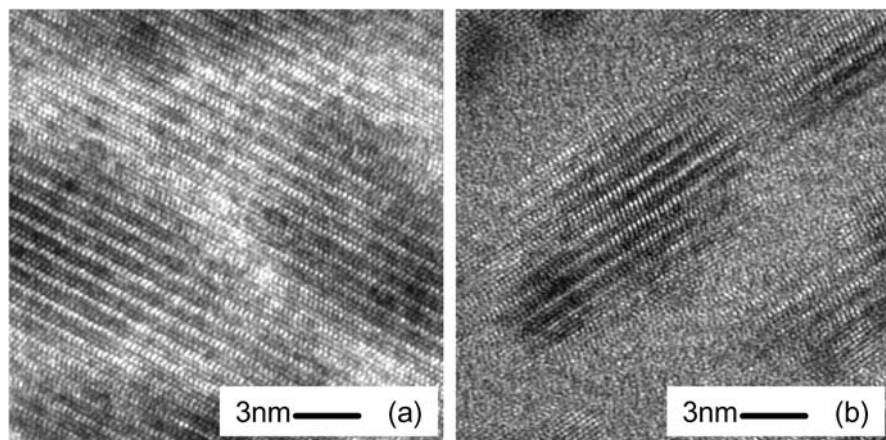


Fig. 25.6. HRTEM micrographs of samples implanted with 2.7×10^{15} 1.1 MeV Al_2^{2+} ions cm^{-2} at 450 K with different ion fluences (a) 2.8 and (b) 5.9×10^{12} $\text{Al}^+ \text{cm}^{-2} \text{s}^{-1}$

cations become amorphous as a result of the gradual accumulation of alpha-recoil collision cascades. These studies rely on the dose rates that can be achieved by radioactive decay and are hence time-consuming. Only limited data for a few sets of experimental conditions can generally be obtained.

More rapid evaluation of radiation effects in pyrochlore materials can be achieved by high-energy heavy-ion irradiation studies [25, 31, 36–39]. Alpha decay of actinide elements produces 4.5 to 5.8 MeV alpha particles and 70 to 100 keV recoil nuclei (alpha recoils). The more massive and lower-energy alpha recoils account for most of the damage produced through elastic-scattering collisions. Because the nuclear stopping of heavy ions is similar to the nuclear stopping of alpha recoils, the damage evolution under ion irradiation can provide a reasonable simulation of the damage evolution behavior due to alpha recoils. For actinide-containing pyrochlores, the radiation damage will be uniformly distributed. The nuclear damage induced in the surface layer by high-energy heavy ions is similar to that from the heavy alpha recoils because the cross sections and energy deposition are similar. MeV ion irradiation can be used to produce a highly damaged or amorphous state that does not differ greatly from that produced by alpha decay over long time periods. The power of this technique is that the highly damaged layers are confined to near-surface regions (up to several μm), and hence the chemical durability of such irradiated samples can be readily tested. The results of irradiation studies, as recently demonstrated for some pyrochlore samples [28], generally confirm the results for the highly damaged states of actinide-doped pyrochlores or natural minerals. Ion-beam irradiation provides a useful method to study amorphization, crystal swelling and dissolution rates and offers a reasonable

representation of the worst-case effect of radiation effects on chemical durability over long time periods for actual actinide-containing waste forms.

Implementation of the technique as a tool for testing pyrochlore materials for high-radiation environments requires quantitative studies of the damage evolution behavior as a function of irradiation dose and the mechanism for amorphization. Recently, the evolution of implantation-induced damage on the Sm and O sublattices from minor disorder to a fully amorphized state in samarium titanate pyrochlore ($\text{Sm}_2\text{Ti}_2\text{O}_7$) has been investigated [40]. In this study, 1 MeV Au^{2+} ions were chosen to simulate the damage production of heavy recoils through alpha decay. Rutherford backscattering spectroscopy (RBS) and $^{16}\text{O}(\text{d}, \text{p})^{17}\text{O}$ nuclear reaction analysis (NRA) along the $\langle 001 \rangle$ direction were used to characterize the relative disorder on the Sm and O sublattices, respectively. The damage accumulation at different irradiation temperatures is shown in Fig. 25.7. The results indicate that the relative

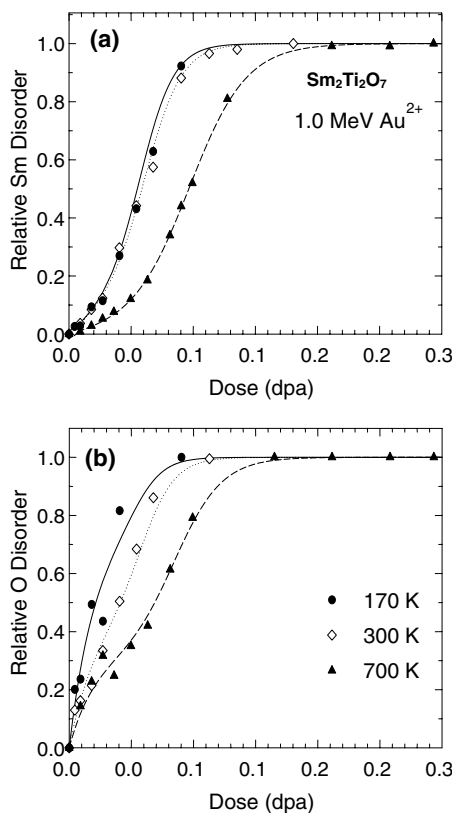


Fig. 25.7. Relative disorder at the damage peak on the (a) Sm and (b) O sublattices as a function of local dose for $\text{Sm}_2\text{Ti}_2\text{O}_7$ single crystals implanted with 1.0 MeV Au^{+} at 170, 300 and 700 K

disorder on each sublattice follows a nonlinear dose dependence and that the relative disorder on the O sublattice is higher than that on the Sm sublattice. There is little difference in damage accumulation on the Sm sublattice at 170 and 300 K. However, dynamic recovery processes dramatically reduce the rate of damage accumulation at 700 K.

The critical dose for amorphization in $\text{Sm}_2\text{Ti}_2\text{O}_7$ under irradiation by Au^{2+} [40] and Bi^+ [28] is shown in Fig. 25.8 as a function of irradiation temperature, together with the results for $\text{Gd}_2\text{Ti}_2\text{O}_7$ irradiated with Bi^+ [28]. Also included is the amorphization data point for alpha decay in $\text{Gd}_2\text{Ti}_2\text{O}_7$ doped with 3 wt% ^{244}Cm [41]. The critical temperature for amorphization, as shown in Fig. 25.8, is close to 975 K, which is similar to the onset temperature for the thermal recrystallization of Cm-doped $\text{Gd}_2\text{Ti}_2\text{O}_7$ [30]. Ion irradiation experiments accelerate the damage rates by six orders of magnitude, as compared with the 3 wt% ^{244}Cm -doped $\text{Gd}_2\text{Ti}_2\text{O}_7$. A good agreement of the amorphization dose at around room temperature (Fig. 25.8) is observed between the results of heavy-ion irradiation and the result obtained in ^{244}Cm -doped $\text{Gd}_2\text{Ti}_2\text{O}_7$ through alpha decay. This indicates that the dose-rate effect is negligible. These results provide some validation for applying models of damage accumulation and amorphization under heavy-ion irradiation to predict the long-term behavior in rare-earth titanates resulting from alpha decay.

One of the most exciting outcomes from fundamental studies of irradiation effects using ion beams has been the discovery of chemically durable and radiation-resistant $\text{Gd}_2\text{Zr}_2\text{O}_7$ and $\text{Er}_2\text{Zr}_2\text{O}_7$ pyrochlores. These materials can readily accommodate Pu on the Gd (or Er) and Zr sites [42]. In

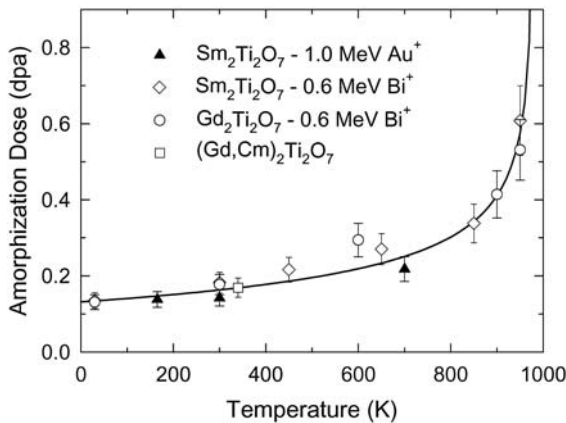


Fig. 25.8. Critical dose for amorphization of $\text{Sm}_2\text{Ti}_2\text{O}_7$ and $\text{Gd}_2\text{Ti}_2\text{O}_7$ irradiated by 1.0 MeV Au^{2+} and 0.6 MeV Bi^+ . Also included is the critical dose for amorphization in $\text{Gd}_2\text{Ti}_2\text{O}_7$ doped with 3 wt% ^{244}Cm . The solid curve is the best fit to the experimental data

general, the radiation resistance depends critically on the ability of the structure to sustain cation disorder on the A and B sites, as well as on a disordering of the oxygen vacancies. Irradiation studies have shown that the titanate–zirconate pyrochlore becomes more radiation-resistant with increasing zirconium concentration [25, 43]. Recent studies [24, 25, 39] demonstrate that pyrochlores of $\text{Gd}_2(\text{Ti}_{2-x}\text{Zr}_x)\text{O}_7$ display a dramatic decrease in susceptibility to radiation-induced amorphization with increasing Zr content. Pure zirconate end members undergo a radiation-induced transition to a disordered fluorite structure, which is highly radiation-resistant and remains crystalline to a high dose, for example 7.0 dpa (displacements per atom; see Appendix 25.A) for $\text{Gd}_2\text{Zr}_2\text{O}_7$ at 25 K [25] and 140 dpa for $\text{Er}_2\text{Zr}_2\text{O}_7$ at room temperature [24, 43]. In contrast, pure titanate end member pyrochlores are sensitive to irradiation damage and readily become amorphous (0.4 dpa [28] for $\text{Gd}_2\text{Ti}_2\text{O}_7$ at 900 K, 0.18 dpa [28] for $\text{Gd}_2\text{Ti}_2\text{O}_7$ and 0.26 dpa [24, 43] for $\text{Er}_2\text{Ti}_2\text{O}_7$ at 300 K).

25.3.3 Radiation Damage and Mixing

Ion beam mixing and the damage induced with MeV ions and keV ions are closely similar. Nuclear scattering leads to atomic displacements so that different layers of materials become ballistically mixed. In addition, non-thermalized defects can also promote diffusion. This provides a convenient method to atomically mix materials even when they form a multiphase structure under thermodynamic equilibrium conditions. An example of the use of MeV ion beam mixing is modification of the band gap in quantum well structures by mixing the quantum well with the adjacent layers, followed by epitaxial regrowth to recover the radiation damage. Figure 25.9 illustrates the mixing of $\text{InP}/\text{Ga}_{0.25}\text{In}_{0.75}\text{As}/\text{InP}$ quantum wells using channeled 10 MeV $^{69}\text{Ga}^+$ ions [44]. This resulted in a blue shift (band gap increase) of about 30 meV (Fig. 25.9(b)). A novel feature was the lateral writing of the degree of blue shift by using a thin Au dechanneling mask to modulate the degree of channeling and hence the degree of ion beam mixing [44] (Fig. 25.9(b)). An alternative route to lateral definition is to use focused ion beams. Low-energy focused 0.16 MeV Si^{2+} ions have been used for direct writing of waveguides in a $\text{GaAlAs}/\text{GaAs}$ superlattice material for distributed-feedback (DFB) lasers [45], whilst 1 MeV Co has been used to synthesize CoSi_2 structures [46].

25.3.4 Electron-Induced Interactions

The strong coupling of the kinetic-energy loss of a penetrating ion with the electronic system of the material may be used to modify materials where electronic excitations can lead to latent or direct material changes. Generally, in metals, the electronic relaxation time is so short that the electronic excitation rapidly dissipates as phonons before atomic displacement takes place.

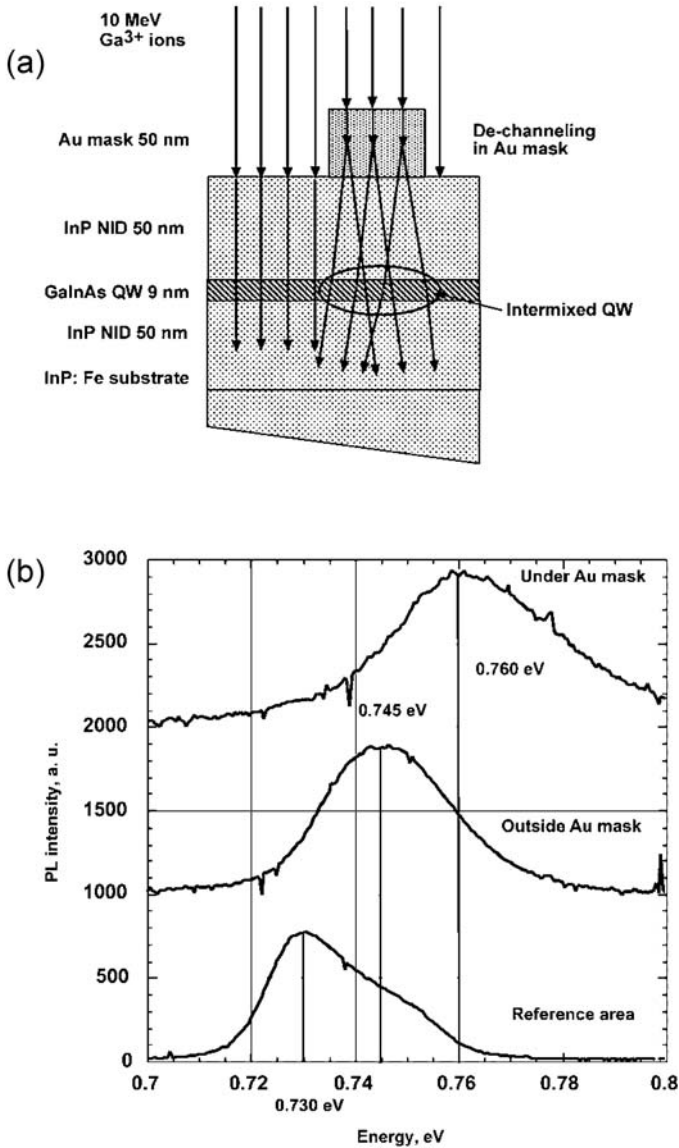


Fig. 25.9. (a) Schematic illustration of switched-channeling MeV ion beam mixing of an InP/Ga_{0.25}In_{0.75}As/InP quantum well structure. (b) Photoluminescence measurement of the band gap change [44]

In materials with a band gap, the excitation from the ion and the secondary electrons and recoils can introduce a number of changes, such as excitation across the band gap as well as to and from defect levels. If the energy available exceeds the amount required to cleave bonds, these may be broken at random. These broken bonds can relax by forming new bond configurations, which in turn change material properties such as refractive index, optical absorption, conductivity, chemical resistance and density. MeV ions have been employed to introduce changes in the refractive index in glasses [47, 48] and introduce photoluminescence centers in silica [49].

An important class of materials where electron-mediated interactions are significant is polymers. These form the basis of the organic resists used in lithography, discussed below. The polymer can be either a positive resist (such as PMMA), where ion irradiation causes chain scission [50] that locally reduces the molecular weight of the polymer chains, or a negative resist (such as epoxies) where crosslinking [50] is induced. After exposure, the latent image is developed in a selective solvent that preferentially dissolves low-molecular-mass chains (positive resists) or noncrosslinked chains (negative resists).

25.3.5 Nanoscale Lithography with MeV ion Beams

Nanoscience and nanotechnology is one of the most exciting and important areas of research today. This is a truly cross-disciplinary research field with a major impact on forefront research in fields as diverse as cell biology and medicine and quantum electronics and optics. One of the big challenges in nanoscale engineering is the development of technologies that can be scaled up from the single devices produced in the research laboratory, first to fabricate large numbers of them in defined configurations to realize circuits, and subsequently to manufacture these circuits in industrial quantities. The active parts of quantum devices based on quantum confinement and tunneling are typically less than 5 nm in size. High-packing-density circuits using these components will require lithographic processing of interconnects on a similar size scale. In semiconductor technology, the well-known Moore's law predicts a halving of feature sizes and doubling of the number of devices per circuit every 18 months. At the time of writing, very large-scale integration (VLSI) devices are routinely produced with a 0.13 μm line width, and 0.09 μm is at the pilot stage. The EU Nanoelectronics Roadmap [51] points to industrial maturity for 70 nm line width features by 2008. Similar trends are seen in magnetic information storage, where the size of the domains (bits) is shrinking to similar sizes, and this implies that new approaches are needed to keep the domains separated. In order to realize this, there is a clear need for new lithographic tools that can extend the writing capability to better resolution than that achievable with the current industry-piloted extreme ultraviolet (EUV) projection lithography (30–50 nm [51]).

Focused MeV ions, such as used in the nuclear microprobe discussed in Chap. 26, can also be used for the lithographic writing of latent or permanent patterns of modified material. Over the past 5 years or so, the proton-beam-writing (PBW) technique that has been pioneered by the Singapore group [53, 54] has attracted interest, not least because of its 3-dimensional writing capability [55]. In PBW, the protons are focused to a 30 nm–2 μ m spot size that is used to write a latent pattern by modifying a resist polymer (discussed above) coated on a substrate material. The writing is done by modulating the intensity of the beam as the spot is scanned over the sample. After developing, the resist is fully developed at points where the dose (25.6) exceeds the *clearing dose*. The clearing dose is the dose needed to just produce a fully developed latent pattern so that, after development, the polymer is either completely removed (positive resist) or completely retained (negative resist).

The smallest feature size that can be written is determined by the extent of the region where the dose exceeds the clearing dose. The minimum size of a feature is then governed by the spreading of the dose about the axis of the beam. This is, in essence, the convolution of an extrinsic contribution from the beam focus profile and an intrinsic contribution associated with the spreading of the ion beam, secondary recoils and electrons in the target. Reference to Fig. 25.2 shows that in the outer few hundred nanometers the radial spreading of the ion beam is a nanometer or so. This, combined with the sharp falloff of the dose from secondary electrons $D(r)$ within a 10 nm radius of the track (Fig. 25.3(b)), allows extremely high-aspect structures with 60 nm wide walls in 10 μ m thick resist to be written [56]. An example is shown in Fig. 25.10, which shows a pattern written in negative SU-8 resist for use as an etch mask for plasma etching of thin silicide lines. In this case a negative resist was chosen to selectively protect the surface during subsequent plasma etching. Another important factor is the proximity exposure effect. If

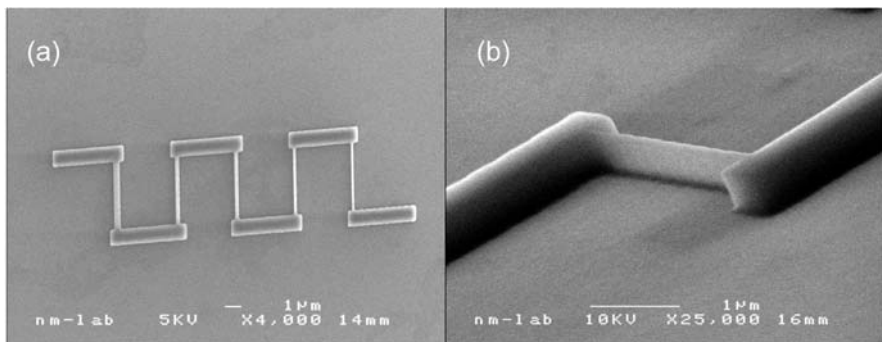


Fig. 25.10. Plasma etch mask produced in SU-8 resist [52] by PBW. The smallest vertical line in (a) is 120 nm wide. (b) Tilted view of the 120 nm line (H.J. Whitlow, I. Maximov, L. Montelius, J. van Kan, A. Bettiol and F. Watt, unpublished data)

the pattern has large, closely spaced exposed areas, the unirradiated region in between receives a dose because of spreading of the dose distribution beyond the edges of the irradiated regions. For PBW, this is much smaller than is the case for the 10–40 keV electrons used in conventional focused-electron-beam lithography (EBL) [57]. This is because, as discussed in Sect. 25.2.4, in PBW the majority of δ -electrons are directed perpendicular to the beam with energies close to zero and hence have short ranges. This gives an extremely sharp radial dose distribution $D(r)$ as shown in Fig. 25.3(b). The small proximity effect for PBW facilitates writing high-spatial-density patterns that would be extremely difficult to write using conventional EBL because of the proximity effect. An example of a high-spatial-density structure that has been written with PBL is shown in Fig. 25.11. This shows a metal pattern produced by metal lift-off using a 170 nm thick positive PMMA resist. Here a positive resist was used, and consequently the area where metal was to be deposited was irradiated. The pattern is a prototype interdigitated electrode array for electrochemical biosensors capable of assaying specific biomolecules such as antigens, antibodies and hormones [57]. The use of PBW enabled wide electrodes with narrow gaps to be written, which is particularly difficult using conventional EBL. Small (~ 100 nm) gaps can also be written using PBW in negative resist, as shown in Fig. 25.12, which shows an Au-coated finger structure in negative SU-8 resist, where the 100 nm gap between the fingers is clearly seen.

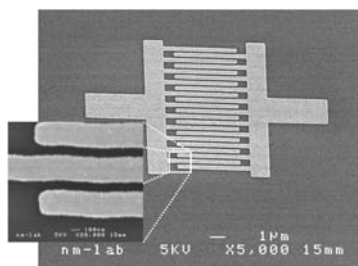


Fig. 25.11. Interdigitated metal pattern with nanoscale electrode gaps for biosensor development, fabricated using PBW. The metal pattern is 30 nm Au/3 nm Ti on a SiO_2/Si substrate. PBW was used to produce an aperture pattern in a 170 nm thick positive PMMA resist prior to metal evaporation, followed by lift-off of metal by dissolution of the resist from unexposed regions in hot acetone [57]

PBW is not restricted to writing latent images. Patterns of direct changes in the refractive index and photoluminescence of glassy materials can be written [47–49]. The changes in molecular structure in polymers subject to MeV ion bombardment also cause direct changes in the density and refractive index [58]. These changes are most pronounced in the end-of-range damage. A spectacular application of this phenomenon is the direct writing of buried

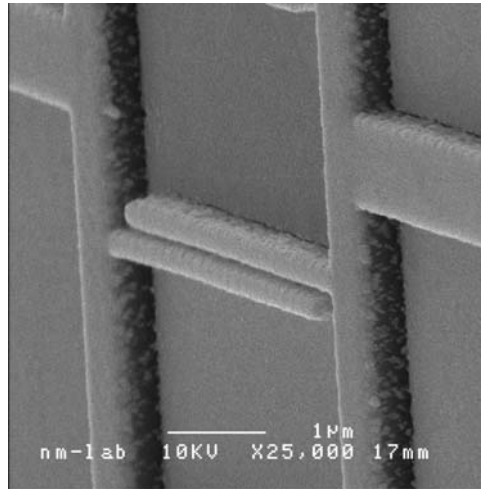


Fig. 25.12. Detail of PBW finger structure in negative resist coated with 100 nm Au (H.J. Whitlow, I. Maximov, L. Montelius, J. van Kan, A. Bettiol and F. Watt, unpublished data)

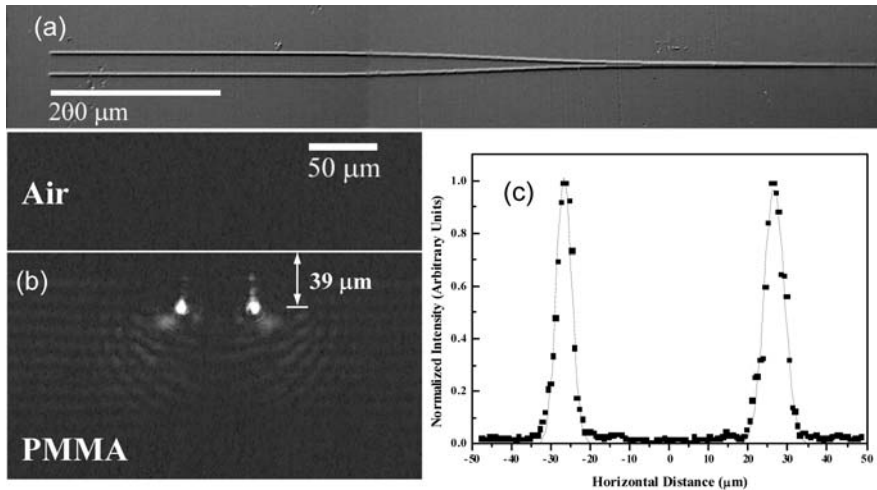


Fig. 25.13. Buried y-branch optical-waveguide structure produced in PMMA by direct writing with 1.5 MeV protons. (a) Differential interference contrast image of the waveguide structure. (b) 633 nm laser light emitted from the branches. (c) Normalized intensity of the light distribution in (b) (Reprinted from [59], copyright 2003, with permission from Elsevier)

optical-waveguide structures [59]. Figure 25.13 shows an optical-waveguide structure fabricated by direct writing with 1.5 MeV $^1\text{H}^+$ in PMMA by the Singapore group [59], the emitted 633 nm light observed with a CCD camera from the branches of a 2 mm y-branch waveguide, and the corresponding intensity distributions.

Acknowledgments

We are grateful to our collaborators at Pacific Northwest National Laboratory, The Nanometer Consortium in Lund and The Centre for Ion Beam Applications at the National University of Singapore for assistance, guidance and permission to use figures. One of the authors (YZ) acknowledges the support of the Division of Materials Sciences and Engineering Office of Basic Science, U.S. Department of Energy.

25.A Appendix: Nomenclature in SI Units

Displacements per atom (dpa). The mean number of times an atom is displaced from its atomic site. This quantity is a measure of the dose in terms of the mean number of displacements from its site that an atom undergoes.

Dose. SI-defined unit, which is the energy absorbed per unit mass of material [60,61]. The unit is the gray (Gy) = 1 J kg^{-1} , or, in terms of SI base units [62], $\text{m}^2 \text{ s}^{-2}$. A related quantity in radiation physics with the same units is the kerma K (kinetic energy released per unit mass) [63] which is the sum of all of the kinetic energy of all charged particles created by uncharged ionizing particles. It is a measure of the dose deposited by uncharged particles, such as X-ray photons.

Electron volt (eV). Energy is often expressed in electron volts; an electron volt is the energy gained by an electron on passing through a potential difference of 1 volt. The eV is not an SI unit but it is accepted for use with the SI [62].

G value. The number of specified chemical events in an irradiated substance produced per 100 eV of energy absorbed from ionizing radiation.

Linear energy transfer (LET). The average energy locally imparted to a medium by a charged particle of specified energy, per unit distance traversed. The LET is not a unit of fundamental significance.

Mass stopping power ϵ^* . This quantity is the stopping power, where the distance is expressed in terms of mass per unit area: $\epsilon^* = (1/\rho) dE/dx$.

Particle flux $\dot{N}$. This is the ratio of the infinitesimal change in the particle number dN to an infinitesimal time interval dt . $\dot{N} = dN/dt$. The units are s^{-1} .

Particle fluence Φ . This is defined [61] as the number of particles dN crossing the surface of a sphere with cross-sectional area dA . $\Phi = dN/dA$. The units are m^{-2} ; however, for convenience quantity is often given in terms of cm^{-2} . For ion beam applications where parallel ions are incident on a planar reference surface, care should be exercised because the particle fluence does not explicitly define the direction. In this case it may be advisable to specify the beam direction, for example “the fluence of ions normally incident on the surface was $3 \times 10^{18} m^{-2}$ ”. It is recommended that the term “particle fluence” or “ion fluence” is used to avoid confusion with the use of “fluence” in radiation chemistry to specify the dose rate [64].

Particle fluence rate ϕ . This is the time differential of the particle fluence; $\phi = d^2N/(dt dA)$. The units are $m^{-2} s^{-1}$. Often the term “particle flux density” has been used; however, “particle flux rate” is preferred [61].

Particle fluence differential in energy Φ . This is the fluence of particles in the energy interval E to $E + dE$. $\Phi(E) = d\Phi/dE$ [61].

Stopping cross section ε . The stopping cross section is the energy loss per atom per unit area. $\varepsilon = dE/(N_a dx)$, where N_a is the number of atoms per unit volume. It is usually given in units of $eV/10^{15} at cm^{-2}$. As a rule of thumb, the value in these units is of the order of the energy loss per monolayer in a typical solid material [64].

Stopping force (stopping power) dE/dx . This quantity is the mean energy loss per unit path length, with units $J m^{-1}$, and has the dimensions of a force and should correctly be termed “stopping force” [65]. For practical applications, the stopping force is often given in units of eV/nm or $eV/\mu m$. The term “stopping power” has been widely used in the literature.

References

1. N. Nastasi, J.W. Mayer, J.K. Hirvonen: *Ion-Solid Interactions: Fundamentals and Applications* (Cambridge University Press, London 1996)
2. J.F. Ziegler, J.P. Biersack, U. Littmark: *The Stopping and Ranges of Ions in Matter*, vol. 1 (Pergamon, New York 1985); SRIM 2003.24, <http://www.srim.org>.
3. M.W. Thompson: *Defects and Radiation Damage in Metals* (Cambridge University Press, London 1969).
4. H.H. Andersen: *Forelæsning over Strålningsbeskadelse i faste stoffer 1. Defektproduktion og defektkonfigurationer* (Det fysiske institut, Aarhus universitet, Aarhus 1984)
5. P. Sigmund: Nucl. Instr. Meth. B **135**, 1 (1998)
6. N. Bohr: Det. Kgl. Danske Videnskabernes Selskab, Matematisk-fysiske Meddelelser **18**, Nr. 8 (1948)
7. R. Spohr: In *Ion Tracks in Microtechnology, Principles and Applications*, ed. by K. Bethge (Vieweg, Braunschweig 1990) p. 112

8. P. Sigmund: Nucl. Instr. Meth. B **164–165**, 401 (2000)
9. P. Hovington, D. Drouin, R. Gauvin: Scanning **19**, 1 (1997)
10. D. Drouin, P. Hovington, R. Gauvin: Scanning **19**, 20 (1997)
11. P. Hovington, D. Drouin, R. Gauvin, D.D. Joy, N. Evans: Scanning **19**, 29 (1997)
12. Y. Zhang, W.J. Weber, W. Jiang, A. Hallén, G. Possnert: J. Appl. Phys. **91**, 6388 (2002)
13. W. Jiang, W.J. Weber: Phys. Rev. B **64**, 125206 (2001)
14. F. Gao, W.J. Weber: Phys. Rev. B **63**, 054101 (2001)
15. L. Henry, M.F. Barthe, C. Corbel, P. Desgardin, G. Blondiaux, S. Arpiainen, L. Liskay: Phys. Rev. B **67**, 115210 (2003)
16. X.D. Chen, C.C. Ling, S. Fung, C.D. Beling, M. Gong, T. Henkel, H. Tanoue, N. Kobayashi: J. Appl. Phys. **93**, 3117 (2003)
17. W.J. Weber, N. Yu, L. M. Wang: J. Nucl. Mater. **253**, 53 (1998)
18. R.E. Avila, J.J. Kopanski, C.D. Fung: J. Appl. Phys. **62**, 3469 (1987)
19. Y. Zhang, W.J. Weber, W. Jiang, C.M. Wang, A. Hallén, G. Possnert: J. Appl. Phys. **93**, 1954 (2002)
20. H.L. Tuller: J. Phys. Chem. Solids **55**, 1393 (1994)
21. B.J. Wuensch, K.W. Eberman, C. Heremans, E.M. Ku, P. Onnerud, E.M.E. Yeo, S.M. Haile, J.K. Stalick, J.D. Jorgensen: Solid State Ionics **129**, 111 (2000)
22. J.B. Goodenough, R.N. Castellano: Solid State Chem. **44**, 109 (1982)
23. S.J. Korf, H.J.A. Koopmans, B.C. Lippens, A.J. Burggraaf, P.J. Gellings: J. Chem. Soc. Faraday Trans. **83**, 1485 (1987)
24. K.E. Sickafus, L. Minervini, R.W. Grimes, J.A. Valdez, M. Ishimaru, F. Li, K.J. McClellan, T. Hartmann: Science **289**, 748 (2000)
25. S.X. Wang, B.D. Begg, L.M. Wang, R.C. Ewing, W.J. Weber, K.V. Godivan Kutty: J. Mater. Res. **14**, 4470 (1990)
26. R. C. Ewing, W. J. Weber, and J. Lian, J. Appl. Phys. **95**, 5949 (2004)
27. R.C. Ewing, W.J. Weber, W. Lutze: In *Diposal of Weapon Plutonium*, ed. by E.R. Merz, C.E. Walter (Kluwer Academic Dordrecht 1996) p. 65
28. B.D. Begg, N.J. Hess, W.J. Weber, R. Devanathan, J.P. Icenhower, S. Thevuthasan, B.P. McGrail: J. Nucl. Mater. **288**, 208 (2001)
29. W.J. Weber, J.W. Wald, Hj, Matzke: Mater. Lett. **3**, 173 (1985)
30. W.J. Weber, J.W. Wald, Hj, Matzke: J. Nucl. Mater. **138**, 196 (1986)
31. B.D. Begg, W.J. Weber, R. Devanathan, J.P. Icenhower, S. Thevuthasan, B.P. McGrail: Ceram. Trans. **107**, 553 (2000)
32. R.C. Ewing, W.J. Weber, F.W. Clinard Jr.: Prog. Nucl. Energy **29**, 63 (1995)
33. W.J. Weber, R.C. Ewing, C.R.A. Catlow, T. Diaz de la Rubia, L.W. Hobbs, C. Kinoshita, Hj. Matzke, A.T. Motta, M. Nastasi, E.K.H. Salje, E.R. Vance, S.J. Zinkle: J. Mater. Res. **13**, 1434 (1998)
34. F.W. Clinard Jr., D.E. Peterson, D.L. Rohr, L.W. Hobbs: J. Nucl. Mater. **126**, 245 (1984)
35. G.R. Lumpkin: J. Nucl. Mater. **289**, 136 (2001)
36. R.C. Ewing, L.M. Wang: Nucl. Instr. Meth. Phys. Res. B **65**, 319 (1992)
37. S.X. Wang, L.M. Wang, R.C. Ewing, G.S. Was, G.R. Lumpkin: Nucl. Instr. Meth. Phys. Res. B **148**, 704 (1999)
38. W.J. Weber, N.J. Hess: Nucl. Instr. Meth. Phys. Res. B **80–81**, 1245 (1993)
39. B.D. Begg, N.J. Hess, D.E. McCready, S. Thevuthasan, W.J. Weber: J. Nucl. Mater. **289**, 188 (2001)

40. Y. Zhang, W.J. Weber, V. Shutthanandan, R. Devanathan, S. Thevuthasan, G. Balakrishnan, D.M. Paul: *J. Appl. Phys.* **95**, 2866 (2004)
41. W.J. Weber, R.C. Ewing: In *Scientific Basis for Nuclear Waste Management XXV*, ed. by B.P. McGrail, G.A. Cragnolino, Mater. Res. Soc. Symp. Proc. 713, (Materials Research Society, Warrendale, PA 2002) p. 443
42. R.E. Williford, W.J. Weber: *J. Nucl. Mater.* **299**, 140 (2001)
43. K.E. Sickafus, L. Minervini, R.W. Grimes, J.A. Valdez, T. Hartmann: *Radiat. Eff. Defects Solids* **155**, 133 (2001)
44. T. Winzell, I. Maximov, L. Landin, Y. Zhang, A. Gustafsson, L. Samulelsson, H.J. Whitlow: *Semicond. Sci. Technol.* **16**, 889 (2001)
45. A. Steckl, P. Chen, H.E. Jackson, A.G. Choo, X. Cao, J.T. Boyd, M. Kumar: *J. Vac. Sci. Technol. B* **13**, 2570 (1995)
46. J. Meyer, U. Weidenmüller, P. Baving, H. Röchen, A. Stephan, H.H. Bukow, C. Rolfs: *Nucl. Instr. Meth. B* **161–163**, 898 (2000)
47. M. Hattori, Y. Ohki, M. Fujimaki, T. Souno, H. Nishikawa, E. Watanabe, M. Oikawa, T. Kamiya, K. Arakawa: *Nucl. Instr. Meth. B* **210**, 272 (2003)
48. I. Rajta, I. Gómez-Morilla, M.H. Abraham, Á.Z. Kiss: *Nucl. Instr. Meth. B* **210**, 260 (2003)
49. T. Souno, H. Nishikawa, M. Hattori, Y. Ohki, E. Watanabe, M. Oikawa, T. Kamiya, K. Arakawa: *Nucl. Instr. Meth. B* **210**, 277 (2003)
50. A. Chapiro: *Nucl. Instr. Meth. B* **32**, 111 (1988)
51. R. Compañó: *Technology Roadmap for Nanoelectronics*, 2nd ed., European Commission IST program Future and Emerging Technologies (2000)
52. Micro Chem, 1254 Chestnut Street, Newton, MA 02464, USA: http://www.microchem.com/products/su_eight.htm
53. J.A. van Kan, T.C. Sum, T. Osipowicz, F. Watt: *Nucl. Instr. Meth. B* **161–163**, 366 (2000)
54. T. Osipowicz, J.A. van Kan, T.C. Sum, J.L. Sanchez, F. Watt: *Nucl. Instr. Meth. B* **161–163**, 83 (2000)
55. J.A. van Kan, J.L. Sanchez, T. Osipowicz, F. Watt: *Microsyst. Technol.* **6**, 82 (2000)
56. J.A. van Kan, A.A. Betiol, F. Watt: *Appl. Phys. Lett.* **83**, 1629 (2003)
57. H.J. Whitlow, M.L. Ng, V. Auželyté, I. Maximov, L. Montelius, J.A. van Kan, A. Bettiol, F. Watt: *Nanotechnology* **15**, 223 (2003).
58. A.A. Betiol, T.C. Sum, J.A. van Kan, F. Watt: *Nucl. Instr. Meth. B* **210**, 250 (2003)
59. T.C. Sum, A.A. Betiol, H.L. Seng, I. Rajta, J.A. van Kan, F. Watt: *Nucl. Instr. Meth. B* **210**, 266 (2003)
60. ICRU Report 33, *Radiation Quantities and Units* (International Commission on Radiation Units and Measurements, 1980)
61. <http://www.npl.co.uk/ionrad/quantities.html>
62. Bureau International des Poids et Mesures, Pavillon de Breteuil, 92312 Sèvres Cedex, France: <http://www.bipm.fr/enus/>
63. <http://rkb.home.cern.ch/rkb/PH14pp/node154.html#153>
64. A.D. McNaught, A. Wilkinson: *Compendium of Chemical Terminology: The Gold Book*, 2nd edn. (International Union of Pure and Applied Chemistry), Internet edition, <http://www.iupac.org/publications/compendium/index.html>
65. P. Sigmund: *ICRU News*, June 2000, p. 5

26 Ion Beam Analysis

K.G. Malmqvist

Division of Nuclear Physics, Lund Institute of Technology at Lund University,
P.O. Box 118, 22100 Lund, Sweden
`klas.malmqvist@pixe.lth.se`

26.1 Introduction

The common denominator in the field of ion beam analysis (IBA) is the use of accelerated charged particles, interacting with a specimen and hence producing radiation suitable for analytical purposes. Here the definition of IBA is limited to charged particles with an initial energy of the order of MeV/u, i.e. ions produced in small and medium-size particle accelerators.

The basic mechanisms behind the IBA techniques vary significantly but all depend strongly on the ion–solid interactions occurring when a specimen is bombarded with an ion beam. The theory and physics behind the atomic collisions that govern the ion trajectories in the specimen are described in detail in Chap. 24. Many decades of careful basic physics studies of such interactions have enabled the IBA field to make use of the databases created on reaction cross sections, energy loss, energy straggling etc. [1,2]. The well-defined ion range and almost straight ion tracks obtained when using MeV ions are utilized in the various IBA methods. The penetration depth in most materials is of the order of tens of μm , and the effectively probed analytical depth will depend strongly on the particular analytical process used. Compared with other surface and near-surface analytical methods, IBA methods must generally be characterized as “surface-oriented” techniques, giving useful information from the first monolayers all through to tens of μm depth.

One serious issue of IBA is that the physical processes involved along the ion track will, in addition to producing useful radiation, also create modification and damage in the material. Such phenomena are utilized in the systematic ion beam modification and processing of materials, as described in Chap. 25. However, these processes could also pose a serious problem when a particular analytical application requires nondestructive analysis.

26.2 Experimental Facilities

There are many different types of IBA facilities, varying from those that are part of a large accelerator complex comprising almost all aspects of both basic and applied physics, to very small IBA labs around a small “turnkey”

accelerator, for instance specializing in specific routine analyses in large numbers.

Apart from a few cyclotron laboratories, two kinds of accelerators dominate IBA: single-ended and tandem electrostatic accelerators (see Chap. 4). The actual ion bombardment of specimens normally occurs in vacuum chambers fitted with an array of detectors suitable for different kinds of radiation. The complex multiparameter systems necessary for implementation of a number of detectors for different radiation components are often designed individually at each accelerator laboratory, while the computer codes used for evaluating spectral information and for simulation of ion-solid interactions are limited to a few well-established computer packages used at most laboratories [3, 4]. Lately, there have been attempts to build general-purpose software tools that will integrate several different analytical methods and the simulation of various compositions of materials. A review on how such a tool can be applied in elemental profiling of thin films is given in [5]. For special applications, for instance very near-surface studies, there may be requirements for sample heating/cooling, other types of specimen modifications, use of ultrahigh vacuum etc. Other specimens may be fragile and radiation-sensitive, requiring a minimization of heating and radiation dose. One approach that has been very successful in combination with valuable art objects is to extract the ion beam out from the vacuum system into air or other nonvacuum atmospheres [6, 7]. This type of analysis is particularly useful when analyzing large and sensitive objects and when investigating compounds of high vapor pressure.

A very important development within IBA is the use of highly collimated and focused ion beams for microscopic investigations. In fact, most of the traditional IBA methods are now applied also in such nuclear microprobes, where the high lateral resolution has resulted in completely new approaches to many analytical problems. The nuclear microprobe can be used to obtain quantitative analytical results from specific points on a specimen, but also for imaging a sample both qualitatively and quantitatively: the narrow ion beam is rapidly scanned over the specimen and, by measuring for a very short time in each pixel with sophisticated computer assistance, the results can, on-line or off-line, be related to the position on the specimen.

The greatest technical challenges in setting up a nuclear-microprobe system with a small beam focus originate from limitations in the present ion sources and technical difficulties in designing strongly focusing devices with low aberrations. One crucial parameter is the beam brightness B , a measure of the number of ions that pass through a given area with a given maximum divergence at a given energy. In a nuclear microprobe, brightness is defined as

$$B = \frac{i}{A_0(A_a/D^2)E_0} \quad (26.1)$$

where i is the beam current at energy E_0 that will pass through an object collimator of area A_0 and an aperture collimator of area A_a at a distance

D from the object collimator. The commonly adopted radio-frequency and duoplasmatron sources both suffer from relatively low brightness and are hence not ideal for nuclear microprobes. The development and adaptation to MeV beam optics of high-brightness sources such as liquid-metal or field-ionization sources have so far been of limited success [8].

A detailed investigation into the various types of focusing devices [9] is beyond the scope of this presentation but the clearly dominating type of device is the magnetic quadrupole in doublet, triplet or quadruplet configurations. A few laboratories use superconducting magnetic solenoids or electrostatic quadrupoles. In order to predict the smallest possible spot size on the specimen, one must consider the detailed action on the ions of the focusing elements and calculate the effects of the imperfect image formation. In addition to inherent components such as chromatic and spherical aberrations, parasitic aberrations originate from various errors in the design of optical elements. The most prominent factors are misalignment, power supply ripple and sextupole field components. Details of how to minimize these effects [10] are beyond the scope of this presentation. Each type of lens configuration has its virtues and drawbacks, and probe formation requires a tedious optimization process. In addition, stringent requirements on vacuum, electromagnetic stray fields, vibrations etc have to be fulfilled for minimum-size probe formation. In practical applications, most laboratories use a beam size $\geq 1 \mu\text{m}$.

Based on lower-energy (200 keV) probes, commercially available high-resolution focused-ion-beam systems are already in use [11] that can perform RBS analysis (see Sect. 26.3) with better than 100 nm lateral resolution. Recently, in order to produce MeV ion beams with sizes $< 100 \text{ nm}$, work has started on the development of new high-demagnification systems in several laboratories [12, 13]. When the IBA methods are extended into such a very high-lateral-resolution mode, it is essential to develop suitable test structures. Such an attempt has been made by the Leipzig group [14], in a laboratory developing a nuclear “nanobeam” system. By growing a semiconductor heterostructure in GaInP on a substrate of GaAs, a sharp edge suitable for beam size calibration was obtained.

To make the most of such systems, the technical requirements on low vibrations etc. mentioned above, are even higher. In addition, the useful ion currents will be very low (tens of pA), which degrades counting statistics in the radiation detectors. Hence, several systems for very large-solid-angle detection are currently under development and implementation [15]. Figure 26.1 shows an eight-crystal X-ray detector (Canberra) that has a very large total solid angle with a maintained state-of-the-art energy resolution. The ion beam direction is shown by the arrow. Although such a detector system requires more complex data acquisition systems, it facilitates trace element determination using very low beam currents.

Also, in a nuclear microprobe it is possible to extract the focused ion beam into an atmosphere, albeit at the cost of degraded beam resolution because of

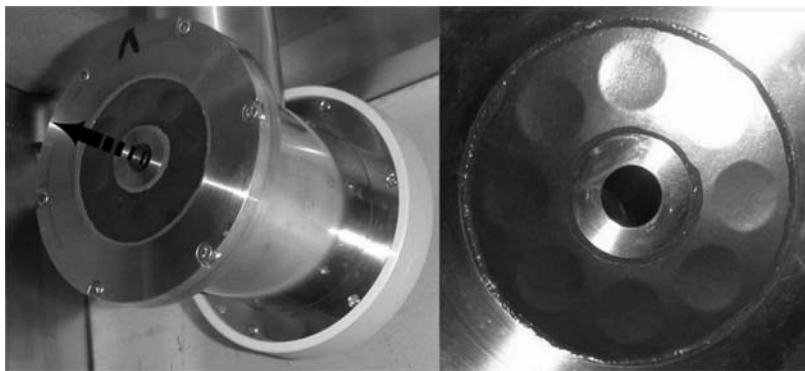


Fig. 26.1. Eight-crystal Ge X-ray detector (Canberra[®]) designed for enhanced X-ray detection efficiency in low-beam-current experiments [15]. The arrow indicates the ion beam direction. The resolution of each individual crystal is better than 150 eV at 5.9 keV

scattering in the exit foil and gas molecules. To reduce this degradation, the distance that the focused ion beam travels outside the vacuum is minimized by arranging very tight geometries between the detectors and the specimen.

26.3 Analytical Techniques

The great diversity of IBA methods makes it necessary to limit the following text to introducing a selection of existing techniques. The IBA techniques make use of a variety of physical interactions taking place after an ion impinges on the specimen surface. Since interactions can be both atomic and nuclear processes, the cross sections for the interactions vary by several orders of magnitude. As will be discussed in some detail later in this chapter, the technical amenities available will play a crucial role for each analytical procedure. The type of results produced varies from qualitative structural information, through semiquantitative elemental information on major constituents, quantitative trace elements and isotopic composition, to highly resolved depth profiling and lateral information. The nuclear microprobe described in Sect. 26.2 can normally integrate most of the existing IBA methods. The choice of IBA method will be determined by the specimens investigated and the specific information required. Owing to their reasonably nondestructive character, IBA techniques are, with great benefit, combined with complementary analytical methods.

26.3.1 Particle-Induced X-ray Emission (PIXE)

Ions of an energy greater than a few MeV/u impinging on a specimen have a high production cross section for inner-shell vacancies in the matrix atoms.

The subsequent deexcitation results either in a characteristic X-ray or in an Auger electron. Hence X-ray spectrometry can be performed for all elements heavier than sodium [16] (Auger electrons dominate over X-rays and attenuation is too high for lighter elements). Cross sections for inner-shell ionization are of importance when using PIXE. Several theoretical approaches have been used to calculate the cross sections, but the plane wave Born approximation model, which applies perturbation theory to a transition from an initial state to a final state, has been developed furthest, particularly by Brandt and coworkers [17, 18]. These authors incorporated a series of modifications into the initial model to correct for its inherent approximations, and this resulted in the so-called ECPSSR treatment of K and L shell ionization cross sections, which takes care of the deflection of the projectile due to the nuclear Coulomb field (C), perturbation of the atomic stationary states (PSS) by the projectile, relativistic effects (R), and energy loss (E) during the collision.

Calculation of the X-ray production cross section from the ionization cross sections involves additional atomic parameters: fluorescence yields and fractional radiative widths in the case of the K shell, and fluorescence yields, Coster–Kronig yields and fractional radiative widths for the L shell. As to the K fractional radiative widths, either the experimental data of Salem et al. [19] or the theoretical values of Scofield [20] are used. For the case of the L X-rays Campbell [21] advocates the use of the radiative widths of Scofield [22] and the fluorescence and Coster–Kronig yields of Chen et al. [23] in combination with the tabulated L ionization cross sections of Chen and Crasemann [24].

A practical alternative to tabulated theoretical ionization and X-ray production cross sections, particularly with computer calculations in mind, is analytical formulas which are obtained by fitting polynomial expressions to theoretical or empirical cross-sectional data [25, 26]. For the proton energy range of 1–4 MeV typically used in PIXE, the ionization (and X-ray production) cross sections increase with increasing proton energy and decrease with increasing atomic number of the target atom (see Fig. 26.2).

Continuous-Photon-Background Production

The characteristic X-ray lines in a PIXE spectrum are superimposed on a continuous background (see Fig. 26.3). The interactions of charged particles and matter occur mainly through inelastic encounters with bound electrons. This results in electron excitation and ionization but also in secondary phenomena that contribute to the background, which is very important since the peak-to-background ratio for each characteristic X-ray line will determine the detection limits for the PIXE method.

Electron Bremsstrahlung

Electron bremsstrahlung is the major background component. In the mid 1970s, Folkmann et al. [27] and various coworkers made very significant

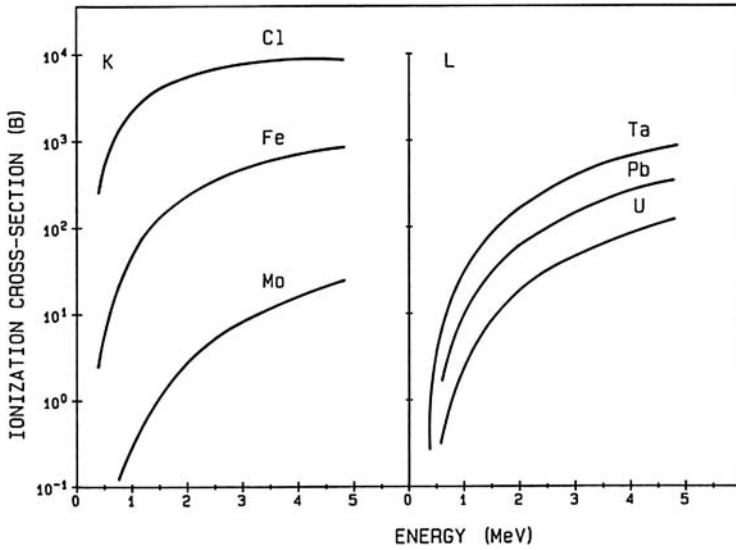


Fig. 26.2. K and L ionization cross sections vs. proton energy for several elements

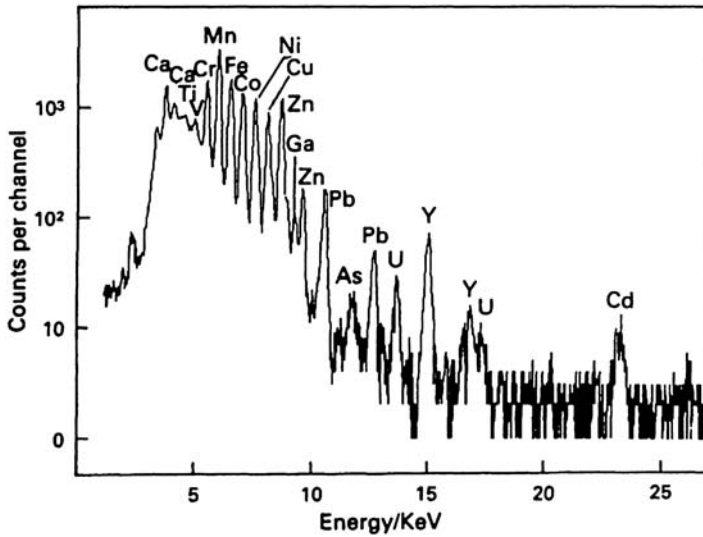


Fig. 26.3. Typical PIXE spectrum, showing several characteristic X-ray lines superimposed on a continuous background

contributions to our understanding of the continuum background in PIXE, which were later complemented and extended by Ishii and Morita [28]. Electron bremsstrahlung in PIXE originates essentially from three processes: quasi-free-electron bremsstrahlung, secondary-electron bremsstrahlung and atomic bremsstrahlung. Besides the energy of the impinging ion, the matrix composition of the target plays a critical role in both the shape and the intensity of the electron bremsstrahlung background. The intensity of this background per unit mass thickness is roughly proportional to the average Z of the matrix. The electron bremsstrahlung is emitted anisotropically and is lower at forward and backward angles than at 90° . For this reason, the X-ray detector is frequently positioned at an angle of 135° relative to the ion beam.

Projectile Bremsstrahlung

While passing through the target, the incident charged particles may be decelerated in the Coulomb field of atomic nuclei and thereby give rise to the emission of projectile bremsstrahlung [27]. Although this process is negligible within the context of the projectile energy loss, it contributes to some extent to the continuum background in the PIXE spectra. It is much less important than the electron bremsstrahlung and only becomes significant at photon energies above 10–20 keV.

Experimental Details

In a typical setup for broad-beam PIXE analysis, several samples are simultaneously inserted into the vacuum, to facilitate a high sample throughput. PIXE spectra have most of their characteristic and continuous X-ray intensity in the low-energy region, so that the spectral shape can advantageously be modulated by placing an X-ray absorber between the sample and the detector. The aim of an absorber is to reduce or eliminate unwanted continuum background and/or intense X-ray peaks and their associated pileup peaks and at the same time to allow bombardment at higher beam intensities, so that the elements of interest can be measured in shorter bombardment times and with fewer spectral interferences. In many cases it is preferable to allow a certain fraction of the low-energy X-rays to pass on onto the detector. This can elegantly be realized by resorting to pinhole absorbers. Also, more sophisticated designs consisting of a combination of various layers with different thicknesses and pinhole diameters have been used [29].

The energy-dispersive detectors used for PIXE are the same as those in other X-ray spectrometry techniques. They typically have a sensitive area from 10 to 80 mm². In ion beam analysis systems based on nuclear microprobes, the ion current for very narrow beams may be very low (tens of pA), and hence new multicrystal detectors are being developed [15] (Fig. 26.1). The ion beam passes through the center of the detector (see arrow) and

the crystals are oriented annularly, facing the specimen surface. The amplifiers/pulse processors in energy-dispersive X-ray spectroscopy require large time constants for optimum energy resolution. Unfortunately, large time constants also imply that pulse pileup becomes a serious problem at relatively low count rates. It is therefore common practice to incorporate an electronic pileup rejector and its associated dead-time correction circuitry in the pulse-processing chain.

The bombardment of an insulating thick (or semithick) specimen in vacuum generally gives rise to charge buildup, and the specimen may reach a positive potential of up to several tens of kV before breakdown and sparking. The high potential accelerates electrons to up to tens of keV, and as a consequence an intense bremsstrahlung background is produced in the PIXE spectrum. The peak-to-background ratios then decrease significantly [30].

Quantitation and Sensitivity

The basis for a quantitative analysis is that there is a relationship between the net area of an element's characteristic K or L X-ray line in the PIXE spectrum and the amount of element present in the sample. For proton bombardment and an infinitely thin specimen (by this is meant a specimen that is sufficiently thin that matrix effects become negligible), the relation is given by

$$Y_p(Z) = \frac{N_0 \sigma_{pZ}^X(E_0) \epsilon_p N C_Z \rho t}{A_Z \sin \theta} \quad (26.2)$$

where $Y_p(Z)$ is the number of counts in a characteristic X-ray line p of the element with atomic number Z , N_0 is Avogadro's number, $\sigma_{pZ}^X(E_0)$ is the X-ray production cross section for line p at the incident proton energy E_0 , ϵ_p is the absolute detection efficiency (including the solid angle) for X-ray line p , N is the number of incident protons, C_Z is the concentration of the particular element in the specimen, ρ is the specimen density, t is the specimen thickness, A_Z is the atomic mass of the element, and θ is the angle between the incident proton beam and the specimen surface. In formulating this equation, it is implicitly assumed that the specimen is uniform and that the beam size is smaller than the specimen area.

In practice, specimens are rarely thin enough for matrix effects to be entirely negligible. For specimens of intermediate thickness and for infinitely thick specimens (the latter are specimens that are thicker than the particle range), the quantitation equation has to include a correction for projectile slowing down and X-ray attenuation.

The minimum detection limits for PIXE depend both on matrix composition and on the element under investigation. By using both K and L X-rays, a good sensitivity across the periodic table is obtained. In routine analysis, the detection limits are of the order of 0.1–1 $\mu\text{g/g}$. With optimization of the experimental parameters, an improvement by a factor of 10 can be achieved.

26.3.2 Backscattering Spectrometry

The acronym RBS (Rutherford backscattering spectrometry) is often used for a variety of techniques based on detection of ions scattered by target nuclei [31]. The energy of the scattered projectile is measured and related to the mass of the scattering nucleus (see Fig. 26.4), as well as to its position in the matrix. The characteristics of RBS make it particularly useful for determination of depth distributions of heavier elements in light-element matrices and for investigation of thin multilayer films. For purely elastic collisions, the incoming ion will not have sufficient energy to penetrate the Coulomb barrier, and classical mechanics can be applied. In this Rutherford energy regime, the energy of the scattered ion and the probability for scattering are described by the Rutherford equations:

$$E = \epsilon_0 \frac{m_1^2}{(m_1 + m_2)^2} \left[\cos \theta + \left(\frac{m_2^2}{m_1^2} - \sin^2 \theta \right)^{0.5} \right]^2 \quad (26.3)$$

$$\frac{d\sigma}{d\Omega} = 1.296 \left(\frac{zZ}{E_0} \right)^2 \left[\sin^{-4} \left(\frac{\theta}{2} \right) - 2 \left(\frac{m_1}{m_2} \right) + \dots \right] \quad (26.4)$$

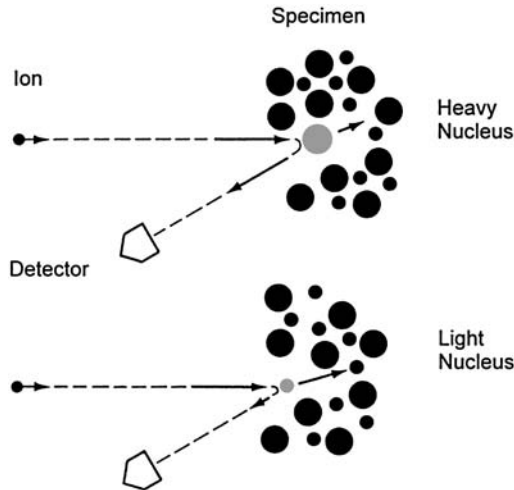


Fig. 26.4. The basic principles of ion backscattering spectrometry

Here m_1 and m_2 are the masses of the incident ion and target nucleus, E_0 is the energy of the ion scattered at a given depth, and z and Z are the charges of the ion and of the scattering nucleus, respectively. By measuring the energy of the scattered ions at a well-defined angle θ , the mass of the scattering nuclei can be determined. As is discussed in Chap. 24, the energy

loss of an ion traversing a relatively well-known matrix can be accurately calculated, as well as the energy loss of the scattered ion on its way through the matrix to the detector. Hence, this method can produce quantitative depth profiling of the isotope studied. Since a realistic analytical situation is much more complex than one single element in a fairly well-characterized matrix, practical analytical procedures require computer-based simulations based on any previous knowledge of the material studied.

In classical RBS, the ions are normally selected to be 2 MeV α -particles. This allows simultaneous PIXE analysis and fulfils the Rutherford requirements. To increase sensitivity and depth resolution, it is possible to use much heavier ions [32] and combine this approach with a time-of-flight detection system [33]. When heavier ions are used, beam-induced damage increases and has to be taken into account. When RBS is combined with PIXE for simultaneous X-ray spectrometry, the ions are normally protons of 2–5 MeV and then the classical Rutherford conditions are not fulfilled. Since protons partially penetrate the Coulomb barrier, they can produce resonant scattering and nuclear reactions that will generate much more complex scattering spectra. Although the interpretation of these spectra will hence be less straightforward, such resonances can also be used to enhance sensitivity for selected isotopes, and the quantitation and interpretation can be assisted by using suitable reference materials.

Experimental Details

The angle at which the scattered ions are detected will determine the characteristics of the RBS measurements. The mass resolution is highest at an angle of 180° relative to the incoming ion beam, and hence the detectors used to measure the ion energy are often placed at a large angle in the backward direction, for instance in the form of an annular detector facing the specimen with the incident ion beam going through a central hole. The resolution is limited by several factors: energy straggling of the scattered ion in the matrix and the intrinsic energy resolution of the detector used. Detecting at approximately 90° increases the scattering cross section, energy straggling and depth resolution, while the mass resolution will be poor. Normally a semiconductor detector is used, but for very high demands on resolution either a time-of-flight (ToF) system or a magnetic spectrometer [34, 35] can be used.

The applications of RBS that dominate by far are within materials science and to study thin layers and depth distributions. Normally good previous information is available on each specimen, which makes it suitable to simulate expected RBS spectra and to compare the simulations with experimental results. Several computer codes are available [3, 36]. For absolute calibration various reference materials are commercially available. In addition, materials science studies often require that experiments are performed in ultrahigh

vacuum and that the specimens can be manipulated during analysis, for instance by heating, cooling or sputtering. Furthermore, since analysis is often performed on single crystals, the specimen needs to be mounted on a multi-axis goniometer. As is discussed in Chap. 24, the ion penetration depth will increase when an incoming ion channels along a crystal axis. Hence, the probability of close encounters such as ion/nucleus scattering is also reduced, and the RBS yield from lattice atoms, as well as substitutional atoms, will be much reduced (by approximately an order of magnitude) along the crystal axis. The yields from other types of close-encounter interaction, for instance PIXE, will also be reduced. By varying the orientation of a crystal mounted on a goniometer and registering RBS signals, the crystalline structure, including potential dislocations and possible interstitial atoms, can be investigated [37]. The quality of crystal lattices, epitaxial layers and lattice strain can be investigated by such ion channeling.

Another type of MeV ion-scattering spectrometry, particle elastic-scattering analysis (PESA), is best applied to thin targets, i.e. targets fully penetrated by the incoming ion. By registering ions scattered into forward angles a high flux (high scattering cross section) is available, but since the mass separation in forward angles is low, this technique only lends itself to measuring low- Z elements. Since several low- Z elements are of interest in many analytical applications and these elements are difficult to measure by other means, the main use of PESA has been as a complementary technique to, for instance, PIXE.

26.3.3 Elastic-Recoil Detection Analysis (ERDA)

Instead of measuring the scattered ions, ERDA [39] exploits the target atom recoil in the forward direction in a glancing-angle geometry. For thin specimens, transmission ERDA is also possible [40]. Recoiling atoms lighter than the incoming ion will receive a higher energy than the corresponding scattered ions, which will hence be easy to eliminate with stopping foils. Hence, ERDA is ideal for profiling light isotopes in a matrix with an abundance of heavier elements. Owing to low cross sections for forward recoiling, the analytical sensitivity is not very high, but the depth distribution is of the order of a few nm. A typical analytical situation, shown in Fig. 26.5, would be to use an alpha particle beam of a few MeV to investigate a hydrogen diffusion profile.

One limitation in the traditional experimental ERDA setups is the energy resolution of the semiconductor detector. ToF systems in conjunction with recoil spectrometry comprise an energy detector and two thin foils used as start and stop detectors, making use of secondary electrons produced in each foil. The measured flight times are then plotted vs. recoil energy. This type of recoil spectrometry can also be applied with high-energy nuclear microprobes although the ion currents are significantly lower than in normal IBA facilities.

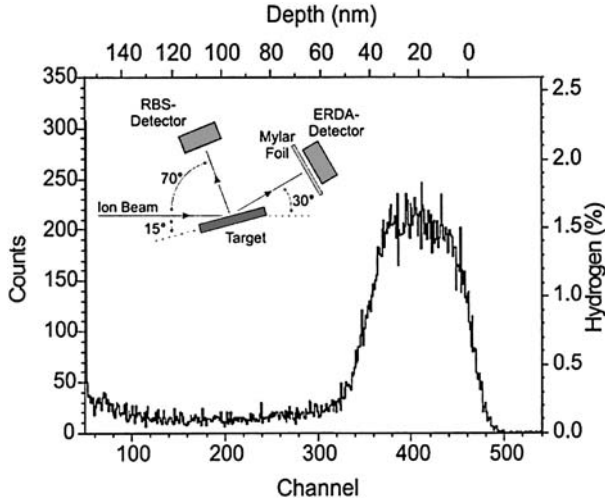


Fig. 26.5. Typical ERDA setup with an additional RBS detector, and an ERDA spectrum showing a hydrogen depth profile based on an alpha particle beam (Reprinted from [38], copyright 2002, with permission from Elsevier)

This requires a detector system with a large solid angle to compensate for the low cross sections for recoils in the forward direction [41].

For experimental facilities with high-energy accelerators and heavier ions, other recoil techniques can be used. Using high-energy-resolution detection, it is then possible to simultaneously profile light to medium-heavy elements in matrices of lighter and/or heavier elements [42].

26.3.4 Nuclear-Reaction Analysis (NRA)

The cross section for a nuclear reaction depends in a complicated manner on the incident ion energy and the internal properties of the particles involved. The Coulomb barrier is approximately

$$E = \frac{zZ}{(a^{1/3} + A^{1/3})} \text{ MeV} \quad (26.5)$$

for an incident ion of atomic number z and mass a and a target nucleus of mass A and atomic number Z . The projectile will normally not interact strongly in a nuclear reaction unless the incident energy is greater than the Coulomb barrier. Clearly, for moderate ion energies, nuclear reactions will only occur with the lightest target element. Hence, a light ion of MeV energy can overcome the nuclear charge of low- Z matrix elements, and inelastic reactions can occur. The existence of particular nuclear reactions with relatively high cross sections represents a potential for accurate trace element determination of specific light isotopes in high- Z matrices [43].

The nuclear reaction occurring is either a direct reaction or a compound-nucleus reaction. In the latter case, an intermediate-state excited nucleus forms that will eventually decay by gamma rays or charged-particle emission. By detecting these radiation quanta, analysis of light elements can be performed. A common example of this type of NRA is the use of $^{19}\text{F}(\text{p}, \alpha\gamma)^{16}\text{O}$ for depth profiling, thereby making use of sharp resonances in the cross section as a function of proton energy [44]. In contrast to compound formation, the direct-mechanism nuclear reactions normally demonstrate a cross section smoothly varying with energy, for instance in the case of $^{13}\text{C}(\text{d}, \text{p})^{14}\text{C}$ [45]. There are many combinations of ions and elements that facilitate analysis by specific nuclear reactions. The possibility of measuring lighter elements, such as boron, beryllium and nitrogen, is a major advantage in combination with techniques such as RBS and PIXE, which are more suited for measuring medium-heavy and heavy elements.

The special case of particle-induced prompt gamma rays is often denoted by the term particle-induced gamma-ray emission (PIGE) [46]. It has been used extensively in simultaneous combination with PIXE and is normally defined as detection of prompt gamma rays produced during MeV ion bombardment of a specimen homogeneous in depth, i.e. resonant reactions are not used for depth profiling. Because of the interaction occurring with the nucleus, the cross sections are lower than for PIXE, and PIGE typically represents a less sensitive analytical technique than PIXE. However, the gamma-ray peaks are generally well isolated and the energy of the emitted gamma-ray photons is high enough that correction for attenuation will not be necessary. PIGE may hence in some situations be favorable compared with PIXE for elements such as calcium and potassium [47].

The high penetration of gamma rays also simplifies the experimental arrangements, and the nuclear interaction facilitates the obtaining of isotopic information. External beams are a standard method in the case of PIGE [48] and offer the same advantages as discussed above for the case of PIXE.

In Tables 26.1 and 26.2, some examples of the detection limits of PIGE are given. They are given only as indications of attainable sensitivities, since the specific experimental conditions will have a strong impact. The values in these tables apply when the experimental arrangements and bombarding energy are optimized for the detection of a specific element.

Table 26.1. Detection limits (atomic fraction) for PIGE using protons ($E_p < 9\text{ MeV}$) [49]

>1%	0.1–1%	<0.1%
Pd, Sm, Gd, Hf, W, Au, Pb	S, K, Sc, Ti, Co, Cu, Ge, Y, Zr, Mo, Ru, Ag, Sn, I, Ta, Pt	Li, Be, B, C, N, Na, Mg Al, Si, Cl, Ca, V, Mn, Fe, Ni, Zn, Nb, Cd, In, Sb

Table 26.2. Detection limits (atomic fraction) for PIGE using α -particles (5 MeV) [50]

>1%	0.1–1%	<0.1%
Sc, Cr, Cu, Zn, Rb Zr, Er, Hf, Hg	O, Mg, Si, Cl, K, Ti Fe, Br, Mo, Ru, Pd Ag, Cd, Ta, W, Re Ir, Pt, Au	Li, B, N, F, Na, Al, P V, Mn, Rh

26.3.5 Scanning Transmission Microscopy (STIM)

When using a nuclear microprobe to investigate nonhomogeneous, thin (ions will pass through) specimens, it is very useful to measure the actual “mass” thickness (mg/cm^2) in each analyzed “pixel”. The well-defined energy loss of ions facilitates calculation of this thickness from the energy loss through the specimen. One can use direct or indirect determination of the energy loss for each individual ion when it passes through a thin specimen [51]. The drawback of STIM in direct mode is that it is a stand-alone technique and cannot be combined with, for example, simultaneous PIXE or RBS measurements, because the ion beam intensity has to be extremely low (1000–10 000 ions/s) not to saturate the detector. With a traditional off-axis geometry, STIM can be performed in parallel to both PIXE and RBS because the beam current is typically 50 pA or more. A drawback with off-axis geometry is that it has an inherent spectral degradation. An alternative geometry, called on-axis/off-axis STIM, combines the advantages of both on-axis and off-axis STIM. This geometry is sensitive only to pure energy loss in the sample. No ions scattered in the sample will influence the spectral shape, because those ion paths are collimated off. Instead, a part of the transmitted beam is scattered into the detector by a thin monoelemental foil situated behind the sample. By varying the scattering angle from the foil, the count rate of the STIM detector can be matched to the actual beam current used, and thus STIM measurements can be done in parallel with, for example, PIXE measurements. STIM offers structural imaging of specimens but also facilitates normalization of the results of other analytical methods to the mass in each analyzed pixel.

Sub-cellular mass determination in organic material or tissues using STIM is crucial for accurate normalization of concentrations by PIXE. A good illustration of what can be obtained by STIM is an investigation where 2 MeV α -particles were used to map an intracellular mass distribution [52]. Freeze-dried cells were investigated and compared with results from cellular sections in epoxy resin, and this facilitated accurate quantitation of the intracellular elemental concentrations. Another example of STIM images from freeze-dried tissue is shown in Fig. 26.6, clearly demonstrating the cellular structure.

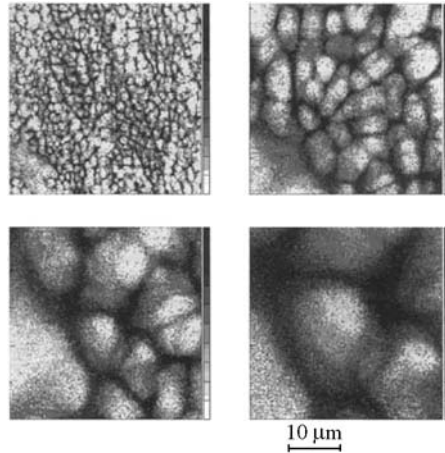


Fig. 26.6. STIM images from a section of freeze-dried brain tissue at different magnifications. The individual cells are easily observed (Courtesy of P. Kristiansson)

Although the focused microbeam does not yield an ideally parallel beam for channeling, the STIM technique can be applied also in channeling mode for crystalline structures. Along the crystal axes, the ion penetration depth will be much increased and less energy will be lost. By rastering the microbeam over the specimen, it is possible to check the lateral variation in the ion channeling intensity due to the existence of regions with imperfect crystalline structure. This method is called channeling contrast microscopy [53] and is useful in materials science applications.

In another development, STIM can be used for tomographic reconstruction [54] of microscopic objects. By rotating and translating a specimen through a narrow ion beam and measuring the energy loss, a computer code can be used for subsequently reconstructing the inner structure of the specimen. When STIM is combined with X-ray detection, PIXE can facilitate reconstruction of 3-dimensional elemental distributions [55]. In Fig. 26.7, ion-tomographic reconstructions of the mass density and manganese content in a scorpion stinger are shown across the stinger. The stinger is very brittle and would break during sectioning, making a noninvasive technique necessary.

26.3.6 Ionoluminescence (IL)

The same mechanisms that create characteristic X-rays, as discussed in Sect. 26.3.1, also produce intense visible or UV light from the outer shells during ion bombardment. In transparent matrices this can be utilized for imaging and analytical purposes, similarly to cathodoluminescence as used in an electron probe [56]. The physical processes creating luminescence are similar for electrons and ions, but since the electron and ion beams have quite different

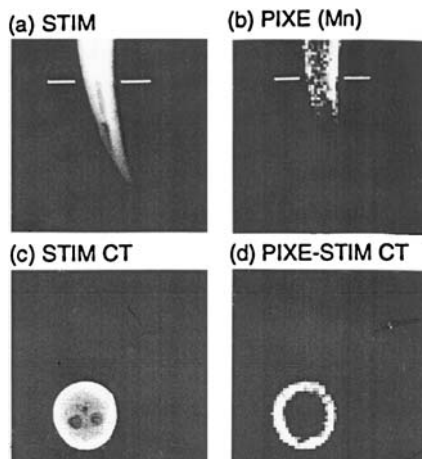


Fig. 26.7. Characterization of a scorpion stinger by a nuclear microprobe: (a) STIM image, (b) PIXE elemental map of manganese, (c) computer-reconstructed mass distribution at the level in the stinger given by the white lines in (a) and (b), (d) computer-reconstructed manganese distribution at the same level. The analysis clearly shows high levels of manganese at the stinger surface as a result of the biomineralization process (Courtesy of R.M.S. Schofield)

matrix penetration characteristics, they are complementary to each other. Luminescence phenomena are common in many solids, yet the use of the light produced is restricted because of the complex physical mechanisms behind the emission. For insulators, to which many minerals belong, crystal field theory can be used to explain luminescence, whereas for semiconductor crystals, band theory is often employed. Luminescence can be divided into two subgroups, intrinsic and extrinsic, according to its origin. In a material with a band gap between the conduction band and valence band, the extrinsic luminescence depends on donor and acceptor levels that are related to impurity ions. The intrinsic luminescence usually contributes to emission through crystal structural defects and hence to crystal lattice properties. Ionoluminescence [57] depends on the characteristics of the specific luminescence centers in the material.

Although IL can yield information on both crystal structure and the chemical state of the atoms present, it is in practice so far limited to semiquantitative analysis. In geochemistry, rare-earth elements (REEs) are often very important and, owing to specific characteristics of the electron configuration for all trivalent and several divalent REE ions, IL represents a particularly good complement to PIXE, which has a particularly low sensitivity for these elements [58]. The narrow luminescence emission bands of the REE ions are caused by electrons in the 4f subshell, which is partially shielded by ($5s^25p^6$) electrons. Since the interaction between the 4f electrons and the crystal field

is weaker than the spin-orbit interaction and the energy levels of REE ions of this configuration are not significantly influenced by crystal fields, the structure of the energy levels in the free REE ions is basically the same in different host matrices.

Experimental Considerations

A nuclear microprobe is normally operated in combination with IL in two ways, spot or scanning mode. Furthermore, the light produced can be analyzed in a spectrometer system with a grating monochromator in combination with a photomultiplier tube (PMT) or directly into the PMT (panchromatic mode). In Fig. 26.8, three elemental maps from PIXE are compared with a photographic image and a panchromatic IL map. Another possibility is to scan the specimen and record the image produced in a very sensitive CCD camera. Using a grating monochromator, high-resolution IL spectra can be acquired in spot mode and be combined with simultaneous PIXE analysis for trace element determination.

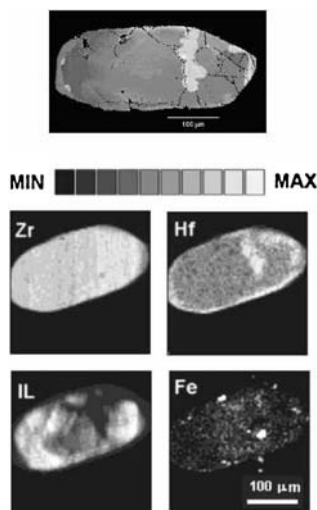


Fig. 26.8. Panchromatic IL image and three PIXE elemental maps from a natural zircon

High-resolution spectra normally require long times of irradiation (minutes), which also means high radiation doses and thus a change in the IL emission during irradiation. By employing a photodiode-array detector with many elements, a few tens of μm each, a rapid high-resolution system can be obtained. It allows the simultaneous recording of all wavelengths, significantly improves the analytical capacity and reduces beam damage effects [59].

26.4 Applications

In the following, a selection of applications is presented in order to illustrate the various characteristics of IBA methods. The selection is also intended to demonstrate typical fields where IBA has had a major impact, although many times in conjunction with other analytical techniques. In fact, the most successful studies involving IBA methods also make use of an arsenal of analytical tools to complement the special characteristics of IBA methods.

In order to illustrate the present status of IBA applications, the latest International Conference on Ion Beam Analysis (IBA 2003, Albuquerque [60]) may offer a deeper insight. As in earlier conferences in this biannual conference series, the contributions are still dominated by traditionally common applications such as thin-film analysis by RBS and ERDA, trace element analysis by PIXE, and light-element analysis by NRA. Altogether these applications represent approximately 2/3 of the contributions submitted and are mainly related to materials analysis. However, novel methods and new developments, for example, nanoscience [61,62], within existing fields are becoming of growing importance. In other conference series, many other fields of application, such as the environment biomedicine and archaeology, are addressed.

26.4.1 Materials Science and Industrial Applications

In a broader sense this field covers such varying areas as electronic devices, metallurgy, geology, the arts and archaeology. However, normally one would limit this field to refer only to the study of technological developments of modern materials.

In the early days of IBA, RBS soon became one of the standard techniques for surface characterization of semiconductors and was also adopted in industrial processing environments. Since the detailed crystalline structure and possible defects were often of great importance, RBS was also combined with crystal channeling phenomena. With the subsequent development of the various analytical techniques in IBA, these techniques were often integrated in various simultaneous combinations (e.g. RBS, PIXE and NRA) for materials characterization. One very good example of this is the characterization of magnetic thin films using a spectrum of various IBA methods. Both α -particles and deuterons are used as bombarding ions, and by combining several methods detailed information has been yielded [63]. In the semiconductor industry today, in-house accelerators are used in R&D and for process quality control. In order to extend the industrial IBA arsenal, NRA based on deuterons is used as a complement to RBS for lighter-element characterization [64].

With the growth during the last decade of complex materials in parallel with a very strong trend toward nanoscience and nanotechnology, IBA techniques have developed accordingly. One such example is an investigation of

quantum well structures in SiGe by using high-resolution channeling contrast microscopy [65]. Using RBS in channeling mode and the high lateral resolution of the nuclear microprobe, it was possible to identify microscopic lateral inhomogeneities associated with a slight lattice tilt present in the specimens studied. Another example of the use of this technique is an investigation of ion-etched silicon microstructures [66]. In Fig. 26.9, images from scanning-electron secondary electrons and channeling-contrast microscopy can be compared.

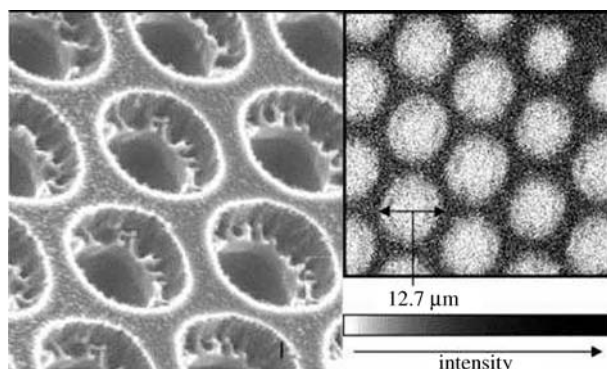


Fig. 26.9. Silicon microstructure imaged by electron secondary electrons and channeling-contrast microscopy (Reprinted from [66], copyright 2002, with permission from Elsevier)

Depth profiling of hydrogen is a common and technologically very important application of IBA. Several methods can be applied, for instance ERDA with He ions and NRA using ^{15}N ions. Samlenski et al. [38] used 2.8 MeV He^+ ions for high-lateral-resolution ERDA measurements on CVD-grown diamond and could relate the hydrogen maxima found to to observed optical structures (see Fig. 26.5). They also complemented these measurements with ^{15}N profiling. This is one example of a very common analytical problem in surface and thin-films analysis where IBA methods have made significant contributions owing to their high depth resolution characteristics. Recently, the theoretical and experimental limitations on depth profiling due to multiple scattering phenomena have been investigated in detail [67].

Another important materials analysis aspect is the use of IBA to monitor the results of another means of using ions in solids, ion implantation. After controlled ion-beam-assisted materials modification, which is regularly used in the production of semiconductor devices (Chap. 25), it is very suitable to use different IBA methods to control the effects of the irradiation. In fact, it is possible to perform this investigation on-line, i.e. to measure during ion beam processing. A Japanese group did this when implanting deuterons by monitoring the yield from the $^2\text{H}(\text{d}, \text{p})^3\text{H}$ reaction [68]. The emitted protons

were counted as a measure of the deuteron fusion intensity. The group were able to show large differences in proton emission intensity between single-crystal and polycrystal areas in tantalum but not in copper matrices.

The industrial applications of IBA were initially limited to materials characterization using primarily RBS, but recently new approaches have been investigated. Wood-based industries are a very economically important industrial field in both Sweden and Finland, which probably has initiated a rather extensive collaboration with this normally rather “low-tech” base industry. In Sweden, in a direct collaboration with the paper industry, PIXE and STIM in a nuclear microprobe were used for studying the details of printing on newsprint paper [69]. The STIM measurements assisted in showing the mass distribution of the fibrous structure, as demonstrated in Fig. 26.10. By combining STIM with PIXE measurements, it is possible to investigate how the ink from printing is distributed in the fibrous structure and to understand how the smearing of the ink on newsprint paper can be reduced [70].

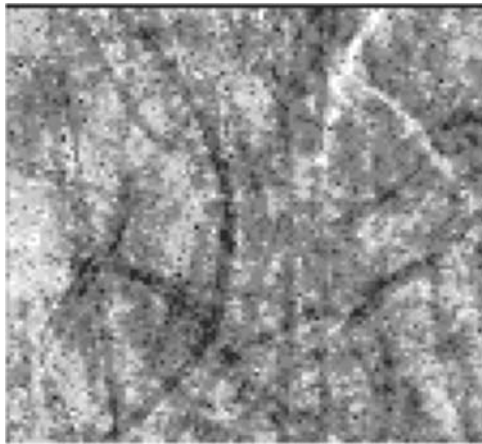


Fig. 26.10. STIM image of newsprint paper revealing the fibrous structure (courtesy of P. Kristiansson)

In a Finnish study, wood material delivered to pulp industries was pre-analyzed by PIXE in order to know the trace element concentrations to minimize discharge of polluting effluents [71] from the processing of pulp.

Another example of innovative “industrial” use of IBA techniques is the study of “hot” radioactive material, as carried out at a dedicated beam line at the Pierre Suë laboratory, Saclay, France [72]. By directing the ion beam into a hot-cell beam line, nuclear-microprobe experiments can be carried out by remote control. IBA methods have also been used to monitor actinide concentrations in drainage from a radioisotope laboratory [73]. By the combined use of a preconcentration technique and PIXE analysis, thorium and

uranium could be analyzed quantitatively down to 5 ppb in the drainage. Another interesting example in the same field is the investigation of actinide transport in granite microstructures in bentonite clay a material favored by many experts as an ideal barrier when storing radioactive waste [74]. The combined use of PIXE and RBS was used to investigate diffusion processes.

26.4.2 Geosciences

The geosciences have traditionally been a very strong field of IBA application, since the beam-induced radiation damage in mineral samples is normally negligible. There are several IBA laboratories that have had a strong focus on geosciences or even have been entirely dedicated to the field and worked both in basic research and in applied ore prospecting. The nuclear microprobe, with its inherent high lateral resolution, is particularly suitable for investigating the microscopic structure of minerals. The CSIRO laboratory in Sydney* [75] has been commendable for a very user-oriented activity that approaches both basic problems and industrial partners. It has also recently upgraded its nuclear microprobe to be even better prepared to approach specifically mineralogical problems [76].

An example of a typical application is a study of trace elements in sulfides to provide insights into various geological processes such as ore formation and magmatic-system evolution [77]. By combining traditional electron microprobe and ion-sputtering mass spectrometry probes with PIXE analysis in a nuclear microprobe, several trace constituents of the sulfides could be accurately quantified and used for understanding, for instance, ore formation processes. In another study of sulfides [78], PIXE results were compared with laser ablation microprobe-ICP measurements and could confirm that the laser ablation analysis gave reasonably accurate quantitative results.

Owing to a lack of suitable methods, it is very difficult to localize and quantitatively analyze boron in minerals. By developing an NRA method based on the reaction $^{11}\text{B}(\text{p}, \alpha)2\alpha$, which has an especially high cross section for 660 keV protons, it was possible to analyze B at low concentrations in calibrated single crystals [79]. This method has later been applied successfully to real geological specimens. Also, PIGE can be successfully applied in geological applications. The combined use of PIXE and PIGE in investigations of fluid inclusions in minerals is a powerful technique, and PIGE contributes with concentrations of light elements, such as Li, Be, B and Na [80]. Using this analytical combination, many disadvantages of the traditionally used destructive methods, for instance laser ablation-ICP, can be avoided, and a good analytical accuracy is obtained.

Also, applications that do not make use of a microbeam but rely only on the high sensitivity of the PIXE method can contribute significantly in prospecting for ores. The geogas technique [81] is based on the principle of

* This lab has just recently (2004) been relocated to Melbourne univ.

tracing the weak signatures of deep mineralizations on the surface of the ground. In geogas prospecting, extremely small concentrations of trace elements, in an upward gas stream collected on thin membranes, are measured using the PIXE technique. The elemental composition of the geogas across geological formations reveals much about the bedrock composition and the possible presence of concealed mineralizations and ore bodies.

26.4.3 Biology and Medicine

Although organic materials are rather sensitive to ion-beam-induced damage, ion beam techniques, in particular PIXE, have had and still have a significant impact on biology and medicine for localized trace element analysis. With the development of the nuclear microprobe, bioscience obtained a versatile tool to add to the existing methods of microanalysis. It is being used for a wide spectrum of biomedical applications, varying from investigations of human kidney stones [82], through the elemental composition of beetles [83], to single marine toxin-producing phytoplanktons [84].

A recent example of how the technique can be applied at single-cell level is a study of clinical oncology using gallium nitrate for antitumor action. Since the reasons for this action are not fully known, ovarian cancer cells were exposed to gallium nitrate and analyzed by a nuclear microprobe [85]. In most cells the gallium was distributed homogeneously in the cells, but in some cells it was concentrated in a well-localized, perinuclear region. This finding contributed considerably to new information on how gallium attacks tumor cells.

Another anticancer treatment relies on specific localization of a boron compound, BPA, in cancer cells, and subsequent neutron exposure and boron neutron capture therapy. The tumor-to-tissue ratio of the boron concentration, as well as the location of boron within the cells – the α -particles produced in the $^{10}\text{B}(\text{n}, \alpha)^7\text{Li}$ reaction only have a range of about $10\text{ }\mu\text{m}$ (a cell diameter) – is critical for the efficiency of this therapy. Using the same reaction for boron detection as described in Sect. 26.4.2, a nuclear microprobe was used to investigate the intracellular boron concentration and distribution [86]. Alpha particles were detected in the forward and backward direction simultaneously. Only the coincidences between the two directions were then considered to be true boron events, resulting in excellent background suppression. The boron abundance was found to be $150 \pm 20\text{ ng/cm}^2$ in normal tissue and $570 \pm 70\text{ ng/cm}^2$ in tumor tissue. However, if the rats were fed with L-dopa prior to the injection of the boron-carrying drug BPA, the uptake in cancer tissue increased 3 to 4 times. Furthermore, the boron seemed to be homogeneously distributed in the cellular structure and no specific accumulation close to the nucleus was shown (see Fig. 26.11, where boron events are marked as black dots and overlaid on the STIM image).

Nuclear-microprobe analysis of plants has had significant implications in plant science in general, but in particular in the understanding of species

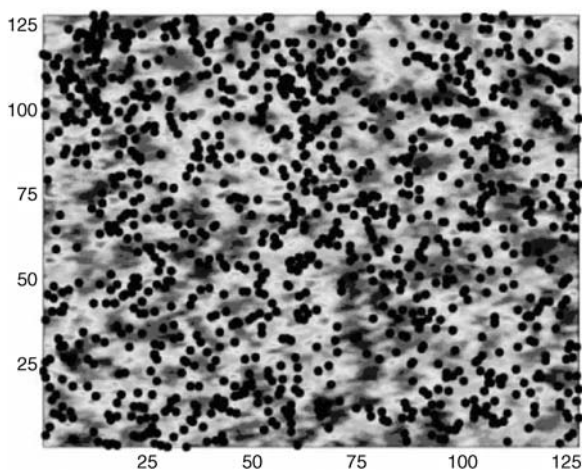


Fig. 26.11. Boron distribution (*black dots*) as measured by a nuclear microprobe in a tumor tissue from a rat brain, superimposed on a STIM map [86]. The rat was infused with the boron-carrying drug BPA after being pretreated with L-dopa

demonstrating hyperaccumulation of metals [87]. The importance of adequate sample preparation was shown in a study of a Ni-hyperaccumulating plant. It was demonstrated that new and more sophisticated techniques of freeze-substitution techniques are required for accurate ultrastructure measurements with a nuclear microprobe [88].

The nuclear microprobe has been very useful in the field of limnology and marine sciences for the investigation of mineralized tissue such as shells or bones. The fixation of various elements during growth generates a time chart of elements that may be utilized for the understanding of changes in the environment surrounding the organism in question. One such specific structure that is being used extensively is the otolith, a biogenic carbonate concretion which forms part of the hearing/balance system in fishes. The radial growth of the otolith and the variation of trace elements along the radius appear to capture important aspects of fishes' environmental history [89,90]. Owing to regional and global overfishing, the migration patterns of fish stock are today of highest interest biologically, ecologically and economically. Hence otoliths are being analyzed in several IBA laboratories, including that at Lund, where, in particular, movement of diadromous species (e.g. eel and brown trout) between fresh and brackish water has been determined by analyzing the Sr/Ca ratio [91]. This technique has also been used to identify fish that were raised in freshwater hatcheries and then released into brackish water.

26.4.4 Environmental Studies

Over recent decades there has been extensive use of IBA methods in environmental studies [92,93]. One reason for this success is probably the very early (1969) use of PIXE for atmospheric samples. Over the three decades following, IBA techniques in combination with many others have had a very important impact on studies of atmospheric aerosols [94], for example in climate change [95] and in urban air pollution [96], and of their regional and industrial impact [97]. The capacity for analyzing many samples for multiple elements in a short time has made PIXE analysis a very important analytical method in these investigations. However, by using microanalytical methods in a nuclear microprobe combined with complementary methods for studying individual aerosol particles, new information can be obtained on the origin of aerosols [98]. In one study, microprobe data were compared with bulk analysis results for large collections of particles [99]. It was estimated that micro-PIXE of aerosols provides data which are accurate to within 10–20%. Possible losses of the elements detected by PIXE in the course of the microprobe bombardment were examined, but were generally found to be insignificant.

26.4.5 Art and Archaeology

One of the best methods for trace element determination in the most sensitive kinds of art and archaeology objects is the use of synchrotron radiation microprobes [100,101], in which photons produce high characteristic X-ray yields without any significant sample degradation. Nevertheless, the available beam time of such microprobes is very limited, and ion beam analysis, with its complementary capabilities, can be very competitive also for sensitive specimens. The characteristics of ion–matter interaction, however, make it essential to select adequate specimens for a successful and nondestructive investigation of an artefact. On one hand, there are significant changes taking place in an irradiated specimen due to ionization and heating; on the other hand, the external-beam approach can limit these changes to an absolute minimum [102] (see also [6]). This is illustrated by the frequent use of IBA analysis for very sensitive ancient documents. By careful monitoring of the ion beam intensity, it is possible to analyze these precious documents without any visible remaining damage.

The same methods that have been developed within more general materials science are also applicable in this field, as illustrated by [103], where the problem of rapid deterioration due to weathering of stained glass windows in European cathedrals was addressed. The hypothesis was that the rapid deterioration is a result of the airborne pollutants of today. Hence, soda–lime glass samples were exposed to atmospheres in which the temperature, relative humidity and pollutant (SO_2 and NO_2) concentration were systematically varied. After exposure, the surfaces of these glass samples were analyzed using NRA (to profile hydrogen) and RBS.

External-beam PIXE and PIGE have been combined in a study of paint structures in a series of layered paint samples prepared according to known late-18th-century techniques [104]. In a recent study, a group at the University of Florence tried to identify pigments in various paint layers from a painting by Leonardo da Vinci (the “Madonna dei fusi”). By a combination of PIXE and PIGE using external-beam irradiation [105] and corresponding analysis of laboratory-made specimens, it was possible to nondestructively investigate specifically the blue pigments. Since such measurements on very precious artefacts have to be genuinely nondestructive, it is crucial to understand how to avoid formation of “brown spots” on the irradiated surface. A group at the Autonomous University of Madrid carried out such a study on common pigments [106] from natural carbonates. They irradiated the various pigments and used a combination of PIXE and IL to follow the degradation of the material, and then combined it with optical absorption and thermoluminescence to understand the formation of the remaining defects induced. Also, PIXE and RBS have been combined to investigate experimental glazes simulating real painted glazes used as translucent layers on old oil paintings [107].

One serious issue in using IBA methods for the analysis of ancient inhomogeneous artefacts is the superficial preference (the first few to tens of micrometers) of the analytical results. The question is: In a corroded object, how representative are IBA results of the bulk? In order to get around this problem, small spots of objects may be polished mechanically, but then at the risk of changing the exterior and lowering the value of the artefact. As a remedy, a group with long experience of using IBA methods on metallic objects has tried to use controlled laser ablation in conjunction with an external-beam microprobe in a study of thick corrosion layers on archaeological metals [108].

The corrosion process of some materials acts as a protection against further deterioration, and by combining a nuclear microprobe (using PIXE) and XRD (micro-X-ray diffraction) it was possible to investigate the structural composition of this layer in 1500-year-old iron [109]. Phosphate phases in the iron were determined by XRD, and they interfered with oxygen phases in the corrosion process. Although a metallic alloy, for instance bronze, may be essentially free of corrosion, analytical results on trace constituents are not easily obtained, owing to the large presence of heavier elements. Normally PIXE analysis would reveal $\mu\text{g/g}$ concentrations of elements that might be used for provenance studies, but since the detectors are saturated by the signal from major matrix elements, practical detection limits will be much higher. This problem may be circumvented or at least reduced by the use of critical absorbers that reduce the major X-ray lines while preserving the interesting lines [110]. In some metallic specimens, the degree of representation of the bulk and corrosion is less critical but there may instead be a need for combining several methods.

26.5 Conclusions

Being within the more successful fields of application of particle accelerators, ion beam analysis has, together with accelerator-based mass spectrometry (Chap. 23), justified the further development of, in particular, small and medium-size accelerators. As shown in this chapter a great variety of ion beam analysis techniques exists and contributes significantly in many fields of science and technology. There are several other types of IBA applications of great relevance in addition to those presented here. Nevertheless, the applications presented here represent a reasonable display of what can be done with IBA analysis, in particular when making use of high analytical resolution both laterally and in depth. Owing to increasing competition in the analytical arena, it is essential for the future of these techniques that the applications will be even more carefully selected to make full use of the inherent properties of IBA and to combine the techniques with complementary techniques in the most productive manner. The technical developments in accelerators, focusing systems and detector acquisition systems will further ensure and even enhance the potential and success of ion beam analysis techniques.

References

1. N. Nastasi, J.W. Mayer, J.K. Hirvonen: *Ion Solid Interactions: Fundamentals and Applications*, (Cambridge University Press, London 1996)
2. J.R. Tesmer, M. Nastasi, J.C. Barbour, C.J. Maggiore, J.W. Mayer: *Handbook of Modern Ion Beam Materials Analysis* (MRS, Pittsburgh, 1995)
3. J.F. Ziegler: SRIM-2003, available from <http://www.srim.org>
4. J.A. Maxwell, J.L. Campbell, W.J. Teesdale: Nucl. Instr. Meth. B **43**, 218 (1989)
5. C. Jeynes, N.P. Barradas, P.K. Marriot, G. Boudreault, M. Jenkin, E. Wendler, R.P. Webb: J. Phys D: Appl. Phys. **36**, R97 (2003)
6. P.A. Mandó: Nucl. Instr. Meth. B **85**, 815 (1994)
7. O. Enguita, M.T. Fernández-Jiménez, G. Garcia, A. Climent-Font, T. Calderón, G.W. Grime: Nucl. Instr. Meth. B **219–220**, 384 (2004)
8. R.A. Colman, G.J.F. Legge: Nucl. Instr. Meth. B **73**, 561 (1993)
9. M.B.H. Breese, D.N. Jamieson, P.J.C. King: Microprobe ion optics. In: *Materials Analysis Using a Nuclear Microprobe* (Wiley, New York 1996) pp. 81–139
10. F. Watt, G.W. Grime: *Principles and Applications of High Energy Ion Microprobes*, ed. by M.B.H. Breese, D.N. Jamieson, P.J.C. King (Hilger, Bristol 1987)
11. R. Mimura, H. Takayama, M. Takai: Nucl. Instr. Meth. B **210**, 104 (2003)
12. F. Watt, J.A. van Kan, I. Rajta, A.A. Bettiol, T.F. Choo, M.B.H. Breese, T. Osipowicz: Nucl. Instr. Meth. B **210**, 14 (2003)
13. P. Kristiansson: private communication
14. D. Spemann, T. Reinert, J. Vogt, D. Dobrev, T. Butz: Nucl. Instr. Meth. B **190**, 313 (2002)

15. A. Shariff, P. Kristiansson, V. Auzelyte, M. Elfman, K.G. Malmqvist, C. Nilsson, J. Pallon, M. Wegdén: Nucl. Instr. Meth. B **219–220**, 494 (2004)
16. S.A.E. Johansson, J. Campbell, K.G. Malmqvist: *Particle-Induced X-ray Spectrometry (PIXE)* (Wiley, New York 1995)
17. W. Brandt, G. Lapicki: Phys. Rev. A **20**, 465 (1979)
18. W. Brandt, G. Lapicki: Phys. Rev. A **23**, 1717 (1981)
19. S.I. Salem, S.L. Panossian, R.A. Krause: At. Data Nucl. Data Tables **14**, 91 (1974)
20. J.H. Scofield: Phys. Rev. A **9**, 1041 (1974)
21. J.L. Campbell: Nucl. Instr. Meth. B **31**, 518 (1988)
22. J.H. Scofield: Phys. Rev. A **10**, 1507 (1974)
23. M.H. Chen, B. Crasemann, H. Mark: Phys. Rev. A **24**, 177 (1981)
24. M.H. Chen, B. Crasemann: At. Data Nucl. Data Tables **41**, 257 (1989)
25. H. Paul: Nucl. Instr. Meth. B **3**, 5 (1984); see also erratum in Nucl. Instr. Meth. B **5**, 554 (1984)
26. Y. Miyagawa, S. Nakamura, S. Miyagawa: Nucl. Instr. Meth. B **30**, 115 (1988)
27. F. Folkmann, J. Borggreen, A. Kjeldgaard: Nucl. Instr. Meth. **119**, 117 (1974)
28. K. Ishii, S. Morita: Int. J. PIXE **1**, 1 (1990)
29. L.E. Carlsson, K.R. Akselsson: Nucl. Instr. Meth. **181**, 531 (1981)
30. M. Ahlberg, G. Johansson, K. Malmqvist: Nucl. Instr. Meth. **113**, 377 (1975)
31. W.-K. Chu, J.W. Mayer, M.-A. Nicolet: *Backscattering Spectrometry* (Academic Press, New York 1978)
32. M.R. Weller, M.H. Mendenhall, P.C. Haubert, M. Döbeli, T.A. Tombrello: Heavy ion Rutherford backscattering. In: *High Energy and Heavy Ion Beams in Materials Analysis*, ed. by J.R. Tesmer, C.J. Maggiore, M. Nastasi, J.C. Barbour (MRS, Pittsburgh 1990) pp. 139–151
33. J.A. Knapp, J.C. Banks, B.L. Doyle: Nucl. Instr. Meth. B **85**, 161 (1994)
34. K. Kimura, Y. Oota, K. Nakajima, M. Suzuki: AIP Conf. Proc. **689**, 373 (2002)
35. W.A. Lanford, S. Bedell, S. Amadon, A. Haberl, W., Skala, B. Hjorvarsson: Nucl. Instr. Meth. B **161–163**, 202 (2000)
36. L.R. Doolittle: Nucl. Instr. Meth. B **15**, 227 (1986)
37. A. Seppälä, R. Salonen, J. Slotte, T. Ahlgren, E. Rauhala, J. Räisänen: Nucl. Instr. Meth. B **161–163**, 520 (2000)
38. R. Samlenski, C. Haug, R. Delto, C. Wild, R. Brenn: Nucl. Instr. Meth. B **190**, 324 (2002)
39. R. Groleau, S.C. Gujrathi, J.P. Martin: Nucl. Instr. Meth. **218**, 11 (1983)
40. R. Ishigami, Y. Ito, K. Yasuda, M. Sasase, S. Hatori: AIP Conf. Proc. **689**, 474 (2002)
41. R. Siegele, I. Orlic, D.D. Cohen: Nucl. Instr. Meth. B **190**, 301 (2002)
42. H.J. Whitlow: Mass and energy dispersive recoil spectrometry: a new quantitative depth profiling technique for microelectronic technology. In: *High Energy and Heavy Ion Beams in Materials Analysis*, ed. by J.R. Tesmer, C.J. Maggiore, M. Nastasi, J.C. Barbour (MRS, Pittsburgh 1990) pp. 73–85
43. G. Vizkelethy: Nuclear reaction analysis: particle–particle reactions. In: *Handbook of Modern Ion Beam Material Analysis*, ed. by J.R. Tesmer, M. Nastar et al. (MRS, Pittsburgh 1995) Chap. 6
44. D. Dieumegard, B. Maurel, G. Amsel: Nucl. Instr. Meth. **168**, 93 (1980)
45. R.A. Jarjis: Nuclear cross-section data for surface analysis Vols. I–III, University of Manchester report (1979)

46. A.K.M. Fazlul Hoque, M. Khaliquzzaman, M.D. Hssain, A.H. Khan: Fluoride **35**, 176 (2002)
47. N.M. Uzunov, I.P. Penev, A.M. Artinian: Nucl. Instr. Meth. B **95**, 276 (1995)
48. J. Räisänen: Nucl. Instr. Meth. B **40/41**, 638 (1989)
49. A. Anttila, R. Hänninen, J. Räisänen: J. Radioanal. Chem. **62**, 441 (1981)
50. I. Giles, M. Peisach: J. Radioanal. Chem. **50**, 307 (1979)
51. J. Pallon, V. Auzelyte, M. Elfman, M. Garmer, P. Kristiansson, K.G. Malmqvist, C. Nilsson, A. Shariff, M. Wegdén: Nucl. Instr. Meth. B **219–220**, 988 (2004)
52. G. Devés, R. Ortega: Anal. Bioanal. Chem. **374**, 390 (2002)
53. M.B.H. Breese, D.N. Jamieson, P.J.C. King (eds.): Spatially resolved ion channelling techniques. In: *Materials Analysis Using a Nuclear Microprobe* (Wiley, New York 1996) pp. 201–246
54. H.W. Lefevre, R.M.S. Schofield, D.R. Ciarlo: Nucl. Instr. Meth. B **54**, 47 (1991)
55. G.S. Bench, K.A. Nugent, M. Cholewa, A. Saint, G.J.F. Legge: Nucl. Instr. Meth. B **54**, 390 (1991)
56. D. Habermann: Mineral. Petrol. **76**, 247 (2002)
57. C. Yang, N.P.-O. Homman, K.G. Malmqvist, L. Johansson, N.M. Halden, V. Barbin: Scanning Microsc. **9**, 43 (1995)
58. K.G. Malmqvist, M. Elfman, G. Remond, C. Yang: Nucl. Instr. Meth. B **109/110**, 227 (1996)
59. C. Yang, K.G. Malmqvist, M. Elfman, P. Kristiansson, J. Pallon, A. Sjöland, R.J. Utui: Nucl. Instr. Meth. B **130**, 746 (1997)
60. G. Vizkelethy, F.O. McDaniel, S. Thevuthasan, J.R. Tesmer: Nucl. Instr. Meth. B **219–220**, (2004)
61. N. Matsumura, T. Haga, S. Muto, Y. Nakata, N. Yokohama, K. Numata, K. Yabuta: Physica E **13**, 1168 (2002)
62. F. Watt, I. Rajta, J.A. Van Kan, A.A. Bettiol, T. Osipowicz: Nucl. Instr. Meth. B **190**, 306 (2002)
63. L. Wei: Mater. Res. Soc. Symp. Proc. **721**, 105 (2002)
64. L. Wei: AIP Conf. Proc. **689**, 464 (2002)
65. T. Osipowicz, H.L. Seng, L.S. Wielunski, E.S. Tok, G. Breton, J. Zhang: Nucl. Instr. Meth. B **190**, 345 (2002)
66. E.J. Teo, M. Alkaisi, A.A. Bettiol, T. Osipowicz, J. Van Kan, F. Watt, A. Markwitz: Nucl. Instr. Meth. B **190**, 339 (2002)
67. G. Amsel, G. Battistig, A. L'Hoir: Nucl. Instr. Meth. B **201**, 325 (2003)
68. N. Kawachi, T. Katabuchi, M. Yamaguchi, Y. Tagishi: Nucl. Instr. Meth. B **190**, 195 (2002)
69. P. Kristiansson, T. Larsson, S. Malmberg, L. Malmqvist, K. Malmqvist, M. Elfman, J. Pallon: Nucl. Instr. Meth. B **158**, 665 (1999)
70. P. Kristiansson, M. Elfman, T. Larsson, S. Malmberg, L. Malmqvist: Ink distribution on newsprints with high and low smearing, Proc. 1998 International Printing and Graphic Art Conference, Quebec
71. K.-E. Saarela, L. Harju, J.-O. Lill, J. Rajander, A. Lindroos, S.-J. Heselius, K. Saari: Holzforschung **56**, 380 (2002)
72. H. Khodja, E. Berthoumieux, L. Daudin, J.-P. Gallien: Nucl. Instr. Meth. B **181**, 83 (2001)
73. H. Yamazaki, Y. Takahashi, K. Ishii, S. Matsuyama, T. Satoh, H. Orihara: J. Nucl. Sci. Tech. Suppl. **3**, 560 (2002)

74. U. Alonso, T. Missana, A. Patelli, V. Rigato, P. Rivas: *J. Contam. Hydrol.* **61**, 95 (2003)
75. C.G. Ryan: *Nucl. Instr. Meth. B* **104**, 377 (1995)
76. C.G. Ryan, D.N. Jamieson, W.L. Griffin, G. Cripps, R. Szymanski: *Nucl. Instr. Meth. B* **181**, 12 (2001)
77. D.D. Hickmott, J. Stimac, A.C.L. Larocque, C. Wettland, A. Brearly: *AIP Conf. Proc.* **689**, 440 (2002)
78. L.J. Cabri, P.J. Sylvester, M.N. Tubrett, A. Peregoedova, J.H.G. LaFlamme: *Can. Mineral.* **41**, 321 (2003)
79. P. Kristiansson, U. Hålenius, H. Skogby, M. Elfman, K.G. Malmqvist, J. Pallon: *Nucl. Instr. Meth. B* **158**, 562 (1999)
80. M. Volfinger: *J. Radional. Nucl. Chem.* **253**, 413 (2002)
81. L. Malmqvist, K. Kristiansson, P. Kristiansson: *Nucl. Instr. Meth. B* **150**, 484 (1999)
82. C.A. Pineda-Vargas, A.L. Rodgers, M.E. Eisa: *Radiat. Phys. Chem.* **71**, 947 (2004)
83. W.J. Przybyłowicz, J. Mesjasz Przybyłowicz, P. Migula, E. Glowacka, M. Nakonieczny, M. Augustyniak: *Nucl. Instr. Meth. B* **210**, 343 (2003)
84. J. Pallon, M. Elfman, P. Kristiansson, K. Malmqvist, E. Graneli, A. Sellborn, C. Karlsson: *Nucl. Instr. Meth. B* **158**, 312 (1998)
85. R. Ortega, A. Suda, E. GeVes: *Nucl. Instr. Meth. B* **210**, 364 (2003)
86. M. Wegdén, P. Kristiansson, C. Ceberg, P. Munck af Rosenschöld, V. Auze-lyte, M. Elfman, K.G. Malmqvist, C. Nilsson, J. Pallon, A. Shariff: *Nucl. Instr. Meth. B* **219–220**, 67 (2004)
87. W.J. Przybyłowicz, J. Mesjasz-Przybyłowicz, V.M. Prozesky, C.A. Pineda: *Nucl. Instr. Meth. B* **130**, 335 (1997)
88. D. Budka, J. Mesjasz-Brzybyłowicz, W.J. Brzybyłowicz: *Radiat. Phys. Chem.* **71**, 785 (2004)
89. G.E. Coote, R.W. Gauldie, I.F. West: *Nucl. Instr. Meth. B* **54**, 144 (1991)
90. J.L. Campbell, J.A. Babaluk, N.M. Halden, A.H. Kristofferson, J.A. Maxwell, S.R. Mejia, J.D. Reist, W.J. Teesdale: *Nucl. Instr. Meth. B* **150**, 260 (1999)
91. M. Elfman, K.E. Limburg, P. Kristiansson, K. Malmqvist, J. Pallon: *Nucl. Instr. Meth. B* **150**, 272 (1999)
92. K.G. Malmqvist: *Nucl. Instr. Meth. B* **85**, 84 (1994)
93. K.G. Malmqvist: *Nucl. Instr. Meth. B* **113**, 336 (1996)
94. E. Swietlicki, B.G. Martinsson, P. Kristiansson: *Nucl. Instr. Meth. B* **109/110**, 385 (1996)
95. T.A. Cahill: *Nucl. Instr. Meth. B* **109/110**, 402 (1996)
96. F. Lucarelli, P.A. Mando, S. Nava, P. Prati, A. Zucchiatti: *Nucl. Instr. Meth. B* **161/163**, 819 (2000)
97. D.D. Cohen, E. Stelcer, D. Garton: *Nucl. Instr. Meth. B* **190**, 466 (2002)
98. E. Perdrix, G. Lecornet, P. Delain, S. Sobanska, H. Guegan, J. Laureyins, A. Robache, A. Morin: *J. Phys.* **107**, 1037 (2003)
99. W. Maenhaut, R. Salomonovic, J. Ptasinski, G.W. Grime: *Nucl. Instr. Meth. B* **130**, 576 (1997)
100. E. Dooryhée, P. Martinetto, Ph. Walter, M. Anne: *Radiat. Phys. Chem.* **71**, 863 (2004)
101. P.M. O'Neill, D.C. Creagh, M. Sterns: *Radiat. Phys. Chem.* **71**, 841 (2004)
102. J. Absil, H.-P. Garnir, D. Strivay, C. Oger, G. Weber: *Nucl. Instr. Meth. B* **198**, 90 (2002)

103. K. Cummings, W.A. Lanford, M. Feldmann: Nucl. Instr. Meth. B **136/138**, 858 (1998)
104. J.-O. Lill, M. Ström, M. Brenner, A. Lindroos: Nucl. Instr. Meth. B **189**, 303 (2002)
105. N. Grassi, A. Migliori, P.A. Mandó, H. Calvo, H. Calvo de Castillo: Nucl. Instr. Meth. B **219–220**, 48 (2004)
106. O. Enguita, T. Calderon, M.T. Fernandez-Jimenez, P. Beneitez, A. Millan, G. Garcia: Nucl. Instr. Meth. B **219–220**, 53 (2004)
107. L. Simonot, A. Zobelli, M. Elias, J. Salomn, J.C. Dran: J. Trace Microprobe Tech. **21**, 35 (2003)
108. M.H. Abraham, G.W. Grime, M.A. Marsh, J.P. Northover: Nucl. Instr. Meth. B **180**, 688 (2001)
109. P. Dillman, R. Balasubramaniam, G. Beranger: Corrosion Sci. **44**, 2331 (2002)
110. K. Morito, S. Nagashima, M. Kato, T. Kotani, M. Uda: Nucl. Instr. Meth. B **109/110**, 569 (1996)

27 Atomic Structure

L.J. Curtis¹ and I. Martinson²

¹ Department of Physics and Astronomy, University of Toledo, Toledo, Ohio
43606-3390, USA

ljc@physics.utoledo.edu

² Atomic Spectroscopy, Department of Physics, Lund University, Box 118, 22100
Lund, Sweden

Indrek.Martinson@fysik.lu.se

27.1 Introduction

The fact that particle accelerators can produce isotopically pure, monoenergetic beams of atomic ions over a wide range of speeds and ionic states means that these machines can be utilized to study both atomic and nuclear properties with no experimental modifications other than beam targeting and detection equipment. By a suitable choice of accelerator energy and ion source capabilities, virtually any energy level of any charge state of any atom can be studied. The excitation can be produced by impinging the beam on a thin foil, a gas cell, or either a crossed or a collinear laser beam. The de-excitation can be interrogated by microwave, optical, or x-ray detection, or by electron spectroscopy, with the possibility of time-of-flight temporal resolution.

One of the most flexible and widely used techniques is that of beam-foil spectroscopy (BFS), in which the accelerated particles impact a thin foil, within which they undergo excitation and ionization in a dense environment, and then suddenly emerge into a low-density, field-free environment. The excitation mechanism has the possibility of providing both time resolution and coherent anisotropic excitation. BFS is now a mature field (having been pursued for about 40 years), and has provided much new data, elucidated fundamental atomic processes, and furthered applications in a variety of related fields. Improved technical capabilities have now broadened its applicability to permit new types of measurements of unprecedented accuracy and scope. The experimental groups in North America and Europe that pioneered the field have now been joined by a new generation of international practitioners, with new and active groups in China, India, Japan, and South Africa, for example.

In this chapter, we shall review some of the successful applications of accelerator-based methods (giving primary emphasis to BFS), and indicate other areas that hold promise for future measurements.

27.2 The Present Status of Atomic-Structure Research

The knowledge of energy levels, transition probabilities, and other atomic-structure data has been greatly advanced by the development of new light sources and technical improvements in the use of traditional light sources, motivated by urgent applicational needs. These light sources include the use of fast ion beams, tokamak- and laser-produced plasmas, orbiting observation of astrophysical objects, recoil ions from ion-atom collisions, electron beam ion sources (EBISs) and ion traps (EBITs), electron cyclotron resonance (ECR) ion sources, ion storage rings, and Fourier transform spectroscopy (FTS), as well as improved use of traditional spark and vacuum arc sources. Since the interpretation of time-resolved measurements requires a thorough knowledge of the radiatively coupled energy-level structure, these studies must be accompanied by high-resolution wavelength and energy-level measurements.

The application of these techniques has produced large blocks of generic data, for example along isoelectronic sequences, in highly excited Rydberg series, in core-excited “hollow-atom” states, and in coherently excited “entangled” states. However, much remains to be done, particularly in the determination of relative intensities of decays from levels that have multiple exit channels (branching-fraction data) in atoms and ions.

27.3 Capabilities and Limitations of the BFS Method

The BFS light source possesses many unique characteristics that permit types of measurements that are not accessible to other methods of excitation. As is often the case, capabilities can also impose restrictions, but many features that were earlier considered as limitations have now been reduced or eliminated. A typical BFS experimental setup is shown in Fig. 27.1.

The primary advantage of BFS is its universal applicability since, given a suitable accelerator and ion source, virtually any ion of any atom can be studied in a time-resolved manner. However, because the time resolution is achieved by time-of-flight methods, this does impose limitations on the minimum and maximum time windows (ranging from the shortest beam length that can be optically resolved to the longest flight path that can be efficiently monitored). Thus the lifetime range for BFS methods is approximately 0.1 ps to 300 ns. Lifetime measurements are usually obtained by translating the foil upstream and downstream stepwise while the optics views a fixed point along the beam. This provides precise time resolution, but it also requires data collection in a multiscaling mode, with the detection apparatus sensitive to only one increment of time-since-excitation at a time. The solid foil provides a region of very high density that causes heavy excitation of all levels, including extensive core excitations. This does, however, involve nonselective excitation, so that decay curves are sums of the many exponentials that are inherent in the repopulation by cascade transitions. The density of the beam

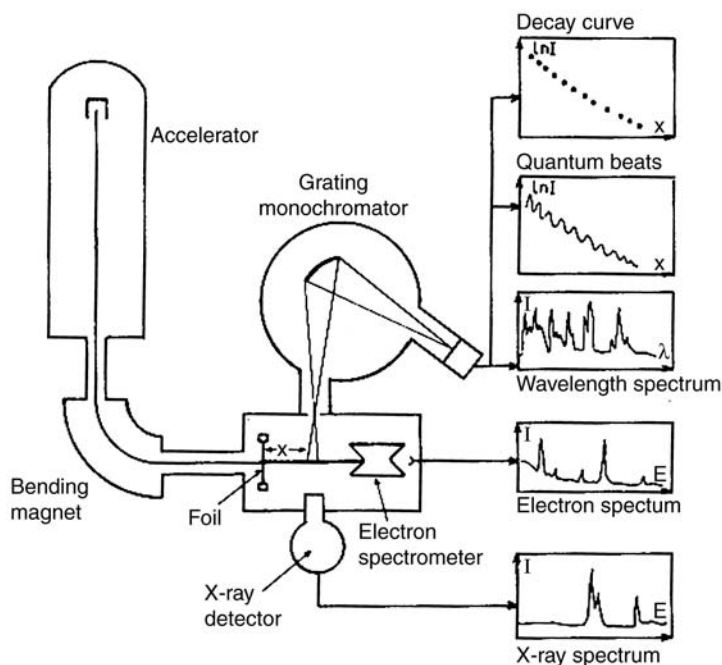


Fig. 27.1. Schematic representation of experimental arrangements for a variety of types of BFS measurements

is very low (much rarer than the residual gas even at ultrahigh vacuum), so there is no significant collisional radiation damping, and there are no significant interionic fields. Because the beam suddenly emerges from the dense foil into the evacuated flight path, the excitation can be coherent, and can exhibit interference and quantum beats. The preferred direction and the possibility of tilting the foil relative to the beam can produce useful excitation anisotropies.

A typical BFS spectrum is shown in Fig. 27.2. Here 24 MeV Fe ions from a tandem accelerator were excited in a carbon foil. The spectrum shows transitions in highly charged Fe only. Several of these lines appear in the spectrum of the solar corona and flares. However, there they are often blended with lines from other elements, for example O, S, and Cl.

27.4 Technical Developments That Enhance Capabilities

The lifetimes of excited states in atoms and ions are of interest for atomic theory, being sensitive indicators of the accuracy of the wave functions applied in calculations. Furthermore, they are also crucial for astrophysics, for example in the determination of element abundances in the Sun and stars.

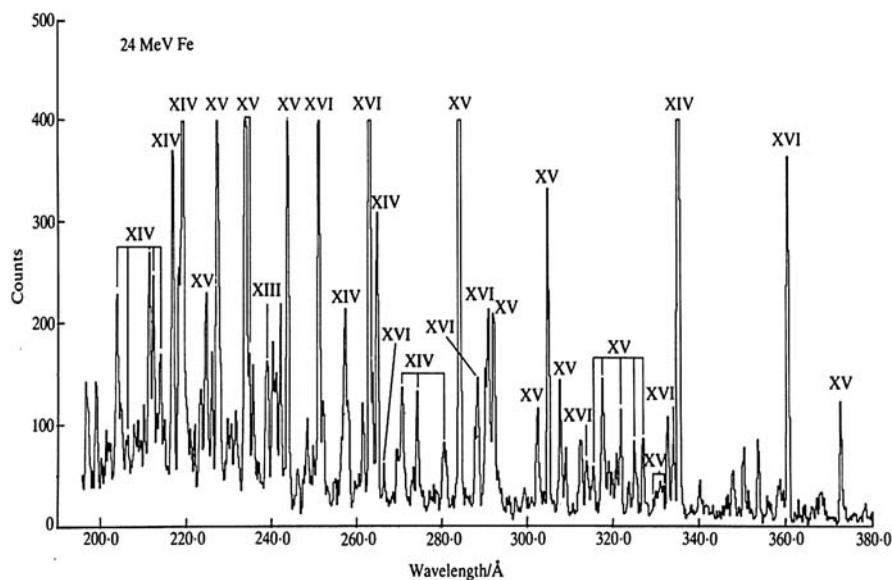


Fig. 27.2. Beam-foil spectrum of iron in the wavelength region 20–38 nm

There exist many techniques which allow such studies in neutral and singly ionized atoms, whereas BFS is the only method for highly charged ions. However, since many levels in an ion can be populated in BFS excitation, the decay curve of a given level is often complicated because of feeding from higher-lying states. In the early days of BFS, when such decay curves were analyzed by multiexponential fitting, large uncertainties (up to 50%) could be encountered. Fortunately, such problems have now been largely eliminated.

27.4.1 Elimination of Cascade Effects

The decay curve distortions introduced by cascade repopulation can be accounted for by a technique that involves measurement, under similar conditions, of the decay curve of the primary transition as well as those of all of its significant cascade feeders. This is known as the method of “adjusted normalization of decay curves” (ANDC).

The instantaneous population change of a given level is governed by the difference in transition rates between the cascades into and the decays out of that level. To within constant factors of the transition probabilities and detection efficiencies, the time dependence of these populations is specified by the variation of the detected intensities along the time-resolved decay curves. Thus, rather than expanding a single decay curve measurement on a set of mathematical exponentials, it is possible to expand the decay curve of the primary level on the measured decay curves of the cascades that repopulate

it. In this formulation, the lifetime and the detection efficiencies can be accurately extracted as linear fitting parameters, with internal-consistency criteria that test and verify whether all significant cascades have been included.

A schematic representation of an ANDC determination of the lifetime of the $6p\ ^2P_{3/2}$ level in Tl III is shown in Fig. 27.3. The population rate equation for this system, reexpressed in terms of measured decay curve intensities, is given by

$$\tau_{6p} \frac{dI_{6p}}{dt}(t_i) = \xi_{6d} I_{6d}(t_i) + \xi_{7p} I_{7p}(t_i) - I_{6p}(t_i) . \quad (27.1)$$

Here the lifetime τ_{6p} of the $6p$ level and the normalizations ξ_{6d} and ξ_{7s} for the $6d$ and $7s$ cascades (relative to that of the primary decay curve) are determined by simultaneous solution of the relationships provided by evaluating (27.1) at each common point t_i on the various decay curves. This method has been used to obtain many lifetimes of high reliability and precision.

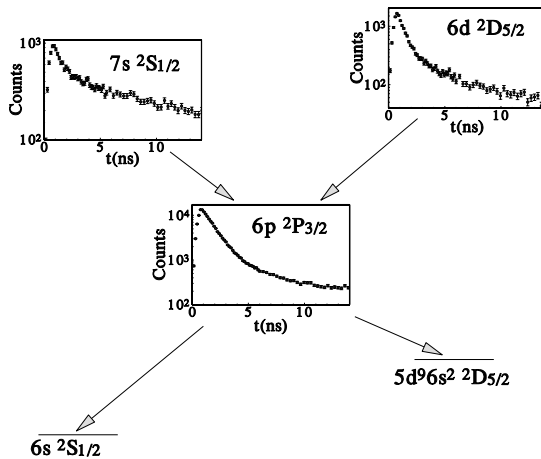


Fig. 27.3. ANDC analysis of cascade-correlated decay curves

The precise determination of this lifetime in Tl III is of application in astrophysics, because the corresponding spectral lines are observed in the spectra of so-called chemically peculiar stars. This ANDC measurement yielded a value 1.06 ± 0.04 ns (4%), as compared with earlier curve-fitted measurements, which had uncertainties of 10–20%.

27.4.2 Multiplexed Detection

Although the low particle density characteristic of BFS has many positive aspects, it also results in low intensities in the emitted light. This is exacerbated by the fact that the time-of-flight method involves single-channel detection, in

which each individual time-since-excitation is separately measured in a multiscaling mode. Thus the count rates for a specific wavelength measured at a specific time-since-excitation can be low, especially on the long-time tails of the decay curve. However, these limitations have been greatly reduced by the development of position-sensitive detectors (PSDs). While detection is still single-channel in the time domain, it is now possible to utilize multiplexed detection in wavelength. By simultaneously measuring a broad range of wavelengths at each time-since-excitation, significant advantages are achieved. A landscape plot showing a multiplexed data set (intensity vs. both wavelength and decay time) is presented in Fig. 27.4.

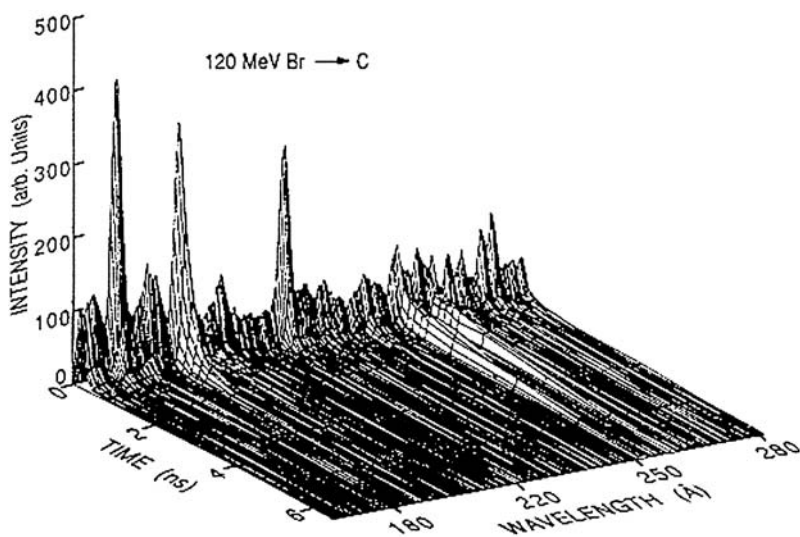


Fig. 27.4. Multiplexed landscape of intensity, wavelength, and decay time

In addition to the obvious gain in data collection efficiency afforded by a sensitivity to a broad range of wavelengths, many other advantages can be cited. By virtue of simultaneous measurement, many possible systematic uncertainties are eliminated by considering subtracted differences between the decay curves. It is also possible to examine the exponential content across a line profile to check for line blending. For systems for which the wavelengths of the primary and cascade transitions lie within the multiplexed wavelength range, it is possible to perform ANDC analyses in real time.

A significant advantage involves differential lifetime measurements, which can reveal small differences in lifetimes among the fine-structure transitions within a multiplet. Consider, for example, the fine structure of the Be-like $2s3s\ ^3S_1-2s3p\ ^3P_J$ transitions. Here the $J=1$ level possesses a strong intercombination relativistic-E1 decay channel to the $2s^2\ ^1S_0$ ground state that is

forbidden to the $J=0$ and 2 levels by angular-momentum conservation. In contrast to the usual situation, in which decay curve measurements determine only lifetimes and not branched transition probabilities, the branching fraction for this ground-state channel can be specified by differential lifetime measurements. A plot of the decay curves for these three levels in Be-like Ne VIII is shown in Fig. 27.5. In addition to their theoretical interest, such intercombination lines are also important for the diagnostics of laboratory and astrophysical plasmas.

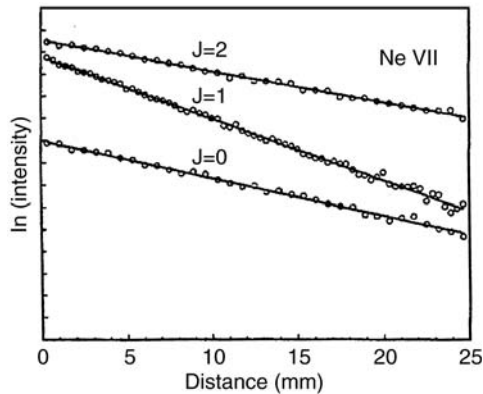


Fig. 27.5. Decay curve measurements within a fine-structure multiplet

Prior to the availability of PSD devices, these decay curves were measured separately and subtracted channel-by-channel. When the emitted radiation from all three of these transitions is detected simultaneously at each distance downstream, many systematic uncertainties cancel in differential measurements, and the precision can be greatly enhanced over previous determinations.

In this example, an extra channel to the ground state is determined by differential lifetime measurements. As will be discussed in Sect. 27.5.4, multiply excited ions sometimes possess a J -dependent autoionization channel that can be determined by using these same differential lifetime measurement methods.

Figure 27.6 shows a measurement of the quantum beats that arise from coherent excitation of fine-structure levels in He I. These frequencies represent fine-structure intervals in atoms and ions, i.e. quantities that are crucial for understanding various relativistic and higher-order effects. The relative frequency measurements in Fig. 27.6 were made with a common time base by using two different spectrometers so as to obtain a precise relative normalization. Similar measurements can now be made through PSD detection, and could provide precise accurate relative values for fine- and hyperfine-structure splittings.

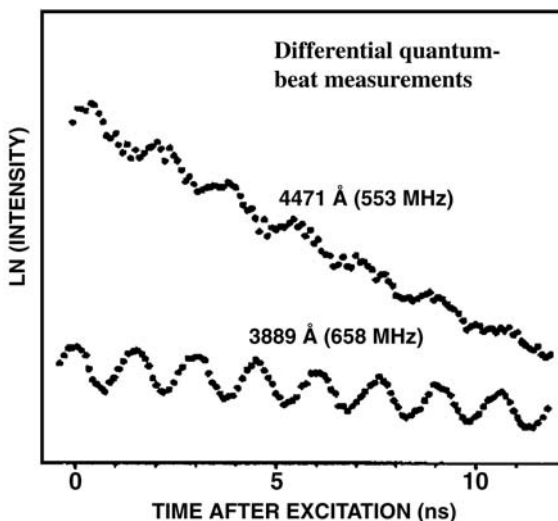


Fig. 27.6. Simultaneously measured quantum beat frequencies in He

27.5 Contributions of Accelerator-Based Studies

The use of accelerator-based measurements has produced large databases that permit comprehensive predictive, semiempirical systematizations.

27.5.1 Isoelectronic Measurements for Alkali-Metal-Like Systems

One particularly valuable application of BFS has been the study of lifetimes along isoelectronic sequences. By utilizing a variety of accelerators to trace a sequence of ions with a fixed number of electrons as a function of nuclear charge Z , it is possible to examine and separate processes that scale differently with Z . For highly charged ions, properties are often different from those observed for neutral and few-times-ionized systems (e.g., magnetic fine structure can exceed electrostatic gross structure, forbidden transitions can exceed allowed transitions, etc.). Rather than studying the lifetimes directly, these quantities can be used converted to line strength factors S_{ik} , defined theoretically in terms of the theoretical dipole transition matrix element as

$$S_{ik} = \sum_{m_i} \sum_{m_k} |\langle i | \mathbf{r} / a_0 | k \rangle|^2 . \quad (27.2)$$

The isoelectronic behavior of S_{ik} is much more regular and slowly varying than that of the lifetime, since the latter contains sums over wavelength factors that vary rapidly as a function of nuclear charge. The line strength factor

can be deduced from the measured lifetime τ_i , the measured transition wavelength λ_{ik} , the statistical weight of the level g_i , and the branching fraction B_{ik} (unity for a decay with only one exit channel) using

$$S_{ik} = [\lambda_{ik}(\text{nm})/126.538]^3 g_i B_{ik} / \tau_i(\text{ns}) . \quad (27.3)$$

This predictive power of these isoelectronic studies is illustrated by Fig. 27.7, which shows a plot of the scaled line strength factors for the lowest resonance transitions of the Na isoelectronic sequence.

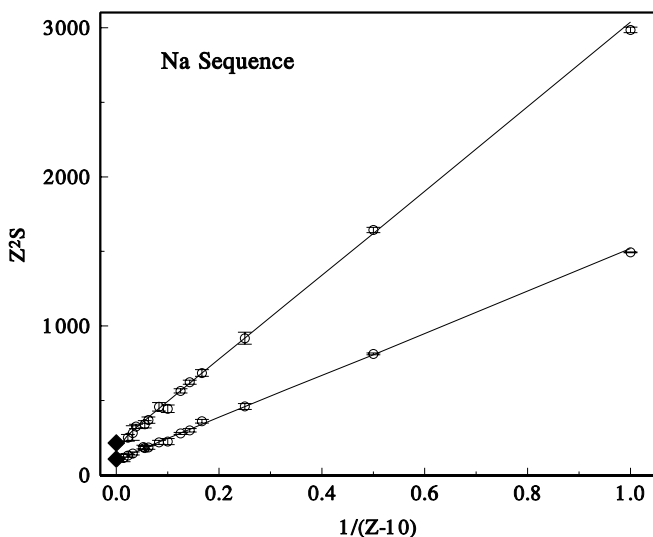


Fig. 27.7. Scaled line strengths for the 3s–3p transitions in the Na isoelectronic sequence

It has been observed that these quantities exhibit a nearly linear behavior when plotted vs. a suitably chosen reciprocal screened charge, and can be characterized by a fit to the empirical function

$$Z^2 S \approx S_H + b/(Z - C) . \quad (27.4)$$

Moreover, this linearity appears to extrapolate to the hydrogenic limit at infinite Z , which is given by

$$S_H \rightarrow 3n^2(n^2 - 1)g_i/4 . \quad (27.5)$$

While the extreme limit of infinite Z is clearly a nonphysical asymptote, tests have been made which indicate that this linearization remains valid through the stable isotopes.

Because of the availability of very high-energy accelerators, BFS measurements for the Li isoelectronic sequence have been extended to very high Z

and include a lifetime determination in three-electron uranium. When this result was put on a plot of the charge-scaled line strength factor, it initially appeared that there was a breakdown in this linear trend, but a subsequent analysis provided an interesting test of relativistic effects. This is shown in Fig. 27.8(a), which reveals a clear dip in value below the linear trend.

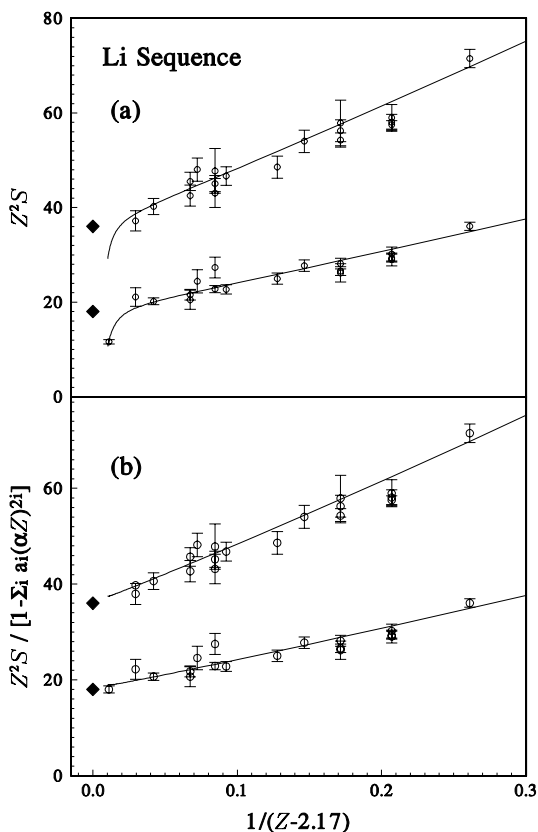


Fig. 27.8. Scaled line strengths for the Li isoelectronic sequence: (a) with no relativistic corrections; (b) with single-particle Dirac relativistic corrections

Two alternative mechanisms were considered to explain the origin of this dip. One possibility is that a many-electron redistribution of the electrons in the complex core could be occurring. A second possibility is that relativistic corrections only to the single-electron formulation of the line strength factor are occurring, with no major modification of the core. It was possible to examine these two models semiempirically. Higher-order calculations were made using the Dirac equation rather than the Schrödinger equation, and used to correct the abscissae of the measured data points for a modified

empirical fitting function given by

$$Z^2 S \left[1 - \sum_i a_i [\alpha(Z - C)]^{2i} \right]^{-1} \approx S_H + b/(Z - C) . \quad (27.6)$$

As shown in Fig. 27.8(b), this second model successfully restored the linearity. Subsequent theoretical calculations have confirmed this result. Because of this linearity, it is now possible to interpolate, extrapolate, and smooth data throughout entire isoelectronic sequences. This can be done on the basis of only a few accurate measurements, sufficient to specify the empirical screening parameter C on the abscissa.

Cancellation Effects

While sufficient numbers of precision measurements now exist to permit accurate semiempirical specification of the lifetimes of the lowest resonance transitions in all alkali-metal-like systems, other transitions in these systems still pose challenges for future measurements. Transition rates among the excited states in alkali-metal-like systems are difficult to specify either experimentally or theoretically. Experimentally, the decay of each upper level is branched to many lower levels, and lifetimes cannot be converted to transition probabilities because of the lack of branching-fraction data. Theoretically, the dipole matrix elements can be affected by strong and highly localized cancellation effects that are extremely sensitive to subtle perturbations not normally accounted for in the basis set wave functions.

The sensitivity of these cancellation effects can be illustrated by the semiempirical quantum defect method, in which the term energy T for an alkali-metal-like ion can be expressed in terms of the Rydberg formula

$$T = Ry \zeta^2 / (n^*)^2 . \quad (27.7)$$

Here Ry is the Rydberg constant, ζ is the net charge of the nucleus and inner core electrons, and n^* is the effective quantum number of a hydrogenlike ion that would have the same energy as the complex ion in question. Theoretical insight can be gained by writing the effective quantum number as $n^* \equiv n - \delta$, where n is the principal quantum number in a hydrogenlike atom, and the quantity δ is known as the effective “quantum defect”. It can be shown that the quantum defect can be related to a phase shift in the radial wave function, whereby the radial coordinate at large r for the wave function of the complex ion is shifted by $\delta\pi$ from a corresponding hydrogenlike wave function. In the isoelectronic variation of a given transition between two excited levels, differential isoelectronic variations in the quantum defects of the upper and lower levels lead to regular cancellations in the dipole matrix element. This is illustrated Fig. 27.9, which is a plot of n_s^* vs. n_p^* for the s-p transitions in the Na isoelectronic sequence. The solid lines correspond to the nodes where

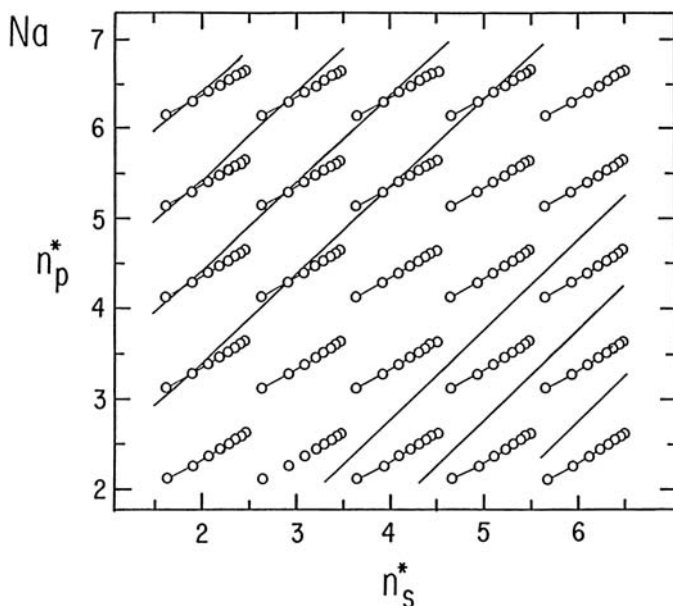


Fig. 27.9. Cancellation plot for the ns - $n'p$ transitions in the Na isoelectronic sequence

these cancellations would be expected to occur, and the connected circles represent the experimental effective quantum numbers.

Wherever the circles fall near a nodal line, an anomalously small transition probability is expected to occur. This exposition clearly illustrates how theory can be hindered by the many cancellations that occur in these transition integrals.

Moreover, experimental measurements of line strength factors between excited levels are almost entirely lacking. These levels undergo branched decay, and lifetimes cannot be reduced to line strength factors without accompanying branching-fraction data. Because of difficulties with relative in-beam intensity calibrations, branching-fraction data for multiply charged ions are virtually nonexistent. However, new methods offer promise in this area, which will be discussed in Sect. 27.6.1.

27.5.2 Isoelectronic Measurements for Alkaline-Earth-Like Systems

These semiempirical methods can also be applied to two-valence-electron atoms. Because of the singlet-triplet nature of transitions of the form ns^2 - $nsnp$, the resonance and intercombination transitions to ground can be related to each other through intermediate-coupling (IC) amplitudes describing the deviation of their angular-momentum states from pure LS coupling. The

IC amplitudes can be specified as mixing angles deduced from spectroscopic energy-level data.

As shown for the $5s5p$ transitions in the Cd sequence in the upper panel of Fig. 27.10, the singlet–triplet mixing angles can be presented on a linearized plot that converges to the jj coupling limit at infinite Z .

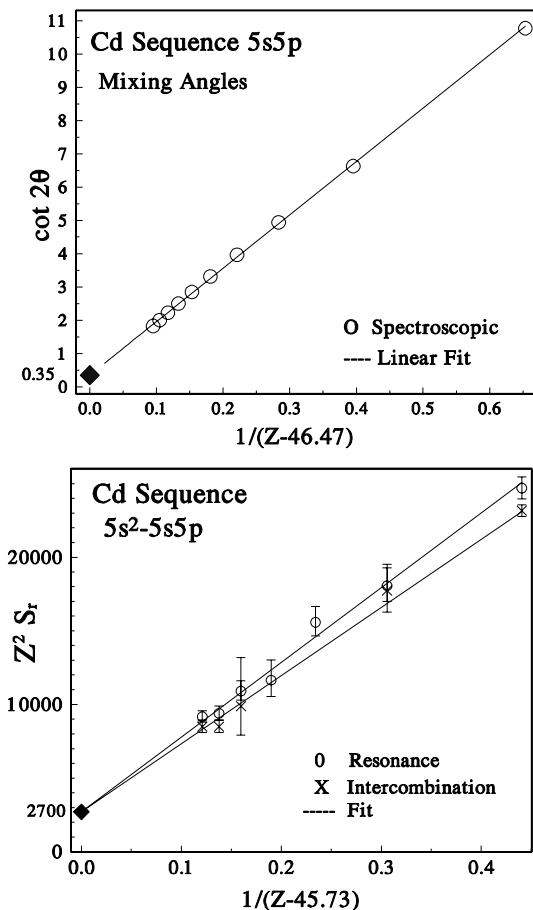


Fig. 27.10. Mixing angle and scaled reduced line strength factors for the Cd iso-electronic sequence

These mixing angles can be used to reduce lifetime measurements for the resonance and intercombination lines to quantities proportional to the square of their radial transition matrix elements. If we denote these quantities as “reduced line strength factors” S_r , they are given by

$$S_r(\text{Res}) \equiv S(\text{Res}) / \cos^2 \theta, \quad (27.8)$$

$$S_r(\text{Int}) \equiv S(\text{Int})/\sin^2 \theta. \quad (27.9)$$

As illustrated for the $5s^2-5s5p$ transitions in the Cd sequence in the lower panel of Fig. 27.10, these scaled reduced line strength factors (like the alkali-metal-like case) exhibit a linearity vs. reciprocal screened charge that converges to the hydrogenic value at infinite Z .

This formulation permits another type of test of relativistic corrections. In the nonrelativistic limit, the Schrödinger equation and LS coupling should be expected to hold, and govern the IC amplitudes. However, for highly complex atoms, the use of the Dirac equation and jj coupling becomes necessary. In that limit, the spin dependence occurs not only in the angular-momentum portion of the wave function, but also in its radial portion. In the Dirac equation, there are two radial wave functions, corresponding to $(j, j') = (\frac{1}{2}, \frac{1}{2})$ and $(\frac{1}{2}, \frac{3}{2})$.

Interestingly, the reformulation relative to the jj limit yields a mixing angle that is very similar to that of the nonrelativistic case, but with a phase shift ξ dictated by the difference between the radial matrix elements. If we denote the radial transition integrals as $R_{2j,2j'}$, the relativistic analogue of (27.8) and (27.9) has been shown to be

$$S_r(\text{Res}) \equiv S(\text{Res})/\cos^2(\theta - \xi), \quad (27.10)$$

$$S_r(\text{Int}) \equiv S(\text{Int})/\sin^2(\theta - \xi), \quad (27.11)$$

where

$$\tan \xi \equiv \sqrt{2}(R_{31} - R_{11})/(2R_{31} + R_{11}). \quad (27.12)$$

MCDF calculations for these quantities have been made that indicate that these effects are negligible for all but the Hg isoelectronic sequence. The correction is most significant for the lower ions of the sequence, where the angular portion of the mixing is smallest. A plot of the mixing-angle linearization, with and without this correction, is shown in the upper panel of Fig. 27.11. The effect on the exposition of the line strength factors is shown in the lower panel of Fig. 27.11. Prior to this correction, the points for the first two charge states for the intercombination transitions dipped below the trend at higher Z . The application of this correction moved the Hg I and Tl II points onto the linear trend. Note that for this sequence only four stable ions exist, but the trend permits accurate predictions for the radioactive members of the sequence.

A similar analysis can be applied to the $2p^6-2p^53s$ transitions in the Ne isoelectronic sequence. Here the fact that the transitions have $\Delta n = 1$ rather than $\Delta n = 0$ results in much shorter lifetimes (because of the wavelength-cubed dependence). For highly charged ions these lifetimes are too short to be measured by time-of-flight and too long to be measured by line widths, but they have important applications in, for example, X-ray laser technology. Thus the accurate extrapolation of the measurements accessible to BFS measurement is valuable.

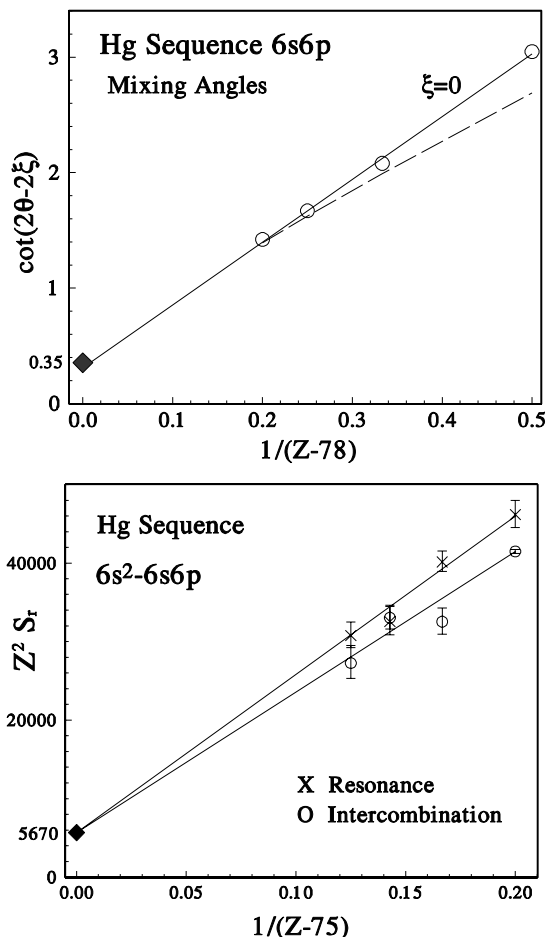


Fig. 27.11. Mixing angles and scaled reduced line strength factors for the Hg isoelectronic sequence

27.5.3 High- n and $-\ell$ States

When a multiply charged fast ion beam is sent through a thin foil it can capture an electron at the foil exit, producing a state of very high n and ℓ . Such states are also formed in many other types of light sources. In thermonuclear fusion plasmas, such states are populated by charge exchange, between multiply charged ions in the plasma and fast beams of neutral H and He injected into the plasma for auxiliary heating. Accurate wavelengths are crucial for the determination of the plasma temperature. However, when formed in denser plasmas, these states tend to be destroyed by Stark quenching (electric-field mixing of the individual ℓ states) or collisional deexcitation. Because of the tenuous nature of the beam-foil plasma, these states persist there until they

decay radiatively, and thus can provide much valuable information about the ion.

The circular-orbit ($\ell = n - 1$) transitions in the so-called “yrast” chain are unbranched, and produce strong features in beam-foil spectra, as do the adjacent “yrare” ($\ell = n - 2$, $\ell = n - 3$) transitions. For sufficiently high n these states are nonpenetrating, and the inner electrons can be treated as a deformable core of charge that can be characterized by effective dipole and quadrupole polarizabilities α_d and α_Q , and a nonadiabatic correlation parameter β . This formulation is known as the “core polarization model”, wherein the term value T of a level n, ℓ can be written as

$$T = T_H + \alpha_d \langle r^{-4} \rangle + (\alpha_Q - 6\beta) \langle r^{-6} \rangle. \quad (27.13)$$

Here T_H is the value for a corresponding hydrogenlike (single-electron) ion with a nuclear charge equal to the net charge of the ion (the nucleus dressed in the electronic core), and $\langle r^{-k} \rangle$ is the expectation value for the radius raised to the power k in this hydrogenlike ion. Wavelength measurements for a sampling of yrast and yrare transitions in a given system permit the determination of the two parameters $A = \alpha_d$ and $B = \alpha_Q - 6\beta$ for its ionic core.

These quantities not only permit the prediction of all other levels and transitions of high n and ℓ in the system, but also provide electric polarizabilities that describe the ionic core in the presence of any external electric field. These nonpenetrating orbitals provide a gentle probe of the inner core, and differences in the interaction energies between orbitals that differ slightly in n and ℓ values can reveal subtle properties of the ionic core.

27.5.4 Multiply Excited States

Because of the high collisional density within the solid foil and the opportunity for electron pickup by the beam particles at the downstream surface of the foil, the beam-foil source produces a copious population of states with multiple electron excitation, core vacancies, and core excitation. Doubly excited states are also populated in laboratory and astrophysical plasmas by the process of dielectronic recombination, i.e. the capture of a free electron and simultaneous excitation of a bound electron. Such states have a different ionization limit from that of the singly excited states, and autoionization (radiationless electron ejection) often leads to very short characteristic lifetimes for these states. There are situations, however, in which autoionization channels are either forbidden or inhibited, and radiative decay is observable.

For example, the quartet levels of the $1s2sn\ell$ and $1s2pn\ell$ ^4L systems in the Li isoelectronic sequence are normally not able to autoionize into the $1s^2\epsilon\ell$ doublet continuum by the Coulomb interaction (because of the $\Delta S = 0$ selection rule), but they can decay to lower quartet states by E1 radiative transitions. However, spin-forbidden E1 transitions to the $1s^2n\ell$ ^2L doublet

system, as well as spin-orbit-induced autoionization, are sometimes possible. An interesting case involves the $1s2p^2\ ^4P_J$ levels, in which a weak mixing of the $1s2p^2\ ^4P_{5/2}$ level with the rapidly autoionizing $1s2p^2\ ^2D_{5/2}$ level causes its lifetime to be substantially shortened relative to that of the other ^4P_J fine-structure levels. This “differential metastability” is similar to that encountered in Sect. 27.4.2 for the Be-like $2s3p\ ^3P_J$ levels, except that here the added decay channel is to the continuum rather than to the ground state.

27.5.5 QED Corrections

The Lamb shift S traditionally refers to the energy difference between the $2s\ ^2S_{1/2}$ and $2p\ ^2P_{1/2}$ levels in hydrogen, caused by quantum electrodynamic (QED) effects. It is particularly interesting to study S along an isoelectronic sequence. It took nearly 20 years before it could be experimentally determined for the doubly ionized system Li^{2+} , using an electrostatic accelerator. Nowadays experimental data exist for many systems, including hydrogenlike U^{91+} .

Lamb shift studies have also been extended to He-like atoms. For instance, the wavelengths of the $1s2s\ ^3S$ – $1s2p\ ^3P$ transitions have been accurately measured for many systems, from He to Xe^{52+} . The agreement between theory and experiment is excellent.

27.5.6 Hyperfine Quenching

An interesting aspect of the theory of highly forbidden transitions is the effect of hyperfine quenching, whereby mixing induced by interactions with the magnetic moment of the nucleus can significantly alter the lifetimes of the levels.

An illustration involving the lifetimes of the $1s2p\ ^3P_J$ levels in the helium isoelectronic sequence is shown in Fig. 27.12. With increasing Z , variously forbidden transitions to the $1s^2\ ^1S_0$ ground state have an increasingly important effect on the lifetimes of the individual fine-structure levels. At low Z all three levels have similar lifetimes, since the E1 transition to $1s2s\ ^3S_1$ dominates. With increasing Z the lifetime of the $^3P_1$ level is drastically shortened, since the $\Delta S \neq 0$ selection rule is valid only for pure relativistic LS coupling, and spin mixing opens the relativistic-E1 transition to ground. A similar but less drastic shortening of the $^3P_2$ lifetime occurs because of its M2 transition to ground. For systems with nonzero nuclear spin I , both the $^3P_2$ and the $^3P_0$ lifetimes are affected by hyperfine-induced E1 transitions to the ground state. This is particularly striking for the $^3P_0$ level, because for a spinless nucleus it would have *no* radiative decay channel to ground, and because for $J=0$ the effect is not smeared over a multiplicity of values for $F = I + J$.

As can be seen from Fig. 27.12, there is a sharp turning on of the hyperfine quenching effect at $Z=21$, and the alternating zero and nonzero I

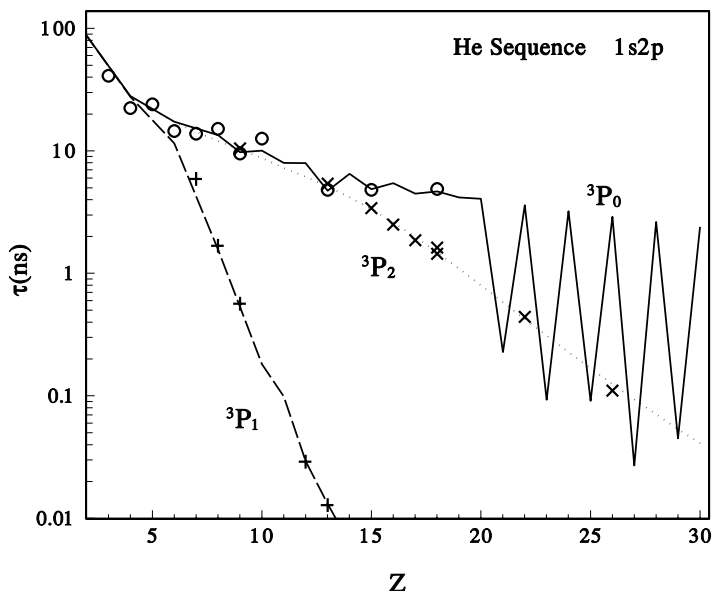


Fig. 27.12. Hyperfine quenching in the He isoelectronic sequence

values in the most abundant isotopes of nuclei of even and odd Z produce a sawtooth pattern in the lifetimes. This plot describes measurements involving only the most abundant isotope, but isotopically enriched samples are available at some accelerator facilities. Thus these studies have been extended to obtain precision measurements of the hyperfine quenching by differential comparisons of lifetimes for systems with the same Z but with isotopes of both even and odd A .

27.6 Future Challenges

While much progress has been made through the application of accelerator-based studies of atomic structure, significant problems remain that must be addressed to achieve fundamental understanding and the application-enabling goals of the field. Many examples exist, and we mention here only a sampling.

27.6.1 Calibration Standards

While great strides have been made in the precision and scope of atomic-lifetime measurements, little work has been done on the specification of branched transition probabilities in multiply charged ions. In order to convert lifetime data for levels with multiple exit channels into transition probabilities, branching fractions are needed. These require the measurement of

relative intensities of the various transitions from a given upper level. For multiply charged ions, virtually no branching-ratio measurements exist. This lack of measured data leads to a corresponding lack of tests of theoretical methods for the computation of transition probabilities in complex ions.

Progress has been made in the determination of branching fractions in neutral and singly ionized atoms through the use of absolute emission, absorption, and dispersion measurements, and through combined measurements of relative branching ratios and lifetimes. However, even this work has been restricted to the visible spectrum because of the limited availability of intensity calibration standards. No comprehensive set of calibration standards is currently available for the ultraviolet region.

The reasons for the dearth of branching-fraction data in multiply charged ions are clear. Their measurement requires an intensity calibration of the detection apparatus as a function of wavelength, which is particularly difficult for multiply charged ions and for wavelengths for which reflective and refractive optics cannot be used. Moreover, it is not sufficient to obtain branching-ratio data. Complete branching fractions which include all exit channels open to a decaying level are necessary, and these may include, in addition to the visible region, significant branching via ultraviolet and infrared transitions. While BFS methods provide a powerful means of producing highly ionized atoms, they involve Doppler broadenings and shifts, polarizations due to anisotropic excitation, time variations due to decay and repopulation, etc. Thus, the use of a lab-fixed standard lamp is not optimal for intensity-vs.-wavelength calibration of a detection system that views a moving beam. A two-step process would be preferable, using a standard lamp to calibrate relative branching ratios in a suitably chosen ion, and then using that ion as an in-beam calibration standard that can be accelerated to a velocity matching that of the desired ion.

The lack of transition probability data is particularly troublesome for the study of isoelectronic sequences. Expositions of line strength data often exhibit slowly varying, nearly linear isoelectronic variations. This is because wavelength and intermediate-coupling factors present in the transition probabilities are removed when converted to effective line strength factors, which have a simple dependence on the nuclear charge Z . In contrast, the reciprocal lifetime for a branched transition is summed over decay channels with different wavelength factors, resulting in a complicated dependence on Z .

Fortunately, isoelectronic measurements and semiempirical parametrizations can be employed to obtain calibration standards. A promising example involves systems with ground configurations ns^2np^2 such as the Si, Ge, Sn, and Pb isoelectronic sequences. The ground term can be specified by two mixing angles (entangling $^3P_2$ with $^1D_2$ and $^3P_0$ with $^1S_0$), and the ns^2npn' 's excited term can be specified in terms of a third mixing angle (entangling $^3P_1^\circ$ with $^1P_1^\circ$). If these are nearly pure configurations (with little CI), the relative transition rates are specified by these three mixing angles. This has been

tested for transitions in the neutral atoms of these sequences, for which precision measurements of the branching fractions exist, and the semiempirical calculations show striking agreement with the measured branching fractions. Moreover, tests utilizing the empirical overdetermination of the mixing angles confirm that CI effects are negligible. Ultraviolet standards can be obtained by isoelectronically extending these formulations to the multiply charged ions.

27.6.2 Coincidence Methods

Coincidence techniques provide a powerful means for measuring lifetimes. However, such methods were difficult to apply to BFS earlier because of the low detection efficiencies in the optical wavelength region, which made accidental coincidences a serious problem. As the acceleration energy (and hence the degree of ionization) is increased, the photon energies become sufficiently large to permit the use of detectors such as those designed for nuclear-physics applications.

An application of coincidence techniques to the decay of the metastable $1s2s\ ^1S_0$ level in He-like bromine is given in Fig. 27.13. Single-E1-photon decay of this level is absolutely forbidden because both the upper and the lower ($1s^2\ ^1S_0$) level have $J = 0$, so the decay occurs by the emission of either a single M1 photon or two E1 photons. Through the use of Si(Li) detectors it is possible not only to gate the coincidence, but also to measure the energies of the photons detected.

Figure 27.13 shows a plot of the intensity of the coincidence events plotted as a function of the energies of the individual photons. The diagonal line represents coincidences between the two simultaneously emitted E1 photons, which sum to the total excitation energy of 12 keV. The peak that occurs when both photons have 12 keV represents accidental coincidences between two unrelated M1 photons. The two rows that occur at a single photon energy of 12 keV represent accidental coincidences between one M1 and one continuum photon. Such applications of coincidence techniques offer many possibilities for new types of measurements of high precision and reliability.

27.7 Conclusion

Accelerator-based atomic spectroscopy has already made significant contributions to the fundamental understanding of atomic structure and its available database, and it has the potential to provide essential new data of unprecedented accuracy and scope. This work can be carried out with existing accelerator facilities, and the diversity of accelerators available is an important part of the strength of the method.

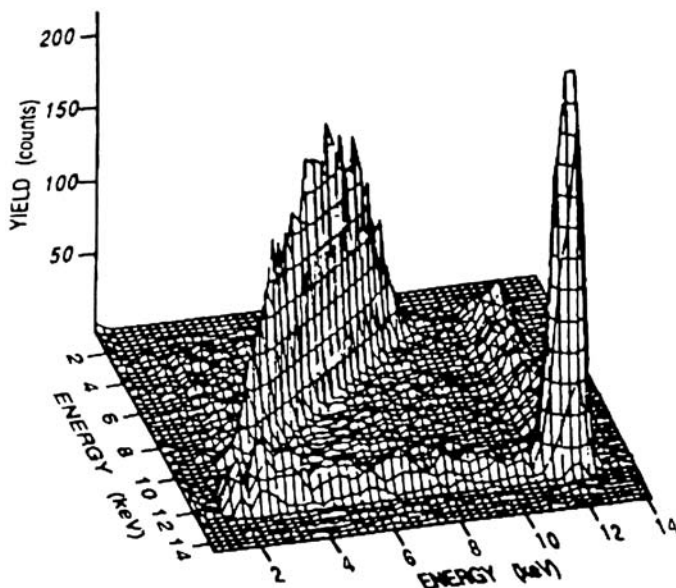


Fig. 27.13. Coincidence measurement of a two-photon decay

References

1. L. J. Curtis, *Atomic Structure and Lifetimes: a Conceptual Approach* (Cambridge University Press, Cambridge, UK 2003)
2. L. J. Curtis and I. Martinson, in *Atomic Physics with Heavy Ions*, H. F. Beyer and V. P. Shevelko, editors (Springer, Heidelberg, 1999) pp. 197–218
3. E. Träbert, in *Accelerator-Based Atomic Physics: Techniques and Applications*, S. M. Shafrath and J. C. Austin, editors (AIP Press, New York, 1998) pp. 567–600
4. E. H. Pinnington, in *Atomic, Molecular, and Optical Physics Handbook*, G. W. F. Drake, editor (AIP Press, New York, 1996) pp. 213–219

28 Industrial Electron Accelerators

M. Letournel

VIVIRAD sa, 23, rue Principale, 67117 Handschuheim, France
vivirad@aol.com

28.1 Introduction

Started in the early 1930s, and in full development during the last 30 years, electron accelerators are now useful tools in industry for many applications. This is due to the numerous capabilities of the electron beam in a wide variety of commercial applications. Several thousands of different products are processed daily throughout the world, creating energy savings and high added value. A rough idea of industrial electron accelerators (of the electrostatic and cascade types) and their use is given in this chapter.

28.2 Interaction of Electrons with Matter

The principal effect of high-velocity electrons is to produce ions in the material treated, resulting in the liberation of orbital electrons. These electrons form the chemical bonds between the atoms of most materials, whether liquids, gases or solids. Since chemical bonding energies are lower than the energy of ionization of an orbital electron, an energetic electron beam can break these bonds, producing ions, free radicals and other excited states. As a result, the original molecule is modified and the free radicals or other excited species can combine to form new molecules. The net effect is the breaking of chemical bonds with attendant chemical reactions, destruction of organisms by disorganization of their DNA chains, or dislocation of atom migration.

28.3 Industrial Applications of Electron Beam Processing

Electron beam processing is now utilized by many major industries, including the plastics, automotive, rubber goods, petrochemical, wire and cable, electrical-insulation, textile, semiconductor, gem, medical, packaging and pollution control industries. A partial listing of proven industrial applications for electron beams includes:

- crosslinking of wire and cable insulation
- disinfection of sewage sludge and waste water
- power plant pollution control (DeSO_x, DeNO_x)
- modification of bulk polymers
- rubber vulcanization
- cellulose depolymerization
- crosslinking of plastic film, foam and tubing
- food preservation
- curing of coatings
- textile modification
- graft polymerization on film, fiber and paper
- semiconductor modification
- sterilization of medical supplies.

The net effect of electron treatment is that physical properties, chemical structures and biological structures can be changed. For example, the molecular weight of bulk polymers can be tailored by polymerization or depolymerization with electron beams. Also, the creation of free-radical sites by electron treatment provides grafting possibilities in the production of copolymers and converted materials.

Disinfection of sludge, pasteurization of foods, and sterilization of surgical supplies, pharmaceutical and packaging containers can all be achieved by the destruction or inhibition of microorganisms by electron treatment. Electrons have the advantage of preserving the nutrient value inherent in food and waste because electron treatment does not cause significant heating of the processed material. Disinfected sludge can be used for its nutrient value as fertilizer for soil and ocean enrichment. Animal feed can be disinfected so that harmful microorganisms are not passed on in the food chain, while maintaining the nutrient value of the feed.

Polymer and elastomer crosslinking by electron beams results in products which are more resistant to heat, stress and environmental decay. In the case of polyolefin film and tubing, electron beam treatment also induces an elastic memory, allowing uses such as shrink-packaging material and heat-shrinkable electrical insulation.

Curing of coatings and adhesives on wood, metals and polymers with electron beam processing eliminates the need for chemical catalysts or solvents and eliminates the pollution and occupational-safety problems associated with corrosive or toxic agents. In addition, the use of heat in conventional curing ovens has become cost-prohibitive and is an uneconomic utilization of a declining resource. Electron treatment is an energy-efficient alternative. Electron beam technology, throughout the scope of industrial applications, offers significant benefits on the basis of speed, convenience, energy economy and the fine control of the process or product.

The potential of electron beam processing can literally make the world food supply feed more people, turn waste into a life-cycle resource, eliminate

sources of environmental pollution and save energy in getting consumer goods to the marketplace. In Table 28.1, the best of the main industrial applications of electron beam processing are summarized.

Table 28.1. Industrial applications of electron beam processing

Industries	Processes	Products
Chemical Petrochemical	Crosslinking	Polyethylene
	Depolymerization	Polypropylene
	Grafting	Copolymers
	Polymerization	Lubricants Alcohol
Coatings Adhesives	Curing	Adhesive tapes
	Grafting	Coated paper products
	Polymerization	Veneered panels
		Thermal barriers Wood/plastic composites
Electrical	Crosslinking	Building, instrument and telephone wires
	Heat-shrink memory	Power cables, insulation tapes
	Semiconductor modification	Shielded cable splices
		Zener diodes, ICs, SCRs, IGBTs
Health Pharmaceutical	Sterilization	Medical disposables
	Polymer modification	Powders and ointments
		Ethical drugs
		Membranes
Plastics Polymers	Crosslinking	Food shrink-wrap
	Foaming	Gymnastic mats, toys
	Heat-shrink memory	Plastic tubing and pipes
		Molded packaging forms
		Flexible packaging laminates
Pollution control	Disinfection	Agricultural fertilizers from sewage sludge
	Precipitation	Safe stack gas emissions
	Monomer entrapment	Safe sludge disposal
	DeSOx/DeNOx	Ocean-life nutrients from sludge
		OSHA and EPA compliances
		Worker safety
Pulp Textiles	Depolymerization	Rayon, permanent-press textiles
	Grafting	Soil release textiles
	Curing	Flocked and printed fabrics
Rubber	Vulcanization	Tire components
	Green strength	Battery separators
	Graded cure	Roofing membranes

28.4 Process

- a. *Features.* An electron beam processing system should offer in-line production capacities that meet the requirements of modern industrial practice. The electron beams are directed at the product and can be spread uniformly over the product area, with minimum energy waste.
- b. *Parameters.* Four major factors require great consideration: dose, penetration, efficiency and production capacity.

28.4.1 Dose

The average dose is the total amount of energy absorbed by a material divided by its mass. The applications of electron beam processing are numerous, and the selection of a proper dose level is important for each. For instance, the treatment of potatoes to prevent sprouting may need only 0.001 of the dose needed for the improvement of the temperature resistance of polyethylene. The dose required for a given process, then, is a major determining factor in establishing processing costs and in selecting equipment power (Fig. 28.1). The unit of dose is the gray, defined as an energy absorption of 1 joule per kilogram of material. Absorbed dose can be transposed into more common industrial terms associated with heat and electrical energy.

28.4.2 Penetration

The penetration is a function of the thickness and density of the product and the energy of the electron beam. The thickness of material that can be penetrated by a high-energy beam is directly proportional to the energy of the beam and inversely proportional to the density of the material being treated. Thus, for a given electron energy, the penetration may be expressed in terms of the weight of material treated per unit area. For this reason, electron beam penetration is commonly expressed in g/cm^2 . The electron beam energies of particular interest for processing applications range from 300 kV to 10 MV. A 1 MV electron beam has a maximum penetration of about 5 mm in water. Figure 28.2 shows the relative range of the electron beam as a function of voltage and its relative dose as a function of depth. The curves in Fig. 28.2 demonstrate two characteristics of electron beams. First, at the surface, the dose is about 60% of the maximum or greater, it builds up to a maximum within the thickness of the material and then it declines. Second, the depth at which the ionization does not fall below 60% of the maximum is between $1/2$ and $3/4$ of the total range. This is often referred to as the “optimum thickness” for electron beam penetration. At this optimum thickness, the best balance is obtained between minimizing overdoses and maximizing the beam’s penetration capability.

If a minimum dose must be delivered to all parts of the product, the optimum thickness – where the dose does not fall below 60% of the maximum –

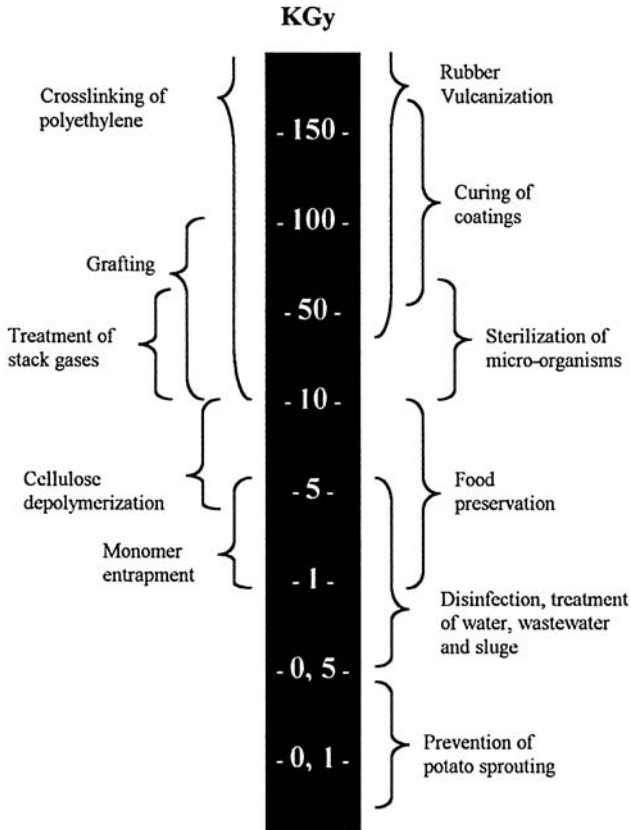


Fig. 28.1. Dose ranges for various applications

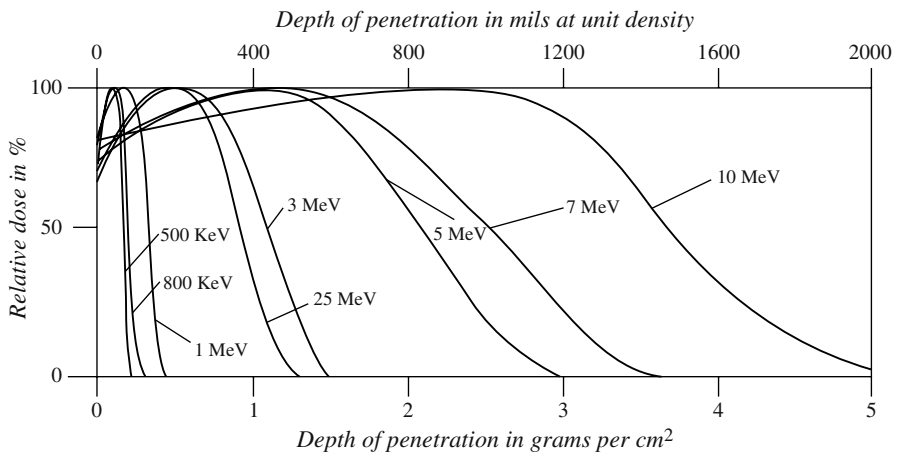


Fig. 28.2. Ranges of electrons for different energies

can be increased by about 2.4 times when electron beam processing is done from both sides of the product – a technique called double-sided treatment. In the case of a solid material, the product can be turned over after exposure on one side and passed through the beam a second time. With liquid or powder products, which could not be “turned over” without internal mixing, simultaneous cross-firing with two electron beams is required. In either case, the effective penetration is more than doubled because the doses obtained from the opposed beams are additive, as shown in the penetration curves for double-sided treatment (Fig. 28.3).

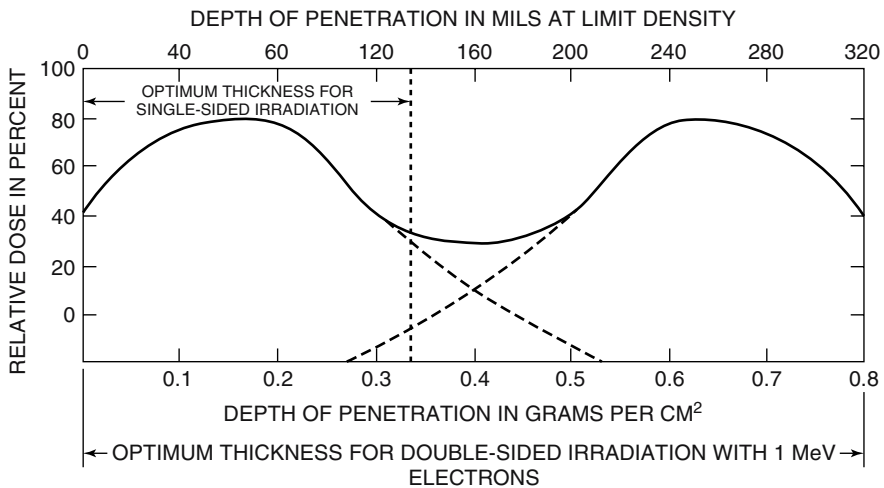


Fig. 28.3. Penetration curves for double sided treatment

28.4.3 Process Efficiency – Production Capacity

The size and shape of the product and the presentation technique influence the efficiency with which the electron beam can be used. For maximum efficiency and minimum cost in an electron beam process, it is necessary that as much of the electron beam power be absorbed in the product and the dose level in all parts of the product be as uniform as possible, with little overdosage or wasted energy.

For maximum efficiency, two factors are of importance:

- first, the uniform distribution of the beam over the product
- second, the absorption of the beam through the product thickness.

There is a dose variation with depth inherent in the process that should be taken in account, and estimated. Thus an estimate of the production capacity for a given situation can be obtained through the formulation

$$\text{Capacity(kg/h)} = 3600 \frac{\text{Power(kW)}}{\text{Dose(kGy)}} \frac{\text{Absorption efficiency in \%}}{100}$$

and similarly when the product is a film.

28.5 Technology

The main parts of an electron beam processing system includes firstly the accelerator, and secondly the scanner and a system adapted to handle the products to be irradiated. On the accelerator side, there exist several ways to obtain the high voltage. For the beam, the electron gun and the accelerator tube are similar to those in a research accelerator. The electron gun is usually a tungsten filament adapted in to a cathode. The optics for the electron beam are similar to ion optics. The delivered beam could be from some 10 μA up to 100 mA or more. Its size, ranging from some mm to tens of mm, does not normally need any adjustment for focusing. In most applications, the electron beam is swept in a scanner chamber in two directions at frequencies of the order of 200 and 1000 Hz, over a thin titanium window of 25 to 50 μm thickness.

The electrons passing through the scanner window then hit the products, which are conveyed at a certain speed under the scanner. Multiple types of conveyors have been adapted to different products, such as cables, film, heat-shrinkable materials and tubing. Depending on the required dose, their speed is automatically linked to the electron beam through a computer. Among the multiple applications, there are some voltages more specifically suitable for each product, as shown in Table 28.2, with a range up to 5 MV.

28.6 Different Types of Electron Accelerators

As mentioned above, the accelerators used differ mainly in the type of their power supply.

28.6.1 Electrostatic Accelerators

In many laboratories, electron beam irradiation projects have started with the use of an electrostatic accelerator. Owing to the voltage easily obtained from this type of accelerator, several of these are still in use. The limited beam current depends on the belt, which allows a beam of up to 1 mA for production in industry, usually at 3 or 4 MeV, and currently beams of 250 μA obtained with smaller 2.5 MV machines are mainly used for producing X-rays on a tungsten target. The X-ray application is broadly utilized for flow detection in welding and radiographic inspection of various products, and even for containers in airports and harbors (Fig. 28.4).

Table 28.2. Usable energy ranges for electron beam applications (MeV)

	0.1–0.3	0.3–0.8	0.8–3	3–5
<i>Wire and cable</i>				
Insulated wire	*	*	*	
Multiconductor cable			*	*
<i>Plastic products</i>				
Food wrapping	*	*		
Electrical tubing		*	*	
Shrinkable tapes	*	*	*	
Molded encapsulation			*	*
Irrigation tubing		*	*	
Plastic foam		*	*	
Molded gaskets			*	
Hot-water pipes			*	*
<i>Solvent-free coatings</i>				
Metal finishing	*			
Wood finishing	*			
Textile finishing	*	*		
Printable surfaces	*			
Magnetic tapes	*	*		
Adhesive tapes	*			
Perselective films	*	*		
Metallized films	*	*		
<i>Rubber products</i>				
Automobile tires		*	*	*
Roofing sheet		*	*	
Elastic bands		*	*	
Conveyor belts			*	*
Industrial hose			*	*
Profile gaskets		*	*	
Rubberized fabrics	*	*		
Weatherstripping			*	*
Latex products	*	*		
<i>Bulk chemicals</i>				
Polyethylene crosslinking		*	*	*
Polypropylene scission			*	*
Teflon degradation			*	*
Rayon pulp aging		*	*	*
Cellulose hydrolysis			*	*
Butyl rubber degradation			*	*
Crude-oil cracking			*	*
High-purity polymers			*	*
Reinforced plastics		*	*	*
Impregnated wood			*	*

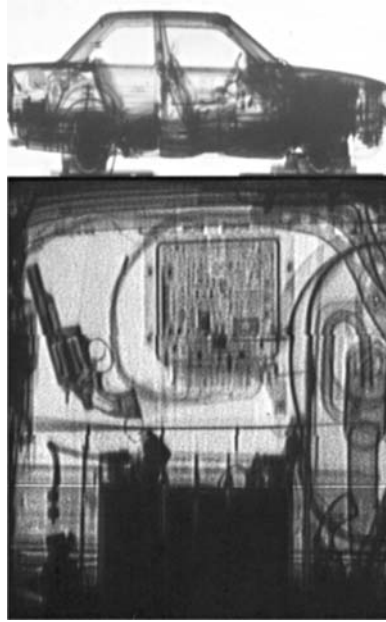


Fig. 28.4. Radiography inspection

28.6.2 Cascade Accelerators

Cascade accelerators of various types (asymmetric (Cockcroft–Walton), symmetric, parallel (Dynamitron), and isolation core transformers (ICTs)) are commonly used for irradiation. These devices are described in Box 3. For the ICT (see Fig. 28.5), the most used version has a three-phase transformer, which operates at 50 or 60 Hz directly from the mains power. The standard power for such a device is between 50 and 200 kW but in principle there is no limitation on power. The high voltage can reach 5 MV with an efficiency of around 85%.

28.6.3 Accelerator Technology

All types of accelerator tubes for electrons are in principle similar and not fundamentally too different from the technology of positive-accelerator tubes. They are constructed from ring insulators placed between highly polished metal electrodes. A uniform voltage gradient along the tube is achieved by column resistors that connect to each of the tube electrodes. The electron beam is focused and accelerated within the evacuated tube, reaching an energy corresponding to the output voltage of the power supply. Usually there is no need for special focusing.

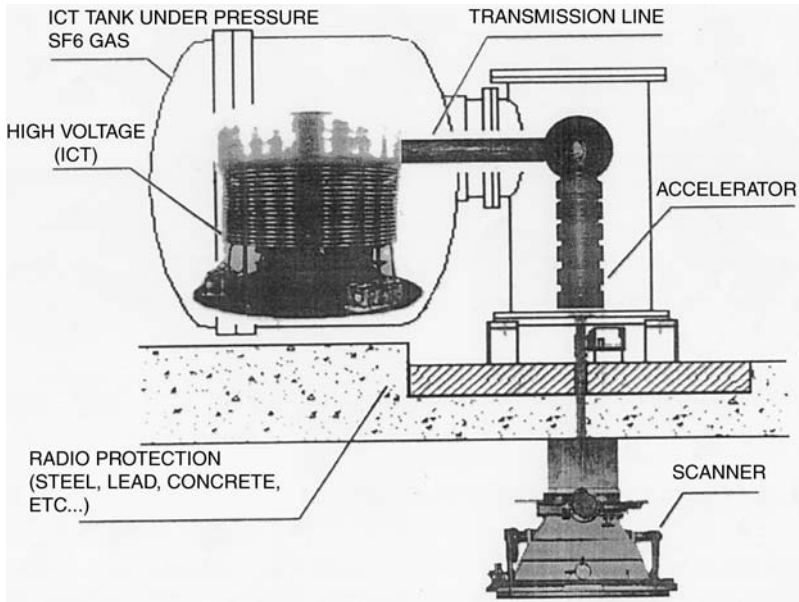


Fig. 28.5. ICT for industrial applications

28.6.4 Arrangement

Various arrangements are used for the complete electron beam system. Usually the accelerator tube and power supply are either together in the same tank under SF_6 pressure, with the tube coaxial with the generator, or separated in different tanks and connected through a transmission line. In one specific arrangement, one or two accelerator tubes and the corresponding scanners are connected to one power supply (Fig. 28.6). For voltages up to 900 kV, one power supply can be connected to one, two or three accelerators through special DC HV cables, to allow more flexibility in the process.

28.6.5 Scanner

After emerging from the accelerator tube, the electron beam enters a stainless steel scanner, where it is subject to an oscillating magnetic field, which causes the beam to be rapidly scanned through a certain angle. This angle can normally be varied up to $\pm 30^\circ$. The scan width is automatically held constant. The scanned beam passes through a thin titanium window and bombards the material being processed. Low pressure is maintained in the accelerator tube and scanner by suitable pumping with ion pumps or turbomolecular pumps.

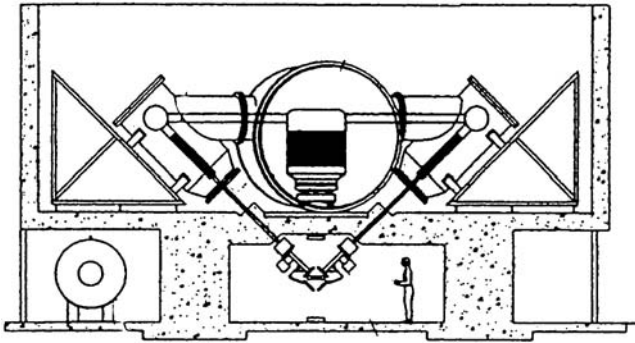


Fig. 28.6. Two scanners on one ICT power supply

28.6.6 The System Under the Beam – Conveyors

Crosslinking of polymers is one of the main applications of electron beams. Various materials such as cables, tubing, films and also heat-shrinkable materials are processed by various techniques and more or less specialized conveyors, as seen in Fig. 28.7.

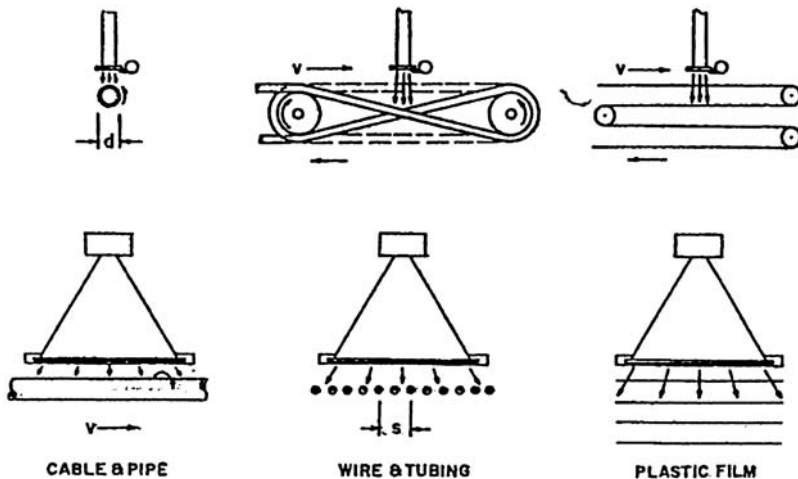


Fig. 28.7. Electron beam processing techniques

An important future application is emerging in the control of environmental pollution by treatment of liquid effluents, waste streams and stack gas, where numerous toxic substances can be removed in an economic process. Many stations are already processing effluents on an experimental basis. In contrast, some flue-gas-cleaning plants using an electron beam process are already operating on an industrial basis in Russia, Japan and China (Fig. 28.8).



Fig. 28.8. A flue-gas-cleaning plant in Chengdu, China (Courtesy of IAEA, publication TECDOC-1386)

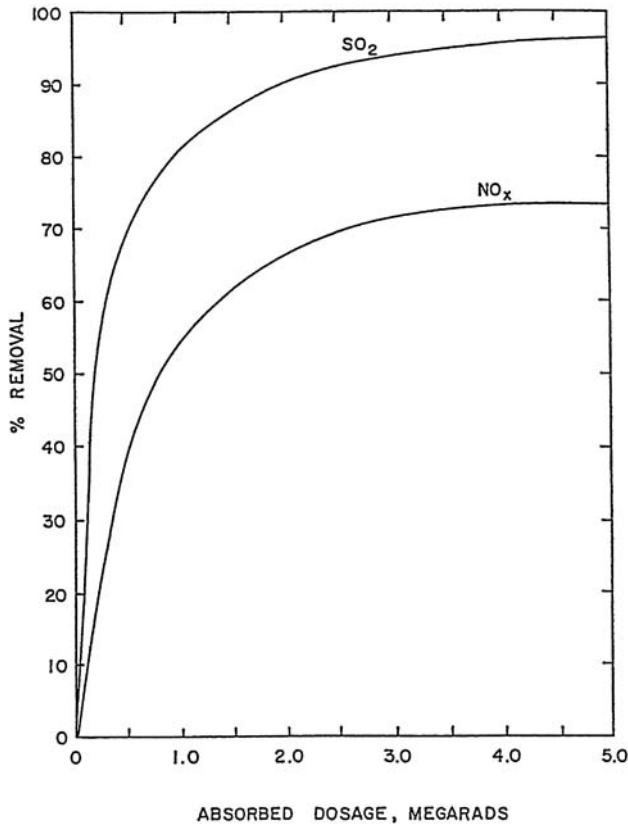


Fig. 28.9. Removal of SO₂ and NO_x as a function of irradiation dose

For example, in China, where coal is the main contributor to thermal power plants, and millions of tons of SO_2 are emitted yearly, a dedicated effort is being made in order to remove simultaneously, by electron beam processes, SO_2 and NO_x , which are the origin of acid rain. The removal of SO_2 and NO_x reaches close to 90 and 65%, respectively, as seen in Fig. 28.9.

It should also be mentioned that in a few cases, sterilization of medical products and food is being performed with an electron beam from electrostatic accelerators.

A Appendix: Electrostatic Accelerators – Production and Distribution

H.R.McK. Hyder¹ and R. Hellborg²

¹ Department of Physics, Oxford University, Denys Wilkinson Building, Keble
Road, Oxford OX1 3RH, England
`r.hyder1@physics.ox.ac.uk`

² Department of Physics, Lund University, Sölvegatan 14, 223 62 Lund, Sweden
`ragnar.hellborg@nuclear.lu.se`

A.1 Invention and Early Development

Van de Graaff's demonstration of a reliable electrostatic generator, capable of 1 MV and with the necessary stability and charging current to act as a particle accelerator, occurred as the need for such a tool was being recognized in nuclear-physics laboratories world wide. The first accelerator-based experiments were, of course, carried out by Cockcroft and Walton using a high-voltage cascade generator, but the prospect of voltages in excess of 1 MV from Van de Graaff's belt generator encouraged him and others to build improved machines in university laboratories and in national and industrial research institutions. A record of these early developments can be found in Bromley's 1974 review. From 1932 until 1946, if you wanted an electrostatic generator you built it yourself. More projects were started than came to fruition; even successful projects were not always recorded in accessible publications, and any list of these endeavors must inevitably be incomplete and inaccurate. Details of some of the more important of these early accelerators are given in Table A.1.

A.2 The War Years: 1939–1945

With a few exceptions, the outbreak of war in 1939 brought accelerator development to a halt. Military research claimed the attention of many who had been developing accelerators before 1940. Herb, Cockcroft and Trump were among those drafted to work on radar. While Herb worked on radar, his accelerators were taken to Los Alamos to provide cross section data. At MIT, small Van de Graaffs were developed to generate high-energy X-rays for examining armor plate and torpedoes. In other laboratories, existing accelerators were pressed into use to provide cross section data, but few resources were available for development and construction of new machines.

Table A.1. Early electrostatic accelerators

Institution	Location	Designer	Year	Voltage (MV)	Insulation	Beam	Layout	Notes
Princeton University	Princeton, NJ	Van de Graaff	1931	1.5	Air	None	Vertical	Positive and negative terminals, no tube
DTM	Washington, DC	Tuve et al.	1932	1.2	Air al fresco	None	Vertical	Breit tube, no ion source
DTM	Washington, DC	Tuve et al.	1933	0.6	Air	p, d	Vertical	First experimental use
MIT	Round Hill, MA	Van de Graaff	1935	+2.4 and -2.7	Air	p	Vertical	Horizontal tube between terminals
University of Wisconsin	Madison, WI	Herb	1934	0.4	Air, 0.4 MPa	p	Horizontal	First pressurized machine
University of Wisconsin	Madison, WI	Herb	1936	2.4	Air, CCl ₄ , 0.6 MPa	p	Horizontal	Hoops round column
University of Wisconsin	Madison, WI	Herb, McKibben	1940	4.5	N ₂ , CCl ₄	p, d	Horizontal	2 intershields

DTM, Department of Terrestrial Magnetism; MIT, Massachusetts Institute of Technology.
Van de Graaff’s first machine was moved to the DTM and equipped with a Breit-type accelerator tube between two terminals.
The Round Hill machine was eventually moved to MIT and reassembled with a single column and a vertical accelerator tube.
The Wisconsin 4.5 MV machine (“Long Tank”) was moved to Los Alamos and held the voltage record for ten years.

A.3 Commercial Production After 1945

Commercial production of DC accelerators started in the late 1930s with the series of Cockcroft–Walton machines built by Philips in Eindhoven. Production of these machines continued for some years after 1945. During the occupation of France in World War II, Felici in Toulouse developed cylindrical high-voltage generators operating in compressed hydrogen. After the war, machines of this type, manufactured by SAMES and capable of supplying currents of $100\mu\text{A}$ or more at voltages up to 1 MV, were widely used until overtaken by improvements in solid-state power supplies and by stricter regulations about the use of compressed hydrogen.

In Switzerland, Hafely developed cascade voltage generators, both air-insulated and pressurized, for industrial use and for such scientific applications as electron microscopes and synchrotron injectors. After the end of the war, an increasing demand for industrial and medical X-ray generators and for neutron sources led Van de Graaff and his colleagues to set up the High Voltage Engineering Corporation in 1946. Electron and ion accelerators with energies ranging from 0.4 to 5.5 MeV went into production and demand was such that in 1958 a European subsidiary, High Voltage Engineering Europa, began operation in the Netherlands. True to their origins, HVEC and its associated companies offered belt-charged electrostatic accelerators for most applications, supplemented by insulated-core transformer power supplies for low-voltage electron beams. Production of tandem accelerators began in 1958 and the first 6 MV EN was delivered to Chalk River Nuclear Laboratories in 1959. Over the next 30 years more than 60 belt-charged tandems were made, with terminal voltages ranging from 1 to 22 MV.

In the USSR, production of a range of belt-charged accelerators, both single-ended and tandem, began in 1955 at the Efremov Electrophysical Research Institute in Leningrad. Single-ended machines with voltages up to 5 MV and a vertical tandem rated at 5–6 MV were designed and supplied within the USSR and exported to Finland and China, and elsewhere.

In 1958 Radiation Dynamics Incorporated began to manufacture high-current accelerators, using the parallel-fed cascade generator developed by Cleland. Initially they made both electron and ion accelerators, mainly for universities and government laboratories, including one 5 MV horizontal tandem. Since the 1970s, they have delivered 250 electron machines for industrial applications.

During this period, Herb at Wisconsin was pursuing a different strategy. He developed the Pelletron chain charging system as an alternative to the insulating belt and emphasized the importance of ultrahigh vacuum in the accelerator tube. In 1964, he founded the National Electrostatics Corporation and began construction of a vertical 8 MV tandem for the University of São Paulo. Subsequently NEC has developed a range of small vertical and horizontal accelerators for analytical and research use and has constructed a small number of very high-voltage vertical tandems for nuclear physics,

including the 25 MV machine at Oak Ridge, which holds the world record for operating voltage.

In 1978 Purser, at General Ionex Corporation in Massachusetts, began to make small horizontal tandems for research and analysis, using the parallel-fed cascade generator invented by Cleland. Under the trade names Tandetron and Singletron, machines based on these solid-state voltage generators are now made by High Voltage Engineering Europa with voltages ranging from 1 to 5 MV.

In 1984 Letournel in Strasbourg set up VIVIRAD to manufacture high-current electron accelerators for industrial use. The lower-voltage models use insulating-core transformer power supplies; belt charging has been retained for voltages above 1 MV.

Records kept by some of these companies enable the numbers, voltages and locations of their products to be compiled with reasonable confidence. However, lack of information about subsequent shutdowns and transfers, and reasons of security and commercial considerations (which exclude some machines from published lists) mean that the tables are inevitably incomplete. Subject to these reservations, lists of research-oriented electrostatic accelerators, grouped by country, age and voltage, are given in Tables A.2, A.3 and A.4. These lists include a selection of home-made accelerators. In some instances the destination country or the voltage is not known. Consequently, the total numbers vary from table to table.

A.4 Noncommercial Developments After 1940

Construction of electrostatic accelerators by noncommercial bodies, mainly universities and government agencies, did not cease in 1946. In many cases, foreign exchange difficulties or shortage of American dollars prompted institutions in Europe and elsewhere to build accelerators similar in design and specification to machines available, at a price, from the American suppliers. In other cases, the desire to develop indigenous accelerator technology led to the formation of design and production teams that might lack experience, but were not under the same constraints of time and expense as the commercial companies. Finally, innovative ideas were not confined to the industrial design teams, and some users wanted machines that went beyond what was specified in the catalogues.

Many small Van de Graaff accelerators of conventional design, some pressurized, some air-insulated, were built in university laboratories in the 1950s and 1960s in support of local research and to provide experience in nuclear techniques for students. Records of these machines are sparse, often confined to internal reports, and most are no longer operating. No attempt has been made to compile a list of them. Some examples of larger projects are listed below. These include machines whose specifications equaled or exceeded what

Table A.2. Distribution of electrostatic accelerators by country

Continent	Number	Continent	Number	Continent	Number
<i>Africa</i>		<i>Europe</i>		<i>Middle East</i>	
Algeria	2	Austria	6	Iran	1
Egypt	2	Belarus	2	Israel	4
Mozambique	1	Belgium	10	Lebanon	1
South Africa	4	Croatia	1	Saudi Arabia	1
TOTAL	9	Czech Republic	1	Turkey	1
		Denmark	8	TOTAL	8
		Finland	2		
<i>Asia</i>		France	48	<i>North America</i>	
Bangladesh	1	Germany	80	Canada	23
China	18	Greece	3	Mexico	4
India	10	Hungary	2	USA	405
Japan	59	Italy	24	TOTAL	432
Korea	4	Netherlands	17		
Siberia	4	Norway	3	<i>South America</i>	
Singapore	1	Poland	4	Argentina	2
Taiwan	4	Portugal	1	Brazil	5
TOTAL	101	Romania	1	TOTAL	7
		Russia	25		
<i>Australasia</i>		Slovenia	1		
Australia	12	Spain	3		
New Zealand	2	Sweden	13		
TOTAL	14	Switzerland	5		
		Ukraine	2		
		UK	61		
		TOTAL	323		

was commercially available at the time, as well as those with innovative designs. The technical reports published by the builders of these machines, especially those from Debrecen, Daresbury and Kyushu, are important contributions to electrostatic-accelerator technology, much of which would still lurk behind the veils of commercial security in their absence. The choice of projects in the list below is arbitrary. Low-voltage accelerators, cascade generators, disk generators and dust generators have generally been excluded. No technical judgment is to be inferred from absence from this list.

A.4.1 Selected Noncommercial Accelerator Projects

Canada

- (i) AECL, Chalk River: “Cambridge” design 4 MV vertical Van de Graaff

Table A.3. Distribution by year of manufacture

Period	Number
Pre-1935	6
1936–1940	7
1941–1945	3
1946–1950	22
1951–1955	61
1956–1960	167
1961–1965	203
1966–1970	130
1971–1975	36
1976–1980	12
1981–1985	21
1986–1990	42
1991–1995	44
1996–2000	40
2001–2004 (part)	27
Total	821

Table A.4. Distribution by voltage and manufacturer

Voltage (MV)	HVEC	HVEE	NEC	Other	Total
0.11–0.50	79	0	6	3	88
0.51–1.00	72	23	22	10	127
1.01–2.00	151	41	55	19	266
2.01–3.00	59	25	32	17	133
3.01–4.00	24	0	6	18	48
4.01–5.00	1	1	7	15	24
5.01–7.50	69	3	1	9	82
7.51–10.00	2	0	3	5	10
10.01–15.00	13	0	5	0	18
15.01–20.00	0	0	2	1	3
>20.00	1	0	1	1	3
Total	471	93	140	98	802

China

- (i) Lanzhou: folded tandem
- (ii) Academia Sinica, Shanghai: 6 MV vertical Laddertron Tandem (Lai)

France

- (i) CEA, Saclay: 5 MV vertical Van de Graaff with liner stabilization (Winter)
- (ii) CNRS, Gif-sur-Yvette: 2 MV horizontal tandem “Aramis” (Chaumont)
- (iii) IReS, Strasbourg: 20 MV horizontal Van de Graaff tandem with multiple intershields and radial insulator posts (Letournel)

Germany

- (i) MPI, Mainz: 6 MV vertical Van de Graaff
- (ii) ZfK, Rossendorf: 5 MV EGP-10 vertical Van de Graaff tandem
- (iii) Siemens, Erlangen: 2.5 MV electron accelerator

Hungary

- (i) KFKI, Budapest: 5 MV vertical Van de Graaff with magnetic tube suppression (Kostka)
- (ii) ATOMKI, Debrecen: 5 MV vertical Van de Graaff with innovative electrostatic design and electrostatic tube suppression (Koltay)

India

- (i) Bhabha Atomic Research Centre, Trombay: 7 MV folded tandem (Singh)

Italy

- (i) CISE, Milan: 3.5 MV vertical Van de Graaff (Iori)
- (ii) CISE, Milan: 4 MV vertical tandem Van de Graaff (Caruso)

Japan

- (i) Kyushu University, Fukuoka: 7 MV vertical pellet-chain accelerator (Isoya)
- (ii) Kyushu University, Fukuoka: 10 MV horizontal pellet-chain accelerator (Isoya)
- (iii) Kyushu University, Fukuoka: 1 MV disk generator for ion implantation (Isoya)

Netherlands

- (i) Groningen University: 5 MV vertical Van de Graaff (Boerma)

United Kingdom

- (i) AERE Harwell: "Cambridge" design 4 MV Van de Graaff (W.D. Allen)
- (ii) AEI Research Laboratory, Aldermaston: "Cambridge" design 4 MV Van de Graaff, incorporating microwave terminal control (Chick)
- (iii) Cavendish Laboratory, Cambridge: "Cambridge (Mass.)" design 4 MV Van de Graaff (Shire)
- (iv) AERE Harwell: 7 MV vertical tandem Van de Graaff (W.D. Allen and K.W. Allen)
- (v) AWRE Aldermaston: 7 MV vertical tandem Van de Graaff, identical to (iv)
- (vi) Nuclear Physics Laboratory, Oxford: 10 MV vertical bipolar Van de Graaff, coupled to HVEC EN tandem (W.D. Allen)
- (vii) Nuclear Physics Laboratory, Oxford: conversion of (vi) to 10 MV folded tandem (K.W. Allen, Hyder)
- (viii) Nuclear Physics Laboratory, Daresbury: 20–30 MV vertical Laddertron tandem with single intershield (Voss, Aitken)

USA

- (i) MIT: 4 MV vertical "Cambridge" Van de Graaff with intershields (Trump and Van de Graaff)
- (ii) Los Alamos National Laboratory: 10 MV vertical Van de Graaff "P-9" with multiple intershields and separation column (McKibben)
- (iii) MIT: 8–10 MV vertical "MIT-ONR" Van de Graaff with one intershield (Trump and Van de Graaff)

USSR

- (i) KphTi, Kharkhov: ESU-2 2 MV horizontal Van de Graaff
- (ii) Kurchatov Institute, Moscow: 3.5 MV vertical tandem
- (iii) IPPE, Obninsk: EGP-15 7.5 MV vertical tandem
- (iv) INR, Kiev, Ukraine: 7 MV vertical tandem (Vishnevsky)

Acknowledgments

The authors of this appendix acknowledge with gratitude the help of the following in compiling the tables: P. Dubbelman, J. Groot and R. Koudijs (HVEE); G.A. Norton (NEC); R. Repnow (MPI, Heidelberg); V.A. Romanov (IPPE, Obninsk); and F.F. Komarov (Minsk).

Literature

- T.W. Aitken: Nucl. Instr. Meth. A **328**, 10 (1993)
 K.W. Allen: Nature **184**, 303 (1959)
 W.D. Allen: Nucl. Instr. Meth. **55**, 61 (1967)
 G.A. Behman: Nucl. Instr. Meth. **3**, 181 (1958) and **5**, 129 (1959)
 D.O. Boerma: Nucl. Instr. Meth. **86**, 221 (1970)
 D.A. Bromley: Nucl. Instr. Meth. **122**, 1 (1974)
 E. Caruso: Report, Centro Informazioni Studi Esperienze (Segrate, Milano), Milan, CISE-N-176 (1975)
 J. Chaumont: Nucl. Instr. Meth. B **62**, 416 (1992)
 D.R. Chick: Proc. IEEE **103b**, 132 and 152 (1955)
 H.R. McK. Hyder: Nucl. Instr. Meth. A **184**, 9 (1981)
 I. Iori: Energia Nucleare **8**, 770 (1961)
 A. Isoya: Proc. First Int. Conf. Electrostatic Accelerator Technology, Daresbury, DNPL/NSF/R5, p. 89 (1973)
 E. Koltay: Proc. First Int. Conf. Electrostatic Accelerator Technology, Daresbury, DNPL/NSF/R5, p. 200 (1973)
 P. Kostka: IEEE Trans **NS-18**, 82 (1971)
 W.Q. Lai: Nucl. Instr. Meth. A **382**, 89 (1996)
 M. Letournel: IEEE Trans. **NS-30**, 2713 (1983)
 J.L. McKibben: Nucl. Instr. Meth. **122**, 81 (1974)
 H. Naylor: Nucl. Instr. Meth. **63**, 61 (1968)
 V.A. Romanov: Proc. European Particle Accelerator Conference Location – Stockholm 1998, 696
 U. Schmidt-Rohr: *Die Deutschen Teilchenbeschleuniger*, Max-Planck-Institut für Kernphysik, Heidelberg (2001)
 E.S. Shire: Br. J. Appl. Phys. Suppl. **2**, S56 (1953)
 P. Singh: Ind. J. Pure Appl. Phys. **35**, 172 (1997)
 J.G. Trump: Elec. Eng., September 1951, p. 1
 R.J. Van de Graaff: Rev. Sci. Instr. **12**, 534 (1941)
 I.N. Vishnevsky: Nucl. Instr. Meth. A **328**, 39 (1993)
 R.G.P. Voss: Nucl. Instr. Meth. A **184**, 1 (1981)
 S.D. Winter: Onde Elec. **35**, 995 (1955)
 M.H. Ye, J.P. Chen: *Electrostatic Accelerators* (in Chinese), probably State Science Publisher, Beijing (1965)

B Appendix: SI Units and Other Units

R. Hellborg

Department of Physics, Lund University, Sölvegatan 14, 223 62 Lund, Sweden
`ragnar.hellborg@nuclear.lu.se`

Throughout this book, the International system of Units (SI system) is used. This system was adopted in 1960 by the Conférence Générale des Poids et Mesures (CGPM), which can be roughly translated as “General Conference on Weights and Measures”.

In many accelerator laboratories, a broad variety of equipment, meters etc., produced in different countries and with different ages, are in use. This results in a variegated set of units – SI units and non-SI units, some of them very old – being in use. Also, a great deal of the existing literature in physics and technology has been expressed in terms of older systems. It is thus necessary to understand the relationships between SI and these systems if the literature is to be fully utilized. The presentation in this Appendix is of course not intended to be a complete review of these systems, its only purpose is to provide a basis for their translation into SI.

The SI system is a coherent system based on seven basic units, listed in Table B.1. In a coherent system, the derived units are expressed in terms of the base units by relations with numerical factors equal to unity. The present definitions of the various basic units are available in the literature from the International Union of Pure and Applied Physics (IUPAP).

Table B.1. SI base units

Base Quantity	SI Restricted Name	SI Symbol
Length	meter	m
Mass	kilogram	kg
Time	second	s
Electric current	ampere	A
Thermodynamic temperature	kelvin	K
Amount of substance	mole	mol
Luminous intensity	candela	cd

From these seven basic units, several coherent, derived SI units have been obtained. Specific names and symbols have been given to several of these; some of them are listed in Table B.2. SI units related to ionizing radiation are not included, as they are discussed in detail and defined in Chap. 17.

Table B.2. Derived SI units with special names

Quantity	SI Name	SI Symbol	Expression in Terms of Base Units	Expression in Terms of Other SI Units
Plane angle	radian	rad	m m^{-1}	
Solid angle	steradian	sr	$\text{m}^2 \text{m}^{-2}$	
Frequency	hertz	Hz	s^{-1}	
Force	newton	N	m kg s^{-2}	J m^{-1}
Pressure	pascal	Pa	$\text{m}^{-1} \text{kg s}^{-2}$	$\text{N m}^{-2}, \text{J m}^{-3}$
Energy, work, quantity of heat	joule	J	$\text{m}^2 \text{kg s}^{-2}$	N m
Power, radiant flux	watt	W	$\text{m}^2 \text{kg s}^{-3}$	J s^{-1}
Electric charge	coulomb	C	A s	
Electric potential difference	volt	V	$\text{m}^2 \text{kg s}^{-3} \text{A}^{-1}$	$\text{W A}^{-1}, \text{J C}^{-1}$
Capacitance	farad	F	$\text{m}^{-2} \text{kg}^{-1} \text{s}^4 \text{A}^2$	C V^{-1}
Electric resistance	ohm	Ω	$\text{m}^2 \text{kg s}^{-3} \text{A}^{-2}$	V A^{-1}
Conductance	siemens	S	$\text{m}^{-2} \text{kg}^{-1} \text{s}^3 \text{A}^2$	$\text{A V}^{-1}, \Omega^{-1}$
Magnetic flux	weber	Wb	$\text{m}^2 \text{kg s}^{-2} \text{A}^{-1}$	V s
Magnetic flux density	tesla	T	$\text{kg s}^{-2} \text{A}^{-1}$	Wb m^{-2}
Inductance	henry	H	$\text{m}^2 \text{kg s}^{-2} \text{A}^{-2}$	Wb A^{-1}

The CGPM has recognized certain units that are important and widely used, but which do not properly fall within the SI. The special names and symbols of those units that have been accepted for continuing use and the corresponding units of the SI are listed in Table B.3. Although the use of these units is acceptable, their combination with SI units to form incoherent compound units should be authorized only in limited cases.

The CGPM has also accepted a few units that must be obtained by experiment. The energy unit electronvolt is such a unit. The symbol is eV, and it is defined as $1 \text{ eV} = (e \text{ C}^{-1}) \text{ J}$. The atomic mass unit is another. The symbol

Table B.3. Commonly used non-SI units

Quantity	Name	Symbol	Definition
Plane angle	degree	$^{\circ}$	$1^{\circ} = \pi (180)^{-1} \text{ rad}$
	minute	$'$	$1' = 1^{\circ} (60)^{-1}$
	second	$''$	$1'' = 1' (60)^{-1}$
Time	minute	min	$1 \text{ min} = 60 \text{ s}$
	hour	h	$1 \text{ h} = 60 \text{ min} = 3600 \text{ s}$
	day	d	$1 \text{ d} = 24 \text{ h} = 86\,400 \text{ s}$
Volume	liter	l, L	$1 \text{ l} = 1 \text{ dm}^3 = 10^{-3} \text{ m}^3$
Mass	tonne	t	$1 \text{ t} = 1000 \text{ kg}$

Table B.4. Conversion factors between the SI system and other systems

Given	Multiply by	To obtain	Given	Multiply by	To obtain
<i>Length</i>					
in (US)	25.4	mm	ft	304.8	mm
yd	914	mm			
<i>Area</i>					
cmil	0.0005067	mm ²	in ²	645.2	mm ²
ft ²	0.09290	m ²			
<i>Volume</i>					
fl oz	29.57	cm ³	gal (US)	3.785	dm ³
in ³	16.39	cm ³	ft ³	0.02832	m ³
<i>Speed</i>					
ft per min	5.080	mm s ⁻¹			
<i>Mass</i>					
oz	28.35	g	lb	0.4536	kg
short ton	0.9072	metric ton			
<i>Density</i>					
lb ft ⁻³	16.02	kg m ⁻³			
<i>Pressure</i>					
lb in ^{-2a}	6.895	kPa	mb	100.0	Pa
mmHg	133.322	Pa	Torr	133.322	Pa
μ	0.133322	Pa	atm	1.013 × 10 ⁵	Pa
<i>Power</i>					
hp	745.7	W	erg s ⁻¹	10 ⁻⁷	W
ft lb s ⁻¹	1.356	W			
<i>Energy</i>					
eV	1.60219 × 10 ⁻¹⁹	J	erg	10 ⁻⁷	J
cal	4.1868	J	ft lb	1.356	W
<i>Force</i>					
dyn	10 ⁻⁵	N	lb	4.448	N
<i>Magnetic flux</i>					
V s	1	Wb	Mx	10 ⁻⁸	Wb
<i>Magnetic flux density</i>					
Wb m ⁻²	1	T	G	10 ⁻⁴	T

^a lb in⁻² is often abbreviated to psi.

is u, and it is defined as $1\text{ u} = m(^{12}\text{C})(12)^{-1}$. Both units are accepted for continuing use with the SI units.

Several old units belong to a group whose use may be discontinued. To this group belong the length unit angstrom, the area unit barn, the pressure units bar and torr, the quantity-of-heat unit calorie, the activity unit curie, the exposure unit röntgen, the absorbed-dose unit rad and the dose-equivalent unit rem.

Conversion factors and equivalents between the SI system and older units in the metric system, as well as nonmetric and related units, which may be useful for people in an accelerator laboratory, are to be found in Table B.4.

Index

11th of September 2001, 445

aberration, 279, 531

absorption, 429

accelerator mass spectrometry, 33, 140, 461

accelerator tube, 123–126, 128–134, 136, 138, 140, 142, 143, 145, 147–150, 276, 284–287, 289, 291, 292, 294, 296, 302

assembly, 140

breakdown, 125

conditioning, 143

design, 136

electrodes, 139

entrance/exit aperture lens, 284–286

fault diagnosis, 144

functions, 123

gluing technology, 140, 147, 148

ideal, 124

inclined-field, 133, 276, 295, 296

model, 292, 295, 296

insulators, 138

limiting gradient, 74

mechanical design, 136

model, 289, 292

operating procedure, 142

physical processes in, 124

radiation levels, 143

summary of performance, 145

vacuum, 140

accelerator tube beam optics, 132, 278, 280

aberration, 132

analytic calculation, 132

effect of suppression systems, 136

effect of thick electrodes, 133

emittance, 133

entrance lens, 132

entrance/exit aperture lens, 278

finite-element calculations, 132

inclined-field, 292

preacceleration, 132

acceptance, 306

adjusted normalization of decay curves (ANDC), method, 563–565

adsorbed gas, 129

aerial effects, 110, 118, 119

AGATA, 416

air, dielectric strength, 64

alpha (α) cluster properties, 431

alternating-gradient focusing, 19

aluminum-26, 474

Alvarez, 5

ambipolar diffusion, 514

ammonium nitrate (AN), 452

amorphization, 507, 508

amplifier, 380

analytical techniques, 530

analyzing magnet, 165

Ancore, 455

ANDC method, 563–565

aperture lens, 285–287

archaeology, 478, 547, 553

area effect (on breakdown voltage), 78

Ariel, 383

art objects, 531, 553

astigmatism, 282, 287, 288

asymptotic breakdown gradient, 84, 85, 87

atomic

branching fractions, 561, 566, 568, 570, 571, 577–579

energy levels, 560

ions, 560

- lifetimes, 562, 564–567, 570, 571, 573, 575–579
- line strength factors, 567, 568, 571–574, 578
- transition probabilities, 561, 563, 566, 570, 577, 578
- attachment
 - coefficient, 77
 - time, 77
- attenuation coefficient, 341
 - energy absorption, 342
 - energy transfer, 342
 - photons, 361
- Auger electrons, 338, 534
- automatic beam tuning, 334
- automatic inspection technologies, 445
- average energy loss, 487
- band termination, 417
- Barkas effect, 491
- barn ($= 10^{-28} \text{ m}^2$), 447
- barrier distributions, 437
- beam
 - aberration, 279
 - brightness, 225, 317, 531
 - current, 317–322, 325, 326
 - diagnostics, 317
 - envelope, 300
 - focal constraints, 288
 - loading, 175–178
 - matching, 280, 285–289
 - profile, 317, 324–326
 - profile monitor, 324, 326
 - stopper, 326
 - tandem, 285–288
 - transport, 278–294, 296, 300
 - aberration, 279
 - axes, 280
 - coupling, 278, 285, 286
 - focal constraints, 282, 283
 - matrix, 280
 - single-stage, 284
 - tandem, 285
 - waist, 280, 283, 284
- beam transport
 - tandem, 286, 287
- beam-foil spectroscopy (BFS), 560–564, 567, 568, 573, 578, 579
- bearings, 91, 95, 99, 100
- belt, 89–95, 97, 98, 101–103
 - guides, 92–95
- Berkeley, 26
- beryllium, 367
- beryllium-10, 472
- betatron, 11, 24
- BGO (bismuth germanate oxide), 452
- binary-collision approximation, 487, 488, 508
- biomedicine, 547
- biomolecules, 524
- biosensors, 524
- BK model, 493
- Bohr velocity, 184
- boron trifluoride, 351
- brachytherapy, 24
- Bragg's rule, 498
- breakdown, 77
 - gas insulation, 84, 85
 - products, 115, 121
 - voltage, 84, 85, 87
- breakup, 431, 439
- bremsstrahlung, 130
 - electron, 536
 - projectile, 536
- brightness, 225, 317, 531
- bunching, 380
- C4 explosive, 456
- cadmium, 367
- calcium-41, 476
- CAMAC, 331
- cancellation effects, 570
- capacitive pickoff (CPO), 91–93
- capacitive pickup, 160
- carbon-11, 32, 408
- carbon-14, 33, 471
- carbon buildup, 175
- carbon ions, 30
- carbon stripper foils, 182, 187–189
- cascade accelerator, 8, 104
- cascade generator circuit
 - asymmetrical, 104
 - parallel-driven, 106
 - symmetrical, 105
- CASINO, 514
- catalysts, 516
- CERN, 20
- chain, 89, 94–100

- chain scission, 522
- Chalk River, 56
- channeling, 544
 - contrast microscopy, 544
- charge, 89–94, 96–100
 - exchange, 181, 183, 231
 - exchanger, 192
 - selection, 288
 - selector, 288, 296
 - state, 166–169, 172, 173, 175–179, 181–185
 - distribution, 182
 - equilibrium, 182
- charging efficiency, 92, 97–100
- charging system, surge protection, 81
- chemical agents, 445
- Child–Langmuir relation, 223
- chiral symmetry, 420
- chlorine-36, 473
- clearing dose, 523
- clinical oncology, 551
- close encounter, 540
- closed-loop control, 334
- clumps, initiating breakdown, 130
- clustering phenomena, 419
- cobalt-60, 24
- Cockcroft–Walton accelerator, 5, 64, 105
- coherent system of units, 604
- coincidence counting
 - techniques, 579
- collector, 382
- collector screen, 91, 94
- colliding-beam system, 20
- collimator, beam, 280
- collision cascade, 499, 510
- column
 - dead section, 74
 - internal field distribution, 73
 - structure, 74
- complex materials, ion beam analysis of, 547
- Compton
 - scattering, 378
 - suppression shields, 415
- computed tomography (CT), 28
- computer control
 - response time, 332
 - system, 164
- conditioning, 167, 334
- conductance, pumping, 169, 171–174
- confined space, 369–371
- confinement time, 194
- contact band, 96, 99, 100
- controlled corona discharge, 154, 160
- controlled down charge, 160
- conveyor, 587, 588, 591
- Cooper pair, 416, 417, 421
- core polarization model, 575
- Coriolis forces, 417
- corona, 110–115
 - current, 77
 - needle assembly, 161
 - point, 153
 - points, 154, 164, 165
 - stabilization, 77, *see also* controlled corona discharge
- corrosion process, 554
- corrugated waveguide, 383
- Cosmotron, 18
- Coulomb explosion, 168, 169
- coupled-channel calculations, 431, 438
- coupling, beam, 278, 285, 286
- CPO (capacitive pickoff), 91–93
- CPU (capacitive pickup), 160
- Cranberg theory of breakdown, 130
- CREOL, University of Central Florida, 384
- crosslinking, 522, 582, 583
- cross section, 341, 447, 448, 451, 534
- crystallography, 46
- current recirculation, 384
- cyclotron, 11, 24
 - AVF, 17
- cyclotron frequency, 194
- d, T, 450
- damage peak, 515
- Daresbury, 59, 287, 418
- dead section, 169, 178
- dead-time correction, 537
- Debye length, 195
- deep inelastic, 435
- deep inelastic collisions, 441
- definition
 - dose, 526
 - electron volt, 526

- fluence, 527
- fluence rate, 527
- G value, 526
- linear energy transfer (LET), 526
- mass stopping power, 526
- Particle flux, 527
- stopping cross section, 527
- stopping force (power), 527
- delta (δ) electrons, 511
- density effect, 184
- Department of Terrestrial Magnetism, 54
- detection limit, 534
- detector, 350, 531
- deuteron, 340
- deuteron beam, 455
- device interface, 329
- dielectric constant
 - alumina, 126
 - glass, 126
- dielectric materials, 507
- differential pumping, 173
- differential-pumping tube, 142
- diffractive effects in scattering, 430
- diffusion bonding, 140
- diffusion models, 441
- dipole
 - matrix, 282
- dipole magnet, 275
- dipole radiation, 36
- direct voltage technique, 8
- discharges, high-voltage in vacuum, 128
- disinfection, 582, 583
- dispersion, beam, 279, 283, 289
- dissolution rates, 516
- distorted-wave Born approximation, 431
- distributed-feedback (DFB) lasers, 520
- divergence, 168, 172
- DLC foils, 190, 191
- dose, 584, 586, 587, 592
 - about ion track, 513
 - absorbed, 342
 - definition, 527
 - effective, 342
 - equivalent, 342
 - measurement, 351
 - neutron, 363
 - personal, 343
 - personal monitors, 352
- drift matrix, 281
- drugs, 445
- dust, 78, 92, 93
- dynamic recovery, 516
- Dynamitron, 106
- earthquake
 - protection, 372–374
 - protection system, 374
 - sensor, 375
- ECPSSR treatment of ionization cross sections, 534
- effective charge, 182, 493
- einzel lens, 276, 287
- elastic collisions, 489
- elastic-recoil detection analysis (ERDA), 540
- electrical breakdown, 77–80, 84, 85
- electrical components, surge protection, 81
- electrode material properties, 139
- electromagnetic spin–orbit coupling, 424
- electron, 338, 340, 581–591, 593
- electron affinity, 225
- electron and hole transport, 514
- electron beam lithography (EBL), 524
- electron capture, 181, 182, 184, 185
- electron cascade, 511, 513
- electron cyclotron resonance heating (ECRH), 384
- electron–hole recombination, 514
- electron loading, 167, 169
- electron loss, 181, 182, 185
- electron optics, 278, 382
- electron–phonon interactions, 507
- electron storage rings, 20
- electron suppression, 133, 177
- electron surface emission, 128, 507
- electron temperature, 193
- electron tunneling, 128
- electronic excitation, 510, 520
- electronic excitation in dielectric materials, 514
- electronic personal dosimeters, 353
- electronic stopping, 490
- electronic stopping cross section, 490

- electronic stopping in channels, 499
- electronic stopping power, 487
- electrostatic accelerator, 9, 299
 - air-insulated, 64
 - FEL (EA-FEL), 380
 - nuclear structure, 413
- electrostatic deflector, 275
- electrostatic field, 67, 84
 - distribution along surfaces, 86
 - systematic errors, estimate, 84
- electrostatic lens, axially symmetric, 312
- electrostatic mirror, 326
- electrostatic suppression, 318–320
- elemental composition, 445
- elemental content, 456
- elemental features, 446
- ellipse, phase space, 280, 283, 284
- emittance, 168, 175, 224, 300, 317
- emittance measurement, 317
- empirical scaling rule, 493
- EN tandem, 56, 116, 382, 414
- energy
 - analyzing system, 153–156
 - balance, 508
 - dispersion, 156
 - loss, 486
 - retrieval, 384
 - spread, 317
 - stored, 78
 - stragging, 495
- entrance/exit lens, tube, 132, 278, 284–288, 296
- environment, 547
- $\text{Er}_2\text{Zr}_2\text{O}_7$, 519
- ethylene-cracked foils, 189
- EUROBALL, 414
- evaporation–condensation, 188, 190
- evaporation residue, 440
- EXOAM, 427
- expansion cup, 204
- Experimental Physics and Industrial Control System (EPICS), 330
- explosive materials, 366
- explosives, 445
- exposure-age dating, 480
- extraction, 202
 - Bayly and Ward type, 214
 - beam, 223
 - efficiency, 386
 - Thonemann type, 214
- extremely high frequency, 389
- far infra-red (FIR), 384
- Faraday cup, 318–321, 323, 326
 - retractable, 321
- fast ion beams, 561
- fast-neutron analysis (FNA), 450
- FEL (free-electron laser), 378
- Feldmühle, 157
- FEM (free-electron maser), 378
- field distribution, 67
 - column, 73
 - cylindrical geometry, 67
 - hoops, 72
 - intershield, 69
 - single-ended accelerator, 66
 - spherical geometry, 69
 - terminal, 72
- film badge, 352
- filter, beam, 279, 288
- flammable materials, 366
- flat-topping, 333
- fluence, 318, 327, 340, 526
- fluorine-18, 32, 409
- FN tandem, 56, 116, 414
- focal constraints, 282
- focus, beam, 280–289
- focusing device, 531
- folded tandem, 67, 285, 288, 372
- FOM Institute, 384
- form factor, 435
- Fowler–Nordheim law, 128
- fragmentation of molecular ions, 181
- Fraunhofer diffraction, 430
- free-electron laser (FEL), 378
- free-electron maser (FEM), 378
- Frenkel pairs, 510
- fret corrosion, 99
- fringing field tube, 286
- fringing field, tube, 278
- full-width half-maximum (FWHM), 386
- fusion, 389
- fusion–fission, 440
- gamma flash, 456

- gamma rays, 338, 456
 - characteristic elemental, 445
 - multidetector systems, 414
- GAMMASPHERE, 415
- gap lens, 285, 286
- gas discharge, 197
 - arc, 197
 - glow, 197
 - high-frequency, 198
 - linear, 198
 - ring, 198
- Townsend-type, 197
- gas insulation, 84, 85
- gas or foil, 182
- gas stripping, 182, 183
- gas-filled magnet, 470
- Gd₂Ti₂O₇, 519
- Ge (germanium), 452
- General Ionex Corporation, 107
- generating voltmeter (GVM), 153, 155, 157, 158, 160, 162, 164, 468
 - amplifier, 159
- glazes, 554
- gradient bar, 93, 95
- grading bars, 74
- gray, 584
- Greinacher, 105
- gridded lens, 285, 286
- gridded windows, 404
- Group3, 331
- GSI, 30
- GVM, *see* generating voltmeter
- half-value layer, 361
- hazardous materials, 445
- hazards
 - electrical, 365
 - fire and explosion, 366
 - mechanical, 366
 - toxic, 367
- Heavy Ion Accelerator Technology
 - Conferences, 62
- Herb, Ray, 6, 89, 95
- high-resolution transmission electron microscope (HRTEM), 516
- high-temperature superconductors (HTSCs), 388
- High Voltage Engineering Corporation, 55
- high-voltage (HV) terminal, 381
- high-voltage DC accelerator, 8
- high-voltage supplies, surge protection, 82
- Hiroshima and Nagasaki, 346
- hoop design, 72
- HVEE, 108
- hydrocarbons, 170, 171, 173, 175
- hyperdeformation, 419
- hyperfine quenching, 576, 577
- Ice Man, 479
- ICRP (International Commission on Radiological Protection), 360
- idler, 99
- image point, 281, 283, 284, 288, 289
- image slit, 156, 165
- imaging, portal, 29
- immobilization of actinide-containing nuclear waste, 516
- impact parameter, 429
- inclined-field tube *see* accelerator tube, inclined-field 133
- incomplete fusion, 439
- induced radiation, 357
- inductor, 96, 97, 99, 100
- industrial applications, 549, 581–584, 588, 591
- inelastic process, 448
- inelastic scattering, 431
- infrared (IR), 380
- instrument protection, 333
- insulating gas, 75, 369, 371
 - carbon tetrachloride, 75
 - compressed air, 75
 - nitrogen/carbon dioxide, 75
 - sulfur hexafluoride *see* sulfur hexafluoride 75
- insulating-core transformer, 107
- insulators, 80
 - properties, 138
 - surface shape, 127, 139
 - tracking length, 127, 139
- intensity-modulated radiotherapy (IMRT), 29
- interacting-boson model, 418
- interaction potential, 429
- interaction quantities, 341
- interatomic potentials, 489

- International Union of Pure and Applied Physics (IUPAP), 604
- intershield, 69
 - effect on maximum voltage, 70
- interstitial, 510
- interstitials and vacancies, 516
- iodine-129, 475
- ion beam, 299
- ion beam analysis (IBA), 530
- ion beam mixing, 520
- ion metastable, 229
- ion-optical calculation, 311
- ion optics, 278, 285, 287, 299
- ion range, 508
- ion-solid interactions, 530
- ion source, 200, 274, 531
 - ANIS, 261
 - Cs-sputter, 244
 - diode, 231
 - duoPIGatron, 204
 - duoplasmatron, 200, 231
 - ECR, 216
 - external-oven, 250
 - gas field ionization (GFIS), 219
 - high frequency (RF)
 - capacitively coupled, 212
 - inductively coupled, 212
 - high-frequency (RF), 212
 - inverted sputter, 247
 - liquid-metal (LMIS), 219
 - multiple sample, 258
 - of single-ended machines, 192
 - Penning-type (PIG), 205
 - plasma-sputter, 260
 - RF plasma-sputter, 263
 - SNICs, 251
- ion spectrum, 317
- ion straggling, 510
- ion track dose model, 511
- ion trajectories, 300
- ionization
 - chamber, 469
 - cross section, 193
 - electron impact, 195
 - field, 196
 - in plasma column, 514
 - ion impact, 196
 - multiple, 195
 - surface, 196
- ionizer
 - conical, 252
 - cylindrical, 254
 - ellipsoidal, 256
 - spherical, 256
 - spiral-wound, 255
- ionoluminescence, 545
- iron-60, 478
- irradiation, 587, 589, 592
- irradiation lifetime, 187
- irradiation-induced damage, 500
 - in pyrochlores, 516
 - in SiC, 514
- Ising, Gustaf, 6
- isochronal annealing, 515
- isochronous cyclotron, 18
- isoelectronic sequences, 561, 567, 570, 578
- isospin, 421
- isospin-breaking effects, 424
- isospin mixing, 425
- Israeli FEL, 384
- Japan Atomic Energy Research Institute, 372
- K* isomer, 418
- Kapchinskiy–Vladimirskiy density distribution, 311
- kerma, 343
- Kerst, 16
- kinematic coincidence, 441
- Kobe University of Mercantile Marine, 372
- Korean FEM, 384
- LabVIEW, 330
- laddertron, 59, 97, 98, 100
- Lamb shift, 576
- Laplace's equation, analytical solution, 66
- large-angle scattering, 496
- laser, 378
- laser plasma ablation–deposition, 190
- lateral spreading, 510
- lattice disorder, 514
 - in pyrochlores, 516
 - in SiC, 514

- Lawrence, 5
- lead, 367
- leakage current, 318, 321
- LED display, 323, 324
- lens, accelerator tube, 278, 284–287
- lens matrix, 281
- LHC, 7, 21
- lifetime
 - of belt, 101, 102
 - of ion, 194
 - of ion source, 192, 198, 202, 206, 209, 214
- light ions, 30
- linac (linear accelerator), 12, 26, 378
- Lindhard–Scharff–Shött (LSS) model, 492
- linear accelerator (linac), 12, 26, 378
- liner, 163
- linewidth, 386
- Liouville’s theorem, 224
- liquid-drop model, 441
- Livingston, 6, 19
- Livingston plot, 6
- LNT (linear-no-threshold) hypothesis, 347
- local-density approximation, 492
- logarithmic amplifier, 157
- Long Tank accelerator, 10, 55, 65, 66
- Los Alamos, 55
- low-voltage arc breakdown, 128

- magic numbers, 423
- magnetic resonance imaging (MRI), 28
- magnetic rotation, 420
- magnetic spectrometer, 433, 539
- magnetic suppression, 320
- magnetostatic wiggler, 382
- manganese-53, 478
- mass asymmetries, 441
- Massachusetts Institute of Technology, 54
- Massey adiabaticity criterion, 233
- matching, beam, 280, 285–289
- material discrimination, 446
- material processing, 389
- material-specific inspection technologies, 445
- materials engineering, 506
- materials science, 526, 539

- matrix
 - accelerator tube, 284
 - beam ellipse, 283
 - beam transport, 280
 - dipole, 282
 - drift, 281
 - thin/thick lens, 281
 - transfer, 303
- maximum field, safe working value, 68
- McMillan, 5
- mean free path, 193
- mechanical fuse, 375
- medicine, 24, 25, 27, 29, 31, 33, 35
- mercury, 367
- metal oxide resistors, 116–118, 120, 121
- metastability, 576
- MeV ion implantation, 514
- microbeam, 31
- microdischarges, 130
- microparticles, 131
- microscopy, 45
- microtron, 11
- mid-column lens, 285, 286
- mineral, 550
- mineralized tissue, 552
- MINIBALL, 427
- minimum, beam, 283–285
- mirror energy differences, 424
- mirror nuclei, 424
- mm wavelengths, 380
- mode competition, 386
- modulation, beam, 289
- molecular dynamics, 487
- molecule, 168, 169, 173
- MP tandem, 57, 68
- multileaf collimator (MLC), 26, 29
- multimodal decay, 440
- multiparameter detector systems, 531
- multiphonon excitations, 438
- multiple scattering, 167, 168, 175, 176, 548
- multiply excited states, 575
- multistage depressed collector, 384

- $N = Z$ nuclei, 422
- n, γ , 448, 450, 451
- $n, n'\gamma$, 451
- NaI (sodium iodide), 452
- nanobeam, 532

- nanoscale engineering, 506
- nanoscience, 547
- National Electrostatics Corporation, 57
- negative-ion injector, 286
- negative resist, 522
- neutron, 340, 351, 362, 445
 - 14 MeV, 450
 - fast, 449
 - thermal, 445, 447
- neutron-based technologies, 445
- neutron capture, 356, 451
- neutron capture cross section, 447
- neutron flash, 456
- neutron generator, 356, 449, 450
 - electronic (ENG), 450
 - sealed, 449, 450
- neutron-proton pairing, 421
- neutron source, 448
- neutron therapy, 24
- newsprint paper, 549
- nickel-59, 478
- nickel-63, 478
- nitrogen-13, 32
- nondestructive analysis, 530
- nonintrusive inspection, 445
- Nottingham effect, 131
- nuclear displacement, 507
- nuclear material, 445
- nuclear microprobe, 531
- nuclear rainbow, 430
- nuclear reaction, 327, 445
 - $^2\text{H}(\text{d}, \text{n})^3\text{He}$, 456
- nuclear-reaction analysis (NRA), 518
- nuclear scattering, 510
- Nuclear Science Centre, New Delhi, 373
- nuclear stopping cross section, 490
- nuclear stopping power, 487
- nucleon evaporation, 435
- nucleon correlations, 432

- Oak Ridge, 58, 288, 373
- object point, 285, 286, 288
- object slit, 156, 165
- occupation probability, 434
- Occupational Safety and Health
 - Administration, 369, 370
- oil paintings, 554
- optical waveguide, 526

- optics, accelerator, 278, *see also*
 - accelerator tube beam optics, 285, 287
- organic elements (hydrogen, carbon, nitrogen, oxygen), 446
- organic resists, 522
- oscillations, Z_1 , 497
- oscillations, Z_2 , 497
- oscillator, 380
- outgassing rate, 142
- oxygen-15, 32, 407

- pair condensate, 421
- pairing interaction, 416, 431, 435
- parallel beam, 278, 281–284, 292–295
- particle elastic-scattering analysis (PESA), 540
- particle flux, 526
- particle-induced gamma-ray emission (PIGE), 542
- particle-induced X-ray emission (PIXE), 534
- parting agent, 188
- Paschen curves, 197
- peak-to-background, 534
- Pelletron, 58, 89, 95, 98–100, 306, 372, 380, 453
- permanent magnet, 275
- permittivity, 223
- perovskite-type oxides, 516
- personnel safety, 333
- perveance, 223
- phase space, 224, 279–281, 283
- phase-stabilized acceleration, 16
- phonons, 514
- photoluminescence, 524
- photon, 338, 339
- pickup, 317, 318
- pickup electrode, 317
- Pierce-type electron gun, 381
- pigments, 554
- pixel, 456
- planes, focal, 281, 284
- plant science, 551
- plasma
 - column, 514
 - density, 193
 - electron density, 193
 - flare, 128

- frequency, 194
 - electron, 194
 - ion, 194
- ion density, 193
- sheath, 195
- state, 193
- temperature, 193
- plutonium, 476
- plutonium-239, 446
- Poisson's equation, 223
- polarization effects, 423
- pollution, 581–583
- poly(methylmethacrylate) (PMMA), 513
- polyvinyl acetate, 140
- ponderomotive force, 380
- portico intershield, 59, 71
- position-sensitive detector (PSD), 565
- positive resist, 522
- positron emission tomography (PET), 28, 32, 396
- positron emission tomography combined with computed tomography (PET/CT), 28
- postaccelerator, 166, 177, 179
- potential divider, 74
- potential-drop accelerator, 8
- potential-energy surfaces, 441
- prebreakdown processes, 128
- precious artefacts, 554
- prompt α -decay, 426
- prompt emissions, 508
- prompt proton decay, 426
- proton, 340, 354
- proton beam writing (PBW), 523
- proton decay, 425
- proton storage rings, 20
- proton therapy, 26, 27, 30
- provenance, 554
- proximity exposure effect, 523
- pulse pileup, 537
- pulsed fast-neutron analysis (PFNA), 450, 453
- pulsed-neutron inspection (PNI), 450, 452
- pulsing, ns, 452
- pyrochlore materials, 516
- Q snout, 132
- quantitative analysis, 531, 537
- quantum beats, 562, 566
- quantum defects, 570
- quantum well structures, 520
- quasi-classical scattering, 511
- quasi-elastic collisions, 435
- quasi-fission, 440
- quasi-optical delivery system, 383
- Röntgen, 24
- radiation
 - damage, 516
 - dose, 531
 - hazards, 326, 327
 - ionizing, 337
 - nonionizing, 337
 - therapy, 27
- radiation cooling, 321
- radiation effects, 344, 518
 - in materials, 507
 - late
 - cancer, 345
 - hereditary, 348
 - leukemia, 347
 - pregnancy, 344
 - skin, 344
 - threshold, 345
 - whole-body, 344
- radiation field quantities, 341
- radiation protection, 337
- radiation user facility, 384
- radiative energy transmission, 389
- radioactive decay, 338
- radioactive ion beam, 414, 427
- radiocarbon calibration, 479
- radiography, 457
- radiopharmaceuticals, 31
- radiotherapy, 27
- radium, 24
- rare-earth elements, 545
- ray vector, 283
- recirculating gas stripper, 468
- recoil, 540
- recombination, 510
- recovery stages, 515
- reference particle, 279
- refractive effects in scattering, 430
- refractive index, 524
- relativistic effects, 569

- residual gas, 130, 141, 318–322, 326
 - analyzer, 143
 - ionization, 318–322, 326
- resistor, 110, 112, 115–122
 - surge protection, 82
- resonance acceleration, 11
- resonant excitation, 511
- resonator, 386
- respiratory system, 367
- RF discharge source, 274
- Righi, Augusto, 52
- RISING, 426
- rotating shaft, 169
- rotational bands, 438
- rotational motion of nuclei, 416
- Round Hill, 54, 64
- round-trip reflectivity, 386
- Rubbia, 5
- Rutherford, 4, 538
- Rutherford backscattering (RBS), 4, 538
 - spectroscopy, 518
- safety
 - administrative, 358
 - confined space, 369
 - sulfur hexafluoride, 369, 370
 - technical, 359
- saturation current, 319
- scaling, 334
- scanner, 587, 590, 591
- scattering, 496
- scattering integral, 489
- Schwinger, 5, 16
- scintillation detectors, 448
- screen, 90–94
- screening function, 502
- screening length, 493
- second stripper, 166, 176–178, 181
- secondary-electron-induced modification, 511
- secondary electrons, 125, 318–321, 324, 325, 507, 511–514
- secondary reactions, 431
- semiclassical model, 434
- sheaves, 100
- shielded resistors, 118, 119
- shorting rod, 143
- SI system, 526, 604
 - base units, 604
 - conversion factors from other systems, 606
 - derived units, 605
 - non-SI units, 605
- SiC, 514
- SiC polytypes, 514
- signatures, nuclear, 446
- silicon-32, 477
- silicon nitride window, 469
- simulation of treatment, 28
- single-particle motion of nuclei, 416
- slot aperture, 287
- slowing down, 486, 487
- $\text{Sm}_2\text{Ti}_2\text{O}_7$, 519
- small-angle scattering, 168, 496
- space charge sheath, 128
- spark, 110–112, 115–121
- spark gap, 80, 140
- spectrum
 - gamma-ray, 456
 - ion, 317
- sputtered foils, 189
- sputtering, 175, 499
- stabilization, 152–154, 157, 164, 165
- stabilization system, 153
- sterilization, 34, 582, 583, 593
- STIM, 543
- stopping force, 508, 526, 527
- stopping power, 342, 526, 527
- stopping power for a heavy ion, 493
- stored energy, 78
- strength
 - dielectric, 101, 103
 - interlayer connection, 101, 102
 - mechanical, 101, 102
- stripper, beam scattering, 133
- stripper density, 317
- stripper gas recirculator, 142
- stripping, 182
 - electron, 181
 - foil, 184, 185
 - gas, 184, 185
 - second, 166, 176–178, 181
- strontium-90, 478
- subcascades, 510
- sublattices, 515
- sulfur hexafluoride, 75

- biological effects, 369
- breakdown vs. pressure, 76
- superconducting magnets, 21
- superdeformed nuclei, 418
- superheavy elements, 441
- suppression electrode, 318–322
- suppression system, 133
 - alternating inclined electrodes, 133
 - axial-field modulation, 136
 - compressed geometry, 137
 - electron trajectories, 137
 - spirally inclined electrodes, 134
 - transverse magnetic, 134
- surface contaminants in accelerator tubes, 130
- surface tracking of tube insulators, 126
- surge damage, 80
- Swedish Work Environment Authority, 369, 370
- Symposium of North Eastern Accelerator Personnel, 62
- synchrotron, 18
- synchrotron radiation, 31
 - coherence, 39
 - facilities, 39
 - monochromators, 40–42
 - power, 37
 - spectral range, 37
- synchrotron undulator radiation, 378
- Talbot reflector, 383
- tandem accelerator, 10, 107
- tandem accelerator geometry, 67
- Tandetron, 107
- tank geometry, 67
- tank, soft elastic suspension, 375
- Tel Aviv University, 382
- tension of belt or chain, 89, 92, 93, 98–100
- terminal
 - impedance, 152, 160
 - magnet, 285
 - pumping, 142
 - shape, 72
- therapy
 - electron beam, 26
 - microbeam, 31
 - neutron, 24
 - photon activation, 31
 - proton beam, 26
- thermal neutrons, 445, 451
- thermal-neutron analysis (TNA), 450
- thermalized neutrons, 447
- Thomas–Fermi effective-charge model, 182, 184
- Thomas–Fermi velocity, 184
- Thomson scattering, 378
- threats, 445
- time of flight, 433, 450, 453, 454, 456, 470, 539, 560, 561, 564, 573
- tin-100, 423
- TL dosimeter, 352
- tomographic reconstruction, 544
- total-voltage effect, 67, 130, 148
- trajectory, particle/ray, 293, 294, 296
- transfer matrix, 303
- transient arc current, 79
- transient voltage, 127
- transport, beam, 278
- transport efficiency, 385
- transport, beam, *see* beam, transport
- triple junction, 126
- tritium, 350, 357, 477
- Trump, John, 24
- tunneling, 436
- Turin Shroud, 479
- Ubitron (undulating-beam interaction), 378
- undulator, 37, 379
- University of California Santa Barbara (UCSB), 380
- University of Hawaii, 384
- University of Tsukuba, 372, 373
- University of Wisconsin, 54
- unmanned airborne vehicles, 389
- Uppsala, 26
- upright ellipse (beam waist), 283
- uranium-235, 446
- uranium-236, 477
- vacancy, 510
- vacuum, 166, 167, 169, 170, 172, 173, 175
- vacuum breakdown, particle-induced, 130
- vacuum conductance, 141
- Van de Graaff, Robert, 6, 24, 89

- Van der Meer, 5
- vehicle explosives detection systems
 (VEDS), 453
- Veksler, 16
- ventilation, 369, 370
- very large-scale integration (VLSI)
 devices, 522
- VIVIRAD, 60, 108
- VIVITRON, 59, 288, 414
- voltage surges, calculation, 79
- volume, phase space, 279

- waist, beam, 280, 283, 284
- wave packet, 379
- waveguide, 379, 526
- weakly bound nuclei, 431
- wear resistance, 101–103
- Weizmann Institute, 382

- Widerøe, Rolf, 6
- Wien filter, 192
- wiggler, 379
- Wigner term, 422
- Wimshurst machine, 52
- wobbling mode, 420

- X-ray detector, 532
- X-ray emission, 338
- X-spectrometry, 534

- y-branch waveguide, 526
- Yale University, 57
- yrastransitions, 575

- ZBL (Ziegler–Biersack–Littmark)
 parametrization, 494