

Advances in
IMAGING and ELECTRON PHYSICS

NUMERICAL FIELD CALCULATION
FOR CHARGED PARTICLE OPTICS

Volume 116



ACADEMIC PRESS

**ADVANCES IN IMAGING AND
ELECTRON PHYSICS**

VOLUME 116

**NUMERICAL FIELD CALCULATION
FOR CHARGED PARTICLE OPTICS**

EDITOR-IN-CHIEF

PETER W. HAWKES

*CEMES—Centre National de la Recherche Scientifique
Toulouse, France*

ASSOCIATE EDITORS

BENJAMIN KAZAN

*Xerox Corporation
Palo Alto Research Center
Palo Alto, California*

TOM MULVEY

*Department of Electronic Engineering and Applied Physics
Aston University
Birmingham, United Kingdom*

Advances in Imaging and Electron Physics

Numerical Field Calculation
for Charged Particle Optics

ERWIN KASPER

*Institut für Angewandte Physik
der Universität Tübingen, Germany*

VOLUME 116



ACADEMIC PRESS

A Harcourt Science and Technology Company

San Diego San Francisco New York Boston
London Sydney Tokyo

This book is printed on acid-free paper. ∞

Copyright © 2001 by Academic Press

All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher.

The appearance of the code at the bottom of the first page of a chapter in this book indicates the Publisher's consent that copies of the chapter may be made for personal or internal use of specific clients. This consent is given on the condition, however, that the copier pay the stated per-copy fee through the Copyright Clearance Center, Inc. (222 Rosewood Drive, Danvers, Massachusetts 01923), for copying beyond that permitted by Sections 107 or 108 of the U.S. Copyright Law. This consent does not extend to other kinds of copying, such as copying for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale. Copy fees for pre-1998 chapters are as shown on the title pages: if no fee code appears on the title page, the copy fee is the same as for current chapters. 1076-5670/01 \$35.00

Explicit permission from Academic Press is not required to reproduce a maximum of two figures or tables from an Academic Press article in another scientific or research publication provided that the material has not been credited to another source and that full credit to the Academic Press article is given.

ACADEMIC PRESS

A Harcourt Science and Technology Company
525 B. St. Suite 1900, San Diego, California 92101-4495, USA
<http://www.academicpress.com>

Academic Press
Harcourt Place, 32 Jamestown Road, London NW1 7BY, UK
<http://www.academicpress.com>

International Standard Serial Number: 1076-5670
International Standard Book Number: 0-12-014758-0

Printed in the United States of America

01 02 03 04 05 QW 9 8 7 6 5 4 3 2 1

CONTENTS

PREFACE	xi
FUTURE CONTRIBUTIONS	xiii
ACKNOWLEDGMENTS	xvii
INTRODUCTION	xix

Chapter I Basic Field Equations 1

1.1 Maxwell's Equations	1
1.2 Electromagnetic Potentials	3
1.2.1 Electrostatic Fields	4
1.2.2 Vector Potentials	5
1.2.3 Magnetic Scalar Potentials	6
1.2.4 Coefficient Transformation	7
1.3 Variational Principles	8
1.3.1 Scalar Potentials	9
1.3.2 Vector Potentials	9
1.3.3 The Magnetic Energy Density	11
1.4 Wave Equations and Hertz Vectors	12
1.5 Boundary Conditions	15
1.5.1 Electric Material Conditions	16
1.5.2 Magnetic Material Conditions	17
1.6 Integral Equations for Electrostatic Fields	19
1.6.1 Dirichlet Problems	22
1.6.2 Linear Material Equations	23
1.6.3 Integral Equation for Surface Sources	24
1.7 Integral Equations for Magnetic Fields	25
1.7.1 Scalar Integral Equations	25
1.7.2 Vector Integral Equation	27
1.8 Integral Equations for Wave Fields	28
1.8.1 Dirichlet Problem	28
1.8.2 Neumann Problem	29
References	29

Chapter II Reducible Systems 31

2.1 Azimuthal Fourier-Series Expansions	31
2.1.1 Vectors Fields	34
2.2 Rotationally Symmetric Boundaries	36
2.2.1 Mathematical Form	37
2.2.2 Fourier Analysis of Boundary Conditions	38

2.3	Magnetic Round Lenses	39
2.3.1	The Flux Potential	40
2.3.2	Differential Equations	43
2.3.3	Boundary Conditions	43
2.3.4	Variational Principle	45
2.4	Series Expansions	45
2.4.1	Symmetry Conditions	46
2.4.2	Repeated z -Differentiations	46
2.4.3	Paraxial-Series Expansion	48
2.4.4	Series Expansion for the Inhomogeneous Equation	49
2.4.5	Series Expansion for the Flux Potential	49
2.4.6	Fourier-Bessel Expansions	50
2.5	Planar Fields	51
2.5.1	Cauchy-Riemann Equations and Conformal Mapping	52
2.5.2	Basic Analytical Functions	54
2.5.3	Analytic Continuation	55
	References	57

Chapter III Basic Mathematical Tools 59

3.1	Orthogonal Coordinate Systems	59
3.1.1	Line Element and Lamé Coefficients	59
3.1.2	Vectors in Curvilinear Coordinates	61
3.1.3	Differential Forms	62
3.1.4	Differential Forms of Second Order	67
3.1.5	The Surface-Adapted Coordinate System	69
3.1.6	The Discretization of Maxwell's Equations	72
3.2	Interpolation and Numerical Differentiation	74
3.2.1	Basic Rules for Interpolation	74
3.2.2	Hermite Interpolation	77
3.2.3	Hermite Splines	82
3.3	Modified Interpolation Kernels	86
3.3.1	Basic Relations	86
3.3.2	The Recurrence Algorithm	88
3.3.3	Extrapolation	92
3.3.4	Nonequidistant Intervals	94
3.4	Mathematical Representation of Curves	96
3.4.1	Differential Geometrical Functions	97
3.4.2	Determination of Sampling Arrays	98
3.4.3	Rounding-off Corners	101
3.5	Mathematical Representation of Surfaces	102
3.5.1	Rectangular Meshes	102
3.5.2	Bivariate Hermite Interpolation	104
3.5.3	Bicubic Splines	105
3.5.4	Some Remarks	107

3.6 Numerical Integration	107
3.6.1 Gauss-Legendre Quadrature	108
3.6.2 Bessel-Hermite Quadratures	109
3.6.3 Newton-Cotes Formulas and Adaptative Procedures	110
3.6.4 Euler Maclaurin Formulas	111
3.6.5 Concluding Remarks	113
References	114
 Chapter IV The Finite-Difference Method (FDM)	 115
4.1 Two-Dimensional Meshes	115
4.1.1 General Coordinate Transforms	117
4.1.2 Variational Principles	118
4.1.3 Orthogonal Meshes	121
4.1.4 Sources and Nonlinearities	124
4.1.5 Classification of Configurations	126
4.2 Five-Point Configurations	127
4.2.1 The Taylor Series Method	127
4.2.2 The Ring-Integral Method	129
4.2.3 Some Remarks	131
4.2.4 Generalization of the Method	133
4.3 Nine-Point Configurations	134
4.3.1 Approximation in One Mesh	135
4.3.2 The Complete Mesh Formula	137
4.3.3 Special Cases	141
4.3.4 The Regularization of Meshes	143
4.4 The Cylindrical Poisson Equation	145
4.4.1 The Radial Discretization	147
4.4.2 Discretization of Separable Differential Equations	149
4.4.3 Accuracy of the Discretization	153
4.4.4 The Radial Power Transform	156
4.4.5 Correction of the Functional	157
4.4.6 The Implicit Algorithm	158
4.4.7 Poisson Equation in Spherical Meshes	159
4.5 Irregular Configurations	167
4.5.1 Inner Mesh Points	167
4.5.2 Edge or Corner Singularities	170
4.5.3 Mesh Points on Boundaries of Materials	172
4.5.4 Evaluation of Series Expansions	173
4.5.5 Harmonic Functions	177
4.5.6 Applications of the General Method	181
4.5.7 Discretization Errors	184
4.6 Subdivision of Meshes	185
4.7 Concluding Remarks	189
References	190

Chapter V The Finite-Element Method (FEM)	193
5.1 Generation of Meshes	193
5.2 Discretization of the Variational Principle	200
5.3 Analysis in Triangular Elements	204
5.3.1 General Relations and Area Coordinates	204
5.3.2 Integration Over Triangular Domains	207
5.3.3 Trial Functions	209
5.3.4 Quadrilateral Elements	213
5.3.5 Differentiation in Systems of Triangles	213
5.4 The Finite-Element Method in First Order	216
5.4.1 Self-Adjoint Partial Differential Equations	216
5.4.2 Error Analysis and Improvements	220
5.4.3 Quadrilateral Meshes	223
5.4.4 The Magnetic Lens	223
5.5 Field Interpolation	229
5.5.1 Determination of the Mesh Position	230
5.5.2 Interpolation in Rectangular Meshes	232
5.5.3 Improved Hermite Interpolation	233
5.5.4 The Paraxial Interpolation	237
5.5.5 Interpolation in Trigonal Meshes	240
5.6 Solutions of Large Systems of Equations	242
5.6.1 Direct Solution Methods	242
5.6.2 The Conjugate Gradient Method	247
5.6.3 Relaxation Methods	249
5.6.4 Successive Line Overrelaxation	253
5.6.5 Nonlinear Systems of Equations	258
References	259
 Chapter VI The Boundary Element Method	 263
6.1 Discretization of Integral Equations	264
6.1.1 General Methods	264
6.1.2 Surface-Coulomb Integrals	267
6.1.3 The Far-Field Approximation	276
6.1.4 The Complete Procedure	279
6.1.5 The Normal Derivative	281
6.2 Axially Symmetric Integral Equations	284
6.2.1 Fourier Analysis of Integral Equations	284
6.2.2 Properties of the Fourier-Green Functions	288
6.3 Numerical Solution of Integral Equations	301
6.3.1 Basic Collocation Techniques	302
6.3.2 Collocation Techniques Using Splines	304
6.3.3 The Galerkin Method	308
6.3.4 A Fast Method for Symmetric Integral Equations	311

6.3.5 The Solution of Dirichlet Problems	315
6.3.6 Generalizations	317
6.4 Special Techniques for Asymmetric Integral Equations	321
6.4.1 Integral Equation for Round Lenses	322
6.4.2 Integral Equation for Deflection Systems	323
6.4.3 The Fast Method for Asymmetric Integral Equations	326
6.4.4 The Conservation of Total Lens Current	329
6.4.5 The Complete Field Calculation	331
6.5 The Calculation of External Fields	335
6.5.1 The Evaluation of Particular Integrals	335
6.5.2 Application to Rotationally Symmetric Fields	337
6.5.3 Coils with Rectangular Cross Sections	340
6.5.4 Magnetic Fields of Deflection Systems	344
6.5.5 Special Cases of Deflection Systems	349
6.6 Other Applications of Integral Equations	350
6.6.1 Planar Fields	350
6.6.2 Wave Fields	351
References	354

Chapter VII Hybrid Methods

357

7.1 Combination of the FEM with the BEM	357
7.2 Combination of the FDM with the BEM	361
7.2.1 The General Procedure	361
7.2.2 The Modified Galerkin Method	365
7.3 The Charge Simulation Method (CSM)	367
7.3.1 The General Procedure	367
7.3.2 Pointed Cathode Models	369
7.3.3 Charged Aperture Plates	377
7.3.4 Systems of Charged Aperture Plates	382
7.4 The Current Simulation Model	387
7.4.1 Magnetic Mirror Properties	387
7.4.2 Local Properties	389
7.4.3 A Simple Model for Cylindrical Coils	391
7.4.4 Generalization of the Method	393
7.4.5 Comparison with Correct Calculations	395
7.5 The General Alternation Method	397
7.5.1 Formulation of the Method	397
7.5.2 Practical Examples	400
7.5.3 Systems with Several Different Materials	404
7.5.4 Nonoverlapping Domains	407
7.6 Fast Field Calculation	408
7.6.1 Radial Interpolation	409
7.6.2 Two-Dimensional Interpolation	412

7.6.3 Three-Dimensional Interpolation	415
7.6.4 Variation of Parameters and Perturbations	417
7.7 Calculation of Equipotentials	418
7.7.1 Equipotentials in FEM Grids	418
7.7.2 Determination of Intersection Points	419
7.7.3 The General Search Algorithm	421
7.7.4 Magnetic Flux Lines	424
References	426
Appendix	429
Index	433

PREFACE

Occasionally, I am approached by an author whose work is entirely suitable for these Advances but is sufficiently long to fill an entire volume. Although such volumes must remain the exception, no valuable material is ever refused on the grounds of length, although just occasionally, the publishers recommend publication in a monograph series rather than in these Advances.

After this preamble, I am delighted to welcome the full account of methods of calculating static electric and magnetic fields by E. Kasper that forms the subject of this volume. He and his colleagues and students at the University of Tübingen have made major contributions to the theory of field calculation, and numerous specialized programs have emerged from their endeavours. Here all this work, which could only be presented in much more condensed form in *Principles of Electron Optics* by E. Kasper and myself (Academic Press, London, 1989), is set out in full detail, including of course many developments that have been made in the past decade.

The first three chapters present the basic material on which the later chapters repose: the field equations, symmetry, and mathematical tools to be used. Then come four long chapters on each of the principal methods, the finite-difference method, the finite-element method, the boundary-element method, and the hybrid methods, which often enable the user to benefit from the attractive features of more than one approach.

I have no doubt that this manual of field-calculation methods will be much appreciated and am most grateful to E. Kasper for agreeing to publish it in these Advances. A list of contributions to forthcoming volumes follows.

Peter Hawkes

This Page Intentionally Left Blank

FUTURE CONTRIBUTIONS

G. Abbate

New developments in liquid-crystal-based photonic devices

D. Antzoulatos

Use of the hypermatrix

M. Barnabei and L. Montefusco (Vol. 119)

Algebraic aspects of signal and image processing

L. Bedini, E. Salerno, and A. Tonazzini (Vol. 119)

Discontinuities and image restoration

I. Bloch

Fuzzy distance measures in image processing

R. D. Bonetto

Characterization of texture in scanning electron microscope images

G. Borgefors

Distance transforms

A. van den Bos and A. den Dekker (Vol. 117)

Resolution

Y. Cho

Scanning nonlinear dielectric microscopy

E. R. Dougherty and Y. Chen (Vol. 117)

Granulometries

G. Evangelista (Vol. 117)

Dyadic warped wavelets

R. G. Forbes

Liquid metal ion sources

E. Förster and F. N. Chukhovsky

X-ray optics

A. Fox

The critical-voltage effect

L. Frank and I. Müllerová

Scanning low-energy electron microscopy

P. Hartel, D. Preikszas, R. Spehr, H. Mueller, and H. Rose (Vol. 119)

Design of a mirror corrector for low-voltage electron microscopes

P. W. Hawkes

Electron optics and electron microscopy: conference proceedings and abstracts as source material

M. I. Herrera

The development of electron microscopy in Spain

K. Hiraga

Structural analysis of quasicrystals based on recent work

K. Ishizuka

Contrast transfer and crystal images

I. P. Jones (Vol. 119)

ALCHEMI

W. S. Kerwin and J. Prince (Vol. 119)

The kriging update model

G. Kögel

Positron microscopy

W. Krakow

Sideband imaging

C. L. Matson

Back-propagation through turbid media

J. C. McGowan (Vol. 118)

Magnetic transfer imaging

S. Mikoshiba and F. L. Curzon

Plasma displays

K. A. Nugent, A. Barty, and D. Paganin (Vol. 118)

Non-interferometric propagation-based techniques

E. Oesterschulze (Vol. 118)

Scanning tunnelling microscopy

M. A. O'Keefe

Electron image simulation

N. Papamarkos and A. Kesidis

The inverse Hough transform

J. C. Paredes and G. R. Arce

Stack filtering and smoothing

C. Passow

Geometric methods of treating energy transport phenomena

E. Petajan

HDTV

F. A. Ponce

Nitride semiconductors for high-brightness blue and green light emission

H. de Raedt, K. F. L. Michielsen, and J. Th. M. Hosson

Aspects of mathematical morphology

H. Rauch

The wave-particle dualism

D. Saad, R. Vicente, and A. Kabashima

Error-correcting codes

G. Schmahl

X-ray microscopy

S. Shirai

CRT gun design methods

T. Soma

Focus-deflection systems and their applications

I. Talmon (Vol. 119)

Study of complex fluids by transmission electron microscopy

I. R. Terol-Villalobos (Vol. 118)

Morphological image enhancement and segmentation

M. Tonouchi

Terahertz radiation imaging

T. Tsutsui and Z. Dechun

Organic electroluminescence, materials and devices

Y. Uchikawa

Electron gun optics

D. van Dyck

Very high resolution electron microscopy

C. D. Wright and E. W. Hill

Magnetic force microscopy

M. Yeadon (Vol. 119)

Instrumentation for surface studies

ACKNOWLEDGMENTS

The concepts for the contents of this volume arose mainly over a long span from the author's experience, gained during his work in the Institute of Applied Physics at the University of Tübingen, Germany, often in cooperation with his coworkers, who contributed useful ideas, and with program-users, who tested the programs by practical applications in electron optical designs. To all my former co-workers and students who challenged me to improve the numerical techniques, I am thankful for their help.

Special acknowledgments are owed to Dr. P. W. Hawkes for his generosity in improving and correcting the text, to my wife Rose, to Mrs. Robert for the tedious work of typing the manuscript and to Mrs. Joan E. Wolk for the editorial work.

Figures 5.26, 5.27 and Table 6.6 are reprinted from: E. Kasper, "An advanced boundary element method for the calculation of magnetic lenses", *Nuclear Instruments and Methods A* **450** (2000), pp 173–178, with kind permission from Elsevier Science, London.

The reproduction of Figures 3.9, 5.5, 7.5, 7.16, 7.23 and of Tables A2, A3 from publications in the *Optik* was kindly permitted by Wissenschaftliche Verlagsgesellschaft, Stuttgart, Germany.

This Page Intentionally Left Blank

INTRODUCTION

GENERAL CONSIDERATIONS

The field of charged particle optics as a whole and the application of modern computers in its various branches have developed rapidly in the past decades and gained great importance in modern technology. Computers are now used at every stage: in the design of new devices, in the adjustment and control of their components, and also in the online evaluation of measurements. It is hardly possible to cover all the various aspects of this widespread field in one single volume and therefore it is necessary to be selective; in this volume we are solely concerned with computer methods for the first stage, the design of new devices.

Characteristics of Computer-Aided Design

The object of computer-aided design (CAD) in charged particle optics is to predict the optical properties of a new device, before a test model is actually manufactured. The attraction of such a procedure is evident: because a preliminary design by guesswork is rarely satisfactory in any respect, it has to be gradually improved with suitable design modifications, and this is far easier with CAD than with real prototypes; construction of a model is then the final step, after the best shape has been found.

The CAD itself passes through different stages. In the beginning, some fundamental decisions concerning number and function of the components of a device, the nature and quality of its output, limitations imposed by technical constraints, and so forth, have to be made. Thereafter a coarse optimization is attempted by use of simple models for the different components such as lenses and deflectors. When this procedure has led to a satisfactory answer, the final shapes of the electrodes or pole pieces and coils are selected, following which the electromagnetic field in the device can be calculated correctly by numerical techniques. The relevant particle–optical properties, the position of a focus or an image, the magnification, and the aberrations can then be determined by accurate ray tracing. It may then become obvious that some parts of the design should be altered, whereupon the calculation cycle is repeated with suitable modifications until a finally acceptable result is found.

It is not self-evident that such a procedure will always lead to a satisfactory solution, because the technical demands and constraints may contradict each other. A still larger difficulty lies in the fact that the various branches of

charged particle optics have developed quite different concepts, goals, and tools. For instance, the terminology in electron microscopy is entirely different from that used in high-energy physics, and the performance of an electron microscope cannot usefully be compared with that of a ring accelerator. It therefore makes little sense to deal with CAD in too general a manner.

For this reason the techniques of optimization will not be considered in this volume, although some kind of optimization is certainly always the final goal. A review of specialized models and techniques is given by Hawkes and Kasper [1] and by the research group at Delft University [2]. The present volume is concerned with one essential part of the design procedure that certainly consumes a major part of the computation time : the calculation of electromagnetic fields with given boundary conditions on the surfaces of yokes and pole pieces. This implies that the geometric shapes of these surfaces must already be defined in advance. Even this task is already so vast that not all kinds of problems can be dealt with. The discussion of radio frequency fields, for example, is very limited, and these are dealt with only as far as they fit the general schemes.

The final goal of field calculations is in very many cases the development and repeated activation of a program element, which enables the electromagnetic field vectors $\mathbf{E}(r, t)$ and $\mathbf{B}(r, t)$ to be calculated at any relevant position $\mathbf{r} = (x, y, z)$ in space, so that systematic ray tracing is possible by numerical solution of the Lorentz equation of particle motion. Another possible form of the output is a set of series expansion coefficients, which make it possible to calculate aberrations by perturbation theory. For reasons of space these latter topics cannot be dealt with in the present volume, although they are the logical continuation of the field calculation. In this respect we must refer to the relevant literature, for instance, reference ref. [1], where many more references can be found.

Here we shall be concerned with the various methods of field calculation and finally with their useful combinations. Chapter VII shows that, in the author's experience, these hybrid methods are very suitable for overcoming many of the drawbacks of 'pure' methods. This concept requires a survey of all relevant techniques.

Survey of the Topics Examined

The presentation of the topics proceeds in a natural way, from general contents to more specialized ones. Thus Chapter I, Basic Field Equations, starts with Maxwell's equations. These are followed by basic concepts: variational principles, differential and integral equations for various potentials, and kinds of sources. In principle, the whole edifice of classical electrodynamics could be

considered here as the foundation, but the presentation is kept as concise as possible.

Chapter II then concentrates on systems with *rotationally symmetric surfaces* as these are of special importance in particle optics; many essential elements, notably lenses and magnetic deflectors, have such surfaces. Furthermore, other components such as multipole correctors with many poles can often be approximated fairly well by systems with rotationally symmetric surfaces. This brings an important gain in simplicity, as it is then possible to carry out field calculations in only two relevant dimensions. Certainly this is not always possible; when it is not possible, a considerable increase of the computational effort must be accepted.

Chapter III, Basic Mathematical Tools, is concerned with those techniques that are later frequently used in the development of the different field calculation methods. It is essential to recognize that numerical calculations of initially unknown functions are quite often performed by *discretization*. This means that a set of discrete arguments or positions is first chosen, and then the corresponding function values, the sampling values, are obtained from the solution of a suitable system of equations. Thereafter the function values in the interior of intervals can be found by interpolation. Known but complicated functions can also be treated in this manner. Consequently, emphasis is put on all those techniques that support this concept; these are the various kinds of splines. It must be pointed out that this account of numerical analysis is far from being comprehensive, but it cannot be the purpose of the present volume.

Chapter IV is devoted to the first important method of field calculation, the finite difference method (FDM). This historically oldest technique was thought to be less attractive than the finite element method, because it is more difficult to match it to configurations with arbitrary surfaces. This is, however, only partly true, as suitable transforms of the mesh can reduce the number of irregular points for which systematic approximations are available. The chapter presents different ways of deriving formulas for the potentials in the nodes of a rectangular mesh that may be distorted, but emphasis is put on nine-point formulas, those that relate the potential in any node to those at its eight closest neighbors. Apart from the fact that these are mostly very accurate, it is of importance to realize that they are compatible with finite-element approximations and that they can hence be incorporated in corresponding programs.

Chapter V deals with the finite element method (FEM). This method has been investigated in great detail, as is shown in the comprehensive work of Zienkiewicz [3] and is quite popular in many branches of engineering and architecture. The first applications of the FEM to field calculation in electron optics were carried out by Munro [4], by Lencova [5], and by Mulvey and

Tahir [6]; numerous other publications followed and gradual improvements were made.

A characteristic requirement in particle optics, which is not so stringent in mechanical engineering, is the smoothness of the solution obtained on the boundary between adjacent finite elements. However, not only the potential but also the field strength must remain continuous there; otherwise ray tracing would become very complicated. This requires some special considerations.

For conciseness, the presentation of the FEM in this chapter is confined to those aspects that are of importance in charged particle optics. The construction of meshes and the definition of trial functions in these are worked out in such a form that these can be used again in the next chapter without any repetition and this demonstrates the generality of these concepts. After an account of the necessary interpolation techniques, a brief review of linear algebra is given; these techniques can be applied to the systems of equations arising in the FDM as well as in the FEM and likewise in the context of the techniques subsequently described.

Although the FDM and the FEM have in common the concept that the whole domain of solution is covered by suitable elements in which a locally valid solution is approximated, the boundary element method (BEM), the subject of chapter VI, is based on an entirely different concept. It relies on the representation of potentials as Coulomb integrals over charges or currents located on the surfaces of the device in question (apart from contributions coming from external sources such as coils).

Because these surface sources are initially unknown, they are now approximated locally in terms of suitably chosen trial functions. This means that the surfaces are dissected into suitable area elements, known as *boundary elements*, and that the coefficients of the trial functions are now determined from the condition that the potential or the normal component of its gradient shall satisfy prescribed conditions. This requires the numerical solution of integral equations. Because these have singular kernels, a major part of this chapter is devoted to the problem of carrying out multiple integrations over singular functions.

The BEM is outlined here in two versions. The first version is more general: the concepts of triangulation, familiar in the FEM, are now applied to the surface of a device and the surface source density is then determined in the linear approximation. This method is feasible for the solution of truly three-dimensional, that is, irreducible, boundary value problems. The second version of the BEM is specialized to configurations with rotationally symmetric surfaces—see, for instance, the early work of Singer and Braun [7]. The dissection of the latter into conical elements, called *rings*, makes it necessary to evaluate elliptic integrals of various kinds. Because these have to be evaluated quite frequently, some effort is necessary to carry out the procedure

efficiently. The gain is now a very smooth field, so that ray tracing becomes straightforward. The main drawback of the BEM is that it is restricted in practice to configurations without spatial source distributions, as Coulomb integrations in all three dimensions become very tedious.

It will emerge in the course of the presentation that all three basic methods have essential difficulties or even limitations as well as certain advantages. Chapter VII, Hybrid Methods, is therefore devoted to the task of overcoming these disadvantages by suitable combination of the basic techniques and some other methods, not outlined earlier, such as the general alternation method. At the end of the volume the reader will have the impression that a spectrum of different techniques in various possible combinations is now available. It is the author's assessment that none of the so-called 'pure' methods can be satisfactory in any respect, so that hybrid techniques should be used. This is obviously in conflict with the familiar requirement to have only one unique program structure, but it is just this demand that leads to essential limitations.

The Form of the Mathematical Representation

This volume is *not* written in a rigorous mathematical style but in a simplified shortened form that is still understandable. This implies that all functions introduced are continuously differentiable as often as needed. This is assumed implicitly, whenever the contrary is not stated explicitly. Similarly, the domain of definition is not stated explicitly whenever this is obvious from the context. Well-known proofs of existence and uniqueness such as those for the solutions of the Dirichlet problem for Poisson's equation in closed domains and other such general theorems are also not presented here. The reader who is interested in these topics is referred to the corresponding mathematical literature. In general, the presentation of proofs and derivations is kept as concise as possible, and straightforward calculation procedures are sketched only verbally. From a mathematical standpoint, this is certainly unsatisfactory, but otherwise the survey of the various methods within the given frame would not be possible.

This Page Intentionally Left Blank

**ADVANCES IN IMAGING AND
ELECTRON PHYSICS**

VOLUME 116

**NUMERICAL FIELD CALCULATION
FOR CHARGED PARTICLE OPTICS**

This Page Intentionally Left Blank

CHAPTER I

Basic Field Equations

In this chapter, we shall start with the presentation of Maxwell's equations in their most general form. Although a program that could solve them for arbitrary initial boundary and material conditions would be of interest, this is hardly feasible in practice, as the amount of necessary data and computer operations would be tremendous. Therefore, we shall gradually specialize Maxwell's equations to cases that are of importance in charged particle optics and that comprise the majority of computation problems. Throughout this volume we shall follow the standard notation in electrodynamics, presented in Table 1.1; any necessary deviations from it will be mentioned explicitly.

1.1 MAXWELL'S EQUATIONS

Maxwell's equations are partial differential equations for vectorial field functions, which all depend on the spatial position $\mathbf{r} = (x, y, z)$ and the time t . In the notation that is familiar in vector analysis, they are given by

$$\text{curl} \mathbf{E}(\mathbf{r}, t) = -\partial \mathbf{B}(\mathbf{r}, t) / \partial t, \quad (1.1)$$

$$\text{curl} \mathbf{H}(\mathbf{r}, t) = \partial \mathbf{D}(\mathbf{r}, t) / \partial t + \mathbf{j}(\mathbf{r}, t), \quad (1.2)$$

$$\text{div} \mathbf{D}(\mathbf{r}, t) = \rho(\mathbf{r}, t), \quad (1.3)$$

$$\text{div} \mathbf{B}(\mathbf{r}, t) = 0 \quad (1.4)$$

These vector functions are interrelated by material equations, which describe the electromagnetic properties of matter on a phenomenological basis.

The electric properties are characterized by a *polarization* $\mathbf{P}(\mathbf{r}, t)$, the spatial density of electric dipoles:

$$\mathbf{D}(\mathbf{r}, t) = \varepsilon_0 \mathbf{E}(\mathbf{r}, t) + \mathbf{P}(\mathbf{r}, t). \quad (1.5)$$

Similarly, the analogous magnetic property is the *magnetization* $\mathbf{M}(\mathbf{r}, t)$, the magnetic dipole density, usually defined by

$$\mathbf{B}(\mathbf{r}, t) = \mu_0 \mathbf{H}(\mathbf{r}, t) + \mu_0 \mathbf{M}(\mathbf{r}, t). \quad (1.6)$$

TABLE 1.1
STANDARD NOTATIONS

<i>Electrodynamics</i>	
E	Electric field strength
D	Electric displacement vector
B	Magnetic field strength
H	Magnetic excitation vector
P	Electric polarization
M	Magnetization
$\varepsilon_0, \varepsilon$	Dielectric coefficient
μ_0, μ	Magnetic permeability
$\nu = 1/\mu$	Magnetic reluctivity
ρ	Space charge density
σ	Surface charge density
j	Electric current density
J, ω	Surface current density
κ	Conductivity
A	Vector potential
ψ	Magnetic flux potential
Scalar potentials have different notations in different contexts.	
<i>Mathematics</i>	
(x, y, z)	Cartesian coordinates
(z, r, φ)	Cylindric coordinates
(R, ϑ, φ)	Spherical coordinates
$(u, v, w), (q_1, q_2, q_3)$	General coordinates
ξ_1, ξ_2, ξ_3	Area coordinates
∇	Vector differentiation symbol
$\Delta = \nabla^2$	Laplace operator
Δ_α	Cylindric Laplace operator

With respect to applications in charged particle optics the polarization $\mathbf{P}$ is of little importance, as electrodes are usually conductors, and consequently all static electric fields vanish in these. In the few cases in which insulator materials are present, the assumption of *proportionality* is sufficient:

$$\mathbf{P}(\mathbf{r}, t) = (\varepsilon(\mathbf{r}) - \varepsilon_0)\mathbf{E}(\mathbf{r}, t), \quad (1.7)$$

whereupon Eq. (1.5) reduces to

$$\mathbf{D}(\mathbf{r}, t) = \varepsilon(\mathbf{r})\mathbf{E}(\mathbf{r}, t). \quad (1.8)$$

Even in configurations with a constant dielectric coefficient ε , this becomes a spatial function because it alters discontinuously at any surface. It is now possible to eliminate the vector field $\mathbf{D}(\mathbf{r}, t)$ from Maxwell's equations, and we shall do this here.

The analogous linearizations with respect to magnetic fields would result in

$$\mathbf{M}(\mathbf{r}, t) = (\mu(\mathbf{r})/\mu_0 - 1)\mathbf{H}(\mathbf{r}, t), \quad (1.9)$$

$$\mathbf{B}(\mathbf{r}, t) = \mu(\mathbf{r})\mathbf{H}(\mathbf{r}, t). \quad (1.10)$$

It would be a considerable simplification if these relations were valid throughout. We then speak of *linear* or *unsaturated* media.

Unfortunately, this assumption does *not* always hold, and there are then different steps of generalization. The simplest one is the case of *isotropic nonlinear* media, most favorably presented by

$$\mathbf{H}(\mathbf{r}) = \nu(\mathbf{r}, |\mathbf{B}|)\mathbf{B}(\mathbf{r}). \quad (1.11)$$

This equation means that $\mathbf{H}$ and $\mathbf{B}$ always have the same *direction*, but the material factor $\nu := \mu^{-1}$ called the *reluctivity* depends on the norm of $\mathbf{B}$. This assumption is simplistically made in most finite-element programs for the calculation of magnetic lenses, for instance, those written by Munro [4] and Lencova [5].

This form, however, is not always sufficient. For instance, it cannot be applied to devices with permanent magnets or with magnetically anisotropic materials. In such cases, we must start from the more general material equation

$$\mathbf{H}(\mathbf{r}) = \mu_0^{-1}\mathbf{B}(\mathbf{r}) - \mathbf{M}(\mathbf{r}, \mathbf{B}), \quad (1.12)$$

in which $\mathbf{M}$ does not necessarily vanish for $\mathbf{B} = 0$.

The system of basic equations is completed by a relation between the current density $\mathbf{j}$ and the electromagnetic field. Its simplest and most familiar form is

$$\mathbf{j} = \kappa\mathbf{E}, \quad (1.13)$$

where κ is the conductivity. We shall, however, hardly ever need this equation in charged particle optics, because here we are mainly interested in the spatial distribution $\mathbf{j}(\mathbf{r})$, producing a designed magnetic field. The determination of the voltage, to be applied to the coils, is an elementary task.

1.2 ELECTROMAGNETIC POTENTIALS

In the main course of this volume, we shall specialize to *stationary*, that is, time-independent fields, as these comprise most cases of technical importance; if we have to consider configurations with time-dependent fields, this will be stated explicitly.

1.2.1 *Electrostatic Fields*

If we disregard the electric field in the coils of magnetic devices, as is usually done, the system of Maxwell's equations becomes uncoupled, leading to an important simplification. The electrostatic part then reduces to

$$\operatorname{curl} \mathbf{E} = 0 \quad (1.14a)$$

$$\operatorname{div} \mathbf{D} = \rho \quad (1.14b)$$

$$\mathbf{D} = \varepsilon \mathbf{E}. \quad (1.14c)$$

Equation (1.14a) can be integrated once by the introduction of an electrostatic potential $V(\mathbf{r})$,

$$\mathbf{E}(\mathbf{r}) = -\operatorname{grad} V(\mathbf{r}). \quad (1.15)$$

It is favorable to eliminate the $\mathbf{D}$ field completely, whereupon the system (1.14) reduces to one partial differential equation of second order for $V(\mathbf{r})$,

$$\operatorname{div}(\varepsilon(\mathbf{r}) \operatorname{grad} V(\mathbf{r})) = -\rho(\mathbf{r}). \quad (1.16)$$

Because $\varepsilon > 0$, this has the basic form of a *self-adjoint elliptic* equation. In the course of this volume, we shall encounter such a mathematical form frequently in various contexts but with different physical meanings of its variables.

Inhomogeneous dielectric properties are of little importance in charged particle optics (the discontinuity of ε at the surface between different materials must be considered by boundary conditions). With $\varepsilon = \text{const.}$ Eq. (1.16) simplifies to *Poisson's* equation. As this will appear quite frequently, we shall introduce a simplified notation for the *Laplace* operator,

$$\Delta := \nabla^2 \equiv \operatorname{div} \operatorname{grad}. \quad (1.17)$$

In cartesian coordinates (and only in these) this operator takes the simple form

$$\Delta = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2. \quad (1.18)$$

We shall use the simpler symbol Δ where this is not misleading: otherwise we shall use the notation ∇^2 . Poisson's equation now takes the familiar form

$$\Delta V(\mathbf{r}) \equiv \nabla^2 V(\mathbf{r}) = -\rho(\mathbf{r})/\varepsilon, \quad (1.19)$$

and for $\rho \equiv 0$, this reduces to Laplace's equation.

1.2.2 Vector Potentials

With respect to magnetic fields, the situation becomes far more complicated. Maxwell's equations now specialize to

$$\text{curl } \mathbf{H} = \mathbf{j} \quad (1.20a)$$

$$\text{div } \mathbf{B} = 0 \quad (1.20b)$$

$$\mathbf{H} = \mu_0^{-1} \mathbf{B} - \mathbf{M}(\mathbf{r}, \mathbf{B}). \quad (1.20c)$$

We can integrate Eq. (1.20b) once by the introduction of a *vector potential* $\mathbf{A}(\mathbf{r})$, giving

$$\mathbf{B}(\mathbf{r}) = \text{curl } \mathbf{A}(\mathbf{r}), \quad (1.21)$$

but now the elimination of the $\mathbf{H}$ -field is far more complicated. Moreover, we face another difficulty. The choice of the electrostatic potential is unique (apart from an unimportant constant), but for the vector potential we can choose different *gauges*. This means that all fields $\mathbf{A}'(\mathbf{r})$ with

$$\mathbf{A}'(\mathbf{r}) = \mathbf{A}(\mathbf{r}) + \text{grad } \chi(\mathbf{r}) \quad (1.22)$$

satisfy Eq. (1.21) equally well. We shall impose the additional condition

$$\text{div } \mathbf{A}(\mathbf{r}) = 0, \quad (1.23)$$

which often brings a simplification but is not really necessary; sometimes we will have to abandon it. If the material is linear, that means $\nu(\mathbf{r})$ independent of $\mathbf{B}$, the relation $\mathbf{H} = \nu \mathbf{B}$ can be favorably used to eliminate this field from Eqs. (1.20). After some short vector analytical calculations we arrive at

$$\nu \{ \text{grad div } \mathbf{A} - \nabla^2 \mathbf{A} \} + \text{grad } \nu \times \text{curl } \mathbf{A} = \mathbf{j}. \quad (1.24)$$

For inhomogeneous materials, for which $\nu \neq \text{const.}$, this equation is hardly ever used in practice. Since then the use of energy functionals (see next section) is more advantageous. In fact, inhomogeneities usually result from hysteresis and owing to Eq. (1.11), the preceding assumptions would not hold.

For truly homogeneously linear media, however, Eq. (1.24) simplifies with Eq. (1.23) and $\mu = 1/\nu$ to the *vector Poisson equation*

$$\text{curl curl } \mathbf{A} = -\Delta \mathbf{A}(\mathbf{r}) = \mu \mathbf{j}(\mathbf{r}). \quad (1.25)$$

We emphasize that Eq. (1.18) can be used only if all vector fields are represented in cartesian form.

1.2.3 Magnetic Scalar Potentials

Although the reduction to this equation means an essential simplification, its numerical solution will become quite complicated unless it is possible to reduce it further to one essential component. An alternative way is the introduction of a magnetic *scalar* potential, which is possible in *current-free* and simply connected domains, in which no current is ever encircled. We then always have $\nabla \times \mathbf{H} = 0$ and can write

$$\mathbf{H}(\mathbf{r}) = \text{grad } W(\mathbf{r}) \quad (1.26)$$

in analogy to Eq. (1.15). For constant permeability, Eq. (1.20b) immediately leads to Laplace's equation for the so-called *total* scalar potential:

$$\Delta W(\mathbf{r}) = 0. \quad (1.27)$$

This procedure is mathematically correct but has the severe difficulty that the domain of solution must be appropriately dissected to exclude current-conducting parts and that it is thereafter hardly possible to find the appropriate boundary values at the surface of the reduced domain. Nevertheless, there are important examples of the feasibility of this method.

If all materials in a magnetic device can be assumed to be linear, there is a way out of this difficulty, as this linearity allows us to separate the magnetic field into a *driving* field $\mathbf{H}_0(\mathbf{r})$ and a contribution $\mathbf{H}_M(\mathbf{r})$ from the materials, which superimpose linearly. The driving field is the field that would be produced by all coils in the absence of any material and therefore satisfies

$$\text{curl } \mathbf{H}_0(\mathbf{r}) = \mathbf{j}(\mathbf{r}), \quad \text{div } \mathbf{H}_0(\mathbf{r}) = 0. \quad (1.28)$$

Together with the natural condition that $\mathbf{H}_0$ vanishes at infinity, this field is uniquely defined and can be found by integrating Biot-Savart's law.

The remainder $\mathbf{H}_M(\mathbf{r})$ is current-free and can hence be represented as the gradient of a new potential $U(\mathbf{r})$, which is now called the *reduced* scalar potential. Altogether we have

$$\mathbf{H}(\mathbf{r}) \equiv \mathbf{H}_0(\mathbf{r}) + \mathbf{H}_M(\mathbf{r}) \equiv \mathbf{H}_0(\mathbf{r}) + \text{grad } U(\mathbf{r}). \quad (1.29)$$

The combination of this equation with $\mathbf{B} = \mu \mathbf{H}$ and $\text{div } \mathbf{B} = 0$ results in the self-adjoint elliptic equation

$$-\text{div}(\mu(\mathbf{r}) \text{grad } U(\mathbf{r})) = \mathbf{H}_0 \cdot \text{grad } \mu =: \rho_m(\mathbf{r}) \quad (1.30)$$

in analogy to Eq. (1.16). The function $\rho_m(\mathbf{r})$, defined by this equation, is a *formal* magnetic source density; it has no physical reality but serves only to

reflect this analogy. In the most frequent case $\mu = \text{const.}$, this “charge density” vanishes and we again arrive at Laplace’s equation

$$\Delta U(\mathbf{r}) = 0. \quad (1.31)$$

The advantage of this concept lies in the fact that here the necessary boundary conditions are more easily satisfied, because there is no need for cuts through the domains of solution: this can simply be the entire space R_3 . The gain with respect to the vector potential formalism arises from the fact that one scalar function can be used to describe the magnetic field instead of three coupled ones.

1.2.4 Coefficient Transformation

In self-adjoint elliptic differential equations such as Eqs. (1.16) or (1.30), the coefficient $\varepsilon(\mathbf{r})$ or $\mu(\mathbf{r})$ can always be removed by a simple coefficient transformation. Because this coefficient must certainly be positive, we can write it in the form of $a^2(\mathbf{r})$; in a slight generalization we have then the self-adjoint equations of the form

$$\nabla \cdot (a^2(\mathbf{r}) \nabla V(\mathbf{r})) + b(\mathbf{r})V(\mathbf{r}) + c(\mathbf{r}) = 0. \quad (1.32)$$

If we introduce a new potential $\Phi(\mathbf{r})$ by

$$\Phi(\mathbf{r}) := a(\mathbf{r})V(\mathbf{r}), \quad (1.33)$$

we find in turn by partial differentiations:

$$a^2 \nabla V = a \nabla \Phi - \Phi \nabla a, \quad (1.34)$$

and thereafter (with $\nabla^2 \equiv \Delta$)

$$\nabla \cdot (a^2 \nabla V) = a \Delta \Phi - \Phi \Delta a. \quad (1.35)$$

Finally, we obtain the transformed self-adjoint equation

$$\Delta \Phi(\mathbf{r}) + \bar{b}(\mathbf{r})\Phi(\mathbf{r}) + \bar{c}(\mathbf{r}) = 0 \quad (1.36)$$

with the new coefficients

$$\bar{b}(\mathbf{r}) = b/a^2 - \Delta a/a, \quad (1.37)$$

$$\bar{c}(\mathbf{r}) = c(\mathbf{r})/a(\mathbf{r}). \quad (1.38)$$

This transformation can be useful in the derivation of discretization formulas for self-adjoint equations, especially for the finite-differences method (see Chapter IV).

1.3 VARIATIONAL PRINCIPLES

In classical mechanics, it is usual to derive the equations of motion in two different but equivalent ways: (1) by transformation of Newton's or Einstein's law or (2) by derivation from Hamilton's law of least action as a variational principle. Similarly, in classical electrodynamics, there are two analogous procedures: (1) the direct formulation of Maxwell's equations as we have presented it so far; and (2) derivation from a variational principle for three-dimensional fields, which is the topic of this chapter.

We shall state here the general rule without mathematical proof; the latter can be found in any comprehensive mathematical textbook.

Let us consider an n -dimensional space with coordinates $x_1, x_2, \dots, x_n$, which need not necessarily be cartesian. Moreover, we consider m independent and sufficiently differentiable field functions $y_1(x_1, \dots, x_n)$, $y_2(x_1, \dots, x_n)$, $y_m(x_1, \dots, x_n)$ in this space. They are defined within a domain G and satisfy prescribed and invariable boundary conditions on its $(m-1)$ -dimensional surface ∂G .

The main task is now the formulation of an appropriate Lagrangian or Lagrange density $L(x_1, \dots, x_n; y_1, \dots, y_m; \partial y_1/\partial x_1, \dots, \partial y_i/\partial x_k, \dots, \partial y_m/\partial x_n)$, a function that may depend on the coordinates, the field functions, and all their partial derivatives of first order.

The variational principle is now the statement that a *functional* F becomes stationary:

$$F := \iiint_G \cdots \int L dx_1 dx_2 \cdots dx_n = \min. \quad (1.39)$$

Generally, the stationary value could also be a maximum or a saddle point, but such a solution is of little interest with respect to the finite-element method, which is the main field of application. The differential equations for the functions $y_1, \dots, y_m$ are now obtained from the familiar Euler equations:

$$\frac{\partial L}{\partial y_i} = \sum_{k=1}^n \frac{\partial}{\partial x_k} \left(\frac{\partial L}{\partial (\partial y_i / \partial x_k)} \right) \quad i = 1, \dots, m. \quad (1.40)$$

The derivatives, appearing in these, are to be understood as *explicit* ones.

We now specialize to those cases that are of importance for practical field calculation; the most general form of the variational principle, leading to

Eqs. (1.1)–(1.4), can be found in any comprehensive textbook on electrodynamics and in Arfken [8].

1.3.1 Scalar Potentials

We set $n = 3$ and $m = 1$ and identify $y_1(x_1, x_2, x_3)$ with $V(\mathbf{r})$.

The Lagrangian consists of an *energy density* and a *potential density* p in the form

$$\mathbf{L} = \frac{1}{2}\varepsilon(\mathbf{r})(\nabla V(\mathbf{r}))^2 - p(\mathbf{r}, V) \quad (1.41)$$

in analogy with a familiar representation in terms of kinetic energy and the potential in classical mechanics. Here it is advantageous to define the source density ρ by

$$\rho(\mathbf{r}, V) := \partial p / \partial V. \quad (1.42)$$

Then, after a short calculation, the evaluation of Eq. (1.40) results in

$$\nabla \cdot (\varepsilon \nabla V) = -\rho,$$

which is the same as Eq. (1.16). The more specialized case that ρ is independent of V is obtained from the potential density

$$p(\mathbf{r}, V) = \rho(\mathbf{r})V(\mathbf{r}). \quad (1.43)$$

With corresponding changes of notation, the variational principle for the potentials $U(\mathbf{r})$ and $W(\mathbf{r})$ in Section 1.2.3 can be found easily.

1.3.2 Vector Potentials

The potential term is here simply $\mathbf{j} \cdot \mathbf{A}$, but the energy-density term is more complicated if the general material equation (1.12) is required. It must then be possible to construct a function $\Lambda(\mathbf{r}, \mathbf{B})$ such that

$$H_i(\mathbf{r}) = \partial \Lambda(\mathbf{r}, \mathbf{B}) / \partial B_i, \quad i = 1, 2, 3 \quad (1.44)$$

is valid in cartesian representation. This implies that the magnetization $\mathbf{M}$ cannot be prescribed arbitrarily but only in agreement with Eq. (1.44) and hence as a gradient in $\mathbf{B}$ -space.

When we have found this function, the appropriate Lagrange density is given by

$$L(\mathbf{r}, \mathbf{A}, \mathbf{B}) = \Lambda(\mathbf{r}, \mathbf{B}) - \mathbf{j}(\mathbf{r}) \cdot \mathbf{A}(\mathbf{r}). \quad (1.45)$$

The left-hand side of (1.40) now gives immediately

$$\partial L / \partial A_i = -j_i(\mathbf{r}), \quad i = 1, 2, 3. \quad (1.46)$$

The evaluation of the right-hand side leads us first to expressions of the form

$$\frac{\partial L}{\partial(\partial A_i / \partial x_k)} = \sum_{n=1}^3 \frac{\partial L}{\partial B_n} \frac{\partial B_n}{\partial(\partial A_i / \partial x_k)}. \quad (1.47)$$

Expressing $\mathbf{B} = \text{curl} \mathbf{A}$ in cartesian coordinates and considering the expressions $\partial A_i / \partial x_k$ as formally independent variables, we find, for example,

$$\frac{\partial B_1}{\partial(\partial A_3 / \partial x_2)} = -\frac{\partial B_1}{\partial(\partial A_2 / \partial x_3)} = 1, \quad (1.48)$$

whereas all other derivatives of B_1 vanish identically. The derivatives of B_2 and B_3 are obtained by corresponding cyclic permutation of the labels. Considering (1.44), we then find

$$\frac{\partial L}{\partial(\partial A_3 / \partial x_2)} = -\frac{\partial L}{\partial(\partial A_2 / \partial x_3)} = H_1(\mathbf{r}) \quad (1.49)$$

and two other such relations with cyclic permutation of the labels. Finally, the second differentiation on the right-hand side of (1.40) simply gives $\text{curl} \mathbf{H}$ in cartesian coordinates. Hence we arrive at Maxwell's equation $\text{curl} \mathbf{H} = \mathbf{j}$, as we should do; this shows that we have constructed the Lagrangian correctly.

The energy functional to be minimized finally takes the compact form

$$F = \iiint_G (\Lambda(\mathbf{r}, \nabla \times \mathbf{A}) - \mathbf{j}(\mathbf{r}) \cdot \mathbf{A}(\mathbf{r})) d^3 r = \min. \quad (1.50)$$

This value is gauge invariant, although the second term of the integrand contains the gauge potential χ of Eq. (1.22) explicitly. To show this, we can carry out a partial integration of the corresponding term

$$F' := \iiint_G \mathbf{j}(\mathbf{r}) \cdot \nabla \chi(\mathbf{r}) d^3 r = - \iiint_G \chi \nabla \cdot \mathbf{j} d^3 r + \iint_{\partial G} \chi \mathbf{j} \cdot d\mathbf{a}. \quad (1.51)$$

The volume integral vanishes identically because $\text{div} \mathbf{j} \equiv 0$. The surface integral may give some contribution if the domain G is chosen unreasonably. But in every reliable field calculation the boundary is chosen far enough away to enclose all currents; we then have $\mathbf{j}(\mathbf{r}) = 0$ on ∂G , and consequently the surface integral vanishes too.

1.3.3 The Magnetic Energy Density

It remains to find a function $\Lambda(\mathbf{r}, \mathbf{B})$ that satisfies Eq. (1.44) in agreement with material equation. This is quite easy in the case of *linear* media: we then obtain

$$\Lambda(\mathbf{r}, \mathbf{B}) = \frac{1}{2} \nu(\mathbf{r}) B^2(\mathbf{r}) = \mathbf{H} \cdot \mathbf{B} / 2, \quad (1.52)$$

which is the familiar energy density of the magnetic field. The differentiation according to Eq. (1.44) then gives $\mathbf{H} = \nu \mathbf{B}$ with $\nu = 1/\mu$. Moreover, it is easy to find Λ for linear but anisotropic media; ν is then to be replaced by a symmetric tensor; hence

$$\Lambda = \frac{1}{2} \sum_{i=1}^3 \sum_{k=1}^3 \nu_{ik}(\mathbf{r}) B_i B_k = \frac{1}{2} \mathbf{H} \cdot \mathbf{B}, \quad (1.53)$$

which comprises Eq. (1.52) as a special case.

If the medium is nonlinear, it may become quite difficult to find the appropriate function Λ . We specialize therefore to the most important case of *isotropic* media, which satisfy Eq. (1.11). Then the integral

$$\Lambda(\mathbf{r}, B) = \int_0^B H(\mathbf{r}, B') dB' \quad (1.54)$$

with $B = |\mathbf{B}|$ is the correct function. In fact, differentiation with respect to B , the upper limit of this integral, and consideration of the isotropy result in

$$\mathbf{H} = \frac{\mathbf{B}}{B} \frac{\partial \Lambda}{\partial B} = \nu(\mathbf{r}, B) \mathbf{B}(\mathbf{r}); \quad (1.55)$$

hence

$$\nu(\mathbf{r}, B) = B^{-1} \partial \Lambda / \partial B. \quad (1.56)$$

This function $\Lambda(\mathbf{r}, B)$ can be interpreted as the area under the inverted hysteresis curve as is sketched in Fig. 1.1. The dependence on $\mathbf{r}$ is a consequence of the fact that, owing to the use of different materials in a realistic device, the hysteresis curve depends on the local position.

Concluding Remarks

The minimization of energy functionals is in some sense equivalent to the solution of partial differential equations for potentials, as both procedures lead to the same final solution if the same consistent boundary conditions are obeyed. Yet there is an important difference: the variational principles presented so

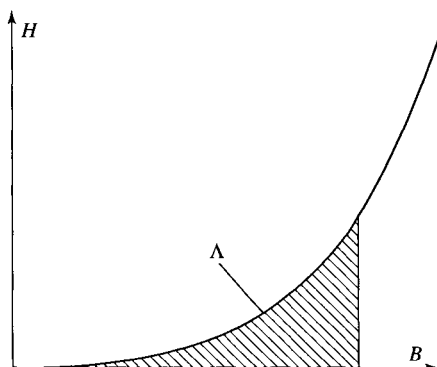


FIGURE 1.1 Inverse hysteresis curve $H(B)$. Λ is the area under this curve up to the given abscissa B .

far require only partial derivatives of *first* order. This has the consequence that, if a device consists of different materials, it suffices to dissect the total domain G into corresponding subdomains $G_1, G_2, \dots$, to evaluate the functionals $F_1, F_2, \dots$, in these and to sum them up. Different material properties are thus considered in a natural way.

In contrast to this, the partial differential equations for potentials are of *second* order, and it is often more complicated to calculate derivatives of second order than those of first order. Moreover, at *inner* surfaces with different materials on both sides, special boundary conditions must be satisfied, which will be the topic of Sections 1.5–1.8.

Comparing these alternatives, the reader might get the impression that the application of variational principles is the more advantageous method, but this is not always true. The appropriate answer to this question depends on the particular problem to be solved and can be given only at the end, once the numerical tools have been developed in some detail.

1.4 WAVE EQUATIONS AND HERTZ VECTORS

In this section, we shall assume *constant* material coefficients ϵ , μ , and κ , at least in each particular medium. In the context of wave propagation, the simple relation (1.13) may be too specialized and is now generalized to

$$\mathbf{j} = \kappa \mathbf{E} + \rho \mathbf{v}, \quad (1.56)$$

$\mathbf{v}(\mathbf{r}, t)$ being the local velocity of the moving charge. However, the continuity condition

$$\operatorname{div} \mathbf{j} + \partial \rho / \partial t = 0 \quad (1.57)$$

must be satisfied throughout the space. The conductivity term in (1.56) is caused by imperfect insulating properties in dielectrics, and the charge term is caused by a driving particle beam. Both terms are spatial functions, which are nonvanishing only in different domains.

We shall consider first a general time dependence of all field functions and denote derivatives with respect to time, as usual by dots. The fields $\mathbf{H}(\mathbf{r}, t)$ and $\mathbf{D}(\mathbf{r}, t)$ can be completely eliminated from Maxwell's equations. By suitable differentiations the wave equations can be decoupled. Thus, we find in turn

$$\nabla \times (\nabla \times \mathbf{E}) = -\nabla \times \dot{\mathbf{B}} = -\mu(\varepsilon \ddot{\mathbf{E}} + \partial \mathbf{j} / \partial t) \quad (1.58)$$

and

$$\nabla(\nabla \cdot \mathbf{E}) = \varepsilon \nabla \rho. \quad (1.59)$$

Putting both together by using the vector identity for the double cross product and considering (1.56), we arrive at

$$\Delta \mathbf{E} - \varepsilon \mu \ddot{\mathbf{E}} - \kappa \mu \dot{\mathbf{E}} = \mu \frac{\partial}{\partial t}(\rho \mathbf{v}) + \varepsilon \nabla \rho. \quad (1.60)$$

In an analogous way, we find the second wave equation

$$\Delta \mathbf{B} - \varepsilon \mu \ddot{\mathbf{B}} - \kappa \mu \dot{\mathbf{B}} = -\mu \nabla \times (\rho \mathbf{v}). \quad (1.61)$$

These differential equations describe the propagation of damped electromagnetic waves with "sources" or "driving terms," given by the expressions on their right-hand side. Apart from optics, their solutions comprise the extensive field of high-frequency techniques.

In cartesian coordinates the wave equations for the individual components have the same constant coefficients $\varepsilon \mu$ and $\kappa \mu$. Yet these equations cannot be solved independently and subsequently composed to form vectors, as they are coupled by $\nabla \cdot \mathbf{E} = \varepsilon \rho$, $\nabla \cdot \mathbf{B} = 0$, and boundary conditions.

Hertz Vectors

For the *source-free* wave propagation *in vacuo*, the coupling by divergence relations can be removed by the introduction of a Hertz potential $\mathbf{Z}(\mathbf{r}, t)$. We now set $\varepsilon = \varepsilon_0$, $\mu = \mu_0$, and introduce the speed of light

$$c = (\varepsilon_0 \mu_0)^{-1/2}. \quad (1.62)$$

The charge density $\rho(\mathbf{r}, t)$ may be so small that the corresponding driving term may be ignored. The wave equations then simplify to

$$c^2 \Delta \mathbf{E} = \ddot{\mathbf{E}}, \quad c^2 \Delta \mathbf{B} = \ddot{\mathbf{B}}. \quad (1.63)$$

A possible solution of Maxwell's equations can be found by assuming the relation

$$\mathbf{E} = \nabla \times (\nabla \times \mathbf{Z}) = \text{grad div } \mathbf{Z} - \Delta \mathbf{Z}, \quad (1.64)$$

which already satisfies $\text{div } \mathbf{E} = 0$. We presume that the sequence of partial differentiations can be exchanged arbitrarily. Differentiating (1.64) further and considering Maxwell's equation (1.2) we then find

$$\dot{\mathbf{E}} = c^2 \nabla \times \mathbf{B} = \nabla \times (\nabla \times \dot{\mathbf{Z}}). \quad (1.65)$$

To satisfy this relation, it is sufficient to assume

$$\mathbf{B} = c^{-2} \nabla \times \dot{\mathbf{Z}}, \quad (1.66)$$

which guarantees that $\text{div } \mathbf{B} = 0$. Finally, the wave equation for $\mathbf{Z}(\mathbf{r}, t)$ is obtained from

$$-\nabla \times \mathbf{E} = \nabla \times (\Delta \mathbf{Z}) = \dot{\mathbf{B}} = c^{-2} \nabla \times \ddot{\mathbf{Z}}. \quad (1.67)$$

This is satisfied if $\mathbf{Z}$ obeys

$$c^2 \Delta \mathbf{Z} = \ddot{\mathbf{Z}}. \quad (1.68)$$

This wave equation has the same form as Eq. (1.63); yet the divergence condition $\text{div } \mathbf{Z} = 0$ is now *not* necessary, and this is a simplification. On the other hand, the boundary conditions on the surfaces of conductors become more complicated, so that this method is most suitable for wave propagation in free space.

The wave modes obtained in this way are referred to as “electric” ones, as in the transition to static fields, only the electric part survives. In analogy to these, we can also define “magnetic” modes by

$$\mathbf{B} = \nabla \times (\nabla \times \mathbf{Z}_M(\mathbf{r}, t)), \quad (1.69)$$

$$\mathbf{E} = -\nabla \times \dot{\mathbf{Z}}_M, \quad (1.70)$$

and again

$$c^2 \Delta \mathbf{Z}_M = \ddot{\mathbf{Z}}_M. \quad (1.71)$$

In many cases, especially in the physics of resonant cavities, it is permissible to assume *harmonic* time dependence, which means

$$\mathbf{F}(\mathbf{r}, t) = \Re[f(\mathbf{r}) \exp(-i\omega t)] \quad (1.72)$$

for all field functions in question, whereupon D'Alembert's wave equation simplifies to the Helmholtz equation

$$\Delta f(\mathbf{r}) + k^2 f(\mathbf{r}) = 0, \quad (1.73)$$

with $k = \omega/c$. This wavenumber k then becomes an eigenvalue and the accurate numerical calculation of it is an additional difficult task.

It may sometimes be interesting to study the time-independent solution. All field functions then satisfy Laplace's equation. Obviously, Eq. (1.64) can be brought into agreement with $\mathbf{E} = -\text{grad } V$ by writing

$$V(\mathbf{r}) = -\text{div } \mathbf{Z}(\mathbf{r}). \quad (1.74)$$

From Eq. (1.69), it becomes obvious that the vector potential is given by

$$\mathbf{A}(\mathbf{r}) = \text{curl } \mathbf{Z}_M(\mathbf{r}). \quad (1.75)$$

But by analogy with Eq. (1.64), with $\Delta \mathbf{Z}_M = 0$, we can also write $\mathbf{B} = \mu_0 \text{grad } W$ with

$$W(\mathbf{r}) = \mu_0^{-1} \text{div } \mathbf{Z}_M(\mathbf{r}). \quad (1.76)$$

The two representations are equivalent.

1.5 BOUNDARY CONDITIONS

In the preceding sections, we have derived various forms of partial differential equations of second order and mentioned briefly that they must obey certain boundary conditions, but we have not yet presented these explicitly; this is the task of the present section. As a consequence, we shall encounter a new class of physical quantities, the *surface source densities*—electric and magnetic ones.

In configurations with constant material coefficients ε and μ in each medium, knowledge of the surface source densities offers an alternative way of calculating electromagnetic fields: the evaluation of integral equations instead of the solution of differential equations or the minimization of energy functionals.

This is the basis of the boundary-element method. Here we shall be concerned only with the formulation of *integral equations*. The far more difficult task of determining the surface source densities is deferred to Chapter VI.

Generally, the solution of an elliptic partial differential equation is specified uniquely if a linear combination of the potential in question and its normal derivative assume prescribed values on a *closed outer* surface (exceptional conditions will be mentioned explicitly). On *inner* boundaries between

different materials, the potential itself, and consequently all tangential derivatives, must remain continuous; otherwise a discontinuity of the potential would cause infinite derivatives. The normal derivatives on the two sides of such a surface are linked by a so-called *material condition*, whose derivation is the task of the present section.

1.5.1 Electric Material Conditions

The situation is sketched in Fig. 1.2 showing a local surface normal $\mathbf{n}$ and one of the surface tangents $\mathbf{t}$, which are all unit vectors. The normal $\mathbf{n} \equiv \mathbf{n}_{1,2}$ is directed from the material 1 toward material 2.

When the electric field is not obtained from $\mathbf{E} = -\nabla V$, the continuity of its tangential component $\mathbf{t} \cdot \mathbf{E}_2 = \mathbf{t} \cdot \mathbf{E}_1$ for *all* possible tangents $\mathbf{t}$ implies that

$$\mathbf{n} \times (\mathbf{E}_2 - \mathbf{E}_1) = 0. \quad (1.77)$$

If it is possible to make use of an electric potential $V(\mathbf{r})$, Eq. (1.77) is already satisfied, as mentioned previously.

The nontrivial material condition arises from $\text{div } \mathbf{D} = \rho$: if we first distribute the charge in a thin layer along the surface, carry out the integration over it in the normal direction, and finally shrink the thickness of this layer to zero with conservation of charge, we arrive at the familiar condition

$$\mathbf{n} \cdot (\mathbf{D}_2 - \mathbf{D}_1) \equiv \mathbf{n} \cdot (\epsilon_2 \mathbf{E}_2 - \epsilon_1 \mathbf{E}_1) = \sigma(\mathbf{r}) \quad (1.78)$$

for the electric surface charge density $\sigma(\mathbf{r})$. Generally, this surface function cannot be chosen arbitrarily but must be determined consistently, and this is the task of the boundary-element method. Only in the special case of wave propagation through a surface between two dielectrics do we know in advance that then $\sigma(\mathbf{r})$ must vanish.

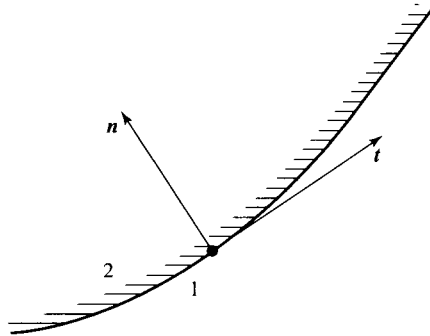


FIGURE 1.2 Directions of local tangent $\mathbf{t}$ and surface normal $\mathbf{n}$.

In the context of potentials, we shall frequently encounter *normal derivatives*. The corresponding operator is

$$\partial_n := \mathbf{n} \cdot \nabla = n_x \partial/\partial x + n_y \partial/\partial y + n_z \partial/\partial z. \quad (1.79)$$

Hence, in combination with $\mathbf{E} = -\text{grad } V$, we can rewrite (1.78) as

$$\varepsilon_1(\partial_n V)_1 - \varepsilon_2(\partial_n V)_2 = \sigma(\mathbf{r}). \quad (1.80)$$

In many cases, notably in all devices with electrostatic lenses, one of the materials, say medium 1, is a conductor that does *not* carry any electric current. Then the condition $\mathbf{j} = \kappa \mathbf{E}_1 = 0$ leads to $V = \text{const}$. This is quite a simple form of a Dirichlet problem. Then Eq. (1.80) remains valid with $(\partial_n V)_1 \equiv 0$.

1.5.2 Magnetic Material Conditions

From $\text{div } \mathbf{B} = 0$ we know that magnetic charges cannot exist; hence, in analogy to (1.78), we now have

$$\mathbf{n} \cdot (\mathbf{B}_2 - \mathbf{B}_1) \equiv \mathbf{n} \cdot (\mu_2 \mathbf{H}_2 - \mu_1 \mathbf{H}_1) \equiv 0. \quad (1.81)$$

It is usual to assume that the tangential component of the $\mathbf{H}$ -field is also continuous, so that in analogy to Eq. (1.77) the relation

$$\mathbf{n} \times (\mathbf{H}_2 - \mathbf{H}_1) = 0 \quad (1.82)$$

is valid. In combination with Eq. (1.81), this leads to the law of ‘magnetic refraction,’

$$\tan \alpha_2 / \tan \alpha_1 = \mu_2 / \mu_1, \quad (1.83)$$

which is sketched in Fig. 1.3.

Certainly, real current-conducting coils cannot be made infinitely thin; yet it might be favorable to introduce a *surface current density* $\mathbf{J}(\mathbf{r})$ to simplify calculations that would be more complicated otherwise. This is a vector function that is locally perpendicular to the plane that contains the surface normal $\mathbf{n}$ and the difference vector $\mathbf{H}_2 - \mathbf{H}_1$; the generalization of Eq. (1.82) hence becomes

$$\mathbf{n} \times (\mathbf{H}_2 - \mathbf{H}_1) = \mathbf{J}(\mathbf{r}). \quad (1.84)$$

This can be proven by assuming that a current-conducting coil of thickness Δh and breadth Δs lies on the surface with current $\Delta I = J \Delta s$ (Fig. 1.4). If

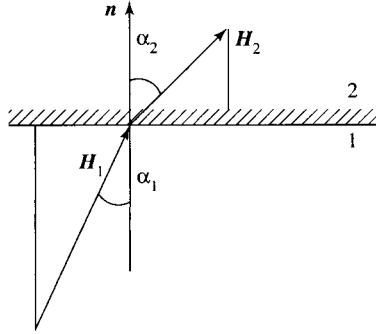


FIGURE 1.3 Law of magnetic refraction.

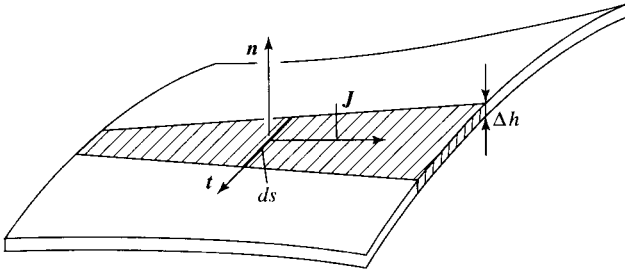


FIGURE 1.4 Surface current element.

we then evaluate Ampere's law for a closed loop around the surface of this coil, we find

$$\oint \mathbf{H} \cdot d\mathbf{s} = \mathbf{t} \cdot (\mathbf{H}_2 - \mathbf{H}_1) \Delta s = \Delta I = J \Delta s \quad (1.85)$$

with some unimportant contributions from the side faces of the loop, which are proportional to Δh and that vanish in the limit $\Delta h \rightarrow 0$.

The breadth Δs finally cancels out and we obtain $J = \mathbf{t} \cdot (\mathbf{H}_2 - \mathbf{H}_1)$, which is in agreement with (1.84). The components of the vector function $\mathbf{J}(\mathbf{r})$ cannot be chosen independently. The conservation of current in stationary fields requires that the total current

$$I = \int_c \mathbf{J} \times \mathbf{n} \cdot d\mathbf{s} = \text{const.} \quad (1.86)$$

through any cross section of a stripe, formed by flux lines of the $\mathbf{J}$ field, always have the *same* value (see Figs. 1.4 and 1.5).

A simple example of the favorable application of surface currents is the calculation of the magnetic field outside a superconducting device, for instance,

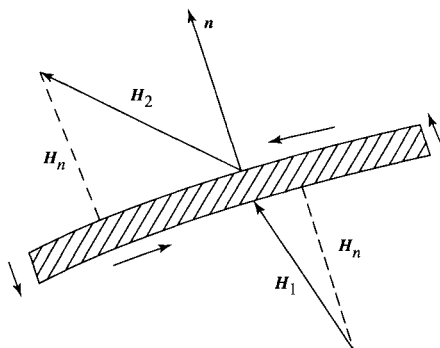


FIGURE 1.5 Discontinuity of the $\mathbf{H}$ -field produced by surface currents (flowing perpendicular to the plane of the drawing); the arrows indicate the direction of the line integration; the normal component H_n is conserved.

a magnetic electron lens with a superconducting pole piece. Owing to the Meissner-Ochsenfeld effect, the magnetic field is expelled from the interior of the superconductor and screened by electric currents that flow in a very thin sheet near the surface. The thickness of this sheet is not important for the physics of this lens, and we can integrate over it. Then, if medium 1 is the superconductor, Eq. (1.84) holds with $\mathbf{H}_1 = 0$ and also Eq. (1.81) with $\mathbf{B}_1 = 0$, $\mathbf{n} \cdot \mathbf{B}_2 = 0$. The field on the vacuum side hence has a purely tangential component. Other examples of the use of surface currents will follow in Section 1.7.2.

1.6 INTEGRAL EQUATIONS FOR ELECTROSTATIC FIELDS

For conciseness, the following considerations are formulated for electrostatic fields, but after appropriate exchange of constants they also hold for any other potential fields that obey a Poisson equation.

We consider two scalar potential functions $U(\mathbf{r})$ and $V(\mathbf{r})$, defined in a spatial domain Γ with *closed* surface $\partial\Gamma =: S$. The familiar Green's identity tells us that (with $\Delta = \nabla^2$)

$$\int_{\Gamma} (U \Delta V - V \Delta U) dv = \int_S (U \partial_n V - V \partial_n U) da \quad (1.87)$$

is valid, $dv \equiv d^3r$ denoting the volume element and da the surface element. The surface normal appearing in the operator ∂_n [see Eq. (1.79)] points in the *outward* direction. This identity holds for any pair of functions for which the required differentiations can be carried out; hence we are free to

impose Poisson's equation (1.19) on V , whereas for $U(\mathbf{r})$ we can choose the *unbounded* Green function [8]

$$G(\mathbf{r}, \mathbf{r}') = (4\pi|\mathbf{r} - \mathbf{r}'|)^{-1}. \quad (1.88)$$

Evidently this function is symmetric in $\mathbf{r}$ and $\mathbf{r}'$ and becomes singular as $\mathbf{r} \rightarrow \mathbf{r}'$. Because a Coulomb potential G satisfies Laplace's equation, apart from the singularity, we can write

$$\Delta G(\mathbf{r}, \mathbf{r}') = \Delta' G(\mathbf{r}, \mathbf{r}') = -\delta(\mathbf{r}, \mathbf{r}'), \quad (1.89)$$

with $\delta(\mathbf{r}, \mathbf{r}')$ denoting the three-dimensional Kronecker symbol.

To simplify integral expressions, we shall consider further the point $\mathbf{r}$ as variable of integration and $\mathbf{r}'$ as "reference"—or "observation"—point. The normal derivative (referring always to the *second* argument)

$$\partial_n G = \mathbf{n} \cdot \nabla G = \frac{(\mathbf{r}' - \mathbf{r}) \cdot \mathbf{n}(\mathbf{r})}{4\pi|\mathbf{r}' - \mathbf{r}|^3} =: p(\mathbf{r}', \mathbf{r}) \quad (1.90)$$

can be interpreted as the potential of a normalized *dipole* oriented in the direction of the local surface normal $\mathbf{n}$; this function is *not* antisymmetric with respect to $\mathbf{r}$ and $\mathbf{r}'$ if $\mathbf{n} \neq \mathbf{n}'$. Its singularity is so strong that a sphere or a spherical sector must be excluded from the domain P of definition, as demonstrated in Figs 1.6 and 1.7. The integration over such a spherical sector with solid angle Ω or edge angle α gives

$$\int_C p(\mathbf{r}', \mathbf{r}) da =: \beta(\mathbf{r}') = \Omega/4\pi = \alpha/2\pi. \quad (1.91)$$

Because this is independent of the radius, the limit for vanishing radius is not critical. The different situations, shown in Fig. 1.6 are now specified by the

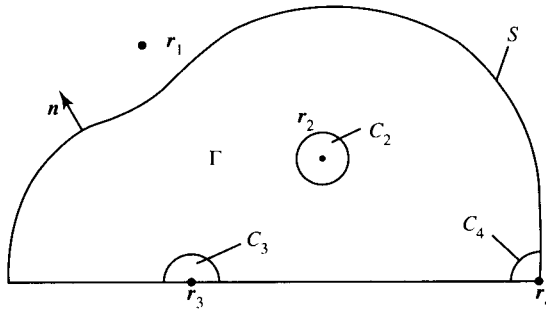


FIGURE 1.6 Domain Γ with surface S and different cases of excluded spherical sectors.

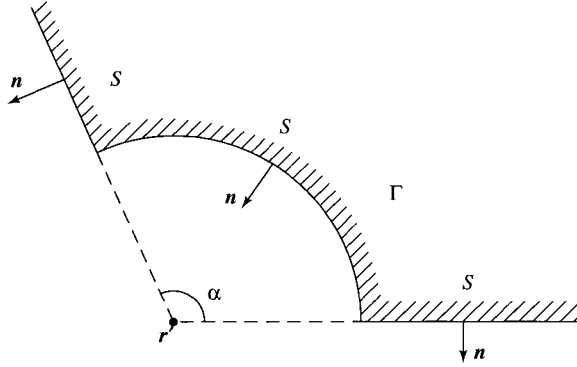


FIGURE 1.7 Excluded spherical sector of angle α near an edge $\mathbf{r}'$; its radius will be made zero in the limit.

following table of values:

$$\begin{aligned}
 \Omega &= 0, & \beta &= 0 & \text{at } \mathbf{r}_1: \text{ external point,} \\
 \Omega &= 4\pi, & \beta &= 1 & \text{at } \mathbf{r}_2: \text{ internal point,} \\
 \Omega &= 2\pi, & \beta &= 1/2 & \text{at } \mathbf{r}_3: \text{ regular boundary,} \\
 \Omega &= 2\alpha; & \beta &= \alpha/2\pi & \text{at } \mathbf{r}_4: \text{ edge point.}
 \end{aligned} \tag{1.92}$$

Putting all this together, we finally arrive at

$$\begin{aligned}
 \beta(\mathbf{r}')V(\mathbf{r}') &= \varepsilon^{-1} \int_{\Gamma} G(\mathbf{r}', \mathbf{r})\rho(\mathbf{r}) dv + \int_S G(\mathbf{r}', \mathbf{r})\partial_n V(\mathbf{r}) da \\
 &\quad - \oint_S p(\mathbf{r}', \mathbf{r})V(\mathbf{r}) da,
 \end{aligned} \tag{1.93}$$

the symbol $\oint$ denoting the *principal value* of the integral. The three terms on the right-hand side can in turn be interpreted as a spatial Coulomb integral, a surface Coulomb integral, and a surface polarization term. The first one is usually considered as *given* (although the determination of space charges in electron guns is a difficult task).

The remaining terms are coupled by an additional *boundary condition* of the general form

$$a(\mathbf{r})V(\mathbf{r}) + b(\mathbf{r})\partial_n V(\mathbf{r}) = c(\mathbf{r}), \quad (\mathbf{r} \in S); \tag{1.94}$$

hence, we have two linear relations (1.93) and (1.94) to determine the functions $V(\mathbf{r})$ and $\partial_n V(\mathbf{r})$ on the surfaces. This is a feasible, albeit complicated, task; we shall therefore now present some simple special cases.

1.6.1 Dirichlet Problems

The boundary values $V = \bar{V}(\mathbf{r})$ on all surfaces are prescribed; hence, we are free to choose

$$a(\mathbf{r}) = 1, \quad b(\mathbf{r}) = 0, \quad c(\mathbf{r}) = \bar{V}(\mathbf{r}) \quad (1.95)$$

in Eq. (1.94). The polarization term in (1.93) is now known in principle, but the numerical evaluation of such integrals is rather complicated; we hence try to eliminate it.

To achieve this, we write down the integral equation (1.93) for the complementary domain $\Gamma^* = \mathfrak{R}_3 - \Gamma$, which is the outer domain. Because this extends to infinity, it is necessary to assume the *natural* boundary conditions

$$V \rightarrow 0, \quad \nabla V \rightarrow 0 \quad \text{for } |\mathbf{r}| \rightarrow \infty. \quad (1.96)$$

Then for any point $\mathbf{r}' \in \Gamma^*$, including its surface, the integral equation

$$\begin{aligned} (1 - \beta(\mathbf{r}'))V^*(\mathbf{r}') &= \varepsilon^{*-1} \int_{\Gamma^*} G(\mathbf{r}', \mathbf{r}) \rho^*(\mathbf{r}) dv - \int_S G(\mathbf{r}', \mathbf{r}) \partial_n V^*(\mathbf{r}) da \\ &\quad + \oint_S p(\mathbf{r}', \mathbf{r}) V(\mathbf{r}) da \end{aligned} \quad (1.97)$$

is valid; the changes of sign are a consequence of the inverted direction of the surface normal $\mathbf{n}$. The polarization integral and the term with factor β now cancel out if we add this integral equation to (1.93).

A further simplification is achieved if it is possible to assume $\varepsilon = \varepsilon^* \approx \varepsilon_0$, as is justified for systems of metallic electrodes and vacuum domains. We can then use the surface charge density $\sigma(\mathbf{r})$ [Eq. (1.80)] and write

$$V(\mathbf{r}') = V_e(\mathbf{r}') + \int_S \frac{\sigma(\mathbf{r})}{4\pi\varepsilon_0|\mathbf{r}' - \mathbf{r}|} da, \quad (1.98)$$

which is valid *everywhere* in space, if we rename $V^* = V$ in Γ^* . The contribution $V_e(\mathbf{r}')$ is the sum of all spatial Coulomb integrals. As a further generalization, we can include any external or “driving” potential in this term.

A possible method of field calculation becomes obvious. First, we evaluate Eq. (1.98) at the boundary, which means that $\mathbf{r}' \in S$, $V(\mathbf{r}') = \bar{V}(\mathbf{r}')$.

This requires the solution of an integral equation for the unknown $\sigma(\mathbf{r})$. When this function has been determined, we can use Eq. (1.98) for field calculation in whole space; moreover, the gradient can be calculated from

$$\nabla' V(\mathbf{r}') = \nabla' V_e(\mathbf{r}') + \int_S \frac{(\mathbf{r} - \mathbf{r}')\sigma(\mathbf{r})}{4\pi\varepsilon_0|\mathbf{r} - \mathbf{r}'|^3} da. \quad (1.99)$$

1.6.2 Linear Material Equations

We now consider two homogeneous materials with dielectric constants ε_1 and ε_2 .

The domains of solution will now be the whole $\mathfrak{R}_3$, and at infinity the potential again satisfies the natural boundary conditions.

As before, the surface normal is directed from medium 1 into medium 2. In contrast to the familiar mathematical formulation of boundary-value problems, the surface S between the two materials is not the boundary in the familiar sense but an *inner* surface.

With the notation $\Gamma_1 = \Gamma$, $\Gamma_2 = \mathfrak{R}_3 - \Gamma$, we have to consider the two Poisson equations

$$\Delta' V(\mathbf{r}') = -\rho_k(\mathbf{r}')/\varepsilon_k, \quad \mathbf{r}' \in \Gamma_k \quad (k = 1, 2). \quad (1.100)$$

Moreover, we shall now assume Eq. (1.80) with $\varepsilon_1 \neq \varepsilon_2$. For reasons of conciseness, we shall exclude sharp corners on the surface S ; hence $\beta = 1/2$ for $\mathbf{r} \in S$. Writing down Eq. (1.93) for both media, we obtain for the *unknown* surface potential $\bar{V}(\mathbf{r})$

$$\begin{aligned} \frac{1}{2}\bar{V}(\mathbf{r}') &= \varepsilon_k^{-1} \int_{\Gamma_k} G(\mathbf{r}', \mathbf{r}) \rho_k(\mathbf{r}) dv - (-1)^k \int_S G(\mathbf{r}', \mathbf{r}) \partial_n V_k(\mathbf{r}) da \\ &+ (-1)^k \oint_S p(\mathbf{r}', \mathbf{r}) \bar{V}(\mathbf{r}) da \quad (k = 1, 2). \end{aligned} \quad (1.101)$$

Here the notation is simplified in the sense that we have written $\partial_n V_k(\mathbf{r})$ instead of $(\partial_n V(\mathbf{r}))_k$.

These surface derivatives can be eliminated by means of Eq. (1.80), if the two integral equations are multiplied in turn by ε_k and subsequently added, giving

$$\begin{aligned} \frac{1}{2}(\varepsilon_1 + \varepsilon_2)\bar{V}(\mathbf{r}') + (\varepsilon_1 - \varepsilon_2) \oint_S p(\mathbf{r}', \mathbf{r}) \bar{V}(\mathbf{r}) da \\ = \sum_{k=1}^2 \int_{\Gamma_k} G(\mathbf{r}', \mathbf{r}) \rho_k(\mathbf{r}) dv + \int_S G(\mathbf{r}', \mathbf{r}) \sigma(\mathbf{r}) da \quad (\mathbf{r}' \in S). \end{aligned} \quad (1.102)$$

This integral equation can be applied in different ways. One way is the prescription of the boundary values $\bar{V}(\mathbf{r})$. This is again a Dirichlet problem and is the natural generalization of Eq. (1.98) for $\varepsilon_1 \neq \varepsilon_2$; in fact, for $\varepsilon_1 = \varepsilon_2 = \varepsilon_0$ we again obtain Eq. (1.98), which is then a Fredholm equation of the *first* kind for the unknown σ .

Alternatively, we can prescribe $\sigma(\mathbf{r})$, which corresponds to the assumption that we have *charged* dielectric surfaces (which, of course, includes $\sigma = 0$ for

a neutral surface as special case). Then Eq. (1.102) is a Fredholm equation of the *second* kind for the unknown $\bar{V}(\mathbf{r})$. Once it has been solved, we have reduced our task to a Dirichlet problem. The disadvantage here is that the polarization terms cannot be completely eliminated.

1.6.3 Integral Equation for Surface Sources

The solution of Eq. (1.102) is always the first step of a two-step procedure: thereafter an integral equation for surface sources has to be solved. It is therefore advantageous to combine both into one, and that will be done now. We set out from

$$V(\mathbf{r}') = V_e(\mathbf{r}') + \int_S G(\mathbf{r}', \mathbf{r}) \eta(\mathbf{r}) da, \quad (1.103)$$

$$\nabla' V(\mathbf{r}') = \nabla' V_e(\mathbf{r}') + \int_S \nabla' G(\mathbf{r}', \mathbf{r}) \eta(\mathbf{r}) da, \quad (1.104)$$

in which $\eta = \sigma/\varepsilon_0$ is *not* necessarily valid.

This surface function $\eta(\mathbf{r})$ is initially unknown and has to be determined in such away that Eq. (1.80) holds in the variable $\mathbf{r}'$.

Because the differentiation cannot be carried out exactly on the surface and we do not have excluding half-spheres as in Figs 1.6 and 1.7, we choose two reference points slightly outside the surface: $\mathbf{r}_k = \mathbf{r}' \pm \mathbf{n}'h$ and finally proceed to the limit as $h \rightarrow 0$. The corresponding surface integrals are then

$$\begin{aligned} I_k &= \int_S \mathbf{n}' \cdot \nabla' G(\mathbf{r}_k, \mathbf{r}) \eta(\mathbf{r}) da \\ &= \int_S \frac{\mathbf{n}' \cdot (\mathbf{r} - \mathbf{r}' \pm \mathbf{n}'h)}{4\pi|\mathbf{r} - \mathbf{r}' \pm \mathbf{n}'h|^3} \eta(\mathbf{r}) da, \quad (\mathbf{r}' \in S), \end{aligned} \quad (1.105)$$

the upper sign holding for medium $k = 1$ and the lower one for medium $k = 2$. The term without h in the numerator is not critical, as $\mathbf{n}' \perp (\mathbf{r} - \mathbf{r}')$ in the limit; this gives a principal value, but with exchanged arguments; hence,

$$I_k = \oint p(\mathbf{r}, \mathbf{r}') \eta(\mathbf{r}) da \pm \eta(\mathbf{r}') \cdot \oint \frac{h}{4\pi|\mathbf{r} - \mathbf{r}' \pm \mathbf{n}h|^3} da. \quad (1.106)$$

In the second term we have taken the slowly varying factor $\eta(\mathbf{r}')$ outside the integral. The remaining second integral can be evaluated over the tangential plane at the point $\mathbf{r}'$. This is elementary and gives the value $1/2$. In the limit $h \rightarrow 0$, the integration over the bent surface leads to the same value; hence,

$$I_k = \oint p(\mathbf{r}, \mathbf{r}') \eta(\mathbf{r}) da \pm \frac{1}{2} \eta(\mathbf{r}'). \quad (1.107)$$

Considering this together with Eqs. (1.104) and (1.80), we find

$$\begin{aligned} \frac{1}{2}(\varepsilon_1 + \varepsilon_2)\eta(\mathbf{r}') + (\varepsilon_1 - \varepsilon_2) \oint_S p(\mathbf{r}, \mathbf{r}') \eta(\mathbf{r}) d\mathbf{a} \\ = \sigma(\mathbf{r}') - (\varepsilon_1 - \varepsilon_2) \partial'_n V_e(\mathbf{r}'), \quad (\mathbf{r}' \in S). \end{aligned} \quad (1.108)$$

This Fredholm equation of the second kind has a structure similar to that of Eq. (1.102); however, the integral kernel is now *transposed*. Once it has been solved, Eqs. (1.103), (1.104) can immediately be used for field calculation.

1.7 INTEGRAL EQUATIONS FOR MAGNETIC FIELDS

The basic mathematical tools—Green's integral theorem and limitation processes for its appropriate application—are the same and will not be repeated. Yet, in general, no global scalar potential exists, and we shall therefore not try to formulate an integral equation for it. However, the *reduced* scalar potential can always be defined, and we shall start our next consideration with integral equations for it, as these are analogous to those already obtained.

1.7.1 Scalar Integral Equations

We consider first the external field $\mathbf{H}_0(\mathbf{r})$. Equation (1.28) together with the natural boundary conditions can be integrated by Biot-Savart's law, and it is even easier to find a corresponding vector potential:

$$\mathbf{A}_0(\mathbf{r}') = \mu_0 \int_{\mathcal{R}_3} G(\mathbf{r}', \mathbf{r}) \mathbf{j}(\mathbf{r}) d\mathbf{v}, \quad (1.109)$$

$$\mathbf{H}_0(\mathbf{r}') = \frac{1}{\mu_0} \nabla' \times \mathbf{A}_0 = \int_{\mathcal{R}_3} \nabla' G(\mathbf{r}', \mathbf{r}) \times \mathbf{j}(\mathbf{r}) d\mathbf{v}. \quad (1.110)$$

These integrations may be quite tedious but are intrinsically straightforward. Moreover, it is well known that $\text{div} \mathbf{j} = 0$ implies that $\text{div} \mathbf{A}_0 = 0$. In the subsequent calculations these functions are regarded as known.

We shall now assume constant permeability in each medium; hence (1.29) together with (1.31) is valid. Moreover, it is necessary to assume that the surface currents $\mathbf{J}$ vanish, because Eq. (1.84) conflicts with the use of a scalar potential; Eq. (1.82) is automatically satisfied. The relevant boundary condition (1.81) now takes the form

$$\mathbf{n} \cdot (\mu_2 \nabla U_2 - \mu_1 \nabla U_1) = (\mu_1 - \mu_2) \mathbf{n} \cdot \mathbf{H}_0. \quad (1.111)$$

In analogy with Eq. (1.101) we obtain now for the boundary potential without any space source term

$$\begin{aligned} \frac{1}{2}\bar{U}(\mathbf{r}') &= -(-1)^k \int G(\mathbf{r}', \mathbf{r}) \partial_n U_k(\mathbf{r}) da \\ &+ (-1)^k \oint_S p(\mathbf{r}', \mathbf{r}) \bar{U}(\mathbf{r}) da, \quad (k = 1, 2) \end{aligned} \quad (1.112)$$

and from these, by linear combination with (1.111), we find

$$\begin{aligned} \frac{1}{2}(\mu_1 + \mu_2)\bar{U}(\mathbf{r}') + (\mu_1 - \mu_2) \oint_S p(\mathbf{r}', \mathbf{r}) \bar{U}(\mathbf{r}) da \\ = (\mu_2 - \mu_1) \int_S G \mathbf{n} \cdot \mathbf{H}_0 da. \end{aligned} \quad (1.113)$$

In analogy with Eq. (1.104) formal surface sources $\tau(\mathbf{r})$ can also be introduced, and they have no immediate physical meaning but simply serve to facilitate the field calculation. We rewrite Eq. (1.29) in the form of

$$\mathbf{H}(\mathbf{r}') = \mathbf{H}_0(\mathbf{r}') + \int_S \nabla' G(\mathbf{r}', \mathbf{r}) \tau(\mathbf{r}) da. \quad (1.114)$$

After considerations similar to those in Section 1.6.3, we find

$$\begin{aligned} \frac{1}{2}(\mu_1 + \mu_2)\tau(\mathbf{r}') + (\mu_1 - \mu_2) \oint_S p(\mathbf{r}, \mathbf{r}') \tau(\mathbf{r}) da \\ = (\mu_2 - \mu_1) \mathbf{n}(\mathbf{r}') \cdot \mathbf{H}_0(\mathbf{r}') \quad (\mathbf{r}' \in S). \end{aligned} \quad (1.115)$$

This integral equation not only facilitates the calculations in the sense that a subsequent solution of a Dirichlet problem is saved but also has an even simpler structure than Eq. (1.113), as no integration is required on the right-hand side. Moreover, we can conclude that the “net charge” must vanish. To show this, we also integrate Eq. (1.115) over the surface in the variable $\mathbf{r}'$. For regular surface points, Green’s identity (1.93) gives $\beta = 1/2$, $V = 1$, and reversed notation for $\mathbf{r}$ and $\mathbf{r}'$

$$\oint_S p(\mathbf{r}, \mathbf{r}') da' = -1/2. \quad (1.116)$$

Considering this in the integration over $\tau(\mathbf{r}')$ in (1.115), we find after a suitable change of notation,

$$\mu_2 \oint_S \tau da = (\mu_2 - \mu_1) \int_S \mathbf{n} \cdot \mathbf{H}_0 da = 0, \quad (1.117)$$

which follows from $\text{div } \mathbf{H}_0 = 0$.

1.7.2 Vector Integral Equation

The scalar integral equations are not always satisfactory, as they become ill-conditioned for round magnetic lenses with $\mu_1 + \mu_2 \gg |\mu_1 - \mu_2|$ as will be shown in Section 6.4. Vector equations, that do not have this short-coming are therefore necessary. Such vector integral equations were first derived by Adamiak [9] and later improved essentially by Ströer [10]. Here we shall give a simpler derivation that holds for systems with constant permeabilities. We introduce a surface current density $\boldsymbol{\omega}(\mathbf{r})$, the properties of which are similar to those of the function defined by Eq. (1.84), but it is *not* the same. It is a new variable, which is initially unknown, and will once again serve to facilitate the calculations. The basic equation is now

$$\mathbf{A}(\mathbf{r}') = \mathbf{A}_0(\mathbf{r}') + \mu_0 \int_S G(\mathbf{r}', \mathbf{r}) \boldsymbol{\omega}(\mathbf{r}) d\mathbf{a}, \quad (1.118)$$

with $\mathbf{A}_0$ given by (1.109). Outside the surface S , the integral term is a solution of the vector Laplace equation. On the surface itself, all components are continuous and consequently so are all their *tangential* derivatives. This is quite analogous to the properties of a scalar potential. Consequently, the condition (1.81) is already satisfied, and we can write

$$\mathbf{n} \cdot \mathbf{B} = \mathbf{n} \cdot (\nabla \times \mathbf{A}) = (\mathbf{n} \times \nabla) \cdot \mathbf{A}, \quad (1.119)$$

and the operator $\mathbf{n} \times \nabla$ contains only differential operators in tangential directions.

There remains the task of satisfying Eq. (1.82). First, we take the differentiation under the integral in Eq. (1.118) and obtain

$$\begin{aligned} \mathbf{B}(\mathbf{r}') &= \text{curl}' \mathbf{A}(\mathbf{r}') \\ &= \mu_0 \mathbf{H}_0(\mathbf{r}') + \mu_0 \int_S \nabla' G(\mathbf{r}', \mathbf{r}) \times \boldsymbol{\omega}(\mathbf{r}) d\mathbf{a}. \end{aligned} \quad (1.120)$$

With the abbreviation $\mathbf{n}' := \mathbf{n}(\mathbf{r}')$ the condition $\mu_2 \mathbf{B}_1 \times \mathbf{n}' = \mu_1 \mathbf{B}_2 \times \mathbf{n}'$ leads (cf. Section 1.6.3) to the relation

$$\mu_2 \mathbf{I}_1 - \mu_1 \mathbf{I}_2 = (\mu_1 - \mu_2) \mathbf{H}_0(\mathbf{r}') \times \mathbf{n}', \quad (1.121)$$

with the two vector integrals

$$\mathbf{I}_k = \int_S (\nabla' G(\mathbf{r}_k, \mathbf{r}) \times \boldsymbol{\omega}(\mathbf{r})) \times \mathbf{n}' d\mathbf{a}, \quad (k = 1, 2) \quad (1.122)$$

to be evaluated at positions $\mathbf{r}_k = \mathbf{r}' \pm h\mathbf{n}'$.

The double cross product can be rewritten as

$$\mathbf{I}_k = \int_S \mathbf{n}' \cdot \nabla' G(\mathbf{r}_k, \mathbf{r}) \boldsymbol{\omega}(\mathbf{r}) d\mathbf{a} - \int_S \mathbf{n}' \cdot \boldsymbol{\omega} \nabla' G d\mathbf{a}. \quad (1.123)$$

The first integral is analogous to the former ones and gives an additive term $\pm \boldsymbol{\omega}(\mathbf{r}')/2$, whereas the second remains regular, because $\boldsymbol{\omega}$ is perpendicular to $\mathbf{n}'$ at position $\mathbf{r}'$. Putting all this together, we obtain the integral equation

$$\begin{aligned} & \frac{1}{2}(\mu_1 + \mu_2) \boldsymbol{\omega}(\mathbf{r}') + (\mu_2 - \mu_1) \oint_S (\nabla' G(\mathbf{r}', \mathbf{r}) \times \boldsymbol{\omega}(\mathbf{r})) \times \mathbf{n}(\mathbf{r}') d\mathbf{a} \\ & = (\mu_1 - \mu_2) \mathbf{H}_0(\mathbf{r}') \times \mathbf{n}(\mathbf{r}'). \end{aligned} \quad (1.124)$$

The application of this integral equation becomes very favorable if it is possible to represent the vector function $\boldsymbol{\omega}(\mathbf{r}')$ by only one scalar amplitude. The most important class of examples for which this is permissible consists of round magnetic lenses.

1.8 INTEGRAL EQUATIONS FOR WAVE FIELDS

The calculation of wave fields is far more complicated than that of static fields, as now we have to distinguish between steady and transient waves and between different irradiation conditions in open structures. Moreover, we have to consider damping and possible phase shifts and different polarizations in vector fields. It would go far beyond the scope of this volume to take all this into account; we therefore deal here only with two very simple special cases: these are the solutions of a Dirichlet problem and a Neumann problem for a scalar wave in a *resonance cavity*.

1.8.1 Dirichlet Problem

We consider the scalar Helmholtz equation (1.73) for a *real* function $f(\mathbf{r})$, whose values $\bar{f}(\mathbf{r})$ on the boundary of the cavity are prescribed. In analogy with the electrostatic Dirichlet problem, we can introduce a surface source representation

$$f(\mathbf{r}) = \int_S \sigma(\mathbf{r}') G_k(\mathbf{r}, \mathbf{r}') d\mathbf{a}' \quad (1.125)$$

with the *real* Green function

$$G_k(\mathbf{r}, \mathbf{r}') = \frac{\cos(k|\mathbf{r} - \mathbf{r}'|)}{4\pi|\mathbf{r} - \mathbf{r}'|} = G_k(\mathbf{r}', \mathbf{r}) \quad (1.126)$$

It is easy to verify that for $\mathbf{r}' \neq \mathbf{r}$ the scalar Helmholtz equation is satisfied by this function. For $\mathbf{r} \rightarrow \mathbf{r}'$ it has the *same* strength of singularity as $G(\mathbf{r}, \mathbf{r}')$ from Eq. (1.88) and consequently, the differential equation

$$\Delta G_k + k^2 G_k = \Delta' G_k + k^2 G_k = -\delta(\mathbf{r} - \mathbf{r}') \quad (1.127)$$

is now satisfied. For a *closed* cavity, the boundary values $\bar{f}(\mathbf{r})$ must *vanish*. A nontrivial solution is then obtained only if k becomes an *eigenvalue* of (1.125). Once this eigenvalue problem has been solved, the same integral Eq. (1.125) can be used to determine the wavefield in the interior. The solution is unique up to a free normalization factor.

1.8.2 Neumann Problem

The integral relation (1.125) is now completed by

$$\nabla f(\mathbf{r}) = \int_S \sigma(\mathbf{r}') \nabla G_k(\mathbf{r}, \mathbf{r}') d\mathbf{a}'. \quad (1.128)$$

Consequently, we form the normal derivative by scalar multiplication with $\mathbf{n}$

$$\partial_n f(\mathbf{r}) = \int_S \sigma(\mathbf{r}') \partial_n G_k(\mathbf{r}, \mathbf{r}') d\mathbf{a}'. \quad (1.129)$$

The Neumann problem requires that the left-hand side must vanish by definition. Again, the differentiation under the integral cannot be carried out immediately at the surface. We therefore choose the point $\mathbf{r}_1 = \mathbf{r} - \mathbf{n}h$ and again determine the limit for $h \rightarrow 0$, the result being

$$\oint_S \sigma(\mathbf{r}') \partial_n G_k(\mathbf{r}, \mathbf{r}') d\mathbf{a}' + \frac{1}{2} \sigma(\mathbf{r}) = 0. \quad (1.130)$$

This again requires the solution of an eigenvalue problem. After it is solved, Eq. (1.125) can again be used throughout the interior of the cavity.

REFERENCES

1. Hawkes, P. W. and Kasper, E. (1989). *Principles of Electron Optics*, Volumes 1, 2, London: Academic Press.
2. Van der Stam, M. A. J. (1996). Computer Assistance for the Pre-design of Charged Particle Optics, Dissertation, University of Delft, Netherlands.
3. Zienkiewicz, O. C. (1977). *The Finite Element Method*, New York & London: McGraw-Hill.
4. Munro, E. (1971). Computer-Aided Design Methods in Electron Optics, Dissertation, University of Cambridge, UK.

5. Lencova, M. (1980). Numerical computation of electron lenses by the finite-element method, *Comput. Phys. Commun.* **80**: 127–132.
6. Mulvey, T. and Nasr, H. (1980). An improved finite element program for calculating the field distribution in magnetic lenses, *Proc. in 7th Eur. Cong. Electron Microscopy*, The Hague, 64–65.
7. Singer, B. and Braun, M. (1970). Integral equation method for computer evaluation of electron optics, *IEEE Trans. Electron Dev.* **ED-17**: 926–934.
8. Arfken, G. (1985). *Mathematical Methods for Physicists*, 3rd ed., London & New York: Academic Press.
9. Adamiak, K. (1985). Applications of integral equations to solving inverse problems of stationary electromagnetic fields, *Int. J. Num. Math. Eng.* **21**: 1447–1458.
10. Ströer, M. (1987). Eine Galerkin-Methode mit singulären Formfunktionen und ihre Anwendung auf die Berechnung magnetostatischer Felder, *Optik* **77**: 15–25.

CHAPTER II

Reducible Systems

In reality, all electromagnetic fields are three-dimensional and it would be desirable to calculate them as such. However, in numerical computations it is often necessary to make far-reaching simplifications to keep the computational effort reasonably small. In fact, this effort increases rapidly with every new dimension; hence, we shall try to reduce the number of necessary dimensions as far as possible.

One important class of two-dimensional fields are those in configurations with rotationally symmetric boundaries. These occur, for instance, in round electromagnetic lenses. It is, however, not necessary to assume that the potential is completely independent of azimuth: only the *geometric* shape of the system must be rotationally symmetric.

2.1 AZIMUTHAL FOURIER-SERIES EXPANSIONS

In the subsequent presentation we shall adopt *cylindrical* coordinates (z, r, φ) with the usual definition

$$x = r \cos \varphi, \quad y = r \sin \varphi \quad (2.1)$$

and assume that all functions are 2π -periodic with respect to the azimuth φ for reasons of regularity. The following considerations will hold for any linear self-adjoint differential equation of the fairly general form Eq. (1.32). A reasonable use of cylindrical coordinates can be made only if the coefficients $a(\mathbf{r})$ and $b(\mathbf{r})$ are rotationally symmetric, that is, independent of φ ; this implies a rotationally symmetric configuration of the materials. Equation (1.32) becomes more explicitly

$$\nabla \cdot (a^2(z, r) \nabla V(z, r, \varphi)) + b(z, r) V(z, r, \varphi) + c(z, r, \varphi) = 0. \quad (2.2)$$

By writing out the differentiation operators in cylindrical coordinates, we can again obtain a self-adjoint form after multiplication of the whole equation by r :

$$\begin{aligned} \frac{\partial}{\partial z} \left(ra^2(z, r) \frac{\partial V}{\partial z} \right) + \frac{\partial}{\partial r} \left(ra^2 \frac{\partial V}{\partial r} \right) + \frac{\partial}{\partial \varphi} \left(\frac{a^2}{r} \frac{\partial V}{\partial \varphi} \right) \\ + rb(z, r) \cdot V(z, r, \varphi) + rc(z, r, \varphi) = 0. \end{aligned} \quad (2.3)$$

The azimuth now appears only in V and c , and it is therefore advantageous to expand these functions in the form:

$$V(z, r, \varphi) = \sum_{m=-\infty}^{\infty} V_m(z, r) \exp(-im\varphi), \quad (2.4)$$

$$c(z, r, \varphi) = \sum_{m=-\infty}^{\infty} c_m(z, r) \exp(-im\varphi). \quad (2.5)$$

We shall assume that these two Fourier-series expansions and even all those obtained by differentiations converge absolutely. In all reasonable physical configurations these conditions are satisfied. The reality of V and c requires

$$V_{-m}(z, r) = V_m^*(z, r), \quad c_{-m}(z, r) = c_m^*(z, r), \quad (2.6)$$

the asterisk denoting complex conjugation. If we introduce Eqs. (2.4) and (2.5) into (2.3) and consider the linear independence of different exponentials, we obtain a set of *two-dimensional* differential equations in the variables z and r :

$$\begin{aligned} \frac{\partial}{\partial z} \left(ra^2 \frac{\partial V_m}{\partial z} \right) + \frac{\partial}{\partial r} \left(ra^2 \frac{\partial V_m}{\partial r} \right) + \left(rb - \frac{m^2 a^2}{r} \right) V_m + rc_m(z, r) = 0 \\ (m = 0, \pm 1, \pm 2, \dots). \end{aligned} \quad (2.7)$$

Note that these differential equations are independent of each other (unless the boundary conditions do not conflict with this). This offers the possibility of solving them sequentially, and in this form the three-dimensional problem reduces to a sequence of two-dimensional ones. An approximation error appears only owing to the fact that the series expansions have to be truncated at a reasonably large but finite value of $|m|$.

These Fourier-series expansions are not yet quite favorable: it should be possible to separate factors of the form

$$(x \pm iy)^m = r^m \exp(\pm im\varphi) \quad (m > 0) \quad (2.8)$$

to facilitate differentiation. We therefore define new functions by:

$$V_m(z, r) = r^{|m|} U_m(z, r), \quad (2.8a)$$

$$c_m(z, r) = r^{|m|} S_m(z, r). \quad (2.8b)$$

The regularity conditions imply that these new functions U_m and S_m remain *finite* as $r \rightarrow 0$.

It is now advantageous to introduce the abbreviation

$$\alpha := 2|m| + 1, \quad (2.9)$$

as this expression will appear quite frequently. Then, after introducing Eqs. (2.8a,b) into (2.7) and multiplying the whole resulting equation by $r^{|m|}$, we can cast it in the very concise self-adjoint form

$$\begin{aligned} \frac{\partial}{\partial z} \left(r^\alpha a^2 \frac{\partial U_m}{\partial z} \right) + \frac{\partial}{\partial r} \left(r^\alpha a^2 \frac{\partial U_m}{\partial r} \right) \\ + r^\alpha \left(b + \frac{|m|}{r} \frac{\partial a^2}{\partial r} \right) U_m(z, r) + r^\alpha S_m(z, r) = 0. \end{aligned} \quad (2.10)$$

The term $r^{-1} \partial a^2 / \partial r$ remains *finite* as $r \rightarrow 0$, as for reasons of regularity $a^2(z, r)$ must be an *even* function of r .

Just as in Section 1.2.4, the material coefficient $a^2(z, r)$ can be removed from the differential equations by a transform

$$\Phi_m(z, r) := a(z, r) U_m(z, r), \quad |m| = 0, 1, 2, \dots \quad (2.11)$$

On introducing this into Eq. (2.10), we first obtain the alternating form:

$$\begin{aligned} \frac{\partial}{\partial z} \left[r^\alpha \left(a \frac{\partial \Phi_m}{\partial z} - \Phi_m \frac{\partial a}{\partial z} \right) \right] + \frac{\partial}{\partial r} \left[r^\alpha \left(a \frac{\partial \Phi_m}{\partial r} - \Phi_m \frac{\partial a}{\partial r} \right) \right] \\ + r^\alpha \left(b/a + 2|m| \frac{1}{r} \frac{\partial a}{\partial r} \right) \Phi_m + a^{-1} r^\alpha S_m = 0. \end{aligned} \quad (2.12)$$

With respect to completing the differentiations, it becomes obvious that it is favorable to introduce the differential operator

$$\Delta_\alpha := \frac{\partial^2}{\partial z^2} + \frac{\partial^2}{\partial r^2} + \frac{\alpha}{r} \frac{\partial}{\partial r}. \quad (2.13)$$

Then, after cancelling out a common factor r^α , we finally obtain a differential equation

$$\Delta_\alpha \Phi_m(z, r) + \left[\frac{b(z, r)}{a^2(z, r)} - \frac{\Delta_1 a(z, r)}{a(z, r)} \right] \Phi_m(z, r) + \frac{S(z, r)}{a(z, r)} = 0. \quad (2.14)$$

Note that the label m does not appear in the transformed coefficient; this is a consequence of its assumed rotational symmetry and has the advantage that

the total coefficient function, now written compactly as $\bar{b}(z, r)$ in agreement with Eq. (1.37), needs to be computed only *once* for all Fourier components.

It is possible to remove the factor r^α from Eq. (2.10) by an analogous transformation, so that only the ordinary Laplace operator would appear finally. This is, however, not favorable, at least not for small values of r , because then a strong singularity would appear in the coefficient.

The field calculation is most important in the vicinity of the z -axis, because in most classes of particle-optical devices, the beam remains in this zone. The z -axis is therefore often referred to as the “optic” axis, and we shall do this too. For technical reasons the medium in this domain must be the vacuum. If we exclude high-frequency devices here, the differential equations then become simplified with $b = 0$, $a = 1$, $\Phi_m = U_m$ to

$$\Delta_\alpha U_m(z, r) = -S_m(z, r). \quad (2.15)$$

This class of differential equations is useful not only for field calculations but also for power-series expansions.

2.1.1 Vectors Fields

For conciseness, we shall assume now a constant permeability μ , so that Eqs. (1.23) and (1.25) together with $\text{div } \mathbf{j} = 0$ are valid; the most important exceptional case—the magnetic round lens with saturation effects—is dealt with in Section 2.3.

It is unfavorable to transform the cartesian representation of $\mathbf{A}(\mathbf{r})$ and $\mathbf{j}(\mathbf{r})$ immediately into cylindrical coordinates, as this would lead to unnecessarily complicated expressions. Instead, we introduce the vectorial Fourier-series expansions

$$\mathbf{A}(\mathbf{r}) = \sum_{m=-\infty}^{\infty} r^{|m|} \mathbf{C}_m(z, r) \exp(-im\varphi), \quad (2.16)$$

$$\mathbf{j}(\mathbf{r}) = \sum_{m=-\infty}^{\infty} r^{|m|} \mathbf{J}_m(z, r) \exp(-im\varphi). \quad (2.17)$$

The reality conditions require that

$$\mathbf{C}_{-m}(z, r) = \mathbf{C}_m^*(z, r), \quad \mathbf{J}_{-m}(z, r) = \mathbf{J}_m^*(z, r) \quad (2.18)$$

must be valid. In cartesian form the vector-Poisson equation is valid component-wise. Hence by analogy with the preceding calculations, now

specialized to $a \equiv 1$, $b \equiv 0$, we finally arrive at

$$\Delta_\alpha \mathbf{C}_m(z, r) = -\mu \mathbf{J}_m(z, r), \quad (2.19)$$

with $\alpha := 2|m| + 1$ and Δ_α again given by Eq. (2.13).

This means that we now have *three* two-dimensional differential equations for each Fourier order m .

The conditions $\text{div} \mathbf{j} = 0$ and $\text{div} \mathbf{A} = 0$ now lead to a *coupling* of each of the three subsequent Fourier orders $m - 1$, m , $m + 1$ ($m > 0$). To show this, we rewrite Eq. (2.17) as

$$\mathbf{j}(\mathbf{r}) = \mathbf{J}_0(z, r) + \sum_{m=1}^{\infty} (\mathbf{J}_m(z, r)(x - iy)^m + \text{c.c.}), \quad (2.20)$$

c.c. denoting the complex conjugate. Recalling that $\mathbf{J}_0$ is real and that $r^2 = x^2 + y^2$, the differentiation in cartesian coordinates is straightforward.

The function $\text{div} \mathbf{j}$ can be cast in an analogous form:

$$\text{div} \mathbf{j} = d_0(z, r) + \sum_{m=1}^{\infty} (d_m(z, r)(x - iy)^m + \text{c.c.}). \quad (2.21)$$

To represent the coefficients in a concise form, it is favorable to introduce complex combinations of the transversal components:

$$L_m(z, r) := J_{m,x} - iJ_{m,y} \quad (m \geq 0), \quad (2.22)$$

whereupon we obtain

$$d_0 = \frac{\partial}{\partial z} J_{0,z} + \Re \epsilon (2L_1 + r \partial L_1 / \partial r), \quad (2.23a)$$

$$d_m = \frac{1}{2r} \cdot \frac{\partial}{\partial r} L_{m-1}^* + \frac{\partial}{\partial z} J_{m,z} + \frac{r}{2} \frac{\partial}{\partial r} L_{m+1} + (m+1)L_{m+1} \quad (m \geq 1). \quad (2.23b)$$

Owing to the linear independence of all functions $(x \pm iy)^m$, the condition $\text{div} \mathbf{j} = 0$ can hold identically only if all the coefficients $d_0, d_1, d_2, \dots$ vanish identically, too. This is the earlier-mentioned coupling. Similar considerations, not given here, must hold for the vector potential $\mathbf{A}$. These render the use of vector potentials *unfavorable*, unless special decoupling conditions are given. The relations given earlier are equivalent to those published in Hawkes and Kasper [1].

2.2 ROTATIONALLY SYMMETRIC BOUNDARIES

The calculations in the previous section were based only on the validity of Eq. (2.2) and some regularity conditions and are thus fairly general.

However, the numerical solution of the obtained sequence of partial differential equations is feasible only if the boundary conditions, that are to be satisfied do not conflict with this separation method. A necessary condition for this is that the boundaries should be rotationally symmetric. This is true in all *round* electromagnetic lenses and, for instance, also in magnetic deflectors with toroidal coils.

Sometimes a device does not have wholly rotationally symmetric surfaces, such as in the electric deflector shown in Fig. 2.1. The *correct* field calculation

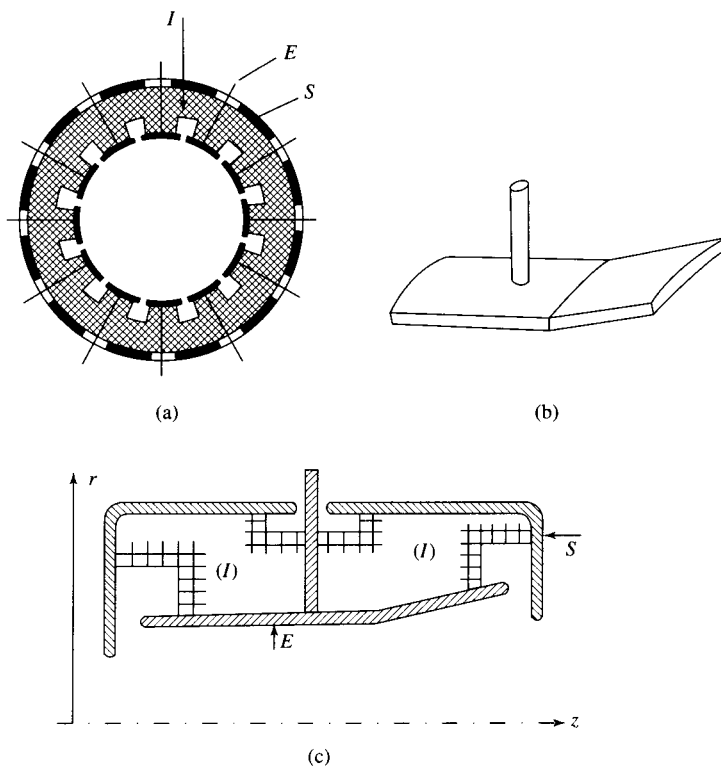


FIGURE 2.1 Electrostatic multipole element: E : electrodes, S : screening, I : insulator. (a) cross section, (b) one single electrode, (c) half of an axial section. Near the optic axis various multipole fields can be generated by choosing corresponding electrode potentials. The surface can be closed to form a rotationally symmetric boundary by choosing interpolated potentials in the gaps.

in such a case is then a truly three-dimensional problem, whose solution is very tedious. However, it is a reasonable and essential simplification to close the gaps between neighboring electrodes to create a round surface and then interpolate the potential in these gaps. The solution thus obtained is, of course, not correct, but comes fairly close to it. Other examples are Klemperer lenses used in cathode-ray tubes [2].

2.2.1 Mathematical Form

Generally, the representation of surfaces in R_3 require two parameters. In the present case it is advantageous to select the azimuth φ round the optic axis as one of these. The other is some *curve parameter* τ (not the time) along any *meridional* section $\varphi = \text{const.}$ through the surface. Thus we obtain a mathematical form

$$\begin{aligned} \bar{x}(\tau, \varphi) &= \bar{r}(\tau) \cos \varphi, & \bar{y}(\tau, \varphi) &= \bar{r}(\tau) \sin \varphi, & \bar{z} &= \bar{z}(\tau) \\ (0 \leq \varphi < 2\pi, 0 \leq \tau \leq T). \end{aligned} \quad (2.24)$$

In practice, a realistic boundary will consist of several distinct parts, as sketched in Fig. 2.1 for example. This provides no essential difficulty, as the interval (O, T) can be dissected into correspondingly many subintervals that are then associated with the different parts.

Although the arc length s is not always the most favorable choice for the curve parameter τ , we shall use it here for reasons of conciseness and uniqueness of the representation. We shall also exclude sharp edges that require special considerations: these are given later. Hence, we have a sequence of smooth contour lines specified by functions $\bar{z}(s)$ and $\bar{r}(s)$. Without loss of generality, it is possible to choose a *positive* orientation, as is sketched in Fig. 2.2.

The normalized curve tangent in the *meridional* plane is then given by

$$\mathbf{t}(s) \equiv (t_z, t_r) = (\bar{z}'(s), \bar{r}'(s)) \quad (2.25)$$

and the two-dimensional normal to it by

$$\mathbf{m}(s) = (m_z, m_r) = (\bar{r}'(s), -\bar{z}'(s)), \quad (2.26)$$

primes denoting derivatives with respect to s . This curve normal is clearly to be distinguished from the surface normal $\mathbf{n} = (n_x, n_y, n_z)$ with components

$$\mathbf{n}(s, \varphi) = (-\bar{z}'(s) \cos \varphi, -\bar{z}'(s) \sin \varphi, \bar{r}'(s)), \quad (2.27)$$

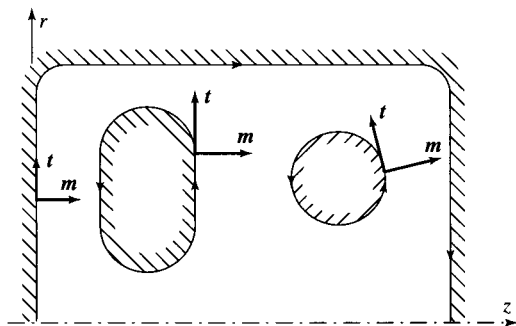


FIGURE 2.2 Example of a closed domain with two inner exclusions (materials). The inner boundary curves are positively oriented, whereas the outer one has the opposite orientation; potentials must be invariant with respect to this choice.

with $\bar{z}^2 + \bar{r}^2 \equiv 1$, as follows from the choice of the arc length s . (Note the different sequence of coordinates).

2.2.2 Fourier Analysis of Boundary Conditions

For the following considerations, we shall assume a linear boundary condition of the form (1.94), now rewritten as

$$\bar{a}(\mathbf{r})V(\mathbf{r}) + \bar{b}(\mathbf{r})\partial_n V(\mathbf{r}) = \bar{c}(\mathbf{r}) \quad (\mathbf{r} \in S) \quad (2.28)$$

to avoid confusion with the coefficients in Eq. (2.2). Here too, the coefficients $\bar{a}$ and $\bar{b}$ must not depend on the azimuth φ , because otherwise we would not have a rotationally symmetric configuration of materials. Hence Eq. (2.28) becomes more explicitly

$$\bar{a}(z, r)V(z, r, \varphi) + \bar{b}(z, r)\partial_n V(z, r, \varphi) = \bar{c}(z, r, \varphi) \quad (\mathbf{r} \in S). \quad (2.29)$$

For simplicity we have dropped the bars on z and r , because only surface coordinates are considered here; later they will be necessary again. Using the transform

$$\frac{\partial}{\partial x} = \cos \varphi \frac{\partial}{\partial r} - r^{-1} \sin \varphi \frac{\partial}{\partial \varphi}, \quad (2.30a)$$

$$\frac{\partial}{\partial y} = \sin \varphi \frac{\partial}{\partial r} + r^{-1} \cos \varphi \frac{\partial}{\partial \varphi}, \quad (2.30b)$$

the evaluation of the operator ∂_n from Eq. (1.79) gives the simple result

$$\partial_n \equiv \mathbf{n} \cdot \nabla = -z' \partial / \partial r + r' \partial / \partial z, \quad (2.31)$$

in which the azimuth φ does not appear. This operator commutates with the Fourier-series expansion.

On introducing Eq. (2.4) in context with Eq. (2.31) into Eq. (2.29), we soon notice that it is necessary to define the Fourier-series expansion for $\bar{c}$:

$$\bar{c}(z, r, \varphi) = \sum_{m=-\infty}^{\infty} \bar{c}_m(z, r) \exp(-im\varphi). \quad (2.32)$$

Then Eq. (2.29) separates into a sequence of boundary conditions for the different Fourier components:

$$\begin{aligned} \bar{a}(z, r)V_m(z, r) + \bar{b}(z, r) \left(r' \frac{\partial V_m}{\partial z} - z' \frac{\partial V_m}{\partial r} \right) &= \bar{c}_m(z, r) \\ (m = 0, \pm 1, \pm 2, \dots). \end{aligned} \quad (2.33)$$

These, together with the partial differential equations (2.7) for the interior, now define a *sequence* of boundary value problems that can be solved in turn, if the functions $\bar{c}_m(z, r)$ are known. The latter are obtained from the Fourier integrals

$$\bar{c}_m(z, r) = \frac{1}{2\pi} \int_0^{2\pi} \bar{c}(z, r, \varphi) \exp(im\varphi) d\varphi. \quad (2.34)$$

This general presentation of boundary conditions contains the following simple special cases:

$$\begin{aligned} \bar{a} = 1, \quad \bar{b} = 0, \quad \bar{c} = \bar{V}(r): & \text{ Dirichlet Problem,} \\ \bar{a} = 0, \quad \bar{b} = 1, \quad \bar{c} = \partial_n \bar{V}: & \text{ Neumann Problem.} \end{aligned}$$

Moreover, as the coefficients are not necessarily constants, these types can be mixed: on some surfaces a Dirichlet condition may hold, whereas on others a Neumann condition is given. Finally, the material conditions on *internal* surfaces commute with the Fourier-series expansion, because for both sides of such a surface, relations of the form (2.33) can be derived.

2.3 MAGNETIC ROUND LENSES

Generally, the use of vector potentials for numerical calculations should be avoided if possible, especially if the reduced scalar potential (Eqs. (1.29–1.31)) can be employed. There is, however, one important exception to this advice—the field calculation for magnetic round lenses. In

this case the equations, resulting from the application of the scalar potential, become ill-conditioned. On the other hand, the vector potential has only one relevant component, the azimuthal one, and the equations obtained from it are well-conditioned.

2.3.1 The Flux Potential

The basic structure of such a magnetic lens is shown in Figs. 2.3 and 2.4. The lens consists essentially of one or more rotationally symmetric coils and a round ferromagnetic yoke, which encapsulates them but has a bore for the particle beam to be focused and a gap for the generation of a magnetic field near the optic axis.

The current distribution in the circular winding is described by a vector function

$$\mathbf{j}(\mathbf{r}) = j(z, r)\mathbf{e}_\varphi(\varphi), \quad (2.35)$$

$$\mathbf{e}_\varphi(\varphi) = \mathbf{e}_y \cos \varphi - \mathbf{e}_x \sin \varphi \quad (2.36)$$

being the unit vector in the azimuthal direction. The scalar amplitude $j(z, r)$ is defined and often constant in the cross section C through the coil, whereas

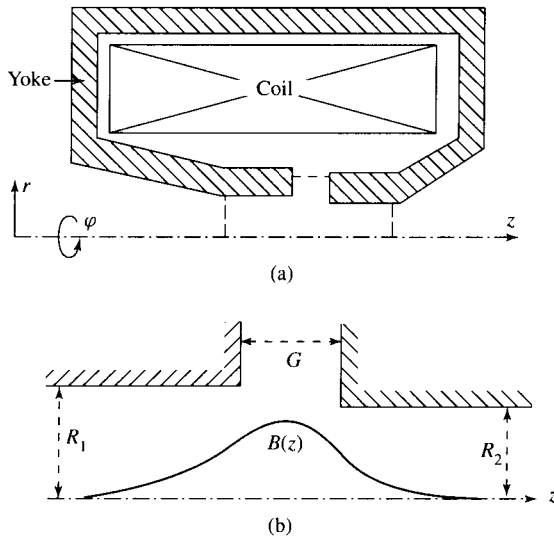


FIGURE 2.3 Schematic structure of a round magnetic lens: (a) half axial section, showing the yoke and the coil; (b) optically relevant domain near the gap with bore radii R_1 and R_2 and the gap width G . The curve $B(z)$ is the axial magnetic field strength.

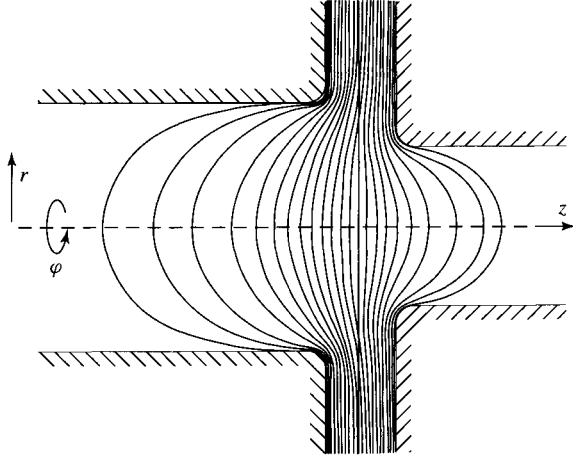


FIGURE 2.4 Approximation by a scalar magnetic potential in the vicinity of a *narrow* gap. This does *not* hold globally. The figure shows equipotentials of this scalar potential; the flux lines are perpendicular to them.

it vanishes outside. The total lens current I is here defined by

$$I = \int_C j(z, r) da. \quad (2.37)$$

We shall not need the number of winding that is often used.

The symmetry conditions for the round lens are most simplified by the assumption

$$\mathbf{A}(\mathbf{r}) = A(z, r) \mathbf{e}_\varphi(\varphi) \quad (2.38)$$

in accordance with Eq. (2.35). These vector fields already satisfy $\text{div} \mathbf{j} = 0$, $\text{div} \mathbf{A} = 0$, as can be easily verified. Any other gauge, although possible in principle, would lead to more complicated equations. The evaluation of $\mathbf{B} = \text{curl} \mathbf{A}$ in cylindrical coordinates results in $B_\varphi \equiv 0$ and

$$B_z(z, r) = \frac{\partial A}{\partial r} + \frac{A}{r} = \frac{1}{r} \frac{\partial}{\partial r}(rA(z, r)), \quad (2.39a)$$

$$B_r(z, r) = -\partial A(z, r)/\partial z. \quad (2.39b)$$

The determination of the amplitude $A(z, r)$ and its derivatives hence suffices to describe the magnetic field.

Owing to the special form (2.38) of the vector potential, its scalar amplitude A can be calculated from the axial component B_z of $\mathbf{B}$. To obtain this relation, we apply Stokes's integral theorem to a coaxial circular disc D of radius r ,

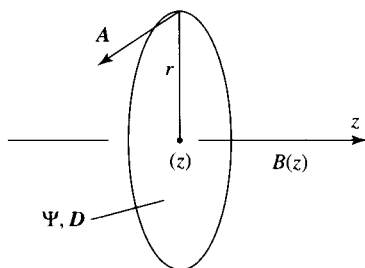


FIGURE 2.5 Magnetic flux Ψ . $\Psi(z, r)$ is the flux through a circular disc of radius r in the plane specified by the coordinate z .

located in a plane $z = \text{const}$. This z -value serves then as an argument of the potential function.

This configuration is sketched in Fig. 2.5. The evaluation of Stokes' theorem in turn with $\mathbf{B} = \text{curl} \mathbf{A}$ now gives the *flux potential*

$$\begin{aligned} \Psi(z, r) &:= \int_D \mathbf{B} \cdot \mathbf{n} \, da = 2\pi \int_0^r r' B_z(z, r') \, dr' \\ &\equiv \oint_{\partial D} \mathbf{A} \cdot d\mathbf{s} = 2\pi r A(z, r), \end{aligned} \quad (2.40)$$

which will be quite useful in numerical calculations.

Because B_z remains confined on the optic axis, the vector potential must vanish there linearly and the flux potential even in second order of r . Sometimes it may therefore be favorable to introduce a third kind of potential, which does not vanish there identically; this is given by

$$\Pi(z, r) := \frac{2}{r} A(z, r) = \frac{1}{\pi r^2} \Psi(z, r). \quad (2.41)$$

The components of $\mathbf{B}$ are then given by the equivalent differentiation formulas

$$B_z(z, r) = \frac{\partial A}{\partial r} + \frac{A}{r} = \Pi + \frac{r}{2} \frac{\partial \Pi}{\partial r} = \frac{1}{2\pi r} \frac{\partial \Psi}{\partial r}, \quad (2.42a)$$

$$B_r(z, r) = -\frac{\partial A}{\partial z} = -\frac{r}{2} \frac{\partial \Pi}{\partial z} = \frac{-1}{2\pi r} \frac{\partial \Psi}{\partial z}. \quad (2.42b)$$

Evidently, on the optic axis, the relation

$$B_z(z, 0) = \Pi(z, 0) =: B(z) \quad (2.43)$$

is valid. The function $B(z)$ is of particular importance for the electron optical properties of the lens and is sketched in Fig. 2.3.

2.3.2 Differential Equations

The partial differential equations for the potentials follow from the Maxwell equation

$$\text{curl} (v(\mathbf{r})\mathbf{B}(\mathbf{r})) = \mathbf{j}(\mathbf{r}), \quad (2.44)$$

which has only one relevant component here. Its evaluation results in

$$\frac{\partial}{\partial z} \left(v \frac{\partial A}{\partial z} \right) + \frac{\partial}{\partial r} \left(v \frac{\partial A}{\partial r} + v \frac{A}{r} \right) = -j(z, r) \quad (2.45)$$

for $A(z, r)$ and

$$\frac{\partial}{\partial z} \left(\frac{v}{r} \frac{\partial \Psi}{\partial z} \right) + \frac{\partial}{\partial r} \left(\frac{v}{r} \frac{\partial \Psi}{\partial r} \right) = -2\pi j(z, r) \quad (2.46)$$

for the flux potential Ψ . In spite of the factor $1/r$, Eq. (2.46) is often to be preferred, as it is *self-adjoint*, whereas Eq. (2.45) does not have this advantage. Also, the differential equation for $\Pi(z, r)$ is not self-adjoint.

In the case of a linear yoke material with $v = 1/\mu$, we obtain differential equations formed with the operator Δ_α of Eq. (2.13), especially

$$\Delta_\alpha \Pi(z, r) = -2\mu j(z, r)/r \quad (\alpha = 3), \quad (2.47)$$

and

$$\left(\frac{\partial^2}{\partial z^2} + \frac{\partial^2}{\partial r^2} - \frac{1}{r} \frac{\partial}{\partial r} \right) \Psi(z, r) = -2\pi r \mu j(z, r), \quad (2.48)$$

which suggests that the parameter α should be generalized to include the case $\alpha = -1$. Both forms are equally well suited, whereas in the equation for $A(z, r)$ a term A/r^2 survives.

In the case of nonlinear ferromagnetic materials the relation (1.11) is usually assumed and favorably evaluated using the flux potential, as

$$|\mathbf{B}| = \frac{1}{2\pi r} \left[\left(\frac{\partial \Psi}{\partial z} \right)^2 + \left(\frac{\partial \Psi}{\partial r} \right)^2 \right]^{1/2} \quad (2.49)$$

is the simplest form of $|\mathbf{B}|$ to be introduced into $v(z, r, |\mathbf{B}|)$.

2.3.3 Boundary Conditions

Although the magnetic field in a conventional magnetic lens is almost wholly confined by the yoke, this is not exactly true, and in open structures the far-reaching fringe fields cannot be ignored. Asymptotically, the field must behave

like that of a magnetic *dipole* oriented in the z direction. This implies that the flux must then decrease as $r^2 \cdot (r^2 + z^2)^{-3/2}$ and hence satisfy the natural boundary conditions at infinity.

The boundary conditions on the yoke surface result from Eqs. (1.81) and (1.82). To formulate them concisely, we first cast Eqs. (2.42a,b) in a convenient vector form

$$\mathbf{B}(z, r, \varphi) = \frac{1}{2\pi r} \text{grad } \Psi(z, r) \times \mathbf{e}_\varphi(\varphi). \quad (2.50)$$

This shows that the $\mathbf{B}$ -field has no azimuthal component (which we already know) and the more important fact is that it is also perpendicular to $\nabla\Psi$. Hence the equipotential lines $\Psi = \text{const.}$ are those lines to which the $\mathbf{B}$ -field is tangential; these are commonly interpreted as flux lines; this relation is *not* true for the vector potential $A(z, r)$.

It is now easy to determine components in any direction by vector operations. Let $N = (N_z, N_r)$ be any vector in the meridional plane, then we find

$$N \cdot \mathbf{B} = \frac{1}{2\pi r} N \cdot (\nabla\Psi \times \mathbf{e}_\varphi) = \frac{1}{2\pi r} \left(N_z \frac{\partial\Psi}{\partial r} - N_r \frac{\partial\Psi}{\partial z} \right). \quad (2.51)$$

In particular the choice $N = \mathbf{m}(s)$ from Eq. (2.26) gives

$$B_n = \mathbf{m} \cdot \mathbf{B} = \frac{1}{2\pi r} \left(r' \frac{\partial\Psi}{\partial r} + z' \frac{\partial\Psi}{\partial z} \right) = \frac{1}{2\pi r} \frac{\partial\Psi}{\partial s}. \quad (2.52)$$

This formula shows that the continuity of B_n is already ensured if the flux potential as a function of s is continuous, as it must be. With the tangent vector $\mathbf{t}$ from Eq. (2.25) we find

$$B_t = \mathbf{t} \cdot \mathbf{B} = \frac{-1}{2\pi r} \left(-z' \frac{\partial\Psi}{\partial r} + r' \frac{\partial\Psi}{\partial z} \right) = \frac{-1}{2\pi r} \mathbf{n} \cdot \nabla\Psi, \quad (2.53)$$

as this is just the operator given by Eq. (2.31).

The relevant condition, that $H_t = \nu B_t$ must remain continuous, now leads to the formula

$$\nu_1 (\partial_n \Psi)_1 = \nu_2 (\partial_n \Psi)_2, \quad (2.54)$$

which is the simplest possible form for this.

There are two cases of extremal boundary conditions.

1. “*Ideal*” ferromagnetic material (1): This means $\mu_1 \gg \mu_2$ or equivalently $\nu_1 \ll \nu_2$. We can then try setting $\nu_1 = 0$; hence $(\partial_n \Psi)_2 = 0$ on the vacuum side. This defines a *Neumann problem* for the outer domain. Equation (2.53)

shows that the $\mathbf{B}$ -field does not have a relevant tangential component, and this means that the flux lines are practically orthogonal to the yoke surface.

2. *Superconducting material* (1): The condition $\mathbf{B} = 0$ in the superconductor can agree with Eq. (2.50) only if $\Psi = \text{const.}$; the superconductor is hence a domain of constant flux potential. This implies that we have a *Dirichlet problem* for the outer side. Equation (2.54) becomes invalid. Instead of this, we have to consider Eq. (1.84) with $\mathbf{H}_1 = 0$. The surface currents thus obtained flow in the azimuthal direction.

2.3.4 Variational Principle

The appropriate variational principle for magnetic lenses is given by Eq. (1.50). This can be formulated in terms of $A(z, r)$ or of $\Psi(z, r)$. The two possibilities are equivalent and both have been used in practice. Because it is often necessary to consider nonlinear material properties, the function $\Lambda(z, r, B)$ of Eq. (1.54) must be used with $B = (B_z^2 + B_r^2)^{1/2}$, determined from Eqs. (2.42a,b). Hence there are the following equivalent forms of the variational principle

$$F_A = 2\pi \iint_G (\Lambda(z, r, B) - j(z, r)A(z, r)) r dz dr = \min, \quad (2.55)$$

$$F_\Psi = \iint_G (2\pi r \Lambda(z, r, B) - j(z, r)\Psi(z, r)) dz dr = \min. \quad (2.56)$$

The second form may be preferred, as Eq. (2.49) is simpler than the formula for B obtained with the vector potential. The behaviour of the integrand near the optic axis is the same in both forms: The second term vanishes within the whole bore, whereas the first one has a zero of first order:

$$2\pi r \Lambda(z, r, B) \rightarrow \pi r B^2(z)/\mu_0, \quad (2.57)$$

$B(z)$ being the field strength at the optic axis. The differential equations resulting from Eqs. (2.55) or (2.56) are (2.45) or (2.46), respectively.

2.4 SERIES EXPANSIONS

In many cases we encounter linear partial differential equations of the basic form

$$\Delta_\alpha V(z, r, \alpha) \equiv \frac{\partial^2 V}{\partial z^2} + \frac{\partial^2 V}{\partial r^2} + \frac{\alpha}{r} \frac{\partial V}{\partial r} = -g(z, r, \alpha), \quad (2.58)$$

as became obvious in the preceding considerations. Formally, even the case $\alpha = 0$ is not forbidden, we should then rename z and r as x and y and interpret these as cartesian coordinates.

In the presentation, some series expansions for the solutions of Eq. (2.58) near the z -axis will be given; these are of great importance for particle-optical calculations, as the coefficients obtained essentially determine the optical properties of a device [1,3]. Moreover, a knowledge of them is helpful for the derivation of suitable formulas for the numerical field calculation itself.

2.4.1 Symmetry Conditions

We seek *regular* solutions, because only these have physical significance. This implies that for $\alpha \neq 0$ the function

$$C(z, \alpha) = \lim_{r \rightarrow 0} r^{-1} \frac{\partial V}{\partial r} \quad (2.59)$$

must remain *finite*; hence $\partial V / \partial r$ must then vanish *linearly* on the z -axis. Moreover, all radial derivatives of higher orders must also remain finite, leading to the condition that all derivatives of *odd* orders must vanish on the optic axis. This means that V is an *even* function of r that could be rewritten as

$$V(z, r, \alpha) = \hat{V}(z, r^2, \alpha) \equiv \hat{V}(z, x^2 + y^2, \alpha). \quad (2.60)$$

The latter form demonstrates that the derivatives with respect to x and y also cause no problems: expressions like $\partial r / \partial x = x/r$, which would survive with nonvanishing odd-order derivatives, are undefined for $x = r \equiv 0$; these must not appear.

Equation (2.58) can be satisfied only if its right-hand side obeys the same symmetry condition; hence,

$$g(z, r, \alpha) = \hat{g}(z, r^2, \alpha) \equiv \hat{g}(z, x^2 + y^2, \alpha). \quad (2.61)$$

These restrictions are unnecessary for $\alpha = 0$ (planar two-dimensional fields); however, we shall still impose them for reasons of conciseness.

2.4.2 Repeated z -Differentiations

The differential equation (2.58) is inhomogeneous and linear. Therefore, according to the familiar mathematical theory, its general solution is obtained by the superposition of *one particular* solution on the general solution

$U(z, r, \alpha)$ of the corresponding *homogeneous* differential equation

$$\Delta_\alpha U(z, r, \alpha) = 0. \quad (2.62)$$

Here and also in the next section, we are concerned only with this homogeneous equation; we shall consider only *analytic* solutions, that is, those which can be differentiated arbitrarily often with respect to z and r . Since the differentiation with respect to z commutes with the radial terms, it is easy to verify that

$$\Delta_\alpha (\partial^n U / \partial z^n) = 0 \quad (2.63)$$

is true. Still, more generally, any converging series expansion is again a solution:

$$\Delta_\alpha \left(\sum_{n=0}^{\infty} A_n \partial^n U(z, r, \alpha) / \partial z^n \right) = 0. \quad (2.64)$$

The reverse step must be considered with care. To find a new solution by integration over z , we must assume that

$$W(z, r, \alpha) = \int U(z, r, \alpha) dz + F(r^2, \alpha) \quad (2.65)$$

and then determine the initially unknown function F such that $\Delta_\alpha W = 0$ is valid.

This rule is helpful for the construction of *polynomial* solutions $P_n(z, r, \alpha)$ for $\alpha \neq -1$, as can easily be verified; we find then in turn:

$$\left. \begin{aligned} P_0 &= 1 & P_1 &= z \quad (\text{start}) \\ P_2 &= \frac{z^2}{2} - \frac{r^2}{2(1+\alpha)} \\ P_3 &= \frac{z^3}{6} - \frac{zr^2}{2(1+\alpha)} \\ P_4 &= \frac{z^4}{24} - \frac{z^2 r^2}{4(1+\alpha)} + \frac{r^4}{24(1+\alpha)(1+\alpha/3)} \\ P_5 &= \frac{z^5}{120} - \frac{z^3 r^2}{12(1+\alpha)} + \frac{r^4 z}{24(1+\alpha)(1+\alpha/3)} \end{aligned} \right\} \quad (2.66)$$

and so on. The structure of the denominator is clearer in the next section.

2.4.3 Paraxial-Series Expansion

We now seek a *radial*-series expansion for the solutions of $\Delta_\alpha U = 0$. In practice this will be useful and needed only for *small* values of r , as particle beams usually remain confined to a very narrow domain close to the optic axis. This is the meaning of the notation “paraxial.” Moreover, we shall see that it becomes difficult to extend this domain for reasons of convergence, even if this seemed desirable.

As a generalization of the formulas (2.66), we can try to find a solution of Eq. (2.62) in the form

$$U(z, r, \alpha) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} F_n(z, \alpha) r^{2n}, \quad (2.67)$$

which must converge absolutely. On introduction into Eq. (2.62), differentiation within the sum and comparison of terms with equal power of r yields a recurrence relation (for $\alpha \neq -1$):

$$(1 + \alpha/(2n + 1))F_{n+1}(z, \alpha) = F_n''(z, \alpha), \quad n = 0, 1, 2, 3, \dots \quad (2.68)$$

This can be satisfied for any *analytical* start function $F_0(z, \alpha)$ having *confined* derivatives. This function is simply the *axial potential*

$$\phi(z, \alpha) := U(z, 0, \alpha). \quad (2.69)$$

We obtain then in turn:

$$F_1(z, \alpha) = (1 + \alpha)^{-1} \phi''(z, \alpha), \quad (2.70)$$

$$F_2(z, \alpha) = (1 + \alpha/3)^{-1} F_1''(z, \alpha) = (1 + \alpha)^{-1} (1 + \alpha/3)^{-1} \phi^{(4)}(z, \alpha), \quad (2.71)$$

and so on. Putting all this together, we finally arrive at the formula

$$\begin{aligned} U(z, r, \alpha) = & \phi(z, \alpha) - \frac{r^2}{2!(1 + \alpha)} \phi''(z, \alpha) + \frac{r^4 \phi^{(4)}(z, \alpha)}{4!(1 + \alpha)(1 + \alpha/3)} \\ & - \frac{r^6 \phi^{(6)}(z, \alpha)}{6!(1 + \alpha)(1 + \alpha/3)(1 + \alpha/5)} \pm \dots \end{aligned} \quad (2.72)$$

valid for $\alpha \neq -1$. The case $\alpha = -1$ is dealt with in Section 2.4.5.

The radius of convergence of this series expansion is rather limited. The reason is that the derivatives are *not* so well confined as may be desirable. This will be demonstrated by considering a simple example. Let us assume that

$$|\phi^{(n)}(z, \alpha)| \leq n! a^{-n} C, \quad (2.73)$$

C being a fixed positive constant, and a the distance from the closest singularity, generally the closest point on the boundary. For $\alpha = 1$ we then obtain

$$|U| \leq C \left(1 + \frac{r^2}{2a^2} + \frac{3r^4}{8a^4} + \frac{5r^6}{16a^6} + \dots \right) = C(1 - r^2/a^2)^{-1/2}, \quad (2.74)$$

which shows that $r < a$ is necessary. This example might be too pessimistic, as the signs in Eq. (2.72) will often alternate. However, the assumption (2.73) is not unrealistic, at least not for large values of n . A factor $(n-1)!$ instead of $n!$ does not alter the poor convergence essentially.

2.4.4 Series Expansion for the Inhomogeneous Equation

We now return to Eq. (2.58) and assume that the function $g(z, r, \alpha)$ is given and *regular*. Because it must satisfy Eq. (2.61), a series expansion of the form

$$\begin{aligned} g(z, r, \alpha) = & g_0(z, \alpha) - \frac{r^2}{2!(1+\alpha)} g_2(z, \alpha) + \frac{r^4 g_4(z, \alpha)}{4!(1+\alpha)(1+\alpha/3)} \\ & - \frac{r^6 g_6(z, \alpha)}{6!(1+\alpha)(1+\alpha/3)(1+\alpha/5)} \pm \dots \end{aligned} \quad (2.75)$$

is appropriate. The fact that the denominators correspond to those in Eq. (2.72) is not a necessity, but it facilitates the recurrence scheme finally obtained. The function $g_n(z, \alpha)$ in the numerators must be known and will generally not be derivatives of $g_0(z, \alpha)$. Evidently, $\alpha \neq -1$ is necessary.

The potential $V(z, r, \alpha)$ must obey a similar series expansion with $V_n(z, \alpha)$ instead of $g_n(z, \alpha)$. Now the coefficients $V_2, V_4, \dots$ are unknowns, to be determined from the requirement, that Eq. (2.58) must be satisfied. The result is the very simple recurrence relation

$$V_{2n+2}(z, \alpha) = V''_{2n}(z, \alpha) + g_{2n}(z, \alpha), \quad n = 0, 1, 2, 3, \dots \quad (2.76)$$

Again we can prescribe the axial potential $V_0(z, \alpha)$, and then all coefficients of higher orders are determined by solving Eq. (2.76). For the special case with $V_0 = \phi$ and $g \equiv 0$, Eq. (2.72) is again obtained, as it must be.

2.4.5 Series Expansion for the Flux Potential

As already mentioned, these series expansions cannot be used for $\alpha = -1$. This is, however, no problem, as the potential $\Pi(z, r)$, given by Eq. (2.41)

and satisfying Eq. (2.47) with $\alpha = 3$ can be used. There are practically no configurations with azimuthal currents in the paraxial domain; hence we can use the homogeneous series expansions. Considering that the axial potential is given by $B(z)$ because of Eq. (2.43), we find

$$\Psi(z, r) = \pi r^2 \left(B(z) - \frac{r^2}{8} B''(z) + \frac{r^4}{192} B^{(4)}(z) - \frac{r^6}{9216} B^{(6)}(z) \pm \dots \right). \quad (2.77)$$

Although the differential equation $\Delta_{-1}\Psi = 0$ also allows solutions of the form $\Psi = c_1 + c_2 z$, these must be excluded for reasons of regularity.

2.4.6 Fourier-Bessel Expansions

We now return to the evaluation of the homogeneous paraxial series expansion (2.72). For values $\alpha = 2m + 1$, ($m \geq 0$), this can be rewritten in compact form as

$$U(z, r, \alpha) = \sum_{n=0}^{\infty} \left(-\frac{r^2}{4} \right)^n \frac{m!}{(m+n)!} \phi^{(2n)}(z, \alpha). \quad (2.78)$$

It is now interesting to assume a harmonic axial potential

$$\phi(z, \alpha) = \phi_0 \cos(kz + \beta). \quad (2.79)$$

The evaluation of Eq. (2.78) is then quite simple and gives an absolutely converging result:

$$U(z, r, \alpha) = \phi_0 \cos(kz + \beta) \sum_{n=0}^{\infty} \left(\frac{kr}{2} \right)^{2n} \frac{m!}{(m+n)!}.$$

This is essentially the well-known series expansion of the modified Bessel function $I_m(kr)$; we can hence write

$$U(z, r, \alpha) = \phi_0 \cos(kz + \beta) m! \left(\frac{kr}{2} \right)^{-m} I_m(kr). \quad (2.80)$$

Finally, we can obtain a more general formula by linear superposition of several harmonic axial potentials in the form of a Fourier integral:

$$\phi(z, \alpha) = \int_{-\infty}^{\infty} A_m(k) \exp(ikz) dk, \quad (2.81)$$

$$U(z, r, \alpha) = m! \int_{-\infty}^{\infty} A_m(k) \exp(ikz) \left(\frac{kr}{2} \right)^{-m} I_m(kr) dk. \quad (2.82)$$

This is an alternative form of analytic continuation of the axial potential, which shifts the problem of repeated differentiations to that of Fourier integrals. The functions $x^{-m}I_m(x)$ remain *finite* as $x \rightarrow 0$ so that it is not favorable to separate the factor x^{-m} . Some applications of the use of Fourier-Bessel expansions in electron optics are reported by Plies [4], who also gives many references to other topics.

2.5 PLANAR FIELDS

Planar Fields are defined as those that are independent of one of the cartesian coordinates; without loss of generality, we can always choose the coordinate system so that this is the z -coordinate. Such fields are certainly a simplifying approximation because in reality every device is of finite extent in the z -direction, but it is presumed that the fringe fields, produced by these ends, can be ignored. This is justified, for instance, in the vicinity of the pole faces of magnetic deflection prisms, as is shown in Fig. 2.6 and some similar examples. The importance of fringing fields for the deflection of particle beams in prisms,

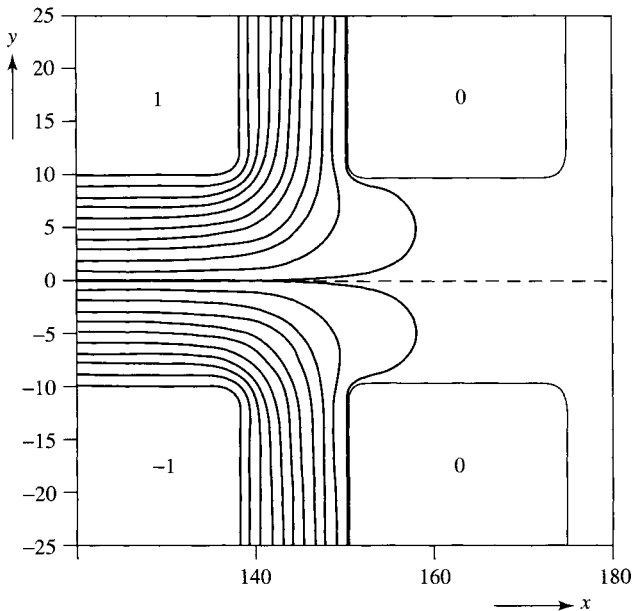


FIGURE 2.6 Cross section through a magnetic deflection prism with two yokes and two screening plates. The figure shows equipotentials of the magnetic scalar potential; the deflecting $\mathbf{B}$ field, not shown here, is perpendicular to these lines. In a confined domain, this field is practically independent of the coordinate z .

including the case of curved end faces, is discussed in Wollnik [5]. Here we shall consider planar fields mainly because they furnish mathematical tools that might later be helpful.

2.5.1 Cauchy-Riemann Equations and Conformal Mapping

Here we state only briefly the essentials of the theory of complex analytical functions with respect to field calculation.

Let $t = x + iy$ be the variable in the complex plane and

$$w(t) = \tilde{u}(t) + i\tilde{v}(t) = u(x, y) + iv(x, y) \quad (2.83)$$

any analytical function. Then outside any singularities the Cauchy-Riemann equations

$$\partial u / \partial x = \partial v / \partial y, \quad \partial v / \partial x = -\partial u / \partial y \quad (2.84)$$

are valid. From these the orthogonality relations

$$\frac{\partial u}{\partial x} \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \frac{\partial v}{\partial y} = \nabla u \cdot \nabla v = 0 \quad (2.85)$$

and the conformity relations

$$(\partial u / \partial x)^2 + (\partial u / \partial y)^2 = (\partial v / \partial x)^2 + (\partial v / \partial y)^2 \quad (2.86a)$$

or

$$|w'(t)| = |\text{grad } u(x, y)| = |\text{grad } v(x, y)| \quad (2.86b)$$

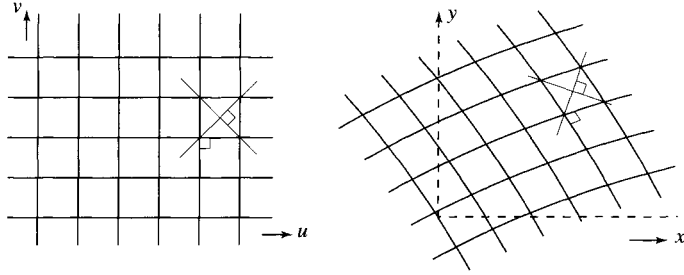
can be derived. Because the functions can be differentiated arbitrarily often in their domain of regularity, the Laplace equations

$$\Delta u = \partial^2 u / \partial x^2 + \partial^2 u / \partial y^2 = 0, \quad \Delta v = \partial^2 v / \partial x^2 + \partial^2 v / \partial y^2 = 0 \quad (2.87)$$

are obtained by differentiating Eq. (2.83).

The function $w(t)$ of Eq. (2.83) can be interpreted as a mapping of the (u, v) plane into the (x, y) plane and vice versa. This is sketched in Fig. 2.7 and has some interesting properties:

- (i) All angles between intersecting lines remain unaltered.
- (ii) Consequently the *ratio* of scales in different directions, but at the same point, remains unaltered. The local magnification factor is just $|w'(t)|$.
- (iii) The inverse transformation has the same essential properties.

FIGURE 2.7 Conformal mapping of an u - v plane into an x - y plane and vice versa.

These are the characteristics of a conformal mapping. As solutions of Laplace's equation the analytic functions can be used to describe planar electrostatic or magnetostatic fields without spatial sources. Moreover, they can supply suitable coordinate systems for better fitting of the boundary conditions of other fields, which do not necessarily satisfy a planar Laplace equation.

With respect to magnetic fields, there are some interesting physical interpretations of the Cauchy-Riemann equations. First of all, we can rewrite Eq. (1.26) as

$$B_x(x, y) = \mu_0 \partial W(x, y) / \partial x, \quad B_y = \mu_0 \partial W / \partial y. \quad (2.88)$$

Next, we can introduce a vector potential $A(x, y)$ having only a z -component, whereupon $\text{div } A = 0$ is identically satisfied. The evaluation of $\mathbf{B} = \text{curl } A$ then gives

$$B_x = \partial A(x, y) / \partial y, \quad B_y = -\partial A / \partial x. \quad (2.89)$$

These two representations must be in agreement, and from Eq. (2.84), we see that this is just the case if we define a complex potential

$$w(t) = \tilde{u}(t) + i\tilde{v}(t) = \mu_0 W(x, y) + iA(x, y) \quad (2.90)$$

as an analytical function of $t = x + iy$. The field components must be combined in *conjugate* form as

$$\tilde{B}(t) = B_x(x, y) - iB_y(x, y) \equiv w'(t) \quad (2.91)$$

to agree with the complex derivative. Similar considerations can be carried out for electrostatic fields, but these are not given here as they do not give anything essentially new.

The theory of analytic functions can be applied only to the inner or outer regions of source-free domains with homogeneous material properties. As soon as points on a boundary are considered, Eq. (2.84) become invalid, as

every boundary carries electromagnetic *sources*. Then a real representation must be used.

2.5.2 Basic Analytical Functions

In this section some elementary functions will be presented, and this will later be useful in various ways.

Complex Powers

$$w(t) = t^n = (x + iy)^n = r^n \exp(in\varphi), \quad (2.92a)$$

$$u(r, \varphi) = r^n \cos n\varphi, \quad v(r, \varphi) = r^n \sin n\varphi. \quad (2.92b)$$

These appear quite frequently and have already been used in the step leading from Eq. (2.7) to Eq. (2.10). The power n may also become *fractional* and the functions are consequently then *singular* at $t = 0$. Such a function must be made unique by introducing a *cut* starting at the singularity and running to infinity. This cut must not be crossed by any integration loop, and outside it the function is then analytic in the familiar sense.

An example of physical relevance is the field in the vicinity of a sharp corner, as shown in Fig. 2.8. In this example, a Dirichlet condition $U = U_0 = \text{const.}$ is to be satisfied on the surface of this wedge and in its interior. The cut can be located in the interior, where it then does not affect the calculation. These conditions are satisfied by a series expansion

$$U(r, \varphi) = U_0 + \sum_{m=1}^{\infty} a_m r^{\mu m} \sin(\mu m \varphi) \quad (2.93a)$$

with the fraction

$$\mu = \pi/(2\pi - \gamma). \quad (2.93b)$$

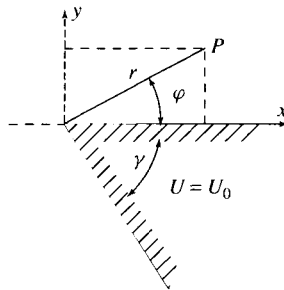


FIGURE 2.8 Cartesian and polar coordinates of an arbitrary point P in the vicinity of a wedge with angle γ .

The coefficients are to be determined from boundary conditions at other surfaces. The essential conclusion from this result is that for $\gamma < \pi$ the radial derivative $\partial U/\partial r$ becomes singular as $r^{\mu-1}$ with $\mu - 1 < 0$. For a rectangular corner, $\gamma = \pi/2$, this gives a singularity as $r^{-1/3}$.

Complex Logarithm and Exponential

The logarithm is a singular function that must be made unique by a cut running from the origin to infinity, usually along the line $y = 0$, $x \leq 0$. It is most favorably represented in polar coordinates (r, φ) and then given by

$$\log(t) = \log(x + iy) = \ln r + i\varphi. \quad (2.94)$$

The function

$$g(r) = -\frac{1}{2\pi} \ln r = -\frac{1}{4\pi} \ln(x^2 + y^2) \quad (2.95)$$

is the potential of an infinitely long straight wire carrying unit “charge” per unit length interval, and as such it will later become of importance for the purposes of singularity analysis.

Another interesting possibility is the introduction of an exponentially increasing radial coordinate, $r = \exp(\rho)$. Then the inverse of the function $\log(t)$ becomes

$$t = x + iy = \exp(\rho + i\varphi), \quad (2.96a)$$

$$x = e^\rho \cos \varphi, \quad y = e^\rho \sin \varphi. \quad (2.96b)$$

This describes a conformal mapping with cylindrical symmetry, as is shown in Fig. 4.3. This is certainly helpful if essential parts of the boundaries are circular arcs. It is interesting that Laplace’s equation becomes simpler:

$$r^2 \partial^2 V / \partial r^2 + r \partial V / \partial r + \partial^2 V / \partial \phi^2 = \partial^2 V / \partial \rho^2 + \partial^2 V / \partial \phi^2 = 0. \quad (2.97)$$

There are many other elementary functions, which are not discussed here for reasons of space.

2.5.3 Analytic Continuation

Sometimes, the problem of determining the field in the vicinity of the optic axis from the potential on the axis itself arises, as in Section 2.4.3; the main difference from the former considerations is that the optic axis now lies in the x -direction. Such a configuration is given, for instance, in cylindric lenses with slit electrodes, as shown in Fig. 2.9. We can use the formulas of Section 2.4.3

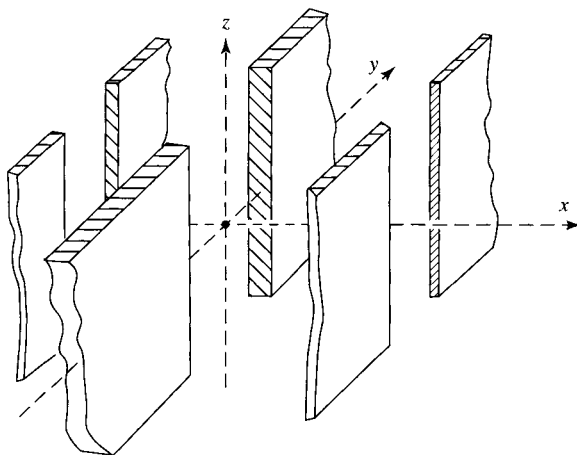


FIGURE 2.9 Perspective sketch of a symmetric cylindrical *three-slit* lens. The term “cylindrical” means that this lens has a focusing effect only in one direction (y). (The slit length in the z -direction should be much longer.)

after making the replacements $z \rightarrow x$, $r \rightarrow y$, $\alpha \rightarrow 0$, whereupon the series expansion for a *positively* symmetric potential becomes

$$U(x, y) = \sum_{n=0}^{\infty} \frac{1}{(2n)!} (-y^2)^n \phi^{(2n)}(x). \quad (2.98)$$

The analytic continuation by a Fourier integral is also fairly simple: with

$$\phi(x) = \int_{-\infty}^{\infty} A(k) \exp(ikx) dk, \quad (2.99a)$$

the result is given by

$$U(x, y) = \int_{-\infty}^{\infty} A(k) \exp(ikx) \cosh(ky) dk. \quad (2.99b)$$

It is interesting that *antisymmetric* potentials such as that for the configuration in Fig. 2.6 and other cylindric deflection systems can also be reduced to this scheme. In such cases, we have $V(x, 0) = 0$, but the derivative $F(x) := \partial V / \partial y$ for $y = 0$ is of physical significance. We obtain a symmetric function $U(x, y)$ by setting

$$V(x, y) = xU(x, y). \quad (2.100)$$

The ordinary planar Laplace equation for $V(x, y)$ then leads to $\Delta_\alpha U = 0$ with $\alpha = 2$. The paraxial series expansion now becomes

$$U(x, y) = \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} (-y^2)^n F^{(2n)}(x), \quad (2.101)$$

and the continuation by a Fourier integral takes the form

$$F(x) = \int_{-\infty}^{\infty} B(k) \exp(ikx) dk, \quad (2.102a)$$

$$U(x, y) = \int_{-\infty}^{\infty} B(k) \exp(ikx) \frac{\sin h(ky)}{ky} dk. \quad (2.102b)$$

We have thus found a number of useful relations, which will be needed later on at various occasions.

REFERENCES

1. Hawkes, P. W. and Kasper, E. (1989). *Principles of Electron Optics*, Volume 1, Chapters 6 and 7, London: Academic Press.
2. Franzen, N. (1984). Computer programs for analyzing certain classes of 3-D electrostatic fields with two planes of symmetry. In *Electron Optical Systems*, J. J. Hren et al., eds., pp. 115–126, Scanning Electron Microscopy, Chicago.
3. Ximen Jiye (1986). Aberration theory in electron and ion optics, *Advances in Electronics and Electron Physics*, Suppl. 17, London & New York: Academic Press.
4. Plies, E. (1994). Electron optics of low-voltage electron beam testing and inspection, Part I: Simulation tools, *Advances in Optical and Electron Microscopy*, **13**: 123–242.
5. Wollnik, H. (1987). *Optics of Charged Particles*, Orlando: Academic Press.

This Page Intentionally Left Blank

CHAPTER III

Basic Mathematical Tools

In this chapter, we shall present several formulas that are well known in numerical mathematics and that are proven there in detail. It is, however, necessary to give them here too, as they will later be referred to in various contexts and the notation employed must be specified. For reasons of space, we shall not reproduce the proofs, which can be found in all comprehensive mathematical textbooks.

3.1 ORTHOGONAL COORDINATE SYSTEMS

Orthogonal curvilinear coordinate systems play an important role in the solution of partial differential equations by means of the *separation method*. Although this method is familiar and is in widespread use in theoretical physics, for instance, in quantum mechanics, we shall *not* discuss it here in detail, but shall present some brief examples later. The reason is that it is hardly possible to find orthogonal coordinate systems that fit *all* boundaries of a realistic technical device. It is, however, possible to find coordinate systems that are better adapted to the configuration than cartesians: an example is the cylindric coordinate system that we have already used frequently in the preceding chapter.

3.1.1 Line Element and Lamé Coefficients

As nonorthogonal coordinate systems will soon be excluded, the formalism of co- and contravariant coordinate systems does not bring an essential gain and is therefore not used in the subsequent presentation.

Let q_1, q_2, q_3 be three independent coordinates that are defined by their law of transformation of the cartesian system x_1, x_2, x_3 :

$$x_i = x_i(q_1, q_2, q_3), \quad i = 1, 2, 3. \quad (3.1)$$

(For conciseness, we shall not distinguish here the names of the coordinates and their corresponding transformation functions, although this is not quite correct. In addition, we assume tacitly that all postulated derivatives and

integrals do really exist). The coordinate differentials are then given by

$$dx_i = \sum_{k=1}^3 \frac{\partial x_i}{\partial q_k} dq_k, \quad i = 1, 2, 3, \quad (3.2)$$

and the line element is then

$$ds^2 := \sum_{i=1}^3 (dx_i)^2 = \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 \frac{\partial x_i}{\partial q_j} \frac{\partial x_i}{\partial q_k} dq_j dq_k. \quad (3.3)$$

This can be rewritten in the familiar form

$$ds^2 = \sum_{j=1}^3 \sum_{k=1}^3 g_{jk} dq_j dq_k, \quad (3.4)$$

whereupon a comparison of equal coefficients gives the elements of the *metric tensor*

$$g_{jk} = g_{kj} = \sum_{i=1}^3 \frac{\partial x_i}{\partial q_j} \frac{\partial x_i}{\partial q_k} \quad (3.5)$$

as functions of q_1 , q_2 , and q_3 .

Orthogonal coordinate systems are defined by the postulate that all off-diagonal elements must vanish identically; hence

$$g_{jk} = g_{kk} \delta_{jk} = L_k^2 \delta_{jk}, \quad (3.6)$$

where δ_{jk} denotes the Kronecker symbol. The Lamé coefficients, defined as the square roots of the diagonal elements, are then

$$L_k(q_1, q_2, q_3) := \sqrt{g_{kk}} = \left[\sum_{i=1}^3 (\partial x_i / \partial q_k)^2 \right]^{1/2}; \quad (3.7)$$

hence, the line element now becomes

$$ds^2 = \sum_{k=1}^3 ds_k^2 = \sum_{k=1}^3 (L_k dq_k)^2, \quad (3.8)$$

which makes it appropriate to interpret the terms

$$ds_k = L_k(q_1, q_2, q_3) dq_k \quad (k = 1, 2, 3) \quad (3.9)$$

as length differentials along coordinate lines, which are such lines on which only one coordinate varies. From these, the transformation of the volume element

$$dv = dx_1 dx_2 dx_3 = J dq_1 dq_2 dq_3 \quad (3.10)$$

with the Jacobian determinant

$$J(q_1, q_2, q_3) := \frac{\partial(x_1, x_2, x_3)}{\partial(q_1, q_2, q_3)} = L_1 L_2 L_3 \quad (3.11)$$

can be derived and easily understood.

Surfaces that are obtained by keeping one coordinate q_i constant will be referred to succinctly as coordinate surfaces. The *area elements* on such surfaces are given by

$$da_i = L_j L_k dq_j dq_k \quad (i \neq j \neq k \neq i) \quad (3.12)$$

with cyclic permutations of the labels.

3.1.2 Vectors in Curvilinear Coordinates

In general mathematical theory, it is usual to distinguish between co- and contravariant basic vectors, but here we shall not do this and we introduce *normalized* vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$. To distinguish them from the cartesian basis, we shall denote the latter by $\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3$ in this section.

The basic vectors can be obtained in two different but equivalent ways:

- (i) by differentiation of the vector $\mathbf{r} = (x_1, x_2, x_3)$ as a function of q_1, q_2, q_3 according to Eq. (3.1) and normalization:

$$\mathbf{e}_j(q_1, q_2, q_3) = \frac{\partial \mathbf{r}}{\partial q_j} \bigg/ \left| \frac{\partial \mathbf{r}}{\partial q_j} \right|, \quad j = 1, 2, 3, \quad (3.13)$$

- (ii) as the gradient of coordinate surfaces $q_j(\mathbf{r}) = \text{const.}$ and normalization:

$$\mathbf{e}_j(x_1, x_2, x_3) = \text{grad } q_j / |\text{grad } q_j|, \quad j = 1, 2, 3. \quad (3.14)$$

The arguments of these vectors as functions are different, but the vectors themselves are the same at the same point in space. By construction, these vectors satisfy the orthonormality relations $\mathbf{e}_j \cdot \mathbf{e}_k = \delta_{jk}$. The presentation of any other vector $\mathbf{v}$ is now simply

$$\mathbf{v} = \sum_{j=1}^3 v_j \mathbf{e}_j = \sum_{k=1}^3 \hat{v}_k \mathbf{i}_k. \quad (3.15a)$$

with the components

$$v_j = \mathbf{v} \cdot \mathbf{e}_j, \quad \hat{v}_k = \mathbf{v} \cdot \mathbf{i}_k. \quad (3.15b)$$

3.1.3 Differential Forms

From Eq. (3.9) in the context of Eq. (3.2), it becomes clear that the Nabla operator transforms as

$$\nabla = \sum_{k=1}^3 \mathbf{i}_k \frac{\partial}{\partial x_k} = \sum_{j=1}^3 \mathbf{e}_j \frac{\partial}{\partial s_j} = \sum_{j=1}^3 L_j^{-1} \mathbf{e}_j \frac{\partial}{\partial q_j} \quad (3.16)$$

and consequently the gradient of any differentiable function $V(q_1, q_2, q_3)$ as

$$\text{grad } V = \sum_{j=1}^3 L_j^{-1} \mathbf{e}_j \partial V / \partial q_j. \quad (3.17)$$

With the representation

$$\mathbf{n} = \sum_{k=1}^3 n_k \mathbf{e}_k \quad (3.18)$$

of any normalized vector, the corresponding normal derivative becomes

$$\partial_n V \equiv \mathbf{n} \cdot \nabla V = \sum_{k=1}^3 L_k^{-1} n_k \partial V / \partial q_k. \quad (3.19)$$

So far, the considerations are straightforward. Anticipating later necessary discretizations, however, we have also to consider the situation in which the derivatives have to be approximated by divided finite differences of function values. This is extensively discussed in Chapter IV, which is devoted to the method of finite differences; here we discuss some elementary aspects.

In a grid formed by coordinate surfaces and coordinate lines, the components of the gradient are most frequently needed along these lines. For example, a line with varying coordinate q_1 will be chosen here and hence q_2 and q_3 are kept constant. For a small but finite distance h_1 the first component of the gradient is

$$\langle D_1 \rangle V \equiv \langle \mathbf{e}_1 \cdot \nabla \rangle V = h_1^{-1} L_1^{-1} (V_{1+} - V_{1-}) = \mathbf{e}_1 \cdot \nabla V + O(h_1^2), \quad (3.20a)$$

with the two function values

$$V_{1\pm} = V(q_1 \pm h_1/2, q_2, q_3). \quad (3.20b)$$

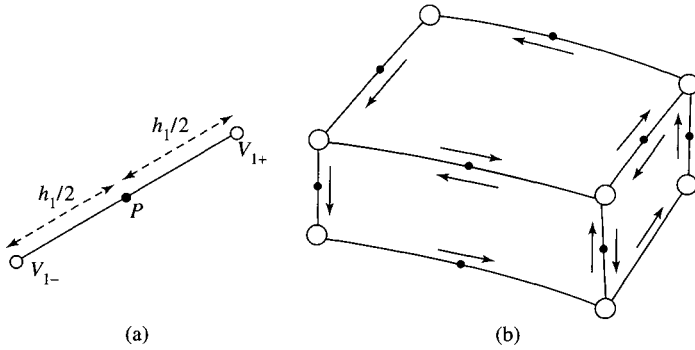


FIGURE 3.1 Discretization of the gradient: (a) in each direction, the discretized projection refers to the midpoint P ; (b) in each orthogonal volume element, whose edges are sampling points of the potential, the line integral over the approximated gradient vanishes for any surface.

Note that this mean value does not refer to one of the two end points but to the *midpoint* between these (see Fig. 3.1a,b); this ensures that the approximation error is of *second* order. This approximation is also in agreement with the familiar midpoint rule of integration:

$$\begin{aligned} V_{1+} - V_{1-} &= \int_{P_-}^{P_+} \nabla V \cdot d\mathbf{s} = \int_{P_-}^{P_+} \mathbf{e} \cdot \nabla V ds \\ &= h_1 L_1 \langle P_1 \rangle V + O(h_1^3). \end{aligned} \quad (3.21)$$

Analogous formulas hold for the other components. It is then obvious that for all closed loops, formed by such elementary integrals along coordinate lines, the relation

$$\oint \langle \mathbf{t} \cdot \text{grad} \rangle V ds = 0 \quad (3.22)$$

is satisfied in this approximation, $\mathbf{t}$ being the local tangent, as each function value appears twice with opposite signs. The idea of evaluating the derivatives at positions different from those of the potentials was first introduced by Weiland [1] and is summarized in Section 3.1.6.

The curl operator can also be introduced by a finite approximation of integrals and their limits, and this is even simpler and more instructive than a direct evaluation by means of the product rule.

The way the derivative is obtained is shown in Fig. 3.2. Let $\mathbf{e}$ be any normalized vector and A any small area element perpendicular to it with positively traced circumference ∂A . Then the component of $\text{curl } \mathbf{v}$ in the direction

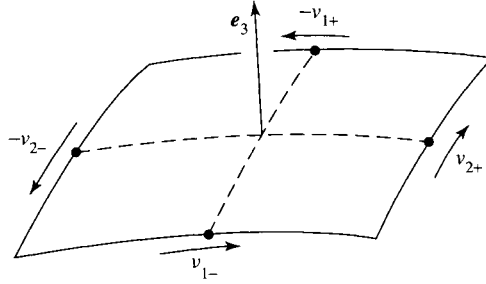


FIGURE 3.2 Discretization of the curl by an oriented ring integral. The result is the component in the direction of $\mathbf{e}_3$ and refers to the midpoint. It is compatible with Fig. 3.1b.

of $\mathbf{e}$ is approximated by

$$\langle \mathbf{e} \cdot \text{curl} \rangle \mathbf{v} = \frac{\int \mathbf{e} \cdot \text{curl} \mathbf{v} \, da}{\int da} = \frac{1}{A} \oint_{\partial A} \mathbf{v} \cdot d\mathbf{s}. \quad (3.23)$$

In the limit of vanishing diameter of the element, this mean value converges to $\mathbf{e} \cdot \text{curl} \mathbf{v}$. For a more detailed evaluation of this expression, we take a rectangular element with midpoint (q_1, q_2, q_3) in the coordinate surface $q_3 = \text{const.}$; hence $\mathbf{e} = \mathbf{e}_3$. The sides of the rectangle are located on coordinate lines at distances $\pm h_1/2$, and $\pm h_2/2$, (see Fig. 3.2), and the vector field is evaluated at the four midpoints of these lines. The integrations can then be carried out by means of the midpoint-rule by analogy with Eq. (3.21). We then obtain for the right-hand part of the integral

$$\int_{P_{+,-}}^{P_{+,+}} \mathbf{v} \cdot d\mathbf{s} = h_2 L_2 \left(q_1 + \frac{h_1}{2}, q_2, q_3 \right) v_2 \left(q_1 + \frac{h_1}{2}, q_2, q_3 \right) + O(h_2^3), \quad (3.24)$$

$v_2 = \mathbf{e}_2 \cdot \mathbf{v}$ being the tangential component at the midpoint. The other three sides of the rectangle are to be treated analogously. If we abbreviate the preceding expression by $h_2 L_{2+} v_{2+}$ and consider the correct sign of integration on the other side, we can cast the result in the convenient form

$$\oint_{\partial A} \mathbf{v} \cdot d\mathbf{s} = -h_1 (L_{1+} v_{1+} - L_{1-} v_{1-}) + h_2 (L_{2+} v_{2+} - L_{2-} v_{2-}). \quad (3.25a)$$

According to Eq. (3.12), the area A appearing in the denominator of Eq. (3.23) can be approximated in the two equivalent forms

$$A = h_1 h_2 L_1 L_2 = h_1 h_2 (L_{1+} + L_{1-})(L_{2+} + L_{2-})/4 + O(h^4), \quad (3.25b)$$

where the factors L_1 and L_2 refer to the center of the rectangle. Putting these expressions together, we finally arrive at the approximation

$$\langle \mathbf{e}_3 \cdot \text{curl} \rangle \mathbf{v} = \frac{L_{2+}v_{2+} - L_{2-}v_{2-}}{h_1 L_1 L_2} - \frac{L_{1+}v_{1+} - L_{1-}v_{1-}}{h_2 L_1 L_2} + O(h^2). \quad (3.26)$$

The result of proceeding to the limits $h_1 \rightarrow 0$, $h_2 \rightarrow 0$ is now obviously the well-known formula

$$\langle \mathbf{e}_3 \cdot \text{curl} \rangle \mathbf{v} = \frac{1}{L_1 L_2} \left[\frac{\partial}{\partial q_1} (L_2 v_2) - \frac{\partial}{\partial q_2} (L_1 v_1) \right]. \quad (3.27)$$

The result for the other two components are obtained in an analogous manner and can be written down readily by cyclic permutation of the labels, if the basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ is positively oriented as we assume here. The discretized operators have one interesting and important property: if we assume $\mathbf{v} = \text{grad } V$, approximate this in the sense of Eq. (3.21), and introduce this into Eq. (3.25a), we find that this integral vanishes *exactly*. These operations can be performed for the other components. The result can be cast in the concise symbolic form

$$\langle \text{curl} \rangle \langle \text{grad} \rangle \equiv 0. \quad (3.28)$$

The corresponding infinitesimal identity $\text{curl grad } V \equiv 0$ is also satisfied by the given expressions.

The divergence operator is approximated as a mean value by volume-integration over a finite box B and transformed by means of Gauss's integral theorem (see Fig. 3.3):

$$\langle \text{div} \rangle \mathbf{v} = \frac{\int_B \text{div } \mathbf{v} d\tau}{\int_B d\tau} = \frac{1}{V_B} \oint_{\partial B} \mathbf{n} \cdot \mathbf{v} da. \quad (3.29)$$

(The volume element of integration is here denoted by $d\tau$ and the full volume of the box by V_B). To facilitate the integrations, we choose a rectangular box with side lengths h_1, h_2 , and h_3 in the q -space, the reference point being located in its center. The volume of the box is then approximated by

$$V_B = J h_1 h_2 h_3 + O(h^5), \quad (3.30)$$

the Jacobian J being given by Eq. (3.11) and referring to the center. The surface integral consists now of six integrals over the side faces, and the normal $\mathbf{n}$ is always to be taken in outward direction. The evaluation will

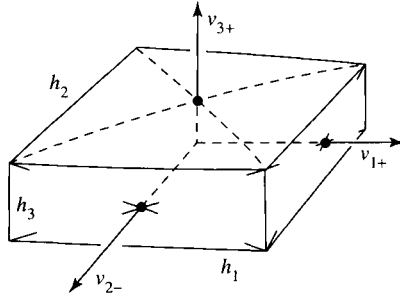


FIGURE 3.3 Discretization of the divergence by an oriented surface integral. The vector function is to be evaluated at the surface midpoints, and the result then refers to the volume midpoint.

again be demonstrated by considering one of these, the rectangle R_1^+ , in the positive q_1 -direction for instance. This integral is again evaluated by means of the midpoint rule and gives

$$J_1^+ = \int_{R_1^+} v_1 \left(q_1 + \frac{h_1}{2}, q_2, q_3 \right) da_1 = h_2 h_3 L_2^+ L_3^+ v_1^+ + O(h^4), \quad (3.31)$$

in which the superfix $+$ denotes the evaluation at the rectangle midpoint $(q_1 + h_1/2, q_2, q_3)$. The remaining five surface integrals are evaluated analogously and the total integral becomes

$$\begin{aligned} \oint_{\partial B} \mathbf{n} \cdot \mathbf{v} da &= h_2 h_3 (L_2^+ L_3^+ v_1^+ - L_2^- L_3^- v_1^-) + h_3 h_1 (L_3^+ L_1^+ v_2^+ - L_3^- L_1^- v_2^-) \\ &\quad + h_1 h_2 (L_1^+ L_2^+ v_3^+ - L_1^- L_2^- v_3^-) + O(h^5). \end{aligned} \quad (3.32)$$

If we now divide by the volume V_B (Eq. (3.30)), we obtain the result

$$\langle \operatorname{div} \mathbf{v} \rangle = J^{-1} \{ (L_2^+ L_3^+ v_1^+ - L_2^- L_3^- v_1^-) / h_1 + \text{cycl.} \} + O(h^2), \quad (3.33)$$

the abbreviation cycl. denoting the cyclic permutation, as in Eq. (3.32). In the limit of vanishing box size this formula clearly converges to the well-known form

$$\operatorname{div} \mathbf{v} = J^{-1} \left\{ \frac{\partial}{\partial q_1} (L_2 L_3 v_1) + \text{cycl.} \right\}. \quad (3.34)$$

This can more concisely be rewritten as

$$\operatorname{div} \mathbf{v} = J^{-1} \sum_{k=1}^3 \frac{\partial}{\partial q_k} (J v_k / L_k). \quad (3.35)$$

It is easy to verify that $\text{div curl } \mathbf{w} \equiv 0$ for any twice continuously differentiable vector field $\mathbf{w}(q_1, q_2, q_3)$, as it must be. Moreover, it is important that the corresponding rule also holds exactly for the discretized form. To verify this, we have to know the longitudinal components of the function $\mathbf{w}(q_1, q_2, q_3)$ at the midpoints of the 12 edge lines, for example $w_{1+} = w_1(q_1, q_2 + h_2/2, q_3 + h_3/2)$ and so on. Then each surface integral

$$\int \mathbf{n} \cdot \mathbf{v} \, da = \int \mathbf{n} \cdot \text{curl } \mathbf{w} \, da = \oint \mathbf{w} \cdot d\mathbf{s} \quad (3.36)$$

can be evaluated by means of the discretized curl formulas. In this context, it becomes clear that the *first* form of the area-approximation in Eq. (3.25b) is the correct one, because the factors $L_1 L_2$ in Eq. (3.26) now appear as function values $L_1^+ L_2^+$ or $L_1^- L_2^-$ at the midpoints of the corresponding rectangle in the last term of Eq. (3.32) and cancel out. Each line integral appears twice with opposite signs and thus cancels out. Hence we finally obtain

$$\langle \text{div} \rangle \langle \text{curl} \rangle \equiv 0.$$

Altogether this means that the set of approximations is *self-consistent* and never causes an error of less than second order.

3.1.4 Differential Forms of Second Order

We consider now a self-adjoint linear differential equation, like Eq. (1.16)

$$\text{div}(\varepsilon(\mathbf{r}) \text{grad } V(\mathbf{r})) + \rho(\mathbf{r}) = 0. \quad (3.37)$$

According to Eq. (3.17), the vector $\varepsilon \text{grad } V$ has the components

$$v_k := (\varepsilon \text{grad } V)_k = \varepsilon L_k^{-1} \partial V / \partial q_k \quad (k = 1, 2, 3) \quad (3.38)$$

in curvilinear coordinates. Introducing this into Eq. (3.35), we find

$$J^{-1} \sum_{k=1}^3 \frac{\partial}{\partial q_k} \left(J \varepsilon L_k^{-2} \frac{\partial V}{\partial q_k} \right) + \rho = 0. \quad (3.39a)$$

This is again of the self-adjoint form

$$\sum_k \frac{\partial}{\partial q_k} \left(\bar{\varepsilon}_k \frac{\partial V}{\partial q_k} \right) + \bar{\rho} = 0, \quad (3.39b)$$

with the transformed coefficients

$$\bar{\varepsilon}_k = J \varepsilon L_k^{-2}, \quad (k = 1, 2, 3), \quad \bar{\rho} = \rho J. \quad (3.40)$$

Evidently, the isotropy is lost if the Lamé coefficients become unequal.

Since the domains of integration are now no longer necessary, as we have already evaluated integrals like those in Eq. (3.29), we can now directly use the scheme of discrete points and simplify the notation as shown in Fig. 3.4. The central point will have the label 0 and the neighboring points in the negative coordinate direction will have the labels 4, 5, and 6, respectively. The gradient and the material coefficient ε are to be evaluated at the *midpoints* between adjacent grid points (see Fig. 3.1a); these are now given the labels $1', \dots, 6'$. With $h_4 = h_1$, Eq. (3.33) is now rewritten as

$$\langle \text{div} \rangle v = J_0^{-1} \{ L_2' L_3' v_{1'} / h_1 - L_5' L_6' v_{4'} / h_4 + \text{cycl.} \}. \quad (3.41)$$

According to Eq. (3.21), the discretized components of the vector $v = \varepsilon \text{ grad } V$ now become

$$v_{1'} = \varepsilon_{1'} (V_1 - V_0) h_1^{-1} L_1^{-1}, \quad v_{4'} = -\varepsilon_{4'} (V_4 - V_0) h_4^{-1} L_4^{-1}. \quad (3.42)$$

After adaptation of the notation, and with $L_2' L_3' = J_{1'}/L_{1'}$, $L_5' L_6' = J_{4'}/L_{4'}$, we notice that it is possible to express the final discretization formula in a very compact form:

$$\sum_{k=1}^6 \frac{J_{k'} \varepsilon_{k'}}{J_0 L_{k'}^2 h_k^2} (V_k - V_0) = -\rho_0 + O(h^2). \quad (3.43)$$

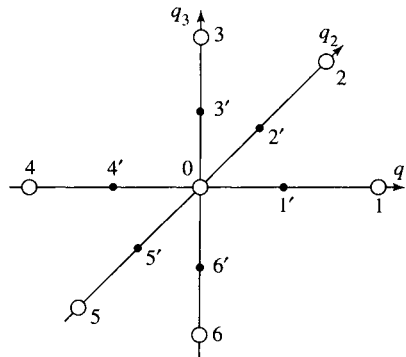


FIGURE 3.4 Simplified notation for a spatial seven-point configuration: the potentials refer to the points $0, \dots, 6$, vector components and material coefficients to the midpoints $1', \dots, 6'$. See also Fig. 3.1.

The coefficients in all six directions now have the same mathematical form and are positive, because $\varepsilon \leq 0$ is not allowed for elliptic equations. This is already a basic formula in the method of finite differences. The latter however, goes, beyond this derivation, as it has also to consider the way in which a global solution can be found.

From the general Eq. (3.39a), the transformation of the Laplace operator with $\varepsilon \equiv 1$ is now simplified as follows:

$$\Delta V \equiv \nabla^2 V = J^{-1} \sum_{k=1}^3 \frac{\partial}{\partial q_k} \left(J L_k^{-2} \frac{\partial V}{\partial q_k} \right). \quad (3.44)$$

This equation will be needed in the following section. It would also be possible and cause no essential problem to write down the transform of expressions like $\text{curl}(\nu(\mathbf{r})\mathbf{A}(\mathbf{r}))$ and the corresponding discretization. We shall not do this here as the resulting expressions would become rather unwieldy, and it seems better to leave this task to the computer.

The following section deals with one important unfamiliar example of a curvilinear coordinate system for which the actual evaluation is possible. The results for the familiar systems—cylindrical and spherical coordinates—can be found in any collection of mathematical formulas.

3.1.5 The Surface-Adapted Coordinate System

We return now to Eqs. (2.24–2.27) and the vectors $\mathbf{t}$ and $\mathbf{m}$ as shown in Fig. 2.2. Because the curvilinear coordinates s and φ served to describe a rotationally symmetric surface in parametric form, it is natural to complete this system by a third coordinate w in normal direction. In the simplest case, this is the distance from the surface, so that $w = 0$ is the equation of the surface itself where as for $w \neq 0$, the sign w indicates the side on which the corresponding point is located (see Fig. 3.5).

We now define the following labeling:

$$x_1 = x, \quad x_2 = y, \quad x_3 = z; \quad q_1 = w, \quad q_2 = s, \quad q_3 = \varphi \quad (3.45)$$

and the transformation

$$z = \bar{z}(s) + w \bar{r}'(s), \quad (3.46a)$$

$$r = \bar{r}(s) - w \bar{z}'(s), \quad (3.46b)$$

$$x = r \cos \varphi, \quad y = r \sin \varphi, \quad (3.46c)$$

which demonstrates simultaneously its close relation to the familiar cylindrical coordinates. After a short calculation, it turns out that this system is orthogonal,

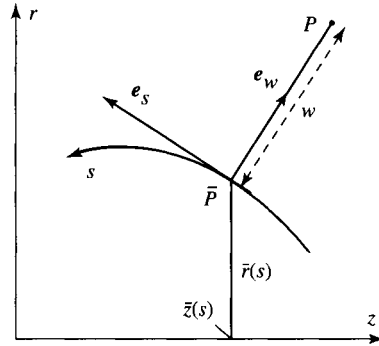


FIGURE 3.5 Definition of the surface-adapted coordinate system. The figure shows the coordinates of an arbitrary point P and its associated foot point $\bar{P}$ in a meridional section. For the position shown, the coordinate w has the positive sign.

and the Lamé coefficients become

$$L_w \equiv 1, \quad L_\varphi = r, \quad L_s = 1 + \kappa(s)w, \quad (3.47)$$

$$\kappa(s) = \bar{z}'(s)\bar{r}''(s) - \bar{z}''(s)\bar{r}'(s) \quad (3.48)$$

being the *curvature* at position s on the boundary curve. Because $L_s > 0$ is necessary, this means that the domain of use of this coordinate system is restricted. On the convex side ($\kappa w > 0$) there is no limitation, but on the concave side the coordinate w must satisfy $|w| < |\kappa|^{-1}$. However, in the vicinity of sharp edges difficulties will arise. The main purpose is the derivation of relations for a field in the vicinity of a boundary, that is, for $w = 0$ or quite small values of $|w|$.

The basic vectors of this system become the surface normal $\mathbf{e}_w = \mathbf{n}(s, \phi)$ from Eq. (2.27), the common azimuthal unit vector $\mathbf{e}_\varphi$, and the *rotated* meridional tangent $\mathbf{e}_s$. Altogether the cartesian components of these vectors are given by

$$\mathbf{e}_w = (-\bar{z}'(s) \cos \varphi, -\bar{z}'(s) \sin \varphi, \bar{r}'(s)), \quad (3.49a)$$

$$\mathbf{e}_s = (\bar{r}'(s) \cos \varphi, \bar{r}'(s) \sin \varphi, \bar{z}'(s)), \quad (3.49b)$$

$$\mathbf{e}_\varphi = (-\sin \varphi, \cos \varphi, 0). \quad (3.49c)$$

(We recall that $\bar{z}^2 + \bar{r}^2 \equiv 1$). In the preceding equations, this basis of vectors is positively oriented. According to Eq. (3.17) the formula for a gradient is now explicitly

$$\text{grad } V = \mathbf{e}_w \frac{\partial V}{\partial w} + L_s^{-1} \mathbf{e}_s \frac{\partial V}{\partial s} + r^{-1} \mathbf{e}_\varphi \frac{\partial V}{\partial \varphi} \quad (3.50)$$

and the curl of a vector function

$$\mathbf{v}(w, s, \varphi) = v_w \mathbf{e}_w + v_s \mathbf{e}_s + v_\varphi \mathbf{e}_\varphi. \quad (3.51)$$

As a result of Eq. (3.27) and its cyclic permutations,

$$\begin{aligned} \text{curl } \mathbf{v} = & \frac{\mathbf{e}_w}{rL_s} \left(\frac{\partial}{\partial s}(rv_\varphi) - \frac{\partial}{\partial \varphi}(L_s v_s) \right) + \frac{\mathbf{e}_s}{r} \left(\frac{\partial}{\partial \varphi}v_w - \frac{\partial}{\partial w}(rv_\varphi) \right) \\ & + \frac{\mathbf{e}_\varphi}{r} \left(\frac{\partial}{\partial w}(L_s v_s) - \frac{\partial}{\partial s}v_w \right). \end{aligned} \quad (3.52)$$

For a simple example, we may consider the vector $\mathbf{v} = \mathbf{A}(\mathbf{r})$ of Eq. (2.38) for a round magnetic field and the associated flux potential ψ defined by Eq. (2.40). The evaluation of $\mathbf{B} = \text{curl } \mathbf{A}$ with Eq. (3.52) then results in $B_\varphi \equiv 0$ and

$$B_w \equiv B_n = (2\pi r L_s)^{-1} \partial \Psi / \partial s, \quad (3.53a),$$

$$B_s \equiv B_t = -(2\pi r)^{-1} \partial \Psi / \partial w, \quad (3.53b)$$

which is a generalization of Eqs. (2.52 and 2.53) for points on the vicinity of the yoke surface. With $J = rL_s$ the formula for the divergence of a vector field can be written as

$$rL_s \text{div } \mathbf{v} = \frac{\partial}{\partial w}(rL_s v_w) + \frac{\partial}{\partial s}(rv_s) + L_s \frac{\partial}{\partial \varphi}v_\varphi. \quad (3.54)$$

The self-adjoint differential Eq. (3.39) now becomes

$$\frac{\partial}{\partial w} \left(rL_s \varepsilon \frac{\partial V}{\partial w} \right) + \frac{\partial}{\partial s} \left(\frac{r\varepsilon}{L_s} \frac{\partial V}{\partial s} \right) + \frac{\partial}{\partial \varphi} \left(\frac{\varepsilon L_s}{r} \frac{\partial V}{\partial \varphi} \right) = -rL_s \rho. \quad (3.55)$$

In one simple but frequently occurring case, this representation leads to a useful conclusion. Let us consider the electrostatic field in the source-free vacuum domain between rotationally symmetric electrodes; we have then $\rho \equiv 0$ and $\varepsilon = \varepsilon_0 = \text{const.}$, and there is no dependence on φ , so that Eq. (3.55) simplifies to

$$\frac{\partial}{\partial w} \left(rL_s \frac{\partial V}{\partial w} \right) + \frac{\partial}{\partial s} \left(\frac{r}{L_s} \frac{\partial V}{\partial s} \right) = 0. \quad (3.56)$$

On the surface of the conducting electrodes ($w = 0$), the potential is constant and hence $\partial V / \partial s = 0$; thus

$$\frac{\partial}{\partial w} \left(rL_s \frac{\partial V}{\partial w} \right) = 0 \quad (\text{at } w = 0). \quad (3.57)$$

The differentiation of this product, recalling Eqs. (3.46b) and (3.47), finally results in

$$\frac{\partial^2 V}{\partial w^2} = \left(\frac{\bar{z}'(s)}{\bar{r}(s)} - \kappa(s) \right) \frac{\partial V}{\partial w} \quad (\text{at } w = 0). \quad (3.58)$$

This means that at the surfaces of such electrodes the derivative of second order is simply proportional to the field strength.

Sometimes the *inverse* transformation to the one given in Eq. (3.46) is needed. In spite of the simplicity of Eq. (3.46), this problem may become rather complicated and can have multiple solutions, as it implies that the *footpoint* on a surface must be determined for a given point in space. The calculation of footpoints can always be performed numerically, but this topic is not suitable for analytic considerations and we therefore omit it here.

3.1.6 The Discretization of Maxwell's Equations

Maxwell's equations are an important example of the application of the method of discretization outlined earlier. This procedure needs to be completed by an approximation of derivatives with respect to time; it is appropriate to define this again as a central finite difference according to

$$[\dot{f}(\mathbf{r}, t)] := \tau^{-1} \left(f\left(\mathbf{r}, t + \frac{\tau}{2}\right) - f\left(\mathbf{r}, t - \frac{\tau}{2}\right) \right) + O(\tau^2) \quad (3.59)$$

for any differentiable function $f(\mathbf{r}, t)$. Then we can immediately write

$$\begin{aligned} \langle \text{curl} \rangle \mathbf{E} &= -[\dot{\mathbf{B}}] & \langle \text{div} \rangle \mathbf{D} &= \rho, \\ \langle \text{curl} \rangle \mathbf{H} &= \mathbf{j} + [\dot{\mathbf{D}}] & \langle \text{div} \rangle \mathbf{B} &= 0 \end{aligned} \quad (3.60)$$

together with the material equations. This set of approximations is self-consistent if the variables in it are evaluated at *different* positions, as shown in Fig. 3.6. If the nodes in a —not necessarily cubic— grid are denoted by triples of integers (m, n, l) , then the following locations are to be chosen:

nodes $(m, n, l,)$: ρ , potential V

side midpoints $(m + 1/2, n, l), (m, n + 1/2, l), (m, n, l + 1/2)$:

polar vectors $\mathbf{E}, \mathbf{D}, \mathbf{A}, \mathbf{j}$, coefficients ε, κ

face midpoints $(m + 1/2, n + 1/2, l), (m, n + 1/2, l + 1/2),$

$(m + 1/2, n, l + 1/2)$: *axial vectors* $\mathbf{B}, \mathbf{H}$, coefficient μ

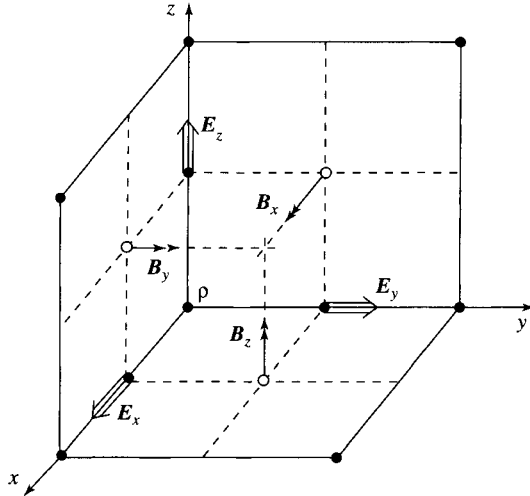


FIGURE 3.6 Dislocated discretization of vector fields: scalars (potential, charge density) refer to the nodes of a spatial grid, components of ‘polar’ vectors ($\mathbf{E}, \mathbf{j}, \mathbf{A}$) to the side midpoints, and components of “axial” vectors ($\mathbf{B}$) to the face midpoints. These are distinguished by different vector symbols.

spatial midpoint $(m + 1/2, n + 1/2, l + 1/2)$:

magnetic potential W , (if used).

The self-consistent application of the central finite differences with respect to time implies that not all quantities can refer simultaneously to the same value t . If we assume that the system will be solved in equidistant time-steps $t_k = k\tau$ with integer k , then the following coordination is to be used:

integer steps $k\tau$: $V, \rho, \mathbf{E}, \mathbf{D}$ (electric)

half integer $(k + 1/2)\tau$: $\mathbf{B}, \mathbf{H}, \mathbf{A}, \mathbf{j}, (W)$ (magnetic).

With this choice, the continuity equation

$$(\text{div})\mathbf{j} + [\dot{\rho}] = 0 \quad (3.61)$$

is consistently discretized. However, the material equation $\mathbf{j} = \kappa\mathbf{E}$ does *not* fit this system. This relation, causing attenuations, requires special considerations that are not discussed here.

The idea of this kind of approximation was introduced first by Weiland [1] and implemented in a big computer program called *MAFIA* [2]. It is obvious that a concept of such generality requires a very large memory to accommodate

the numerous field data and should therefore be used only in big and fast computer workstations. The special techniques needed there are beyond the scope of this volume, and we refer here to the corresponding literature [2].

3.2 INTERPOLATION AND NUMERICAL DIFFERENTIATION

In the practice of numerical calculations, we frequently encounter functions for which an analytical law is not known or is very complicated. Quite often the whole function is initially unknown and has to be determined by a complicated solution procedure. Typical examples are the calculation of potentials by means of the finite-element method or the calculation of surface charges by the boundary-element method. The characteristic of numerical methods that distinguish them from analytical ones is that all such problems must be *discretized*, that is, reduced to a *finite set* of data and that interpolations between these data become necessary. We therefore examine interpolation techniques and numerical differentiation as they are both related.

A full study of these topics would fill a volume of its own and would go well beyond the scope of this book; we therefore confine our considerations to a necessary minimum and refer to specialized textbooks for more details. The reader, who is interested in discussion of numerical techniques in combination with short programs in FORTRAN or C, is referred to *Numerical Recipes* [3].

3.2.1 Basic Rules for Interpolation

The general problem can be described in the following way: Let $x_0, x_1, \dots, x_N$ be a set of $N + 1$ *sampling* coordinates; without loss of generality we can assume that this array is ordered in a strictly monotonic sequence: $x_i > x_{i-1}$ for $i \leq N$. We consider now a function $y(x)$, for which the sampling values $y_i := y(x_i)$ ($i = 0, \dots, N$) are given. The problem is then to find a function $y(x)$ that satisfies these conditions.

A general solution can be found in the form

$$y(x) = \sum_{i=0}^N y_i T_i(x), \quad (3.62)$$

where $T_i(x)$ being called the *trial or test functions* or *interpolation kernels*. Evidently these have to satisfy the conditions

$$T_i(x_k) = \delta_{ik} \quad (i, k = 0, \dots, N), \quad (3.63)$$

and this is already sufficient to meet the earlier described requirements. However, the results may still be unacceptable, if the functions oscillate strongly between the sampling points. A natural condition that should be satisfied by general interpolation routines is that these must reproduce any constant exactly; hence

$$\sum_{i=0}^N T_i(x) \equiv 1. \quad (3.64)$$

As long as these requirements are satisfied, the choice of the trial functions is free, and it remains up to us to make an optimal choice. This, however, depends on the characteristics of the actual data set; varieties of different methods are hence in use.

Among these functions, two main classes are of particular interest: trigonometric functions and polynomials. The trigonometric interpolation leads to a discrete form of Fourier analysis and synthesis; this cannot be discussed here for reasons of space.

The choice of polynomials is certainly to be preferred, if N is quite small, say $N < 4$. The trial functions are then the Lagrange polynomials

$$T_i(x) = \prod_{k=0}^N{}' [(x - x_k)/(x_i - x_k)], \quad (3.65)$$

the prime signifying that the factor for $k = i$ must be skipped. In spite of its closed form, this formula is unfavorable, owing to the rather numerous operations necessary to evaluate it. There are equivalent and more economic algorithms—the familiar ones found by Newton and Neville, for example; these can be found in any textbook on numerical mathematics.

The formulas for $N = 2$ are given here in an alternative form, which also furnishes the finite difference approximations for the derivatives that will be required more often. For conciseness we introduce the mesh lengths

$$a = x_1 - x_0, \quad b = x_2 - x_1, \quad (3.66)$$

and then have

$$y_1' = \frac{a^2(y_2 - y_1) + b^2(y_1 - y_0)}{ab(a + b)}, \quad (3.67)$$

$$\frac{1}{2}y_1'' = \frac{a(y_2 - y_1) - b(y_1 - y_0)}{ab(a + b)}. \quad (3.68)$$

By means of these expressions, it is then quite easy to evaluate $y(x)$ and $y'(x)$ at any position $x \neq x_1$. For $a = b$ these formulas simplify to

$$2y'_1 = (y_2 - y_0)/a, \quad y''_1 = (y_2 - 2y_1 + y_0)/a^2. \quad (3.69)$$

These were the approximations made in Section 3.1.3.

The *general scheme* for interpolation and numerical differentiation using the Neville algorithm can be cast in the combined form

$$p_{i,0}^{(m)} = y_i \delta_{i,0} \quad (i = 0, 1, \dots, N; m = 0, \dots, M) \quad (\text{start}) \quad (3.70)$$

for $(m = 0 \text{ to } M \text{ step } 1)$

$\left[\text{for } (i = 1 \text{ to } N \text{ step } 1)$

$\left\{ \text{for } (k = 1 \text{ to } i \text{ step } 1)$

$$p_{ik}^{(m)} = p_{i-1,k-1}^{(m)} + \frac{x - x_{i-k}}{x_i - x_{i-k}} (p_{i,k-1}^{(m)} - p_{i-1,k-1}^{(m)}) \\ + \frac{m}{x_i - x_{i-k}} (p_{i,k-1}^{(m-1)} - p_{i-1,k-1}^{(m-1)}) \quad \left. \right\} \quad (3.71)$$

$$y^{(m)}(x) = p_{N,N}^{(m)} \quad (m = 0, \dots, M). \quad (3.72)$$

Here superscripts denote the order of differentiation and M is hence the highest order. It can be proved by induction that this recurrence scheme satisfies all conditions for interpolation. Yet, in spite of its general validity and relative simplicity, it must be applied with great care, for it is easy to reach unreasonable results. To achieve an acceptable accuracy, the sampling points should be selected in such a way that the reference abscissa x remains centered as well as possible. The recurrence scheme can then, of course, not always start with the label 0, but with a label $n \geq 0$, and we have to replace i and k by $i + n$ and $k + n$, respectively. With $N = 9$, this number n is to be determined in such a way that $x_{n+4} \leq x < x_{n+5}$ holds. In the vicinity of the margins, where this is not possible, a deterioration cannot be avoided.

A consequence of this technique is that as the abscissa x increases from x_0 toward its maximum, the selection of neighboring points necessarily alters, and the interpolation curve is therefore not a unique polynomial but a combination of several of them, some kind of spline. At the sampling points the function values $y(x)$ remain continuous, but not the derivatives; their discontinuities become larger with increasing order. Moreover, their accuracy soon becomes poor, so that it is not reasonable to go far beyond $M = 2$. Alternative techniques are hence necessary.

3.2.2 Hermite Interpolation

The disadvantages mentioned earlier can be avoided by means of the so-called *Hermite interpolation*. This technique uses function values and derivatives upto order $M > 0$ at the endpoints of each interval, to determine the interpolation polynomial and its derivatives in its interior. Since all these derivatives are the same at the junction between neighboring intervals, the global interpolation function is M -times continuously differentiable by construction.

Because the general theory is not very familiar, it is outlined here in some detail. We consider an interval $x_a \leq x \leq x_b$ with midpoint $x_c = (x_a + x_b)/2$ and half-interval-size $h = (x_b - x_a)/2$. It is favorable to introduce a normalized coordinate

$$u = (x - x_c)/h \equiv (2x - x_a - x_b)/(x_b - x_a), \quad (3.73)$$

with $|u| \leq 1$. The condition that all derivatives up to order M are prescribed for $u = \pm 1$, is most easily matched by a polynomial of the odd degree $2M + 1$. This polynomial and its derivatives can be cast in the convenient form

$$y^{(n)}(x) = \sum_{m=0}^M h^{m-n} [A_m^{(n)}(u)y^{(m)}(x_a) + B_m^{(n)}(u)y^{(m)}(x_b)], \quad (3.74)$$

the derivatives of y referring to x and those of A_m and B_m to u . These coefficients are called *the normalized form functions*.

The boundary conditions now take the standard form

$$A_m^{(n)}(-1) = \delta_{m,n}, \quad A_m^{(n)}(+1) = 0, \quad (3.75a)$$

$$B_m^{(n)}(-1) = 0, \quad B_m^{(n)}(+1) = \delta_{m,n}, \quad (3.75b)$$

$$(0 \leq m \leq M, \quad 0 \leq n \leq N),$$

because powers of h cancel out by the transform from x to the variable u . If we replace u by $-u$, then the two boundaries and consequently the coefficients A_m and B_m are exchanged. Under this mirror operation, all even derivatives are invariant, whereas all order derivatives change their sign. This implies

$$A_m(\mp u) = (-1)^m B_m(\pm u) \quad (m = 0, 1, \dots, M). \quad (3.76)$$

All these requirements can be satisfied with a set of *symmetric* functions $S_{m,M}(u)$ and *antisymmetric* functions $T_{m,M}(u)$ if we set

$$T_{m,M}(u) = \dot{S}_{m+1,M+1}(u) \quad (3.77a)$$

$$A_m(u) = \frac{1}{2}(S_{m,M}(u) - T_{m,M}(u)) \quad (3.77b)$$

$$B_m(u) = \frac{1}{2}(S_{m,M}(u) + T_{m,M}(u)) \cdot (-1)^m. \quad (3.77c)$$

The T -functions are derivatives of the S -functions in the *next higher order* $M + 1$, and it is therefore necessary to use a second label here.

All the boundary conditions in Eq. (3.75) are simultaneously satisfied with

$$S_{m,M}^{(n)}(-1) = \delta_{m,n}, \quad S_{m,M}^{(n)}(+1) = (-1)^m \delta_{m,n}, \quad (3.78)$$

as can be easily verified because the T -functions obey the same conditions after appropriate index shift. To determine these functions $S_{m,M}(u)$ it is favorable to introduce new basic functions

$$p_k(u) := 2^{-k}(1 - u^2)^k/k! \quad (k \geq 0), \quad (3.79)$$

which are symmetrix in u and have zeroes of order k on both sides. Since $S_{m,M}(u)$ must have zeroes of order m , a power-series expansion

$$S_{m,M}(u) = \sum_{k=m}^M U_{mk} p_k(u) \quad (3.80)$$

is appropriate. Generally, the coefficients should also have a label M , but it will become clear that they do not depend on it; hence $S_{m,M}$ depends on M only over the length of the series expansion.

It suffices now to satisfy the normalized boundary conditions in Eq. (3.78) only on one side, say $u = -1$, as the symmetry properties imply that they are then simultaneously satisfied on the other side. The repeated differentiations of the functions $p_k(u)$, necessary for this task, can be cast in the form

$$\begin{aligned} \dot{p}_k &= -u p_{k-1} \\ \ddot{p}_k &= -p_{k-1} + u^2 p_{k-2} \\ \dddot{p}_k &= 3u p_{k-2} - u^3 p_{k-3} \end{aligned} \quad (3.81)$$

and so on, and in general,

$$p_k^{(n)} = (-u)^n \left[p_{k-n} - \left(\frac{n}{2}\right) u^{-2} p_{k-n+1} + 3 \left(\frac{n}{4}\right) u^{-4} p_{k-n+2} \pm \dots \right] \quad (3.82)$$

or ordered in reversed sequence

$$p_k^{(n)} = \sum_{j=\alpha}^{\beta} (-1)^n p_{k-j} L_{nj} u^{2j-n} \quad (3.83)$$

with

$$\alpha = \text{int}[(n+1)/2], \quad \beta = \min(n, k)$$

$$L_{nj} = (-1)^{n-j} \left(\frac{n}{2n-2j} \right) (2n-2j+1)!! \quad (3.84)$$

for $\alpha \leq j \leq n$, 0 else. These numbers form a *lower* triangular matrix, the subdiagonal part of the matrix **C**, given by Eq. (3.87). At the boundaries $u = \pm 1$, only the term with $j = k$ gives a nonvanishing contribution; hence,

$$p_k^{(n)}(-1) = L_{nk}, \quad p_k^{(n)}(+1) = (-1)^n L_{nk}. \quad (3.85)$$

Introducing this into Eq. (3.80) and then into Eq. (3.78), we obtain a linear system of equations for the coefficients U_{mk}

$$\sum_{k=m}^M U_{mk} L_{nk} = \delta_{m,n} \quad (3.86a)$$

or in matrix notation with the unit matrix **I**:

$$\mathbf{U}\mathbf{L}^+ = \mathbf{I}. \quad (3.86b)$$

This is solved by $\mathbf{U} = (\mathbf{L}^+)^{-1}$, which gives an *upper* triangular matrix. The fact that the matrix **U** is triangular has the consequence that increasing its rank to $M+1$ does not affect the elements for which both indices are less than M . With respect to Hermite-interpolation this has the advantage that increasing the accuracy by considering polynomials and boundary-derivatives of higher order does not require the recalculation of the polynomials already determined.

Because both triangular matrices have unit diagonal elements, it is possible to combine them without conflict into one coefficient matrix $\mathbf{C} = (\mathbf{L}\mathbf{U})$, which for $M = 6$ takes the form

$$\mathbf{C} = \left(\begin{array}{c|cccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 3 & 15 & 105 & 945 \\ 0 & -1 & 1 & 3 & 15 & 105 & 945 \\ 0 & 0 & -3 & 1 & 6 & 45 & 420 \\ 0 & 0 & 3 & -6 & 1 & 10 & 105 \\ 0 & 0 & 0 & 15 & -10 & 1 & 15 \\ 0 & 0 & 0 & -15 & 45 & -15 & 1 \end{array} \right) \quad (3.87)$$

Note that the labeling starts with zero. In practical programs the first row and column that have no relevant information apart from $C_{00} = 1$ can be omitted. The remaining matrix then has positive indexing, as usual, and is of rank 6.

This matrix enables us to calculate all Hermite-interpolation polynomials up to degree 11, which are quite accurate enough for practical purposes. All the polynomials of lower degrees are also contained in this scheme, and we shall now discuss the most frequently appearing cases.

$$M = 0$$

With $U_{00} = U_{11} = 1$, $U_{01} = 0$ we obtain from Eqs. (3.79) and (3.80)

$$S_{00} = 1, \quad S_{11} = p_1 = (1 - u^2)/2, \quad \dot{S}_{11} = T_{00}(t) = -u, \quad (3.88)$$

and from Eq. (3.77)

$$A_0(u) = (1 - u)/2, \quad B_0(u) = (1 + u)/2, \quad (3.89)$$

and finally with Eq. (3.73)

$$y(x) = [(x_b - x)y_a + (x - x_a)y_b]/(x_b - x_a). \quad (3.90)$$

This is the familiar linear interpolation, the lowest possible approximation.

$$M = 1$$

We now obtain with the abbreviation $p(u) := p_1(u)$

$$\begin{aligned} S_{01} &= 1, & S_{11} &= p, & S_{12} &= p + p_2, & S_{22} &= p_2 = p^2/2, \\ T_{01} &= \dot{S}_{12} = -u(1 + p), & T_{11} &= \dot{S}_{22} = -up \end{aligned} \quad (3.91)$$

From these, we obtain the form functions

$$\begin{aligned} A_0(u) &= (1 - u)/2 - pu/2, & A_1(u) &= p(1 - u)/2, \\ B_0(u) &= (1 + u)/2 + pu/2, & B_1(u) &= -p(1 + u)/2, \end{aligned} \quad (3.92)$$

which are depicted in Fig. 3.7a.

This *cubic* Hermite interpolation is the most frequently used technique, especially with the application of cubic *splines*, which will be outlined in the next section.

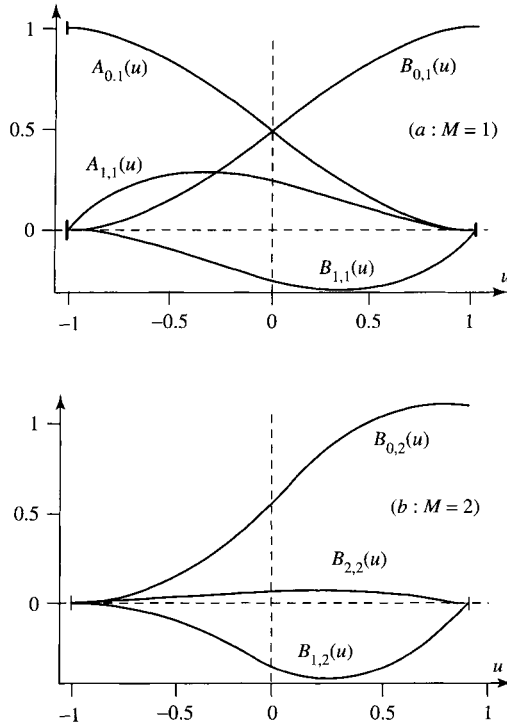


FIGURE 3.7 Form functions of Hermite interpolation: (a) cubic order; (b) quintic order; in this case only the half referring to the right-hand side is shown.

$M = 2$

Here we shall present only the final results and use the notation $A^- = A$, $A^+ = B$ to combine the formulas; we then find:

$$\left. \begin{aligned} A_0^\pm &= (1 \pm u)/2 \pm u(p + 3p_2)/2 \\ A_1^\pm &= \mp(1 \pm u)p/2 + (\pm 1 - 3u)p_2/2, \\ A_2^\pm &= (1 \pm u)p_2/2. \end{aligned} \right\} \quad (3.93)$$

Comparison with Eq. (3.92) shows that the terms without p_2 have indeed not altered. The functions referring to the right-hand side ($-$) are shown in Fig. 3.7b. The other ones are obtained by mirror operations as in Fig. 3.7a.

Approximation Error

We assume here that the marginal derivatives are correct. Then the dominant error is given by a polynomial of degree $2M + 2$ with zeroes of order $M + 1$

on both sides. This polynomial is given by

$$\varepsilon_M(u) = \frac{h^{2M+2}}{(2M+2)!} (1-u^2)^{M+1} f^{(2M+2)}(\xi), \quad (3.94)$$

where ξ being an abscissa near the midpoint x_c . Evidently the largest error occurs at the midpoint x_c . This formula demonstrates that with high orders and *confined* derivatives the interval can be made fairly large. For instance with $h = 1$, $x_a = -1$, $x_b = 1$, $M = 5$ and $y(x) = \exp(x)$ we obtain $|\varepsilon| < 2 \cdot 10^{-9}$.

Another example: $x_a = 0$, $x_b = \pi/2$, $M = 4$ for $y = \cos x$ gives $|\varepsilon| < 2 \times 10^{-8}$. Such high accuracies are lost if the marginal derivatives are not accurate.

3.2.3 Hermite Splines

Generally, spline functions are defined as interpolation functions that have a separate definition in each interval and are joined together at the interval ends as smoothly as possible. Most frequently, polynomials are used but trigonometric or exponential functions are also suitable. In this sense the polynomials of Section 3.2.1 are already splines but not very smooth ones, because even the derivatives of first order may become slightly discontinuous.

The most popular and very favorable kind of splines is the *cubic* Hermite splines, which we shall now deal with briefly: these use the cubic Hermite polynomials, as given by Eq. (3.92). Traditionally, the cubic splines are described by a tridiagonal system of equations for the *second-order* derivatives that are derived from a variational principle. Here we shall *not* follow this route but derive such a system for the *first-order* derivatives that are equivalent, but easier to evaluate.

The form Eq. (3.92) can be differentiated twice and then $y''(x)$ evaluated according to Eq. (3.74). The results of this elementary calculation are the familiar formula

$$y''(x_c) = [y'(x_b) - y'(x_a)]/(2h), \quad (3.95)$$

and more interesting

$$\begin{aligned} y''(x_a) &= 1.5[y(x_b) - y(x_a)]/h^2 - [2y'(x_a) + y'(x_b)]/h, \\ y''(x_b) &= -1.5[y(x_b) - y(x_a)]/h^2 + [2y'(x_b) + y'(x_a)]/h. \end{aligned} \quad (3.96)$$

A check for the accuracy of the calculation is that the value of $y''(x_c) = [y''(x_a) + y''(x_b)]/2$ must give Eq. (3.95) as $y''(x)$ is a linear function in cubic approximation.

We adopt as before the notation of Section 3.2.1 and consider two adjacent intervals $[x_{i-1}, x_i]$ and $[x_i, x_{i+1}]$. Then the derivative $y''(x)$ can be evaluated

in both intervals for the common node x_i . Using the abbreviations

$$q_i := (x_i - x_{i-1})^{-1}, \quad q_{i+1} := (x_{i+1} - x_i)^{-1}, \quad (3.97)$$

this gives the two values

$$\begin{aligned} y''_{i(-)} &= -6q_i^2(y_i - y_{i-1}) + 2q_i(2y'_i + y'_{i-1}), \\ y''_{i(+)} &= 6q_{i+1}^2(y_{i+1} - y_i) - 2q_{i+1}(2y'_i + y'_{i+1}). \end{aligned} \quad (3.98)$$

Unless y'_{i-1} , y'_i , and y'_{i+1} do not satisfy a certain condition, as given here, the values $y''_{i(-)}$ and $y''_{i(+)}$ will be different. Therefore, a natural postulate is their equality or $y''_{i(-)} - y''_{i(+)} = 0$. This can be rewritten as

$$\begin{aligned} q_i y'_{i-1} + 2(q_i + q_{i+1})y'_i + q_{i+1}y'_{i+1} &= 3q_i^2(y_i - y_{i-1}) + 3q_{i+1}^2(y_{i+1} - y_i), \\ (i = 1, 2, \dots, N-1). \end{aligned} \quad (3.99)$$

This equation is the earlier-mentioned tridiagonal system of equations, as it links each of the three subsequent unknown derivatives.

The system is to be completed by equations for the marginal derivatives y'_0 and y'_N . These choices are analogous for both endpoints; hence we shall confine the discussion to the endpoint at $x = x_0$. The most frequent choices are the following:

- (i) y'_0 is *prescribed*, for example as the end slope of another function to which the spline must be matched, or simply $y'_0 = 0$ at a local extremum
- (ii) *Inflection* point or “natural” boundary condition: $y''_{i(-)} = 0$ then gives (also formally with $q_0 = 0$ in Eq. (3.99))

$$q_1(2y'_0 + y'_1) = 3q_1^2(y_1 - y_0). \quad (3.100)$$

It is unfavorable to drop a factor q_1 , as this destroys the symmetry of the tridiagonal matrix

- (iii) *Periodic* spline: this means

$$y'_{N+k} = y'_k, \quad y_{N+k} - y_N = y_k - y_0, \quad (3.101)$$

for any positive or negative labels k . The rule for y_k is a generalization—a periodicity with *heap*. The elimination of all terms with negative labels or with labels $> N$ results in

$$\begin{aligned} q_1 y'_N + 2(q_1 + q_2)y'_1 + q_2 y'_2 &= 3q_1^2(y_1 - y_0) + 3q_2^2(y_2 - y_1), \\ q_N y'_{N-1} + 2(q_N + q_1)y'_N + q_1 y'_1 &= 3q_1^2(y_1 - y_0) + 3q_N^2(y_N - y_{N-1}). \end{aligned} \quad (3.102)$$

If these equations are used as the first and last ones, respectively, the matrix has the rank N and remains symmetric; it has become cyclically tridiagonal.

The system matrix, obtained here, has the same favorable properties as the familiar one; it is symmetric, positive definite, and diagonally dominant; hence the well-known solution techniques such as the elimination algorithm can be applied. Moreover, this unfamiliar formulation of the spline technique has the advantage that the results are immediately the boundary values (y_i, y'_i) ($i = 0, \dots, N$) needed for the application of cubic Hermite polynomials, as outlined in the earlier section.

The concept can be extended to Hermite splines of the fifth order. The technique is the same: the polynomials must be differentiated three or four times; the derivatives are evaluated at the margins and are required to remain continuous at the transition to the neighboring interval. The corresponding calculations are not given here for reasons of space. The requirement of continuous third-order derivatives results in

$$\begin{aligned} -q_i y''_{i-1} + 3(q_i + q_{i+1})y''_i - q_{i+1}y''_{i+1} &= 20q_i^3(y_{i-1} - y_i) \\ + 20q_{i+1}^3(y_{i+1} - y_i) + 8q_i^2(1.5y'_i + y'_{i-1}) - 8q_{i+1}^2(1.5y'_i + y'_{i+1}). \end{aligned} \quad (3.103)$$

Now all first-order derivatives can be prescribed, and the solution of Eq. (3.103) then yields the consistent second-order derivatives, whereupon the technique of quintic Hermite polynomials can be started. This procedure is straightforward.

It is possible to go further and to determine the first-order derivatives by the additional demand that the fourth order should also be continuous; this results in a system

$$\begin{aligned} 7q_i^3 y'_{i-1} + 8(q_i^3 + q_{i+1}^3)y'_i + 7q_{i+1}^3 y'_{i+1} &= 15q_i^4(y_i - y_{i-1}) \\ + 15q_{i+1}^4(y_{i+1} - y_i) + q_i^2(1.5y''_i - y''_{i-1}) + q_{i+1}^2(y''_{i+1} - 1.5y''_i). \end{aligned} \quad (3.104)$$

These systems are to be completed by formulas for y'_0, y'_N and y''_0, y''_N . These can be derived in the same way as the former cases by assuming a given boundary value or by setting $q_0 = 0$ or by assuming cyclic properties. These formulas are not given here.

In principle, this concept can be carried on to arbitrarily high orders with rapidly increasing difficulties. Thus for practical reasons a spline of seventh order with prescribed derivatives up to second order at the sampling points is an upper limit. The demand that even the fourth order should remain continuous leads to the following tridiagonal system of equations:

$$\begin{aligned}
q_i y_{i-1}''' + 4(q_i + q_{i+1}) y_i''' + q_{i+1} y_{i+1}''' &= 15q_i^2 (-y_{i-1}'' + 2y_i'') \\
+ 15q_{i+1}^2 (-2y_i'' + y_{i+1}'') - 30q_i^3 (3y_{i-1}' + 4y_i') &- 30q_{i+1}^3 (4y_i' + 3y_{i+1}') \\
+ 210q_i^4 (-y_{i-1} + y_i) + 210q_{i+1}^4 (-y_i + y_{i+1}), & \quad (3.105)
\end{aligned}$$

with corresponding formulas for both endpoints. As the coefficients on the right-hand side increase rapidly for decreasing interval sizes, some care must be taken in the application of this formula.

The accuracy of different splines has been tested by the author with an example, published by Schwarz [4]. The test function is

$$y(x) = 5 + 2.5(\cos 2\pi x/16 - \sin 4\pi x/16), \quad (3.106)$$

in the interval $1 \leq x \leq 17$, corresponding to one period. The sampling points and the results are shown in Fig. 3.8. The spline $s_5(x)$ was calculated by the combination of Eqs. (3.103) and (3.104) and lies between the exact curve and the cubic spline. The spline $\bar{s}_5(x)$ is obtained from Eq. (3.103) alone; the derivatives of first order were given exactly, whereas $s_7(x)$ was obtained from Eq. (3.105), all terms on the right-hand side being prescribed exactly. The functions $\bar{s}_5(x)$ and $\bar{s}_7(x)$ are not shown in the graph because they are too close to the exact curve. The worst error is assumed at $x = 7.2$ and has the following values:

$$s_3 : -0.693, \quad s_5 : -0.373, \quad \bar{s}_5 : -0.0437, \quad s_7 : -0.0028.$$

Obviously, the accuracy is improved by prescribing as many sampling derivatives as possible.

A field of application of Hermite splines is the description of boundary curves, as already introduced in Section 2.2.1; this is worked out in detail in Section 3.4. It will then be shown that the definition of derivatives up to

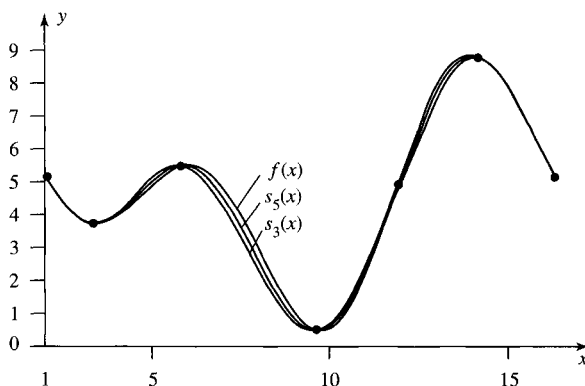


FIGURE 3.8 Test of periodic splines of third and fifth order, $s_3(x)$ and $s_5(x)$ for a superposition $f(x)$ of trigonometric functions. The sampling points are marked.

the second order is not too complicated. The solution of tridiagonal systems of equations is an important special case of the solution of general systems and is dealt with in Section 5.6. In the next section we shall consider a third class of interpolations, which is in some sense a combination of Lagrange interpolation and Hermite splines.

3.3 MODIFIED INTERPOLATION KERNELS

In many cases, for example, in the applications of the boundary-element method (see Chapter VI), the trial functions $T_i(x)$, defined in Section 3.2.1 are needed *directly*. This is the case if the sampling values y_i in Eq. (3.62) are initially unknown and are to be determined from the solution of a system of equations for them. Then, neither the recurrence scheme (3.70–72) nor the Hermite interpolation with splines can be used. The Lagrange functions (Eq. (3.65)) that are an explicit representation are unsuitable because of their strong oscillations. It is the goal in the present section to construct a new class of trial functions or interpolation kernels that are well behaved all over the x -scale.

Although the procedure outlined here also works for general sequences of x -intervals, it becomes reasonably simple only for *equidistant* intervals; we hence assume that

$$x_n = x_0 + nh, \quad y_n := y(x_n), \quad (n \text{ integer}). \quad (3.107)$$

The interval size h is positive, but the integer n may have either sign. Moreover, we postpone the problem of reaching a boundary to Section 3.3.3; it will turn out that this does not make severe difficulties.

The basic idea is quite simple: We determine the derivations of the function to be interpolated at the sampling points from the given sampling data in the neighborhood. This is done for every sampling point, and thereafter the Hermite interpolation can be used in each interval and has the required continuity properties by construction. The only novel aspect here is that, this will not be done *explicitly* but in implicit form, in which all intermediate steps are finally removed.

3.3.1 Basic Relations

We set out from the well-known relations

$$\begin{aligned} y(x_0 + a) - y(x_0 - a) &= \sum_{n=1}^{\infty} \frac{2}{(2n-1)!} y_0^{(2n-1)} a^{(2n-1)} \\ &= 2ay'_0 + a^3 y_0'''/3 + a^5 y_0^{(5)}/60 + \dots, \end{aligned} \quad (3.108)$$

$$y(x_0 + a) + y(x_0 - a) - 2y_0 = \sum_{n=1}^{\infty} \frac{2}{(2n)!} y_0^{(2n)} a^{2n} \\ = a^2 y_0'' + a^4 y_0^{(4)}/12 + a^6 y_0^{(6)}/360 + \dots \quad (3.109)$$

valid for any regular function $y(x)$. Finite difference expressions for y_0' and y_0'' can be obtained if we set in turn $a = h, 2h, 3h, \dots$, and eliminate as many derivatives of higher orders as possible by linear combinations of these formulas. Thus the truncation after the first term on the right-hand side results in Eq. (3.69) with a corresponding change in the notation. The next better approximation is found from

$$y_1 - y_{-1} = 2hy_0' + h^3 y_0'''/3 + h^5 y_0^{(5)}/60, \\ y_2 - y_{-2} = 4hy_0' + 8h^3 y_0'''/3 + 32h^5 y_0^{(5)}/60. \quad (3.110)$$

Elimination of y_0''' results in

$$y_0' = [8(y_1 - y_{-1}) - (y_2 - y_{-2})]/(12h) + h^4 y^{(5)}(\xi)/30, \quad (3.111)$$

in which the last term is the discretization error in the general form.

In an analogous way we find

$$y_0'' = [-30y_0 + 16(y_1 + y_{-1}) - (y_2 + y_{-2})]/(12h^2) + h^4 y^{(6)}(\xi)/90. \quad (3.112)$$

We are now in a position to introduce the modified interpolation kernel in a nontrivial example $M = 3$. It is favorable to introduce a normalized variable $\tau = (x - x_0)/h$ and to study the interpolation in the interval $0 \leq \tau \leq 1$, as other intervals produce only an index shift. The basic equation is now

$$y(x) = \sum_{n=1-M}^M y_n F_M(\tau - n). \quad (3.113)$$

Note that all interpolation kernels F_M have the *same form* and are only *shifted* in their argument. The interpolating condition, which requires that the interpolation formula collapses to the correct value at the sample points, implies that

$$F_M(k) = \delta_{0,k} \quad (k \text{ integer}). \quad (3.114)$$

Moreover, we impose the condition that $F_M(t)$ be differentiable twice continuously at the sampling points. Hence, if we set $x = x_0$ and differentiate

Eq. (3.113) twice, we find

$$y'_0 = \frac{1}{h} \sum_{n=1-M}^M y_n F'_M(-n), \quad (3.115a)$$

$$y''_0 = \frac{1}{h^2} \sum_{n=1-M}^M y_n F''_M(-n). \quad (3.115b)$$

Apart from the discretization errors, these must agree with Eqs. (3.111) and (3.112), respectively. The comparison of coefficients then result in

$$F'_3(0) = 0, \quad F'_3(\mp 1) = \pm 2/3, \quad F'_3(\mp 2) = \mp 1/12, \quad F'_3(\mp 3) = 0, \quad (3.116a)$$

$$F''_3(0) = -5/2, \quad F''_3(\mp 1) = 4/3, \quad F''_3(\mp 2) = -1/12, \quad F''_3(\mp 3) = 0. \quad (3.116b)$$

The confinement to the interval $0 \leq \tau \leq 1$ can now be abandoned, and we then have a function $F_M(t)$ (with $t = \tau - n$), piecewise defined by fifth-order Hermite polynomials in the entire interval $-3 \leq t \leq 3$.

Because all derivatives vanish at $t = \pm 3$, the basic requirements are still satisfied if we set $F_3(t) = 0$ for $|t| \geq 3$. Moreover, the data given in Eqs. (3.116a,b) show that $F_3(t)$ is a *symmetric* function. These rules can be shown to be general, hence:

$$F_M(t) = F_M(-t) = F_M(|t|), \quad (3.117)$$

$$F_M^{(k)}(\pm M) = 0 \quad \text{for } k = 0, 1, \dots, M-1, \quad (3.118)$$

$$F_M(t) \equiv 0 \quad \text{for } |t| \geq M. \quad (3.119)$$

Generally, the recipes for the construction of modified interpolation kernels of order M are as follows:

- (i) Consider $2M - 1$ points and write down the expressions for the central difference-approximation of the derivatives $y'_0, \dots, y_0^{(M-1)}$.
- (ii) Use the coefficients of these formulas with $h = 1$ as derivatives of the kernel function, as in Eqs. (3.116a,b), and consider them as sampling values.
- (iii) Perform the Hermite interpolation of order $2M - 1$ in each interval.

The results for $M \leq 4$ are shown in Fig. 3.9. The oscillations of the functions are strongly damped, so that they can really be used as *kernel* functions. In comparison to these, Fig. 3.10 shows a conventional Lagrange kernel with its very strong oscillations.

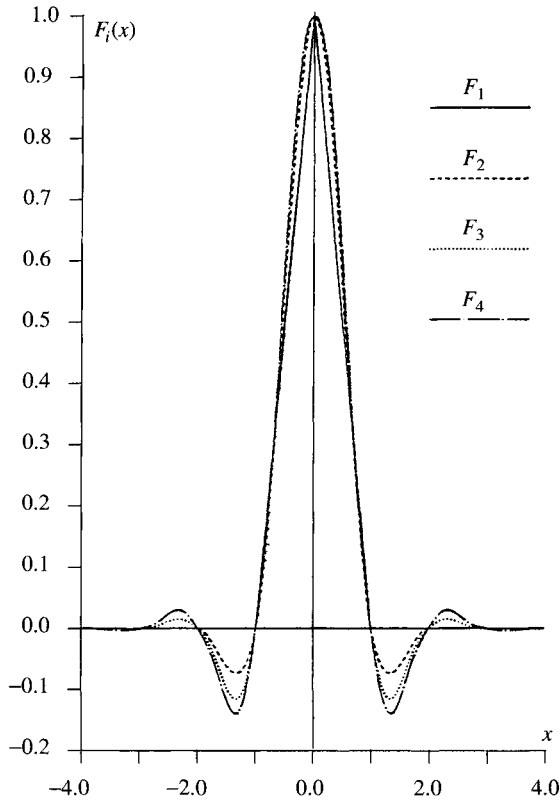


FIGURE 3.9 Modified interpolation kernels up to order $M = 4$ (after Kasper [5]).

3.3.2 The Recurrence Algorithm

The interpolation of a function in one of the unit intervals requires the calculation of all those shifted kernels $F_M(t - n)$ that do not vanish in this particular interval. It is advantageous to determine these simultaneously. The shift of the argument, as in Eq. (3.113), and the shift of the interval, as in the composition of Fig. 3.9, are equivalent operations. Owing to the symmetry in Eq. (3.117), it is sufficient to consider only positive values of t . The integer i satisfying $0 \leq i - 1 \leq t < i \leq M$ is then the number of the interval counted from the center.

The elementary way, whereby the boundary values of the derivatives are determined from central differences, after which the Hermite polynomials are calculated for each interval and the results are stored, is very tedious. The author [5, 6] has carried out this procedure explicitly for some low orders and was successful in finding an efficient recursive algorithm. This can be

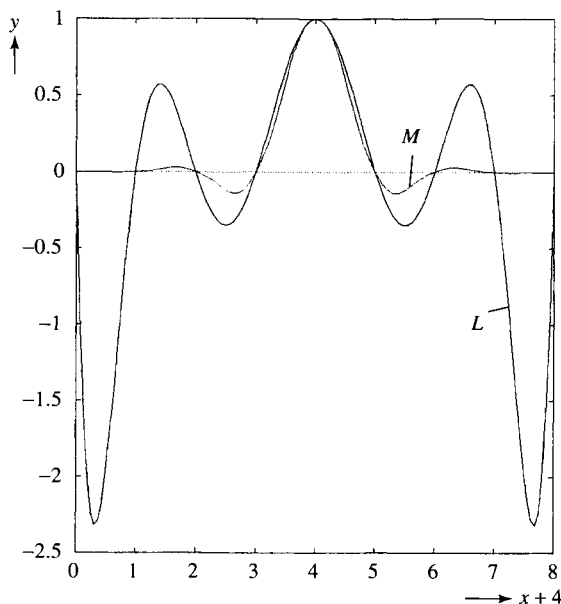


FIGURE 3.10 Comparison of a Lagrange polynomial of 8th order, (L), with a modified interpolation kernel F_4 , (M). At integer values of the abscissa x , both functions have the same sampling values. Note the very strong oscillations of the Lagrange polynomial.

checked by tedious explicit calculations or by precise numerical tests, but a general mathematical proof is not yet known.

Let $i - 1 \leq t \leq i$. Then we define new variables u and v by

$$u := 2(t - i) + 1, \quad -1 \leq u \leq 1, \quad (3.120a)$$

$$v := (1 - u^2)/8 = p_1(u)/4, \quad (3.120b)$$

which corresponds to Eqs. (3.73) and (3.79). We now define a set of functions $G_n(v)$ by

$$G_1 = 1/2 \quad \text{and} \quad G_{n+1}(v) = -G_n(v) \cdot [v + n(n-1)/2] \cdot [n(2n-1)]^{-1}, \\ (n \geq 1). \quad (3.121)$$

The beginning of this sequence is explicitly

$$G_1 = 1/2, \quad G_2 = -v/2, \quad G_3 = v(1+v)/12, \\ G_4 = -v(1+v)(3+v)/180, \quad G_5 = v(1+v)(3+v)(6+v)/5040. \quad (3.122)$$

These polynomials are well behaved, as they converge rapidly to zero for $n \rightarrow \infty$. Besides these we need a second sequence of functions $U_n(v)$, which can unfortunately not be given by such a simple law

$$\begin{aligned} U_1 &= 1/2, & U_2 &= -v/2, & U_3 &= v(1 + 5v)/12, \\ U_4 &= -v(3 + 16v + 89v^2)/180, \\ U_5 &= v(18 + 99v + 550v^2 + 3569v^3)/5040. \end{aligned} \quad (3.123)$$

The coefficients in these functions increase rapidly so that it is not favorable to go far beyond U_5 .

Now, if a rank $M \leq 5$ is given, the following algorithm can be executed:

$$\left. \begin{aligned} P_{kk} &= G_k, & Q_{kk} &= U_k & \text{for } k &= 1, 2, \dots, M, \\ P_{ik} &= G_k W_{ik} + P_{i,k-1}, \\ Q_{ik} &= U_k W_{ik} + P_{i,k-1}/(2i-1) \end{aligned} \right\} \begin{aligned} & \text{for } i = 1, \dots, k-1, \\ & k = 2, \dots, M. \end{aligned} \quad (3.124)$$

First a single loop is carried out for the diagonal elements, followed by a double loop, in which i is the label of the inner loop. The matrix W is given by

$$W = \begin{pmatrix} 1 & -1 & 2 & -5 & 14 \\ -3 & 1 & -3 & 9 & -28 \\ 10 & -5 & 1 & -5 & 20 \\ -35 & 21 & -7 & 1 & -7 \\ 126 & -84 & 36 & -9 & 1 \end{pmatrix}. \quad (3.125)$$

The subdiagonal part is used to calculate the matrix Q ; these elements can easily be identified as

$$W_{ki} = (-1)^{i+k} \binom{2k-1}{k-i} \quad (k \geq i). \quad (3.126a)$$

The coefficients W_{ik} in the upper diagonal part, needed to calculate the matrix P , are a little more complicated:

$$W_{ik} = (-1)^{i+k} \left[\binom{2k-3}{k-i} - \binom{2k-3}{k-i-2} \right] \quad (i \leq k, k \geq 2), \quad (3.126b)$$

where the second term must be dropped for $k < i + 2$. If this recurrence algorithm is completed, then the interpolation kernel $F_k(t)$ in the interval $i-1 \leq t \leq i$ is given by

$$F_k(t) = P_{ik} - u Q_{ik} \quad (3.127)$$

TABLE 3.1
TABLE OF ACCURACIES ($h = 0.1$, $M = 5$)

x	$\exp(x)$	$\cos(x)$	$\log(1 + x^2)$	$1/(1 + x^2)$
0.500	0	0	0	0
0.525	-5.0033e-13	1.4435e-13	6.2401e-09	-3.1136e-08
0.550	4.1762e-14	-2.0519e-14	4.2273e-10	-9.1870e-09
0.575	5.6080e-13	-1.7739e-13	-5.6368e-09	1.7992e-08
0.600	0	0	0	0

This means that all functions up to order $k = M$ are simultaneously given in their positive domain of definition and, because of Eq. (3.117), in the negative domain as well. (In the original paper, [5], some errors occurred and these are corrected here.) The recurrence algorithm has the advantage that its rank can easily be increased until some realistic criteria for accuracy are satisfied. If the final rank $k = M$ is prescribed in Eq. (3.127), then only the matrix elements Q_{iM} are needed. This implies that it is immediately sufficient to calculate only U_M and to set $k = M$ in the loop for the Q -coefficients; the algorithm then becomes very economic.

The figures in Table 3.1 are the interpolation errors in an interval between sampling points at $x = 0.5$ and 0.6 , respectively.

As can be seen by inspection of these numbers and, as can be proven generally, the interpolation error does not assume its absolute maximum at the interval midpoint but at some points in between. With kernels of rank M , any polynomial of degree $2M - 2$ (here 8) is interpolated exactly, and at the midpoint, any polynomial of degree $2M - 1$. This is the maximum possible with $2M$ sampling points.

The table also shows that functions that have a singularity in the complex plane, like $x = \pm i$ in the last two examples, are approximated far less accurately. This is a consequence of the fact that, in such cases, the derivatives of higher orders are not absolutely confined, and this is true for any kind of approximation that does not fit these singularities.

3.3.3 Extrapolation

Quite often the problem arises that a function $y(x)$ is defined only in an interval $a \leq x \leq b$ and that the interpolation is to be carried out near the bounds of this interval. We presume that the quotient $(b - a)/h$ has an integer value, which can always be achieved by appropriate choice of h . The technique outlined so far requires a minimal distance of $(M - 1)h$ from both endpoints, and hence the function $y(x)$ must be extrapolated in some way if x comes closer to one of the endpoints.

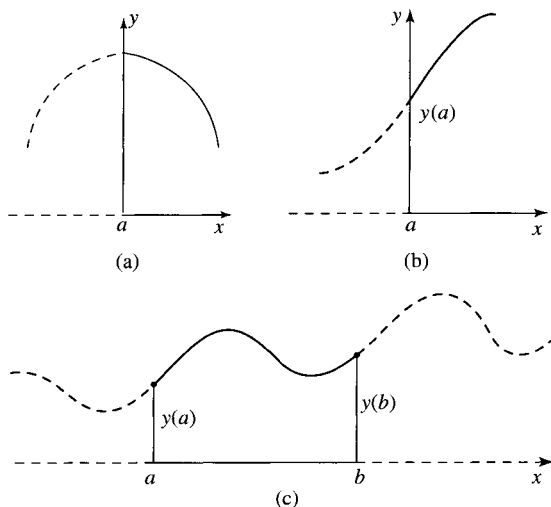


FIGURE 3.11 Elementary symmetry properties of functions: (a) positive mirror symmetry; (b) negative symmetry with shift $y(a)$; (c) periodicity with jump discontinuity $y(b) - y(a)$.

This task is quite easy if the function has specified symmetry properties — a case that was already considered in the theory of splines. Let s be the necessary extrapolation length; then we find the simple extrapolation rules (see Fig. 3.11)

(i) *Periodicity* (with possible superimposed linear function)

$$\begin{aligned} y(b+s) &= y(a+s) + y(b) - y(a), \\ y(a-s) &= y(b-s) - y(b) + y(a). \end{aligned} \quad (3.128)$$

(ii) *Positive symmetry*

$$y(a-s) = y(a+s) \quad \text{or} \quad y(b+s) = y(b-s). \quad (3.129)$$

(iii) *Negative symmetry* (with possible shift)

$$y(b+s) = 2y(b) - y(b-s), \quad (3.130a)$$

$$y(a-s) = 2y(a) - y(a+s). \quad (3.130b)$$

These rules can be used easily in context with Eq. (3.113) and are compatible with it if the sampling points are either the direct endpoints

$$x_n = a + nh \quad (x_0 = a, n \leq (b-a)/h) \quad (3.131)$$

or the midpoints

$$x_n = a + (n + 1/2)h, \quad (x_0 = a + h/2, n < (b - a)/h). \quad (3.132)$$

Sometimes none of the symmetry rules is valid. Then it is possible to try an extrapolation by a general functional law

$$y(x) = P_N(x)A(x) + B(x), \quad (3.133)$$

where $P_N(x)$ is a polynomial of degree N , and $A(x)$ and $B(x)$ are the given functions. The polynomial is generally not known in explicit form, but we can easily determine its sampling values

$$Y_n := P_N(x_n) = (y_n - B(x_n))/A(x_n) \quad (L - N \leq n \leq L) \quad (3.134)$$

and extrapolate then with these.

3.3.4 Nonequidistant Intervals

The basic requirement of Eq. (3.107) cannot always be satisfied or may not be adequate. If we do not impose it, the algorithm naturally becomes more complicated. The necessary modifications are outlined only very briefly.

A fairly simple generalization is a smooth transform of the x -scale of the form

$$x = \bar{x}(t), \quad y(x) = y(\bar{x}(t)) = \bar{y}(t). \quad (3.135)$$

Without loss of generality the parameter t can be chosen such that the intervals have unit length; hence Eq. (3.107) is now replaced by

$$x_n = \bar{x}(t_0 + n), \quad y_n = \bar{y}(t_0 + n), \quad (n \text{ integer}). \quad (3.136)$$

This modification has the advantage that the recurrence algorithm can again be executed up to high orders ($M = 5$), and this yields a good accuracy. The constant t_0 can be chosen freely; depending on the actual problem, $t_0 = 0$ or $t_0 = 1/2$ is advantageous. The only novel aspect is that here the value t has to be determined from the inverse function $t = \bar{t}(x)$ if x is prescribed.

There may be cases in which a function $\bar{x}(t)$ cannot be defined in a reasonable manner because the ratios of neighboring interval lengths differ too much from unity. Then, however, the algorithm becomes far more complicated. A reasonably simple formalism is possible only for *four-point* interpolation ($M = 2$). Without loss of generality we can always shift the labeling in such a way that the interpolation is to be carried out in the interval $x_2 \leq x \leq x_3$ by

use of the sampling data with labels $1 \leq n \leq 4$. It is favorable to define the interval lengths

$$h_n = x_{n+1} - x_n, \quad (n = 1, 2, 3). \quad (3.137)$$

Next, we apply the Eq. (3.67) to the points with labels 2 and 3:

$$y'_n = \frac{h_{n-1}^2(y_{n+1} - y'_n) + h_n^2(y_n - y_{n-1})}{h_{n-1}h_n(h_{n-1} + h_n)}, \quad (n = 2, 3). \quad (3.138)$$

The interpolation polynomial is now uniquely defined as the cubic Hermite polynomial formed with the boundary values y_2, y_3 and the boundary derivatives y'_2, y'_3 . By construction, it remains continuously differentiable on passing to the neighboring intervals. It can be easily cast in a more explicit form if we introduce a normalized variable

$$t := (x - x_2)/h_2, \quad 0 \leq t \leq 1, \quad (3.139)$$

and the cubic polynomials

$$p_1(t) = (1 + 2t)(1 - t)^2, \quad p_2(t) = t^2(3 - 2t), \quad (3.140)$$

$$q_1(t) = t(1 - t)^2, \quad q_2(t) = t^2(1 - t), \quad (3.141)$$

which satisfy normalized boundary conditions. We can now cast the interpolation in the convenient form

$$y(x) = \sum_{n=1}^4 y_n f_n(t), \quad (3.142)$$

with the kernels

$$f_1(t) = \frac{-h_2^2 q_1(t)}{h_1(h_1 + h_2)}, \quad f_4(t) = \frac{-h_2^2 q_2(t)}{h_3(h_3 + h_2)} \quad (3.143a,b)$$

for the outer points and

$$f_2(t) = p_1(t) + \left(\frac{h_2}{h_1} - 1\right) q_1(t) + \frac{h_3}{h_2 + h_3} q_2(t), \quad (3.143c)$$

$$f_3(t) = p_2(t) + \left(\frac{h_2}{h_3} - 1\right) q_2(t) + \frac{h_1}{h_1 + h_2} q_1(t) \quad (3.143d)$$

for the inner points. An example is shown in Fig. 3.12. This representation is particularly favorable if the kernel functions themselves are explicitly needed.

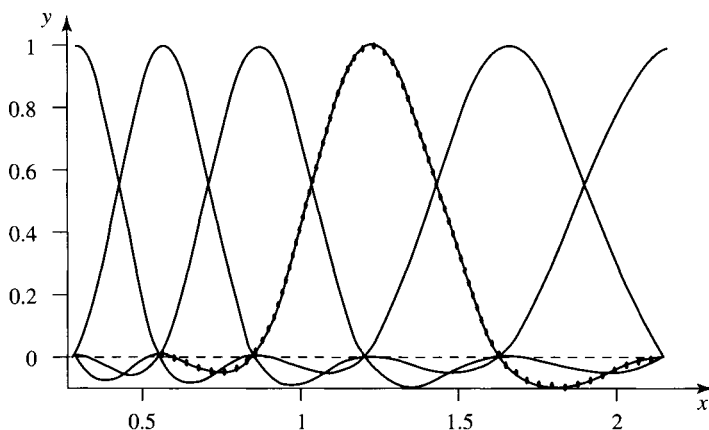


FIGURE 3.12 Modified interpolation kernels of order $M = 2$, with exponentially increasing intervals on the x -axis. One of these functions is marked for clarity.

Such a situation arises in calculations that imply the determination of an *unknown* function $y(x)$. A possible approach is then a discretization and the application of the kernel-approximation to each interval. The determination of a *continuous* function is then reduced to the solution of a system of *discrete* equations for the unknown sampling values y_n ($n = 1, \dots, N$). Examples of such problems appear in the boundary-element method (see Chapter VI).

There still remains the problem of carrying out the interpolation in the first or last interval where one neighboring interval is missing. This problem can be solved by appropriate extrapolation (see Section 3.3.3). Sometimes the boundary derivative y'_1 is given, or it can serve as an additional variable. Then we can again use y'_2 from Eq. (3.138) and determine the kernels of the Hermite polynomials with boundary values (y_1, y_2) and (y'_1, y'_2) . In the simplest case, we can eliminate y'_1 by means of the approximation $y'_1 = 2(y_2 - y_1)/h_1 - y'_2$. Similar considerations hold for the upper end at $x = x_N$.

3.4 MATHEMATICAL REPRESENTATION OF CURVES

In numerical procedures for charged-particle optics the mathematical representation of curves in space or in a plane is necessary in the following contexts:

- (i) as sections through electrodes or yoke surfaces (see Sections 2.2 and 3.1.5),
- (ii) as windings of current-conducting coils,
- (iii) as the curved optic axis of a deflected charged particle beam.

In all these cases, the most favorable representation depends on particular properties, which can be taken into account to improve the corresponding mathematical description. Here we shall try to find a fairly general method.

Apart from some very special cases, the cartesian representation of a curve by functions $x(z)$, $y(z)$ is not sufficient because these will not always be unique, and therefore we shall not consider this representation here. However, the *parametric* form $\mathbf{r}(p) = (x(p), y(p), z(p))$ is always feasible, where p is a suitable curve parameter. Here we shall consider only *smooth* curves; this does not imply any loss of generality, as curves with sharp corners can be represented as a *sequence* of smooth ones. Favorably, each corner is then treated as a *double point*—the first point then being the end of the preceding section and the second one the beginning of the next one. Furthermore, we shall assume that the basic information for a smooth curve is a sequence of *sampling points* with derivatives

$$\mathbf{r}_n^{(m)} := [d^m \mathbf{r}(p)/d p^m]_{p=p_n}, \quad (0 \leq m \leq M, 0 \leq n \leq N), \quad (3.144)$$

which can then be used for suitable interpolations. These can be carried out with trigonometric functions or with Hermite polynomials; the latter are more suitable, because with these it is easier to satisfy the necessary smoothness conditions at the sampling points.

3.4.1 Differential Geometrical Functions

We now assume that the vector function $\mathbf{r}(p)$ is differentiable at least thrice continuously ($M = 3$), as it is a spline. It is then possible to define all differential geometrical functions, which can be done in terms of the parameter p or the arc length s . For conciseness, we shall denote derivatives with respect to p by dots and those with respect to s by primes. With respect to the transformation between these, it is favorable to introduce a “velocity”

$$v(p) = |\dot{\mathbf{r}}(p)| \equiv (\dot{x}^2 + \dot{y}^2 + \dot{z}^2)^{1/2} > 0. \quad (3.145)$$

Then for any smooth function $f(p) \equiv \hat{f}(s)$, this transform takes the concise form

$$\hat{f}'(s) = \dot{f}(p)/v(p). \quad (3.146)$$

Using this and the corresponding transforms of higher orders, we obtain in turn the normalized curve tangent (in which the circumflex $\hat{}$ has been dropped):

$$\mathbf{t} = \mathbf{r}'(s) = \dot{\mathbf{r}}(p)/v(p) \quad (3.147)$$

and the *unsigned* curvature

$$c = |\mathbf{r}''(s)| = |\dot{\mathbf{r}} \times \ddot{\mathbf{r}}|/v^3. \quad (3.148)$$

These are the most important functions. For $c > 0$ we can introduce a principal normal $\mathbf{n}$, the binormal $\mathbf{b}$, and the torsion w by

$$\mathbf{b} = \dot{\mathbf{r}} \times \ddot{\mathbf{r}}/|\dot{\mathbf{r}} \times \ddot{\mathbf{r}}| \equiv v^3 \dot{\mathbf{r}} \times \ddot{\mathbf{r}}/c, \quad (3.149)$$

$$\mathbf{n} = \mathbf{r}''(s)/c = \mathbf{b} \times \mathbf{t}, \quad (3.150)$$

$$w = c^{-2} \det(\mathbf{r}', \mathbf{r}'', \mathbf{r}''') \equiv \det(\dot{\mathbf{r}}, \ddot{\mathbf{r}}, \dddot{\mathbf{r}})/|\dot{\mathbf{r}} \times \ddot{\mathbf{r}}|^2. \quad (3.151)$$

The vector-triplet $(\mathbf{t}, \mathbf{n}, \mathbf{b})$ is positively oriented and satisfies Frenet's equations

$$\mathbf{t}' = c\mathbf{n}, \quad \mathbf{b}' = -w\mathbf{n}, \quad (3.152a,b)$$

$$\mathbf{n}' = -c\mathbf{t} + w\mathbf{b}, \quad (3.152c)$$

which may be useful for Hermite interpolations. It is obvious that the choice of the arc length as the curve parameter brings some simplification.

These equations have the disadvantage that they do not hold at points of inflection ($c = 0$), and that they become useless if a bent curve is joined smoothly to a straight line. Fortunately this difficulty does not arise for *planar* curves that are two-dimensional ones. We now set $y \equiv 0$ and obtain agreement with the formulas in Section 2.2 if we identify x with r . The torsion w vanishes identically, and the curvature may now have a *sign*

$$\kappa = z'x'' - x'z'' \equiv (\dot{z}\ddot{x} - \dot{x}\ddot{z})/v^3 = \pm c. \quad (3.153)$$

The principal normal can always be defined as

$$\mathbf{m} = \mathbf{t} \times \mathbf{e}_y = (-z', 0, x'), \quad (3.154)$$

whereupon Fresnet's equations specialize to

$$\mathbf{t}' = -\kappa\mathbf{m}, \quad \mathbf{m}' = \kappa\mathbf{t}. \quad (3.155)$$

No difficulties or ambiguity arises if κ changes its sign.

3.4.2 Determination of Sampling Arrays

With respect to the development of a fairly simple and generally applicable code, it is desirable to present the information about the curve in the form

of an array, as given in Eq. (3.144). It is then straightforward to carry out Hermite interpolations of order M . The number N of necessary intervals and the order M depend on the particular requirements for accuracy; in any event, a reasonable compromise between these demands has to be found.

The situation is fairly simple if the curve in question, is the *axis* of a particle beam (case (iii)). This is usually obtained as the result of a numerical ray tracing program, which supplies the data as sampling points, the endpoints of integration steps. The parameter p is the time or some variable related to it. The velocity components are immediately given by the output and the acceleration is easily found from the Lorentz force, and hence we then have $M = 2$. This is, however, not sufficient for an accurate determination of aberrations. The accuracy can be improved by solution of Eq. (3.105) for the derivatives of third order (with appropriate adjustment of the notation). We then have $M = 3$, and even the derivatives of fourth order remain continuous. If the curved axis remains in the plane $y = 0$, it is easy to define an adapted curvilinear coordinate system (p, u, y) by a transform

$$\hat{z} = z(p) + u \dot{x}(p)/v(p), \quad \hat{x} = x(p) - u \dot{z}(p)/v(p), \quad (3.156)$$

which might be useful for tracing neighboring rays.

Another simple case is the treatment of curves that are given entirely in *analytical* form. Here we have two choices: either to maintain this form or to determine a suitable array (Section 3.4.1) from it and to store this. The latter alternative is preferred with respect to a unified program structure and also if a frequent evaluation of the analytical functions would be too time-consuming. The appropriate decision depends on the particular situation.

A very simple case is the presentation of *polygons* entirely consisting of *straight* lines and without the requirement to round off the corners. It is then not necessary to store these corners as double points, as each pair of successive points is the beginning and end of the straight line between them. This case arises in the treatment of simple technical structures and in unsmoothed line plots.

We now focus our attention on more sophisticated cases. These arise if only the points $\mathbf{r}_0, \dots, \mathbf{r}_N$ themselves are given, which will be sampling points of a smooth curve, and a suitable parameter p is *not* known initially. For reasons of uniqueness it is then favorable to choose the arc length s as parameter, but this too must be approximated. In the lowest approximation, to start with, we can choose the sequence of chord lengths

$$\mathbf{d}_i := \mathbf{r}_i - \mathbf{r}_{i-1} \quad (i = 1, \dots, N), \quad (3.157a)$$

$$d_i := |\mathbf{d}_i|; \quad q_i = d_i^{-1} \quad (i = 1, \dots, N), \quad (3.157b)$$

and then

$$p_0 = 0, \quad p_i = p_{i-1} + d_i \quad (i = 1, \dots, N). \quad (3.158)$$

The subsequent procedure depends on the requirements for accuracy.

Lowest Approximation

If only a smooth line plot is required but not the accurate arc length on the curve, it is sufficient to determine the *cubic* Hermite spline for each cartesian coordinate as a function of p . However, it is better to use each tangent vector as a boundary condition if it is known exactly.

Improved Splines

If all normalized tangent vectors are also given as

$$\mathbf{r}'_i = \mathbf{r}'_i = \mathbf{t}_i, \quad |\mathbf{t}_i| = 1 \quad (i = 0, \dots, N), \quad (3.159)$$

then the *quintic* spline Eq. (3.103) can be applied. It is, however, necessary to improve the calculation of the arc length

$$s_0 = 0, \quad s_i = s_{i-1} + \int_{p_{i-1}}^{p_i} |\dot{\mathbf{r}}(p)| dp \quad (i = 1, \dots, N). \quad (3.160)$$

When these integrations are complete, we have to replace p_i by s_i , ($i = 0, \dots, N$). Thereafter the whole procedure has to be iterated until it has converged sufficiently. The accuracy finally achieved depends on the angle between subsequent tangents. It is usually sufficient to ensure that the arcs are not bent too strongly, that is, $|\mathbf{t}_{i-1} \times \mathbf{t}_i| < 0.2$.

If the arcs between adjacent points are too large, they must be subdivided by spline interpolation. In any case, the normalization (Eq. (3.159)) must be satisfied. If tangents are not known, they can be found from the circle by passing through three adjacent points ($i-1, i, i+1$). The corresponding formula is given by

$$\mathbf{t}_i = \frac{d_{i+1}^2 \mathbf{d}_i + d_i^2 \mathbf{d}_{i+1}}{d_{i+1} d_i |\mathbf{d}_i + \mathbf{d}_{i+1}|}, \quad |\mathbf{t}_i| = 1. \quad (3.161)$$

If the vicinity of the point $\mathbf{r}_i$ is to be approximated by the osculating circle for technical reasons, we can easily determine the curvature c_i and binormal $\mathbf{b}_i$ by

$$c_i \mathbf{b}_i = \frac{2\mathbf{d}_i \times \mathbf{d}_{i+1}}{d_i d_{i+1} |\mathbf{d}_i + \mathbf{d}_{i+1}|}. \quad (3.162)$$

The unsigned curvature is the norm of this vector. The vectorial derivatives of second and third order are now simply given by

$$\mathbf{r}_i'' = c_i \mathbf{b}_i \times \mathbf{t}_i, \quad (3.163)$$

$$\mathbf{r}_i''' = -c_i^2 \mathbf{t}_i. \quad (3.164)$$

It must be emphasized that Eq. (3.164) implies vanishing torsion so that a solenoid cannot be approximated in this way. We then have to omit this formula and determine the third-order derivatives from the spline Eqs. (3.103) or (3.105).

For curves in a plane—here the (z, x) plane—the vector products are unnecessary and the curvature can have a sign. Equation (3.162) is then replaced by

$$x_i = \frac{2(\Delta z_i \Delta x_{i+1} - \Delta x_i \Delta z_{i+1})}{d_i d_{i+1} |\mathbf{d}_i + \mathbf{d}_{i+1}|}, \quad (3.165)$$

where Δz , Δx are the cartesian components of the corresponding vector $\mathbf{d}$. The derivatives of second order are now simply given by

$$z_i'' = -\kappa_i x_i', \quad x_i'' = \kappa_i z_i'. \quad (3.166)$$

3.4.3 Rounding-off Corners

Sharp edges and corners are unphysical, as the electromagnetic fields diverge at them. Also, a current-conducting wire cannot be bent in an infinitely sharp manner. They are hence always a simplifying idealization, which is sometimes allowed and sometimes not. The technique of Hermite splines provides a simple method of smoothing that is shown in Fig. 3.13.

At first the original corner $\mathbf{r}_E$ is introduced as a *double* point $\mathbf{r}_A = \mathbf{r}_B$, but with different derivatives, for instance $\mathbf{r}'_A \neq \mathbf{r}'_B$. Then for both smooth parts of

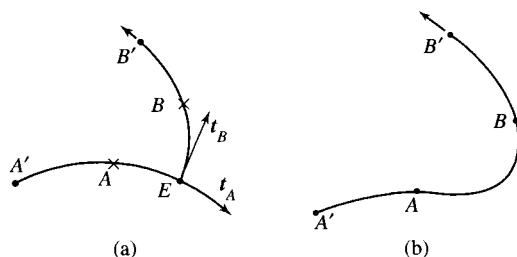


FIGURE 3.13 Rounding-off procedure: (a) form of the curve with sampling points A' , E , and B' , E being a corner (double point) with tangent vectors $\mathbf{t}_A = \mathbf{r}'_A$ and $\mathbf{t}_B = \mathbf{r}'_B$. The positions A and B (marked by cross) are determined by Hermite interpolation. (b) Hermite interpolation between points A and B , after elimination of the corner E .

the curve the Hermite splines described earlier is determined. After that, some reasonable spacing h from the edge is chosen and the points and derivatives

$$\mathbf{r}_A^{(m)} = \mathbf{r}^{(m)}(s_E - h), \quad \mathbf{r}_B^{(m)} = \mathbf{r}^{(m)}(s_E + h) \quad (m = 0, \dots, M) \quad (3.167)$$

are determined by Hermite interpolation.

Now the original part between the points A and B is abandoned and is replaced by the smooth arc obtained by Hermite interpolation using the boundary conditions (Eq. (3.167)). If it is essential to have the arc length as curve parameter, some points must be inserted in the new arc, and the integrations (Eq. (3.160)) are then repeated. In this way it is possible to construct very smooth technical curves consistently.

3.5 MATHEMATICAL REPRESENTATION OF SURFACES

The mathematical representation of surfaces and of area elements is necessary in various fields of application, for example,

- (i) as the necessary dissection of domains for purposes of interpolations and integrations in the methods of finite differences and finite elements, (see Chapters IV and V, respectively);
- (ii) as the representation of curved surfaces in R_3 applications of the boundary-element method (see Chapter VI).

Sometimes a representation in rectilinear coordinates (x, y) or (z, r) might be adequate, but in general this is too restrictive; we therefore choose the parametric form (Fig. 3.14)

$$\mathbf{r}(u, v) = (x(u, v), y(u, v), z(u, v)). \quad (3.168)$$

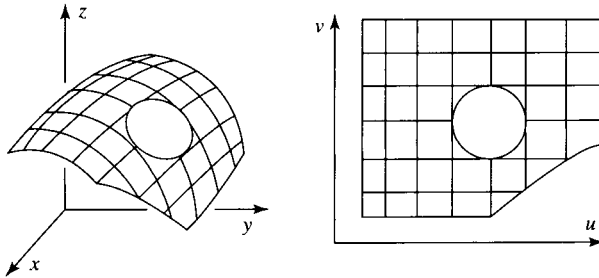


FIGURE 3.14 Example of the parametric representation of a curved surface.

Its evaluation is often a repetition of the same kind of algorithm for the different coordinates; hence for conciseness we shall consider only scalar functions, in which the vector character of the surface function is not essential.

In every case, the surface is mapped on some domain D in the plane of the parameters (u, v) . This domain will generally have curved boundaries (see Fig. 3.14) and might be multiply connected. The general case can become extremely complicated. Hence for reasons of space we can deal here only with fairly simple situations.

3.5.1 Rectangular Meshes

We now consider the fairly simple case shown in Fig. 3.15: The whole domain or at least a major part of it is dissected into rectangles by coordinate lines $u_i = \text{const.}$ and $v_k = \text{const.}$ Any function $f(u, v)$ can now be discretized as a two-dimensional array

$$f_{ik} := f(u_i, v_k) \quad (i = 0, \dots, J, k = 0, \dots, L) \quad (3.169)$$

of the function values at the nodes. This situation arises mainly if the function is initially not known, and the sampling array (Eq. (3.169)) will then be the result of a major numerical calculation.

The bivariate interpolation can be written down in very concise form if *kernel functions* for the individual coordinates, $F_i(u)$ and $G_k(v)$, respectively, are explicitly known and satisfy

$$F_i(u_m) = \delta_{i,m} \quad G_k(v_n) = \delta_{k,n} \quad (3.170)$$

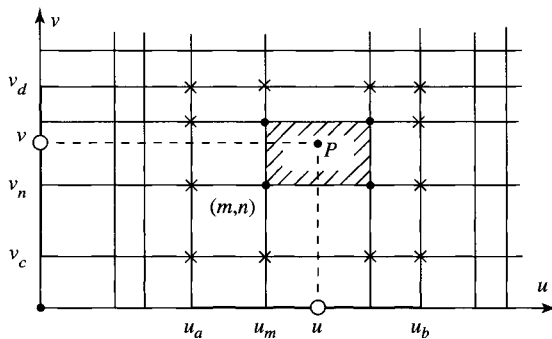


FIGURE 3.15 Bivariate modified Lagrange interpolation at a position P satisfying $u_m \leq u \leq u_{m+1}$ and $v_n \leq v \leq v_{n+1}$. The case $M = N = 1$ leads to a 16-points interpolation.

for all combinations of labels. A straightforward calculation shows that with appropriate labels and choice of M , N , m , and n , $\mu = 2M + 1$, $\nu = 2N + 1$,

$$f(u, v) = \sum_{i=a}^b \sum_{k=c}^d F_i(u) G_k(v) f_{ik} \quad (3.171)$$

is then a correct interpolation formula; the limits $a = m - M$, $b = m + M + 1$, $c = n - N$, and $d = n + N + 1$ are then to be used. It is often sufficient to choose equal ranks $M = N$, but this is not absolutely necessary. The highest power in the interpolated functions appears in a mixed term $u^\mu v^\nu$. This might be unnecessary, but attempts to avoid this effect generates formulas of more complicated structure. The main advantage of Eq. (3.171) is the fact that its programming can be made very efficient and its generalization to three dimensions is straightforward. The most time-consuming part—the calculation of the kernels—increases only *linearly* with the dimension.

The main drawback of Eq. (3.171) is the difficulty of applying it near a boundary of the domain D . The solution of this problem is fairly easy if one of the extrapolation formulas of Section 3.3.3 can be applied, either by explicit continuation of the sampling array (Eq. (3.169)) beyond its given bounds or by incorporation of the extrapolation rules into the kernel functions.

The interpolation of a vector function like that of Eq. (3.168) is now simply a threefold evaluation of the same algorithm for the cartesian components and can be organized in an efficient form, because the functions $F_i(u)$ and $G_k(v)$ need to be determined only once per surface point.

3.5.2 Bivariate Hermite Interpolation

The concept of Hermite interpolation can be generalized for two and more dimensions. It has the advantage that only the function values and derivatives at the corners of the corresponding element are used and no array data outside it are used. On the other hand, it is now necessary to store sampling values of partial derivatives as well, unless the latter can be calculated easily. Examples of such a favorable exceptional situation are given later in Section 5.5.3.

Here we shall confine our considerations to the most important general case: the bicubic interpolation that uses cubic polynomials in both coordinate directions. To construct these, we must know the function values, the partial derivatives of first order, and the *mixed* derivative of second order at the nodes.

For conciseness, we now define formal vectors

$$\mathbf{V}(u, v) = [f(u, v), \partial f / \partial u, \partial f / \partial v, \partial^2 f / \partial u \partial v] \quad (3.172)$$

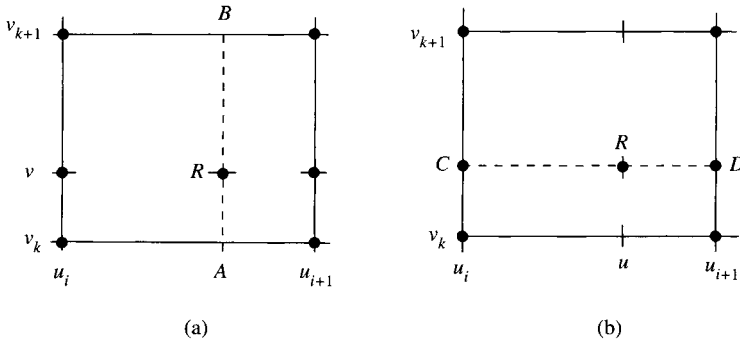


FIGURE 3.16 Bivariate Hermite interpolation: (a) interpolations in the u -direction at points A and B and subsequently in the v -direction between A and B at the reference point R ; (b) reversed sequence of directions involving points C , D , and R .

and use the abbreviations $V_{ik} = V(u_i, v_k)$ for their values at the nodes. The task is then to determine $V(u, v)$ from the four vectors V_{ik} , $V_{i+1,k}$, $V_{i,k+1}$, and $V_{i+1,k+1}$ in a consistent manner. Since these are altogether 16 values, they suffice to determine uniquely the 16 terms of the bicubic polynomial.

A possible way of calculating this polynomial and its derivatives is shown in Fig. 3.16; we first carry out two Hermite interpolations in the u -direction to find the vectors $V(u, v_k)$ and $V(u, v_{k+1})$ at points A and B , respectively. One Hermite interpolation in v -direction is then carried out to determine the vector $V(u, v)$ at the reference point R from these vectors. This procedure can be expressed as

$$\left. \begin{aligned} V(u, v_k) &= H_u(u, V_{ik}, V_{i+1,k}) \\ V(u, v_{k+1}) &= H_u(u, V_{i,k+1}, V_{i+1,k+1}) \\ V(u, v) &= H_v(v, V(u, v_k), V(u, v_{k+1})) \end{aligned} \right\} \quad (3.173)$$

Here H_u and H_v are vector functions, which each interpolate two components as cubic Hermite polynomials and the other ones as their derivatives. For example, H_u interpolates f and $\partial f / \partial v$ as cubic functions of u , while $\partial f / \partial u$ and $\partial^2 f / \partial u \partial v$ are then their derivatives, as they must be. Likewise, in H_v the components f and $\partial f / \partial u$ are now functions of v and $\partial f / \partial v$, $\partial^2 f / \partial u \partial v$ are their derivatives.

We could also choose the other alternative, interpolating first in the v -direction at points C and D , and subsequently in the u -direction between these. The result is the same, as the polynomial is uniquely determined by the 16 edge values and these are assumed continuous in both cases.

The fact that $\partial^2 f / \partial u \partial v$ remains continuous along the mesh lines implies that all *normal* derivatives remain continuous, an advantage that can hardly

be achieved by two-dimensional Lagrange interpolation. The algorithm is very flexible as *equidistant* meshes are not assumed.

3.5.3 Bicubic Splines

The bicubic Hermite interpolation requires only a unique definition for the derivatives at the nodes; the interpolated function $f(u, v)$ will then always have the stated smoothness properties. The introduction of spline functions results from the further-going demand that the pure derivatives of second order $\partial^2 f / \partial u^2$ and $\partial^2 f / \partial v^2$ should also remain continuous. Such a demand arises in a natural way if, for example, $f(u, v)$ represents function values of a potential that has continuous sources.

The construction of bicubic splines is straightforward: we first evaluate the one-dimensional Eq. (3.99) together with special conditions for the boundary derivatives along the rows and columns of the grid, thus obtaining consistent node-values of the first-order derivatives. Thereafter, one sequence of spline formulas is applied to their derivatives in order to determine the node values of the mixed second derivative.

In order to cast this procedure in a concise form, we introduce the abbreviations

$$\left. \begin{aligned} p_i &:= (u_i - u_{i-1})^{-1} \quad (1 \leq i \leq J) \\ q_k &:= (v_k - v_{k-1})^{-1} \quad (1 \leq k \leq L) \end{aligned} \right\} \quad (3.174)$$

and attach a third label n ($0 \leq n \leq 3$) to the sampling vectors V_{ik} for the numbering of their components. Thus $V_{ik0} = f_{ik}$, $V_{ik1} = (\partial f / \partial u)_{u_i, v_k}$, and so on.

Two different kinds of systems of tridiagonal equations can now be obtained from Eq. (3.99) and the additional conditions by appropriate adjustment of the notation. In particular, we have

$$\begin{aligned} p_i V_{i-1, k, m} + 2(p_i + p_{i+1}) V_{i, k, m} + p_{i+1} V_{i+1, k, m} &= 3p_i^2 (V_{i, k, \mu} - V_{i-1, k, \mu}) \\ &+ 3p_{i+1}^2 (V_{i+1, k, \mu} - V_{i, k, \mu}), \quad (1 \leq i \leq J-1, 0 \leq k \leq L, \mu = m-1). \end{aligned} \quad (3.175)$$

For differentiations in the u -direction, the *normal* derivatives $V_{0, k, m}$ and $V_{J, k, m}$ are to be specified additionally. In an analogous way, we find for the v -direction

$$\begin{aligned} q_k V_{i, k-1, n} + 2(q_k + q_{k+1}) V_{i, k, n} + q_{k+1} V_{i, k+1, n} &= 3q_k^2 (V_{i, k, v} - V_{i, k-1, v}) \\ &+ 3q_{k+1}^2 (V_{i, k+1, v} - V_{i, k, v}) \quad (1 \leq k \leq L-1, 0 \leq i \leq J, v = n-2) \end{aligned} \quad (3.176)$$

again with specifications of the normal derivatives $V_{i,0,n}$ and $V_{i,L,n}$. All these normal derivatives may be predefined, if known, or may be incorporated into the corresponding system of tridiagonal equations by one of the adjusted conditions Eq. (3.100) or (3.102).

The actual calculations now proceed in the following way:

- (1) Set all known values $V_{ik0} = f_{ik}$.
- (2) Solve Eq. (3.175) with $m = 1$ and $\mu = 0$ to find the node-values of $\partial f / \partial u$.
- (3) Solve Eq. (3.176) with $n = 2$ and $\nu = 0$ to find the node-values of $\partial f / \partial v$.
- (4) Determine the values of $\partial^2 f / \partial u \partial v$ at the four corners $i = 0, J$ and $k = 0, L$ in some reasonable manner.
- (5) Calculate the boundary values of $\partial^2 f / \partial u \partial v$ on two parallel sides of the rectangle; this means solving Eq. (3.175) with $m = 3$ and $\mu = 2$ for $k = 0$ and $k = L$, respectively.
- (6) Calculate the remaining values of $\partial^2 f / \partial u \partial v$ by solving Eq. (3.176) for $n = 3$ and $\nu = 1$, which means differentiation of $\partial f / \partial u$ with respect to v .

The calculation of array data is now complete, and it is reasonable to store all these data. In the steps (4) to (6) we could exchange the sequence of directions; this would give the same final results, as can be shown from the uniqueness of the spline function. We are now in a position to carry out the bicubic Hermite interpolation in any rectangular domain, and this gives very smooth results. This is of some importance for programs using quadrilateral finite elements. Similarly, surface charge distributions on curved surfaces can be interpolated very favorably in this way. Details will follow in the corresponding chapters.

3.5.4 Some Remarks

The spline techniques described earlier are the most important tools of fairly general applicability. The equally important techniques for interpolation in triangular nets are deferred to Chapter V on the finite-element method, where they are indispensable.

3.6 NUMERICAL INTEGRATION

It is a common situation that the numerical solution of physical problems involves the evaluation of integrals for which an analytical expression is not known, and hence a suitable numerical technique must be employed. There is a wide variety of integration formulas that can be found in practically every comprehensive textbook on numerical methods, and we shall therefore not derive them but just state some of them without proof and concentrate more on their range of applicability. The basic forms are integrals in one dimension.

In practically all cases, the configuration of sampling points and integration weights is symmetric, as this cancels out integration errors. We therefore shift the origin of the x scale to the center of the integration interval; the basic form then becomes

$$J := \int_{-h}^h f(x) dx = 2h \sum_{n=1}^N W_n f_n + \sum_{m=1}^M ' c_m h^{m+1} (f_{(-h)}^{(m)} + (-1)^m f_{(h)}^{(m)}) + R_L, \quad (3.177a)$$

$$R_L = h^{L+1} f^{(L)}(\xi)/D_L, \quad -h \leq \xi \leq h. \quad (3.177b)$$

(The prime means that this term is not included if $M = 0$).

The integration weights W_n are dimensionless numbers, should preferably be positive, and must satisfy the normalization

$$\sum_{n=1}^N W_n = 1. \quad (3.178)$$

The integrand $f(x)$ is to be evaluated at positions $x_n = p_n h$ with $|p_n| \leq 1$; hence,

$$f_n := f(x_n) \equiv f(p_n h) \quad n = 1, \dots, N. \quad (3.179)$$

The symmetry of the formula means that

$$W_n = W_{N+1-n}, \quad p_n = -p_{N+1-n} \quad (n \leq (N+1)/2). \quad (3.180)$$

Sometimes it is more favorable to represent the formula in terms of the full interval size $H = 2h$. We now examine the most important classes of integration formulas.

3.6.1 Gauss-Legendre Quadrature

These formulas use the zeros of Legendre polynomials $P_n(x)$ and do not use boundary derivatives; hence $M = 0$. The sets of zeros p_n and weights W_n are given in comprehensive tables and are therefore not reproduced here in general; the first few data sets are

$$\left. \begin{aligned} N = 1 : & \quad p_1 = 0, \quad W_1 = 1 \\ N = 2 : & \quad p_2 = -p_1 = 1/\sqrt{3}, \quad W_1 = W_2 = 1/2 \\ N = 3 : & \quad p_2 = 0, \quad p_3 = -p_1 = \sqrt{0.6} \\ & \quad W_2 = 4/9, \quad W_3 = W_1 = 5/18 \end{aligned} \right\} \quad (3.181)$$

The Gauss quadrature is favored because in practice it minimizes the integration error for a fixed number N of evaluations. In fact, the order $L = 2N$ is the largest possible, and a reliable estimate for the denominator D_L is

$$D_L \geq 2^{2N+1}(2N+1)/(16N). \quad (3.182)$$

This does *not* mean that Gauss quadratures are always the best choice, because that expectation is based on the tacitly made assumption that the derivatives $f^{(L)}(\xi)$ remain *confined* for increasing order L . An example demonstrating the opposite is the integrand $f(x) = \ln(x+d+h)$, to be integrated between $-h$ and h . Then $d > 0$ is the closest distance from the singularity and the derivatives increase as $|f^{(L)}| = (L-1)!/d^L$.

The remainder is

$$|R| \approx \frac{4h}{2N+1} \left(\frac{h}{2d} \right)^{2N}, \quad (3.183)$$

that does *not* vanish for $N \rightarrow \infty$ if $h > 2d$. In this case, it is essential to subdivide the integration interval, but Gauss quadratures are not very favorable for adaptive procedures. They are preferred only when the appropriate subdivision is known in advance.

3.6.2 Bessel-Hermite Quadratures

These formulas become favorable if the integrand is obtained by *Hermite interpolation*, because only the sampling data already known are then used.

In this set, the fixed choice $N = 2$, $W_1 = W_2 = 1/2$, $p_2 = -p_1 = 1$ is made; hence, the term without derivatives is always

$$J_T = h(f(-h) + f(h)) - \frac{2}{3}h^3 f''(\xi), \quad (3.184)$$

which is the well-known trapezoidal rule. This is now improved by terms involving boundary derivatives in Eq. (3.177):

$$M = 1: \quad c_1 = 1/3, \quad L = 5, \quad D_5 = 22.5, \quad (3.185)$$

$$M = 2: \quad c_1 = 0.4, \quad c_2 = 1/15, \quad L = 7, \quad D_7 = -787.5, \quad (3.186)$$

$$M = 3: \quad c_1 = 3/7, \quad c_2 = 2/21, \quad c_3 = 1/105, \quad L = 9. \quad (3.187)$$

This sequence can be continued, but with increasing order it becomes more and more difficult to calculate the boundary derivatives with the necessary accuracy. It is then better to use hybrid formulas, for example $N = 3$, $M = 1$,

$p_1 = -1, p_2 = 0, p_3 = 1$:

$$J = \frac{h}{15}(7f_1 + 16f_2 + 7f_3) - \frac{h^2}{15}(f'_1 - f'_3) + h^7 f^{(6)}(\xi)/4725. \quad (3.188)$$

The Hermite interpolation also provides an elegant tool to calculate integral functions

$$I(x) = \int_{-h}^x f(x') dx' \quad (x \leq h) \quad (3.189)$$

by evaluating first the complete integral J and using it then as a boundary value together with those of the function and its derivatives.

3.6.3 Newton-Cotes Formulas and Adaptive Procedures

Apart from the trapezoidal rule, the formulas with even numbers of sub-intervals (N odd) are of special importance, because formulas enable adaptive algorithms to be developed. The most important cases are the familiar Simpson rule

$$J_S = h[f(-h) + 4f(0) + f(h)]/3 - h^5 f^{(4)}(\xi)/90, \quad (3.190)$$

and Bode's rule

$$J_B = h[7f(-h) + 32f(-h/2) + 12f(0) + 32f(h/2) + 7f(h)]/45 + 8h^7 f^{(6)}(\xi)/945. \quad (3.191)$$

For a fixed number N of function values to be determined, these formulas are certainly less favorable than the corresponding Gauss quadratures. They have, however, the great advantage that they allow the interval to be halved without recalculating all the values and are hence suitable for adaptive integrations.

The basic idea is as follows. First, a reasonable error limit ε is specified. Then the integration is carried out over the full interval and compared with the sum of the integrals over the subdivided intervals. If the absolute difference between these results is less than ε , the final integral can be improved by extrapolation and the process is finished; otherwise, the interval is repeated halved until convergence is attained or the sequence is halted. As the interval bounds are not fixed here, we give them explicitly and denote the formulas used by the labels T , S , B according to Eqs. (3.184), (3.190), and (3.191), respectively. The beginning of the algorithm could then be expressed as

follows:

$$\left. \begin{aligned} J_1 &= J_T(-h, h), \quad J_2 = J_T(-h, 0) + J_T(0, h) \\ J &= J_S(-h, h) \equiv (4J_2 - J_1)/3 \\ &\quad \text{if } |J_1 - J_2| < \varepsilon \text{ stop} \\ J_1 &= J_S(-h, h), \quad J_2 = J_S(-h, 0) + J_S(0, h) \\ J &= J_B(-h, h) \equiv (16J_2 - J_1)/15 \\ &\quad \text{if } |J_1 - J_2| < \varepsilon. \text{ stop} \end{aligned} \right\} \quad (3.192)$$

By adopting suitable storage techniques, the recalculation of values that are already known can be avoided. The interval is sufficiently short if one of the two stop conditions is satisfied. If this is *not* the case, the process of interval halving and comparing the results is to be continued. However, the extrapolation of Bode's rule to the nine-point formula already generates some negative weights; hence, some care must be taken.

Certainly, the classical nine-point formula of Stokes' closed type [3] gives the best result if the integrand can be calculated with high precision and the interval size is fairly large. A slightly less-accurate but more stable formula uses the unnormalized weights

$$\begin{aligned} w'_1 = w'_9 &= 351, & w'_2 = w'_8 &= 1792, & w'_3 = w'_7 &= 288, \\ w'_4 = w'_6 &= 2304, & w'_5 &= -20, & S &= 9450, \end{aligned} \quad (3.193)$$

and the normalized weights $W_k = w'_k/S$. This formula has an error term of the order h^9 .

If the accuracy limit requires further halving of the interval, it is better to use the results for the halved intervals with the nine-point formula and *not* extrapolate further, because the latter procedure becomes sensitive to rounding errors.

3.6.4 Euler Maclaurin Formulas

These formulas can be used if the integrand is given at equidistant sampling points. Since this is a common case, the Euler Maclaurin formulas are a tool that are worth considering. They exist in two versions: for function values given at the interval endpoints and at their midpoints. Both formulas are special cases of Eq. (3.177), but it is better to write them down explicitly. Let H be the full spacing between adjacent abscissae and $0 \leq x \leq NH$ the total integration interval. For conciseness, we introduce the abbreviation

$$q := H^2/24. \quad (3.194)$$

Then the two versions are

$$J_A := \int_0^{NH} f(x) dx = \frac{H}{2}[f(0) + f(NH)] + H \sum_{n=1}^{N-1} f(nH) \\ + \left[-2qf' + \frac{4}{5}q^2 f''' - \frac{16}{35}q^3 f^{(5)} \pm \dots \right]_0^{NH} \quad (3.195)$$

and

$$J_B := \int_0^{NH} f(x) dx = H \sum_{n=1}^N f((n-1/2)H) \\ + \left[qf' - \frac{7}{10}q^2 f''' + \frac{31}{70}q^3 f^{(5)} \pm \dots \right]_0^{NH}, \quad (3.196)$$

respectively. The coefficients are related to the Bernoulli numbers.

In general, these formulas are not very favorable, because the boundary values of the derivatives must be somehow approximated, and even if they are all known exactly, the series expansion is semiconvergent. This means that it must be truncated before the correction terms pass through their absolute minimum. There is, however, one important exception, and this is the integration over a *periodic* function: if the interval of length NH is a full period of the function, then all terms with derivatives cancel out *exactly*, and the integration is reduced to the simplest form of summation. This is also a justification for the methods employed in numerical Fourier analysis.

If the function is *symmetric* with respect to the position $x = 0$, which implies that $f(-x) = f(x)$, then the derivatives of odd orders vanish at this point. If the function is periodic and NH is a *half* period, then the same holds also for the upper end, and again the integration reduces to a simple summation. In the case of smooth analytic functions, it is also very accurate. For instance, all functions $\cos^{2m}(x)$ or $\sin^{2m}(x)$ can be integrated exactly over the interval $[0, \pi/2]$ with only $N = m$ subintervals; the double number is needed for the full period π , so that it is worth making use of the symmetry properties.

In general, the error of the Maclaurin formulas is difficult to assess. If the integrand is not periodic, so that some boundary terms do not vanish, the error is determined by these. In the simplest case, we can eliminate the contributions from qf' by a linear combination $J = (2J_B + J_A)/3$ that gives the so-called extended Simpson rule. In the case of periodic functions that have a singularity somewhere in the complex plane, it can be shown that the error is asymptotically proportional to $\exp(-2\pi d/H)$, d being the (absolute) distance from the singularity.

A practical accuracy control is then given by the fact that the errors of J_A and J_B are nearly of equal value and of opposite signs. Hence the checks

$$|J_A - J_B| < \varepsilon, \quad J = (J_A + J_B)/2 \quad (3.197)$$

are certainly reliable. The averaging corresponds to the evaluation of J_A with number $2N$, which implies a fast convergence.

An interesting task is the evaluation of integrals of the form

$$J_M = \int_{-MH}^{MH} F_M\left(\frac{x}{H}\right) P(x) dx = H \int_{-M}^M F_M(t) P(Ht) dt, \quad (3.198)$$

in which $P(x)$ is any nonsingular analytic function and $F_M(t)$ a modified interpolation kernel, as was defined in Section 3.3. We refer back to Eqs. (3.117–3.119) and recall that F_M is $(M-1)$ -times continuously differentiable at the integer interval endpoints.

Because $F_M(t)$ is a function that is joined together by polynomials in the different unit intervals, we now apply Eq. (3.195) to these intervals in turn and sum up the results. It becomes obvious that only the terms at $t = 0$ survive, so that

$$J_M = HP(0) + R_M(H), \quad (3.199)$$

where $R_M(H)$ is the remainder resulting from the uncanceled derivatives. It is obvious that, owing to the continuity conditions of the derivatives up to order $M-1$ and their vanishing at $t = \pm M$, the corresponding terms in Eq. (3.195) cancel out, but careful numerical integrations over polynomials $P(x)$ show that the order of the remainder is considerably higher:

$$R_M(H) = H^{2M+1} C_M P^{(2M)}(0)/(2M)! \quad (3.200)$$

with the coefficients [5]

$$C_1 = 1/6, \quad C_2 = -0.3, \quad C_3 = 1.702, \quad C_4 = -19.9, \quad C_5 \approx 385. \quad (3.201)$$

Hence, for sufficiently small interval size H and confined derivatives, the remainder can be ignored and Eq. (3.199) is then the simplest possible integration formula.

3.6.5 Concluding Remarks

The methods outlined so far are only a very brief and incomplete selection from the field of numerical analysis. Further presentations of numerical mathematics

in general and on splines in particular in a more rigorous manner can be found in references [3] and [8–14]. More material will follow in the next chapters, when the context for its presentation becomes more clearer. For example, the methods of multidimensional integrations and the solution of large systems of equations that are omitted here are easier to understand in connection with the finite-element method.

REFERENCES

1. Weiland, T. (1977). A discretization method for the solution of Maxwell's equations for six-component fields, *Electronics and Communications* **31**: 116–120.
2. Weiland, T. (1986). Die Diskretisierung der Maxwell-Gleichungen, *Physikalische Blätter* **42**: Nr. 7, 191–201.
3. Press, W. H., Flannery, B. P., Teukolsky, S. A. and Vetterling, W. T. (1988). *Numerical Recipes*, 3rd ed., Cambridge University Press.
4. Schwarz, H. R. (1986). *Numerische Mathematik*, Stuttgart: Teubner-Verlag, pp 135–139.
5. Kasper, E. (1987). An advanced method of field calculation in electron optical systems with axisymmetric boundaries, *Optik* **77**: 3–12.
6. Kasper, E. and Ströer, M. (1990). A new method for the calculation of magnetic lenses, *Nucl. Instrum. Meth. A* **298**: 1–9.
7. Abramowitz, M. and Stegun, I. A. (1970). *Handbook of Mathematical Functions*, p. 886, New York: Dover.
8. Ahlberg, J., Nilson, E. and Walsh, J. (1967). *The Theory of Splines and Their Applications*, New York & London: Academic Press.
9. Boor, de C. (1978). *A Practical Guide to Splines*, Berlin: Springer.
10. Hamming, R. W. (1962). *Numerical Methods for Scientists and Engineers*, New York: McGraw-Hill.
11. Hämmerlin, G. and Hoffmann, K. H. (1994). *Numerische Mathematik*, 4th ed., Berlin, Heidelberg, New York: Springer.
12. Schumaker, L. L. (1981). *Spline Functions, Basic Theory*, New York: Wiley.
13. Schwarz, H. R. (1986). *Numerische Mathematik*, Stuttgart: Teubner.
14. Watson, G. A. (1980). *Approximation Theory and Numerical Methods*, New York: Wiley.

CHAPTER IV

The Finite-Difference Method (FDM)

The finite-difference method (FDM) is historically the oldest numerical technique for solving boundary value problems for elliptic differential equations such as those derived in Chapter II. It was introduced by H. Liebmann as early as 1918 and is thus often called “*Liebmann’s method*” or the “*method of meshes*.” The associated mathematical theory is exhaustively studied in the literature, see for instance references [1] to [5]. Survey articles on the application of the FDM to field calculations in charged particle optics are given in the references [6] to [9]. In the past decades the FDM has lost much of its attraction with the introduction of the method of finite elements (FEM) into electron optics by E. Munro [10] and the development of powerful computers that has made it feasible.

The reason why we still deal here with the FDM in some detail is that it is often used in too restricted a manner and that it can be more powerful than is usually believed. Moreover, in the FEM, the full generality of arbitrarily irregular triangular grids is hardly ever used in practice because this would be rather circumstantial. If, however, the triangular grids are obtained by systematic subdivision of topologically quadrilateral ones, these can be *combined* with the FDM grids, and the fundamental difference between these two methods dwindles away.

With both methods, the discretization results in a large system of equations for the potentials at the nodes. Traditionally, the FDM equations are solved by iterative methods, whereas the FEM equations are solved by conjugate gradient techniques. However, this is not specific and is not absolutely necessary for these two methods, and, in fact, we shall find some common features of a general cell theory.

4.1 TWO-DIMENSIONAL MESHES

For reasons of conciseness we consider here a general elliptic partial differential equation (PDE) in two rectilinear coordinates (x, y) or (z, r) typically of the types dealt with in Chapter II. This equation will be solved for prescribed boundary conditions. The basic idea in both the FDM and the FEM is the following: if it is not possible to obtain an analytic solution, an approximation to it can be obtained by discretization. This means that we cast a

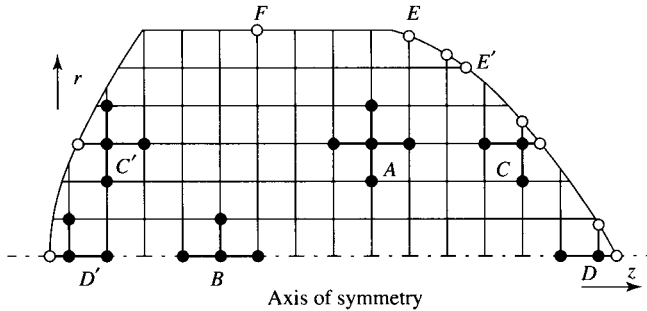


FIGURE 4.1 Square-shaped mesh with different kinds of mesh points: A regular inner point; B regular axial point; C , C' irregular inner points; D , D' irregular axial points; E , E' irregular boundary points; F regular boundary point. (See also Section 4.1.5 for more details).

two-dimensional grid over the whole domain and try to determine the potential at all its internal nodes, see Fig. 4.1. Of course, the accuracy would become better with decreasing mesh width—provided that the algorithm employed is numerically stable—but for practical reasons there are upper limits on the amount of information that can be stored in the computer. This means that the algorithm for discretizing the PDE should be as accurate as possible to achieve a given precision with an acceptable number of meshes.

A more serious problem is the fact that regular rectangular meshes hardly ever fit the given boundaries, so that many irregular meshes are obtained in their vicinity (see Fig. 4.1). It is therefore desirable to introduce curvilinear coordinates (u, v) that are adapted to these boundaries (see Fig. 4.2) and to solve the PDE in these coordinates. It may not be possible to remove all irregular meshes in this way, but an improvement will already have been achieved, if the irregular meshes are removed from sensitive parts, the surface of a curved cathode for example.

After the transformation, the grid in the (u, v) plane is rectangular and all its nodes can be specified by two labels i and k (see Fig. 4.2). Hence, any function $f(u, v)$ is then replaced by a two-dimensional array

$$f_{ik} := f(u_i, v_k), \quad (0 \leq i \leq M, 0 \leq k \leq N). \quad (4.1)$$

If this is known, then the techniques of Section 3.5, especially the bicubic spline algorithm, can be applied to determine any value within the grid by interpolation.

The FDM in three dimensions does not exhibit any essentially new aspects, apart from the obvious fact that we have to consider three coordinates (x, y, z) and (u, v, w) , and consequently three labels (i, j, k) . The main problem is now the strongly increased demand for memory. Some three-dimensional

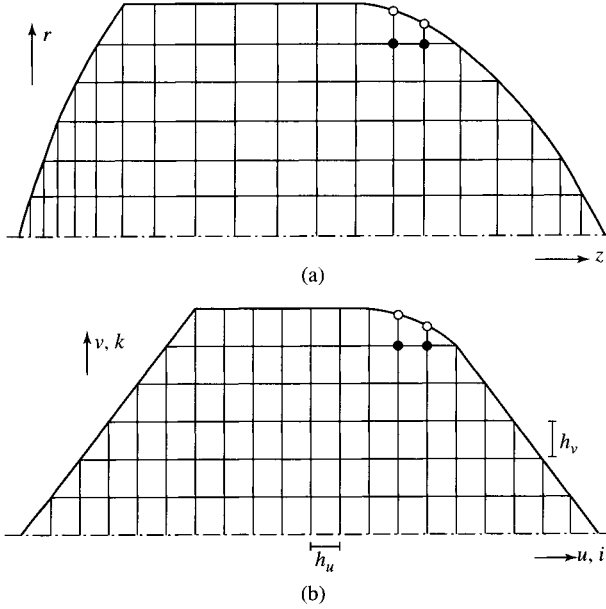


FIGURE 4.2 Removal of irregular points by mesh transformation: (a) deformed mesh in the original (z, r) -plane; (b) equidistant rectangular mesh in the (u, v) -plane, there remain only two irregular internal points. Any regular position can be specified by $u_i = ih_u$, $v_k = kh_v$.

calculations have been reported by Franzen [11] and Rouse and Munro [12], for example.

4.1.1 General Coordinate Transforms

Most frequently, the coordinate systems employed in practice are orthogonal, and a special case of the formalism is outlined in Section 3.1. Certainly this is highly advantageous and even necessary in the familiar five-point approximation. This demand however, may, be too restrictive and may not be correctly satisfied for machine-generated meshes. We therefore start now from a general nonsingular coordinate transform

$$x = \bar{x}(u, v), \quad y = \bar{y}(u, v) \quad (4.2)$$

and shall later drop the bars denoting the distinction between a function and its value, so that this does not cause confusion. Any scalar *invariant*, for example a potential, is then transformed according to

$$\Phi(x, y) = \Phi(\bar{x}(u, v), \bar{y}(u, v)) = \phi(u, v). \quad (4.3)$$

To facilitate the next task, the transformation of partial derivatives, we introduce some simplifying notation, for example,

$$\phi_{|u} \equiv \partial_u \phi := \partial \phi(u, v) / \partial u \quad (4.4a)$$

or $\phi_{|v}$, $\Phi_{|x}$, $\Phi_{|y}$, or any other derivative of first order. In an analogous manner the derivatives of second order are abbreviated as

$$\begin{aligned} \phi_{|uu} &\equiv \partial_{uu} \phi := \partial^2 \phi(u, v) / \partial u^2, \\ \phi_{|uv} &\equiv \partial_{uv} \phi := \partial^2 \phi(u, v) / \partial u \partial v, \end{aligned} \quad (4.4b)$$

and so on. The transformation of derivatives requires the nonsingular matrix

$$\mathbf{H} = \begin{pmatrix} x_{|u} & x_{|v} \\ y_{|u} & y_{|v} \end{pmatrix} \quad (4.5)$$

with the Jacobi determinant

$$J(u, v) = \det(\mathbf{H}) = x_{|u} y_{|v} - x_{|v} y_{|u} \neq 0. \quad (4.6)$$

Although the case $J < 0$ would be allowed, we shall exclude such transforms here and assume that J is *strictly* positive, which means that the orientation in the plane is conserved. The inverse matrix

$$\mathbf{H}^{-1} = \begin{pmatrix} u_{|x} & u_{|y} \\ v_{|x} & v_{|y} \end{pmatrix} = J^{-1} \begin{pmatrix} y_{|v} & -x_{|v} \\ -y_{|u} & x_{|u} \end{pmatrix}, \quad (4.7)$$

then certainly exists and can be used to transform the gradient according to:

$$\begin{aligned} \Phi_{|x} &= \phi_{|u} u_{|x} + \phi_{|v} v_{|x} = J^{-1} (\phi_{|u} y_{|v} - \phi_{|v} y_{|u}), \\ \Phi_{|y} &= \phi_{|u} u_{|y} + \phi_{|v} v_{|y} = J^{-1} (-\phi_{|u} x_{|v} + \phi_{|v} x_{|u}). \end{aligned} \quad (4.8)$$

The corresponding expressions for the derivatives of second order are not given here in their most general form as these circumstantial expressions can be avoided by means of a variational principle. Special cases of practical significance will be given in Section 4.1.3.

4.1.2 Variational Principles

In Chapter II, we have frequently encountered self-adjoint elliptic PDEs, which are ultimately a consequence of the variational principles (1.39) with (1.41) or

(1.50) with (1.45). The basic idea is now to carry out the Eq. (4.2) not in these PDEs but in the corresponding Lagrangian itself as the knowledge of Eq. (4.8) is already sufficient. This is then an application of the formalism (1.39), (1.40) with the specializations $n = 2$, $m = 1$, $x_1 = x$, $x_2 = y$, $y_1 \hat{=} \Phi(x, y)$.

Because the Eq. (4.2) has no special properties, we must consider general quadratic forms of the derivatives:

$$\bar{L} := \frac{1}{2}(\bar{A}\Phi_{|x}^2 + \bar{B}\Phi_{|y}^2) + \bar{C}\Phi_{|x}\Phi_{|y} + \bar{W}(x, y, \Phi) \quad (4.9)$$

and analogously

$$L := \frac{1}{2}(A\phi_{|u}^2 + B\phi_{|v}^2) + C\phi_{|u}\phi_{|v} + W(u, v, \phi), \quad (4.10)$$

where the coefficients $\bar{A}$, $\bar{B}$ and $\bar{C}$ are functions of (x, y) , whereas A , B , and C are then functions of (u, v) . The transform between these two Lagrangians must be carried out in such a way that the functional F , to be minimized, remains invariant. Considering the transformation of area elements by $dx dy = J(u, v) du dv$ we obtain

$$F := \iint \bar{L} dx dy = \iint \bar{L} J du dv = \iint L du dv = \min. \quad (4.11)$$

This implies $L = \bar{L}J$ for the function values of the integrands and consequently

$$W(u, v, \phi) = J(u, v)\bar{W}(x, y, \Phi). \quad (4.12)$$

The coefficients of the quadratic forms can be combined to form symmetric matrices, and these transform according to the tensor law

$$\begin{pmatrix} A & C \\ C & B \end{pmatrix} = J\mathbf{H}^{-1} \begin{pmatrix} \bar{A} & \bar{C} \\ \bar{C} & \bar{B} \end{pmatrix} \mathbf{H}^{-1T}, \quad (4.13)$$

where the symbol T denotes the transposed matrix. The evaluation of these matrix products yields

$$\begin{aligned} A &= J^{-1}(\bar{A}y_{|v}^2 + \bar{B}x_{|v}^2 - 2\bar{C}x_{|v}y_{|v}), \\ B &= J^{-1}(\bar{A}y_{|u}^2 + \bar{B}x_{|u}^2 - 2\bar{C}x_{|u}y_{|u}), \\ C &= J^{-1}(-\bar{A}y_{|u}y_{|v} - \bar{B}x_{|u}x_{|v} + \bar{C}(x_{|u}y_{|v} + y_{|u}x_{|v})). \end{aligned} \quad (4.14)$$

From Eqs. (4.13) or (4.14), it can be concluded that the determinant

$$AB - C^2 = \bar{A}\bar{B} - \bar{C}^2, \quad (4.15)$$

is an invariant. This must be strictly *positive* in the case of elliptic PDEs.

The evaluation of the Euler equations (1.39) for the Lagrangians given in Eqs. (4.9) and (4.10) now results in

$$\partial_x(\bar{A}\Phi_{|x} + \bar{C}\Phi_{|y}) + \partial_y(\bar{B}\Phi_{|y} + \bar{C}\Phi_{|x}) = \partial\bar{W}/\partial\Phi, \quad (4.16)$$

and analogously,

$$\partial_u(A\phi_{|u} + C\phi_{|v}) + \partial_v(B\phi_{|v} + C\phi_{|u}) = \partial W/\partial\phi, \quad (4.17)$$

which means that this general expression of the PDE is *form invariant*. If we carry out these product differentiations, we soon notice that the coefficients $\bar{C}$ or C are related to the derivatives $\Phi_{|xy}$ or $\phi_{|uv}$, respectively, so that it would be advantageous to remove them. However, this is possible only under very restrictive conditions.

Before we start making specializations and approximations, we rewrite Eq. (4.17) in another useful form. It is advantageous to introduce the *conjugate gradient* $\mathbf{G}(u, v)$ by means of its components

$$\begin{aligned} G_u(u, v) &:= A\phi_{|u} + C\phi_{|v}, \\ G_v(u, v) &:= B\phi_{|v} + C\phi_{|u}. \end{aligned} \quad (4.18)$$

Then the Lagrangian (4.10) can be rewritten as

$$L = (G_u\phi_{|u} + G_v\phi_{|v})/2 + W(u, v, \phi). \quad (4.19)$$

A more important fact is the possibility of rewriting Eq. (4.17) as

$$\operatorname{div} \mathbf{G} = G_{u|u} + G_{v|v} = \partial W/\partial\phi =: S(u, v, \phi). \quad (4.20)$$

This can be integrated once by means of Gauss's theorem for any domain D with closed boundary ∂D in the u - v plane:

$$\iint_D \operatorname{div} \mathbf{G} \, du \, dv = \iint_D S \, du \, dv = \oint_{\partial D} \mathbf{G} \cdot \mathbf{n} \, ds, \quad (4.21)$$

$\mathbf{n}$ being the normal in the outward direction and ds , as usual, the line element on the positively oriented curve ∂D . The curve normal has the components $\mathbf{n} = (v', -u')$ so that we finally have the integral expression

$$\oint_{\partial D} (v'(s)G_u - u'(s)G_v) \, ds = \iint_D S(u, v, \phi) \, du \, dv. \quad (4.22)$$

This is valid exactly, regardless of the choice of the coordinates (u, v) a cartesian interpretation is not necessary.

4.1.3 Orthogonal Meshes

As mentioned earlier, it would be advantageous to remove the coupling coefficients $\bar{C}$ and C from the variational calculus as this would obviously lead to a simplification. In fact, the coefficient $\bar{C}(x, y)$ never appeared in the PDEs derived in the preceding chapters, and it was included here only for reasons of completeness of the theory. If we set $\bar{C}(x, y) \equiv 0$, Eq. (4.14) still give

$$C = -J^{-1}(\bar{A}y_{|u}y_{|v} + \bar{B}x_{|u}x_{|v}), \quad (4.23)$$

which does not vanish in general as the condition for orthogonality is

$$x_{|u}x_{|v} + y_{|u}y_{|v} \equiv 0. \quad (4.24)$$

This is a special case of the formalism outlined in Section 3.1.1, if we identify

$$x_1 = x, \quad x_2 = y, \quad q_1 = u, \quad q_2 = v, \quad (x_3 = q_3 = z). \quad (4.25)$$

There are only two choices that enable us to obtain $C \equiv 0$:

Isotropic Material Coefficients

$$\bar{A}(x, y) \equiv \bar{B}(x, y) = \bar{\varepsilon}(x, y) \quad (4.26)$$

The Eq. (4.14) then simplifies to

$$\begin{aligned} A &= \bar{\varepsilon}(x_{|v}^2 + y_{|v}^2)/J \equiv \bar{\varepsilon} L_v^2/J, \\ B &= \bar{\varepsilon}(x_{|u}^2 + y_{|u}^2)/J \equiv \bar{\varepsilon} L_u^2/J, \end{aligned} \quad (4.27)$$

because the expressions in parentheses are just the components of the metric tensor (3.5) with (3.7). However, these coefficients can be simplified still further because in two dimensions the Jacobian becomes $J = L_u L_v$, hence, we arrive at

$$A = \bar{\varepsilon} L_v/L_u, \quad B = \bar{\varepsilon} L_u/L_v. \quad (4.28)$$

The invariance Eq. (4.15) is evidently satisfied as it is obvious that $AB = \bar{\varepsilon}^2$. This kind of transformation is the one most frequently employed because in most practical applications the material properties *are* isotropic, but care must be taken if the condition Eq. (4.26) is not satisfied.

Conformal Mappings

These are still more restrictive special cases as the Cauchy-Riemann equations lead to $L_u \equiv L_v$. Hence $A = B = \varepsilon$ is then an *invariant* material coefficient.

The PDE now simplifies to

$$\partial_u(\varepsilon\phi|_u) + \partial_v(\varepsilon\phi|_v) = S(u, v, \phi), \quad (4.29)$$

which is form invariant with respect to conformal mappings.

Although complex functions are commonly used to find solutions of the two-dimensional Laplace equation under given constraints, this is not the only way to use them. An alternative is the construction of suitable curvilinear meshes. An example of this is the exponential function

$$x + iy = \exp(u + iv) \rightarrow x = e^u \cos v, \quad y = e^u \sin v, \quad (4.30)$$

which is a favorable form of polar coordinates, see Fig. 4.3. The Laplace equation then has the strikingly simple form

$$\phi|_{uu} + \phi|_{vv} = 0 \quad (4.31)$$

Affine Distortions

Another way of obtaining $C \equiv 0$ in Eq. (4.23) is to choose $y|_u y|_v \equiv 0$ and $x|_u x|_v \equiv 0$. This is easily achieved with a transformation of the kind

$$x = \bar{x}(u), \quad y = \bar{y}(v), \quad (4.32)$$

which means that the alterations of the scales in the two directions are independent of each other, see Figs. 4.2a,b.

Such a grid can easily be adapted to systems of electrodes or pole pieces with rectangular shapes if the data configurations are not too complicated. To construct it, we first count all the corners in the x -direction sequentially from 0 to L and give them *mesh numbers* $i_0 \dots i_L$ with $i_0 = 0$ and $i_L = M$ in a strictly monotonic sequence. With a reasonable choice of mesh size $\Delta u = h$ we now have the sampling data

$$u(i_n) = i_n h, \quad \bar{x}(u(i_n)) = \tilde{x}_n, \quad n = 0 \dots L \quad (4.33)$$

for a cubic spline, $\tilde{x}_n$ being the given x -coordinates of the corners. This spline is now evaluated for all values $u_i = ih, i = 0 \dots M$, and gives the real mesh lines at $x_i = \bar{x}(u_i)$. This spline is constructed correctly if $\bar{x}'(u) > 0$ for all u . In the same manner we can construct the grid in the y -direction.

The affine transformation (4.32) is so simple that it is not necessary to start from a variational principle although this would be favorable. The PDE need not be self-adjoint but must not contain a mixed derivative of second order. The PDE

$$\bar{A}\Phi|_{xx} + \bar{B}\Phi|_{yy} + \bar{a}\Phi|_x + \bar{b}\Phi|_y = \bar{S}(x, y, \Phi) \quad (4.34)$$

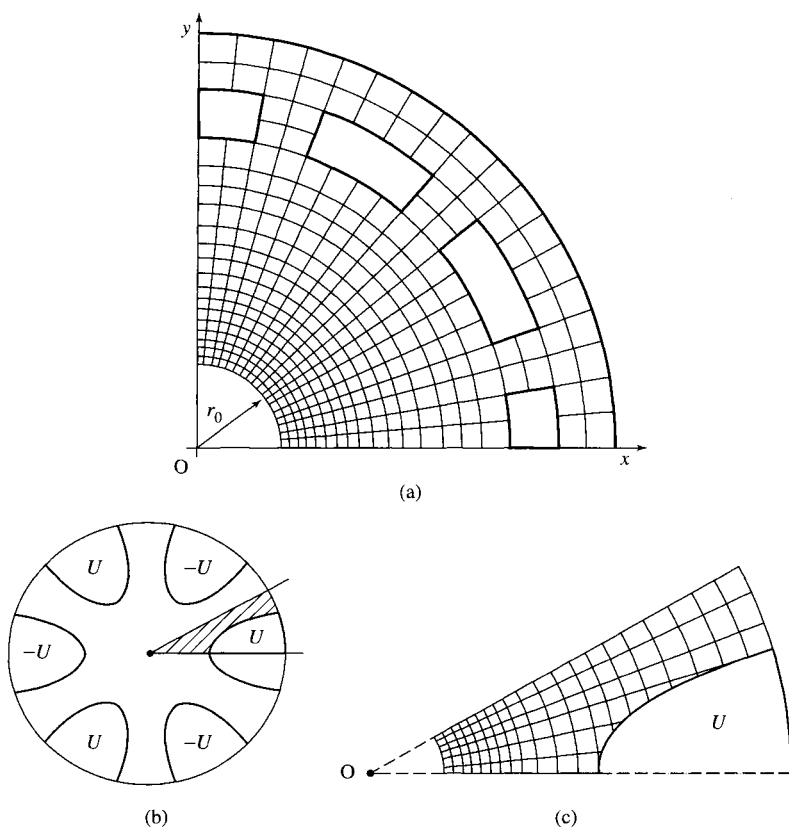


FIGURE 4.3 Applications of an exponentially expanding mesh: (a) quarter of the cross section through a twelve pole element; (b) cross section through a hexapole element; (c) expanding mesh in one relevant section. (Note the problem of reaching the origin O ! see Chapter VII).

with coefficients depending on x and y is again form invariant in the sense that

$$A\phi_{|uu} + B\phi_{|vv} + a\phi_{|u} + b\phi_{|v} = S(u, v, \phi) \quad (4.35)$$

holds the coefficients being functions of u and v . Their transform becomes now

$$A = \bar{A} \bar{y}' / \bar{x}', \quad B = \bar{B} \bar{x}' / \bar{y}', \quad S = \bar{S} \bar{x}'(u) \bar{y}'(v), \quad (4.36)$$

which are in agreement with Eqs. (4.12) and (4.14), (here we have dropped some arguments for conciseness). The two new transformations are

$$\begin{aligned} a &= \bar{a} \bar{y}' - \bar{A} \bar{x}'' \bar{y}' / \bar{x}'^2, \\ b &= \bar{b} \bar{x}' - \bar{B} \bar{y}'' \bar{x}' / \bar{y}'^2. \end{aligned} \quad (4.37)$$

From these equations, it becomes obvious that the functions $\bar{x}(u)$ and $\bar{y}(v)$ must be at least twice continuously differentiable.

Self-adjoint PDEs are now characterized by the additional relations

$$\bar{a} = \bar{A}|_x, \quad \bar{b} = \bar{B}|_y, \quad (4.38)$$

which lead to the corresponding equations

$$a = A|_u, \quad b = B|_v, \quad (4.39)$$

so that Eq. (4.35) can be rewritten as

$$\partial_u(A\phi|_u) + \partial_v(B\phi|_v) = S, \quad (4.40)$$

as must be possible.

4.1.4 Sources and Nonlinearities

So far we have not specified the function $W(x, y, \Phi)$ and its transform $\bar{W}(u, v, \phi)$ with derivative $S = \partial W / \partial \phi$. The system of equations resulting from the FDM and from the FEM can be solved by means of linear algebra only if S is a linear function of ϕ and W consequently a quadratic one. We therefore assume now that

$$W(u, v, \phi) = \frac{1}{2}D(u, v)\phi^2 + Q(u, v)\phi + \text{const.}, \quad (4.41)$$

$$S(u, v, \phi) = \partial W / \partial \phi = D(u, v)\phi + Q(u, v). \quad (4.42)$$

A special problem arises if the coefficient D contains a free parameter that must become an *eigen value*; this problem cannot be discussed here for reasons of space. Another complicated situation arises in the cases of electron guns with strong *space charge* because the latter depends very sensitively on the potential. This problem cannot be dealt with here, because the necessary techniques exceed the topics of this volume.

A quite different kind of complication arises in the calculation of magnetic lenses with saturation effects in the yokes. The appropriate variational principle is then Eq. (1.50) with the Lagrangian (1.45). In the linear approximation, the function $\Lambda(r, B)$ is given by (1.52). For round lenses, the flux potential $\Psi(z, r)$ defined by Eq. (2.40) and satisfying the PDE (2.46) is adequate. In the case of nonlinearity the material function $\nu(z, r, B)$ depends on the absolute field strength $B = |B|$, given by Eq. (2.49). Apart from the nonlinearity of ν , this problem fits the general mathematical scheme. The basic structure is as follows.

The original PDE can be cast in the form

$$\partial_x (\bar{\varepsilon}(x, y) f(p) \Phi_{|x}) + \partial_y (\bar{\varepsilon}(x, y) f(p) \Phi_{|y}) = \bar{Q}(x, y) \quad (4.43a)$$

with

$$p := \Phi_{|x}^2 + \Phi_{|y}^2 \quad (4.43b)$$

in which $\bar{\varepsilon}(x, y)$ and $\bar{Q}(x, y)$ are well-known functions and the structure of the function $f(p)$ is also given, but the appropriate argument is known only after solution of the PDE. This is a nonlinear case. The transformed PDE becomes

$$\partial_u (A f(p) \phi_{|u} + C f(p) \phi_{|v}) + \partial_v (B f(p) \phi_{|v} + C f(p) \phi_{|u}) = Q(u, v). \quad (4.44)$$

We now introduce the metric tensor (3.5); here

$$\begin{aligned} s_u &:= g_{uu}^{1/2} = (x_{|u}^2 + y_{|u}^2)^{1/2} \\ s_v &:= g_{vv}^{1/2} = (x_{|v}^2 + y_{|v}^2)^{1/2} \\ s_u s_v \cos \beta &= g_{uv} = x_{|u} x_{|v} + y_{|u} y_{|v} \\ J &= x_{|u} y_{|v} - y_{|u} x_{|v} = s_u s_v \sin \beta \end{aligned} \quad (4.45)$$

(see Fig. 4.4) and obtain the coefficients

$$A = \bar{\varepsilon} \frac{s_v}{s_u \sin \beta}, \quad B = \bar{\varepsilon} \frac{s_u}{s_v \sin \beta}, \quad C = -\bar{\varepsilon} \cot \beta. \quad (4.46)$$

These are not altered by the nonlinearity. The additional task here is the calculation of p :

$$p = (A \phi_{|u}^2 + B \phi_{|v}^2 + 2C \phi_{|u} \phi_{|v}) J^{-1} \bar{\varepsilon}^{-1}, \quad (4.47)$$

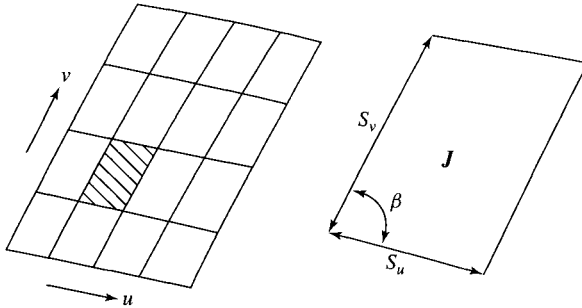


FIGURE 4.4 Locally parallelepipedal mesh cells showing the notation; J is the area of the cell.

which is to be introduced into the function $f(p)$. This problem can be solved only in an iterative way.

4.1.5 Classification of Configurations

Many different situations can arise if a square-shaped grid is spread over an arbitrary domain. Such situations are already shown in Fig. 4.1. The vast majority of points are located in the interior, and each has four closest neighbors and four next-nearest neighbors in the diagonal direction; we shall call them *regular* points. Special cases arise on *symmetry* axes (situations A or B). These do not cause difficulties, as the missing neighborhood can be completed by symmetry operations.

Exceptional situations arise if a mesh point is located near a boundary, so that some of the grid points are not available. We then have to use the intersections of the grid with the boundary in order to obtain at least four neighbors (situation C). This can also happen near a symmetry axis (situation D). We shall call all these *irregular* points.

Quite often, the domain of solution is composed of different materials, separated from each other by *internal* boundaries (see Fig. 4.5). The internal boundary points do not fit the grid, but the potentials in these must be considered as additional *variables*. We shall see that this case is no obstacle for the application of the FDM. In this context, it should be noted that the generation of general triangular meshes for the FEM is also no easy task.

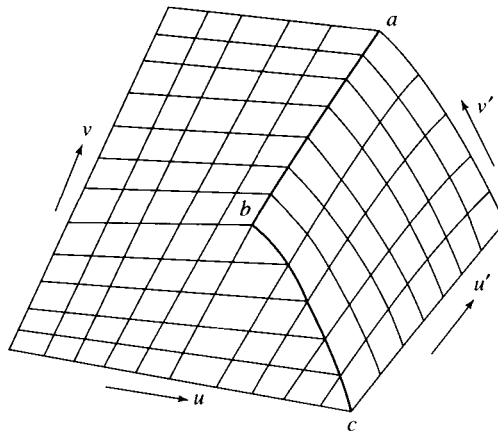


FIGURE 4.5 Two connected sets of meshes: the line (abc) may be an *inner* boundary between two materials of different properties. It consists entirely of irregular points.

4.2 FIVE-POINT CONFIGURATIONS

Consideration of the four closest neighbors of each internal mesh point is the lowest possible approximation and also the simplest one. It requires orthogonal coordinates u, v , hence $C \equiv 0$. In the derivation of the general formula it is not necessary to distinguish between regular and irregular configurations. The former will finally be obtained as a special case of the general formula.

Depending on the choice of the PDE, there are two different methods: the Taylor-series method for the discretization of Eq. (4.35) and the integration method for self-adjoint Eq. (4.40). We shall start with the former more general case and assume Eq. (4.42) for the source term.

4.2.1 The Taylor Series Method

In order to cast the calculations in a concise form, we shall adopt the simplified notation, shown in Fig. 4.6a: the central point (i, k) has the label 0 and its

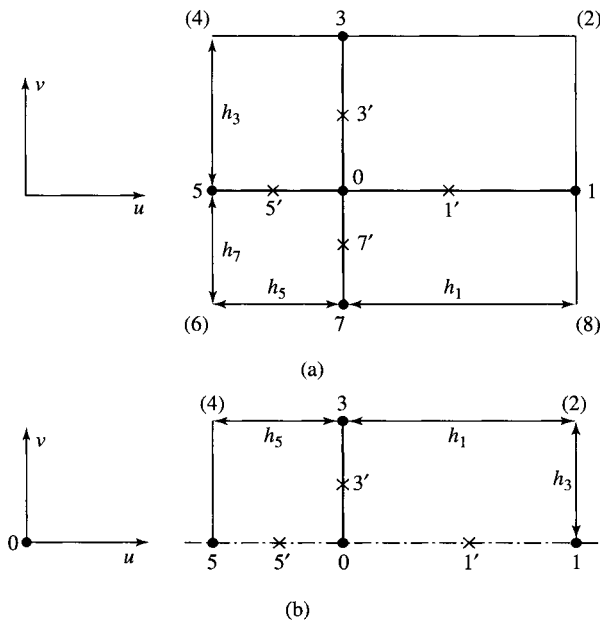


FIGURE 4.6 (a) Asymmetric five-point configuration. The positions with a prime on the number are side midpoints. (b) The corresponding axial four-point configuration.

neighbors have the labels 1 to 8 in anticlockwise orientation. Here we need only the *odd* labels.

The approximation consists simply in replacing the partial derivatives by their three-point approximations according to Eqs. (3.67) and (3.68), respectively. Thus,

$$\phi_{|u} = \{h_5^2\phi_1 - h_1^2\phi_5 + (h_1^2 - h_5^2)\phi_0\} / \{h_1h_5(h_1 + h_5)\}, \quad (4.48)$$

$$\phi_{|uu} = 2\{h_5\phi_1 + h_1\phi_5 - (h_1 + h_5)\phi_0\} / \{h_1h_5(h_1 + h_5)\}, \quad (4.49)$$

with an error of second order for $\phi_{|u}$ and of first order for $\phi_{|uu}$. Similar approximations hold for $\phi_{|v}$ and $\phi_{|vv}$. These are now introduced into Eq. (4.35), which hence becomes a finite-difference equation for the five points considered. It can be solved for the *central* value ϕ_0 :

$$\phi_0 = c_0 + c_1\phi_1 + c_3\phi_3 + c_5\phi_5 + c_7\phi_7, \quad (4.50)$$

with the coefficients

$$\left. \begin{aligned} c_1 &= \frac{2A_0 + a_0h_5}{h_1(h_1 + h_5)N}, & c_3 &= \frac{2B_0 + b_0h_7}{h_3(h_3 + h_7)N}, & c_0 &= \frac{-Q_0}{N} \\ c_5 &= \frac{2A_0 - a_0h_1}{h_5(h_1 + h_5)N}, & c_7 &= \frac{2B_0 - b_0h_3}{h_7(h_3 + h_7)N} \\ N &= D_0 + \frac{2A_0 + a_0(h_5 - h_1)}{h_1h_5} + \frac{2B_0 + b_0(h_3 - h_7)}{h_3h_7}. \end{aligned} \right\} \quad (4.51)$$

The specialization to *equal* mesh sizes h for regular points brings some simplifications of the coefficients, but — more important — a gain of accuracy because the approximation error for $\phi_{|uu}$ and $\phi_{|vv}$ becomes now of second order. Hence the local discretization error of the Eq. (4.50) becomes of the order h^4 . Because this concerns the vast majority of regular points, smooth deformations should be made in the *grid*, as in Section 4.1.1, and not by choosing unequal mesh sizes in (4.50).

It is fairly easy to incorporate symmetry conditions into this approximation. If, for example, the function $\phi(u, v)$ must have the property $\phi(u, v) = \phi(-u, v)$ and we choose $u_0 = 0$, then the value ϕ_5 can either be set by means of this property or, better, eliminated. This is simply done by writing $h_5 = h_1$ and doubling the coefficient c_1 . A more serious problem arises on the *axis of rotational symmetry*, where the coefficient of $V_{|r}$, here identified with $\phi_{|v}$ becomes *singular*; the corresponding configuration is shown in Fig. 4.6b.

On generalizing Eq. (2.58), we consider here a PDE in which we have

$$b(u, v) = \tilde{b}(u, v)/v, \quad (4.52)$$

$\tilde{b}(u, v)$ being a finite and even function of v for any value of u . Thanks to this symmetry property, we can make the approximation

$$\phi_{|v}(u, 0) = 0, \quad \lim_{v \rightarrow 0} [v^{-1} \phi_{|v}(u, 0)] \equiv \phi_{|vv}(u, 0) = 2(\phi_3 - \phi_0)/h_3^2 \quad (4.53)$$

and find $c_7 = 0$ and

$$\left. \begin{aligned} c_3 &= 2(B_0 + \tilde{b}_0)N^{-1}h_3^{-2} \\ N &= D_0 + \frac{2A_0 + a_0(h_5 - h_1)}{h_1 h_5} + 2(B_0 + \tilde{b}_0)h_3^{-2}. \end{aligned} \right\} \quad (4.54)$$

The accuracy of this approximation is practically the same as for the general five-point formula.

4.2.2 The Ring-Integral Method

The ring-integral method is applicable to self-adjoint PDEs for which it has certain advantages over the Taylor-series method. We now choose a rectangular domain R of integration with a positively oriented periphery C , as shown in Fig. 4.7a,b: this periphery passes through the midpoints of the meshlines.

We now apply Gauss's integral theorem to this domain. In the case of Eq. (4.40) with Eq. (4.42), the general Eq. (4.22) simplifies to

$$\oint [A\phi_{|u} dv - B\phi_{|v} du] = \iint_R (D\phi + Q) du dv. \quad (4.55)$$

So far, this formula is still exactly valid, but its practical evaluation requires some simplifications, which mean discretizations. Thus on the right-hand side, the factor $\phi = \phi_0$ is considered as constant and taken in front of the corresponding integral. Likewise, the normal derivatives $\phi_{|u}$ and $\phi_{|v}$ are assumed to be constant on the corresponding periphery line and taken out of the integral. To cast the notation in a reasonably concise form, we denote the side midpoints by $1', 3', 5', 7'$; for example, we then have

$$\begin{aligned} 1' &\triangleq [(u_i + u_{i+1})/2, v_k] \triangleq [u_{1'}, v_k], \\ 3' &\triangleq [u_i, (v_k + v_{k+1})/2] \triangleq [u_i, v_{3'}], \end{aligned} \quad (4.56)$$

and so on. The corresponding normal derivatives are then simply

$$\phi_{|u}(1') = (\phi_1 - \phi_0)/h_1, \quad \phi_{|v}(3') = (\phi_3 - \phi_0)/h_3, \quad (4.57)$$

and other analogous expressions.

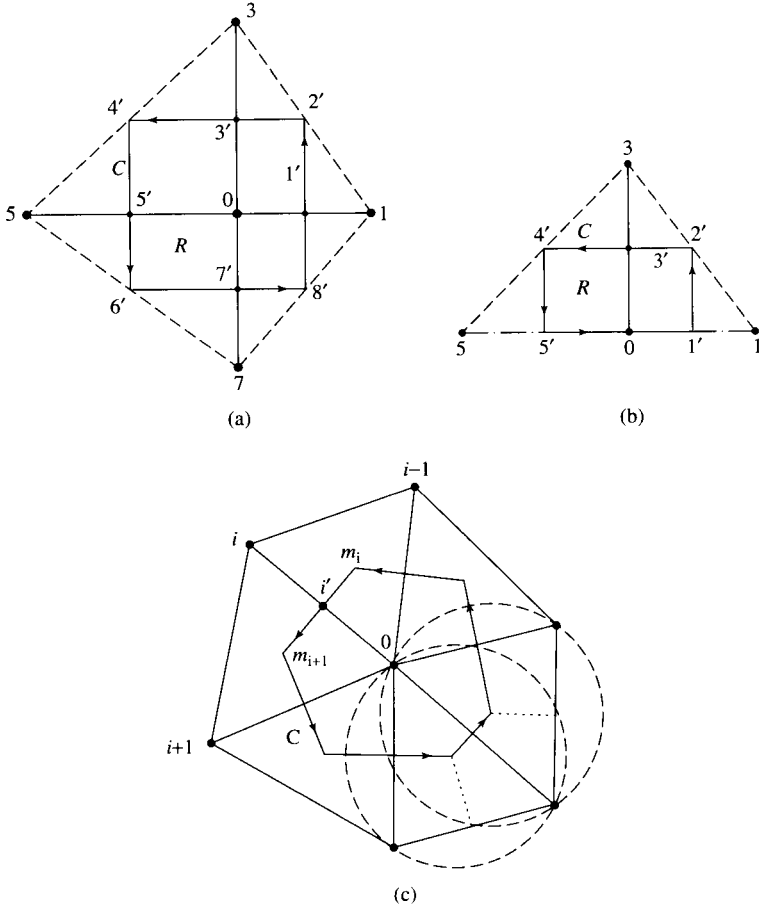


FIGURE 4.7 Closed path for the line integral: (a) for an arbitrary five-point configuration; (b) for an axial four-point configuration; (c) for an arbitrary polygonal configuration of $N + 1$ points, here $N = 6$ for example.

Considering that the integration on two sides of the periphery proceeds in the negative direction and exchanging the limits and signs, we find

$$\begin{aligned}
 & \frac{\phi_1 - \phi_0}{h_1} \int_{v_{7'}}^{v_{3'}} A(u_{1'}, v) dv + \frac{\phi_3 - \phi_0}{h_3} \int_{u_{5'}}^{u_{1'}} B(u, v_{3'}) du \\
 & + \frac{\phi_5 - \phi_0}{h_5} \int_{v_{7'}}^{v_{3'}} A(u_{5'}, v) dv + \frac{\phi_7 - \phi_0}{h_7} \int_{u_{5'}}^{u_{1'}} B(u, v_{7'}) du \\
 & = \phi_0 \int_{u_{5'}}^{u_{1'}} \int_{v_{7'}}^{v_{3'}} D(u, v) du dv + \int_{u_{5'}}^{u_{1'}} \int_{v_{7'}}^{v_{3'}} Q(u, v) du dv. \quad (4.58)
 \end{aligned}$$

Because $A > 0$, $B > 0$, and $u_{1'} > u_{5'}$, $v_{3'} > v_{7'}$, the integrals on the left-hand side are always strictly *positive*, a property that is advantageous with respect to the problem of solving the whole set of finite-difference equations.

The remaining integrations may become tedious, and it is often not necessary to carry them out exactly because we can assume that the coefficients A , B , D , and Q should not vary strongly if the derivatives are approximated by simple finite differences. Hence, we apply again the midpoint rule of integration, for example,

$$\int_{v_{7'}}^{v_{3'}} A(u_1, v) dv = A(u_{1'}, v_1) \int_{v_{7'}}^{v_{3'}} dv = \frac{1}{2} A_{1'}(h_3 + h_7), \quad (4.59)$$

the coefficient $A_{1'}$ now being evaluated at the midpoint $1'$ of the mesh line. Continuing in this manner with the other terms and collecting up all these expressions, we arrive at:

$$\begin{aligned} & \frac{h_3 + h_7}{2h_1} A_{1'}(\phi_1 - \phi_0) + \frac{h_1 + h_5}{2h_3} B_{3'}(\phi_3 - \phi_0) \\ & + \frac{h_3 + h_7}{2h_5} A_{5'}(\phi_5 - \phi_0) + \frac{h_1 + h_5}{2h_7} B_{7'}(\phi_7 - \phi_0) \\ & = \frac{1}{4}(h_1 + h_5)(h_3 + h_7)(D_0\phi_0 + Q_0). \end{aligned} \quad (4.60)$$

In the case of an axial mesh point (Fig. 4.7b), only the upper half part of the integral expressions (4.58) (the part with $v \geq 0$) must be used. It is generally *not* permissible to consider the coefficients as slowly varying, as in the case of the cylindric multipole Eq. (2.10) they contain a strongly varying common factor r^α .

4.2.3 Some Remarks

The approximation (4.60) can also be obtained directly from the PDE (4.40), as the derivatives on the left-hand side can immediately be replaced by finite differences:

$$\begin{aligned} (A\phi_{|u})_{1'} &= A_{1'}(\phi_1 - \phi_0)/h_1 + O(h_1^2), \\ (A\phi_{|u})_{5'} &= A_{5'}(\phi_0 - \phi_5)/h_5 + O(h_5^2). \end{aligned} \quad (4.61)$$

Next, the expression $\partial(A\phi_{|u})/\partial u$ is again replaced by a difference expression, the spacing now being $(h_1 + h_5)/2$, hence

$$\frac{\partial}{\partial u}(A\phi_{|u}) = \frac{2}{h_1 + h_5} [h_1^{-1} A_{1'}(\phi_1 - \phi_0) + h_5^{-1} A_{5'}(\phi_5 - \phi_0)] + O(h). \quad (4.62)$$

The order again becomes h^2 if $h_1 = h_5$. If we approximate the term $\partial(B\phi_v)/\partial v$ analogously and put all this together, we again arrive at Eq. (4.60), (all terms are to be multiplied with a common factor $(h_1 + h_5)(h_3 + h_7)/4$). The approximation for ϕ_0 is again of third order and of fourth order in the case of equal mesh widths, hence Eqs. (4.50) with (4.51) and (4.60) are of comparable accuracy.

In *both* approximations (4.58) and (4.60), the resulting system of equations is symmetric, and for $D(u, v) \geq 0$, also positively definite. To see this, we introduce the two-dimensional numbering that is the mathematically correct one:

$$\begin{aligned} 0 &:= (m, n), & 1 &:= (m+1, n), & 3 &:= (m, n+1), \\ 5 &:= (m-1, n), & 7 &:= (m, n-1). \end{aligned} \quad (4.63)$$

The mesh widths then becomes $h_1 = u_{m+1} - u_m$, $h_3 = v_{n+1} - v_n$, etc. For reasons of conciseness we also introduce the side midpoints $p_m := (u_m + u_{m+1})/2$, $q_n := (v_n + v_{n+1})/2$, whereupon we obtain the coefficients

$$A_{m,n} = (u_{m+1} - u_m)^{-1} \int_{q_{n-1}}^{q_n} A(p_m, v) dv, \quad (4.64a)$$

$$B_{m,n} = (v_{n+1} - v_n)^{-1} \int_{p_{m-1}}^{p_m} B(u, q_n) du, \quad (4.64b)$$

$$Q_{m,n} = \int_{p_{m-1}}^{p_m} \int_{q_{n-1}}^{q_n} Q(u, v) du dv, \quad (4.64c)$$

$$D_{m,n} = \int_{p_{m-1}}^{p_m} \int_{q_{n-1}}^{q_n} D(u, v) du dv, \quad (4.64d)$$

which can also be approximated by means of the midpoint rule of integration. The coefficients $A_{m,n}$ and $B_{m,n}$ are certainly positive, and for the coefficients $D_{m,n}$ we shall assume $D_{m,n} \geq 0$. Then the total system of five-point equation can be rewritten concisely as

$$\begin{aligned} &A_{m-1,n}(\phi_{m-1,n} - \phi_{m,n}) + A_{m,n}(\phi_{m+1,n} - \phi_{m,n}) \\ &+ B_{m,n-1}(\phi_{m,n-1} - \phi_{m,n}) + B_{m,n}(\phi_{m,n+1} - \phi_{m,n}) \\ &= D_{m,n}\phi_{m,n} + Q_{m,n}, \quad (1 \leq m \leq M-1, 1 \leq n \leq N-1). \end{aligned} \quad (4.65)$$

Although this is certainly not recommended, it would be possible to rewrite this system in a one-dimensional form but, if we were to do this, some elementary calculation would reveal that the system matrix is symmetric and diagonally dominant. Without rearrangement, it is already obvious that these properties

hold for the rows and columns separately. This means that if one of the labels m or n is kept fixed, the subsystem resulting for the other label has a positively definite tridiagonal structure. This property might be helpful for the solution of the system. These algebraic properties are studied in great details and are a standard topic in the mathematical literature; see, for instance, references [1–5].

4.2.4 Generalization of the Method

The method of ring-integration can easily be generalized for *irregular* configurations, as is shown in Fig. 4.7c. Then, for each considered node (0) and its N closest neighbors, the corresponding system of *triangular* cells must be determined. We presume here that these are oriented positively and that the numbering of neighbor points is *cyclical modulo N* . Thereafter, for each triangle its including circle is to be determined, as is shown for two examples. The path C of integration becomes now the periphery of the polygon P , obtained by joining subsequent centers of such circles, for instance $\mathbf{m}_i$ and $\mathbf{m}_{i+1}$ with distance $s_i = |\mathbf{m}_{i+1} - \mathbf{m}_i|$. Because of the elementary geometrical rules, this line passes perpendicularly through the midpoint i' of the common triangle side having the length $h_i = |\mathbf{r}_i - \mathbf{r}_0|$. This construction is reasonable only if none of the angles in the configuration is obtuse, so that never does any center fall out of the corresponding triangle.

This discretization becomes reasonably simple only for the *isotropic* PDE (4.29), Eq. (4.55) then rewritten as

$$\oint \varepsilon \phi_{|n} ds = \iint_p (D\phi + Q) du dv.$$

Now (with number i) the approximation

$$\int_{m_i}^{m_{i+1}} \varepsilon \phi_{|n} ds = \varepsilon_{i'} (\phi_i - \phi_0) s_i / h_i$$

is made on each polygon side, while the double integral over the sources on the right-hand side is simply approximated by $(D_0\phi_0 + Q_0) A_P$, A_P denoting the *area* of the polygonal domain. Hence, altogether we arrive at

$$\sum_{i=1}^N \varepsilon_{i'} (\phi_i - \phi_0) s_i / h_i = (D_0\phi_0 + Q_0) A_P.$$

Again, the corresponding system matrix is symmetric and diagonally dominant. The discretization error is again of third order and of fourth one in symmetrical configurations.

With $A = B = \varepsilon$ and $N = 8$, Eq. (4.60) now becomes a simple special case of this result, as in triangles with one rectangular corner now the center of the including circle becomes the midpoint of the hypotenuse. Hence, all centers are located on diagonals (points $2'$, $4'$, $6'$, and $8'$), so that $s_2 = s_4 = s_6 = s_8 = 0$ and the corresponding potentials cancel out.

In the next section, we shall derive formulas in which the diagonal points will give a non-vanishing contribution, so that the accuracy of the discretization can become better. Other improvements will be presented in Section 4.5 and in Chapter V.

4.3 NINE-POINT CONFIGURATIONS

We have outlined the derivation of five-point formulas in their quite general form as they can be employed in a fairly large class of problems. We shall see later that even irregular internal boundaries with different material coefficients on either side and consequently discontinuous normal derivatives are no obstacle. However, because a high accuracy is required, it is desirable to find more accurate approximations, so that larger mesh sizes would give as good results, and the demand for memory could be reduced correspondingly.

In this context, it is appropriate to consider the neighboring points in diagonal directions as well, which produces a nine-point configuration, as shown in Fig. 4.8. The necessary calculations will of course become more complicated, and the number of exceptional cases will become larger: if a diagonal point is missing in the neighborhood of a boundary, the corresponding five-point formula must then be evaluated.

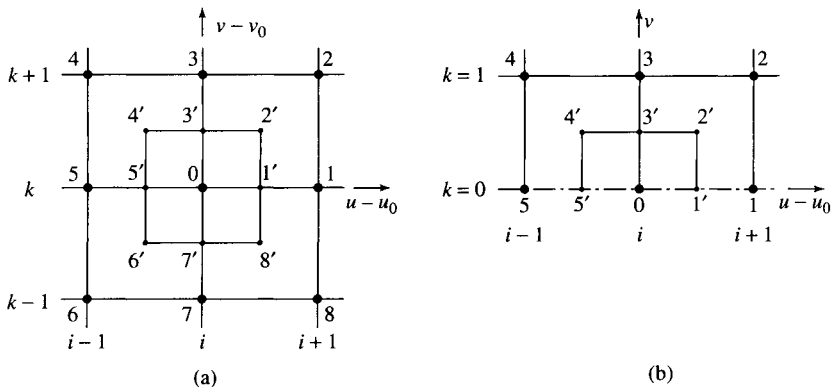


FIGURE 4.8 Connection between the local numbering with the global double indexing by i and k : (a) for a complete nine-point scheme; (b) for an axial configuration.

For a given structure of the PDE, the nine-point approximations can assume different forms, and in fact a great variety of them have been published (see refs [13–18]). This is not surprising as the task of determining a central value ϕ_0 from its neighbors $\phi_1 \dots \phi_8$ by some kind of interpolation has no unique answer. Different formulas for the same configuration can be considered as equivalent if their discretization error is of the same order of magnitude.

Traditionally, the FDM is applied only to orthogonal meshes, and in the five-point approximation this assumption is even necessary. However, if we consider configurations of seven or more points, this restriction can be given up, and we shall do this here because orthogonality is often too restrictive. As far as the author knows, the resulting approximation has not been published before. The method is based on the minimization of the functional F given in Eq. (4.11) in context with Eq. (4.10) and shows that the five-point approximation is a special case.

4.3.1 Approximation in One Mesh

Strictly speaking, all mesh points would need double indexing, and the mid-points between these would even require half-integer values. To cast the calculations in a reasonably concise form, we shall adopt the notation shown in Figs. 4.8 and 4.9 : the central node will have the label 0 and its neighbours are counted from 1 to 8 in the positive sense. The *inner* midpoints in the main or diagonal directions will be given the labels $1', \dots, 8'$. The *outer* midpoints are similarly given the labels $1'', \dots, 8''$ in the anticlockwise sense, but those points will finally not appear in the mesh formula.

We now consider the approximation of the functional F in one of the rectangular meshes, for example, the one shown in Fig. 4.9 and introduce

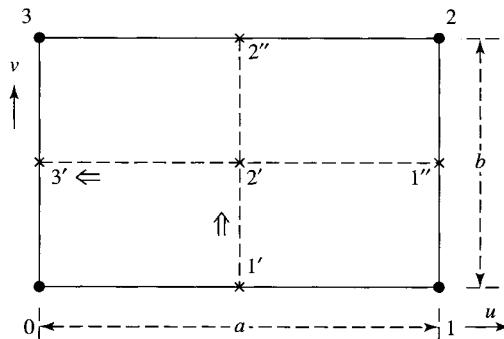


FIGURE 4.9 Local notation in a particular rectangular element (see also Fig. 4.8). The arrows indicate the direction of accumulation with respect to the mesh formula referring to the node 0.

the temporary abbreviations

$$a := h_1 = u_1 - u_0, \quad b := h_3 = v_3 - v_0 \quad (4.66)$$

for the mesh widths. The algorithm will be set up in such a way that only data for these points are used and none from outside. This has the great advantage that each mesh can be processed separately and those contributions belonging to the same coefficient need to be summed up. This is quite analogous to the method used in the FEM.

The partial derivatives $\phi_{|u}$ and $\phi_{|v}$ must now be approximated by central finite differences, and this is possible only in the following manner.

$$\begin{aligned} \phi_{|u}(1') &= (\phi_1 - \phi_0)/a, & \phi_{|u}(2'') &= (\phi_2 - \phi_3)/a, \\ \phi_{|u}(2') &= (\phi_1 - \phi_0 + \phi_2 - \phi_3)/2a \end{aligned} \quad (4.67)$$

and similarly

$$\begin{aligned} \phi_{|v}(3') &= (\phi_3 - \phi_0)/b, & \phi_{|v}(1'') &= (\phi_2 - \phi_1)/b, \\ \phi_{|v}(2') &= (\phi_3 - \phi_0 + \phi_2 - \phi_1)/2b. \end{aligned} \quad (4.68)$$

These all have an approximation error of *second* order. The integration over the gradient-part of the Lagrangian (4.10) must now be replaced by a weighted summation, and to avoid errors of third order, this must be done in a *symmetric* way. Hence, the basic structure of this term must be

$$\begin{aligned} \delta F_G &= \frac{ab}{2} [w_u A_{1'} \phi_{|u}^2(1') + w_u A_{2''} \phi_{|u}^2(2'') + (1 - 2w_u) A_2 \phi_{|u}^2(2')] \\ &+ \frac{ab}{2} [w_v B_{3'} \phi_{|v}^2(3') + w_v B_{1''} \phi_{|v}^2(1'') + (1 - 2w_v) B_2 \phi_{|v}^2(2')] \\ &+ ab C_{2'} \phi_{|u}(2') \phi_{|v}(2') + O(4). \end{aligned} \quad (4.69)$$

The mixed product can be calculated only at the center, if asymmetric terms are to be avoided. By Taylor-series expansion, all functions with respect to the center (2') can be shown that the error-order is at least 4, no matter how the positive weights w_u and w_v are chosen, if we introduce the approximations (4.67) and (4.68) into (4.69). It is, however, not true that the weights $w_u = w_v = 1/6$ of Simpson's rule would give the best results.

After introduction of Eqs. (4.67) and (4.68) into Eq. (4.69) and reordering thereafter we obtain a quadratic form

$$\delta F_G = -\frac{1}{2} \sum_{m=0}^3 \sum_{n=0}^3 \delta g_{mn} \phi_m \phi_n. \quad (4.70)$$

The coefficients with common label $m = 0$ are

$$\left. \begin{aligned} \delta g_{01} &= \frac{-a}{4b}(1 - 2w_v)B_{2'} + \frac{b}{4a}[4w_u A_{1'} + (1 - 2w_u)A_{2'}] \\ \delta g_{03} &= \frac{-b}{4a}(1 - 2w_u)A_{2'} + \frac{a}{4b}[4w_v B_{3'} + (1 - 2w_v)B_{2'}] \\ \delta g_{02} &= \frac{b}{4a}(1 - 2w_u)A_{2'} + \frac{a}{4b}(1 - 2w_v)B_{2'} + \frac{1}{2}C_{2'} \\ \delta g_{00} &= -\delta g_{01} - \delta g_{02} - \delta g_{03}. \end{aligned} \right\} \quad (4.71)$$

The other coefficients can be determined in an analogous manner but are not needed now.

The integration over the term $W(u, v, \phi)$ in Eq. (4.10) must similarly be approximated by a weighted sum. We find

$$\begin{aligned} \delta F_s &= \iint_R W(u, v, \phi) du dv \\ &= ab[p(W_0 + W_1 + W_2 + W_3)/4 + (1 - p)W_{2'}] \end{aligned} \quad (4.72)$$

with a free parameter p and the abbreviation

$$W_n := W(u_n, v_n, \phi_n), \quad n = 0, 1, 2, \dots \quad (4.73)$$

for the function values at the four corners. To calculate the function value at the center $2'$, we define

$$\phi_{2'} = (\phi_0 + \phi_1 + \phi_2 + \phi_3)/4. \quad (4.74)$$

Though this approximation has an error of second order, the resulting error in Eq. (4.72) is again of fourth order.

4.3.2 The Complete Mesh Formula

The above outliner procedure is to be carried out for all four rectangular domains with common node 0 and the labeling adjusted in a cyclical manner. All coefficients that refer to the same pair ϕ_0, ϕ_n are then to be summed up. The resulting form of the functional is now (with $0' = 0, 1 - p = q$)

$$F = \sum_{n=1}^8 g_n(\phi_0^2/2 - \phi_0\phi_n) + \sum_{v=0}^4 ab q_{2v} W_{2v'} + F_R, \quad (4.75a)$$

F_R being the remainder, which is independent of ϕ_0 . The requirement that this expression shall be minimized with respect to ϕ_0 leads to the condition

$\partial F/\partial \phi_0 = 0$, or:

$$S_G := \sum_{n=1}^8 g_n(\phi_n - \phi_0) = \sum_{v=0}^4 ab q_{2v} \partial W_{2v}/\partial \phi_0 =: \hat{T}_G, \quad (4.75b)$$

which is the required mesh formula. The source terms on the right-hand side can be further evaluated by means of Eq. (4.72) where we have to consider that $\partial \phi_2/\partial \phi_0 = 1/4$, $\partial \phi_4/\partial \phi_0 = 1/4$ and so on. We shall *not* write down the mesh formula in its completely explicit form as this would not give a general expression however elementary. It must be kept in mind that the procedure outlined above can be performed *independently* in each rectangular mesh. This means that in each of them the free parameters w_u , w_v , and p can be chosen in a different manner.

The mesh sizes a and b can also become different in a compatible manner, as for the asymmetric five-point formula. The latter is simply obtained by setting $w_u = w_v = 1/2$, $p = 1$, $a = u_1 - u_0$ or $u_0 - u_5$, $b = u_3 - u_0$ or $u_0 - u_7$, respectively. The result is then Eq. (4.60).

Even hybrid forms of mesh formulas with configuration numbers between 5 and 9 can be generated easily. The only obstacle is to generate the explicit form, whereas the calculation of the coefficients is an elementary task for a computer subprogram.

The assessment of the discretization error is a difficult task. This error is certainly not worse than h^3 (h being an average mesh size), so that the exact solution is always approached as $h \rightarrow 0$. In favorable symmetric configurations an order h^6 can be achieved, and this is our next task. To keep these calculations reasonably limited and concise, it is necessary to assume *constant* coefficients A , B , and C ; moreover we shall set $w_u = w_v =: w$ because there is then no reason to allow for a principal asymmetry. Finally, it is favorable to define new coefficients by writing

$$\hat{A} := Ab/a, \quad \hat{B} := Ba/b, \quad \hat{C} \equiv C. \quad (4.76)$$

After some minor calculation for a regular configuration, the left-hand side of Eq. (4.75b) can then be cast in the form

$$\begin{aligned} S_G &= \sum_{n=1}^8 g_n(\phi_n - \phi_0) = \hat{A}(\phi_1 + \phi_5 - 2\phi_0) + \hat{B}(\phi_3 + \phi_7 - 2\phi_0) \\ &\quad + \frac{1-2w}{4}(\hat{A} + \hat{B})\{\phi_2 + \phi_4 + \phi_6 + \phi_8 - 2(\phi_1 + \phi_3 + \phi_5 + \phi_7) + 4\phi_0\} \\ &\quad + \frac{1}{2}C(\phi_2 - \phi_4 + \phi_6 - \phi_8). \end{aligned} \quad (4.77)$$

Evidently, all diagonal points cancel out for $w = 1/2$, but this is possible only for $C = 0$, which means orthogonal meshes. The discretization error is found by Taylor series expansions of all potentials ϕ_n with respect to the central node 0, after which they are introduced into Eq. (4.77). All contributions of odd orders are then canceled out by symmetry. The remainder is favorably represented in terms of a function

$$T(u, v) := A\phi_{|uu} + B\phi_{|vv} + 2C\phi_{|uv}, \quad (4.78)$$

which is nothing but the right-hand side of the PDE, the partial derivatives here being functions of u and v . Our goal is to express all terms with *mixed* derivatives as derivatives of $T(u, v)$ as far as possible. The result of the elementary calculation is then, with a free parameter λ :

$$\begin{aligned} S_G = & abT(u_0, v_0) + \frac{ab}{12}[a^2T_{|uu} + b^2T_{|vv} + \lambda abT_{|uv}] \\ & + \frac{ab}{12}[a^2(2C - \lambda\hat{A})\partial_{uu} + b^2(2C - \lambda\hat{B})\partial_{vv}]\phi_{|uv} \\ & + \frac{1}{6}a^2b^2[(1 - 3w)(\hat{A} + \hat{B}) - \lambda C]\partial_{uu}\partial_{vv}\phi, \end{aligned} \quad (4.79)$$

where all derivatives to be taken at the center. The terms in the second line cannot be cancelled out simultaneously if $C \neq 0$. A reasonable choice is to cancel out the arithmetic mean of both, which results in

$$\lambda = C/\hat{A} + C/\hat{B}. \quad (4.80)$$

The expression in the last line can now be eliminated completely with the choice

$$w = \frac{1}{3} \left(1 - \frac{C^2}{\hat{A}\hat{B}} \right) \equiv \frac{1}{3} \left(1 - \frac{C^2}{AB} \right), \quad (4.81)$$

which cannot become negative. The final result is now

$$\begin{aligned} S_G = & ab \left[T_0 + \frac{1}{12}(a^2T_{|uu} + b^2T_{|vv} + \lambda abT_{|uv}) \right]_{(u_0, v_0)} \\ & + \frac{ab}{12}C\{a^2(1 - \hat{A}/\hat{B})\partial_{uu} + b^2(1 - \hat{B}/\hat{A})\partial_{vv}\}\phi_{|uv}(u_0, v_0). \end{aligned} \quad (4.82)$$

If it is still possible to choose the mesh sizes unconditionally, this should be done in such a way that the relations

$$\hat{B} = \hat{A} \quad \text{or} \quad b^2/a^2 = B/A \quad (4.83)$$

are satisfied because the remainder, not represented in terms of T , then vanishes. In this case, we shall speak of *adapted* grids.

If the PDE is homogeneous, the equation $S_G = 0$ with Eq. (4.77) is already the finite difference approximation for it as then we have $T(u, v) = 0$ and hence $T_{|uu} = T_{|vv} = T_{|uv} = 0$. If, however, the source term does not vanish identically, we have

$$T(u, v) = S(u, v, \phi(u, v)) = Q(u, v) + D\phi(u, v) \quad (4.84)$$

and then the Eq. (4.82) must be brought into agreement with the right-hand side $\hat{T}_G$ of Eq. (4.75b). For a regular rectangular grid the determination of the weights $q_0 \dots q_8$ from the accumulated sums of Eq. (4.72) gives the following:

$$q_0 = p, q_2 = q_4 = q_6 = q_8 = 1 - p. \quad (4.85)$$

We then find, using Eq. (4.74) and analogous approximations:

$$\begin{aligned} \hat{T}_G &= ab[pT_0 + (1 - p)(T_{2'} + T_{4'} + T_{6'} + T_{8'})/4] \\ &= ab[(1 + 3p)T_0/4 + (1 - p)(T_1 + T_3 + T_5 + T_7)/8 \\ &\quad + (1 - p)(T_2 + T_4 + T_6 + T_8)/16]. \end{aligned} \quad (4.86)$$

The Taylor-series expansion of this expression results in

$$\hat{T}_G = ab[T_0 + (1 - p)(a^2T_{|uu} + b^2T_{|vv})/4] + O(6), \quad (4.87)$$

which can be brought into agreement with the main term in Eq. (4.82) for $p = 2/3$.

For $\lambda = 0$, that implies $C = 0$, the mesh formula is now complete and has a discretization error of *sixth* order. However, in the case of tilted meshes, at least the term with $\lambda \neq 0$ must be compensated, and this is not possible with any symmetrical approximation like Eq. (4.72). We therefore define an additional contribution

$$\delta F_A := \sigma ab[T_{2'}(\phi_0 - \phi_1 + \phi_2 - \phi_3) - T_0\phi_0 + T_1\phi_1 - T_2\phi_2 + T_3\phi_3] \quad (4.88)$$

to be considered together with δF_s in Eq. (4.72), where σ being a free fitting parameter. After summation and differentiation with respect to ϕ_0 , there remains an additive contribution to $\hat{T}_G$ in Eq. (4.75b):

$$\begin{aligned} \hat{T}_A &= \sigma ab[T_{2'} - T_{4'} + T_{6'} - T_{8'}] = \sigma ab[T_2 - T_4 + T_6 - T_8]/4 \\ &= \sigma a^2 b^2 T_{|uv} + O(6). \end{aligned} \quad (4.89)$$

On comparison with Eq. (4.82) it becomes now obvious that the appropriate choice is $\sigma = \lambda/12$, whereupon all the free parameters are optimized.

It should be mentioned that Eq. (4.86) is not the only possible form of the source term, because the compensation of the derivatives in Eq. (4.82) can also be achieved with other ratios between the sampling values. In addition, the elimination of $T_2, T_4 \dots$ is not really necessary if $D \equiv 0$, and the function $Q(u, v)$ can be evaluated at any position; then $p = 1/3$ is the appropriate choice. Hence, there is a wide variety of equivalent discretizations.

In the case of nonconstant coefficients, it is generally not possible to eliminate all the errors of fourth order, but they remain small if the coefficient functions vary slowly. The parameters w and λ are then to be evaluated at the area midpoints $2', 4', 6'$, and $8'$.

4.3.3 Special Cases

The general formalism simplifies considerably if some reasonable assumptions for the coefficient functions can be made, which are often satisfied in practice.

(a) Orthogonality.

The assumption $C(u, v) = 0$ leads to $\lambda = 0$ and $w = 1/3$ according to Eqs. (4.80) and (4.81). The uncompensated terms in Eq. (4.82) then vanish identically, even for $\hat{B} \neq \hat{A}$. The expression $\hat{T}_A$ (Eq. (4.89)) is similarly unnecessary.

A very favorable mesh formula is obtained if the (u, v) grid was constructed by *conformal mapping* of a square-shaped grid in the (x, y) plane. Then we have $A = B = \varepsilon(u, v)$, and it is appropriate to choose equal mesh sizes $a = b =: h$. We then find the constant factors (see Fig. 4.8)

$$\begin{aligned}\mu_1 &= \mu_3 = \mu_5 = \mu_7 = 1/5, \\ \mu_2 &= \mu_4 = \mu_6 = \mu_8 = 1/20,\end{aligned}\tag{4.90}$$

and with these

$$S_G = \sum_{n=1}^8 g_n(\phi_n - \phi_0) = \frac{10}{3} \sum_{n=1}^8 \mu_n \varepsilon_{n'}(\phi_n - \phi_0).\tag{4.91}$$

The material factors $\varepsilon(u, v)$ are to be evaluated at the midpoints between the corresponding nodes and the center (see Fig. 4.8). Equation (4.86) yields a fixed ratio 2:1 of the T sums in the main and diagonal directions, respectively, but this is not an absolute necessity because any multiple of the quantity

$$\begin{aligned}\hat{T}_D &= ab[2(T_1 + T_3 + T_5 + T_7) - (T_2 + T_4 + T_6 + T_8) - 4T_0] \\ &= a^3 b^3 \partial_{uu} \partial_{vv} T\end{aligned}\tag{4.92}$$

can be added without causing an error lower than the sixth order, even in the general case. Hence, it is possible to achieve the ratio 4:1. With this option Eq. (4.91) can finally be completed to give the very concise result

$$\sum_{n=1}^8 \mu_n \left[\varepsilon_{n'}(\phi_n - \phi_0) - \frac{h^2}{12} T_n \right] = \frac{13}{60} h^2 T_0. \quad (4.93)$$

(b) Rhombic Meshes

We now adopt the coefficients A , B , and C given by Eq. (4.46), in which we shall assume $S_u = S_v$ and $\beta = \text{const.}$, however $\bar{\varepsilon}(x, y) = \varepsilon(u, v)$ may not be constant. Moreover, we again assume that $a = b = h$. This corresponds to a *rhombic* grid in the (x, y) plane. The mesh angle $\beta = \pi/3$ then leads to a *hexagonal* structure (see Fig. 4.10b). Even for $\beta \neq \pi/3$, the grid is adapted, now with $A = B \neq \varepsilon$:

$$A = B = \varepsilon(u, v) / \sin \beta, \quad C = -\varepsilon(u, v) \cot \beta. \quad (4.94a)$$

In this special case the followings constants are obtained:

$$\lambda = -2 \cos \beta, \quad w = \frac{1}{3} \sin^2 \beta. \quad (4.94b)$$

It is also possible to cast the mesh formula in a form analogous to Eq. (4.93), and again the weights will be normalized to unit sum. A minor calculation gives

$$\left. \begin{aligned} \mu_1 &= \mu_3 = \mu_5 = \mu_7 = \sin^2 \beta / (3 + 2 \sin^2 \beta) \\ \mu_4 &= \mu_8 = (1 + \cos \beta)(1 + 2 \cos \beta) / (12 + 8 \sin^2 \beta) \\ \mu_2 &= \mu_6 = (1 - \cos \beta)(1 - 2 \cos \beta) / (12 + 8 \sin^2 \beta). \end{aligned} \right\} \quad (4.95)$$

As before, the term $\hat{T}_D$ from Eq. (4.92) can be used to achieve proportionality in the source factors. It is convenient to transform back to the functions in the (x, y) plane by means of Eqs. (4.3) and (4.12), from which it can be concluded that $T(u, v) = J \bar{T}(x, y)$, the Jacobian here being $J = (s/h)^2 \sin \beta$ (see Fig. 4.10a). After some elementary calculation, we then arrive at

$$\sum_{n=1}^8 \mu_n \left[\varepsilon_{n'}(\Phi_n - \Phi_0) - \frac{s^2}{12} \sin^2 \beta \bar{T}_n \right] = \bar{\mu}_0 \frac{s^2}{12} \sin^2 \beta \bar{T}_0, \quad (4.96)$$

$$\bar{\mu}_0 := \frac{15 - 2 \sin^2 \beta}{3 + 2 \sin^2 \beta}.$$

For $\beta = \pi/2$ this specializes to the back transformed expression of Eq. (4.93) with Eq. (4.90) as it must.

For the angles $\beta = \pi/3$ or $2\pi/3$ corresponding to a regular *hexagonal* grid, we obtain *equal* weights of value $1/6$ for the six closed neighbors, and *vanishing* weights for the two more distant ones, as must be the case for reasons of symmetry. For the sources, this result emerges unconditionally without using Eq. (4.92). If we renumber the closest neighbors from 1 to 6 (see Fig. 4.10b), we obtain the well-known equation

$$\frac{1}{6} \sum_{n=1}^6 \left[\varepsilon_{n'} (\Phi_n - \Phi_0) - \frac{s^2}{16} \bar{T}_n \right] = \frac{3s^2}{16} \bar{T}_0. \quad (4.97)$$

This demonstrates that all the basic requirements are satisfied.

4.3.4 The Regularization of Meshes

In Section 4.1, we have discussed the generation of two-dimensional meshes by means of analytic coordinate transforms. In this way, it was possible to adapt the mesh at least at some parts of the boundary, which are considered as very important, but on other parts irregular configurations may still appear. It is of course possible to accept these and to use asymmetric formulae like Eqs. (4.50), (4.51), or (4.60).

Another possibility is to remove the irregularities by *deformation* of the mesh. The first step consists in selecting an acceptable location of the boundary points so that there are no discontinuities in their mutual distance (see Fig. 4.11). These then remain unaltered, but the distances to the inner

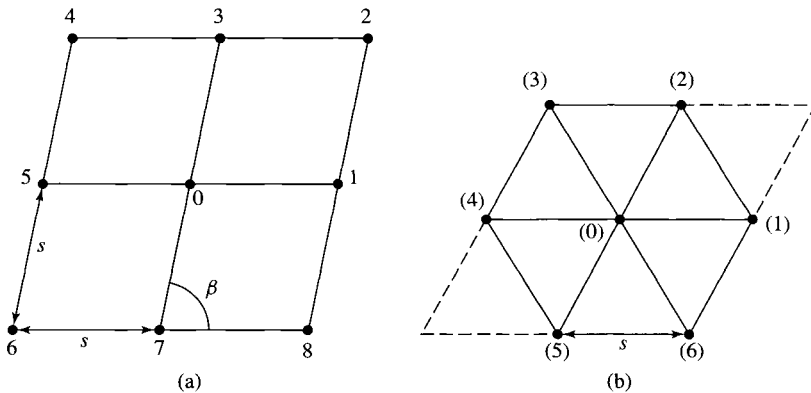


FIGURE 4.10 Rhombic mesh (a) and hexagonal mesh (b) in the (x, y) plane with readjustment of the numbering.

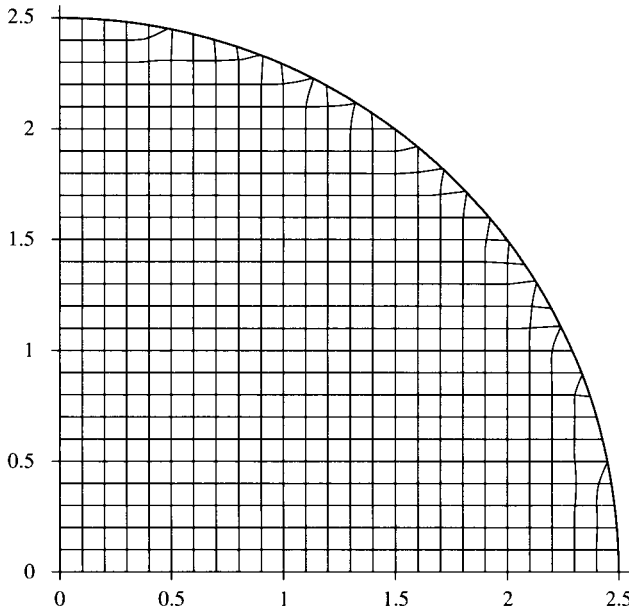


FIGURE 4.11 Regularization of a square-shaped mesh in a quadrant. The result demonstrates that the deformations remain confined to the vicinity of the boundary, but in some points they are strong for topological reasons. Note that nowhere are distances less than $h/2 = 0.05$ obtained.

neighbors are unacceptable. To remove this deficiency, the whole mesh is now slightly deformed until all distances between nearest neighbors have become nearly equal. This process is called *regularization*.

In the lowest order of approximation this can be achieved by solution of the whole set of five-point equations

$$\begin{aligned} x_{ik} &= (x_{i+1,k} + x_{i,k+1} + x_{i-1,k} + x_{i,k-1})/4, \\ y_{ik} &= (y_{i+1,k} + y_{i,k+1} + y_{i-1,k} + y_{i,k-1})/4 \end{aligned} \quad (4.98)$$

for every internal points. In a better approximation, the nine-point formulas

$$\begin{aligned} x_{ik} &= (x_{i+1,k} + x_{i,k+1} + x_{i-1,k} + x_{i,k-1})/5 \\ &\quad + (x_{i+1,k+1} + x_{i-1,k+1} + x_{i-1,k-1} + x_{i+1,k-1})/20 \end{aligned} \quad (4.99)$$

and a corresponding formula for y_{ik} can be used, which have an approximation error of eight order. The resulting mesh is almost conformal but not exactly. The solution of PDEs in such a grid can be achieved using the general nine-point formulas outlined in this chapter. In this context, it is important that the partial derivatives $x_{|u}$, $x_{|v}$ and so forth needed for this purpose can easily be

obtained by finite differences expressions:

$$\begin{aligned} x_{|u} &= (x_{i+1,k-1} + 4x_{i+1,k} + x_{i+1,k+1})/12h \\ &\quad - (x_{i-1,k-1} + 4x_{i-1,k} + x_{i-1,k+1})/12h + O(h^4) \end{aligned} \quad (4.100)$$

and three other analogous formulas for the derivatives at the position (i, k) . These are so simple that it is not necessary to calculate them in advance and store the results for later uses.

4.4 THE CYLINDRICAL POISSON EQUATION

In Chapter II, we frequently encountered a PDE of the form (2.58), here rewritten as

$$\Delta_\alpha V(z, r, \alpha) \equiv V_{|zz} + V_{|rr} + \alpha r^{-1} V_{|r} = -g(z, r, \alpha), \quad (4.101)$$

or equivalently,

$$\partial_z(r^\alpha V_{|z}) + \partial_r(r^\alpha V_{|r}) = -r^\alpha g(z, r, \alpha), \quad (4.102)$$

which is certainly of the self-adjoint form considered for $r > 0$, regardless of the value of the constant parameter α . Therefore, for $r > 0$, the formulas derived for five-or nine-point configurations can be applied to this PDE. There is, however, one important new aspect: for $r \rightarrow 0$, this PDE becomes *singular*, which means that the coefficients $A = B = r^\alpha$ *vanish* if $\alpha > 0$, which causes problems.

Unfortunately, the vicinity of the z -axis, where these difficulties arise, in the most important domain in the field because this is also the domain occupied by particle rays. Therefore, some further discussion is necessary to overcome these problems, but before we begin this, we shall briefly state the results of the preceding theory for a square-shaped mesh in the (z, r) -plane. We thus identify $x \equiv u \hat{=} z$, $y \equiv v \hat{=} r$.

The special formula, resulting from Eqs. (4.50) and (4.51) with $A = B = 1$, $D = 0$, $a = 0$, and $b = \alpha/r$, $Q = -g$ in double index form is

$$\begin{aligned} V_{i,k} &= \frac{1}{4}(V_{i+1,k} + V_{i,k+1} + V_{i-1,k} + V_{i,k-1}) \\ &\quad + \frac{\alpha}{8k}(V_{i,k+1} - V_{i,k-1}) + \frac{h^2}{4}g_{ik}, \quad (k \geq 1), \end{aligned} \quad (4.103)$$

as the central node has the radial coordinate $r = k \cdot h$. The axial formula resulting from Eqs. (4.52), (4.53), and (4.54) is then

$$V_{i,0} = \frac{1}{2(\alpha + 2)} [V_{i+1,0} + V_{i-1,0} + h^2 g_{i,0} + 2(\alpha + 1)V_{i,1}], \quad (\alpha \geq 0). \quad (4.104)$$

The evaluation of Eq. (4.58) requires more attention. After canceling out a factor $r^\alpha = (kh)^\alpha$ there remains for $k \geq 1$ and $\alpha > 1$:

$$\begin{aligned} (V_{i+1,k} + V_{i-1,k} - 2V_{i,k} + h^2 g_{i,k}) & \left[\left(1 + \frac{1}{2k}\right)^{\alpha+1} - \left(1 - \frac{1}{2k}\right)^{\alpha+1} \right] \frac{k}{\alpha + 1} \\ & + (V_{i,k+1} - V_{i,k}) \left(1 + \frac{1}{2k}\right)^\alpha + (V_{i,k-1} - V_{i,k}) \left(1 - \frac{1}{2k}\right)^\alpha = 0, \end{aligned} \quad (4.105)$$

whereas Eq. (4.60), then gives

$$\begin{aligned} V_{i+1,k} + V_{i-1,k} - 2V_{i,k} + h^2 g_{i,k} + (V_{i,k+1} - V_{i,k}) \left(1 + \frac{1}{2k}\right)^\alpha \\ + (V_{i,k-1} - V_{i,k}) \left(1 - \frac{1}{2k}\right)^\alpha = 0. \end{aligned} \quad (4.106)$$

Evidently, the exact integration over r^α that result in the complicated factor in the first line of Eq. (4.105) is *unnecessary*; a Taylor series expansion of this expression results in $1 + O(k^{-2})$. If we continue with series expansions of the remaining factors, we arrive at Eq. (4.103). Numerical checks (see refs. [8, 19]) have shown that all three approximations are practically equivalent with respect to the discretization error. The Eq. (4.103) has the disadvantage that, for $\alpha > 2$, some coefficients become negative, which is numerically unfavorable, while Eq. (4.105) is not applicable for $\alpha = -1$. Hence Eq. (4.106) in combination with Eq. (4.104) is the most favorable discretization.

It must be emphasized again that this numerically stable form was obtained by cancelling out a common factor v^α . This destroys the positively symmetric form of the Eq. (4.65), with the result that the conjugate gradient techniques for their solution *cannot* be applied. On the other hand, if we had not done this, the dramatically increasing factor k^α for $\alpha > 3$ would have given rise to an ill-conditioned matrix. The functional is then accurately minimized in the far off-axis domain but not in the paraxial domain, which is frequently the most important. The same difficulty arises with nine-point formulas and also in the FEM, since all are based on the same unfavorable functional

$$F = \iint r^\alpha \left[\frac{1}{2} (V_{|z}^2 + V_{|r}^2) - gV \right] dr dz = \min. \quad (4.107)$$

We now seek for an alternative method by which a more accurate mesh formula can be derived.

4.4.1 The Radial Discretization

Temporarily, the consideration of the PDE Eq. (4.101) is postponed. Instead we now deal with the ordinary differential equation (ODE)

$$y''(r) + \frac{\alpha}{r} y'(r) = q(r), \quad (4.108)$$

which has the same kind of singularity as Eq. (4.101). The general solution of Eq. (4.108) is the linear superposition of a particular solution and the general solution of the associated homogeneous ODE. With later applications in mind, we consider here only solutions that remain finite and regular at the position $r = 0$. These can contain only even powers of r , hence

$$y(r) = y_0 + \frac{r^2}{2} y_0'' + \frac{r^4}{24} y_0^{(4)} + \frac{r^6}{720} y_0^{(6)} + O(r^8). \quad (4.109)$$

From Eq. (4.108) we find

$$q(r) = (1 + \alpha) y_0'' + \frac{r^2}{6} (3 + \alpha) y_0^{(4)} + \frac{r^4}{120} (5 + \alpha) y_0^{(6)} + O(r^6), \quad (4.110)$$

which means that — apart from the trivial case $y = y_0 = \text{const}$ — there are no *regular homogeneous* solutions.

For reasons of conciseness, we consider here only equidistant intervals, by which we mean positions $r_n = nh$ with integral value of n , and write $y_n := y(nh)$. Because we are searching for a modified nine-point formula in the (z, r) plane, we consider a *three*-point formula and start with a form

$$R_k := \eta_k^+ (y_{k+1} - y_k) + \eta_k^- (y_{k-1} - y_k), \quad (4.111)$$

in which

$$\eta_k^\pm = 1 + \gamma_k \pm (\alpha - \gamma_k)/2k \quad (4.112)$$

for all $k > 0$. The parameter γ_k is still undetermined and will be used to optimize the discretization of the ODE Eq. (4.108). To find this optimum, we evaluate the Taylor series expansion of the expression R_k with respect to r at the central position $r_k = kh$. For conciseness, the label k will be dropped in this calculation, after $k = r_k/h$ has been substituted. All derivatives present will refer to this value of the radial coordinate.

We then find

$$R = h^2 [(1 + \gamma)y'' + (\alpha - \gamma)y'/r] + h^4 [(1 + \gamma)y^{(4)} + 2(\alpha - \gamma)y'''/r] / 12 + O(h^6). \quad (4.113)$$

From the terms with factor γ , it is clearly appropriate to introduce a “perturbation”

$$S := r^{-2}(y'' - y'/r) = -\frac{1}{3}(y^{(4)} - 2y'''/r) + O(h^6). \quad (4.114)$$

The second part of this equation is a consequence of Eq. (4.109), as can be verified easily. The remaining terms can be represented with a differential operator D_α and its iteration, thus,

$$D_\alpha y := \left(\frac{d^2}{dr^2} + \frac{\alpha}{r} \frac{d}{dr} \right) y = y'' + \frac{\alpha}{r} y' \equiv q(r),$$

$$D_\alpha^2 y = \left(\frac{d^2}{dr^2} + \frac{\alpha}{r} \frac{d}{dr} \right)^2 y = y^{(4)} + \frac{2\alpha}{r} y''' + \alpha(\alpha - 2)S. \quad (4.115)$$

We can now rewrite Eq. (4.113) concisely as

$$R = h^2 D_\alpha y + \frac{h^4}{12} D_\alpha^2 y + h^2 S (\gamma r^2 - h^2 \gamma / 4 - h^2 \alpha(\alpha - 2) / 12), \quad (4.116)$$

from which it is obvious that the whole term involving the perturbation S cancels out for

$$\gamma = \frac{h^2 \alpha(\alpha - 2)}{12r^2 - 3h^2} = \frac{\alpha(\alpha - 2)}{12k^2 - 3} \equiv \gamma_k, \quad (4.117)$$

whereupon we obtain

$$R = h^2 q + \frac{h^4}{12} D_\alpha q + O(h^6). \quad (4.118)$$

The second term can be eliminated by introduction of a finite difference expression for the function $q(r)$ like R_k in Eq. (4.111).

Putting all this together, we finally arrive at

$$\begin{aligned} & \eta_k^+ \left(y_{k+1} - y_k - \frac{h^2}{12} (q_{k+1} - q_k) \right) + \eta_k^- \left(y_{k-1} - y_k - \frac{h^2}{12} (q_{k-1} - q_k) \right) \\ & = h^2 q_k + O(h^6). \end{aligned} \quad (4.119)$$

The coefficients $\eta_k^\pm$ from Eq. (4.112) with Eq. (4.117) can be simplified further. After a short elementary calculation we find

$$\eta_k^\pm = 1 \pm \frac{\alpha}{2k} + \frac{\alpha(\alpha - 2)}{12k(k \pm 1/2)}. \quad (4.120)$$

In terms of a dimensionless variable

$$t := (r - r_k)/r_k, \quad (4.121a)$$

and a general function

$$\eta(t) = 1 + \alpha t + \frac{\alpha(\alpha - 2)t^2}{3(1 + t)}, \quad (4.121b)$$

we obtain the simple result $\eta_k^\pm = \eta(\pm 0.5/k)$.

The discretization is to be completed by an “axial” formula for $k = 0$. Its derivation from Eqs. (4.109) and (4.110) results in

$$\begin{aligned} y_1 - y_0 &= \frac{h^2}{4(1 + \alpha)(3 + \alpha)} \\ &\times ((5 + \alpha)q_0 + (1 + \alpha)q_1) + O(h^6), \quad (\alpha \geq 0). \end{aligned} \quad (4.122)$$

With respect to the discretization in the z -direction, it is necessary to find a three-point formula for the solutions of the ODE.

$$x''(z) = s(z). \quad (4.123)$$

Apart from the necessary change of notation, this is simply the special case $\alpha = 0$ of the formulas derived earlier. The result can be cast in the concise form

$$x_{i+1} - 2x_i + x_{i-1} = h^2(s_{i+1} + 10s_i + s_{i-1})/12 + O(h^6), \quad (4.124)$$

which is well known in numerical analysis. All the formulas for the discretization of Eq. (4.101) are now available.

4.4.2 Discretization of Separable Differential Equations

The cylindrical Poisson equation is a special case of a PDE for a potential $\phi(u, v)$ of the form

$$\underbrace{a_2(u)\phi_{|uu} + a_1(u)\phi_{|u}}_{A(u,v)} + \underbrace{b_2(v)\phi_{|vv} + b_1(v)\phi_{|v}}_{B(u,v)} = Q(u, v), \quad (4.125)$$

where A and B are introduced for convenience. In such a case it is fairly easy to find the appropriate nine-point formula, provided that the one-dimensional discretizations are known in both coordinate directions, as we have determined in the earlier section, even in nonuniform meshes. For reasons of conciseness we consider three monotonically increasing sampling coordinates $u_1 < u_2 < u_3$ and $v_1 < v_2 < v_3$ and use again the notation $\phi_{ik} = \phi(u_i, v_k)$.

We now assume that the ODE

$$a_2(u)U''(u) + a_1(u)U'(u) = f(u) \quad (4.126a)$$

may be discretized in the form

$$\sum_{i=1}^3 (X_i U_i - S_i f_i) = 0. \quad (4.126b)$$

Similarly, the ODE

$$b_2(v)V''(v) + b_1(v)V'(v) = g(v) \quad (4.127a)$$

may be discretized thus:

$$\sum_{k=1}^3 (Y_k V_k - T_k g_k) = 0. \quad (4.127b)$$

In these, we need to find all the *coefficients* X_i, S_i and Y_k, T_k , 12 numbers altogether.

Now Eq. (4.125) is brought into the form of Eq. (4.126a) by setting $U(u) = \phi(u, v)$, the inhomogeneity then being $f(u) = Q(u, v) - B(u, v)$. Evaluation at the positions v_k , ($k = 1, 2, 3$) by means of Eq. (4.126b) results in

$$\sum_{i=1}^3 X_i \phi_{ik} = \sum_{i=1}^3 S_i (Q_{ik} - B_{ik}), \quad (k = 1, 2, 3). \quad (4.128a)$$

In quite an analogous manner, Eqs. (4.127a,b) are used to find

$$\sum_{k=1}^3 Y_k \phi_{ik} = \sum_{k=1}^3 T_k (Q_{ik} - A_{ik}), \quad (i = 1, 2, 3). \quad (4.128b)$$

To eliminate the matrix elements A_{ik} and B_{ik} , which contain derivatives, Eq. (4.128a) is multiplied by T_k and Eq. (4.128b) by S_i , and then each is

summed over the free index. We then obtain

$$\sum_{i=1}^3 \sum_{k=1}^3 (X_i T_k + Y_k S_i) \phi_{ik} = \sum_{i=1}^3 \sum_{k=1}^3 S_i T_k (2Q_{ik} - A_{ik} - B_{ik}). \quad (4.129)$$

However, from Eq. (4.125), we have $A_{ik} + B_{ik} = Q_{ik}$ and hence finally

$$\sum_{i=1}^3 \sum_{k=1}^3 [(X_i T_k + Y_k S_i) \phi_{ik} - S_i T_k Q_{ik}] = 0. \quad (4.130)$$

This kind of discretization is very favorable because it can be easily used for distorted meshes as defined by Eq. (4.32). The whole effort consists in the calculation of the coefficients for the one-dimensional discretizations. In the general case, an approximation with error of *fifth* order can be found, whereas for equidistant meshes the formula will be of *sixth* order, as was demonstrated in the earlier section. The product coefficients in Eq. (4.130) are so easily calculated that it is unnecessary to store them, hence only four *one-dimensional* arrays are sufficient for these coefficients.

With respect to the cylindrical Poisson equation, we now identify

$$\begin{aligned} u = z, \quad v = r, \quad \phi(u, v) = V(z, r), \quad Q(u, v) = -g(z, r) \\ a_2(z) = 1, \quad a_1(z) = 0, \quad b_2(r) = 1, \quad b_1(r) = \alpha/r. \end{aligned} \quad (4.131)$$

The labels are shifted to the ranges $i - 1, i, i + 1$ and $k - 1, k, k + 1$, respectively. Then we find, by comparing Eq. (4.126b) with Eq. (4.124),

$$\begin{aligned} X_{i-1} = X_{i+1} = 1, \quad X_i = -2 \\ S_{i-1} = S_{i+1} = h^2/12, \quad S_i = 5h^2/6. \end{aligned} \quad (4.132)$$

Similarly, comparison of Eq. (4.127b) with Eq. (4.119) yields the coefficients

$$\begin{aligned} Y_{k\pm 1} = \eta_k^\pm, \quad Y_k = -\eta_k^+ - \eta_k^- \equiv -2(1 + \gamma_k) \\ T_{k\pm 1} = h^2 \eta_k^\pm / 12, \quad T_k = h^2 (5 - \gamma_k) / 6. \end{aligned} \quad (4.133)$$

If we introduce all this into Eq. (4.130) and cancel out a common factor $h^2/6$, we find the simple coefficient matrix $P_{ik} \sim X_i T_k + S_i Y_k$ with (see Fig. 4.8)

$$P^{(k)} = \begin{pmatrix} \eta^+ & 4\eta^+ & \eta^+ \\ 4 - 2\gamma & -20 - 8\gamma & 4 - 2\gamma \\ \eta^- & 4\eta^- & \eta^- \end{pmatrix} \quad (4.134)$$

for the potentials, in which we have dropped the omnipresent label k for reasons of conciseness. The matrix for the sources, $G_{ik} \sim S_i T_k$, is found to be

$$\mathbf{G}_I^{(k)} = \frac{h^2}{24} \begin{pmatrix} \eta^+ & 10\eta^+ & \eta^+ \\ 10 - 2\gamma & 100 - 20\gamma & 10 - 2\gamma \\ \eta^- & 10\eta^- & \eta^- \end{pmatrix}. \quad (4.135)$$

However, this can be modified easily by forming linear combinations with a second matrix

$$\mathbf{G}_{II}^{(k)} = \frac{h^2}{24} \begin{pmatrix} \eta^+ & -2\eta^+ & \eta^+ \\ -2(1 + \gamma) & 4(1 + \gamma) & -2(1 + \gamma) \\ \eta^- & -2\eta^- & \eta^- \end{pmatrix}. \quad (4.136)$$

Any linear combination $\mathbf{G}_I^{(k)} + \lambda \mathbf{G}_{II}^{(k)}$ can be used, thereby causing an additional error of about $\lambda h^6 D_\alpha g_{|z|} / 24$ that can be ignored. A favorable choice is $\lambda = -1$, giving

$$\bar{\mathbf{G}}^{(k)} = \frac{h^2}{2} \begin{pmatrix} 0 & \eta^+ & 0 \\ 1 & 8 - 2\gamma & 1 \\ 0 & \eta^- & 0 \end{pmatrix}, \quad (4.137)$$

and another one is $\lambda = +1$, which leads to

$$\mathbf{G}^{(k)} = \frac{h^2}{12} \mathbf{P}^{(k)} + 6h^2 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (4.138)$$

In the latter case, it is obviously advantageous to introduce a modified potential

$$W(z, r) := V(z, r) + h^2 g(z, r) / 12, \quad (4.139)$$

whereupon Eq. (4.130) can be written very concisely as

$$\sum_{m=-1}^1 \sum_{n=-1}^1 P_{m+2, n+2} W_{i+m, k+n} = -6h^2 g_{ik} + O(h^6), \quad (4.140)$$

where the matrix elements depend on the temporarily dropped label k , as mentioned earlier. This discretization is to be completed by a suitable axial formula, which should not be less accurate. The derivation is straightforward, and the combination of the coefficients in Eqs. (4.122) and (4.124) now leads to a 3×2 matrix

$$\mathbf{P}^{(0)} = N^{-1} \begin{pmatrix} 6 + 7\alpha + \alpha^2, & 24 + 34\alpha + 10\alpha^2, & 6 + 7\alpha + \alpha^2 \\ 12 - \alpha - \alpha^2, & -60 - 46\alpha - 10\alpha^2, & 12 - \alpha - \alpha^2 \end{pmatrix}, \quad (4.141a)$$

with the frequently appearing denominator

$$N = (1 + \alpha)(3 + \alpha). \quad (4.141b)$$

Here a factor $h^2/12$ has been dropped. The corresponding $\mathbf{G}$ -matrices become in turn

$$\begin{aligned} \mathbf{G}_I^{(0)} &= \frac{h^2}{4N} \begin{pmatrix} 1 + \alpha & 10(1 + \alpha) & 1 + \alpha \\ 5 + \alpha & 10(5 + \alpha) & 5 + \alpha \end{pmatrix} \\ \mathbf{G}_{II}^{(0)} &= \frac{h^2}{4N} (1 + \alpha) \begin{pmatrix} 1 & -2 & 1 \\ -1 & 2 & -1 \end{pmatrix} \\ \overline{\mathbf{G}}^{(0)} &= \mathbf{G}_I^{(0)} - \mathbf{G}_{II}^{(0)} = \frac{h^2}{2N} \begin{pmatrix} 0 & 6(1 + \alpha) & 0 \\ 3 + \alpha & 24 + 4\alpha & 3 + \alpha \end{pmatrix}, \end{aligned} \quad (4.142)$$

and

$$\mathbf{G}^{(0)} = \mathbf{G}_I^{(0)} + (1 + \alpha/3)\mathbf{G}_{II}^{(0)}, \quad (4.143)$$

giving

$$\mathbf{G}^{(0)} = \frac{h^2}{12} \mathbf{P}^{(0)} + h^2 \begin{pmatrix} 0 & c_1 & 0 \\ 0 & c_0 & 0 \end{pmatrix} \quad (4.144a)$$

with

$$c_0 = (18 + 7\alpha + \alpha^2)/N, \quad c_1 = -\alpha(1 + \alpha)/N. \quad (4.144b)$$

The mesh formula takes the form

$$\sum_{m=-1}^1 \sum_{n=0}^1 P_{m+2,n} W_{i+m,n} = -h^2(c_0 g_{i0} + c_1 g_{i1}) + O(h^6). \quad (4.145)$$

The discretization is now complete, and our next task is to study its properties.

4.4.3 Accuracy of the Discretization

The cylindrical Poisson equation appears so frequently in particle-optical field calculations that the properties of its discretization will be presented here in some essential details but now without derivation.

On comparing the simple formulas Eqs. (4.103), (4.104), and (4.106) with the much more complicated Eq. (4.134) to Eq. (4.145) the first question that arises is whether the enhanced computational effort will pay off. The author has given an answer in the form of tables [8], which show that the more

complicated formulas are indeed worthwhile but, instead of reproducing these, we give some simple error formulas.

The *local* discretization error is found by introducing an exact analytical solution of the PDE into the potentials of the neighbor points of the mesh formula, solving for the central value and comparing this with the corresponding analytical value. Because this is most important in the vicinity of the optic axis, it is sufficient to use the polynomials arising from the paraxial series expansion especially those from Eq. (2.66) combined with Eq. (2.72). For reasons of conciseness, we shall consider here only the homogeneous PDE, for which $g(z, r) \equiv 0$.

In this case, the error of Eq. (4.103) becomes

$$\delta V = \frac{\alpha^2 + 10\alpha + 6}{48(\alpha + 1)(\alpha + 3)} h^4 \phi^{(4)}(z), \quad (4.146)$$

$\phi(z)$ being the axial potential. The error of Eq. (4.106) is a more complicated expression but has the same order of magnitude. The axial formula is a little more accurate but is finally dominated by the off-axis errors. As a thumb rule the formula $\delta V = h^4 \phi^{(4)}/20$ is adequate. In the nine-point approximation the axial formula has the largest error; we therefore confine the discussion to this case. For ordinary rotationally symmetric Laplace fields ($\alpha = 1$) this formula becomes

$$V_{i,0} = \frac{5}{58}(V_{i+1,0} + V_{i-1,0}) + \frac{7}{58}(V_{i+1,1} + V_{i-1,1}) + \frac{34}{58}V_{i,1} - h^6 \phi^{(6)}/2420.87, \quad (\alpha = 1). \quad (4.147)$$

This is usually much more accurate than the corresponding five-point formula: if we assume that the derivatives remain confined according to $|\phi^{(n)}| \leq SR^{-n}$. S being a fixed number and R a radial constant, we obtain a ratio of about $(h/10R)^2$ between the local discretization errors, which gives 10^{-4} for $h = 0.1 R$.

The next interesting case consists of dipole or deflection fields ($\alpha = 3$). Then the axial formula becomes astonishingly simple:

$$V_{i,0} = \frac{1}{8}(V_{i+1,1} + V_{i-1,1}) + \frac{3}{4}V_{i,1} - h^6 \phi^{(6)}/4051, \quad (\alpha = 3). \quad (4.148)$$

This formula has the remarkable property that the neighbor points on the optic axis; $(i \pm 1, 0)$ do not appear in it. This is quite favorable with respect to the field calculation, as we shall see in the next section.

For $\alpha > 3$ the difficulty arises that, in the axial formula and also in the neighboring rows ($k > 0$), some coefficients become negative although the

accuracy remains of the same order of magnitude. This has the consequence that the simple overrelaxation technique for the solutions of the system of equations cannot be applied, but this disadvantage can be overcome. For $\alpha = 0$ we obtain $\eta^+ = \eta^- = 1$, $\gamma = 0$, and hence the nine-point formula simplifies to Eq. (4.99), the notation to be adjusted accordingly. This formula has an error of eighth order. The corresponding axial formula

$$V_{i,0} = \frac{1}{5}(V_{i+1,0} + V_{i-1,0}) + \frac{1}{10}(V_{i+1,1} + V_{i-1,1}) + \frac{2}{5}V_{i,1} + \frac{h^8 \phi^{(8)}}{10080}, \quad (\alpha = 0) \quad (4.149)$$

is simply a symmetrized version of the general one and has hence the same accuracy.

It is of some importance that the difference formula for the *flux potential* Ψ with $\alpha = -1$ also has a discretization error of order h^8 . This can be shown by introducing the polynomial

$$\Psi = r^2(8z^4 - 12z^2r^2 + r^4) \quad (\alpha = -1) \quad (4.150)$$

into the mesh formula, which is then exactly satisfied. The consequences of this property will be studied in the next section.

Finally, it is of some interest to apply the formula (4.93) to the solutions of the cylindric Poisson equation. In this case we have to set $\varepsilon(z, r) = r^\alpha$ with $r = k \cdot h$ or $(k \pm 1/2)h$. After canceling out a factor $(kh)^\alpha$, the formula for the homogeneous case ($T \equiv 0$) can be cast in the form

$$V_0 = N^{-1} \left[4(V_1 + V_5) + \left(1 + \frac{1}{2k}\right)^\alpha (V_2 + 4V_3 + V_4) + \left(1 - \frac{1}{2k}\right)^\alpha (V_6 + 4V_7 + V_8) \right] \\ N = 8 + 6 \left(1 + \frac{1}{2k}\right)^\alpha + 6 \left(1 - \frac{1}{2k}\right)^\alpha. \quad (4.151)$$

The Taylor series expansion of the coefficients gives

$$\left(1 \pm \frac{1}{2k}\right)^\alpha = 1 \pm \frac{\alpha}{2k} + \frac{\alpha(\alpha-1)}{8k^2} + O(k^{-3}), \quad (4.152)$$

and these converge for $k \rightarrow \infty$ to the coefficients $\eta_k^\pm$ of Eq. (4.120). The consequence is an error of asymptotically h^4/r^2 . Unfortunately, this becomes

small where it is unimportant. Near the optic axis it remains confined. A more rigorous calculation shows that this error is about

$$\delta V = \frac{\alpha(2-\alpha)}{40(1+\alpha)(3+\alpha)} h^4 \phi^{(4)}(z), \quad (\alpha \neq -1). \quad (4.153)$$

For $\alpha = 1$ this is smaller by a factor of about 14 than the error of the five-point formula. However, with the rigorous formula several orders of magnitude can be gained. All this shows that undoubtedly larger computational effort does indeed pay off.

4.4.4 The Radial Power Transform

The ODE (4.108) from which we set out has a useful property: if we introduce the function $y(r) = r^\lambda x(r)$ and $g(r) = r^\lambda s(r)$ into it, carrying out the product differentiations and canceling out a common factor r^λ , we find

$$x''(r) + (2\lambda + \alpha)r^{-1}x' + \lambda(\lambda + \alpha - 1)r^{-2}x = s(r). \quad (4.154)$$

If we choose the value $\lambda = 1 - \alpha$, the term in r^{-2} cancels out, and the ODE becomes *form invariant* in the sense that here

$$x''(r) + \bar{\alpha}r^{-1}x'(r) = s(r), \quad \bar{\alpha} = 2 - \alpha \quad (4.155)$$

holds. This property can now be transferred to Eqs. (4.101) and (4.102), which means that with

$$\bar{V}(z, r) = r^{1-\alpha}V(z, r), \quad \bar{g}(z, r) = r^{1-\alpha}g(z, r), \quad (4.156)$$

a new solution is found for the parameter $\bar{\alpha} = 2 - \alpha$.

The pair $(\alpha = 0, \bar{\alpha} = 2)$ is of little interest; the pair $(\alpha = 1, \bar{\alpha} = 1)$ represents the neutral element of this transform; but the pair $(\alpha = 3, \bar{\alpha} = -1)$ or similarly $(\alpha = -1, \bar{\alpha} = 3)$ is of some importance: this represents the equivalence of the PDEs for *dipole* fields and for *flux* fields.

In fact, this is the equivalence of Eqs. (2.47) and (2.48) with (2.41) for round magnetic lenses.

Fortunately, the discretization is in exact agreement with this transformation law, as we can rewrite Eqs. (4.121) in the form

$$\begin{aligned} \eta(t) &= \frac{1}{1+t}, & \eta_k^\pm &= \frac{r_k}{r_k \pm h/2} \quad (\alpha = -1) \\ \eta(t) &= \frac{(1+2t)^2}{1+t}, & \eta_k^\pm &= \frac{r_k}{r_k \pm h/2} \left(\frac{r_k \pm h}{r_k} \right)^2 \quad (\alpha = 3). \end{aligned} \quad (4.157)$$

In practice this means that the use of the potential $\Pi(z, r)$ in Eq. (2.41) has no particular benefit, although in Eqs. (2.42) the division by r could be avoided by its use. The *flux potential* has the advantages that no axial formula is necessary and that the contributions to the functional (4.107) are proportional to r^{-1} instead of r^3 in the far off-axis zone.

With *square*-shaped meshes, the discretization is of order h^8 , and so there is then practically *no* loss of accuracy if we use Eq. (4.148) for purposes of differentiation: on identifying $V(z, r)$ with $\Pi(z, r)$ for $\alpha = 3$ in Eq. (4.148) and considering Eqs. (2.41) and (2.42a) we obtain immediately

$$B_{i0} = \frac{1}{8\pi h^2} (\Psi_{i+1,1} + 6\Psi_{i,1} + \Psi_{i-1,1}) + O(h^6) \quad (4.158)$$

for the axial magnetic field strength.

4.4.5 Correction of the Functional

Many methods of field calculation rely on the minimization of a functional as we have outlined in Section 4.3, but the very important functional from Eq. (4.107) leads to the Eq. (4.151) that is only asymptotically accurate. A fairly simple method of overcoming this difficulty is to modify the integration factors in Eq. (4.69) in such a manner that the correct coefficient matrix is obtained when Eq. (4.75) are evaluated. This is a longer but elementary calculation. Here we shall present only the results for the most important cases, $\alpha = \pm 1$, corresponding to rotationally symmetric fields. In these cases we have to make the identifications $u = z$, $v = r$, $A(z, r) = B(z, r) = \varepsilon(r)$, $C(z, r) \equiv 0$, and $w_u = v_v = 1/3$.

We can assume here that the meshes are rectangular of width a in the axial and b in the radial direction. The midpoint of the rectangle considered for the derivation of Eqs. (4.69ff) then has the radial coordinate

$$\rho := r_{2'} = (k + 1/2)b. \quad (4.159)$$

The simplest way of obtaining the correct result is to modify the coefficient $\varepsilon(r)$ according to

$$\begin{aligned} \alpha = 1: \quad \varepsilon(r) &= r + \frac{1}{2\rho}(r - \rho)^2 - \frac{b^2}{12\rho}, \quad |r - \rho| \leq b/2 \\ \alpha = -1: \quad \varepsilon(r) &= \frac{1}{r} - \frac{(r - \rho)^2}{2r^2\rho} \end{aligned} \quad (4.160)$$

for $r > 0$. These functions are discontinuous at the mesh lines. The summation procedure, outlined in Section 4.3.2 leads to the arithmetic mean of the two

values referring to the same position, and this is just $r(1 - \gamma/2)$ or $(1 - \gamma/2)/r$, respectively.

Thus it is easy to verify that the coefficients obtained are indeed proportional to those of Eq. (4.91) if $a = b = h$. Even for $a \neq b$, formulas with an error of sixth order are obtained.

For $\alpha = -1$, an axial formula is unnecessary; it suffices to assume a nonzero value of r , which does not contribute to the functional. For $\alpha = 1$ Eq. (4.160) is still valid at $r = 0$. The axial mesh formula obtained is less accurate than Eq. (4.147) but is still better than Eq. (4.104).

4.4.6 The Implicit Algorithm

Although the mesh formulas, derived so far, serve mainly for the numerical solution of Eq. (4.101) in cylindric coordinates (z, r) , this is not their only field of application.

We may also consider conformal transforms to new coordinates (u, v) and then a more general self-adjoint PDE.

$$\partial_u(p^2 v^\alpha \phi_{|u}) + \partial_v(p^2 v^\alpha \phi_{|v}) + p^2 v^\alpha (q\phi + s) = 0. \quad (4.161)$$

Here, ϕ , p , q , and s are smooth functions of u and v , which must satisfy the regularity condition that they are *even* functions of v in the vicinity of the "optic" axis $v = 0$. Moreover, we assume that $p(u, v) > 0$, so that there is no other singularity of the PDE, than this axis.

By analogy with the transform from Eq. (2.10) to Eq. (2.14), the factor $p(u, v)$ can here be removed by introduction of a new potential $V = p\phi$, whereupon Eq. (4.161) simplifies to

$$\Delta_\alpha V(u, v) := -g(u, v) = -qV + Vp^{-1}\Delta_\alpha p - ps. \quad (4.162)$$

Here the differential operator Δ_α refers to the new coordinates u and v instead of z and r . The factor $\Delta_\alpha p$ can be approximated by a five-point formula if it is not known analytically. We can therefore regard the function $g(u, v)$ as linear in V with *known* coefficients. The mesh formula does not require any special information about this function: it must be defined and regular on the axis. It is therefore permissible and advantageous to introduce the modified potential

$$W(u, v) = V(u, v) + h^2 g(u, v)/12, \quad (4.163)$$

in analogy to Eq. (4.139). The mesh formulas (4.140) and (4.145) can now be used for the numerical field calculation. In this context, it is favorable to eliminate the ancillary function $V(u, v)$ completely from Eq. (4.162) and

(4.163). The source function to be used is then

$$g(u, v) = (\hat{q}W + ps)/(1 + h^2\hat{q}/12) \quad (4.164a)$$

with

$$\hat{q}(u, v) = q - p^{-1}\Delta_\alpha p. \quad (4.164b)$$

The solution technique is now an *iterative* one, familiar in the practical application of the FDM. At the beginning, some reasonable guesses for all the unknown functions are made. These do not need to be accurate but must be in agreement with the prescribed boundary conditions. Thereafter we start the iterations, each of which consists of two steps: (1) solution of Eqs. (4.140) and (4.145) with the current values of the sources g_{ik} ; (2) recalculation of this array of sources from the current values of the potentials. The iteration technique is studied in some detail in Section 5.6. Finally, when convergence has been achieved, the original potential $\phi(u, v)$ is obtained from

$$\phi(u, v) = (W - h^2g/12)/p, \quad (4.165)$$

whereupon the whole procedure is finished.

This method is equivalent to the one published by Hawkes and Kasper [9] and contains the formulas of reference [18] as special cases.

4.4.7 Poisson Equation in Spherical Meshes

As a practical example of the application of the method outlined above, we study here the rotationally symmetric Poisson equation in spherical coordinates R , ϑ , and φ . For reason of conciseness we shall confine the presentation to fields that are independent of φ , which is the most important practical case.

In cylindrical coordinates, the PDE can be cast in the self-adjoint form

$$\partial_z(r\bar{\epsilon}\bar{\phi}_{|z}) + \partial_r(r\bar{\epsilon}\bar{\phi}_{|r}) = -r\bar{\rho}(z, r), \quad (4.166)$$

$\bar{\epsilon}(z, r)$ being the dielectric coefficient and $\bar{\rho}(z, r)$ the charge density. As already mentioned earlier the conformal transform to an exponentially expanding mesh is quite favorable, and we therefore introduce the complex exponential

$$\begin{aligned} t(u + i\vartheta) &= R_0 \exp(u + i\vartheta) = z + ir, \\ z &= R_0 e^u \cos \vartheta, \quad r = R_0 e^u \sin \vartheta, \end{aligned} \quad (4.167)$$

where $R = R_0 \exp(u)$ is the transformed spherical coordinate. Because any conformal transformation leaves the numerical values of the coefficients invariant,

we find immediately

$$\partial_u(\varepsilon e^u \sin \vartheta \phi|_u) + \partial_\vartheta(\varepsilon e^u \sin \vartheta \phi|_\vartheta) = -R_0^2 e^{3u} \rho \sin \vartheta \quad (4.168)$$

with $\phi(u, \vartheta) = \bar{\phi}(z, r)$, and so on.

This can be in the form of Eq. (4.161) with

$$\begin{aligned} \alpha &= 1, \quad q \equiv 0 \\ v &= \vartheta \text{ for } \vartheta \leq \pi/2, v = \pi - \vartheta \text{ else} \\ p(u, v) &= (\varepsilon e^u v^{-1} \sin v)^{1/2} \end{aligned} \quad (4.169)$$

and

$$s(u, v) = R_0^2 \varepsilon^{-1} e^{2u} \rho(u, v) = R^2 \rho / \varepsilon. \quad (4.170)$$

In the case of constant dielectric coefficient ε , it is possible to carry out the analytical calculations some steps further. The coefficient $\hat{q}$ in Eq. (4.164a) then becomes $\hat{q} = p^{-1} \Delta_1 p$, giving

$$\hat{q}(v) = \frac{1}{4 \sin^2 v} - \frac{1}{4v^2} = \frac{1}{12} \left(1 + \frac{v^2}{5} + \frac{2v^4}{63} \right) + O(v^6). \quad (4.171)$$

This coefficient is always positive and does not exceed the value 0.15; the implicit procedure is hence very stable. Its practical applicability and potentially high accuracy were demonstrated by Killes [20], who applied the method to field calculation in thermionic electron guns with space charge and a pointed cathode. These formulas are significantly more accurate than the corresponding five-point formulae [21–23].

The nine-point formula for Poissons equation can also be derived *directly*, and the resulting expression is even simpler. We assume here $\varepsilon = 1$ and set out from the familiar form

$$\nabla^2 V(R, \vartheta, \varphi) = -\rho(R, \vartheta, \varphi) \quad (4.172)$$

in spherical coordinates R, ϑ , and φ . The first step is then a product separation with respect to the azimuth φ , equivalent to a Fourier analysis of the form

$$U = \sum_m \operatorname{Re}[U_m(R, \vartheta) \sin^{|m|} \vartheta \exp(-im\varphi)] \quad (4.173)$$

to be applied to $U = V$ or $U = \rho$. On introducing these into Eq. (4.172) and considering the linear independence of trigonometric functions, the separated

factors can be cancelled out on both sides and we obtain

$$\begin{aligned} R^{-1} \frac{\partial^2}{\partial R^2} (RV_m(R, \vartheta)) + R^{-2} (V_{m|\vartheta\vartheta} + (2|m| + 1) \cot \vartheta V_{m|\vartheta}) \\ = -\rho_m(R, \vartheta). \end{aligned} \quad (4.174)$$

We now introduce once again the abbreviation $\alpha = 2|m| + 1$ and the transform $R = R_0 \exp(u)$, R_0 being a free parameter. The Fourier potentials then become $V_m(R, \vartheta) = \phi_m(u, \vartheta)$, but we shall drop the ever-present subscript m for reasons of conciseness. We then arrive at the favorable final form

$$\phi_{|uu} + \phi_{|u} + \phi_{|\vartheta\vartheta} + \alpha \cot \vartheta \phi_{|\vartheta} = -Q(u, \vartheta) = -R^2 S_m(R, \vartheta). \quad (4.175)$$

This is a special case of Eq. (4.125) with $v = \vartheta$, and hence the method developed in Section (4.4.2) can be applied directly with $a_2 = 1$, $b_2 = 1$, $a_1 = 1$ and $b_1 = \alpha \cot \vartheta$. Here we consider temporarily the more general case that $a_1 = \beta$; hence,

$$U''(u) + \beta U'(u) = f(u). \quad (4.176)$$

Because the coefficients are *constants*, the same must hold for those of the discretization. These are easily determined by the requirement that the homogeneous ODE has the two solutions $U = \text{const.}$ and $U = \exp(-\beta u)$, which must be reproduced exactly. For equidistant meshes, $u_2 - u_1 = u_3 - u_2 = h$, this results in

$$X_2 = -2, \quad X_{1,3} = 1 \mp \tan h(\beta h/2), \quad (4.177)$$

and the two outer source coefficients can be chosen to be proportional to these:

$$S_{1,3} = h^2 X_{1,3}/12. \quad (4.178a)$$

The central source coefficient is obtained by a Taylor series expansion of $U(u)$ and $f(u)$ around the positions $u = u_2$ and evaluation of these for $u = u_1$ and u_3 , respectively. The result can be cast in the form

$$S_2 = \frac{h^2}{6} (5 - p), \quad p = \frac{1}{2} (h\beta)^2 + \frac{1}{15} (h\beta)^4. \quad (4.178b)$$

The first contribution to p results from the condition that all errors of order h^4 shall be cancelled out, the second one from the additional requirement that the terms of sixth order shall be cancelled out simultaneously for the functions $U = \exp(\beta u)$ and $U = \exp(-2\beta u)$. The resulting discretization

include Eq. (4.124) as the special case for $\beta = 0$, as it must be; subsequently we shall use the formulas for $\beta = 1$.

The second ODE, corresponding to Eq. (4.127), is now

$$V''(v) + \alpha \cot v V'(v) = g(v), \quad (4.179)$$

with $v \equiv \vartheta$. This is a generalization of Eq. (4.108), where $\cot v = v^{-1}$ for $v \ll 1$; hence its discretization must necessarily become more complicated. One possible way is to make a suitable modification of the method outlined in Section (4.4.1). Another is the method of repeated differentiation of the ODE. In order to cast this in a concise form, we write

$$c := \cot(v), \quad c'(v) = -(1 + c^2); \quad (4.180)$$

hence the repeated differentiation of the cotangent leads to a polynomial in it. Derivatives of higher orders of $V(v)$ can similarly be reduced to the first order by means of the ODE. We thus obtain in turn

$$\begin{aligned} V'' &= g - \alpha c V' \\ V''' &= g' - \alpha c g + [\alpha + c^2 \alpha(\alpha + 1)] V' \\ V^{(4)} &= g'' - \alpha c g' + [2\alpha + c^2 \alpha(\alpha + 2)] g \\ &\quad - [\alpha(2 + 3\alpha)c + \alpha(\alpha + 1)(\alpha + 2)c^3] V' \end{aligned} \quad (4.181)$$

and so on. It is sufficient to truncate this sequence after the fourth order as higher orders cannot be considered correctly in the discretization. These relations hold for any position v , and we now choose the central position v_2 in the mesh. Thus with $\tau := v - v_2$ we can write

$$V(v_2 + \tau) = \sum_{n=0}^4 \tau^n V^{(n)}(v_2)/n! \equiv V_2 + \frac{\tau^2}{2} \bar{g} + \tau V'_2/\eta(c, \tau) \quad (4.182)$$

with $\bar{g}$ containing all sources terms and the last term comprising all contributions that are proportional to V'_2 . From these we obtain

$$\begin{aligned} \eta(\tau, c) = & \left[1 - \frac{\alpha c \tau}{2} + \frac{\alpha \tau^2}{6} + \frac{\alpha(\alpha + 1)c^2 \tau^2}{6} \right. \\ & \left. - \frac{\alpha(2 + 3\alpha)c \tau^3}{24} - \frac{\alpha(\alpha + 1)(\alpha + 2)c^3 \tau^3}{24} \right]^{-1}. \end{aligned} \quad (4.183)$$

This reciprocal is again expanded as a power series, which can be truncated after the third order. In this series, the cubic terms can be replaced as follows

$$c^2\tau^2 - c^3\tau^3/2 \rightarrow c^2\tau^2/(1 + c\tau/2), \quad (4.184)$$

whereupon we arrive at

$$\eta(\tau, c) = 1 + \frac{\alpha c \tau}{2} - \frac{\alpha^2 \tau^2}{12} + \frac{\alpha(\alpha - 2)(c^2 + 1)\tau^2}{12 + 6c\tau}. \quad (4.185)$$

With $\tau \rightarrow 0$, $c \rightarrow \infty$, in such a way that $t = c\tau/2$ remains confined, we find Eq. (4.121b) as a special case, as we must.

Equation (4.182) can be used in different ways. One possibility, not pursued here, is to solve it for V'_2 that gives a one side finite difference approximation for this derivative. The other way is to evaluate it for $\tau = \pm h$. Then the left-hand side is known from V_3 or V_1 . The condition that the derivative V'_2 must be continuous and which is now eliminated, results in the required mesh formula. This procedure leads immediately to the coefficients

$$\left. \begin{aligned} Y_1 &= \eta(-h, c), & Y_3 &= \eta(+h, c) \\ Y_2 &= -Y_1 - Y_3 =: 2(1 + \gamma) \\ \gamma &= \frac{h^2}{12} \left(\frac{\alpha(\alpha - 2)(c^2 + 1)}{1 - h^2 c^2/4} - \alpha^2 c^2 \right). \end{aligned} \right\} \quad (4.186)$$

Again, the two outer source coefficients can be assumed to be proportional to the corresponding Y -coefficients:

$$T_1 = \frac{h^2}{12} Y_1, \quad T_3 = \frac{h^2}{12} Y_3. \quad (4.187)$$

The central one is now obtained from an analysis by Taylor series expansions. The result is then analogous to Eq. (4.178a):

$$T_2 = \frac{h^2}{6} (5 - \gamma - q). \quad (4.188)$$

Already with $q \equiv 0$, all discretization errors up to the fifth order are cancelled out. The quantity q can be determined in such a way that the function $V = \cos v$, being an eigenfunction of (4.179) (since $q = (1 + \alpha) \cos v$), is reproduced with high accuracy. A fairly good approximation is then

$$q = h^4 \tilde{C}_\alpha - h^2 \gamma [2(1 + \alpha)]^{-1}, \quad (4.189)$$

with the empirically found coefficients $\tilde{C}_1 = 11/120$, $\tilde{C}_3 = 209/480$, $\tilde{C}_5 = 1.1$ for $\alpha = 1, 3, 5$.

For $\vartheta = v = 0$ a special axial formula is again necessary. This is easily obtained from the requirement, that the functions $\sin^2(v/2)$ and $\sin^4(v/2)$ must be represented exactly. The formula then takes the form

$$V_1 - V_0 = T_0 g_0 + T_1 g_1 \quad (4.190)$$

with the coefficients

$$T_1 = \sin^2(h/2) / [1 + (2 + \alpha) \cos h], \quad (4.191a)$$

$$T_0 = 2 \sin^2(h/2) / (1 + \alpha) - T_1 \cos h. \quad (4.191b)$$

We are ready to set up the two-dimensional mesh formula. This can again be cast in a form in which all source coefficients (apart from the central one) are proportional to the corresponding potential coefficients, so that the analogy with Eqs. (4.139) and (4.140) is reasonable.

The result of the evaluation of Eq. (4.130) can be cast in the astonishingly simple form

$$P_{m,n}^{(k)} = A_m B_n^{(k)} \quad (1 \leq m \leq 3, 1 \leq n \leq 3) \\ (m, n) \neq (2, 2) \quad (4.192)$$

with the two triples

$$A = [1 - \tan h(h/2), \quad 4 - h^2/2 - h^4/15, \quad 1 + \tan h(h/2)] \quad (4.193a)$$

$$B^{(k)} = [\eta(-h, c_k), \quad 4 - 2\gamma_k - q_k, \quad \eta(+h, c_k)]. \quad (4.193b)$$

The central coefficient is determined by the requirement that

$$\sum_{m=1}^3 \sum_{n=1}^3 P_{m,n}^{(k)} = 0 \quad (4.193c)$$

must be satisfied. The label k appears via

$$c_k = \cot(v_k) = \cot(hk). \quad (4.194)$$

The modified potential now becomes

$$W(u, v) = \phi(u, v) + \frac{h^2}{12} Q(u, v), \quad (4.195)$$

with the source function Q given in Eq. (4.175). The right-hand side of Eq. (4.140) is to be replaced by

$$S_{ik} = -6h^2 Q_{ik}(1 - h^2/12 - h^4/90)(1 - q_k/6), \quad (4.196)$$

the last two factors comprising the correction for the coordinate transform.

The discretization is to be completed by axial formulas for $v = 0$ and $v = \pi$. These have the same form, and we therefore present only one of them. The combination of Eqs. (4.177) and (4.178) with (4.190) and (4.191) results in a scheme

$$\mathbf{P}^{(0)} = \begin{pmatrix} A_1 B_1^{(0)} & \bar{B}_1 & A_3 B_1^{(0)} \\ A_1 B_0^{(0)} & \bar{B}_0 & A_3 B_0^{(0)} \end{pmatrix}, \quad (4.197a)$$

(apart from the shift of labels) with the coefficients

$$\begin{aligned} B_0^{(0)} &= 12T_0/h^2 - 1, & B_1^{(0)} &= 12T_1/h^2 + 1, \\ \bar{B}_0 &= -2(6 - p + B_0^{(0)}), & \bar{B}_1 &= 2(6 - p - B_1^{(0)}). \end{aligned} \quad (4.197b)$$

Again the sum of all the coefficients in Eq. (4.197a) must vanish and this is satisfied by (4.197b) because $A_1 + A_3 = 2$. The source terms can again be cast in a form that is analogous to Eq. (4.145), if Eqs. (4.190) and (4.191) are combined with Eqs. (4.177) and (4.178). After some elementary calculations, the necessary central coefficients are found to be

$$\begin{aligned} \bar{g}_0 &= (12 - 2p)(B_0^{(0)} + 2) \\ \bar{g}_1 &= (12 - 2p)(B_1^{(0)} - 2). \end{aligned} \quad (4.198)$$

This kind of discretization has certain advantages. It describes Coulomb potentials exactly and homogeneous fields and those of source-free dipoles with very high accuracy. In all other cases it is essentially more accurate than the five-point formula for the same configuration. The author has checked this numerically for some analytical test functions given in the table below. The quantity N is the number of equidistant intervals on a semicircle, the mesh size is consequently $h = \pi/N$ in radians. The quotients Q_5 and Q_9 are the convergence rates, obtained by halving the mesh size. $P_2(\cos v)$ denotes the second Legendre polynomial.

Because the local errors do not entirely determine the accuracy finally obtained, the Dirichlet problem was solved for a spherical shell with $R_0 \leq R \leq R_{\max}$ fitting the mesh conditions. The test function was chosen to be a linear combination of those given in Table 4.1 with the coefficients given in Table 4.2, which shows the results obtained.

TABLE 4.1
RELATIVE LOCAL DISCRETIZATION ERRORS ($\alpha = 1$)

Position $R = 5.72778705, v = \theta = 60^\circ$					
$N = 18 : h = 10^\circ$					
No	Value		Error, 5-point	Error, 9-point	Functions
0	+1.000000000		+0.00000	+0.00000	constant 1
1	+0.174587496		-1.93709 e-5	+0.00000	1/R
2	+2.863893525		+1.15990 e-4	-6.29886 e-10	$\cos(v)*R$
3	+0.015240397		+5.72855 e-5	-6.29887 e-10	$\cos(v)/R^2$
4	-4.100943062		+6.23637 e-4	-2.18297 e-6	$P_2(\cos v)*R^2$
5	-0.000665196		-3.48434 e-7	+1.45210 e-9	$P_2(\cos v)/R^3$
6	+5.467924082		+6.21761 e-4	-7.13296 e-6	$R^2/6$
7	+0.030480794		-6.30872 e-7	-2.14504 e-10	$1/R^2$
8	+0.007620198		-1.25628 e-6	-6.64104 e-7	$\cos^2 v/R^2$

$N = 36 : h = 5^\circ$					
No	Q_5	Q_9	Error, 5-point	Error, 9-point	Sources*R ²
0	-	-	+0.00000	+0.00000	0.
1	16.02	-	-1.20884 e-6	+0.00000	0.
2	16.00	256.6	+7.24936 e-6	-2.45468 e-12	0.
3	15.85	256.7	+3.61363 e-6	-2.45462 e-12	0.
4	16.10	63.07	+3.87418 e-5	-3.46107 e-8	0.
5	16.04	63.07	-2.17196 e-8	+2.30229 e-11	0.
6	16.06	64.45	+3.87123 e-5	-1.10673 e-7	-R ²
7	64.20	256.8	-9.82526 e-9	-8.35127 e-13	-2/R ²
8	64.00	63.48	-1.96292 e-8	-1.04613 e-8	(2 cos ² v - 4)/R ²

TABLE 4.2
GLOBAL ABSOLUTE DISCRETIZATION ERRORS ($\alpha = 1$)

Coefficients = (1, 1, 0.5, 0.125, 0, 0.25, 0, -0.25, 0.125)			
Radial extent : $R_0 = 1,$		$R_{\max} = 187.914629$	
MESH : $M = 30, N = 18,$		$h = 10^\circ$	
Five points : $R_m = 1.417743,$	$P_m = 2.156862,$	Error	3.68970 e-3
Nine points : $R_m = 2.009994,$	$P_m = 1.872369,$	Error	1.61241 e-6
MESH : $M = 60, N = 36,$		$h = 5^\circ$	
Five points : $R_m = 1.842024,$	$P_m = 1.923653,$	Error	1.26444 e-3
Nine points : $R_m = 144.6313,$	$P_m = 19.08585,$	Error	1.29366 e-8

In this table, the value R_m is the radius of the point with worst error and P_m the corresponding potential; the angular coordinate was always $v = 0$. Evidently the rate of convergence for the five-point formula is rather poor, whereas for the nine-point formula it is excellent. This demonstrates that the increased effort required to derive and program it finally pays off.

The field of application of this method is the calculation of guns with pointed filaments and space charge. The method of Fourier analysis outlined here also provides a way of calculating space charge distributions that are not rotationally symmetric about an optic axis.

4.5 IRREGULAR CONFIGURATIONS

The irregular configurations appearing in the vicinity of boundaries that do not fit the mesh require special consideration, and it is this difficulty that makes the FDM less attractive than the FEM. Certainly, any loss of symmetry leads to a corresponding loss of accuracy, whereas in the FDM this is confined to a narrow domain near the boundary. In a typical FEM-grid this is spread over the whole domains of solution. The consequences will be discussed further below.

Depending on the different special properties of the boundary value problem to be solved, corresponding different strategies must be followed in the derivation of suitable mesh formulas. This is not very attractive but is not an insurmountable obstacle.

4.5.1 Inner Mesh Points

A typical situation is shown in Fig. 4.12: a mesh formula for the irregular inner node 0 is to be derived, the potentials at the surrounding nodes being given.

If the assumptions for the general five-point formulas from Section 4.2 are satisfied, the problem is already solved by these, and this is the simplest way if a discretization error of third order can be tolerated. The PDEs (4.16) or

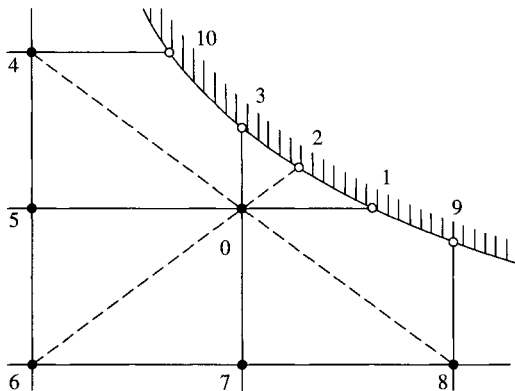


FIGURE 4.12 Example of an irregular configuration obtained by intersection of a rectangular mesh with a curved boundary.

(4.17) cannot be solved in this way because at least one neighboring point in a diagonal direction is needed to approximate the mixed derivative $\phi_{|uv}$ associated with the coefficient C .

Without loss of generality, we can shift the origin of the (u, v) coordinate system to the node 0, and we shall adopt this from now on. The PDE (4.17) can then be rewritten as

$$\begin{aligned} A\phi_{|uu} + B\phi_{|vv} + 2C\phi_{|uv} + A_{|u}\phi_{|u} + B_{|v}\phi_{|v} \\ + C_{|u}\phi_{|v} + C_{|v}\phi_{|u} = D\phi + Q, \end{aligned} \quad (4.199)$$

all function values and derivatives referring to the origin. The terms with coefficients A and B and their derivatives can be approximated in the usual manner, as can the terms with $\phi_{|u}$ and $\phi_{|v}$. The only novel aspect here is the approximation

$$\phi_{|uv} = (\phi_0 + \phi_6 - \phi_5 - \phi_7)/h^2 + O(h) \quad (4.200)$$

for the configuration of Fig. 4.12. Introducing this together with all other approximations into Eq. (4.199) and solving for ϕ_0 , we obtain the required mesh formula, which has a discretization error of third order. Such a procedure is always feasible because there is always a regular mesh on the opposite side of the irregular part.

Another problem that arises near a boundary is caused by the fact that there remain nonrectangular domains for which the accumulation procedure, outlined in Section 4.3, is *not* applicable because one point is missing. A typical situation is shown in Fig. 4.13. In general, only a *linear* approximation for the potential can then be made, which corresponds to the linear approximation in the FEM. The contributions to the functional become simply

$$\begin{aligned} \delta F_G = \frac{b}{4a}A_m(\phi_1 - \phi_0)^2 + \frac{a}{4b}B_m(\phi_3 - \phi_0)^2 \\ + C_m(\phi_1 - \phi_0)(\phi_3 - \phi_0), \end{aligned} \quad (4.201a)$$

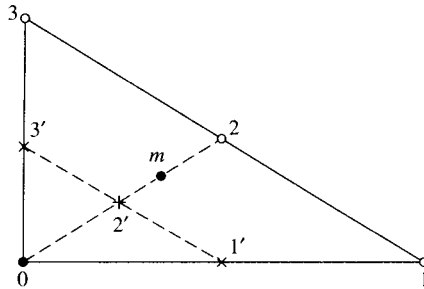


FIGURE 4.13 Degenerate case of a triangle in the functional algorithm.

$$\delta F_S = \frac{ab}{2} W_m = \frac{ab}{2} \left(Q_m \phi_m + \frac{1}{2} D_m \phi_m^2 \right), \quad (4.201b)$$

where $\phi_m = (\phi_0 + \phi_1 + \phi_3)/3$. All functions refer here to the centroid m of the corresponding triangle. The evaluation of these expressions is simple and leads to mesh formulas having a discretization error of third order.

This loss of accuracy can be diminished substantially in the simple but frequently occurring case of isotropic coefficients and square meshes: $A \equiv B \equiv \varepsilon$, $C \equiv 0$, $a = b = h$. It is then reasonable to use the potential at the point $2'$ instead of the missing point 2 in Fig. 4.8. The formula can be cast in the convenient form

$$\sum_{n=1}^8 \frac{\bar{g}_n}{21} [\varepsilon_{n'} (\phi_n - \phi_0) - h^2 T_n / 12] = \frac{17}{84} h^2 T_0 \quad (4.202)$$

with relative weights $\bar{g}_n$ given in Fig. 4.14.

Here ϕ_2 is now the new inserted point and the coefficients $\varepsilon_{2'}$ refers to the midpoint between this at the center 0. The generation of these weights can easily be understood from the elimination process in the incomplete element:

$$\begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix} + \begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ -21 & 3 \end{bmatrix} \quad (4.203)$$

Equation (4.202) has an error of *fourth* order, which is about 1/21 of that for the conventional five-point formula, but more important is the fact that it can be incorporated in the allocation scheme. To achieve this we have to replace

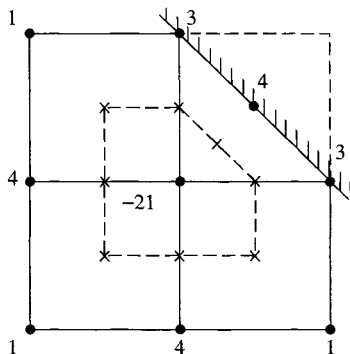


FIGURE 4.14 Nine-point configuration with one missing diagonal point; the numbers are here relative summation weights. The positions for the evaluation of material coefficients are marked by crosses.

Eq. (4.201a) with

$$\delta F_G = \frac{1}{12}[\varepsilon_{1'}(\phi_1 - \phi_0)^2 + 4\varepsilon_{2'}(\phi_2 - \phi_0)^2 + \varepsilon_{3'}(\phi_3 - \phi_0)^2 + R_G], \quad (4.204)$$

R_G being an unimportant remainder that cancels out by differentiation with respect to ϕ_0 . The contribution to the source term requires some short consideration. Because the finite difference $\hat{T}_D$ in Eq. (4.92) cannot be formed here, the allocation scheme leading to Eq. (4.93) must be formulated directly without it. For a square of area h^2 the formula replacing Eq. (4.72) is then

$$\delta F_S = \frac{h^2}{72}\phi_0(13T_0 + 2(T_1 + T_3) + T_2) + R_s, \quad (4.205)$$

R_s being the remainder, which has three analogous terms, found by appropriate change of the labels. In the case of the triangle, the elimination of T_2 and replacement by the value on the midpoint of the *hypotenuse* results in

$$\delta F'_S = \frac{h^2}{72}(12T_0 + T_1 + T_3 + 4T_2) + R'_s \quad (4.206)$$

in analogy to the scheme given earlier.

In every case where a diagonal point is missing, the corresponding quadrant can be treated similarly.

4.5.2 Edge or Corner Singularities

The general series expansion for the Laplace equation is already given by Eqs. (2.93). These are now used to derive mesh formulas. The situation is complicated by the fact that sharp corners cannot appear for physical reasons and must hence be rounded off, but the radius of least curvature is often not known exactly. We shall therefore present here different approaches, which are always simplifications.

The simplest way of smoothing a rectangular edge or corner is shown in Fig. 4.15a; the original edge becomes an internal point, and the potential at it is then given roughly by

$$\phi_0 = \frac{1}{2}(\phi_3 + \phi_4) + K(\phi_1 + \phi_2 - \phi_3 - \phi_4), \quad (4.207)$$

K being a free parameter, for which $K = 0.2$ is a reasonable value.

Another possibility is shown in Fig. 4.15b. The potentials at the nearest neighbors of the edge are corrected by interpolation formulas using four potentials in the corresponding line. To find these, we write down Eqs. (2.93) for $U = \phi$, $\gamma = \pi/2$, $r = nh$, ($n = 1, 2, 3$), and $m = 1, 2$.

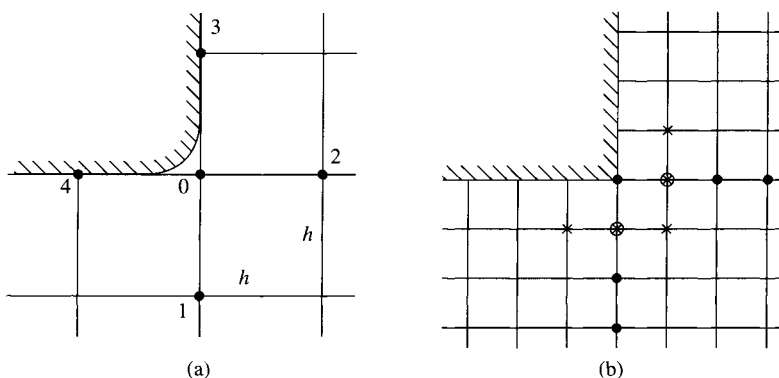


FIGURE 4.15 Corner singularity: (a) treatment by rounding-off the corner, Eq. (4.207); (b) different methods for sharp corners. Encircled crosses and dots: linear interpolation, see Eqs. (4.209a,b). Nine-point formula (4.210). Encircled crosses: $c = 0.528$ and other crosses $c = 0.895$.

The result can be cast in the form

$$\phi_n - \phi_0 = \sum_{m=1}^2 b_m n^{m\mu} \quad (n = 1, 2, 3), \quad (4.208)$$

the coefficients b_m containing fractional powers of h and sines as factors. It is not necessary to know these explicitly. Instead, we write down Eq. (4.208) for $n = 2$ and 3 , the left-hand side being given. These are now two linear equations for b_1 and b_2 . The solution is then introduced into Eq. (4.208) for $n = 1$. The final result is

$$\phi_1 = A\phi_2 + B\phi_3 + (1 - A - B)\phi_0 \quad (4.209a)$$

with

$$A = \frac{3^\mu - 1}{2^\mu(3^\mu - 2^\mu)}, \quad B = \frac{1 - 2^\mu}{3^\mu(3^\mu - 2^\mu)}. \quad (4.209b)$$

In the most important case of a rectangular corner, $\mu = 2/3$, these coefficients become $A = 1.381$, $B = -0.573$, $1 - A - B = 0.192$.

The fact that $B < 0$ causes difficulties during the iterative solution of the FDM equations. Another shortcoming of this approach is the fact that the basic Eqs. (2.93) hold only for the planar Laplace equation. Nevertheless, Lenz [24] applied it successfully to rotationally symmetrical configurations.

A third alternative that can also be combined with the edge-rounding approach, is the use of nine-point formulas. In these, the ratio of the coefficients can be adjusted to eliminate the singularity of lowest order. We consider first the case of a planar field, $\phi_{|xx} + \phi_{|yy} = 0$. The mesh formula then takes the

basic form

$$\begin{aligned} \phi_{ik} = & c_{ik}(\phi_{i+1,k} + \phi_{i-1,k} + \phi_{i,k+1} + \phi_{i,k-1})/4 \\ & + (1 - c_{ik})(\phi_{i+1,k+1} + \phi_{i-1,k+1} + \phi_{i+1,k-1} + \phi_{i-1,k-1})/4 \end{aligned} \quad (4.210)$$

At a large distance from the corner, the appropriate value is $c = 0.8$. If the corner has the labels (i_c, k_c) , then the coefficients c_{ik} with $|i - i_c| \leq 1$ and $|k - k_c| \leq 1$ must be modified. Numerical investigations carried out by the author have shown that $c_{ik} = 0.528$ for closest neighbors and $c_{ik} = 0.895$ for neighbors in diagonal directions are favorable values for an unsmoothed corner, which is thus always accorded a lower relative weight.

The *radial* correction for the solution of the cylindrical Laplace equation $\Delta_\alpha \phi = 0$ can now be found easily by introducing the corresponding factors $1 \pm \alpha/2k$ for the points with labels $(i, k \pm 1)$. This procedure can be straightforwardly incorporated in FDM programs; however, the functional method of Section 4.3 is here less favorable because it requires a post correction.

4.5.3 Mesh Points on Boundaries of Materials

We first consider the simpler case in which the corresponding point is the center of parallel epipedal structure, which is bisected by the surface along a meshline (Fig. 4.16a). This case can be treated quite easily by means of the method outlined in Section 4.3. Because the accumulation of the functional is to be performed for each element without coupling to the neighboring ones, the correct result is obtained by simply making use of the material coefficients that are appropriate for each element in question. This procedure is completely equivalent to that used in the FEM.

The situation is more complicated if the boundary passes through a diagonal (Fig. 4.16b). Then Eqs. (4.201) must be applied to the triangular elements, which makes the final result less accurate and more complicated, though there is no essential obstacle to this procedure.

An essential simplification is possible if the mesh is square-shaped and Eq. (4.29) is to be solved, the coefficient $\varepsilon(u, v)$ being isotropic in each medium but discontinuous at the surface (see Fig. 4.16c,d). It is then not necessary to evaluate the formalism in its general form, but we can use Eq. (4.93) in a modified way: all material coefficients ε_n , T_0 , and T_n , which become discontinuous, are simply replaced with their *arithmetic means*

$$\begin{aligned} \varepsilon'_n &\rightarrow \bar{\varepsilon}_n = (\varepsilon_n^{(+)} + \varepsilon_n^{(-)})/2, \\ T_n &\rightarrow \bar{T}_n = (T_n^{(+)} + T_n^{(-)})/2. \end{aligned} \quad (4.211)$$

The corresponding modification holds also for Eq. (4.97) in hexagonal structures.

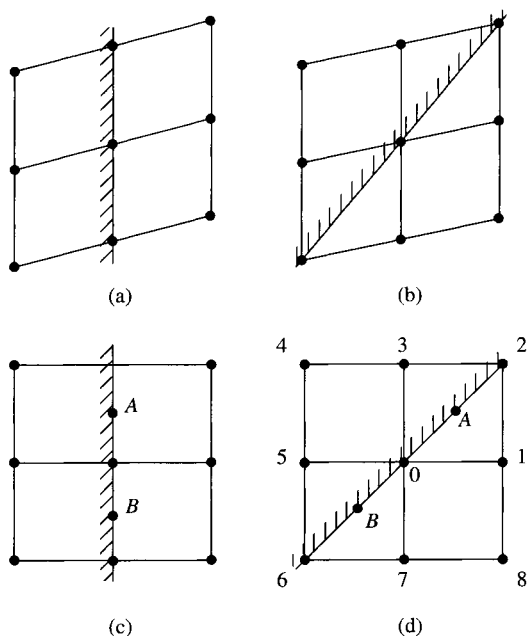


FIGURE 4.16 Different situations that can arise when fitting a material boundary to a mesh: (a) boundary on a mesh line: functional accumulation in the four coherent parallelepipeds; (b) boundary on a diagonal: functional accumulation in two parallelepipeds and four triangles; (c) square-shaped meshes, boundary on a mesh line: use the nine-point formula with the arithmetic means of the material coefficients at the points A and B ; (d) boundary on a diagonal: do the same as (c) with respect to points A , B , and 0 ; apply Eq. (2.212) to points A and B .

This replacement is valid only if the central point lies on a boundary: for inner points (Fig. 4.14) it is necessary to use Eq. (4.202) with continuous coefficients and an inserted diagonal point (A or B in Fig. 4.16d). The potentials at the latter are obtained according to the general rule, for example,

$$\sum_{n=0}^3 \left[\bar{\varepsilon}_n (\phi_n - \phi_A) - \frac{h^2}{2} \bar{T}_n \right] = 0. \quad (4.212)$$

An analogous method for the calculation of the field vector in magnetic circuits is reported in Morizumi [25].

4.5.4 Evaluation of Series Expansions

If the configuration round a mesh point is irregular as in Fig. 4.12, the analytic derivation of mesh formulas of higher than the second order becomes very

tedious. Only very few attempts have been made to overcome this obstacle therefore [8], and usually a discretization error of third order, resulting from irregular boundaries is tolerated as an unsurmountable limit.

A possibility for further development is the series expansion of the potential in terms of suitable *trial functions* with initially *unknown* coefficients. These trial functions $T_n(\mathbf{r})$, ($n = 1, \dots, N$) should satisfy exactly (or at least very accurately) the PDE to be solved. The expansion coefficients are then found from the condition that the given potentials are assumed at the surrounding points $(1, \dots, N)$. The potential at the central point is subsequently obtained by evaluation of the now known expansion at this position. Because this concept is rather general, we shall present it here in a form that could also be used in three dimension though this would become laborious.

Let D be a homogeneously linear partial differential operator of second order and $D\Phi(\mathbf{r}) = 0$ the PDE to be solved. Then a suitable set of trial functions must be chosen, which are linearly independent and satisfy $DT_n(\mathbf{r}) = 0$. The series expansion in question is now

$$\Phi(\mathbf{r}) = \sum_{n=1}^N c_n T_n(\mathbf{r}). \quad (4.213)$$

The condition that the prescribed values $\phi_j := \Phi(\mathbf{r}_j)$ ($j = 1, \dots, N$) are assumed by this expression requires the solution of a linear system of equations

$$\sum_{n=1}^N A_{jn} c_n := \sum_{n=1}^N T_n(\mathbf{r}_j) c_n = \phi_j \quad (j = 1, \dots, N). \quad (4.214)$$

We must assume here that the matrix A is nonsingular. Then we can write

$$c_n = \sum_{k=1}^N A_{nk}^{-1} \phi_k \quad (n = 1, \dots, N), \quad (4.215)$$

and introduce this into Eq. (4.213), whereupon our problem is solved. For the central mesh point $\mathbf{r}_0$, the condition

$$\phi_0 = \Phi(\mathbf{r}_0) = \sum_{k=1}^N \left(\sum_{n=1}^N T_n(\mathbf{r}_0) A_{nk}^{-1} \right) \phi_k \quad (4.216)$$

is obtained. This is none other than the required mesh formula

$$\phi_0 = \sum_{k=1}^N \mu_k \phi_k \quad (4.217a)$$

with the coefficients

$$\mu_k = \sum_{n=1}^N T_n(\mathbf{r}_0) A_{nk}^{-1} \quad (k = 1, \dots, N). \quad (4.217b)$$

Apart from this systematic derivation of mesh formulas, this method also offers a way of calculating derivatives

$$\nabla \Phi(\mathbf{r}) = \sum_{n=1}^N c_n \nabla T_n(\mathbf{r}) \quad (4.218)$$

and even derivatives of higher orders, if necessary. These could be used for purposes of field calculation and interpolation. Another essential task is the derivation of mesh formulas for positions at material boundaries, as dealt with in the preceding section.

The situation is shown in Fig. 4.17: we have a surface with normal $\mathbf{n}$ and different material coefficients ε_1 and ε_2 on either side. The numbers N_1 and N_2 of surrounding points may be different and surface points can be used twice. The set of trial functions here will be different, as a constant for the lowest order is unacceptable here. Instead of Eq. (4.213), there are now *two* series expansions

$$\Phi(\mathbf{r}) = \phi_0 + \sum_{n=1}^{N_k} c_n^{(k)} T_n(\mathbf{r}) \quad (k = 1, 2) \quad (4.219)$$

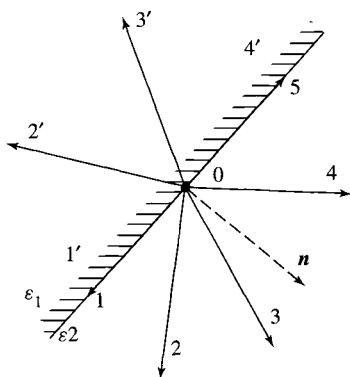


FIGURE 4.17 Arbitrary configuration of nodes and corresponding space vectors in the vicinity of a material boundary. Here the numbers $N_1 = 4$ and $N_2 = 5$ were chosen. The vector $\mathbf{n}$ is the surface normal at point 0.

to be used in the two separate domains. We have to calculate two inverse matrices and express the coefficients in terms of potential differences:

$$c_n^{(k)} = \sum_{j=1}^{N_k} (A_k^{-1})_{nj} (\phi_j - \phi_0) \quad (k = 1, 2). \quad (4.220)$$

It is now straightforward to write down the continuity condition for the position $\mathbf{r}_0$

$$\varepsilon_1 (\mathbf{n} \cdot \nabla \Phi)_1 = \varepsilon_2 (\mathbf{n} \cdot \nabla \Phi)_2, \quad (4.221)$$

leading to

$$\varepsilon_1 \sum_{i=1}^{N_1} c_i^{(1)} \mathbf{n} \cdot \nabla T_i(\mathbf{r}_0) = \varepsilon_2 \sum_{i=1}^{N_2} c_i^{(2)} \mathbf{n} \cdot \nabla T_i(\mathbf{r}_0). \quad (4.222)$$

On introducing the coefficients (4.220) into this equation, we obtain a linear relation of the general form

$$\varepsilon_1 \sum_{j=1}^{N_1} F_j^{(1)} (\phi_j - \phi_0) = \varepsilon_2 \sum_{j=1}^{N_2} F_j^{(2)} (\phi_j - \phi_0) \quad (4.223a)$$

with coefficients

$$F_j^{(k)} = \sum_{i=1}^{N_k} (A_k^{-1})_{ij} \mathbf{n} \cdot \nabla T_i(\mathbf{r}_0) \quad (k = 1, 2). \quad (4.223b)$$

It remains to find the discretization of an inhomogeneous linear PDE

$$D\Psi(\mathbf{r}) = S(\mathbf{r}), \quad (4.224)$$

$S(\mathbf{r})$ being a given source function. This problem can be reduced to the task of finding any convenient particular solution $\Psi_p(\mathbf{r})$ of this PDE. When this is found, we can use the above outlined relations because $\Phi(\mathbf{r}) = \Psi(\mathbf{r}) - \Psi_p(\mathbf{r})$ solves $D\Phi = 0$. This implies that for instance

$$\Psi(\mathbf{r}_0) = \Psi_p(\mathbf{r}_0) + \sum_{k=1}^N \mu_k (\Psi(\mathbf{r}_k) - \Psi_p(\mathbf{r}_k)) \quad (4.225)$$

follows from Eq. (4.217a), and also Eq. (4.223a) must be modified accordingly. From a purely theoretical standpoint, therefore, all the problems have

been solved in principle, but in practice there are numerous obstacles to be overcome.

There are two essential difficulties: the first is that the rank N of the matrix is rather high, and the number of neighbor points is correspondingly large if a certain polynomial order M of the approximation is to be reached: in two dimensions we have $N = 2M + 1$, whereas in three dimensions this number is $N = (M + 1)^2$. For a large number N , it might become difficult to find suitable neighbor points. The other, more serious problem arises from ill-conditioned or even singular matrices, which may result from an unsuitable choice of configurations. Here we shall outline briefly a method, which has been tested by the author and which gave good results. This method will also exhibit some interesting physical relations.

4.5.5 Harmonic Functions

For reasons of conciseness, we shall consider here only two-dimensional fields and start with the ordinary Laplace equation

$$D\Phi(x, y) : \equiv \Delta\Phi(x, y) \equiv \Phi_{|xx} + \Phi_{|yy} = 0. \quad (4.226)$$

It is favorable to shift the origin to the point $\mathbf{r}_0$ in question and to scale the new coordinates u, v in such a way that the maximum distance becomes unity:

$$u = (x - x_0)/H, \quad v = (y - y_0)/H, \quad (4.227a)$$

$$H^2 = \max[(x_n - x_0)^2 + (y_n - y_0)^2], \quad n = 1, \dots, N'. \quad (4.227b)$$

A very suitable set of trial functions is then

$$T_1(u, v) = 1,$$

$$T_{2n}(u, v) = \Re(u + iv)^n / n!, \quad (4.228a)$$

$$T_{2n+1}(u, v) = \Im(u + iv)^n / n! \\ (\text{for } n = 1, 2, \dots, (N' - 1)/2, N' \text{ odd}). \quad (4.228b)$$

These can be evaluated efficiently by pure algebraic operations. The factorials serve to improve the convergence of the search.

It is often necessary to choose a number $N' > N$ at the outset and also the number N'' of neighbor points larger than the minimum because some of the functions may be unsuitable for certain configurations. For example, in configurations without any neighbors in diagonal directions, we have $u_k v_k = 0$ for all k , and hence $T_5(u_k, v_k) = 0$, $T_9(u_k, v_k) = 0$. This does not mean that the

familiar five-point formula could not be found with this method, but simply that T_5 is not necessary for it. Because in a *general* procedure such difficulties are not known in advance, the inverse matrix should be determined with a *totally pivoting* algorithm, which thereby eliminates linearly dependent rows (functions) and columns (points).

The result is the optimum matrix A with rank $N \times N$. As a result of the particular choice Eqs. (4.228) the coefficients μ_k of Eq. (4.217) now simplify to

$$\mu_k := \beta_k = (A^{-1})_{1,k}, \quad (k = 1, \dots, N). \quad (4.229)$$

These are purely geometric quantities, so that for a given grid, this kind of calculations need be carried out only once for all irregular meshes.

The method also furnishes an intrinsic control of the discretization order finally achieved; for trial functions, including those omitted, we can determine the discretization errors

$$E_m := \sum_{k=1}^N \beta_k T_m(u_k, v_k) \quad (m = 2, 3, \dots, N'). \quad (4.230)$$

The first nonvanishing value indicates the error order $m/2$ that should not be < 4 , while of course $E_1 = 1$ must be satisfied. Another criterion for a suitable discretization is $\beta_k > 0$ for all k because negative coefficients are unfavorable with respect to iterative solution techniques. In the subsequent presentation, we shall assume that $E_m = 0$ for at least $2 \leq m \leq 7$. Recalling the special definitions in Eqs. (4.228), we can then write

$$\sum_k \beta_k u_k = \sum_k \beta_k v_k = \sum_k \beta_k u_k v_k = 0, \quad (4.231)$$

$$\sum_k \beta_k u_k^2 = \sum_k \beta_k v_k^2 =: W/2 > 0. \quad (4.232)$$

These sums can be interpreted as moments of a rotationally symmetric distribution. It is now appropriate to define a *radial* coordinate

$$\rho := (u^2 + v^2)^{1/2} \quad (4.233)$$

and then weight factors

$$w_k := \beta_k \rho_k^2 / W \quad (k = 1, \dots, N), \quad (4.234)$$

which are normalized to $\sum w_k = 1$, as required. These can be used to introduce new mean values, which do not vanish identically:

$$\bar{u} := \sum_k w_k u_k, \quad \bar{v} := \sum_k w_k v_k, \quad (4.235a)$$

$$\overline{\rho^2} := \sum_k w_k \rho_k^2 \equiv \sum_k w_k (u_k^2 + v_k^2). \quad (4.235b)$$

From $E_6 = E_7 = 0$, we can now conclude that

$$\sum_k \beta_k u_k v_k^2 = \sum_k \beta_k u_k^3 / 3 = W\bar{u}/4, \quad (4.236a)$$

$$\sum_k \beta_k v_k u_k^2 = \sum_k \beta_k v_k^3 / 3 = W\bar{v}/4, \quad (4.236b)$$

which will be needed later. The midpoint (x_0, y_0) is now the centroid in the above defined sense if $\bar{u} = \bar{v} = 0$, while Eqs. (4.231) are always necessarily satisfied.

We now consider the Poisson equation

$$\Delta \Psi(x, y) = S(x, y) \quad (4.237a)$$

or, equivalently, in terms of u and v ,

$$\Psi|_{uu} + \Psi|_{vv} = Q(u, v) := H^2 S(x, y). \quad (4.237b)$$

In the second order, the source function $Q(u, v)$ can be expanded as

$$Q(u, v) = Q_0 + a_1 u + a_2 v + a_3 \rho^2 + a_4 (u^2 - v^2)/2 + a_5 uv, \quad (4.238)$$

the term with ρ^2 not being a harmonic function. A suitable and simple solution of the Poisson equation is then readily obtained in the form

$$\begin{aligned} \Psi_p(u, v) = & \frac{\rho^2}{4} [Q_0 + (a_1 u + a_2 v)/2 + a_3 \rho^2 / 4 \\ & + a_4 (u^2 - v^2)/6 + a_5 uv / 3], \end{aligned} \quad (4.239)$$

as can be verified easily. The last two terms can be eliminated with $Q(u, v)$, resulting in

$$\Psi_p(u, v) = \frac{\rho^2}{12} [2Q_0 + Q(u, v) + (a_1 u + a_2 v)/2 - a_3 \rho^2 / 4]. \quad (4.240)$$

With $\Psi_p(0, 0) = 0$ and $\mu_k = \beta_k$, Eq. (4.225) now becomes

$$\psi_0 = \sum_k \beta_k (\psi_k - \Psi_p(u_k, v_k)), \quad (4.241)$$

which can be further evaluated by means of Eqs. (4.231) to (4.235b), resulting in:

$$\begin{aligned} \psi_0 = & \sum_k \beta_k (\psi_k - \rho_k^2 Q_k / 12) \\ & - W (8Q_0 + 2(a_1 \bar{u} + a_2 \bar{v}) - a_3 \bar{\rho}^2) / 48. \end{aligned} \quad (4.242)$$

This equation contains all the terms of the source function but is inconvenient because it requires a knowledge of the coefficients a_1 , a_2 , and a_3 , the derivatives of $Q(u, v)$ at the central point. These can be eliminated approximately in the following way:

$$\sum_k \beta_k \bar{u} u_k Q_k = \sum_k \beta_k \bar{u} u_k (Q_0 + a_1 u_k + a_2 v_k) = W a_1 \bar{u} / 2 \quad (4.243)$$

with reference to Eqs. (4.231 and 4.232), and similarly,

$$\sum_k \beta_k \bar{v} v_k Q_k = W a_2 \bar{v} / 2, \quad (4.244)$$

where we have truncated the series expansion (4.238) after the linear terms because the next terms would already lead to the third order.

The third coefficient a_3 is found in another way. Differentiation of Eq. (4.238) leads to the Poisson equation

$$Q_{|uu} + Q_{|vv} = 4a_3 = \text{const.} \quad (4.245)$$

A particular solution of it is simply $Q_p = a_3 \rho^2$. According to the general formula, we then have

$$Q_0 = \sum_k \beta_k (Q_k - Q_p(u_k, v_k)) = \sum_k \beta_k Q_k - a_3 W. \quad (4.246)$$

Now the elimination of a_1 , a_2 , and a_3 , is straightforward; the result can be cast in the concise form

$$\psi_0 = \sum_k \beta_k (\psi_k - q_k S_k) - q_0 S_0 \quad (4.247)$$

with the source coefficients

$$q_k = H^2(\rho_k^2 + \bar{u}u_k + \bar{v}v_k - \bar{\rho}^2/4)/12, \quad (k > 0), \quad (4.248a)$$

$$q_0 = H^2(W/6 + \bar{\rho}^2/48), \quad (4.248b)$$

$$q_0 + \sum_{k=1}^N \beta_k q_k = H^2 W/4 = \frac{1}{4} \sum_{k=1}^N \beta_k |\mathbf{r}_k - \mathbf{r}_0|^2. \quad (4.248c)$$

The coefficients β_k are independent of the scaling factor H , as they must be; they are all purely *geometric* quantities.

As a simple example, we study the results for a regular polygon with N edges. The normalizations (4.227) imply that all radial distances have the equal length H , and consequently $\rho_1 = \rho_2 = \dots \rho_N = 1$. For reasons of symmetry all the β -factors must become $\beta_k = 1/N$, consequently $W = 1$, $\bar{u} = \bar{v} = 0$, $\bar{\rho}^2 = 1$, whereupon Eq. (4.247) specializes to

$$\psi_0 = N^{-1} \sum_{k=1}^N (\psi_k - H^2 S_k/16) - 3S_0/16. \quad (4.249)$$

With $\varepsilon \equiv 1$ and after appropriate adjustment of the notation, we obtain Eq. (4.97) for $N = 6$, as expected.

4.5.6 Applications of the General Method

The method outlined earlier is not limited to the discretization of a Poisson equation (4.237a); this would not justify the explicit presentation of all these calculations. Much more generally, it is the basis for the discretization of general self-adjoint equations.

To cast the following calculations in a convenient form, we introduce the unscaled distances from the central point:

$$d_k := H\rho_k = [(x_k - x_0)^2 + (y_k - y_0)^2]^{1/2}, \quad (k = 1, \dots, N). \quad (4.250a)$$

As a first application we set $S(x, y) = \text{const.}$ in Eq. (4.237a), whereupon Eq. (4.247) with Eq. (4.248) specializes to

$$\psi_0 = \sum_k \beta_k \psi_k - \frac{1}{4} S_0 \sum_k \beta_k d_k^2. \quad (4.250b)$$

Solving for $S_0 = \Delta\Psi$, we obtain a general formula for the Laplacian:

$$\nabla^2\Psi \equiv \Delta\Psi = 4 \sum_k \beta_k (\psi_k - \psi_0) / \sum_k \beta_k d_k^2. \quad (4.251)$$

This simple approximately valid formula holds for *any* twice continuously differentiable function.

Next we consider the relation

$$\Delta[(x - x_0)(\Psi(x, y) - \psi_0)] = (x - x_0)\Delta\Psi + 2\Psi_{|x} \quad (4.252)$$

and an analogous formula for $\Psi_{|y}$. At the position $x = x_0$, $y = y_0$ the term in $\Delta\Psi$ cancels out. Using now the approximation (4.251) for the expression on the left-hand side, we obtain a simple approximation for the gradient:

$$(\nabla\Psi)_0 = 2 \sum_k \beta_k (\mathbf{r}_k - \mathbf{r}_0)(\psi_k - \psi_0) / \sum_k \beta_k d_k^2. \quad (4.253)$$

Again, this formula holds *generally* and hence provides a method of differentiation in irregular meshes. If the β -coefficients for both half-spaces at a material boundary (Fig. 4.17) are known, this formula also provides a general alternative approach for the evaluation of Eqs. (4.222) and (4.223).

A third application is the discretization of a PDE of the form

$$\Delta\Psi(x, y) + \mathbf{a}(x, y) \cdot \nabla\Psi(x, y) + b(x, y)\Psi(x, y) = Q(x, y). \quad (4.254)$$

Using the approximations (4.251) and (4.253), this discretization is straightforward and results in

$$\hat{\beta}_0 \psi_0 = \sum_{k=1}^N \hat{\beta}_k \psi_k - \hat{Q}_0 \quad (4.255)$$

with the coefficients

$$\left. \begin{aligned} \hat{\beta}_k &= \beta_k (1 + \mathbf{a}_0 \cdot (\mathbf{r}_k - \mathbf{r}_0)/2), \quad (k = 1, \dots, N) \\ \hat{\beta}_0 &= 1 - \frac{b_0}{4} \sum_k \beta_k d_k^2 \\ \hat{Q}_0 &= \frac{1}{4} Q_0 \sum_k \beta_k d_k^2. \end{aligned} \right\} \quad (4.256)$$

This discretization is quite simple, but its error can hardly be reduced beyond the third order. A gain of accuracy is possible if we write the PDE in the form

$$\partial_x(p^2(x, y)V_{|x}) + \partial_y(p^2(x, y)V_{|y}) = p^2(x, y)Q(x, y) \quad (4.257)$$

with $p(x, y) > 0$. This coefficient can then be removed from the left-hand side by a transform

$$V(x, y) = \Psi(x, y)/p(x, y), \quad (4.258a)$$

leading to

$$\Delta \Psi = pQ + \eta \Psi \quad (4.258b)$$

with

$$\eta(x, y) := p^{-1} \Delta p(x, y). \quad (4.258c)$$

This latter coefficient can be determined either by analytical differentiation or by application of Eq. (4.251). Now Eqs. (4.247) and (4.248) are applied to this PDE, in which $S(x, y)$ is *not* approximated by a constant. Thereafter the transform (4.258a) is reversed. The result of these calculations is then

$$\sum_{k=1}^N \beta_k p_k [(1 - q_k \eta_k)(V_k - V_0) - q_k Q_k] = p_0 q_0 Q_0, \quad (4.259)$$

which satisfies the necessary condition that for $Q \equiv 0$ the function $V = V_0 = \text{const.}$ must be discretized exactly.

In the case of the cylindrical Poisson equation $\Delta_\alpha V(z, r) = Q(z, r)$, frequently appearing in practical applications, we have to identify x with the axial coordinate z and y with the radial coordinate r . Then the PDE (4.254) is obtained with $\mathbf{a} = (0, \alpha/r)$, $b \equiv 0$. The local coordinates are now $u_k = z_k - z_0$, $v_k = r_k - r_0$. From Eqs. (4.256) we obtain the coefficients $b_0 = 1$ and $\beta_k = 1 + \alpha v_k / 2r_0$ ($k \neq 0$). This approximation can be improved by replacing r_0 by a mean value $\bar{r}$, chosen in such a way that the polynomial P_2 in Eqs. (2.66) is discretized exactly. Using Eqs. (4.231) to (4.236), we find

$$\bar{r} = r_0 + (2 - \alpha)\bar{v}/4 \quad (4.260a)$$

and then

$$\psi_0 = \sum_{k=1}^N \beta_k \left(1 + \frac{\alpha v_k}{2\bar{r}}\right) \psi_k - \frac{1}{4} Q_0 \sum_{k=1}^N \beta_k d_k^2. \quad (4.260b)$$

Equation (4.260a) holds also for $\alpha = -1$, as can be verified with the function $\Psi = r^2$.

To cast the PDE $\Delta_\alpha V = Q$ in the form of Eq. (4.257), we have to set

$$\left. \begin{aligned} p(x, y) &\rightarrow p(r) = (r/r_0)^{\alpha/2} \\ \eta(x, y) &\rightarrow \eta(r) = \alpha(\alpha - 2)/(4r^2) \end{aligned} \right\} \quad (4.261)$$

and to introduce the sampling values of these functions into Eq. (4.259) with q -coefficients given by Eqs. (4.248). This gives a more complicated formula, but this is distinctly more accurate if $N \geq 6$. One of the reasons lies in the fact that the variation of the sources is now taken into account, whereas Eq. (4.260b) considers only the central value Q_0 . Moreover, a *material coefficient* $\varepsilon(z, r)$ or $\nu(z, r)$ can be incorporated in $p^2(z, r)$, which gives a real extension of the class of possible applications. For example, it should be easily possible to apply this method to the calculation of magnetic lenses ($\alpha = -1$), even with nonlinear reductivity $\nu(z, r)$.

Summary

The method requires only the solution of a linear system of Eq. (4.214) for the harmonic powers (4.228) as trial functions. Thereafter, all the mesh coefficients, even for general self-adjoint PDEs, can be determined successively by forming mean values (Eqs. (4.231) to (4.236) and appropriate transformations (Eqs. (4.258)). There is no need to repeat the solutions of the linear systems of equations, if another type of PDE is chosen. As a result of this property the method can work very efficiently.

4.5.7 Discretization Errors

For reasons of conciseness we consider here the approximation (4.250b); the solutions of other PDEs have qualitatively the same behavior. The function $\Psi(x, y)$ is the superposition of the *exact* solution $\Psi^{(e)}(x, y)$ of the given boundary value problem and a *deviation* $\Psi^{(d)}(x, y)$ caused by the discretization. Only the sum of both can be obtained, and so we now write

$$\sum_k \beta_k (\psi_k^{(e)} + \psi_k^{(d)}) - \psi_0^{(e)} - \psi_0^{(d)} = \frac{1}{4} S_0 \sum_k \beta_k d_k^2. \quad (4.262)$$

The exact function satisfies a condition

$$\sum_k \beta_k \psi_k^{(e)} - \psi_0^{(e)} = \frac{1}{4} S_0 \sum_k \beta_k d_k^2 + E(x_0, y_0), \quad (4.263)$$

$E(x_0, y_0)$ being the *local* discretization error that can be obtained from a systematic Taylor series expansion and which must be ignored. Subtracting Eq. (4.263) from (4.262) we find

$$\sum_k \beta_k \psi_k^{(d)} - \psi_0^{(d)} = -E(x_0, y_0). \quad (4.264)$$

This is equivalent to the solution of a Poisson equation with a source $s_d(x, y)$ having the value

$$s_d(x_0, y_0) = 4E(x_0, y_0) \left(\sum_k \beta_k d_k^2 \right)^{-1} \quad (4.265)$$

at position (x_0, y_0) . The function $\Psi^{(d)}$, the *global* discretization error, can now be understood as the solution of Poisson's equation $\Delta \Psi_d = -s_d$ with vanishing boundary values.

Unfortunately, there are no exact solutions of this problem; we therefore discuss here a very simple example, which yields a rough estimate. Let $\Psi(x, y)$ be a polynomial of fourth order, to be calculated by the ordinary five-point formula (4.103) with $\alpha = 0$. Then the error $E(x, y)$ is well known to be

$$E(x, y) = h^4 (\partial^4 \psi / \partial x^4 + \partial^4 \psi / \partial y^4) / 48 = \text{const.}, \quad (4.266)$$

where $h = H$ being the mesh width and s_d is then given by

$$s_d(x, y) = 4E/h^2 = \text{const.} \quad (4.267)$$

For a circle of radius R , the appropriate solution is then

$$\psi_d(x, y) = s_d(R^2 - x^2 - y^2)/4 = E(R^2 - x^2 - y^2)/h^2. \quad (4.268)$$

Its maximum is then $\Psi_d(0, 0) = E(R/h)^2$, which is considerably larger than E for $R \gg h$. Quite generally the power of h in Ψ_d is *two orders less* than the corresponding power in E , so that nine-point formula are favorable from this point of view.

4.6 SUBDIVISION OF MESHES

Limitations on memory might make it impossible or at least unfavorable to cover a domain with a uniform mesh even if the transforms, described in Section 4.1, are used. Such a situation can arise if the boundary has a large extent, and also if there are short sides with sharp corners as shown in Fig. 4.18. It is then better to use a coarse mesh in domains with little inhomogeneity of the field strength and a narrow one near the sharp edges or corners.

This raises the problem of the coupling between the different meshes. So far as the boundary points of the coarse mesh along the lines B_2 are concerned this

points A , B , C , and D . The first result is hence

$$\frac{h}{2}\varepsilon_A\phi_{A|u} + h\varepsilon_B\phi_{B|v} - \frac{h}{2}\varepsilon_C\phi_{C|u} - h\varepsilon_D\phi_{D|v} = \frac{h^2}{2}S_0. \quad (4.270)$$

Now the partial derivatives are approximated by central finite differences, as usual, and the equation obtained is then solved for ϕ_0 , the result being

$$\begin{aligned} \phi_0 = \frac{1}{8(\varepsilon_B + \varepsilon_D)} \{ & \varepsilon_A(\phi_1 + \phi_6) + \varepsilon_C(\phi_3 + \phi_4) - 2h^2S_0 \\ & + (8\varepsilon_B - \varepsilon_A - \varepsilon_C)\phi_2 + (8\varepsilon_D - \varepsilon_A - \varepsilon_C)\phi_5 \} + O(h^4). \end{aligned} \quad (4.271)$$

If the source term S_0 depends on ϕ in the form $S = D\phi + Q$, this should be approximated by

$$S_0 = D_0(\phi_2 + \phi_5)/2 + Q_0 + O(h^2), \quad (4.272)$$

so that the right-hand side becomes completely independent of ϕ_0 ; the potentials $\phi_1, \dots, \phi_6$ belong entirely to the *coarse* mesh.

An analogous formula holds for the case in which the longer side of the rectangle has the v -direction; this is simply a rotation of the configuration by 90 degrees. Moreover it is straightforward to derive corresponding formulas for the general case $h_u \neq h_v$ and *anisotropic* coefficients $A(u, v) \neq B(u, v)$. However, the orthogonality ($C(u, v) \equiv 0$) is essentially necessary.

The subdivision of the mesh should be performed *gradually*, if halving the mesh size in one step is not sufficient. The overlapping must be at least *one* full row or column; the use of two overlapping rows or columns improves the stability of the solution and its accuracy and is hence recommended.

There are situations in which the preceding method is not sufficient, for example in front of a cathode with thermionic electron emission. Quite often the cloud of space charge in front of it varies so strongly in a narrow sheet that it is very difficult to model this in a square-shaped mesh. This case is shown in Fig. 4.20: without loss of generality the cathode can be chosen as the surface $u = 0$. The strong variation of the space charge may be confined to a few columns of the coarse mesh, but the finer one must have an overlap of one column more because of the coupling. This case has been studied by Kumar and Kasper [26, 27]; here we shall deal with it in a more general way. In any case, a new mesh formula is necessary because five-point formulas are not accurate enough, and the familiar nine-point formulas require $H/h \leq 2$, if strongly negative coefficients are to be avoided.

It is now necessary to assume that the coefficient ε does not depend on u , and that all functions vary only slowly with v because otherwise the mesh

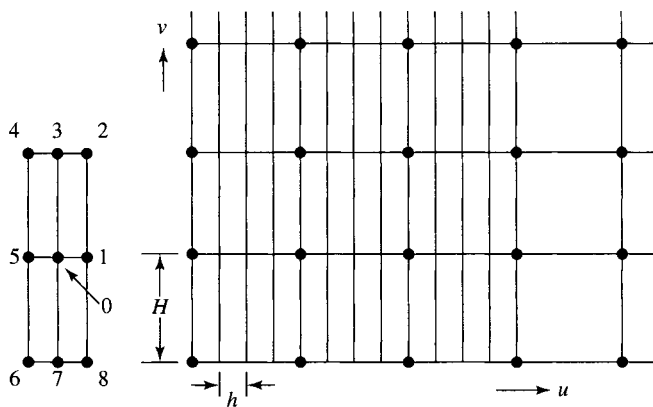


FIGURE 4.20 Subdivision of the mesh in only one direction.

size H must also be diminished. Thus, we can write the PDE in the form

$$\phi_{uu} = -\frac{1}{\varepsilon(v)} \partial_v(\varepsilon \phi_v) - g(u, v) =: -Q(u, v). \quad (4.273)$$

We now consider the configuration shown in Fig. 4.20 and adopt a simplified labelling of the points. Provided that the function $Q(u, v)$ is given, then the mesh formula for the potential ϕ_0 is given straightforward by the Numerov formula (4.124), now rewritten as

$$\phi_0 = (\phi_1 + \phi_5)/2 + h^2(Q_1 + 10Q_0 + Q_5)/24 + O(h^6). \quad (4.274)$$

The evaluation of the function $Q(u, v)$ requires the coefficients

$$\varepsilon_+ := \varepsilon(v_0 + H/2), \quad \varepsilon_- := \varepsilon(v_0 - H/2), \quad \varepsilon_0 = (\varepsilon_+ + \varepsilon_-)/2. \quad (4.275)$$

We have then, for example, the approximation

$$Q_1 = \frac{2}{(\varepsilon_+ + \varepsilon_-)H^2} [\varepsilon_+(\phi_2 - \phi_1) + \varepsilon_-(\phi_8 - \phi_1)] - g_1 \quad (4.276)$$

and similarly for Q_0 and Q_5 . On introducing these expressions into Eq. (4.274), it becomes obvious that the coefficients

$$\lambda_{\pm} = h^2 \varepsilon_{\pm} [12(\varepsilon_+ + \varepsilon_-)H^2]^{-1} \quad (4.277)$$

appear in the final formulas, which takes the concise form

$$[1 + 10(\lambda_+ + \lambda_-)]\phi_0 = \lambda_+(\phi_2 + 10\phi_3 + \phi_4) + \lambda_-(\phi_6 + 10\phi_7 + \phi_8) \\ + (0.5 - \lambda_+ - \lambda_-)(\phi_1 + \phi_5) + h^2(g_1 + 10g_0 + g_5)/24. \quad (4.278)$$

This has an error of order h^6 in the sensitive u -direction and of order H^2h^2 in the insensitive v -direction, the coefficients remaining positive for any ratio $H/h > 1/\sqrt{6}$.

If the system in question is axisymmetric about the axis $v = 0$, then again a special formula is necessary. In complete analogy with earlier such considerations (see Eq. (4.53)) we may assume that $\phi(u, v)$ and $g(u, v)$ are *even* functions of v and $\varepsilon(u, v) = \varepsilon_0(u)v^\alpha$, Eq. (4.273) then specializing to

$$\phi_{|uu} + (\alpha + 1)\phi_{|vv} = -g \quad \text{at } v = 0 \quad (4.279)$$

with the approximation

$$\phi_{|vv}(u, 0) = 2(\phi(u, H) - \phi(u, 0))/H^2 + O(H^4). \quad (4.280)$$

Combining this with Eq. (4.274), we finally arrive at

$$\phi_0(2 + 10\lambda_\alpha) = (1 - \lambda_\alpha)(\phi_1 + \phi_5) + \lambda_\alpha(\phi_2 + 10\phi_3 + \phi_4) \\ + h^2(g_1 + 10g_0 + g_5)/12 \quad (4.281a)$$

with

$$\lambda_\alpha = h^2(1 + \alpha)/(6H^2). \quad (4.281b)$$

This method can be generalized to include the vicinity of bent cathode surfaces. Then the curvature parameter β for the discretization in the normal direction has to be adjusted appropriately, so that the more general formulas (4.176) to (4.178) can be used instead of Eq. (4.274), the notation being adopted correctly. Equation (4.274) is the special case for $\beta = 0$; generally we should choose $\beta > 0$ for convex surfaces and $\beta < 0$ for concave ones. The subsequent derivation of the mesh formula is straightforward.

4.7 CONCLUDING REMARKS

In this chapter the method of finite differences has been outlined in some detail as it is potentially very favorable for field calculations in charged particle optics. The variational method, described in Section 4.3, already comes very close to corresponding techniques in the FEM and can easily be combined

with them. The full power of the FDM becomes obvious in combination with semianalytical methods such as series expansions and the method of boundary elements, by which the obstacles arising from artificially limited field domains and complicated boundaries can be overcome. This will be the topic of later chapters.

Some problems have not been discussed so far. One of them is the solution of truly irreducible *three*-dimensional problems. Apart from the dramatically increased demand for memory and computation time, there is nothing intrinsically new about this. The most frequently appearing PDEs (3.37) or (3.39) are discretized by the seven-point formula (3.43), and it is usually too complicated to go beyond this approximation.

Another problem, not discussed so far, is the numerical solution of the large system of equations obtained by discretization. This will be dealt with in Section 5.6, after all the methods of setting up such systems of equations have been presented.

REFERENCES

1. Forsythe, G. E. and Wasow, W. R. (1960). *Finite Difference Methods for Partial Differential Equations*, New York: Wiley.
2. Mitchell, A. R. (1969). *Computational Methods in Partial Differential Equations*, London: Wiley.
3. Smith, G. D. (1978). *Numerical Solution of Partial Differential Equations: Finite Difference Methods*, Oxford: Clarendon Press.
4. Strikwerda, J. C. (1989). *Finite Difference Schemes and Partial Differential Equations*, New York: Chapman & Hall.
5. Ganzha, V. G. and Vorozhtsov, E. V. (1996). *Computer-Aided Analysis of Difference Schemes for Partial Differential Equations*, New York: Wiley.
6. Weber, C. (1967). Numerical solution of Laplace's and Poisson's equations and the calculation of electron trajectories and electron beams. In *Focusing of Charged Particles*, Volume 1, ed., A. Septier, pp. 45–99, London & New York: Academic Press.
7. Bonjour, P. (1980). Numerical methods for computing electrostatic and magnetostatic fields, *Advances in Electronics and Electron Physics*, Suppl. 13A: 1–44.
8. Kasper, E. (1982). Magnetic field calculation and the determination of electron trajectories, In *Topics in Current Physics*, Volume 13, P. W. Hawkes, ed., pp. 57–118, Berlin, Heidelberg, New York: Springer.
9. Hawkes, P. W. and Kasper, E. (1989). *Principles of Electron Optics*, Volume 1, Chapter 11, London & New York: Academic Press.
10. Munro, E. (1973). Computer-aided design of electron lenses by the finite element method, In *Image Processing and Computer-Aided Design in Electron Physics*, P. W. Hawkes, ed., pp. 284–323, London: Academic Press.
11. Franzen, N. (1984). Computer programs for analyzing certain classes of 3-D electrostatic fields with two planes of symmetry, in *Electron Optical Systems*, J. J. Hren et al., eds., pp. 115–126, Scanning Electron Microscopy, Chicago.
12. Rouse, J. and Munro, E. (1990). Three-dimensional modelling of various aspects of the scanning electron microscope, *Nucl. Instrum. Meth. A* **298**: 78–84.

13. Durand, E. (1966). *Electrostatique, Tome 2*, Paris: Masson.
14. Thomae, H. and Becker, R. (1990). Reduction of discretization errors in the numerical simulation of axisymmetric electrostatic potentials, *Nucl. Instrum. Meth. A* **298**: 407–414.
15. Kasper, E. (1976). On the numerical calculation of static multipole fields, *Optik* **46**: 271–286.
16. Kasper, E. (1984a). Improvements of methods for electron optical field calculations, *Optik* **68**: 341–362.
17. Kasper, E. (1984b). Recent developments in numerical electron optics. In *Electron Optical Systems*, J. J. Hren et al., eds., pp. 63–73, Chicago: Scanning Electron Microscopy.
18. Kasper, E. (1990). Advanced nine-point formulae for the discretization of Poisson's equation, *Nucl. Instrum. Meth. A* **298**: 295.
19. Kasper, E. and Lenz, F. (1980). Numerical methods in geometrical electron optics. In *Electron Microscopy*, P. Brederoo and G. Boom, eds., Volume 1, pp. 10–15, Amsterdam: North Holland.
20. Killes, P. (1985). Solution of Dirichlet problems using a hybrid finite differences and integral-equation method applied to electron guns, *Optik* **70**: 64–71.
21. Kang, N. K., Orloff, J., Swanson, L. W. and Tuggle, D. (1981). An improved method for numerical analysis of point electron and ion source optics, *J. Vac. Sci. Technol.* **19**: 1077–1081.
22. Kang, N. K., Tuggle, D. and Swanson, L. W. (1983). A numerical analysis of the electric field and of space charge for a field electron source, *Optik* **63**: 313–331.
23. Swanson, L. W. (1984). Field emission source optics. In *Electron Optical Systems*, J. J. Hren et al., eds., pp. 137–147, Chicago: Scanning Electron Microscopy.
24. Lenz, F. (1973). Computer-Aided Design of Electron Optical Systems. In *Image Processing and Computer-Aided Design in Electron Physics*, P. W. Hawkes, ed., pp. 274–282, London & New York: Academic Press.
25. Morizumi, Y. (1972). Computer-aided design of an axially symmetrical magnetic circuit and its application to electron-beam devices, *IEEE Trans. Electron Dev.* **ED-19**: 782–797.
26. Kumar, L. and Kasper, E. (1985). On the numerical design of electron guns, *Optik* **72**: 23–30.
27. Kumar, L. (1990). Computer simulation of electron flow in linear-beam microwave tubes, *Nucl. Instrum. Meth. A* **298**: 332–343.

This Page Intentionally Left Blank

CHAPTER V

The Finite-Element Method (FEM)

The finite-element method (FEM) is the most general method for solving differential equations under given boundary conditions, and it is therefore in widespread use in the various fields of science and engineering. There exists an extensive literature on the applications of the FEM to various classes of mathematical problems; this is so huge, indeed, that we can here refer only to some standard works [1–8], where the interested reader can find more details.

The first application of the FEM to field calculations in electron optics was made by E. Munro [9–12], who calculated a magnetic lens with saturation effects. This problem was later reconsidered by many other scientists, notably B. Lencova [13–20] and T. Mulvey [21–24] and their co-workers, whose work led to progressively improved versions of the FEM.

In this chapter we cannot go into all the details but must confine ourselves to the principles. The reason why the FEM is not always the best choice lies in its intrinsic difficulties: it is very hard to achieve continuity of derivatives normal to the mesh lines. This makes ray tracing highly complicated. It therefore makes sense to couple the FEM with other methods of field calculation or to carry out at least some postprocessing of the results in the domains traversed by the particle rays.

5.1 GENERATION OF MESHES

The finite-element method consists essentially in approximating functions in a large number of small elements, which are not regular but in general triangles or quadrangles. This concept provides a great flexibility in fitting such meshes to arbitrary boundaries. On the other hand, some complications arise. Because it is practically impossible to define thousands of irregular elements “by hand,” some automatic or at least semiautomatic procedures for their generation are necessary. In view of the great variety of different situations, which may occur in practice, there is virtually no algorithm that is satisfactory in every respect. We shall therefore briefly sketch a variety of methods. This topic is also of great importance for the three-dimensional version of the boundary element method (Chapter VI), where the problem of discretizing general surfaces in space arises. To avoid unnecessary repetition, we shall consider these cases together.

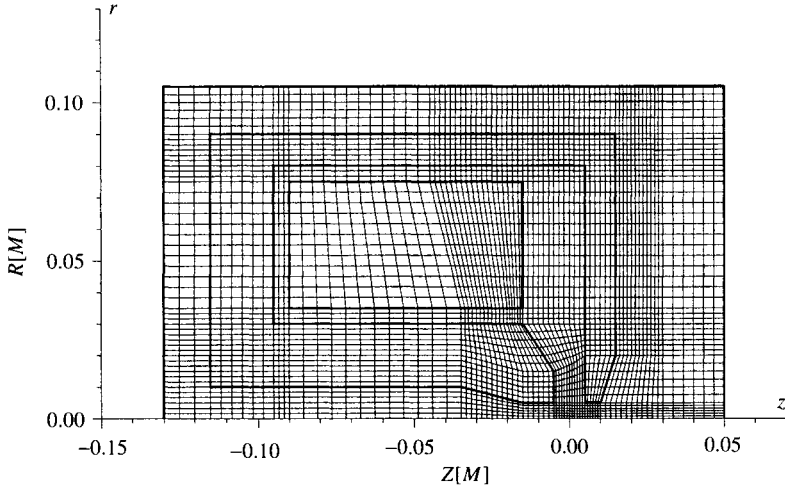


FIGURE 5.1 Semi-axial section through a conventional magnetic lens with discretization by a quadrilateral mesh. This can be converted into a triangular one by bisecting each element along one of its diagonals. Calculations by E. Munro [10] and K. Zeh [25]; for the results, see Figs. 5.24 and 5.25.

The simplest possible method is demonstrated in Fig. 5.1. This was applied by E. Munro [10] to the cross-section through a magnetic lens and then become a standard technique. A coarse mesh of quadrilateral structure, adapted to the polygonal structure of the boundaries, is defined by hand and then uniformly subdivided in both directions. This method certainly has the great advantage that it allows a rigorous double-index numbering of all nodes, as in the FDM, and if necessary, a subdivision into triangular meshes as shown in Fig. 5.2. On the other hand, this method has the essential drawback that abrupt discontinuities of the mesh size appear at locations where they are unmotivated, simply because the lines must be continued through the whole grid. Some smoothing is possible by using graduated meshes, but the unfavorable structures cannot be eliminated satisfactorily. A rough rule for a favorable discretization states that the ratio of side lengths should satisfy $0.5 \leq S_1/S_2 \leq 2$ and that obtuse angles should never appear in triangular elements.

A method that generates almost regular triangular meshes is shown in Fig. 5.3. The whole domain is covered with a regular triangular mesh. Then the mesh points that do not fit the boundary are shifted on to it, so that, after the regularization, either regular triangles are obtained ($G = 1$) or right-angled ones ($G = 2$), see Fig. 5.3. This choice should be made in such a way that the necessary shift is small, to minimize the inevitable distortion. The boundary points thus obtained are then kept fixed and the inner ones are

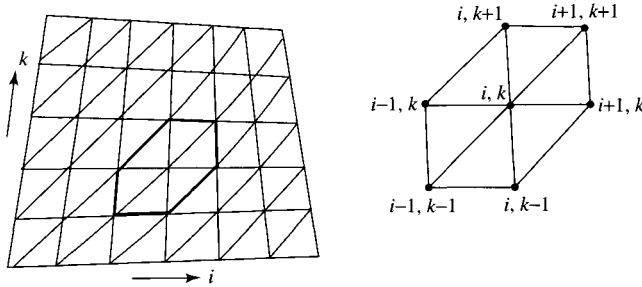


FIGURE 5.2 A trigonal mesh obtained by bisection of quadrangles and double-indexing in seven-point configurations.

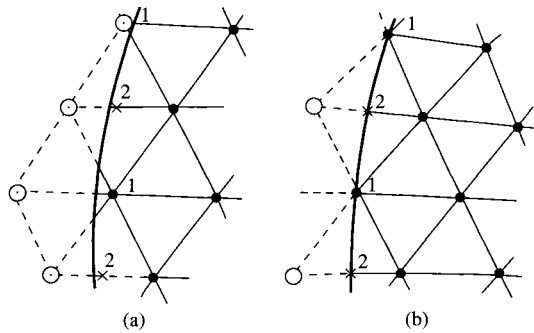


FIGURE 5.3 Regularization of hexagonal meshes: (a) irregular situation; (b) matched meshes. Either nodes, marked by dots ($G = 1$), or midpoints, marked by crosses ($G = 2$) are shifted exactly onto the boundary, whichever gives the least deformation. If midpoints are involved, they must alternate with nodes for topological reasons.

moved iteratively according to the procedure

$$x_0 = \sum_{n=1}^6 G_n x_n / \sum_{n=1}^6 G_n, \quad (5.1)$$

$$y_0 = \sum_{n=1}^6 G_n y_n / \sum_{n=1}^6 G_n$$

in a simplified sequential numbering. This is related to the complete two-dimensional indexing (i, k) by the scheme (see Fig. 5.4)

0	1	2	3	4	5	6	
i	$i+1$	$i+m$	$i-1+m$	$i-1$	$i-1+m$	$i+1$	(5.2)
k	k	$k+1$	$k+1$	k	$k-1$	$k-1$	

with $m = \text{mod}(k, 2) = 0$ or 1 .

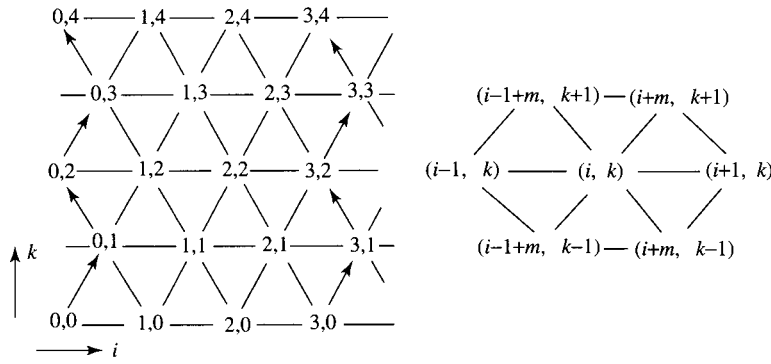


FIGURE 5.4 The alternating indexing in hexagonal meshes with $m = \text{mod}(k, 2)$.

This is an anticlockwise numbering of the six closest neighbors of the point (i, k) . Value $G_n = 2$ can appear only for nodes located on boundaries or symmetry axes; they cannot be avoided in general because without them it is not possible to discretize orthogonal domains in a reasonable manner.

This method is quite analogous to the one outlined in Section 4.3.4. Some flexibility is given by the possibility of applying this method not to the original coordinates (x, y) or (z, r) but to transformed coordinates (u, v) , but generally it is too restrictive. An example in which irregular meshes must be used is shown in Fig. 5.5. For physical reasons, the mesh sizes must be small near the edge of the cylindrical surface, and this requires a different approach.

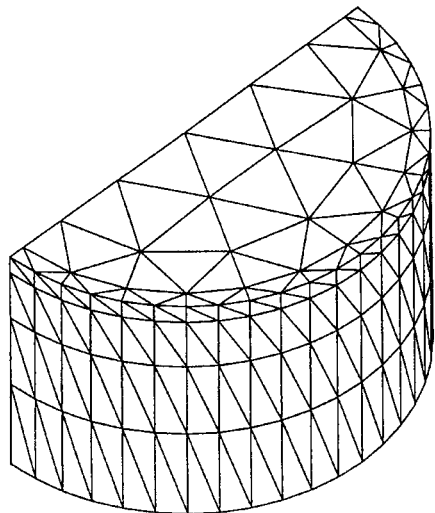


FIGURE 5.5 Triangulation of a part of a cylindrical surface. The elements must be small near the circumference of the circle (from M Str  r [26]).

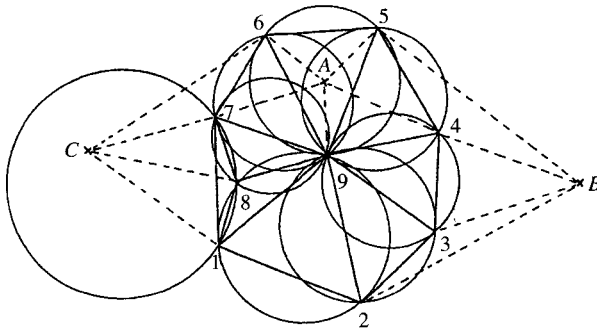


FIGURE 5.6 Different geometrical situations (after M. Eupper [29]).

The discretization can be obtained by some explicit method, which is adapted to the particular geometry. At the other extreme are completely automatic algorithms for triangulation of arbitrary domains. Such methods have been published by Hermeline [27], Thacker [28], and Eupper [29], for instance. Here we shall follow Eupper's method. The basic algorithm is shown in Fig. 5.6.

Let us assume that nine points and nine allowed nonoverlapping triangles with their associated circles have already been found, and we now wish to include the next point. Then, depending on the location of this point, different strategies are necessary.

Situation A: Inner Point

All those triangles are removed for which the point is located in the corresponding circle; these are the triangles (9, 4, 5), (9, 5, 6), and (9, 6, 7). Instead of them, five new triangles are generated by joining point A to points 4, 5, 6, 7, and 9. Thereafter the corresponding circles must be determined.

Situation B: Completely External Point

This is joined to all those points on the convex shell that separate the remainder from the point B. This means, for example, that the straight line through points 3 and 4 separates the whole of the old polygon from position B. Now the triangles (B, 2, 3), (B, 3, 4), and (B, 4, 5) are generated.

Situation C: Partly External Point

This point is located completely outside the polygon but still in one of the circles. Now the corresponding triangle (1, 8, 7) is removed, and then point C joined with points 1, 8, 7, and 6, thereby generating three new triangles.

The whole procedure is demonstrated for the domain D shown in Fig. 5.7. The outer domain must be oriented positively and inner ones in the opposite

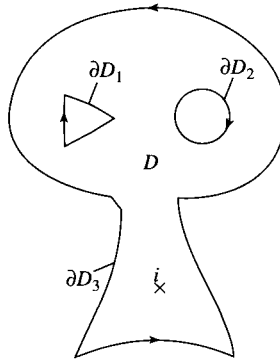


FIGURE 5.7 Definition of a domain D with three boundaries, ∂D_1 , ∂D_2 , and ∂D_3 , for purposes of demonstration (after M. Eupper [29]).

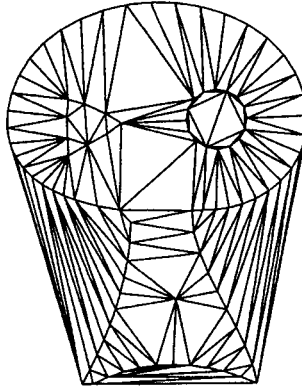


FIGURE 5.8 The triangulation of the boundaries, resulting in the convex shell [29].

sense. At position (i) , a very fine discretization will be obtained. The positions $\mathbf{u}_1, \dots, \mathbf{u}_n$ at the *boundaries* and $\mathbf{u}_{n+1}$ at (i) must be prescribed reasonably, and then the above algorithm starts with $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$, $(\mathbf{u}_1, \dots, \mathbf{u}_4)$, and so on. The result is shown in Fig. 5.8: it supplies the convex shell.

In the next step, all *external* triangles are removed; this can be done by means of their orientation, which is positive for inner triangles and negative for outer ones. The next result is shown in Fig. 5.9: The triangulation is consistent but still impracticable. Hence, a successive refinement is now necessary.

To achieve this, a reasonable *limitation* function $d(\mathbf{u}) > 0$ must be defined. A triangle will be acceptable if the condition

$$s_i := |\mathbf{u}_j - \mathbf{u}_k| \leq \frac{1}{2}(d(\mathbf{u}_j) + d(\mathbf{u}_k)), \quad i, j, k \text{ cyclic} \quad (5.3)$$

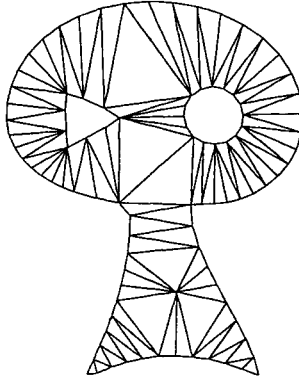


FIGURE 5.9 Removal of all external triangles, that have negative orientation [29].

is satisfied simultaneously for all three side lengths s_i . Initially this will not be the case.

Then a new internal point, defined as a weighted centroid

$$\mathbf{u}_s = \frac{\mathbf{u}_i(w_j + w_k) + \mathbf{u}_j(w_k + w_i) + \mathbf{u}_k(w_i + w_j)}{2(w_i + w_j + w_k)} \quad (5.4)$$

is inserted into the corresponding triangle. Thereafter the process of elimination and creation of triangles is carried out with respect to this new point, and the weights are then recalculated. The different publications concerning this topic use slightly different definitions of the weights. Eupper defined

$$w_i = \max \{0, 2s_i/(d_j + d_k) - 1\}, \quad (i, j, k \text{ cyclic}) \quad (5.5)$$

with $d_j := d(\mathbf{u}_j)$, $d_k = d(\mathbf{u}_k)$ for abbreviation. It is essential that this limitation function $d(\mathbf{u})$ be chosen reasonably. In the case of a bent surface this must be some fraction of the least radius of local curvature; on plane surfaces $d(\mathbf{u})$ is to be confined by a tolerable upper limit, and in the example of Fig. 5.5 it must be sufficiently small near the edge line of the cylinder. In the case of Fig. 5.6, the function

$$d(\mathbf{u}) = \min (12, 1 + |\mathbf{u} - \mathbf{u}_i|/3) \quad (5.6)$$

was chosen.

When the whole process is finished, the shape of the obtained configuration can be improved further by *regularization*. This means here that all boundary points and all invariant positions (the point $\mathbf{u}_i$ in Fig. 5.6) are kept fixed, and all other ones iteratively replaced by the arithmetic centroid of the surrounding



FIGURE 5.10 The result of the triangulation after the regularization (after M. Eupper [29]).

ones, found from the table of contiguity. For the example chosen, the result obtained is shown in Fig. 5.10.

Although not absolutely necessary, it is convenient for the solution of boundary value problems by the FEM to renumber all points with *known* potential in such a manner that they are placed at the *end* of the listing; otherwise the system of equations would have initial labels > 1 , which is feasible but inconvenient.

The mathematical handling of such general patterns requires two kinds of coherence tables. The first one is the *table of nodes*. It contains in sequentially ascending order the number n , the type (internal, axial, or boundary) the coordinates u_n , the list of all triangles that have this node in common, and later all physical properties (potential and potential gradient). This is to be completed by the second one, a *table of elements*, which counts sequentially all triangular elements and contains for each of them the labels of the three nodes and physical properties associated with the elements. By combining these two tables, we are now in a position to handle all local informations efficiently.

5.2 DISCRETIZATION OF THE VARIATIONAL PRINCIPLE

In mechanical and electrical engineering, there is a great variety of methods for the formulation of discrete systems of equations that are approximate solutions of partial differential equations. Among these, two methods are especially well suitable for field calculations in charged particle optics: the variational method and the Galerkin method. The latter will be dealt with in the context of the boundary-element method, whereas the former is discussed in this section. For comparison with the FDM, we shall consider here the functional F of

Eq. (4.11) with all the associated relations; this is the most general case in two dimensions.

We now start with a general series expansions of the potential $\phi(u, v)$ in terms of suitable linearly independent trial functions $N_i(u, v)$. These are to be defined *locally* that is, in a small subset of elements, called the *carrier*, outside which they vanish identically. Examples of such functions will follow in later chapters. The series expansion can be cast in the form

$$\phi(u, v) = \sum_{i=1}^N V_i N_i(u, v), \quad (5.7)$$

N being the total number of degrees of freedom and V_i the initially unknown expansion coefficients. Usually these are the values of the potential at the nodes, but this restriction is not absolute, as derivatives may also be used; the only requirement is that the approximation achieves the highest possible accuracy.

The following procedure is straightforward: We differentiate Eq. (5.7) with respect to u and v and introduce the derivatives into Eq. (4.11); because N is finite, all differentiations and integrations may be exchanged with the sequence of summations. A short elementary calculation leads to

$$F = \frac{1}{2} \sum_{j=1}^N \sum_{k=1}^N \hat{L}_{jk} V_j V_k + F_w(V_1, \dots, V_N), \quad (5.8)$$

with the symmetric matrix elements

$$\hat{L}_{jk} = \iint_{\mathfrak{A}} [A N_{j|u} N_{k|u} + B N_{j|v} N_{k|v} + C(N_{j|u} N_{k|v} + N_{j|v} N_{k|u})] \, du \, dv, \quad (5.9)$$

and the driving term

$$F_w = \iint_{\mathfrak{A}} W \left(u, v, \sum_k V_k N_k(u, v) \right) \, du \, dv, \quad (5.10)$$

$\mathfrak{A}$ denoting the total area of integration. The minimization of the functional requires solution of the system of equations

$$\partial F / \partial V_i = \sum_k \hat{L}_{ik} V_k + \partial F_w / \partial V_i = 0, \quad (i = 1, \dots, N). \quad (5.11)$$

Provided that this system is nonsingular and has a unique solution, the resulting coefficients can be introduced into Eq. (5.7), whereupon the problem is solved.

In spite of the striking simplicity of this general concept, there are numerous problems in practice. The most serious one is the choice of the “best” set of trial functions $N_i(u, v)$; this will be the topic of Section 5.3. Before we come to this subject, we shall briefly outline here the most important special cases.

The general form (4.10) of the Lagrangian usually arises from the coordinate transforms (4.2), which were introduced to match *regular* meshes to given boundaries. If, however, completely irregular meshes are in use, such a transform is not absolutely necessary, though it might still be useful. In cartesian coordinates (x, y) or cylindrical coordinates (z, r) the material properties are usually isotropic, and the driving term is often quadratic. We therefore consider now the less general functional

$$F = \iint_{\mathfrak{A}} \left[\frac{1}{2} \varepsilon(u, v) |\nabla \phi|^2 + \frac{1}{2} D(u, v) \phi^2 + Q(u, v) \phi(u, v) \right] du dv = \min, \quad (5.12)$$

analogous to Eqs. (4.41) and (4.42) after identification of u with x and v with y . The matrix elements (5.9) then simplify to

$$\begin{aligned} \hat{L}_{jk} &= \iint_{\mathfrak{A}} \varepsilon(N_{j|u} N_{k|u} + N_{j|v} N_{k|v}) du dv \\ &= \iint_{\mathfrak{A}} \varepsilon(u, v) \nabla N_j \cdot \nabla N_k du dv. \end{aligned} \quad (5.13)$$

The driving term now becomes

$$F_w = \frac{1}{2} \sum_j \sum_k q_{jk} V_j V_k + \sum_j S_j V_j, \quad (5.14)$$

with the symmetric matrix elements

$$q_{jk} = \iint_{\mathfrak{A}} D(u, v) N_j(u, v) N_k(u, v) du dv \quad (5.15)$$

and the effective source terms

$$S_j = \iint_{\mathfrak{A}} Q(u, v) N_j(u, v) du dv. \quad (5.16)$$

It is now advantageous to define the total system matrix by

$$L_{jk} := \hat{L}_{jk} - q_{jk}, \quad (5.17)$$

whereupon the minimization condition results in a linear system of equations

$$\sum_{k=1}^N L_{jk} V_k = -S_j \quad (j = 1, \dots, N). \quad (5.18)$$

We must assume here that the matrix L with elements (5.17) is *positive* definite. This will be the case for reasonable discretizations.

In almost all applications of the FEM, trial functions are chosen that satisfy the conditions for two-dimensional Lagrange interpolations. To cast these in a convenient form, we assume that the coefficients $V_1, \dots, V_N$ are unknown potentials at internal or axial nodes, whereas $V_{N+1}, \dots, V_M$ are prescribed potentials at the nodes on the boundary. It is then favorable to complete the series expansion (5.7) by writing

$$\phi(u, v) = \sum_{i=1}^M V_i N_i(u, v), \quad (5.19)$$

which differs from Eq. (5.7) only in having the upper bound M instead of N . The interpolation conditions then imply that

$$N_i(u_k, v_k) = \delta_{i,k}, \quad (i = 1, \dots, M, k = 1, \dots, M) \quad (5.20)$$

must be satisfied. The complete functional, containing all contributions by boundary terms, becomes

$$F = \frac{1}{2} \sum_{i=1}^M \sum_{k=1}^M L_{ik} V_i V_k + \sum_{i=1}^M S_i V_i, \quad (5.21)$$

in which the Eqs. (5.13), (5.15), (5.16), and (5.17) are extended to the upper bound M . The Eq. (5.18) must then be replaced by the modified system

$$\sum_{k=1}^N L_{ik} V_k = -S_i - \sum_{j=N+1}^M L_{ij} V_j, \quad (i = 1, \dots, N), \quad (5.22)$$

from which it becomes obvious that the boundary values contribute to the inhomogeneous or driving terms. Often these are assumed to vanish when the field has no finite outer boundary but extends to infinity for example. We then speak of natural boundary conditions if the potential vanishes at infinity. This condition is approximated by vanishing boundary values at a finite, sufficiently distant boundary. In this case, Eq. (5.22) formally agrees with Eq. (5.18).

The important application of the FEM to the field calculation in magnetic round lenses will be deferred to Section 5.4.4 because this requires some special considerations.

5.3 ANALYSIS IN TRIANGULAR ELEMENTS

The practical application of the FEM in its general form requires the development of analytical methods for calculations in triangular domains. These comprise special methods for interpolation, differentiation, and integration of functions in such domains. Moreover, these techniques are also necessary for the application of the boundary element method (BEM) to systems with surfaces that are not circular (see Chapter VI).

5.3.1 General Relations and Area Coordinates

In this section we shall consider one arbitrary triangular element as the representative of all the others and must then distinguish clearly between the *global* numbering (i, j, k) and the *local* numbering $(1, 2, 3)$ of its three corners. Throughout this chapter we shall assume positive orientation and denote cyclic permutations of these three labels by “cyclic” or “cycl”:

$$(1, 2, 3), \quad (2, 3, 1), \quad (3, 1, 2)$$

and correspondingly for permutations of (i, j, k) .

The general notation is demonstrated in Fig. 5.11. In the two-dimensional version of the FEM, two cartesian coordinates (x, y) of a vector $\mathbf{r}$ suffice, whereas in the BEM the full generality $\mathbf{r} = (x, y, z)$ is needed. For reasons of

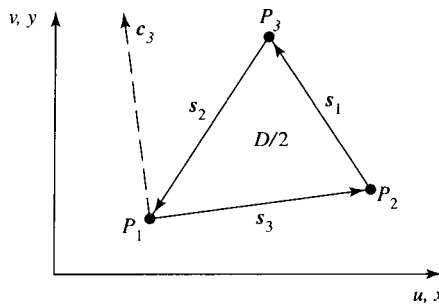


FIGURE 5.11 An arbitrary triangle with corners P_1 , P_2 , and P_3 and oriented side vectors s_1 , s_2 , and s_3 opposite to them. One of the rotated vectors, c_3 , is shown as an example. The area of the triangle is $D/2$.

conciseness we shall present explicit formulae in their two-dimensional form but general ones in the vector analytical notation. For reasons of unification, we again set $u = x$, $v = y$.

As in Section 5.1, we define the positively oriented side vectors s_n by

$$s_1 = r_3 - r_2, \quad s_2 = r_1 - r_3, \quad s_3 = r_2 - r_1. \quad (5.23)$$

From these the *double area* D and the surface normal $\mathbf{n}$ are determined by $D > 0$, $|\mathbf{n}| = 1$ and then

$$\mathbf{n}D = s_1 \times s_2 \quad \text{cycl.} \quad (5.24)$$

Moreover, it is favorable to define the *rotated* normal vectors

$$\mathbf{c}_v = \mathbf{n} \times s_v \quad (v = 1, 2, 3). \quad (5.25)$$

These are located in the plane of the triangle and are directed normal to the corresponding side vector s_v towards the corner P_v , as shown for $\mathbf{c}_3$ in Fig. 5.11. Evidently, the following relations are valid:

$$s_1 + s_2 + s_3 = \mathbf{c}_1 + \mathbf{c}_2 = \mathbf{c}_3 = 0, \quad (5.26)$$

$$|\mathbf{c}_v| = |s_v|, \quad (v = 1, 2, 3), \quad (5.27)$$

$$s_{\mu v} := \mathbf{c}_\mu \cdot \mathbf{c}_v = s_\mu \cdot s_v \quad (\mu, v = 1, 2, 3). \quad (5.28)$$

In *planar* structures the third cartesian coordinate is not used, and the area normal $\mathbf{n}$ can be eliminated; we then obtain the explicit formulas

$$D = (u_1 - u_3)(v_2 - v_3) - (u_2 - u_3)(v_1 - v_3) \quad \text{cycl}, \quad (5.29)$$

$$\mathbf{c}_1 = (v_2 - v_3, u_3 - u_2) \quad \text{cycl}. \quad (5.30)$$

The *area* coordinates, often also called the *barycentric coordinates*, are now defined as ratios of triangular areas, as shown on Fig. 5.12, here

$$\xi_1 = \frac{\text{area}(P, P_2, P_3)}{\text{area}(P_1, P_2, P_3)} \quad \text{cycl}, \quad (5.31)$$

the corresponding corners being denoted by P_1, P_2, P_3 , and P , respectively. Evidently, these coordinates are linearly dependent, as the relation

$$\xi_1 + \xi_2 + \xi_3 = 1 \quad (5.32)$$

must hold. It is hence possible to eliminate one of them, for example ξ_3 , but this is unfavorable as it would destroy the natural symmetry of the ensuing

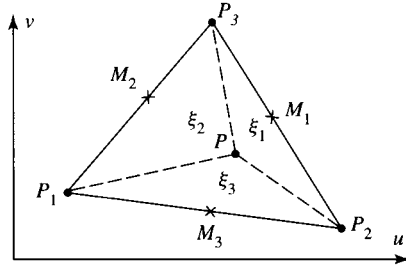


FIGURE 5.12 The area coordinates ξ_1 , ξ_2 , and ξ_3 as relative partial areas and the side midpoints M_1 , M_2 , and M_3 in cyclic notation.

formulas. The areas appearing in Eq. (5.33) can be calculated as determinants, and these are more concisely represented in vector algebraic form. For this purpose, it is favorable to introduce also the position vectors of the *side midpoints* M_ν opposite to P_ν , ($\nu = 1, 2, 3$):

$$\mathbf{m}_1 = (\mathbf{r}_2 + \mathbf{r}_3)/2, \quad \text{cycl.} \quad (5.33)$$

We then obtain the very concise formulas

$$\xi_\nu = (\mathbf{r} - \mathbf{m}_\nu) \cdot \mathbf{c}_\nu / D \quad (\nu = 1, 2, 3), \quad (5.34)$$

which is also valid in three-dimensional form. From this, it can be concluded that ξ_ν vanishes on the whole triangle side with midpoint M_ν and becomes unity at point P_ν . Other special values are

$$\xi_2 = \xi_3 = 1/2 \quad \text{at} \quad \mathbf{m}_1, \quad \text{cycl.}, \quad (5.35)$$

$$\xi_1 = \xi_2 = \xi_3 = 1/3 \quad \text{at} \quad \mathbf{r} = \mathbf{r}_c = (\mathbf{r}_1 + \mathbf{r}_2 + \mathbf{r}_3)/3, \quad (5.36)$$

that is at the centroid $\mathbf{r}_c$ of the triangle.

Another useful construct is the transform between the unit triangle and the given one. The former is shown in Fig. 5.13. Without loss of generality we can choose the normalized coordinates, ξ and η in such a manner that the three corners have the positions:

$$P_3 : (0, 0) \quad P_1 : (1, 0) \quad P_2 : (0, 1);$$

then the relations between (ξ, η) and (ξ_1, ξ_2, ξ_3) are most simply be given by

$$\xi_1 = \xi, \quad \xi_2 = \eta, \quad \xi_3 = 1 - \xi - \eta. \quad (5.37)$$

These relations are useful for the derivation of analytic formulas, especially for integration over triangles.

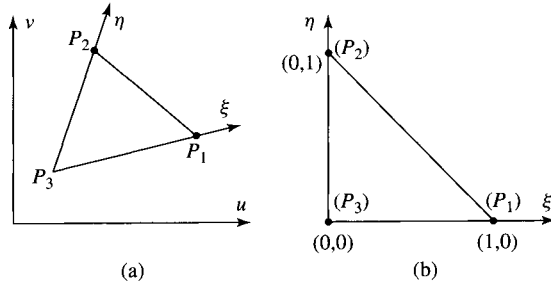


FIGURE 5.13 Two area coordinates (ξ, η) provide a linear transformation of an arbitrary triangle (a) into the unit triangle (b).

It is now a straightforward task to differentiate any differentiable function $f(\xi_1, \xi_2, \xi_3)$ with respect to u and v or even in three dimensions. From Eq. (5.34) we obtain immediately

$$\text{grad } \xi_v = \mathbf{c}_v / D, \quad (v = 1, 2, 3). \quad (5.38)$$

Thereafter the application of the chain rule results in

$$\text{grad } f = D^{-1} \sum_{v=1}^3 \mathbf{c}_v \partial f / \partial \xi_v, \quad (v = 1, 2, 3). \quad (5.39)$$

Quite frequently, the squared norm of this gradient is needed in practical calculations; this function is most favorably determined by means of Eq. (5.28), resulting in

$$|\nabla f|^2 = D^{-2} \sum_{\mu=1}^3 \sum_{v=1}^3 s_{\mu v} \partial f / \partial \xi_\mu \cdot \partial f / \partial \xi_v. \quad (5.40)$$

In this formula the coefficients $s_{\mu v}$ can be calculated entirely as scalar product of side vectors; the $\mathbf{c}$ -vectors are no longer needed.

5.3.2 Integration Over Triangular Domains

In practical calculation, it is often necessary to evaluate integrals of the form

$$I_{m,n,l} := \int_{\Delta} \xi_1^m \xi_2^n \xi_3^l da, \quad (5.41)$$

da denoting the area element, Δ the triangular domain, and m, n, l nonnegative integer exponents. This task is first carried out for the unit triangle, using

Eq. (5.37). The integral is then transformed to the triangle Δ by means of the quantity D , now appearing as Jacobian. The result is the well-known formula

$$I_{m,n,l} = D \frac{m! n! l!}{(m + n + l + 2)!}. \tag{5.42}$$

By means of this, the integration over any two-dimensional polynomial can be carried out exactly, but this requires that all its series expansion coefficients must be given explicitly. However, it is often difficult to determine these coefficients. Therefore, great efforts have been made to derive *quadrature formulas* in analogy with the well-known Gauss quadratures.

Such formulae are known for 3, 7, and 13 points with accuracy of second, fifth, and seventh order, respectively. The latter correspond to the order of two-dimensional polynomials that can still be integrated exactly. These formulas have the general form of a weighted summation:

$$\int_{\Delta} f(\xi_1, \xi_2, \xi_3) da = \frac{D}{2} \sum_{k=1}^M w_k f(p_k, q_k, 1 - p_k - q_k). \tag{5.43}$$

In this formula the numbers p_k and q_k denote special values of the area coordinates and w_k the corresponding weights. We have changed the notation here to avoid double indexing. For $M = 3$ the particular values are given in the following table (Table 5.1), which demonstrates the cyclic invariance.

TABLE 5.1
THREE-POINT QUADRATURE PARAMETERS

k	p_k	q_k	$1 - p_k - q_k$	w_k
1	1/6	1/6	2/3	1/3
2	2/3	1/6	1/6	1/3
3	1/6	2/3	1/6	1/3

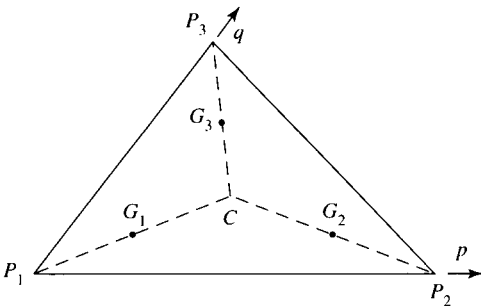


FIGURE 5.14 Positions G_1 , G_2 , and G_3 for the three-points Gauss quadrature. In this order, they are just the midpoints between the centroid C and the corresponding corners.

These points are also shown in Fig. 5.14. The data for the quadrature formulas of higher orders are presented in the appendix. They can also be found in any comprehensive textbook on the FEM.

5.3.3 Trial Functions

Just as in interpolation in one dimension or in rectangular domains, there now arises the task of interpolating arbitrary smooth functions by means of suitable sampling data. This is already obvious from Eq. (5.19). Here we are concerned with the interpolation in only one representative triangular domain; the global interpolation will be the subject of the following section.

As in the case of rectangular elements there are two classes of interpolation, the Lagrange family and the Hermite family. The *Lagrange* family that uses only function values as sampling data is depicted in Fig. 5.15. For reasons of compatibility, the original triangle must be subdivided into L^2 equal subtriangles, which is achieved by choosing L equal intervals on each side. This produces $(L + 1)(L + 2)/2$ nodes, which correspond exactly to the number of coefficients in the complete two-dimensional polynomial of degree L . For reasons of space we shall present here only the functions for $L = 1$ and $L = 2$, respectively, as higher orders are hardly ever considered.

Linear Functions ($L = 1$)

In this case, the trial functions are identical with the area coordinates; hence the interpolation takes the very concise form

$$f(\xi_1, \xi_2, \xi_3) = \sum_{v=1}^3 \xi_v f_v, \quad (5.44)$$

$$\text{grad } f = D^{-1} \sum_{v=1}^3 \mathbf{c}_v f_v = \text{const.}, \quad (5.45)$$

$$|\text{grad } f|^2 = D^{-2} \sum_{\mu} \sum_{\nu} s_{\mu\nu} f_{\mu} f_{\nu}. \quad (5.46)$$

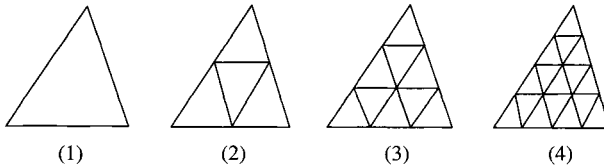


FIGURE 5.15 Sampling patterns of the Lagrange family. The number in parentheses is the corresponding order L .

Because of their striking simplicity, these formulas are the most frequently used in practice, but their disadvantages will gradually become obvious in the following sections.

Quadratic Functions ($L = 2$)

We now adopt the notation shown in Fig. 5.16. This figure demonstrates that for $L > 1$ it is possible to set up FEM relations in *curvilinear* triangular elements by means of so-called *isoparametric* element functions. This means that the area coordinates of a given position are determined in a straight (u, v) coordinate system, and the trial functions are then determined with these. Subsequently, the potential and all cartesian coordinates are interpolated in the same manner. This concept is feasible also in three dimensions, so that it is possible to interpolate on curved surfaces.

For $L = 2$ in particular, the set of trial functions is given by

$$N_k = \xi_k(2\xi_k - 1), \quad (k = 1, 2, 3), \quad (5.47a)$$

$$N_4 = 4\xi_2\xi_3, \quad N_5 = 4\xi_3\xi_1, \quad N_6 = 4\xi_1\xi_2, \quad (5.47b)$$

whereupon we have

$$\phi(u, v) = \sum_{k=1}^6 \phi_k N_k(\xi_1, \xi_2, \xi_3), \quad (5.48)$$

and correspondingly for $\mathbf{r} = (x, y, z)$,

$$\mathbf{r}(u, v) = \sum_{k=1}^6 \mathbf{r}_k N_k(\xi_1, \xi_2, \xi_3). \quad (5.49)$$

This is the curvilinear isoparametric transformation of lowest order. It would be desirable to consider higher orders $L > 2$, as the accuracy would then

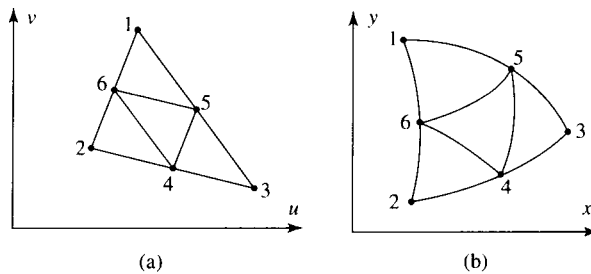


FIGURE 5.16 Isoparametric transform of second order from a straight triangle in the uv -plane to a curved one in the xy -plane with numbering of the sampling points.

increase accordingly, but the limits of memory and computation time are soon reached. Another difficulty is the fact that the transform of the area element as

$$da = J(u, v) du dv \quad (5.50a)$$

with the Jacobian

$$J(u, v) = \left| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right| \quad (5.50b)$$

requires higher orders of the Gauss quadrature formulas.

Hermite Interpolation

As in the case of interpolation techniques for rectangular domains, we can prescribe the values of the potential and of its partial derivatives $\phi|_u$ and $\phi|_v$ at the three nodes of each triangle. The corresponding *local* numbering is shown in Fig. 5.17, the derivatives being indicated by arrows. Each node thus acquires *three* degrees of freedom, making nine in all. However, the complete polynomial of third degree has 10 coefficients, so that it is necessary to consider also the potential at the centroid (u_c, v_c) as the last one.

Such a choice is feasible but inconvenient because the last coefficient requires a separate numerical technique. Attempts have therefore been made to eliminate it from the system, whereupon the polynomial becomes incomplete.

This elimination must be performed in such a way that the following conditions are satisfied:

- (i) The conditions for cubic Hermite interpolation at the three nodes are not violated.
- (ii) The interpolation is invariant with respect to cyclic permutations of the numbering.
- (iii) Any quadratic function of u and v shall still be interpolated exactly.

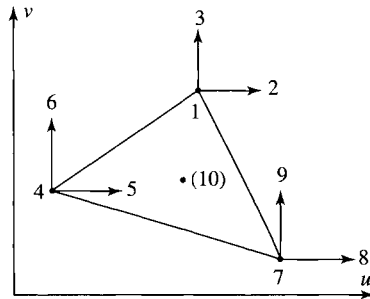


FIGURE 5.17 Numbering of the degrees of freedom in the triangular Hermite interpolation. The arrows indicate the corresponding partial derivatives. Number 10 referring to the centroid can be eliminated.

Such trial functions have been derived by Zienkiewicz [30] (see also reference [31]); they are usually written down for the *unit* element (Fig. 5.13) and then transformed accordingly. Here we shall present them in their final form, which even allows us to choose between use and elimination of the *centroid* data. This latter decision depends on the choice of the coefficients either

$$C_{0,1} = 27 \quad C_{1,1} = -7 \quad C_{2,1} = -3 \quad (5.51a)$$

or

$$C_{0,2} = 0 \quad C_{1,2} = 2 \quad C_{2,2} = 1.5, \quad (5.51b)$$

for the frequently appearing function

$$p(\xi_1, \xi_2, \xi_3) := \xi_1 \xi_2 \xi_3 \quad (5.51c)$$

which vanishes on all three sides of the triangle and has its maximum $1/27$ at the centroid. The trial functions are now given by

$$\left. \begin{aligned} N_{3i-2} &= \xi_i^2 (3 - 2\xi_i) + C_{1,j} p \\ N_{3i-1} &= \xi_i^2 (u - u_i) + C_{2,j} p(u_c - u_i) \\ N_{3i} &= \xi_i^2 (v - v_i) + C_{2,j} p(v_c - v_i) \\ N_{10} &= C_{0,j} p, \quad (i = 1, 2, 3; j = 1 \text{ or } 2). \end{aligned} \right\} \quad (5.52)$$

The requirement (ii) is evidently satisfied. Considering that not only p but also $p|_u$ and $p|_v$ vanish at the three nodes, it is easy to verify that condition (i) is also satisfied irrespective of the coefficients in Eqs. (5.51). The property (iii) requires a longer derivation, which is not given here for reasons of space; it can, however, be verified easily by application of the interpolation to quadratic or cubic polynomials.

This kind of interpolation has some favorable properties: it is very smooth in the vicinity of the nodes, but it is not possible to achieve complete continuity of the *normal* derivative along the side lines. Nevertheless, even in its reduced form, it is more accurate than the quadratic Lagrange interpolation. It is also *economic* with respect to memory. In comparison with the *first-order* finite-element method (FOFEM), the number of variables has increased by a factor of *three* and consequently the memory for the system matrix by a factor 9. However, for the *second* order (SOFEM) using Eqs. (5.47) and (5.48), the corresponding factors are 4 and 16, respectively. The memory for the system matrix can be lowered by suitable techniques for sparse matrices, but this does not reverse the general tendency.

5.3.4 Quadrilateral Elements

These are quite common in charged particle optics, as is shown in Fig. 5.1 for a magnetic lens. In such a case the skew quadrilateral meshes in the (x, y) -plane or (z, r) -plane are mapped to orthogonal ones in the (u, v) -plane. If more points than only the four corners are considered, then curvilinear elements in the (x, y) -plane can also be taken into account.

The lowest order, using only the four corners, allows only the bilinear interpolation:

$$\left. \begin{aligned} \xi &:= (u - u_1)/(u_2 - u_1), & 0 \leq \xi \leq 1 \\ \eta &:= (v - v_1)/(v_3 - v_1), & 0 \leq \eta \leq 1 \end{aligned} \right\} \quad (5.53a)$$

$$\mathbf{r}(u, v) = \mathbf{r}_1 + \xi(\mathbf{r}_2 - \mathbf{r}_1) + \eta(\mathbf{r}_3 - \mathbf{r}_1) + \xi\eta(\mathbf{r}_2 + \mathbf{r}_3 - \mathbf{r}_1 - \mathbf{r}_4), \quad (5.53b)$$

$$\phi(u, v) = \phi_1 + \xi(\phi_2 - \phi_1) + \eta(\phi_3 - \phi_1) + \xi\eta(\phi_2 + \phi_3 - \phi_1 - \phi_4). \quad (5.53c)$$

Because of its simplicity, this isoparametric transform is used quite frequently in FEM programs, although its accuracy is not better than that of Eq. (5.44).

Interpolation with eight points—without the centroid—is the simplest representative of the so-called serendipity family and is mainly used for mesh generation. The complete nine-point pattern in Fig. 5.18a allows complete Lagrange interpolations of second order in both coordinate directions. This possibility is used in the SOFEM programs of Munro et al. (1995). Another choice is the *bivariate Hermite interpolation* (see Section 3.5.2).

5.3.5 Differentiation in Systems of Triangles

We now consider a system of triangular elements with common node 0, as shown in Fig. 5.19a. The surrounding points are numbered sequentially in

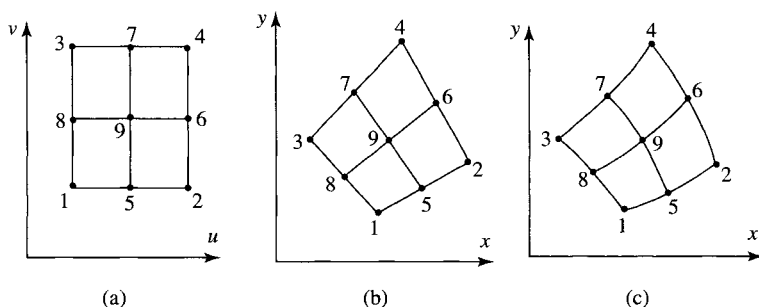


FIGURE 5.18 (a) Isoparametric transform of a regular nine-point pattern; (b) bilinear transform, points 5...9 linearly dependent on points 1...4; (c) biquadratic transform, nine degrees of freedom.

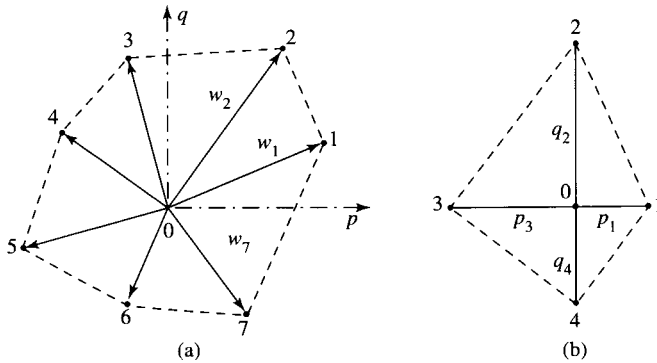


FIGURE 5.19 (a) An arbitrary configuration of neighboring points with associated w -vectors; (b) an orthogonal five-point configuration.

anticlockwise sense from 1 to M , though this is not absolutely necessary for the following procedure. We now derive a simple method for the calculation of partial derivatives of first order, the function values $\phi_i = \phi(u_i, v_i)$, $i = 0, \dots, M$, being given. For reasons of conciseness we introduce the notation

$$w = (p, q) = (u - u_0, v - v_0) \quad (5.54)$$

for the relative positions and

$$g = (\phi_{|u}, \phi_{|v})_0 = (g_u, g_v) \quad (5.55)$$

for the gradient at the central position.

As a problem with discrete data, this task has no unique answer. Hence, we look for a method that is sufficiently accurate and also simple and *stable*. A suitable requirement is then

$$F(g) = \sum_{i=1}^M G_i (\phi_i - \phi_0 - g \cdot w_i)^2 = \min \quad (5.56)$$

with appropriately chosen positive weights G_i . This can be interpreted as the choice of a tangential plane at node (0) with least square deviations at the surrounding points.

The remaining calculation is straightforward: the minimization conditions

$$\partial F / \partial g_u = 0, \quad \partial F / \partial g_v = 0, \quad (5.57)$$

lead to two linear equations for g_u and g_v .

These are found to be

$$\mathbf{A} \cdot \mathbf{g} = \mathbf{b} \equiv \sum_i G_i (\phi_i - \phi_0) \mathbf{w}_i \quad (5.58)$$

with the matrix elements

$$A_{uu} = \sum_i G_i p_i^2, \quad A_{uv} = A_{vu} = \sum_i G_i p_i q_i, \quad A_{vv} = \sum_i G_i q_i^2. \quad (5.59)$$

The solution can now be written explicitly as

$$\begin{aligned} g_u &= (b_u A_{vv} - b_v A_{uv}) / \det \mathbf{A}, \\ g_v &= (b_v A_{uu} - b_u A_{uv}) / \det \mathbf{A}. \end{aligned} \quad (5.60)$$

This can be cast in the form

$$\mathbf{g} = \sum_{i=1}^M \mathbf{D}_i (\phi_i - \phi_0), \quad (5.61)$$

with the differentiation coefficients $\mathbf{D}_i = (D_{i,u}, D_{i,v})$ given explicitly by

$$\begin{aligned} D_{i,u} &= G_i (p_i A_{vv} - q_i A_{uv}) / \det \mathbf{A}, \\ D_{i,v} &= G_i (q_i A_{uu} - p_i A_{uv}) / \det \mathbf{A}, \quad (i = 1, 2, \dots, M). \end{aligned} \quad (5.62)$$

There still remains the task of defining the weights $G_1, \dots, G_M$. From Eq. (5.56) it is obvious, that any choice of positive numbers provides an exact differentiation of linear functions $\phi(u, v)$. But it is possible to achieve at least fairly accurate and often exact results for quadratic functions. To find these, we consider the orthogonal structure shown in Fig. 5.19b. The formulas (5.60) then specialize to $A_{uv} \equiv 0$ and

$$g_u = b_u / A_{uu}, \quad g_v = b_v / A_{vv}, \quad (5.63a)$$

and the differentiation coefficients $\mathbf{D}_i$, of Eq. (5.62) consequently become

$$\begin{aligned} D_{i,u} &= G_i p_i / (G_1 p_1^2 + G_3 p_3^2), \quad (i = 1, 3), \\ D_{j,v} &= G_j q_j / (G_2 q_2^2 + G_4 q_4^2), \quad (j = 2, 4). \end{aligned} \quad (5.63b)$$

These are to be brought into agreement with the common three-point formula (3.67), when the notation has been adopted accordingly. After a short elementary calculation, we obtain complete agreement with the choice

$$G_1 = p_1^{-3}, \quad G_2 = q_2^{-3}, \quad G_3 = |p_3|^{-3}, \quad G_4 = |q_4|^{-3}. \quad (5.64)$$

The weights must be invariant with respect to rotation of the coordinate system; hence the correct general choice is

$$G_i = |\mathbf{w}_i|^{-3} = (p_i^2 + q_i^2)^{-3/2}, \quad (i = 1, \dots, M). \quad (5.65)$$

By means of numerical checks, it is possible to verify that the procedure is exact for quadratic functions if the polygon is regular or if the neighbor points are pairwise located on straight lines through the center, that implies $\mathbf{w}_j \times \mathbf{w}_{j+M/2} = 0$, $j = 1, \dots, M/2$. In the general case, the error remains fairly small.

The coefficients $\mathbf{D}_i$ can easily be calculated for all nodes in a general triangular mesh. This makes it possible to eliminate all partial derivatives as *free variables* from the Hermite interpolation (Section 5.3.3). We then obtain a SOFEM with the advantages of the latter but with only *one* degree of freedom per node. This is certainly a highly advantageous method.

5.4 THE FINITE-ELEMENT METHOD IN FIRST ORDER

The approximation of first-order finite-element method (FOFEM) consists in the use of the area coordinates ξ_1 , ξ_2 , and ξ_3 themselves as trial functions N_i in each triangular element. Because of its simplicity, this approximation was used in Munro's first programs [10] for the calculation of magnetic lenses and is still used by Lencova [18] and many others. This version of the FEM is still simple enough for some analytical conclusions to be derived, while all higher order versions require tedious evaluations by computer. This is another reason why the FOFEM is discussed here.

Here again the notation of Section 5.3.5 is adopted. We shall consider those FEM equations that refer to one central node as representative of all of them; once again its label is chosen as zero without loss of generality; the cyclic numbering of its neighbors is now necessary for reasons of conciseness. The domain in Fig. 5.19a must be considered as the local carrier of the trial function, which is the pyramid shown in Fig. 5.20.

We first discuss the discretization of self-adjoint partial differential equations and then specialize to the magnetic lens as the most important case of application.

5.4.1 Self-Adjoint Partial Differential Equations

We now resume the considerations of Section 5.2. The FOFEM is a method that satisfies Eqs. (5.20) and (5.21). The potential terms on the right-hand side

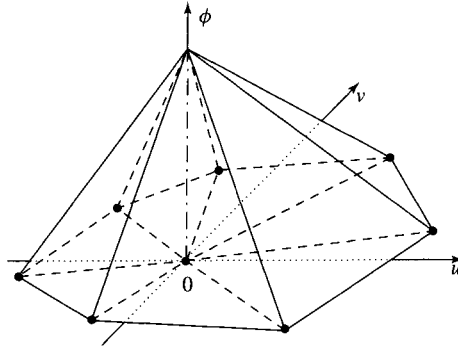


FIGURE 5.20 Perspective graph of linear form functions $\phi(u, v)$ referring to the common node 0 in a carrier of triangles; see also Fig. 5.19a. The height of this pyramid is unity.

of Eq. (5.22) appear only in the vicinity of boundaries. Apart from the fact that they are kept constant, they generate no new features. We hence shift then to the left-hand side and then always have a complete carrier. If we adopt the numbering shown in Fig. 5.19, the equation corresponding to one member of the system (5.22) becomes

$$-\sum_{k=0}^M L_{0k} \phi_k = S_0, \quad (\phi_k \equiv V_k). \quad (5.66)$$

For reasons of conciseness, we shall drop the terms q_{ik} given by Eq. (5.15); their numerical evaluation does not raise any particular problems, and in very many practical cases such terms do not appear.

The evaluation of the matrix elements $L_{0k} = \hat{L}_{0k}$ according to Eq. (5.13) is facilitated by the fact that in each element the gradient is constant and given by Eq. (5.45); this can then be taken in front of the corresponding integrals. The concise presentation of the final formula requires now a redefinition of the *cyclicity* of notation: the neighbors of a position w_i are now $w_{i\pm 1}$ *modulo* M , which means $w_1 = w_{M+1} \bmod M$ and $w_{1-1} = w_M \bmod M$.

For reasons of conciseness, we shall not repeat *mod* M explicitly, but it must be kept in mind.

The elements are now sequentially numbered from 1 to M . The formulas in Section 5.3.1 are to be evaluated in turn with the following shift of labels:

$$\left. \begin{aligned} 1 &\hat{=} 0, & 2 &\hat{=} i, & 3 &\hat{=} i+1 \\ s_3 &= w_i, & s_2 &= -w_{i+1}, & s_1 &= w_{i+1} - w_i \end{aligned} \right\}, \quad (i = 1, \dots, M \bmod M). \quad (5.67)$$

The double areas then become $D_i = |\mathbf{w}_i \times \mathbf{w}_{i+1}|$ and with these the mean material coefficients can be defined as

$$\bar{\varepsilon}_i = 2D_i^{-1} \int_{\Delta_i} \varepsilon(u, v) da, \quad (i = 1, \dots, M). \quad (5.68)$$

Within the frame of the FOFEM, it is permissible to approximate this quantity by the function value at the centroid of the corresponding triangle because any better approximation would make little sense.

The gradient term of the functional for such a triangle Δ_i can now be evaluated easily by combination of Eq. (5.68) with Eq. (5.46), after the numbering has been adapted accordingly; we then obtain

$$\frac{1}{2} \int_{\Delta_i} \varepsilon |\nabla \phi|^2 da = \frac{\bar{\varepsilon}_i}{4D_i} (s_{0,0} \phi_0^2 + 2s_{0,i} \phi_0 \phi_i + 2s_{0,i+1} \phi_0 \phi_{i+1} + R_i), \quad (5.69)$$

R_i denoting the remainder, which does not depend on ϕ_0 . The matrix elements $L_{0,k}$ contributing to Eq. (5.66) are now obtained by summation of all such contributions. Recalling the definition of the factors s_{uv} by Eq. (5.28), we obtain for the integral over the carrier C ,

$$\begin{aligned} \int_C \frac{\varepsilon}{2} |\nabla \phi|^2 da = & \sum_{i=1}^M \frac{\bar{\varepsilon}_i}{4D_i} ((\mathbf{w}_{i+1} - \mathbf{w}_i)^2 \phi_0^2 - 2\mathbf{w}_{i+1} \cdot (\mathbf{w}_{i+1} - \mathbf{w}_i) \phi_0 \phi_i \\ & - \mathbf{w}_i \cdot (\mathbf{w}_{i+1} - \mathbf{w}_i) \phi_0 \phi_{i+1} + R_i). \end{aligned} \quad (5.70)$$

This expression is to be differentiated with respect to ϕ_0 and can then be compared with Eq. (5.66). Each coefficient L_{0k} consists of two terms with the labels $i = k$ and $i = k - 1$, respectively. To cast the following calculation in a concise form, we introduce the frequently appearing normalized coefficients

$$\begin{aligned} \gamma_k^\pm &:= \mathbf{w}_{k\pm 1} \cdot (\mathbf{w}_{k\pm 1} - \mathbf{w}_k) / |\mathbf{w}_{k\pm 1} \times \mathbf{w}_k| \equiv \cot \alpha_k^\pm, \\ (k &= 1, \dots, M, \text{ cycl}), \end{aligned} \quad (5.71)$$

$\alpha_k^\pm$ being the angles opposite to the vector $\mathbf{w}_k$ on the corresponding side, as is shown in Fig. 5.21. These coefficients have evidently a purely geometrical meaning and are invariant with respect to scale transforms. The differentiation of expression (5.70) with respect to ϕ_0 leads to the coefficients

$$-L_{0k} = (\bar{\varepsilon}_{k-1} \gamma_k^- + \bar{\varepsilon}_k \gamma_k^+) / 2, \quad (k = 1, \dots, M, \text{ cycl}), \quad (5.72a)$$

$$L_{00} = \sum_{k=1}^M \bar{\varepsilon}_k |\mathbf{w}_{k+1} - \mathbf{w}_k|^2 / 2D_k = - \sum_{k=1}^M L_{0k} > 0. \quad (5.72b)$$

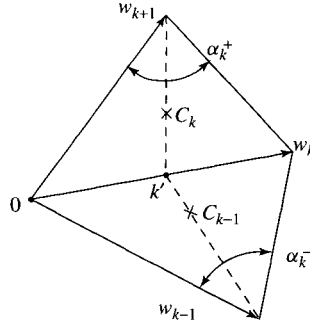


FIGURE 5.21 Two adjacent triangular elements with the corresponding notation. The material coefficients should be evaluated not at the centroids C_{k-1} and C_k but at the side midpoint k' with vector $w_{k/2}$.

Evidently, a coefficient L_{0k} vanishes if the vector w_k is a diagonal in an *orthogonal* structure. The source term (5.16) for a particular triangle Δ_i can easily be evaluated by interpolating $\phi(u, v)$ and $Q(u, v)$ linearly according to Eq. (5.44) and integrating by means of Eq. (5.42). In local notation the result is

$$\int_{\Delta_i} Q\phi da = \frac{D_i}{24} \left[\sum_{\mu=1}^3 Q_\mu \sum_{v=1}^3 \phi_v + \sum_{v=1}^3 Q_v \phi_v \right]. \quad (5.73)$$

On summation of all these integrals for the carrier C and adopting the cyclic notation we arrive at

$$\int_c Q\phi da = \frac{1}{24} \phi_0 \sum_{i=1}^M D_i (2Q_0 + Q_i + Q_{i+1}) + \tilde{R}, \quad (5.74)$$

$\tilde{R}$ being again a remainder independent of ϕ_0 ; we then obtain immediately

$$S_0 = \frac{1}{24} \sum_{i=1}^M D_i (2Q_0 + Q_i + Q_{i+1}). \quad (5.75)$$

The combination of this approximation with that given in Eqs. (5.72) is what is explicitly or implicitly used in FOFEM programs. In general it holds exactly only for *linear* functions $\phi(u, v)$ that is for *homogeneous* fields. Such a strong restriction is already inconsistent with $\varepsilon(u, v) \neq \text{const.}$ and $Q(u, v) \neq 0$. Hence discretization errors of *second* order must appear, unless these cancel out mutually because of special symmetries; this problem is investigated in the following section.

5.4.2 Error Analysis and Improvements

In the case of arbitrary triangular meshes, it is difficult to analyse the discretization errors of the approximation (5.66) with (5.72) and (5.75). However, for meshes that are obtained by subdividing along diagonals, as shown in Fig. 5.22 such an analysis is feasible.

According to Eq. (5.72a), the L -coefficients vanish for all neighbors in diagonal directions, regardless of how the bisections into triangles were made, and consequently a *five-point formula* is found. However, this depends on the mode of subdivision, because the material coefficient ε is evaluated at different centroids of triangles. This inconsistency already causes a discretization error of second order, which can easily be removed.

It is favorable to introduce the coefficients

$$\varepsilon_k^\pm := \varepsilon(u_0 + p_k/2 \mp \eta q_k, v_0 + q_k/2 \pm \eta p_k), \quad (k = 1, \dots, M, \eta \ll 1), \quad (5.76a)$$

if $\varepsilon(u, v)$ is discontinuous and

$$\varepsilon'_k = \varepsilon_k^+ = \varepsilon_k^- = \varepsilon(u_0 + p_k/2, v_0 + q_k/2) \quad (5.76b)$$

for a continuous function. This means that the coefficient is determined not at a centroid but at the *midpoint* (k') on the line from point (0) to point (k), as shown in Fig. 5.21. Equation (5.72a) is now replaced by

$$-L_{0k} = (\varepsilon_k^- \gamma_k^- + \varepsilon_k^+ \gamma_k^+)/2, \quad (k = 1, \dots, M), \quad (5.77)$$

whereupon the second and more general form of Eq. (5.72b) has still to be satisfied. After a corresponding change of notation and some elementary calculations, it is now possible to show that for the *homogeneous* PDE with $A = B = \varepsilon$ and $D \equiv Q \equiv 0$ the five-point formula (4.60) is obtained, as it should be.

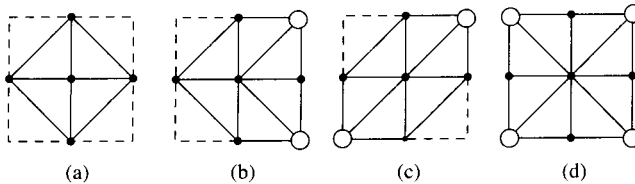


FIGURE 5.22 Different triangulations of a square-shaped mesh. The points in diagonal directions always have vanishing coefficients. In the traditional form of the FEM only configuration (c) gives the correct results.

Another source of errors of second order is Eq. (5.75). This is most easily seen for a constant function $Q(u, v) = Q_0$, Eq. (5.75) then simplifying to

$$S_0 = \frac{1}{6} Q_0 \sum_{i=1}^M D_i = Q_0 a_c / 3, \quad (5.78)$$

a_c being the total area of the carrier. For the configurations shown in Fig. 5.22, this takes *different* values, whereas the L -coefficients remain the same. This is certainly a severe error that must be eliminated by suitable modifications.

A fairly easy way is found by imposing the condition that a function with some free parameters shall become a particular integral of the *inhomogenous* PDE. An example of such a function is

$$\phi(u, v) = \phi_0 + (p^2 + q^2)(C_0 + pC_1 + qC_2)/4, \quad (5.79a)$$

where C_0 , C_1 , and C_2 being the free parameters. We must assume here a linear function $\varepsilon(u, v)$, and we write

$$\varepsilon_0 := \varepsilon(u_0, v_0), \quad \varepsilon' := \varepsilon(u_0 + p/2, v_0 + q/2). \quad (5.79b)$$

The source function corresponding to Eq. (5.79a) is then found by differentiation

$$Q(u, v) := \nabla \cdot (\varepsilon \nabla \phi) = \varepsilon' (3C_0 + 2pC_1 + 2qC_2) - 2\varepsilon_0 C_0 + R, \quad (5.79c)$$

where R is the remainder. The latter consists of quadrupole terms with derivatives of ε as coefficients; such terms can be ignored within the framework of the FOFEM. The source at the center, $Q_0 = \varepsilon_0 C_0$, can now be used to eliminate the free parameters, the result being

$$\varepsilon'(\phi - \phi_0) = (p^2 + q^2)[Q + (2 - \varepsilon'/\varepsilon_0)Q_0]/8. \quad (5.79d)$$

Considering Eqs. (5.76b) and (5.77), the total source term becomes

$$-\sum_{k=1}^M L_{0k}(\phi_k - \phi_0) = S_0 = \sum_{k=0}^M C_{0k} Q_k, \quad (5.80a)$$

with the source coefficients

$$C_{0k} = (p_k^2 + q_k^2)(\gamma_k^- + \gamma_k^+)/16 \quad (k = 1, \dots, M), \quad (5.80b)$$

$$C_{00} = \sum_{k=1}^M (p_k^2 + q_k^2)(\gamma_k^- + \gamma_k^+)(2 - \varepsilon'_k/\varepsilon_0)/16. \quad (5.80c)$$

If the source function $Q(u, v)$ is discontinuous, like the current density at the surface of a coil, for example, then the following replacements should be made:

$$(\gamma_k^- + \gamma_k^+)Q_j \longrightarrow \gamma_k^- Q_j^- + \gamma_k^+ Q_j^+ \quad (j = 0 \text{ or } k). \quad (5.81)$$

All the deficiencies described earlier now vanish, because only those nodes that contribute also to the potential term contribute to the source term. In orthogonal structures the final formulas become *independent* of the mode of bisection into triangles. Moreover, it can be shown that the mesh formula is now consistent with Eq. (4.60) in its inhomogeneous form, as it should be.

In spite of these improvements, which can easily be incorporated in FOFEM programs, this method is still *not* exact in second order, if *tilted* meshes are taken into account, as must be done in realistic applications. The error can then be reduced only by suitable choice of the meshes. Apart from orthogonal ones, there is a second class of configurations that are obtained by smooth and locally *affine* deformations of regular hexagonal elements, as is shown in Fig. 5.23. In such a case, the deformations should be small enough such that obtuse angles never appear, as the latter would cause negative potential coefficients. The optimum consists of completely regular elements, whose construction was dealt with in Section 5.1. If abrupt alterations of the mesh sizes cannot be avoided because of unfortunate geometric constraints, then discretization errors of *second* order must be taken into account.

All these arguments demonstrate that the FOFEM is not very favorable. The only reason for its frequent use in practice is its simplicity when compared with other methods.

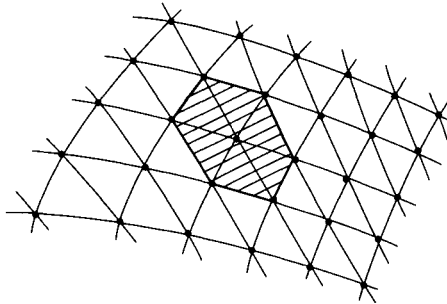


FIGURE 5.23 Appropriate mesh construction for the application of the FOFEM by affine distortion of regular hexagonal structures. The hatched area is an example for one of the hexagonal carriers.

5.4.3 Quadrilateral Meshes

These are quite often used in FEM programs and certainly have many advantages, the most important one being the ease of integration by bivariate Gauss quadratures. Within the FOFEM-approximation, *bilinear* trial functions that correspond to Eqs. (5.53) can be used [32]. The evaluation of the FEM now results in a *nine-point formula* for every internal node. This proves to be equivalent to the general one derived in Section 4.3, the weights being $w_u = w_v = 1/6$.

However, there is no obligation to choose these particular weights because there is no need to use bilinear functions. Such a restriction arises only from the familiar but not always advantageous philosophy that each finite element must be processed completely independently from its surrounding. Because the nine-point formulae in Section 4.3 have a higher accuracy, the computational effort being practically the same, these should be preferred.

Another way of improving the approximation is to use bivariate *cubic Hermite* functions (see Section 3.5.2). These require *four* degrees of freedom per node, as is obvious from Eq. (3.172). This is the same requirement for memory as for the SOFEM, but the accuracy will be better.

5.4.4 The Magnetic Lens

The rotationally symmetric magnetic lens with a ferromagnetic yoke is a standard object of FEM calculations, because nonlinear material properties in the yoke can hardly be evaluated appropriately by other methods. In fact, the first realistic calculations of magnetic lenses by Munro [10] used the FOFEM on the basis of the variational principle (2.55). Here we shall set out from the equivalent form (2.56), which has the advantage that the expression (2.49) for $|\mathbf{B}|$ is simpler than the corresponding expression formed with the vector potential. The factor r^{-1} will not cause any difficulty. Because the essential nonlinearity is included in the gradient term of F , the considerations of Section 5.2 are not directly applicable and require some modifications. In this section, we choose $u = z$ and $v = r$ as coordinates and the fluxes $\psi_i = V_i$ as node values; Eq. (5.7) becomes

$$\psi(z, r) = \sum_{i=1}^N \psi_i N_i(z, r). \quad (5.82)$$

Because $\psi(z, 0) = 0$ and $\psi(z, r) \sim r^2$ is valid in the vicinity of the optic axis, any self-consistent series expansion must satisfy these conditions, but at the moment we postpone the appropriate choice of trial functions and assume that the necessary differentiations can be carried out in all triangular elements.

The essential quantity in the Λ -term of Eq. (2.56) is now

$$\beta := |\mathbf{B}| = (2\pi r)^{-1} \left\{ \sum_i \sum_k \psi_i \psi_k (N_{i/z} N_{k/z} + N_{i/r} N_{k/r}) \right\}^{1/2}, \quad (5.83)$$

and with this the minimization condition can be rewritten as

$$F_\psi = \iint_G \left[2\pi r \Lambda(z, r, \beta) - j(z, r) \sum_{k=1}^N \psi_k N_k(z, r) \right] dz dr = \min, \quad (5.84)$$

which leads to the necessary conditions

$$\partial F_\psi / \partial \psi_i = 0, \quad (i = 1, \dots, N). \quad (5.85)$$

The differentiation of the integrand results in a system of nonlinear equations:

$$\iint_G \left[(2\pi r \beta)^{-1} \partial \Lambda / \partial \beta \sum_k (\nabla N_i \cdot \nabla N_k) \psi_k - j N_i \right] dz dr = 0, \quad (i = 1, \dots, N). \quad (5.86)$$

We now recall Eq. (1.56) with $B = \beta$ and find that the essential nonlinearity is included in the reluctivity $\nu(z, r, \beta) = \mu^{-1}$, as was to be expected; Eq. (5.86) can now be cast in the concise form

$$\sum_{k=1}^N L_{ik}(\beta) \psi_k = J_i, \quad (i = 1, \dots, N), \quad (5.87)$$

with the matrix elements

$$L_{ik} = \iint_G \frac{\nu(z, r, \beta)}{2\pi r} \nabla N_i \cdot \nabla N_k dz dr \quad (5.88)$$

and the driving terms

$$J_i = \iint_G j(z, r) N_i(z, r) dz dr. \quad (5.89)$$

We have reached a result, which is analogous to Eqs. (5.13) and (5.16) if we identify $Q := -j$ and $\varepsilon := \nu/(2\pi r)$, which was to be expected from the form of PDE (2.46). With these settings, the *outer form* for the application of Eqs. (5.77), (5.80), and (5.81) is given, and the evaluation of Eq. (5.89) in

this sense is, indeed, straightforward. Yet, there are now two new and essential problems.

- (i) The system (5.87) is *nonlinear*. This difficulty arises regardless of the choice of the set of trial functions. An exact solution of such systems of equations with realistic magnetization functions is practically impossible and also not necessary because errors were already introduced by the discretization. Within the FOFEM it suffices to evaluate β at the centroids of the elements, then determine ν with these values, and consider this value as constant for the corresponding element. We thus have a piecewise constant function, but this depends on the node values of the corresponding potentials. As these alter during the calculations, the whole system of equations can be solved only iteratively. This procedure will be described in Section 5.6.
- (ii) The second problem arises from the fact that the *linear* trial functions cannot be used for elements adjacent to the optic axis, as is immediately obvious. Moreover, within a certain neighborhood around the axis, this approximation becomes rather poor, but surprisingly the coefficients of Eq. (5.77) with (5.76) and $\varepsilon = \nu/(2\pi r)$ can be used with *good* accuracy!

The reason lies in the fact that the theoretically necessary integration in Eq. (5.88) or Eq. (5.68) is not really carried out but replaced with a mean value. This has the consequence that for any function $f(z, r) = C(z)r^2/2$ the relation

$$\frac{1}{r} \frac{\partial f}{\partial r} = \frac{f(z, r+h) - f(z, r-h)}{2rh} = C(z) \quad (5.90)$$

is *exactly* satisfied. There is hence no need to modify the trial functions at this stage. This is necessary only afterwards when we come to the task of field interpolation because linear approximations of the flux potential are of course certainly wrong.

The boundary conditions for the field calculation problems are easy to understand, but it is often difficult to satisfy them. On the axis we have $\psi = 0$, and these nodes do not even appear on the right-hand side in the expression equivalent to Eq. (5.22). The potentials at the inner surfaces — those of the yokes and the coils — are determined by the electromagnetic boundary conditions. There remain the outer boundary conditions. Theoretically, the flux must vanish at infinity like that of a dipole

$$\psi = \frac{\mu_0 M r^2}{2[r^2 + (z - z_0)^2]^{3/2}}. \quad (5.91)$$

However, it is difficult to determine the strength M and the center z_0 of this dipole. Therefore, a finite, sufficiently distant outer boundary is chosen on

which ψ is assumed to vanish. For closed lenses this causes no problem, but for open lenses it may become difficult for reasons of memory to extend the mesh sufficiently far. Examples of closed and open lenses are shown in Figs. 5.24 and 5.26.

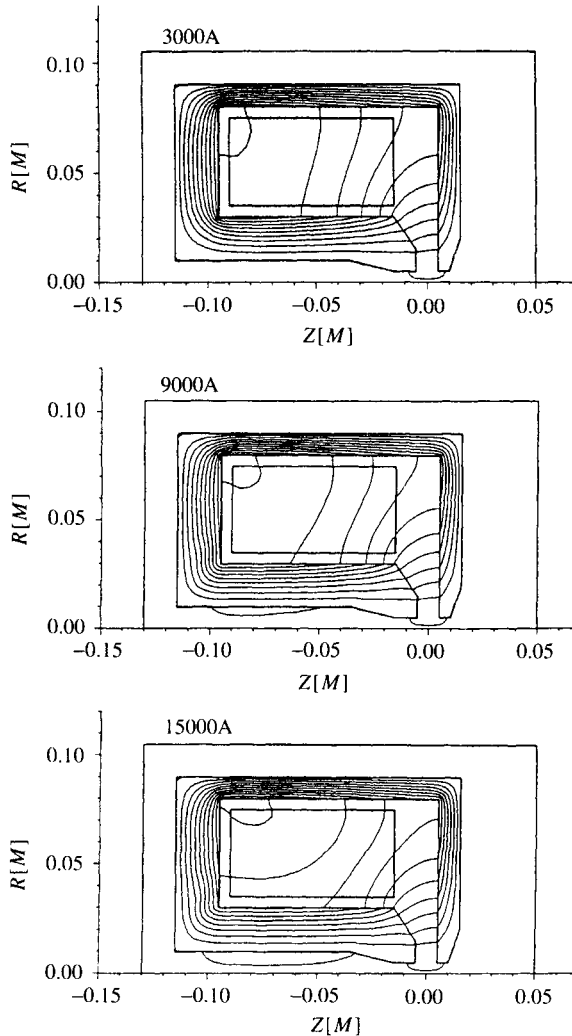


FIGURE 5.24 Half axial section through a closed magnetic lens with increasing coil current and saturation effect. The plots show magnetic flux lines. This lens was calculated first by Munro [9] with a FOFEM. The graphs shown here were determined by K. Zeh [25], who used Lagrange elements of third order (Fig. 5.15).

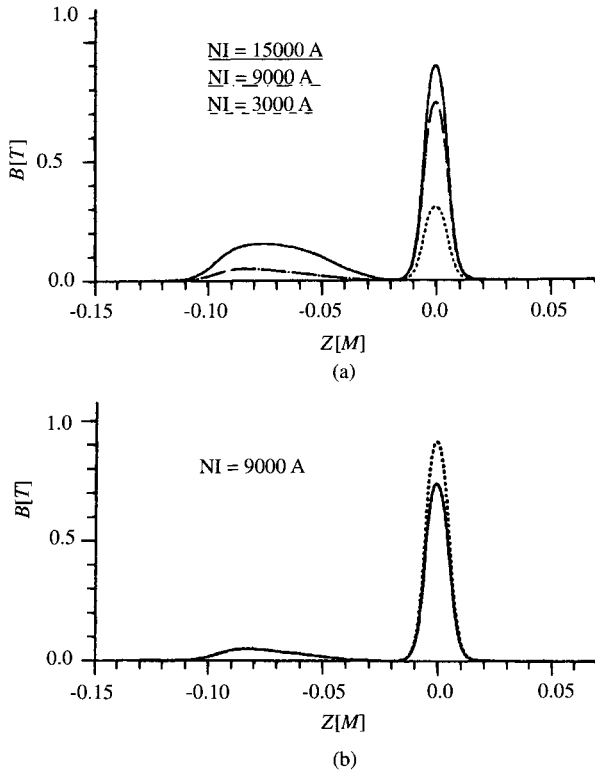


FIGURE 5.25 Axial flux density $B(z)$ in the lens, shown in Figs. 5.1 and 5.24: (a) rigorous nonlinear calculation, showing the growth of side maxima; (b) comparison with a linear calculation ($\mu = \text{const.}$), marked by the dashed line. The central peak has practically the same form, but has a different height, so that the integral under the curve is conserved. Calculation by K. Zeh [25].

Alternative Methods

The application of the FEM in its various forms to field calculation in magnetic lenses is feasible, as is shown in many publications [19–24]. Other possibilities involve its combination with the FDM; then in all domains that can be discretized appropriately by rectangles or squares, the formulas of Section 4.4 with $\alpha = -1$ can be used, whereas the FEM is used in the remainder of the domain. This procedure will certainly be more complicated, but the accuracy will be improved, as mesh formulas for squares have an error of order h^6 for the axial field strength $B(z)$ (see Eq. (4.158)), which can hardly be achieved by any other method.

Finally, it is possible to replace the FEM entirely by the FDM. Then there arises the problem of irregular meshes that is solved using the various methods

described in Section 4.5. For internal mesh points Eq. (4.60) is considered. There is hence a great variety of possible combinations. It will hardly be possible to achieve a high order of the error polynomial in the vicinity of the material boundaries, but a gain is already achieved if this error remains confined to this zone and does not appear in the whole domain of solution.

Figures 5.26 and 5.27 demonstrate the limitation of the FEM and FDM because these graphs were determined using the BEM that will be the subject of Chapter VI. It is extremely difficult to construct a mesh system of 500×500 in such a way that the far-reaching fringe field is included adequately.

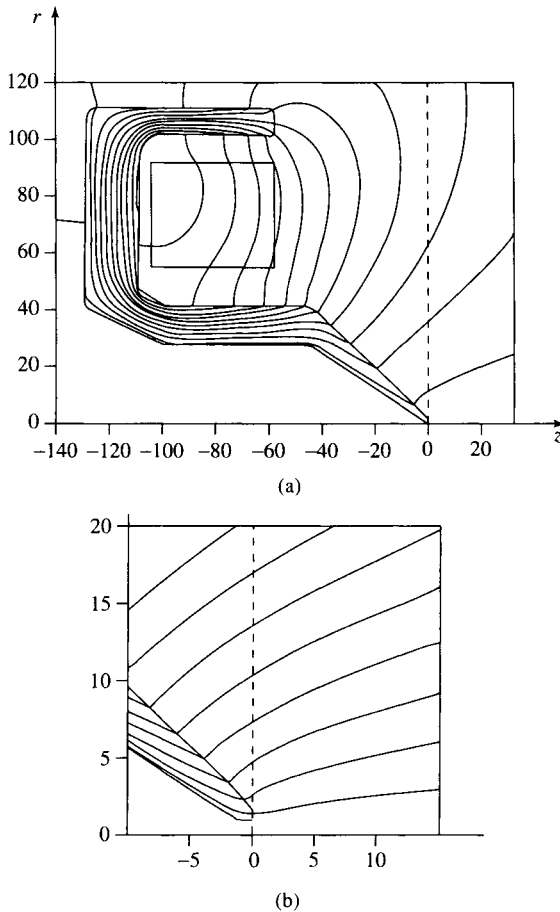


FIGURE 5.26 Half axial section through an open magnetic lens with flux lines $\phi = \text{const.}$: (a) window showing the complete yoke; (b) window near the snorkel close to the optic axis. This lens was calculated by Kasper [33] in cooperation with Knell after a similar design by Shao and Lin [34]. In this case the BEM was used, as the FEM would require too large a mesh.

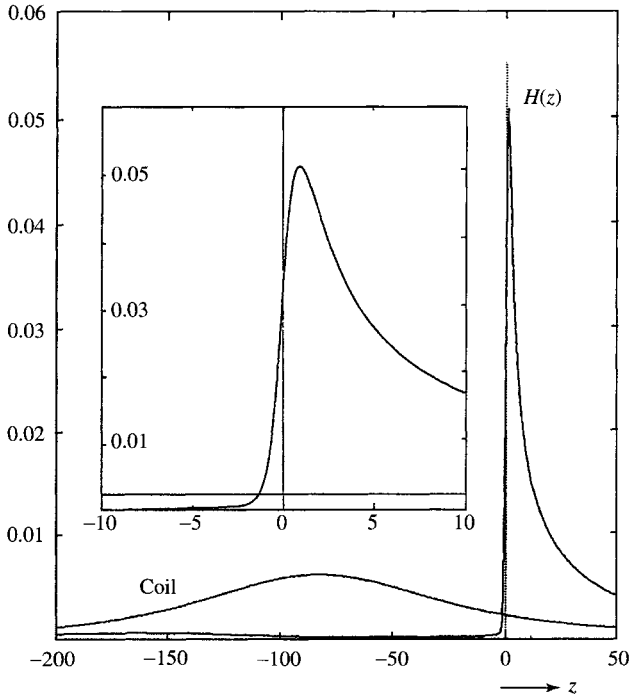


FIGURE 5.27 Axial excitation $H(z)$ of the lens shown in Fig. 5.26. From the figure it can be seen that the contribution from the pure coil is suppressed within the bore and concentrated in the very sharp peak, of which a very narrow window is also depicted. The curves are normalized to unit coil current. Calculation by Kasper [33].

5.5 FIELD INTERPOLATION

The result of calculations using the FDM, the FEM, or any combination of both is a set of discrete potential values referring to the nodes of a grid. This is only the first part, though the most complicated one, of a more complex procedure. The final goal is the determination of the potential and of its partial derivative at any *arbitrary* position $\mathbf{r} = (x, y, z)$ or (u, v, w) within the domain of solution. This requires interpolation and differentiation at any such position, which can become a fairly complicated task. As for physical reasons, *all* derivatives at the inner mesh lines must be continuous. For the longitudinal components, this condition is automatically satisfied in consistent calculations but, for the normal component, it may become a severe obstacle.

Another requirement, particularly in charged particle optics, is the *analytic* behavior of the field strength in the vicinity of the optic axis, since this is the most important domain in the system. This also includes the requirement of

numerical stability, which means that very tiny errors on the position must not cause larger than proportional errors in the field strength. For example, expressions such as $\cos \varphi - x/r$ in polar coordinates must never appear because they are very sensitive to errors and indefinite on the optic axis itself. In this respect, the Fourier series expansion (2.4) is unfavorable, and the separation of factors r^m as in Eqs. (2.8), is to be preferred. For round magnetic lenses, the potential $\Pi(z, r)$, Eq. (2.41), is now to be preferred because this yields a cartesian form of the field strength with $s = r^2$ and

$$B_x = -\frac{x}{2} \frac{\partial \Pi}{\partial z}, \quad B_y = -\frac{y}{2} \frac{\partial \Pi}{\partial z}, \quad B_z = \Pi + s \frac{\partial \Pi}{\partial s}. \quad (5.92)$$

More generally, we shall seek forms of the radial derivatives from which all factors that vanish on the axis have been removed.

Here we shall be concerned with *two-dimensional* interpolation techniques. The need for these arises mainly in the calculation of systems having rotationally symmetric boundaries, which were the topic of Sections (2.1) and (2.2). The interpolation then concerns the field in a meridional section through the system. The exacting task of three-dimensional interpolation is dealt with in Section 7.6.3.

5.5.1 Determination of the Mesh Position

From the input of a coordinate triple (x, y, z) by some outer program part for tracing of particle rays or of equipotentials we can immediately calculate

$$s = x^2 + y^2, \quad r = \sqrt{s}. \quad (5.93)$$

In programs that require the field only in one plane, say the plane $y = 0$, we can also make use of $r = x$ and $s = r^2$, where it may become favorable to allow negative values of r .

The first problem is to determine the mesh cell in which the two-dimensional reference point (z, r) is located, or if no such cell can be found, which procedure should then be followed. This is not yet a problem of interpolation in the proper sense but of a preceding *exploration*. Depending on the kind of mesh generation, this can be almost trivial or very tedious. In this context it is of importance that the method implemented should be as fast as possible because it is executed quite frequently during the tracings. The latter usually proceed in very small steps; it is hence appropriate to check first whether a new reference point is located in the same cell, which is already known or in one of its neighbors. This will be the most common situation, but at least at the beginning a more general procedure is necessary.

The problem of exploration is easily solved in rectangular meshes with *constant* mesh sizes h_u and h_v in both directions. If the mesh was found by a coordinate transformation, this is now required in the form of

$$u = \tilde{u}(z, r), \quad v = \tilde{v}(z, r), \quad (5.94)$$

together with all necessary partial derivatives. If no such transform was necessary, we set $u = z$ and $v = r$ to avoid frequent decisions. After evaluation of Eq. (5.94) we find immediately

$$i - 1 = \text{int}((u - u_0)/h_u), \quad k - 1 = \text{int}((v - v_0)/h_v), \quad (5.95)$$

as the labels of the node on the lower and left-hand side, “int” denoting the truncation to an integer.

A little more complex but still fairly simple is the exploration in meshes obtained by subdivision of coarse grids into finer ones, as is shown in Fig. 5.1. First the appropriate (u, v) -rectangle in the coarse grid is determined, which is quite fast, because the number of such rectangles is small. Subsequently the procedure (5.95) is carried out.

Still more effort is necessary if the mesh is nonuniform but still rectangular (see Fig. 5.28). The grid is then specified by two arrays

$$[u] = \{u_i, i = 0, \dots, M\}, \quad [v] = \{v_k, k = 0, \dots, N\}, \quad (5.96)$$

which must be ordered in strictly monotonic sequence. If no particular knowledge of a new position (u, v) is established, then the *bisection method* is recommended in the literature on numerical analysis [35]. This consists in subsequent halving of the index interval. Starting with the full interval $(0, M)$

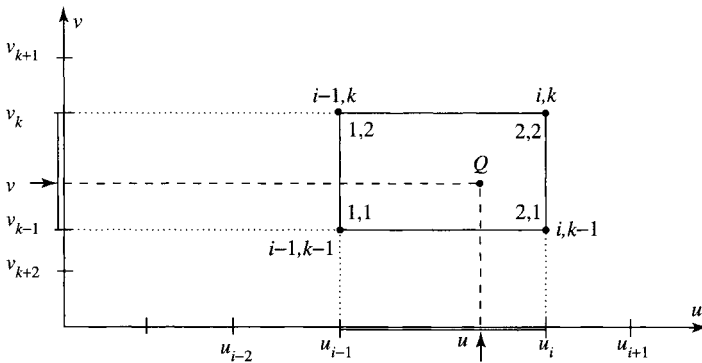


FIGURE 5.28 Determination of the appropriate rectangular domain of interpolation at an arbitrary position Q and connection between local and global labels.

the remaining interval is bisected in each step, and then examined to determine in which of the two corresponding subintervals in the u -array the given u -value is located; the other subinterval is then abandoned. This procedure converges much faster than a sequential scan. Concerning the v -array, the same procedure is carried out. Of course, combination with the procedure of Eq. (5.95) is possible if the grid is constructed accordingly.

Finally, we come to the most complicated problem, the exploration in an arbitrary triangular grid. Without additional information this task is very tedious, and that may be the reason why such grids are rarely used. A feasible strategy is as follows. During the construction of the triangular grid, a rectangular one with two arrays $[u]$ and $[v]$ in the sense of Eq. (5.96) have been introduced; for each pair (i, k) of labels the number $N(i, k, n)$ ($n = 1, 2, \dots$) of all those triangular elements that are located partly in the corresponding rectangle are stored. If the grid is fine enough, there will be only very few of these.

If an arbitrary position (u, v) is prescribed, then first the two one-dimensional searches are carried out, which is straightforward. Now the subarray of numbers $N(i, k, n)$ with *given* labels i and k is known, and it only remains to control of position for a *very few* elements, which can be made quite fast. The given point is located inside an element, if all three barycentric coordinates are *nonnegative*. For the first check, it suffices to determine the numerators of ξ_1, ξ_2, ξ_3 in Eq. (5.31) as determinants. If one of them turns out to be negative, the corresponding element can immediately be abandoned and the next one in the list is then checked. If no element was found in this way, this indicates that the position (u, v) is invalid, and the procedure must then be terminated with an error message. In any case a quick answer is obtained.

5.5.2 Interpolation in Rectangular Meshes

This problem was already dealt with in Section 3.5, and in fact, the methods presented there give fairly accurate results. In routines for tracing equipotentials, these techniques are quite sufficient. If, however, numerical ray tracing is to be carried out, problems arise because cubic Hermite polynomials do *not* satisfy the conditions of analyticity at the optic axis. Unfortunately, the relative error is largest in the vicinity of the axis. To see this, we consider a function

$$f(r) = A_0 + A_1 r^2 + A_2 r^4, \quad (5.97)$$

from which the function values $f(nh)$ and those of the derivative, $f'(nh)$ may be prescribed at equidistant positions $n = 0, 1, 2, \dots$. The error of the cubic Hermite interpolation is largest near the midpoint of each interval and

is determined by

$$\begin{aligned} \varepsilon_n = & \frac{1}{2}f(nh) + \frac{1}{2}f((n+1)h) + \frac{h}{8}f'(nh) - \frac{h}{8}f'((n+1)h) \\ & - f((n+1/2)h), \end{aligned} \quad (5.98)$$

the expression in the upper line representing the interpolated value. The evaluation of this expression results in

$$\varepsilon_n = -h^4 A_2(6n+1)/16. \quad (5.99)$$

This increases only linearly, whereas the corresponding potential term increases with the fourth power. This error implies that the electron optical aberrations, being roughly proportional to $A_2 r^3$, cannot be calculated correctly, however small the mesh size h may be.

A way out of this difficulty is the use of the variable $s = r^2$ for the radial terms in the bicubic Hermite interpolation. This means that, instead of the original potential $P(z, r)$, a transformed function $P(z, r) =: \hat{P}(z, s)$ is now used. For $r > 0$ the transformation of the derivatives is simply

$$P_{|r}(z, r) = 2r\hat{P}_{|s}(z, s), \quad P_{|zr} = 2r\hat{P}_{|zs}(z, s), \quad (5.100a)$$

whereas on the axis we have

$$P_{|rr}(z, 0) = 2\hat{P}_{|s}(z, 0), \quad P_{|zrr}(z, 0) = 2\hat{P}_{|zs}(z, 0), \quad (5.100b)$$

which must be determined by suitable finite differences formula. This procedure removes all difficulties in the vicinity of the optic axis, as we can now easily calculate

$$P_{|x} = 2x\hat{P}_{|s}, \quad P_{|y} = 2y\hat{P}_{|s} \quad (5.100c)$$

without worrying about indefinite expressions. In spite of this advantage, there is still the unfavorable property that in one direction the sixth power of the coordinate is considered, whereas the expansion is already truncated after the third power in the other one.

5.5.3 Improved Hermite Interpolation

The accuracy of the field interpolation is improved by the use of Hermite polynomials of *fifth* order. This makes sense only when the potentials at the nodes were calculated with a nine-point algorithm having a discretization

error of *sixth* order. Otherwise the algorithms outlined here is still feasible but would bring only some smoothing of the results. Quite generally, it is better to determine first the partial derivatives at the nodes by sufficiently precise numerical formula and to store them before embarking on the task of interpolation because the differentiation of polynomials always causes some loss of accuracy.

We now assume equidistant grids in a u - v plane; although it is certainly favorable to have equal spacings $h_u = h_v$ in the two directions, this added assumption is not necessary here. For reasons of conciseness, we introduce the abbreviations

$$U_{ik} := P|_u(u_i, v_k), \quad V_{ik} := P|_v(u_i, v_k), \quad W_{ik} := P|_{uv}(u_i, v_k), \quad (5.101)$$

for the derivatives at the node $u_i = u_0 + ih_u$, $v_k = v_0 + kh_v$.

Because the interpolation polynomial and the accuracy of the FDM calculations are both of fifth order, a seven-point formula for numerical differentiation is here adequate; we hence obtain

$$\begin{aligned} U_{ik} &= [45(P_{i+1,k} - P_{i-1,k}) - 9(P_{i+2,k} - P_{i-2,k}) + P_{i+3,k} - P_{i-3,k}] / 60h_u, \\ V_{ik} &= [45(P_{i,k+1} - P_{i,k-1}) - 9(P_{i,k+2} - P_{i,k-2}) + P_{i,k+3} - P_{i,k-3}] / 60h_v, \\ W_{ik} &= [45(V_{i+1,k} - V_{i-1,k}) - 9(V_{i+2,k} - V_{i-2,k}) + V_{i+3,k} - V_{i-3,k}] / 60h_u, \end{aligned} \quad (5.102)$$

the error being of sixth order in all three cases. In the vicinity of boundaries or margins, where some of the necessary points are missing, the extrapolation rules based on symmetries, as outlined in Section 3.3.3, must be evaluated appropriately. If this is impossible, then asymmetric formulas can be used at the price of some loss of accuracy.

We consider now the configuration shown in Fig. 5.28. To carry out the Hermite interpolation at an arbitrary point Q inside the rectangle, we need to know the partial derivatives at the four corners. The four arrays P , U , V , and W are sufficient only for the bicubic interpolation, outlined in the preceding section. It is, of course, possible to extend the procedure in Eqs. (5.102) to higher orders, but this would require too much memory. An alternative way is shown in Fig. 5.29: the rectangle of Fig. 5.28 now becomes the central one in a configuration of nine rectangles or of 16 points. The exceptional cases of marginal locations will be discussed below.

For reasons of symmetry and continuity, it is necessary to consider only the eight closest neighbors of each of the four inner nodes, because these remain in common with the corresponding neighboring cells. This is shown for the node (0) as an example. The nearest neighbors considered here are numbered sequentially from 1 to 8; the rigorous two-dimensional indexing must of course be used in a practical program.

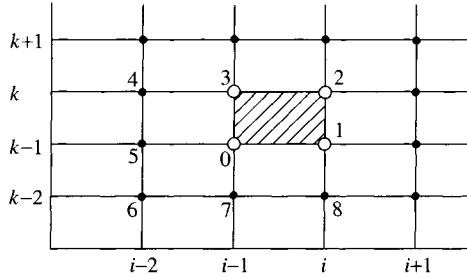


FIGURE 5.29 Connection between the simplified inner labels referring to the node 0 with the global ones. The other three corners are to be treated analogously.

The technique for determining derivatives of higher orders is quite simple. We assume that $f(x)$ is a six times continuously differentiable function from which the function values and derivatives of first order may be given at three positions $x - h$, x , and $x + h$. We can then write down the two Taylor series expansions for $x \pm h$, the derivatives referring to the central position, and form the following linear combinations.

$$\begin{aligned} f(x+h) + f(x-h) - 2f(x) &= h^2 f'' + h^4 f^{(4)}/12 + h^6 f^{(6)}/360 + \dots \\ hf'(x+h) - hf'(x-h) &= 2h^2 f'' + h^4 f^{(4)}/3 + h^6 f^{(6)}/60 + \dots \end{aligned} \quad (5.103)$$

By elimination of $f^{(4)}$ and solving for $f''(x)$ we obtain immediately

$$\begin{aligned} h^2 f''(x) &= 2[f(x+h) + f(x-h) - 2f(x)] \\ &\quad - h[f'(x+h) - f'(x-h)]/2 + h^6 f^{(6)}/360. \end{aligned} \quad (5.104)$$

By repeated application of this formula, we can obtain the matrix elements in the following manner:

$$G_{j,l}^{(m,n)} := h_u^m h_v^n \partial^{m+n} P / \partial u^m \partial v^n |_{(u_{i+j-2}, v_{k+l-2})}, \quad (0 \leq j \leq 3, \quad 0 \leq l \leq 3). \quad (5.105)$$

The elements for $m \leq 1$ and $n \leq 1$ are obtained from the stored arrays, and even the multiplications with h_u and h_v can be saved, if the corresponding multiplied arrays are stored instead of U , V , and W . The differentiation in the u -direction gives

$$\begin{aligned} G_{j,l}^{(2,n)} &= 2(G_{j-1,l}^{(0,n)} - 2G_{j,l}^{(0,n)} + G_{j+1,l}^{(0,n)}) - 0.5(G_{j+1,l}^{(1,n)} - G_{j-1,l}^{(1,n)}), \\ (n &= 0, 1; \quad j = 1, 2; \quad l = (0), 1, 2, (3)) \end{aligned} \quad (5.106a)$$

and similarly for the v -direction:

$$G_{j,l}^{(m,2)} = 2(G_{j,l-1}^{(m,0)} - 2G_{j,l}^{(m,0)} + G_{j,l+1}^{(m,0)}) - 0.5(G_{j,l+1}^{(m,1)} - G_{j,l-1}^{(m,1)}),$$

$$(m = 0, 1; \quad l = 1, 2; \quad j = (0), 1, 2, (3)). \quad (5.106b)$$

The labels in parentheses are those for which the corresponding elements can be calculated but are not needed. The matrix scheme is completed by

$$G_{j,l}^{(2,2)} = (G_{j-1,l-1}^{(1,1)} - G_{j+1,l-1}^{(1,1)} + G_{j+1,l+1}^{(1,1)} - G_{j-1,l+1}^{(1,1)})/4,$$

$$(j = 1, 2; \quad l = 1, 2). \quad (5.106c)$$

These matrix elements are so easy to calculate that it is not necessary to store them permanently.

We now reconsider the configuration of Fig. 5.29. With respect to the interpolation and differentiation at the position Q , it is favorable to introduce the normalized coordinates

$$t_u = (2u - u_i - u_{i-1})/h_u,$$

$$t_v = (2v - v_k - v_{k-1})/h_v. \quad (5.107)$$

The corresponding Hermite polynomials are then given by Eq. (3.93) with

$$H_{1,s}(t) = A_s^-(t), \quad H_{2,s}(t) = A_s^+(t), \quad (s = 0, 1, 2), \quad (5.108a)$$

with replacement of u by t and $t = t_u$ or $t = t_v$, respectively. For the derivatives, we introduce the notation

$$H_{r,s}^{(n)}(t) = 2^n d^n H_{r,s} / dt^n, \quad (5.108b)$$

the power of two arising from the factor 2 in Eqs. (5.107). The interpolation polynomials can now be written in compact form as

$$h_u^m h_v^n P_{(u,v)}^{(m,n)} = \sum_{j=1}^2 \sum_{p=0}^2 \sum_{l=1}^2 \sum_{q=0}^2 H_{j,p}^{(m)}(t_u) H_{l,q}^{(n)}(t_v) G_{j,l}^{(p,q)}, \quad (m + n \leq 2). \quad (5.109)$$

This kind of interpolation furnishes very smooth results as even the normal component of the gradient at the mesh lines is still continuously differentiable, which is difficult to achieve by other techniques. With slight modifications, this kind of interpolation has been used by Killes [36], who obtained very good results with it for ray tracing in electron guns.

There remains the task of interpolation in the vicinity of boundaries. The optic axis is a special case, which is dealt with in the next section. At other symmetry lines, these particular symmetries can be exploited to determine the missing matrix elements. For example if the potential has the property $P(-u, v) = P(u, v)$ and $u = 0, j = 0$ is the lower boundary line, then we know that $G_{0,1}^{(1,n)} = 0$ for all n and l and Eq. (5.106a) is then to be completed by

$$G_{0,l}^{(2,n)} = 4(G_{1,l}^{(0,n)} - G_{0,l}^{(0,n)}) - G_{1,l}^{(1,n)} \quad (5.110)$$

Near an outer boundary to field-free space or to at least a homogeneous field, the missing elements should all become zero, thereby satisfying the natural boundary conditions. In the vicinity of inner boundaries, material surfaces, a precise determination is hardly possible. It is then usually sufficient to assume constant derivatives within the respective mesh and use the values that can be calculated by Eqs. (5.106).

5.5.4 The Paraxial Interpolation

The Hermite interpolation technique outlined earlier is still feasible in the vicinity of the optic axis, if the appropriate values of the partial derivatives on the axis itself are determined by exploiting the even symmetry in the radial direction. Nevertheless, the relations (5.100) are *not* satisfied though the error is certainly less than that of the cubic Hermite interpolation. We must therefore look for some other technique. This should be *compatible* with the Hermite interpolation and similarly fairly general, and we hence impose only the condition of *even symmetry* with respect to the variable v and no other special conditions.

Let $f(v)$ be such a function; the dependence on u is unimportant for the present considerations and will be considered afterwards. It is convenient to introduce a relative variable $\rho = v/h_v$, whereupon we can form the power series expansion

$$f(v) = F(\rho) = a_0 + a_1\rho^2 + a_2\rho^4 + a_3\rho^6 + \dots \quad (5.111)$$

Truncation after the fourth power is too inaccurate: although the sixth power usually has a discretization error, resulting from the FDM approximations, the local error is so small that it is still better to retain this term (see Section 4.4.3). Moreover, the coefficient a_3 is necessary to have sufficiently many degrees of freedom.

Because a power series expansion of the form (5.111) must become inaccurate for large values of ρ , it is necessary to confine its use to the interval $|\rho| \leq 2$ and match it to given values at the position $\rho = 0, 1$, and 2 , as is shown in

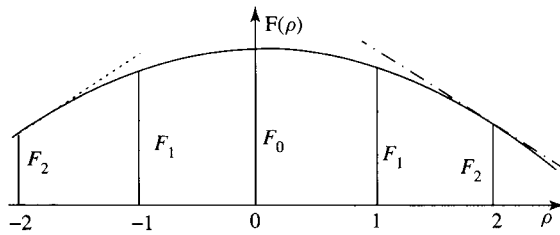


FIGURE 5.30 Interpolation of a symmetric function using three function values and the derivative at $\rho = 2$.

Fig. 5.30. The last free parameter is then favorably the slope $\dot{F}(2) = h_v f'(2h_v)$ at the endpoint, so that a smooth junction with a Hermite polynomial at this point is possible. The coefficients a_n are now well defined and result from the solution of a linear system of equations. In this context, it is favorable to introduce the quantities

$$D_1 = F(1) - F(0), \quad D_2 = F(2) - F(0), \quad D_3 = \dot{F}(2), \quad (5.112a)$$

whereupon we obtain immediately

$$a_0 = f(0) = F(0) \quad (5.112b)$$

and then

$$a_i = \sum_{k=1}^3 C_{ik} D_k, \quad (i = 1, 2, 3), \quad (5.112c)$$

the matrix C being given by

$$C = \frac{1}{144} \begin{pmatrix} 256 & -40 & 12 \\ -128 & 47 & -15 \\ 16 & -7 & 3 \end{pmatrix}. \quad (5.113)$$

Finally, the derivative of second order at the endpoint, necessary for a smooth function, becomes

$$h_v^2 f''(2h_v) \equiv \ddot{F}(2) = (128D_1 - 74D_2)/9 + 31D_3/6. \quad (5.114)$$

The derivative $\dot{F}(1)$ cannot be prescribed independently but is found to be given by

$$h_v f'(h_v) \equiv \dot{F}(1) = 2D_1/3 + 11D_2/24 - \dot{F}(2)/4. \quad (5.115)$$

If the exact value is known, the deviation from it is a measure of the approximation error. It would be possible to extend the approximation to the eighth order, but this makes little sense because the sampling data are not accurate enough for this. Numerical checks show that this error and that of (5.115) are usually smaller than those caused by the FDM calculations.

With respect to simple incorporation in a general interpolation program, it is favorable to introduce *form functions*, so that a shape like that of Eq. (5.109) can be obtained. We therefore write Eq. (5.111) as

$$F(\rho) = \sum_{n=0}^2 \{F(n)K_{n,0}(\rho) + \dot{F}(n)K_{n,1}(\rho)\}, \quad (5.116)$$

where the first label n indicates the radial position and the second one the order of the sampling derivative. However, the functions $K_{0,1}(\rho) \equiv 0$ and $K_{1,1}(\rho) \equiv 0$ could be omitted because the corresponding coefficients do not appear in Eqs. (5.112), but the above given form is more concise.

The evaluation of these kernels is simply a reordering of Eqs. (5.112) and (5.113). If we introduce a formal vector $\mathbf{g} = (g_1, g_2, g_3)$ with

$$g_1 = K_{1,0}, \quad g_2 = K_{2,0}, \quad g_3 = K_{2,1}, \quad (5.117)$$

we obtain then with the *transposed* matrix

$$g_i = \sum_{n=1}^3 C_{ni} \rho^{2n}, \quad (i = 1, 2, 3). \quad (5.118a)$$

The function $K_{0,0}(\rho)$ is then obtained from the condition that any constant must be interpolated *exactly*; this results in

$$K_{0,0}(\rho) = 1 - K_{1,0}(\rho) - K_{2,0}(\rho). \quad (5.118b)$$

Graphs of these functions are shown in Fig. 5.31.

The incorporation of this kind of approximation in the general procedure is now quite easy. The part of the formulas that concerns the interpolation in the longitudinal (u -) direction remains unaltered. This also implies that the recurrence formulas (5.106a) for derivatives with respect to u can be used subsequently. The radial (v -) part is now to be modified as follows:

$$h_u^m h_v^n P_{(u,v)}^{(m,n)} = \sum_{j=1}^2 \sum_{p=0}^2 \sum_{l=0}^2 \sum_{q=0}^1 H_{j,p}^{(m)}(t_u) K_{l,q}^{(n)}(\rho) G_{j,l}^{(p,q)}. \quad (5.119)$$

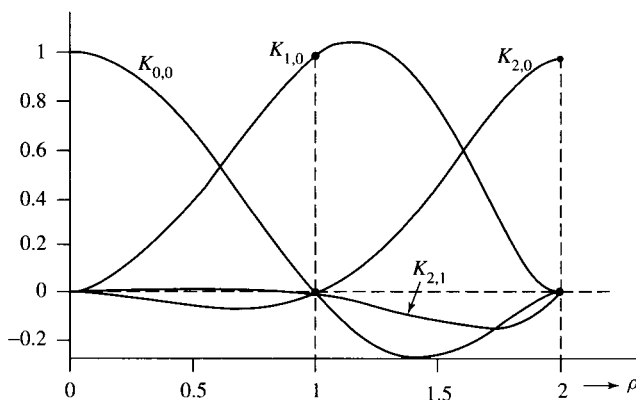


FIGURE 5.31 The form functions of paraxial interpolation.

In comparison with Eq. (5.109), not only has the second Hermite factor been replaced but the ranges of summation are also different. Consequently, the label of v in Eq. (5.105) also must now become v_l instead of v_{k+l-2} .

The radial differentiation formula (5.106b) is still applicable for $l = 2$, but the results will be different from those obtained with Eq. (5.114). The use of this formula is therefore recommended to achieve continuity:

$$\begin{aligned}
 G_{j,2}^{(0,2)} &= -6P_{J,0} + (128P_{J,1} - 74P_{J,2})/9 - h_v V_{J,2}, \\
 G_{j,2}^{(1,2)} &= -6U_{J,0} + (128U_{J,1} - 74U_{J,2})/9 - h_u h_v W_{J,2}, \quad (J = i + j - 2).
 \end{aligned}
 \tag{5.120}$$

The difference between the two kinds of evaluation provides an additional control of the accuracy.

Another version of this kind of interpolation is very favorable if the particle rays to be traced remain entirely in a narrow tube not exceeding the radial extend $2h_v$.

It is then better to evaluate Eqs. (5.112) for the columns $[P_{i,0}, P_{i1}, P_{i2}; V_{i2}]$ ($i = 1, \dots, M$) and $[U_{i,0}, U_{i1}, U_{i2}; W_{i2}]$ ($i = 1, \dots, M$) in turn, and to store the results of these calculations as two arrays A_{ik} and B_{ik} for $i = 1, \dots, M$, $k = 0, \dots, 3$. The Hermite interpolations with respect to the coordinate u can similarly be performed with these arrays and supply field-coefficients $a_0(u), \dots, a_3(u)$ and $b_0(u), \dots, b_3(u)$, which may have some importance for the determination of aberrations. The calculations can be improved further if the second order derivatives with respect to u are not determined as in Eq. (5.104) but by using the given PDE for the field.

There is a great variety of different versions for this task. An improved technique will be presented in the context of the boundary element method. Another has been published by Barth et al. [20].

5.5.5 Interpolation in Trigonal Meshes

This task is inevitably more complicated than those outlined so far. A fairly simple case arises if the mesh is obtained by an affine deformation of regular hexagonal structures, as shown in Fig. 5.32. It is then possible to join two triangles to make a parallel epipedal cell. This can be considered as the affine distortion of a rectangle. We can now introduce a tilted (u, v) coordinate system and the bivariate interpolation techniques, especially the bivariate cubic splines, can then be applied. Thereafter, the transform to the orthogonal (x, y) or (z, r) must be carried out as the last step. The results of such a procedure are fairly accurate.

Unfortunately the required regularity of the mesh rarely occurs in practical applications of the FEM, and we now have to face the most general case. The necessary calculations then proceed as follows:

- (1) Determine the appropriate mesh (see Section 5.5.1)
- (2) Find the numbers of its three nodes from the table
- (3) Determine all data associated with it (potentials, derivatives, etc.)
- (4) Calculate the area coordinates ξ_1, ξ_2, ξ_3 (see Eqs. (5.31))
- (5) Interpolate the potential by means of the appropriate trial function (5.44), (5.47), (5.48), or (5.52), respectively.
- (6) Calculate the gradient from Eq. (5.39).

The results are consistent with the kind of approximations made in the preceding FEM compilations of the potentials. In general, it is impossible to avoid discontinuities of the gradient on the mesh lines. In this respect, the use of the Hermite-trial functions (5.52) is to be preferred. If this has not been the case, there is still the possibility of calculating the node values of the derivatives *afterwards* by means of the procedure outlined in Section 5.3.5.

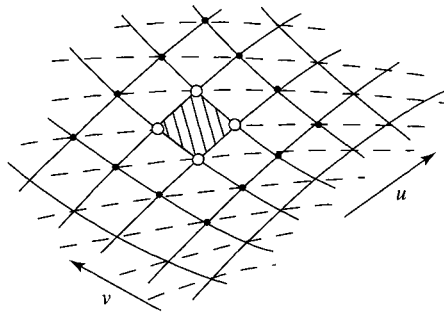


FIGURE 5.32 Slightly deformed triangular meshes can be transformed to parallelogram meshes by joining pairs of triangles, as is indicated for the hatched area. The parallelogram meshes can be mapped onto rectangular ones.

It must be emphasized that although the results will certainly become smoother, they are not necessarily more accurate.

A special problem arises if the FOFEM (Section 5.4) is applied to calculate rotationally symmetric fields or if the method of Section 5.4.4 is used for the computation of a magnetic lens: the use of linear trial functions is invalid in the vicinity of the optic axis. With respect to the solution of the FEM equation, this error is avoided by the modifications in Eqs. (5.76) and (5.77), but it is still present in interpolations. The easiest way to remove it is to use linear form functions in z and $s = r^2$ instead of (z, r) . This is necessary only in a narrow tube near the optic axis; the field then has the correct behavior.

5.6 SOLUTIONS OF LARGE SYSTEMS OF EQUATIONS

We now come to the techniques for the numerical solution of large systems of equations such as those derived in the earlier considerations, and which we shall encounter again in the context of the boundary-element method. This is a standard task in numerical analysis and is extensively dealt with in any comprehensive textbook, where program codes ready to use are even provided, as for example in *Numerical Recipes* [37]. Moreover, these programs are standard elements of program libraries like NAG, IMSL, LINPACK, etc. We can therefore keep the discussion here very short and concentrate more on the nonstandard methods.

In the following presentation, we shall designate matrices by capital letters and vectors by small ones, both in boldface type. More specifically, diagonal matrices will be denoted by $\mathbf{D}$, unit matrices by $\mathbf{I}$, lower triangular ones by $\mathbf{L}$, and upper triangular ones by $\mathbf{U}$, as is usual in the literature. Other types of matrices will be defined explicitly as necessary.

5.6.1 Direct Solution Methods

In the following discussion we consider linear systems of equations at the standard form

$$\mathbf{Ax} = \mathbf{b}, \quad (5.121)$$

and assume that the matrix $\mathbf{A}$ is nonsingular with full rank N . Systems with rank $N \leq 200$ can be considered as small ones and raise no particular problems with respect to memory, even if they are compact. If the system (5.121) is to be solved only once, then the familiar Gauss-elimination technique with row pivotation can be used, but care must be taken that the matrix is *well-conditioned*. Some rough control for this is the requirement that the pivot elements should not fall below a reasonably chosen threshold.

If the system (5.121) is to be solved several times, for example, if the inhomogeneity $\mathbf{b}$ consists in a linear superposition of several vectors, then the repeated application of the Gauss elimination is a waste of CPU, and it is then better to use the LU algorithm. This consists in a decomposition of the form

$$\mathbf{A} = \mathbf{L} \cdot \mathbf{U}, \quad (5.122)$$

with unit diagonal elements of the matrix $\mathbf{L}$. This needs to be performed only once. The trigonal matrices can then occupy the memory for $\mathbf{A}$ that is no longer needed. The solution consists now in the two steps

$$\mathbf{L} \cdot \mathbf{y} = \mathbf{b} \quad (5.123a)$$

in forward direction and

$$\mathbf{U} \cdot \mathbf{x} = \mathbf{y} \quad (5.123b)$$

as backward substitution. In linear combinations for $\mathbf{b}$ only the procedures (5.123) need to be repeated correspondingly many times. We define the *complexity* of an operation as the number of multiplications necessary to perform it and then give only the general values for $N \gg 1$. Then the complexity of the Gauss algorithm and of the LU decomposition (5.122) is $2N^3/3$, whereas that of Eqs. (5.123) is N^2 , which is then not important.

An important gain can be achieved if the matrix $\mathbf{A}$ is *symmetric* and *positive* definite. Then the familiar Cholesky algorithm can be applied. This is *not* the specialization of (5.122) for such symmetric matrices, although we still have

$$\mathbf{A} = \mathbf{L} \cdot \mathbf{U} = \mathbf{L} \cdot \mathbf{L}^T, \quad (5.124)$$

as earlier we had assumed that $L_{ii} = 1$, which clearly contradicts (5.124) but is still compatible with $U_{ii} \neq 1$ in (5.122). The Cholesky algorithm has the following important advantages, which makes it useful for large ranks N .

- (i) The pivotation is unnecessary, without it the algorithm is very stable for positive matrices.
- (ii) The *bandwidth* does not grow during the elimination, hence, with a bandwidth m , less than $m \cdot N$ elements need really to be stored.
- (iii) Apart from the additional requirement to calculate N square roots, the complexity of the Cholesky algorithm is practically half of that of the LU decomposition and correspondingly smaller for sparse matrices.

Although the stability of the algorithm is often quite sufficient, it can be further improved by appropriate scaling. This is here quite simple, as the matrix is not subject to permutations. The original system (5.121) is equivalent to the system

$$\mathbf{A}' \cdot \mathbf{x}' = \mathbf{b}', \quad (5.125a)$$

with a diagonal matrix D and

$$A' = D \cdot A \cdot D, \quad x' = D^{-1} \cdot x, \quad b' = D \cdot b. \quad (5.125b)$$

Any nonsingular matrix D can be chosen for this purpose, but the most favorable one is quite often defined by the demand that the diagonals of A' shall become unity:

$$D_i = (A_{ii})^{-1/2}, \quad (i = 1, \dots, N), \quad (5.126a)$$

$$A'_{ik} = A_{ik} D_i D_k \quad (i, k = 1, \dots, N). \quad (5.126b)$$

This requires only the calculation and storage of N additional square roots before the decomposition.

In the sense that a summation is dropped if the upper bound is less than the lower one, as is the case in modern programming languages, the whole algorithm can be written down very concisely. Although this would not be done in practical applications, we present here the decomposition (5.124) in its original form because this is easier to follow:

for $(k = 1, \dots, N)$

$$\left\{ L_{kk} = \left(A_{kk} - \sum_{j=1}^{k-1} L_{kj}^2 \right)^{1/2}, \right. \\ \left. \text{for } (i = k+1, \dots, N) \quad \left[L_{ik} = \left(A_{ik} - \sum_{j=1}^{k-1} L_{ij} L_{kj} \right) / L_{kk} \right] \right\}. \quad (5.127a)$$

The substitutions (5.123) are then analogously written as

$$\text{for } (i = 1, \dots, N) \quad \left\{ y_i = \left(b_i - \sum_{j=1}^{i-1} L_{ij} y_j \right) / L_{ii} \right\}, \quad (5.127b)$$

$$\text{for } (k = N, \dots, 1, \text{step} - 1) \quad \left\{ x_k = \left(y_k - \sum_{j=k+1}^N L_{jk} x_j \right) / L_{kk} \right\}. \quad (5.127c)$$

In this algorithm only the element A_{ik} with $i \geq k$ are explicitly used and then appear only as start terms in the corresponding summations. It is hence possible to overwrite them. Moreover, it turns out to be favorable to store these matrices sequentially in a *row-wise* order, as will be explained further below. This is of great importance with respect to the calculation of splines (see Section 3.2.3) and in the applications to the systems arising in the FDM or

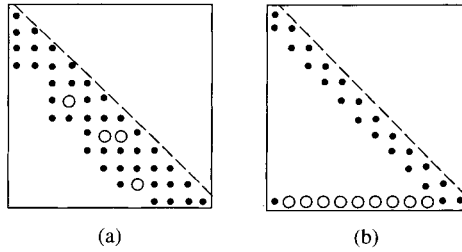


FIGURE 5.33 Examples of the L matrices corresponding to sparse symmetric matrices of rank $N = 12$. Occupied elements are marked by dots, fill-ins by little circles. (a) General case with maximum band width 5 and 45 elements; (b) matrix for a periodical spline with 33 memory locations. The memory-saving effect is more pronounced for large ranks N .

the FEM. The matrix A is then *sparse*, which leads to the following situation (see Fig. 5.33).

The number N of unknown (e.g., potentials at the nodes) is quite huge, often $N > 10,000$. However, each unknown value x_i is linearly connected to only a few other variables x_j within the same row. The *bandwidth* of the row with label i is defined as

$$\beta(i) = 1 + \max |i - j| \quad \text{with } A_{ij} \neq 0. \quad (5.128)$$

Fill-ins, as shown in Fig. 5.33 must be taken into account with respect to memory.

This definition (5.128) implies that the positions $1 \leq k \leq i - \beta_i$ remain completely empty, whereas the positions $i + 1 - \beta_i \leq k \leq i$ may be filled in, so that the corresponding memory must be allocated. A *compact* sequential ordering is then obtained by the following technique. First the labels d_i of the diagonal elements are found as

$$d_1 = 1, \quad d_i = d_{i-1} + \beta_i, \quad (2 \leq i \leq N), \quad (5.129a)$$

so that $M = d_N$ is the length of the array $\mathbf{a}$, which must be allocated. Then the matrix is to be stored according to

$$A_{ik} \triangleq a[d_i + k - i], \quad (i + 1 - \beta_i \leq k \leq i) \quad (5.129b)$$

All that remains to be done is to rewrite Eqs. (5.127) in terms of this array $\mathbf{a}$, A_{ik} and L_{ik} occupying the same location. This is an elementary task that is not presented here and is favorably done by the computer. Care must be taken that the summation indices do not extend beyond their allowed ranges.

The efficient use of this shell-oriented calculation technique requires that the sequence of variables (the nodes in a FEM program) is initially reordered

in such a manner that the total length M is minimized by optimizing the distribution of bandwidths β_i . Different strategies are known for this purpose [38, 39], but for reasons of space we cannot outline these here.

A combination of these techniques is the so-called Gauss-Cholesky algorithm, in which no square roots need to be calculated. It is then also sufficient to assume a symmetric matrix A ; its positiveness is not required. Again a pivotation is not necessary, and the decomposition now takes the form

$$A = L \cdot D \cdot L^T, \quad U = D \cdot L^T, \quad (5.130)$$

where $L_{ii} = 1$ is assumed here. For any nonsingular symmetric matrix, this decomposition is then unique. A disadvantage is that the complexity is again $2N^3/2$, as for the LU algorithm, so that the gain obtained by avoiding the square root calculation is soon overcompensated for large systems, and moreover, the condition of the process is worse. The Gauss-Cholesky algorithm is therefore recommended only for small and stable systems like the *tridiagonal* systems for splines.

Such a system can be cast in the general form

$$a_k x_{k-1} + b_k x_k + c_k x_{k+1} = g_k, \quad (k = 1, \dots, N) \quad (5.131)$$

with $a_1 = 0$ and $c_N = 0$. The system is symmetric if $c_k = a_{k+1}$, but this restriction is not absolutely necessary if the common Gauss algorithm is used. A sufficient criterion for a regular system is the dominance of diagonals:

$$|b_k| > |a_k| + |c_k|, \quad (k = 1, \dots, N); \quad (5.132)$$

then, even when $c_k \neq a_{k+1}$, the pivotation is unnecessary, and the common straightforward elimination procedure now results in

$$b'_1 = b_1, \quad b'_k = b_k - a_k c_{k-1} / b'_{k-1}, \quad (k = 2, \dots, N), \quad (5.133a)$$

$$g'_1 = g_1, \quad g'_k = g_k - a_k g_{k-1} / b'_{k-1}, \quad (k = 2, \dots, N), \quad (5.133b)$$

$$x_N = g'_N / b'_N, \quad x_k = (g'_k - c_k x_{k+1}) / b'_k, \quad (k = N-1, N-2, \dots, 1). \quad (5.133c)$$

If only the right-hand side of Eq. (5.131) is altered, as for example in applications to two-dimensional splines, then Eq. (5.133a) must be carried out once and only Eqs. (5.133b,c) need to be repeated. If the original coefficients in Eq. (5.131) are not needed any longer, it is possible to overwrite them, replacing b_k with b'_k and g_k with g'_k . This is an example of a situation in which the familiar Gauss algorithm already has all favorable properties and hence it makes little sense to apply more sophisticated procedures.

There are many more techniques in the field of linear algebra, the most powerful and sophisticated one being singular value decomposition. For reasons of space we cannot outline this here and refer to *Numerical Recipes* [37], where codes ready to use can be found.

5.6.2 The Conjugate Gradient Method

The direct solution techniques are applicable as long as it is possible to hold the relevant parts of the matrix, including the necessary fill-in elements, in the working memory. For huge systems with many fill-in spaces, this may become impractical. The iterative solution techniques to be dealt with now do not have this disadvantage as they use only the nonvanishing parts of the system matrix. However, they may have the disadvantage that the total amount of calculations can increase drastically.

The conjugate gradient (CG) method is such an iterative technique, which works most favorably if a linear system of the form (5.121) with *positive* matrix A is to be solved because then this solution is equivalent to the minimization problem [40]:

$$f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \cdot A \cdot \mathbf{x} - \mathbf{b} \cdot \mathbf{x} = \min. \quad (5.134)$$

Fortunately, this assumption does hold for linear boundary value problems within the FEM and for many others within the FDM. The calculation of magnetic lenses with saturation effects (Section 5.4.4) is a nonlinear problem, but, nevertheless, the CG method can be used, provided that an approximation of the form of (5.134) can be found at least locally for each iteration. The number of necessary iterations increases then with increasing saturation effect.

Because it is well known that the method of steepest descent towards the minimum is not the best choice, as is demonstrated in Fig. 5.34, the idea is here to choose successive directions $\mathbf{p}_k$ and $\mathbf{p}_{k+1}$ of search that are conjugate to each other, which means here

$$\mathbf{p}_{k+1}^T \cdot A \cdot \mathbf{p}_k = 0. \quad (5.135)$$

For reasons of conciseness we drop here the transposition symbol on the vector $\mathbf{p}_{k+1}$ as $A \cdot \mathbf{p}_k$ is an ordinary vector, and the expression in Eq. (5.135) is an ordinary scalar product. The algorithm can be cast in the following form:

Start

Choose an initial position $\mathbf{x}_0$ and a value $\varepsilon \ll 1$.

Calculate $\mathbf{p}_0 = \mathbf{r}_0 = \mathbf{b} - A \cdot \mathbf{x}_0$

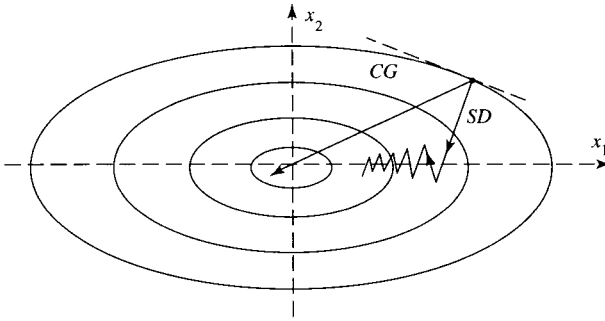


FIGURE 5.34 The methods of steepest descent (SD) and conjugate gradients (CG) for a function with nearly elliptical lines of constant value. Whereas the SD carries out many oscillations near the bottom of the valley, the CG aims directly at a point close to the minimum and would reach it exactly in one step, if the function were exactly quadratic.

$$\begin{aligned}
 &\text{for } (k = 0, 1, \dots, k_{\max}) \\
 &\{ \quad \text{if } (|\mathbf{p}_k| < \varepsilon) \text{ stop} \quad \mathbf{x}_k \text{ is the solution.} \\
 &\quad \text{else} \\
 &\quad [\quad a_k = |\mathbf{r}_k|^2 / \mathbf{p}_k \cdot \mathbf{A} \cdot \mathbf{p}_k, \quad \mathbf{x}_{k+1} = \mathbf{x}_k + a_k \mathbf{p}_k, \\
 &\quad \quad \mathbf{r}_{k+1} = \mathbf{r}_k - a_k \mathbf{A} \cdot \mathbf{p}_k, \quad c_k = |\mathbf{r}_{k+1}|^2 / |\mathbf{r}_k|^2, \\
 &\quad \quad \mathbf{p}_{k+1} = \mathbf{r}_{k+1} + c_k \mathbf{p}_k \quad]. \\
 &\}
 \end{aligned} \tag{5.136}$$

This can be interpreted in the following manner. The vectors $\mathbf{p}_k$ and $\mathbf{p}_{k+1}$ are the conjugate gradients in the sense of Eq. (5.135); the scalars a_k are then the step sizes, which should decrease in the course of the iterations. The vectors $\mathbf{r}_k$ and $\mathbf{r}_{k+1}$ are the *residuals*.

Generally, the following properties of this algorithm can be proven by induction:

$$\begin{aligned}
 &\text{for all } i < k < N : \\
 &\quad \{ \mathbf{p}_i \cdot \mathbf{r}_k = 0, \quad \mathbf{p}_k \cdot \mathbf{r}_k = |\mathbf{r}_k|^2, \\
 &\quad \mathbf{p}_i \cdot \mathbf{A} \cdot \mathbf{p}_k = 0, \quad \mathbf{p}_k \cdot \mathbf{A} \cdot \mathbf{p}_k > 0, \\
 &\quad \mathbf{r}_i \cdot \mathbf{r}_k = 0, \quad \mathbf{r}_k = \mathbf{b} - \mathbf{A} \cdot \mathbf{x}_k \}.
 \end{aligned} \tag{5.137}$$

From the last two relations it becomes obvious that the residuals successively build up an orthogonal basis. Hence, in the absence of nonlinearities and rounding errors, the algorithm must stop after N iterations with $\mathbf{p}_N = \mathbf{r}_N = 0$. However, this will not be reached exactly. Then the last set $\mathbf{p}_N$, $\mathbf{r}_N$, and $\mathbf{x}_N$ are to be used as start set for a next cycle, until the accuracy criterion is satisfied.

The CG algorithm is presented here in its correct mathematical form, but of course, a practical program would not look like this. With some tricks, apart from a few scalar data, only the last updates of the vectors $\mathbf{x}$ and $\mathbf{p}$ need really to be stored. It is essential that the matrix $\mathbf{A}$ itself is never required

explicitly; it is sufficient to calculate the vector $\mathbf{A} \cdot \mathbf{p}_k$ in an efficient manner, which is equivalent to evaluating local mesh formulas. Thus, the algorithm can be made very efficient.

Sometimes it is advantageous or even necessary to improve the condition of the matrix $\mathbf{A}$, to achieve a reasonably fast convergence of the CG algorithm. This can be done with a generalization of Eqs. (5.125), $\mathbf{A} \cdot \mathbf{x}' = \mathbf{b}'$, but now with

$$\mathbf{A}' = \mathbf{H}\mathbf{A}\mathbf{H}^T, \quad \mathbf{x}' = \mathbf{H}^{-1} \cdot \mathbf{x}, \quad \mathbf{b}' = \mathbf{H} \cdot \mathbf{b}. \quad (5.138)$$

This procedure is called a *preconditioning*, and in principle, any nonsingular matrix $\mathbf{H}$ is chosen, which leads to an improvement of the condition. The optimum choice would be $\mathbf{H} = \mathbf{L}^{-1}$ with $\mathbf{L}$ given by Eq. (5.124) because this would result in $\mathbf{A}' = \mathbf{T}$, but this choice is not feasible. Therefore, an *incomplete* Cholesky decomposition is made, which consists in a procedure similar to that of Eq. (5.127a) but with suppression of all fill-in elements for reasons of memory. This can result in singular matrices, which must be circumvented by suitable tricks. We cannot outline these here and refer to the corresponding literature [14, 40].

5.6.3 Relaxation Methods

These are iterative techniques that belong to the standard methods in numerical analysis and are investigated there in detail [41–43]. We therefore keep this presentation fairly short and give the important relations without proof. Thus we do not discuss here the oldest methods, namely, the Jacobian and Gauss-Seidel methods.

The system (5.121) must be rewritten as

$$\mathbf{x} = \mathbf{C} \cdot \mathbf{x} + \mathbf{q}, \quad (5.139)$$

which means, partly ‘solved’ for $\mathbf{x}$. In the practical applications to FDM procedures, this means solution for the central potential value in each five- or nine-point configuration. The successive over relaxation (SOR) consists then in the iteration procedure

$$\mathbf{x}^{(n+1)} = \mathbf{R}(\omega) \cdot \mathbf{x}^{(n)} + \mathbf{q}'(\omega), \quad (5.140)$$

with the iteration matrix $\mathbf{R}(\omega)$, ω being a free parameter. The appropriate choice of the latter is a critical task, which will be discussed further below. The matrix $\mathbf{C}$ in (5.139) is simply decomposed by writing $\mathbf{C} = \mathbf{L} + \mathbf{U}$ with *vanishing* diagonal. Then the iteration matrix $\mathbf{R}$ is to be defined as

$$\mathbf{R}(\omega) = (\mathbf{I} - \omega\mathbf{L})^{-1} \cdot [(1 - \omega)\mathbf{I} + \omega\mathbf{U}] \quad (5.141a)$$

and the inhomogeneity as

$$\mathbf{q}'(\omega) = \omega[\mathbf{I} - \omega\mathbf{L}]^{-1} \cdot \mathbf{q}. \quad (5.141b)$$

This is a formal representation that is impractical; instead, the inverse matrix factor is avoided by means of the partly implicit algorithm

$$\mathbf{x}^{(n+1)} = (1 - \omega)\mathbf{x}^{(n)} + \omega[\mathbf{U} \cdot \mathbf{x}^{(n)} + \mathbf{L} \cdot \mathbf{x}^{(n+1)} + \mathbf{q}]. \quad (5.142)$$

The practical realization with controls by the maximum norm Δ and the sum norm σ can be cast in the following form

Allocation of arrays $\mathbf{x}$ and $\mathbf{q}$	
Choice of error limit ε and maximum $n_{\max}$	
Input or calculation of vector $\mathbf{q}$	
Initial guess for vector $\mathbf{x}$, $\sigma = 0$	
for ($n = 1, 2, \dots; n \leq n_{\max}$)	
{	$\Delta = 0$
	Choice of ω (depending on σ and n), $\sigma = 0$
	for ($j = 1, \dots, N$)
[	
	$S = \sum_{k=1}^{j-1} C_{jk}x_k + \sum_{k=j+1}^N C_{jk}x_k + q_j \quad (5.143)$
	$d = S - x_j$ (Local deviation)
	$\Delta = \max(\Delta, d)$ (Maximum-norm)
	$\sigma = \sigma + d $ (Sum-norm)
	$x_j = x_j + \omega d$ (Overrelaxation)
]	(End of inner loop)
	Control output of n , Δ , σ
	if ($\Delta < \varepsilon$) stop (Convergence)
{	
	(End of outer loop)
	if ($\Delta \geq \varepsilon$) $\longrightarrow$ Error message

In this form, the SOR has a fairly general applicability, not only in the FDM but also in the FEM. Strictly the quantities Δ , σ , S , and D should have labels,

but we have dropped these because in the real program these labels are indeed unnecessary. The calculation of S is nothing but the evaluation of the right-hand side of Eq. (5.139) with the currently available vector components. As with the CG algorithm, it is here not necessary to have the matrix C in its complete form; it is quite sufficient to be able to evaluate the corresponding mesh formula numerically in some efficient way.

The maximum norm Δ being the worst error, is a sensitive criterion for convergence; however, $\Delta < \varepsilon$ does not also imply $|\mathbf{x} - \mathbf{x}_E| < \varepsilon$, $\mathbf{x}_E$ being the unknown exact solution. Because Δ fluctuates strongly, at least in the beginning, this is an unsuitable measure for the determination of ω , and the smoother sum norm σ is therefore calculated too.

We come now to the discussion of the convergence properties of the SOR. A sufficient criterion is the dominance of diagonals. With $C_{jj} = 0$ and the prime indicating that this term is to be skipped, this means

$$\sum_{k=1}^N |C_{jk}| \leq 1 \quad (j = 1, \dots, N). \quad (5.144)$$

Then for $1 < \omega < 2$ a convergence faster than that of the Gauss-Seidel algorithm can be achieved. There is a critical value ω_{opt} , for which the damping factor $a = \sigma^{(n+1)}/\sigma^{(n)}$ has a sharp minimum, but unfortunately it is difficult to determine this value ω_{opt} . The importance of this choice is illustrated in Figs. 5.35a and 5.35b. The forms of these graphs are valid, if the matrix C has only *real* eigenvalues. A sufficient criterion for this is that the matrix C has the so-called ‘property A’, but often it is difficult to verify this theorem, and the SOR still works quite well.

The iteration process in Eq. (5.140) can be interpreted as the superposition of the required stationary solution $\mathbf{x}^{(\infty)}$ with all damped eigenvectors of the matrix R . If we interpret the number n as a time-like parameter increasing

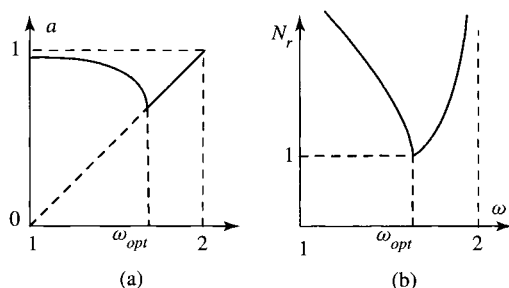


FIGURE 5.35 Properties of the SOR and the SLOR: (a) damping factor a as a function of ω ; (b) relative number N_r to reach a given threshold as a function of ω . Both functions have a sharp minimum at a well-specified value $\omega = \omega_{\text{opt}}$.

in unit steps, then this damping proceeds exponentially but with different attenuation constants for the different eigenmodes. Asymptotically the speed of convergence is determined by the dominant eigenmode, the one with the *largest* eigenvalue λ_1 of $\mathbf{C}$. Then the corresponding eigenvalue μ_1 of $\mathbf{R}$ is related to this by

$$\omega^2 \lambda_1 = (\mu_1 + \omega - 1)^2 / \mu_1, \quad (\omega \leq \omega_{\text{opt}}). \quad (5.145)$$

This equation has two different solutions for μ_1 and ω_{opt} is just that value at which these two solutions become equal:

$$\omega_{\text{opt}} = \frac{2}{1 + \sqrt{1 - \lambda_1^2}} \quad (5.146)$$

For $\omega < \omega_{\text{opt}}$, the damping factor is now $a = \max |\mu_1^\pm|$, whereas for $\omega > \omega_{\text{opt}}$ the two μ_1 values become conjugate complex and we have then $a = \omega - 1$. Hence, the optimum is just

$$a_{\text{opt}} = \omega_{\text{opt}} - 1 \quad (5.147)$$

as shown in Fig. 5.35a.

In general cases, it is difficult to determine the value λ_1^2 , to be introduced into Eq. (5.146). An empirical way, followed by Carré [44] and Winslow [45] is to determine the damping factor $a = |\mu_1|$ from a sequence of rates $\sigma^{(n)} / \sigma^{(n-1)}$, which can then be introduced on the right-hand side of Eq. (5.145). This must be performed with great care, as the influence of higher, not sufficiently damped eigenmodes may lead to wrong guesses and thus to instabilities.

Another way is to calculate λ_1 as a *Rayleigh quotient*. If $\mathbf{x} = \mathbf{e}_1$ is an eigenvector of the matrix $\mathbf{C}$ for the eigenvalue λ_1 , then the latter satisfies exactly

$$\lambda_1 = \mathbf{e}_1^T \cdot \mathbf{C} \cdot \mathbf{e}_1 / \mathbf{e}_1^2 \cdot \mathbf{e}_1 =: Z_1 / Z_2. \quad (5.148)$$

Both λ_1 and $\mathbf{e}_1$ can be determined iteratively. It is not necessary to normalize $\mathbf{e}_1$ to unity. In applications to potential arrays as practical realizations of $\mathbf{e}_1$, these must have vanishing boundary values and correspond roughly to the standing wave inside the domain of solution. Because the iteration process for the eigenvectors converges always to that of the absolutely smallest eigenvalue and λ_1 will be the largest one, the parameter $1/\lambda_1$ must here be used. The elementary procedure tends to oscillations in λ , which can be eliminated by averaging. Then an improved version has the following form

$$\begin{aligned}
& \text{Initial guess for vector } \mathbf{x} \approx \mathbf{e}_1 \\
& \lambda_1 = \Lambda_1 = \Lambda_2 = 1. \\
& \text{for } (n = 1, 2, \dots; \quad n \leq \tilde{n}_{\max}) \quad \text{(Iteration)} \\
& \left\{ \begin{aligned} & Z_1 = Z_2 = 0.; \\ & \text{for } (j = 1, \dots, N) \\ & \left[\begin{aligned} & S = \sum_{k=1}^{j-1} C_{jk} x_k + \sum_{k=j+1}^N C_{jk} x_k \quad \text{(Mesh-formula)} \\ & Z_1 = Z_1 + S x_j \quad \text{(Numerator)} \\ & Z_2 = Z_2 + x_j^2 \quad \text{(Denominator)} \\ & x_j = S / \lambda_1 \quad \text{(Relaxation)} \end{aligned} \right. \\ & \left. \begin{aligned} & \Lambda_1 = \Lambda_2, \quad \Lambda_2 = Z_1 / Z_2 \quad \text{(Eq. (5.148))} \\ & \lambda_1 = (\Lambda_1 + \Lambda_2) / 2. \quad \text{(Averaging)} \end{aligned} \right. \\ & \text{if } (\text{abs } (\Lambda_1 - \Lambda_2) < \tilde{\varepsilon}) \text{ break loop} \quad \text{(Ready)} \end{aligned} \right\} \quad \text{(End of outer loop)} \\
& \omega_{\text{opt}} = 2. / \left(1. + \sqrt{(1. - \lambda_1^2)} \right) \quad \text{(Result)} \quad (5.149)
\end{aligned}$$

The number of algebraic operations in the inner loop is not much greater than in the Eq. (5.143). The essential gain lies in the fact that the Rayleigh quotient converges very much faster to its final value than does the eigenvector. Hence, with $\tilde{\varepsilon} = 10^{-5}$ and $\tilde{n}_{\max} = 20$, the eigenvalue λ_1 and correspondingly ω_{opt} can be determined with sufficient accuracy and in a reliable manner. Moreover, with some refinements, this method provides also a way to calculate the ground state of a resonator by means of the FDM.

5.6.4 Successive Line Overrelaxation

The main advantage of the SOR is its simple formalism, which can easily be coded. However, the condition (5.144) is a strong restriction on its applicability, which is easily seen if we interpret the vector $\mathbf{x}$ as the representative of a potential field. Because a *constant* potential must always be possible as solution of Laplace's equation, this implies then $x_k = \text{const.}$, ($k = 1, \dots, N$)

and then further

$$\sum_{k=1}^N{}' C_{jk} = 1, \quad (j = 1, \dots, N). \quad (5.150)$$

The mesh formulas derived in the frame of the FDM and the FEM are consistent with this condition after appropriate normalization. But Eq. (5.150) together with Eq. (5.144) implies that all C -coefficients must never be negative. This latter and stronger restriction is *not* always satisfied. Convergence can then be achieved with local underrelaxation ($\omega < 1$), but that is not attractive because it slows down the procedure.

A possible solution of this problem is the use of successive *line* iteration (SLOR). This technique belongs to the family of block iteration methods as it combines direct solutions in partitions of the matrix A with subsequent overrelaxations. The condition (5.144) is then not necessary; moreover, the stability of the SLOR is better and the convergence a little faster. It is now not necessary to normalize the central mesh coefficient to unity, as was done in Eq. (5.139), though this might still be favorable. We shall present here the SLOR not in its general abstract form but more practically in its application to the solution of the still fairly general PDE (4.161). For reasons of conciseness, we exclude here the numerical treatment of irregular meshes, which must certainly be done in a realistic program.

The total procedure to be carried out consists in a preparative part, the SLOR in the proper sense and the back transformation; three two-dimensional arrays are needed simultaneously to carry out all operations efficiently. In the preparative part the given PDE (4.161) is transformed to the cylindric Poisson equation for the potential $W(u, v)$. It is advantageous to rewrite Eqs. (4.163) and Eq. (4.164a) in the form

$$G(u, v) := h^2 g(u, v)/12, \quad W = V + G. \quad (5.151)$$

By means of Eq. (4.164b), it is then possible to cast these relations in the form

$$G_{ik} = C_{ik} W_{ik} + S_{ik}, \quad (5.152)$$

with the coefficients

$$d := (12/h^2 + \hat{q}_{ik})^{-1}, \quad C_{ik} = \hat{q}_{ik} d, \quad S_{ik} = p_{ik} s_{ik} d. \quad (5.153)$$

The arrays $[C_{ik}]$ and $[S_{ik}]$ must be stored because they are needed quite frequently. The mesh formulas (4.140) and (4.145) are favorably cast in

the form

$$W_{i,0} = A_{0,1}W_{i,1} + \sum_{j=0}^1 B_{0,j}(W_{i-1,j} + W_{i+1,j}) + A_{0,0}G_{i,0} + \gamma_s(G_{i,0} - G_{i,1}) \quad (5.154a)$$

$$W_{i,k} = A_{k,-1}W_{i,k-1} + A_{k,1}W_{i,k+1} + A_{k,0}G_{i,k} + \sum_{j=-1}^1 B_{k,j}(W_{i-1,k+j} + W_{i+1,k+j}) \quad (k \geq 1). \quad (5.154b)$$

The coefficients are obtained by appropriate normalization of those in Section (4.4.2). Their storage requires only $6N + 1$ real data, N being the maximum mesh size in the radial direction. (In program languages that do not allow negative subscripts, the second one in A and B must increased by $+1$).

The SLOR can now start. Its basic idea is shown in Fig. 5.36. In each column $i = \text{const.}$ of the mesh, the corresponding potentials W_{ik} depend on each other through a tridiagonal system of linear equations, the potentials in the neighboring columns ($i - 1$) and ($i + 1$) being treated as temporarily given. This tridiagonal system can easily be solved by the Gauss algorithm. The results are then overrelaxed, and the procedure steps to the next column. The whole process is repeated until it has converged sufficiently.

The algorithm requires an additional array Y of length $N + 1$ for the temporary storage of the results obtained by the Gauss procedure. These cannot overwrite immediately the old values, as the latter are still needed for the overrelaxation. Another array D of this length is necessary for the transformed diagonals of the tridiagonal system. Equations (5.152) and (5.154) must now be cast in a shape corresponding to Eq. (5.131). Then the terms with

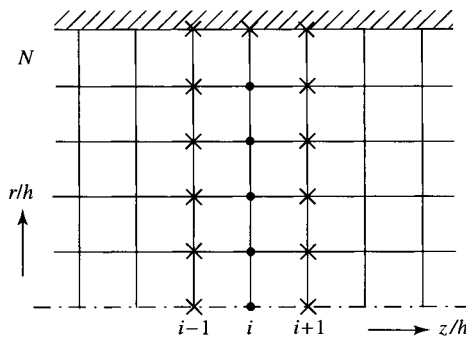


FIGURE 5.36 Schematic presentation of the SLOR: the points, marked by dots, contribute to the tridiagonal system of equations, whereas the next neighbors, marked by crosses, contribute to the driving terms.

B -coefficients, with factor γ_s and with sources altogether form the right-hand side, and we have to identify in turn

$$a_k = -A_{k,-1}, \quad b_k = 1 - A_{k,0}C_{ik}, \quad c_k = -A_{k,1}. \quad (5.155)$$

It is favorable to make use of $D_k = 1/b'_k$ because then only one division per step is necessary. In this way, we obtain the Eq. (5.156). It is presented here, starting at the optic axis $k = 0$ and with a variable upper end $e = k_E(i)$. Of course, it is possible to start at a position $k_a > 0$ if necessary. The normalization in Eq. (5.154) is most favorable as the coupling $Y_e = W_{i,e}$ to the boundary value is the simplest possible one. If this is not directly given but is to be determined by other procedures, this has to be done outside the scheme but still within the whole iteration process.

$$\begin{aligned}
 & \text{(Sweep over the column with label } i) \\
 e &= k_E(i) && \text{(Length of this column)} \\
 D_0 &= 1./ (1. - A_{0,0}C_{i,0}) \\
 Y_0 &= \sum_{j=0}^1 B_{0,j}(W_{1-1,j} + W_{1+1,j}) + A_{0,0}S_{i,0} \\
 & \quad + \gamma_s(C_{i,0}W_{i,0} + S_{i,0} - C_{i,1}W_{i,1} - S_{i,1}) \\
 \text{for } (k &= 1, \dots, e-1) && \text{(Elimination loop)} \\
 \{ \quad t &= A_{k-1}D_{k-1} \\
 D_k &= 1./ (1. - A_{k,0}C_{i,k} - tA_{k-1,1}) \\
 Y_k &= \sum_{j=-1}^1 B_{k,j}(W_{i-1,k+j} + W_{i+1,k+j}) + A_{k,0}S_{ik} + tY_{k-1} \\
 \} & && \text{(End of elimination)} \\
 Y_e &= W_{i,e} && \text{(Known boundary value)} \\
 \text{for } (k &= e-1, \dots, 0, \text{ step } -1) && \text{(Resolution loop)} \\
 \{ \quad Y_k &= D_k(Y_k + A_{k,1}Y_{k+1}) \\
 d &= Y_k - W_{i,k} && \text{(Local deviation)} \\
 \Delta &= \max(\Delta, |d|) && \text{(Maximum-norm)} \\
 \sigma &= \sigma + |d| && \text{(Sum-norm)} \\
 W_{i,k} &= W_{i,k} + \omega d && \text{(Overrelaxation)} \\
 \} & && \text{(End of resolution)} \quad (5.156)
 \end{aligned}$$

This scheme being the central part of a SLOR program has to be incorporated in a double loop. The outer one counts and controls the iterations, as in

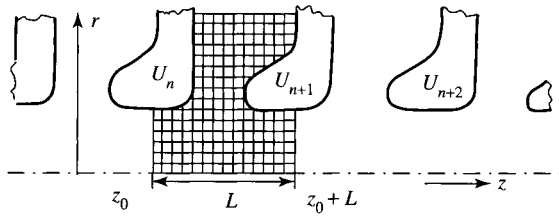


FIGURE 5.37 Half axial section through a periodic system.

Eq. (5.143), whereas the inner one runs over all columns $i = 1, \dots, M - 1$ of the mesh.

It is fairly easy to incorporate *periodic* boundary conditions in this scheme, as shown in Fig. 5.37, showing an electrostatic linear accelerator as an example. This is even *much simpler* than the corresponding task in direct solution techniques. The potential $W(z, r)$ then satisfies the condition

$$W(z + nL, r) = W(z, r) + n(U_1 - U_0), \quad (n = 0, \pm 1, \pm 2, \dots). \quad (5.157)$$

This can be used to reduce the field calculation to one significant period. The periodicity condition appears in the driving terms, where it is used to eliminate the nonallocated array elements for $i = -1$ or $i = M$, referring to the columns $z = z_0 - h$ and $z = z_0 + L$, respectively. This is to be done in each iteration cycle. The periodicity also modifies the optimum relaxation parameter, so that it must be considered in the Rayleigh-quotient procedure.

Certainly, no device is exactly periodic, but it is reasonable to determine first the field in one period and then use this knowledge of the boundary values of the field in the entrance and exit region. We now continue the general considerations.

When convergence has been achieved, the original potential values are to be determined according to the back transformation formula (4.165), here now rewritten as

$$\phi_{ik} = [W_{ik}(1 - C_{ik}) - S_{ik}]/P_{ik}, \quad (5.158)$$

whereupon the calculations are finished. If the calculation of the coefficients P_{ik} requires a two-dimensional field, then first the numerator is calculated, giving the potentials V_{ik} ; thereafter, one of the now unnecessary arrays is used for P_{ik} and then $\phi_{ik} = V_{ik}/P_{ik}$ as the last step.

The convergence properties of the SLOR are similar to those of the SOR, which means that Eqs. (5.145–5.147) are similarly valid. But the eigenvalue λ_1 is now slightly smaller, so that the SLOR needs only about 70% of the iterations necessary with the SOR. Hence, the larger algebraic effort finally pays off. Moreover, the SLOR has two additional advantages. First, as already mentioned; its stability is better; for example, multipole fields with $\alpha > 3$ are

easily calculated. The second advantage is that the residual after reaching the threshold ε is significantly less than with the SOR.

There are more powerful iteration techniques like the ADI [46], the strongly implicit methods [47], and the cyclic reduction methods [48]. These often require special conditions that do not hold generally, for instance, a rectangular domain of solution, and are therefore not outlined here.

5.6.5 Nonlinear Systems of Equations

These arise in the context of applications of the FEM to configurations with nonlinear material properties, such as the hysteresis in ferromagnetic yokes. We shall briefly outline here the general formalism and then turn to this special problem. In the following we assume vectors $\mathbf{x} = (x_1, \dots, x_N)$ and smoothly differentiable functions $f_1(\mathbf{x}), \dots, f_N(\mathbf{x})$ in an N -dimensional space. The problem to be solved can then be written as

$$f_j(\mathbf{x}) = c_j, \quad (j = 1, \dots, N) \quad (5.159)$$

with initially unknown vector $\mathbf{x}$ and given $c_1, \dots, c_N$.

The classical method is Newton's multidimensional iteration method. Let $\mathbf{x}_a$ be an approximate solution. Then a Taylor series expansion, truncated after the linear terms, results in

$$f_j(\mathbf{x}_a + \delta\mathbf{x}) = f_j(\mathbf{x}_a) + \delta\mathbf{x} \cdot \text{grad } f_j(\mathbf{x}_a) \Rightarrow c_j. \quad (5.160)$$

This can be solved for the necessary shift, which implies the solution of the linear system

$$\mathbf{J} \cdot \delta\mathbf{x} = -\mathbf{r}, \quad (5.161a)$$

$$\mathbf{r} = (f_1(\mathbf{x}_a) - c_1, \dots, f_N(\mathbf{x}_a) - c_N)^T, \quad (5.161b)$$

being the residual vector and $\mathbf{J}$ the Jacobian matrix with elements

$$J_{ik}(\mathbf{x}_a) = \partial f_i(\mathbf{x}_a) / \partial x_k. \quad (5.162)$$

We must assume here that this matrix is nonsingular and that its spectral radius, the largest absolute eigenvalue, is less than unity. Then the iteration of these equations converges.

Quite often, the system (5.159) results from the minimization of a functional or a scalar function $F(\mathbf{x})$. We then have

$$f(\mathbf{x}) = \partial F / \partial x_j \quad (j = 1, \dots, N). \quad (5.163)$$

The Jacobian then becomes the Hesse matrix of $F(\mathbf{x})$, with elements

$$J_{ik} = \partial^2 F / \partial x_i \partial x_k, \quad (5.164)$$

which are evidently symmetric. In the case of a minimum, this matrix is also positive, so that the Cholesky decomposition $\mathbf{J} = \mathbf{L} \cdot \mathbf{L}^T$ is applied. In this context, it is not necessary to use the form of Eq. (5.164), as (5.162) is equivalent. It suffices that the Cholesky algorithm be used instead of the LU technique. A further essential gain is achieved if the residuals become so small that a recalculation and new decomposition of the Jacobian is practically unnecessary: the sole iteration of Eq. (5.123) in combination with the recalculation of residuals can speed up the procedure significantly.

Apart from the different notation, the system (5.87) already has this form, and the matrix elements L_{ik} in Eq. (5.88) are then interpreted as those of the Jacobian. Owing to the strong nonlinearity in the reluctivity ν , the iteration process here *cannot* be enhanced; each cycle requires a complete recalculation of the L_{ik} . The same holds for the CG method. For more details, we refer to the corresponding literature [49, 50].

REFERENCES

1. Norrie, D. H. and de Vries, G. (1973). *The Finite Element Method*, London & New York: Academic Press.
2. Fenner, R. T. (1975). *Finite Element Methods for Engineers*, London & New York: Macmillan.
3. Bathe, K. J. and Wilson, E. L. (1976). *Numerical Methods in Finite Element Analysis*, Englewood Cliffs: Prentice Hall.
4. Mitchell, A. R. and Wait, R. (1977). *The Finite Element Method in Partial Differential Equations*, London: Wiley.
5. Zienkiewicz, O. C. (1977). *The Finite Element Method*. London & New York: McGraw-Hill.
6. Chari, M. V. and Silvester, P. P. (1980). *Finite Elements in Electrical and Magnetical Field Problems*, New York: Wiley.
7. Brebbia, C. A., ed. (1982). *Finite Element Systems, A Handbook*, Berlin, Heidelberg, New York: Springer.
8. Schwarz, H. R. (1983). *Methode der Finiten Elemente*, Stuttgart: Teubner.
9. Munro, E. (1971). Computer-Aided Design Methods in Electron Optics, Dissertation, University of Cambridge, UK.
10. Munro, E. (1973). Computer-aided design of electron lenses by the finite element method. In *Image Processing and Computer-Aided Design in Electron Physics*, P. W. Hawkes, ed., pp. 284–323, London & New York: Academic Press.
11. Munro, E. (1987). Computer programs for the design and optimization of electron and ion beam lithography systems, *Nucl. Instrum. Meth. A* **258**: 443–461.
12. Zhu, X. and Munro, E. (1995). Second-order finite element method and its practical application in charged particle optics, *J. Microscopy* **179**: 170–180.
13. Lencova, B. (1980). Numerical computation of electron lenses by the finite element method, *Comp. Phys. Commun.* **20**: 127–132.

14. Lencova, B. and Lenc, M. (1986). *A Finite Element Method for the Computation of Magnetic Electron Lenses*, Scanning Electron Microscopy, Chicago, pp. 897–915.
15. Lencova, B. and Wisselink, G. (1990). Program package for the computation of lenses and deflectors, *Nucl. Instrum. Meth. A* **298**: 56–66.
16. Lencova, B. (1995a). Unconventional lens computation, *J. Microscopy* **179**: 185–190.
17. Lencova, B. (1995b). Computation of electrostatic lenses and multipoles by the first order finite element method, *Nucl. Instrum. Meth. A* **363**: 190–197.
18. Lencova, B. (1999). Accurate computation of magnetic lenses with FOFEM, *Nucl. Instrum. Meth. A* **427**: 329–337.
19. Adamec, P., Delong, A., and Lencova, B. (1995). Miniature magnetic electron lenses with permanent magnets, *J. Microscopy* **179**: 129–132.
20. Barth, J. E., Lencova, B., and Wisselink, G. (1990). Field evaluation from potentials calculated by the finite element method: the slice method, *Nucl. Instrum. Meth. A* **298**: 263–268.
21. Mulvey, T. and Nasr, H. (1981). An improved finite element method for calculating the magnetic field distribution in magnet electron lenses and electromagnets, *Nucl. Instrum. Meth. A* **187**: 201–208.
22. Mulvey, T. (1992). Unconventional lens design, In *Magnetic Electron Lenses*, P. W. Hawkes, ed., Volume **13**, pp. 359–420, Berlin, Heidelberg, New York: Springer.
23. Mulvey, T. (1984). Magnetic electron lenses II, In *Electron Optical Systems*, J. J. Hren et al. (ed.), pp. 15–27, Scanning Electron Microscopy, Chicago.
24. Tahir, K. and Mulvey, T. (1990). Pitfalls in the calculation of the field distribution of magnetic lenses by the finite element method, *Nucl. Instrum. Meth. A* **298**: 389–395.
25. Zeh, K. (1987). *Eine elektronenoptische Anwendung der Methode der finiten Elemente*, unpublished work, Tübingen, Germany.
26. Ströer, M. (1988). The integral equation method for field calculations in three dimensions and its reduction to a sequence of two-dimensional problems, *Optik* **81**: 12–20.
27. Hermeline, F. (1982). Triangulation Automatique d'un Polyèdre en Dimension N, *R.A.I.R.O. Analyse Numerique* **16**: 211–242.
28. Thacker, W. C. (1980). A brief review of techniques for generating irregular computational grids, *Int. J. Num. Meth. Eng.* **15**: 1135–1341.
29. Eupper, M. (1985). Eine verbesserte Integralgleichungsmethode zur numerischen Lösung dreidimensionaler Dirichletprobleme und ihre Anwendung in der Elektronenoptik, Dissertation, Germany: Universität Tübingen.
30. Bazeley, K. J., Cheung, G. P., Irons, Y. K., and Zienkiewicz, O. C. (1965). Triangular elements in bending-conforming and nonconforming solutions, *Proc. of First Conf. Matrix Methods in Struct. Mech.* 547–576.
31. Schwarz, H. R., ref. [8], pp. 96–100.
32. Edgcombe, C. J. (1999). Consistent modelling for magnetic flux in rotationally symmetric systems, *Nucl. Instrum. Meth. A* **427**: 412–416.
33. Kasper, E. (2000). An advanced boundary element method for calculation of magnetic lenses, *Nucl. Instrum. Meth. A* **450**: 173–178.
34. Shao, Z. and Lin, P. S. D. (1989). High resolution low voltage electron optical system for very large specimens, *Rev. Sci. Instrum.* **60**: 3434–3441.
35. *Numerical Recipes*, ref. [37], pp. 89–92.
36. Killes, P. (1985). Solution of Dirichlet problems using a hybrid finite differences and integral-equation method applied to electron guns, *Optik* **70**: 64–71.
37. Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (1988). *Numerical Recipes*, 3rd ed., Cambridge: Cambridge University Press.
38. Collins, R. A. (1973). Bandwidth reduction by automatic mesh renumbering, *Int. J. Num. Meth. Eng.* **6**: 345–346.

39. Bunch, J. R. and Rose, D. J. (eds.) (1976). In *Sparse Matrix Computations*, New York: Academic Press.
40. Hestenes, M. R. (1980). *Conjugate Direction Methods in Optimization*, New York: Springer.
41. Varga, R. S. (1962). *Matrix Iteration Analysis*, Englewood Cliffs: Prentice Hall.
42. Young, D. M. (1971). Iterative solution of large linear systems. *Computer Science and Applied Mathematics*, New York: Academic Press.
43. Kelley, C. T. (1995). *Iterative Methods for Linear and Nonlinear Equations*, Philadelphia: SIAM (Society for Industrial and Applied Mathematics).
44. Carré, B. A. (1961). The determination of the optimum accelerating factor for successive over-relaxation, *Comput. J.* **4**: 73–78.
45. Winslow, A. M. (1966). Numerical solution of the quasilinear Poisson equation in a nonuniform triangular mesh, *J. Comput. Phys.* **1**: 149–172.
46. Peaceman, P. W. and Rachford, H. H. (1955). The numerical solution of parabolic and elliptic differential equations, *SIAM J.* **3**: 28–41.
47. Stone, H. L. (1968). Iterative solution of implicit approximations of multidimensional partial differential equations, *SIAM J.* **5**: 530–558.
48. Buneman, O. (1973). A compact noniterative Poisson solver, *J. Comput. Phys.* **11**: 307–314 and 447–448.
49. Murray, W. (ed.) (1972). *Numerical Methods for Unconstrained Optimization*, New York: Academic Press.
50. Dennis, J. E. (1976). A brief survey of convergence results for quasi-newton methods, *SIAM-AMS Proc.*, **9**: 185–200.

This Page Intentionally Left Blank

CHAPTER VI

The Boundary Element Method

The boundary element method (BEM) is the third important method of field calculation. Whereas the FDM and the FEM consist in the dissection of the area or volume of solution into sufficiently small and numerous elements and the calculation of the potential at their nodes, this dissection is now performed at the *boundary* surfaces. The unknowns to be determined here, are then preferably the *sources* at the nodes of this boundary grid, which may be surface charge densities or surface current densities. These result from an approximate solution of such *integral equations* as were derived in Sections 1.6 and 1.7.

The BEM has two important advantages, which may make it attractive:

- (i) It can be used for truly three-dimensional configurations, that is, in cases where the surfaces are not rotationally symmetric about an optic axis.
- (ii) It is not necessary to use field interpolation techniques: at positions outside the charge carrying surfaces, the potential and all its derivatives are analytic functions; the requirement of continuity is automatically satisfied.

On the other hand, there are also essential disadvantages that must not be overlooked:

- (i) It is very difficult, though not impossible to solve problems with nonlinear material properties; these require the calculation of spatial source distributions, which is very time-consuming.
- (ii) Although the resulting field is very smooth, the calculation is very time-consuming, because it requires the evaluation of very many analytic functions such as square roots.
- (iii) The matrices resulting from the discretization of integral equations are here compact; because memory saving techniques such as those for sparse matrices cannot be applied, the number of possible nodes is limited.

These obstacles are not of an intrinsic nature and can be overcome together with the development of the computer technology, especially the development of vectorizable computers. An important gain is already achieved if parts of the necessary integrations can be carried out analytically, as is the case for rotationally symmetric surfaces.

6.1 DISCRETIZATION OF INTEGRAL EQUATIONS

In the following presentation, we shall denote the boundary surface of a spatial configuration by B ; this boundary may consist of several *disjoint* parts, but for reasons of conciseness this will not be denoted explicitly, unless it is indispensable. According to our general conventions, the integrations here run over all parts of B . Moreover, the boundary may be closed or *open*; in the latter case it is tacitly assumed that all field functions vanish at infinity, which is called the *natural boundary conditions*. A third aspect, to be considered, is the distinction between general boundaries and *rotationally symmetric* ones as in the latter case a more efficient technique can be developed in the context of Fourier series expansions; this topic is deferred to Section 6.2. Here we shall deal now with the general situation.

The boundary B is now to be represented in parametric form: $(x, y, z) = \mathbf{r}(u, v)$, u and v being any suitable variables. Depending on the form of B in the u - v -plane, the discretization can be performed by means of the techniques outlined in Section 3.5 or more generally by using the triangulations explained in Section 5.1, whereupon all the techniques of numerical analysis in triangular meshes (Section 5.3) can be applied here too. The choice of the most suitable method depends essentially on the shape of the boundary in question.

6.1.1 General Methods

The scalar integral equations, derived in Chapter II, are cast in the form of an integral equation of Fredholm's second kind in two dimensions:

$$\int_{\bar{B}} \bar{K}(\mathbf{r}, \mathbf{r}') \bar{Q}(\mathbf{r}') d^2 r' = \bar{\lambda}(\mathbf{r}) \bar{Q}(\mathbf{r}) + \bar{P}(\mathbf{r}) \quad (6.1a)$$

or equivalently in parametric form

$$\iint_B K(u, v; u', v') Q(u', v') du' dv' = \lambda(u, v) Q(u, v) + P(u, v) \quad (6.1b)$$

with $\mathbf{r} \in \bar{B}$ or $(u, v) \in B$. The specialization to Fredholm's equation of the first kind, $\lambda = \bar{\lambda} = 0$ is included in this form. The appropriate discretization proceeds in analogy to that used in the FEM: we start from a series expansion of the general form

$$Q(u, v) = \sum_{j=1}^N S_j N_j(u, v), \quad (6.2)$$

with initially unknown sampling coefficients $S_1, \dots, S_N$. Again, these are not necessarily function values at the nodes; they may also be nodal values of partial derivatives. On introducing Eq. (6.2) into Eq. (6.1b), we now obtain

$$\begin{aligned} \sum_{j=1}^N S_j \iint_B K(u, v; u', v') N_j(u', v') du' dv' \\ = \sum_{j=1}^N (S_j \lambda(u, v) N_j(u, v)) + P(u, v). \end{aligned} \quad (6.3)$$

This is a linear integral relation that cannot be satisfied exactly, except in the exceptional case when the series expansion (6.2) is an exact solution. Therefore, some further approximations are now necessary.

The commonest kind of approximation is the so-called *collocation*. This consists of the requirement that Eq. (6.3) shall be satisfied at all nodes of the surface grid. Because the number of these must agree with the number N of unknowns, this is possible only when the unknowns are the sampling values at the nodes, hence

$$S_j = Q(u_j, v_j) =: Q_j, \quad (j = 1, \dots, N), \quad (6.4a)$$

$$N_j(u_i, v_i) = \delta_{ij}, \quad (i, j = 1, \dots, N). \quad (6.4b)$$

The system matrix A obtained from Eq. (6.3) becomes (after omission of now unnecessary primes)

$$A_{ij} = \iint_B K(u_i, v_i; u, v) N_j(u, v) du dv - \lambda(u_i, v_i) \delta_{ij}, \quad (6.5)$$

and the linear system of equations to be solved is

$$\sum_{j=1}^N A_{ij} Q_j = P_i \equiv P(u_i, v_i), \quad (i = 1 \dots N). \quad (6.6)$$

Evidently, this system is *asymmetric*. We must now assume that the matrix A is nonsingular; the solution can then be found by means of the familiar *LU* algorithm. The accuracy of the collocation technique is not as high as that of the Galerkin method outlined later; the collocation is, however, still used frequently because only *one* pair of surface integrations is necessary. These integrations may become complicated, as we shall see later.

The *Galerkin* method is the second important kind of approximation. It is in widespread use in computational engineering and is very similar to corresponding techniques in quantum mechanics, as it consists in the evaluation of

projections into the basic functions. This means here that Eq. (6.3) is in turn multiplied by $N_i(u, v)$ ($i = 1, \dots, N$) and then integrated over u and v . Thus we obtain the matrix elements

$$K_{ij} \equiv \langle N_i | K | N_j \rangle := \quad (6.7a)$$

$$\iint_B \iint_B N_i(u, v) K(u, v; u', v') N_j(u', v') du dv du' dv'$$

$$\hat{\lambda}_{ij} = \iint_B N_i(u, v) \lambda(u, v) N_j(u, v) du dv, \quad (6.7b)$$

$$\hat{P}_i \equiv \langle N_i | P \rangle := \iint_B N_i(u, v) P(u, v) du dv, \quad (6.7c)$$

whereupon the linear system of equations becomes

$$\sum_{j=1}^N (K_{ij} - \hat{\lambda}_{ij}) S_j = \hat{P}_i \quad (i = 1, \dots, N). \quad (6.8)$$

This approximation has the following advantages:

- (i) In general, the Galerkin technique yields more accurate results as it comes close to a least squares fit.
- (ii) It is possible to use other sampling variables than just the nodal values of the function itself.
- (iii) In important classes of application, the system matrix becomes symmetric and positive definite, so that the Cholesky algorithm can be used.

On the other hand, it is now necessary to perform the fourfold integrations in Eq. (6.7a). For large distances between the two reference points in question, this is straightforward and only a matter of skillful program organization. When $u \rightarrow u'$ and $v \rightarrow v'$, however, the kernel K becomes *singular* so that special integration techniques become necessary.

Because large compact matrices form a serious obstacle, the first successful applications of the BEM to electron optical systems (Singer and Braun [1] and many other similar investigations) concerned rotationally symmetric configurations, for which the memory requirements are fairly small. This topic will be dealt with in Sections 6.2ff. This situation was systematically explored by Harting and Read [2]. With developments in computer technology and the increase in available memory, it became possible to employ three-dimensional methods, different versions of which are to be found in references [3–8]. The benchmark tests of Cubric et al. [9] have shown that in two dimensions the BEM is always the most accurate and that even the three-dimensional form is still very favorable. The mathematical treatment of the equations arising in the

BEM is to be found in references [10–14]. A few analytically soluble cases, which are useful as tests, can be found in Polyanin and Manzhirov [15]. The Galerkin method, described in detail below, has been tested by K. Oehler in collaboration with the author (Tübingen, 1994, unpublished) and proved to be favorable.

6.1.2 Surface-Coulomb Integrals

As a practical application of some importance, we shall now deal with the evaluation of the integral Eq. (1.98) resulting from a Dirichlet problem for Laplace's equation. The external potential $V_e(\mathbf{r})$ may be quite useful, as we shall see, but for the process of discretization this is of no importance because we assume here that the boundary values $U(\mathbf{r}) := V(\mathbf{r}) - V_e(\mathbf{r})$ are known. To eliminate inessential constants, we introduce the surface function $S(\mathbf{r}) = (4\pi\epsilon_0)^{-1}\sigma(\mathbf{r})$, whereupon Eq. (1.98) takes the concise form

$$U(\mathbf{r}) = \int_B S(\mathbf{r}') |\mathbf{r} - \mathbf{r}'|^{-1} d^2\mathbf{r}', \quad \mathbf{r} \in B. \quad (6.9)$$

This is now a simple special case of Eq. (6.1a) with $\bar{\lambda} = 0$ and a symmetric positive kernel $K = |\mathbf{r} - \mathbf{r}'|^{-1}$. In principle, the method of solution, outlined in the previous section, is straightforward, but owing to the *singularity* of the kernel the numerical integration may become difficult. We therefore consider here only the approximation of lowest order: the choice of *triangular* elements on the boundary and the use of the *linear* form function in these. This means that in each element the functions $N_i(u, v) = \xi_1, \xi_2, \xi_3$, defined by Eq. (5.31) are used, the numbering being adjusted accordingly.

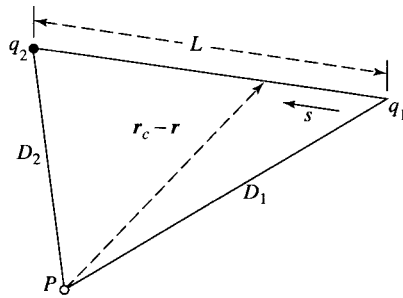
The consequence of this choice is to facilitate the necessary integration because at least *one* of the two integrations can be carried out analytically in a fairly simple manner. Eupper [3,4] has shown that it is also possible to perform the second integration in an analytical way; however, the number of transcendental functions to be determined is then so large that this method becomes slow; here we shall consider it in a modified form.

The first integration over the Coulomb kernel is the calculation of the potential produced by a linearly charged rod of length L (see Fig. 6.1). The *line* charge density $q(s)$ on the rod may be written in the form

$$q(s) = q_1 + s(q_2 - q_1)/L \quad (0 \leq s \leq L) \quad (6.10a)$$

and its position correspondingly by

$$\mathbf{r}_c(s) = \mathbf{r}_1 + s(\mathbf{r}_2 - \mathbf{r}_1)/L \quad (0 \leq s \leq L). \quad (6.10b)$$

FIGURE 6.1 Notation concerning the Coulomb potential of a charged rod of length L .

The Coulomb integral

$$\Phi(\mathbf{r}) = \int_0^L \frac{q(s) ds}{|\mathbf{r} - \mathbf{r}_c(s)|} \quad (6.11)$$

becomes after some elementary calculations:

$$\begin{aligned} \Phi(\mathbf{r}) = & (q_2 - q_1)(D_2 - D_1)/L \\ & + \frac{1}{2}[q_1 + q_2 - (q_2 - q_1)(D_2^2 - D_1^2)/L^2] \\ & \times \ln \left(\frac{D_1 + D_2 + L}{D_1 + D_2 - L} \right), \end{aligned} \quad (6.12a)$$

$$D_1 = |\mathbf{r} - \mathbf{r}_1|, \quad D_2 = |\mathbf{r} - \mathbf{r}_2| \quad (6.12b)$$

being the distances from the two endpoints of the rod. Apart from the evaluation of the logarithm, only two square roots for the distances D_1 and D_2 are necessary to calculate this potential.

By means of this formula, it is easy to calculate the potential $V(\mathbf{r}_0)$ at a corner P_0 of an arbitrary linearly charged triangle. We now choose a local (u, v) -coordinate system with its origin at this point, and the u -axis perpendicular to the opposite side of the triangle as is shown in Fig. 6.2. The integration over the coordinate v leads to a result that is analogous to Eq. (6.12), where D_1, D_2 , and L being replaced by d_1, d_2 , and l , respectively. However, the ratios $d_1/l = D_1/L$ and $d_2/l = D_2/L$, and hence the whole logarithm are *constant* and can be taken out of the integral.

If we assume a linear function $S(u, v)$ having the values S_0, S_1 , and S_2 at the three corners P_0, P_1 , and P_2 , respectively, the surface charge densities on the two lines $\overline{P_0P_1}$ and $\overline{P_0P_2}$ becomes

$$\bar{S}_k(u) = S_0 + (S_k - S_0)u/h, \quad (k = 1, 2). \quad (6.13)$$

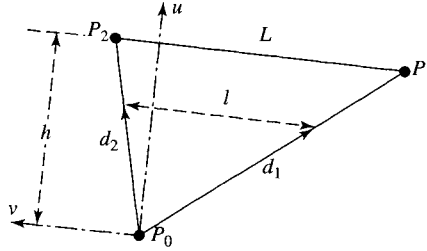


FIGURE 6.2 Notation concerning the calculation of the potential at the edge P_0 of a charged triangle.

These are to be substituted for q_1 and q_2 in Eq. (6.12a), and then integrated over u from $u = 0$ to $u = h$. The result is again a formula of the same type for $V(\mathbf{r}_0) = \Phi(\mathbf{r}_0)$, but now with the charge parameters

$$q_k = h(S_0 + S_k)/2 \equiv (S_0 + S_k)a/L, \quad (6.14)$$

a being the area of this triangle.

Within the frame of collocation, it is now quite easy to determine the diagonal elements A_{ii} according to Eq. (6.5). The configuration is shown in Fig. 6.3. The source values to be used are $S_0 = 1$ and $S_k = 0$. If we adopt again a cyclic notation, which means that $D_{N+1} = D_1$, then the result can be written concisely as

$$A_{ii} \equiv V_0 = \frac{1}{2} \sum_{k=1}^N \frac{a_k}{L_k} \ln \left(\frac{D_k + D_{k+1} + L_k}{D_k + D_{k+1} - L_k} \right). \quad (6.15)$$

In the general situation, particularly if the reference point $\mathbf{r}_0$ is located outside the plane of the triangle, the result of the integration becomes more

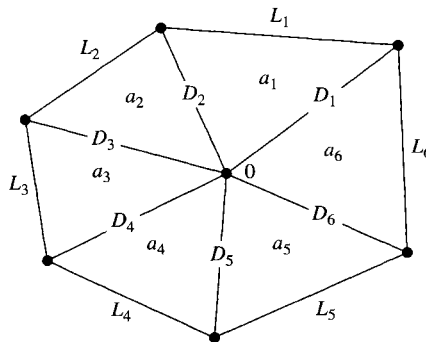


FIGURE 6.3 Notation concerning the potential at the common node 0 of $N = 6$ triangles. Generally this configuration can be a pyramid with top O in space.

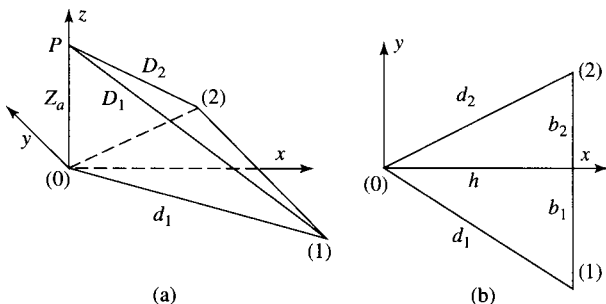


FIGURE 6.4 Notation concerning the calculation of the potential at a point P perpendicularly above the footpoint (0) as an edge of a triangle: (a) perspective graph; (b) projection into the $x - y$ plane.

complicated, but it can still be represented in terms of analytic functions. Because the accuracy of all the numerical integration techniques is poor in the close vicinity of the triangle, we shall now outline briefly this analytical procedure. It is essentially analogous to Eupper's [4] derivation.

We start with a special configuration, shown in Fig. 6.4, in which the reference point is located directly above one of the three edges. Without loss of generality, a *local* cartesian coordinate system (x, y, z) can be chosen in the way shown in the figure. The surface source density may be written as a linear function

$$S(x, y) = S_0 + Ax + By. \quad (6.16)$$

On introducing the *slopes* $\alpha = b_1/h$, $\beta = b_2/h$ and the frequently appearing distance

$$D(x, y, z) = (x^2 + y^2 + z^2)^{1/2}, \quad (6.17)$$

the Coulomb integral for the potential V takes the explicit form

$$V(0, 0, z) = \int_{x=0}^h \left(\int_{y=-\alpha x}^{\beta x} D^{-1}(S_0 + Ax + By) dy \right) dx. \quad (6.18)$$

The integrals with a linear factor in the numerator are easily evaluated, if the inner integration is carried out over this coordinate. The results are most favorably represented in terms of the frequently appearing function values

$$F_0 := \ln \left(\frac{D_2 + b_2}{D_1 - b_1} \right) \equiv \ln \left(\frac{D_1 + D_2 + L}{D_1 + D_2 - L} \right) \quad (6.19a)$$

and

$$F_i := \ln[(D_i + d_i)/z_a]/d_i, \quad (i = 1, 2) \quad (6.19b)$$

with $z_a := |z|$ and the distances $L = b_1 + b_2$ and $D_i = (z^2 + d_i^2)^{1/2}$, as are shown in the figure. The corresponding integrals then take the concise form

$$V_1 := \iint_{\Delta} x D^{-1} dx dy = \frac{1}{2}(z^2 + h^2)F_0 - \frac{z^2}{2}(b_1 F_1 + b_2 F_2), \quad (6.20a)$$

$$V_2 := \iint_{\Delta} y D^{-1} dy dx = \frac{h}{2}(D_2 - D_1) + \frac{hz^2}{2}(F_2 - F_1). \quad (6.20b)$$

The third integral leads to a more complicated expression. The first integration over y is quite simple and results in

$$V_0 = \iint_{\Delta} D^{-1} dy dx = \int_0^h \ln \left(\gamma x + \sqrt{(1 + \gamma^2)x^2 + z^2} \right) dx \Big|_{\gamma=-\alpha}^{\beta}. \quad (6.21)$$

For $z = 0$, the integrand becomes $\operatorname{arsinh}(\gamma x)$, and the integration is then quite simple, but in the general case this does not hold. The next step is now a partial integration, the second factor being unity. Thereafter, the decomposition of the rational integrands into partial fractions and the substitution $u = x^2$ can be carried out, whereupon some other elementary integrals are evaluated, but there remains the more complicated integral

$$\int \frac{\sqrt{pu+q}}{2(u+q)} du = \sqrt{pu+q} + \sqrt{pq-q} \arctan \sqrt{\frac{pq-q}{pu+q}} + C \quad (6.22)$$

with $q = z^2$ and $p = 1 + \gamma^2$. This formula is to be evaluated twice for $\gamma = -\alpha$ and $\gamma = \beta$, respectively. The results of these longer calculations can be cast in the concise form

$$V_0 = hF_0 - z_a(\varphi_1 + \varphi_2) \quad (6.23)$$

with the angular functions

$$\begin{aligned} \varphi_i &:= \arctan(b_i/h) - \arctan \left(\frac{b_i z_a}{h D_i} \right) \\ &= \arctan \left(\frac{h b_i (D_i - z_a)}{h^2 D_i + z_a b_i^2} \right), \quad (i = 1, 2). \end{aligned} \quad (6.24)$$

Although the argument in the second form is more complicated, it has the advantage that the denominator is always positive.

It is now of some importance that the integrals for the components of the gradient can be represented in terms of the same analytic functions, so that

the whole calculations can be cast in a fairly economic form. These integrals are defined by

$$W_{m,n} := \int_{\Delta} x^m y^n D^{-3} dx dy, \quad (0 \leq m+n \leq 2). \quad (6.25)$$

The results of their evaluation are be written as follows:

$$\left. \begin{aligned} z_a W_{0,0} &= \varphi_1 + \varphi_2 \\ W_{1,0} &= -F_0 + b_1 F_1 + b_2 F_2 \\ W_{0,1} &= h(F_1 - F_2) \\ W_{2,0} &= hb_1/(D_1 + z_a) + hb_2/(D_2 + z_a) - z_a(\varphi_1 + \varphi_2) \\ W_{1,1} &= h^2/(D_1 + z_a) - h^2/(D_2 + z_a) \\ W_{0,2} &= V_0 - hb_1/(D_1 + z_a) - hb_2/(D_2 + z_a). \end{aligned} \right\} \quad (6.26)$$

Note that only $z_a = |z|$ appears in these formulas, which implies that they are symmetric with respect to z , as they must be.

The coefficients A and B , appearing in Eqs. (6.16) and (6.18), can easily be determined from the nodal values S_0, S_1 , and S_2 , resulting in

$$B = (S_2 - S_1)/L, \quad A = [(b_1 S_2 + b_2 S_1)/L - S_0]/h. \quad (6.27)$$

With these we obtain now the following linear combinations for the potential v and the gradient $\mathbf{g} = \nabla v$ at the position $(0, 0, z)$.

$$\left. \begin{aligned} v &= S_0 V_0 + A V_1 + B V_2 \\ g_x &= S_0 W_{1,0} + A W_{2,0} + B W_{1,1} \\ g_y &= S_0 W_{0,1} + A W_{1,1} + B W_{0,2} \\ g_z &= -(S_0 W_{0,0} + A W_{1,0} + B W_{0,1})z. \end{aligned} \right\} \quad (6.28)$$

The result for g_z is always finite as $z/z_a = \pm 1$. Moreover, all functions are eventually found to be finite because each singularity is compensated by a vanishing factor in the numerator, but this form is not yet acceptable for a computer. To avoid difficulties, it is sufficient to replace any vanishing denominator with a very tiny *positive* value.

This implies that the singularity cannot be reached exactly, and the over-compensation by the numerator leads then to the correct result.

With respect to later summations, it is necessary to transform the gradient into components in a *global* coordinate system. The local unit vectors

$$\begin{aligned}
u_y &= (\mathbf{r}_2 - \mathbf{r}_1)/L, \\
u_z &= (\mathbf{r}_1 - \mathbf{r}_0) \times (\mathbf{r}_2 - \mathbf{r}_0)/2a \equiv \mathbf{n}, \\
u_x &= u_y \times \mathbf{n},
\end{aligned} \tag{6.29}$$

can easily be represented in global cartesian form, the vector $\mathbf{n}$ is here the surface normal, as $2a$ is twice the area of the triangle. It is now easy to determine the *global coordinates* of

$$\mathbf{g} = g_x u_x + g_y u_y + g_z u_z, \tag{6.30}$$

which must later be summed up.

We now come to the next task, the calculation of the potential $U(\mathbf{r})$ and its gradient $\mathbf{G}(\mathbf{r}) = \nabla U$ at an *arbitrary* position $\mathbf{r}$ in space. The first step is the determination of the corresponding *area* coordinates ξ_1, ξ_2, ξ_3 . The methods outlined in Section 3.5.1 are quite generally applicable and the results refer then to the *footpoint* $\mathbf{r}_F$ in the plane of the triangle as the area coordinates are scalar products formed with *c*-vectors located in this plane. These formulas remain valid even when the footpoint is located outside; then at least one of the area coordinates becomes negative. The footpoint itself is easily obtained by

$$\mathbf{r}_F = \xi_1 \mathbf{r}_1 + \xi_2 \mathbf{r}_2 + \xi_3 \mathbf{r}_3 \tag{6.31a}$$

and consequently $z = \mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_F)$, $z_a = |z|$. The source density S_0 at this point is analogously given by

$$S_0 = S(\mathbf{r}_F) = \xi_1 S_1 + \xi_2 S_2 + \xi_3 S_3. \tag{6.31b}$$

This value is to be introduced into Eqs. (6.16), (6.27), and (6.28). Negative area coordinates imply here a linear extrapolation instead of an interpolation.

After the determination of the footpoint, this is to be considered as the origin of the local system and the previously derived method becomes applicable, if we dissect the true domain of integration into three partial triangles by the lines leading from the origin to the three edges, as shown in Fig. 6.5. With respect to the definition of ξ_1, ξ_2, ξ_3 , a *cyclic* notation is here favorable. With the appropriate adaptation to this notation, the previously derived formulas are now in turn applied to the triangles (0,1,2), (0,2,3), and (0,3,1) and the results summed up in terms of global cartesian components. The method remains valid if it is recognized that one of the projection b_i becomes negative in obtuse triangles. It also remains valid for footpoints located outside as shown in Fig. 6.5b. All those contributions for which the corresponding area coordinate becomes *negative* are to be subtracted. This situation occurs for the triangle (0,2,3) in the figure for which we have $\xi_1 < 0$. In a practical program this condition is easily implemented if we allow the *height* h to become

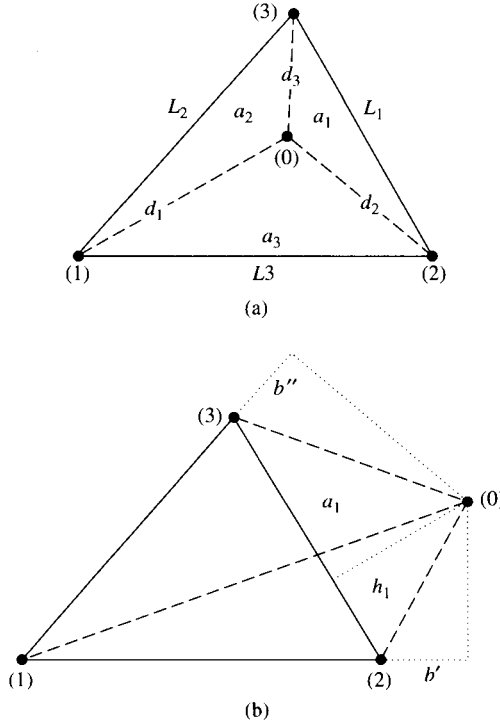


FIGURE 6.5 Dissection into partial triangles: (a) footpoint (0) inside an acute triangle (1,2,3), all quantities are positive; (b) footpoint (0) outside of the triangle: the area a_1 of the triangle (0,2,3), and consequently the height h_1 and the coordinate ξ_1 become negative. Moreover, the projections b' and b'' of the obtuse triangle (0,1,2) and (0,3,1) are negative.

negative; in the situation shown, this arises for $h = h_1 = 2a\xi_1/L_1$, a being the area of the triangle (1,2,3).

The entire procedure can be implemented in a fairly economic form, as it requires altogether only six logarithms, arctangents and nonconstant square roots, and some algebraic calculations. For purposes of demonstration, an example for the resulting potential in the plane $z = 0$ is shown in Fig. 6.6. In this case an asymmetric but positive source function $S(\mathbf{r})$ was chosen, and the potential then calculated along a line passing through two midpoints of the sides of the triangle. It shows the qualitative behavior to be expected: in the interior, it reaches its maximum and far outside it decreases like a Coulomb potential. In the vicinity of the boundaries ($\xi_i = 0$), a weak singularity of the form $\xi_i \ln |\xi_i|$ is superimposed on the regular background. As mentioned earlier, this must be eliminated by confinement. The neighborhood of this singularity is confined to a very narrow zone, as shown in the enlarged window Fig. 6.6a.

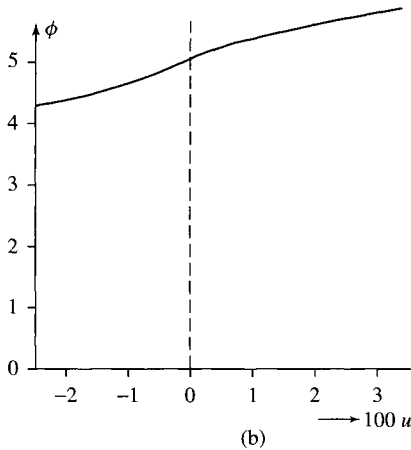
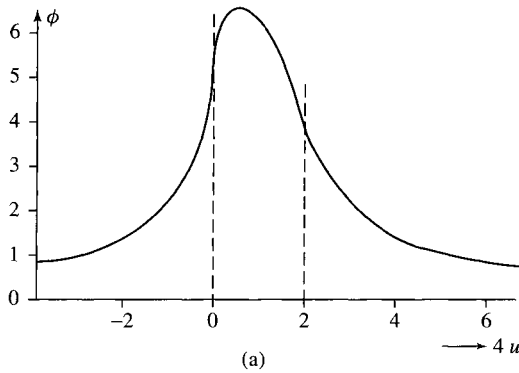


FIGURE 6.6 Potential along a line passing through two midpoints of the sides of a linearly charged triangle. The interval in the interior is marked: (a) larger range; (b) very small window enlarged 25 times.

Within the frame of the BEM, this feature of the potential is quite unimportant because this term is compensated by a corresponding one caused by the adjacent triangle that has the same strength but opposite sign because of the continuity of the surface source. With respect to the Galerkin approximation, it is therefore sufficient to carry out the necessary second integration in Eq. (6.7a) by a 13-point surface Gauss quadrature, which leads to a considerable simplification of this method.

For *larger* distances between the reference point $\mathbf{r}$ and the area of integration, triangular *Gauss quadratures* can be used, which are easier to implement, asymptotically stable, and less expensive, as their order can be reduced with increasing distance. The determination of the appropriate order is a matter of trial and testing; it depends also on the accuracy requirements.

As a reasonable guess that should be checked, Table 6.1, presented in the next section, is suggested.

6.1.3 The Far-Field Approximation

A realistic configuration consists of very many triangular surface elements, and for the vast majority of them, the mutual distance is much larger than the maximal side length L . It is then not always necessary to use the previously outlined method; much computation time can be saved if the subsequently outlined algorithm can be used. This is of particular advantage in ray-tracing programs as the rays usually stay far from charged surfaces. We again start from Eq. (6.9) and consider the integral over one representative triangle, but now in the far-field zone.

The most advantageous method here is the well-known electrostatic *multipole* expansion, referred to the *centroid* $\mathbf{r}_c$ of the actual triangle. It is hence favorable to use the difference vectors $\mathbf{d} := \mathbf{r}' - \mathbf{r}_c$. The multipole elements of a linearly charged triangle are easily determined from the knowledge of the three edge vectors $\mathbf{d}_1, \mathbf{d}_2$, and $\mathbf{d}_3$ and the source values S_1, S_2 , and S_3 at these. Let a be the area of the triangle, then we obtain in turn first the total charge

$$C \equiv aS_c = a(S_1 + S_2 + S_3)/3, \quad (6.32a)$$

S_c being the source density at the centroid, and then the *dipole* vector

$$\mathbf{p} = a \sum_{n=1}^3 (S_n - S_c) \mathbf{d}_n / 4 \quad (6.32b)$$

and the symmetric trace-free quadrupole tensor having the components

$$Q_{ik} = \frac{a}{12} \sum_{n=1}^3 (0.6S_c + 0.4S_n) (3d_{ni}d_{nk} - d_n^2 \delta_{ik}). \quad (6.32c)$$

It is not necessary to use the tensor in this explicit form as in practical calculations only the projection onto an arbitrary vector $\mathbf{u}$ is necessary, giving

$$\mathbf{q} := \mathbf{Q} \cdot \mathbf{u} = \frac{a}{12} \sum_{n=1}^3 (0.6S_c + 0.4S_n) [3\mathbf{d}_n \cdot \mathbf{u} \mathbf{d}_n - d_n^2 \mathbf{u}]. \quad (6.32d)$$

Now the potential $\delta U(\mathbf{r})$ and its gradient are easily determined. The distance R of the reference point and the *unit* vector $\mathbf{u}$ pointing to it are

$$R := |\mathbf{r} - \mathbf{r}_c|, \quad \mathbf{u} := (\mathbf{r} - \mathbf{r}_c)/R, \quad (6.33a)$$

whereupon we obtain

$$\delta U(\mathbf{r}) = C/R + \mathbf{p} \cdot \mathbf{u}/R^2 + 0.5 \mathbf{q} \cdot \mathbf{u}/R^3 + O(R^{-4}), \quad (6.33b)$$

$$\begin{aligned} \text{grad } \delta U = & -C\mathbf{u}/R^2 + (\mathbf{p} - 3\mathbf{u} \cdot \mathbf{p}\mathbf{u})/R^3 \\ & + (\mathbf{q} - 2.5\mathbf{u} \cdot \mathbf{q}\mathbf{u})/R^4 + O(R^{-5}). \end{aligned} \quad (6.33c)$$

Provided that the distance R is so large that the quadrupole terms can already be ignored, and that the surface source has a unique sign, then the dipole terms can be eliminated by choosing the centroid $\mathbf{r}_s$ of the sources instead of the geometric one. The former is defined by

$$\mathbf{r}_s = \mathbf{r}_c + \mathbf{p}/C. \quad (6.34)$$

Only the monopole terms are retained; these are now given by

$$\delta U = C/|\mathbf{r} - \mathbf{r}_s|, \quad \nabla \delta U = C \frac{\mathbf{r}_s - \mathbf{r}}{|\mathbf{r}_s - \mathbf{r}|^3}. \quad (6.35)$$

Within the frame of the *Galerkin* approximation, the matrix elements K_{ij} must be determined from Eq. (6.7a) in an efficient manner. As the trial functions N_i , and N_j are now the area coordinates, these matrix elements specialize to

$$K_{i,j} = \sum_{\mu} \sum_{\nu} \int_{\Delta_{\mu}} \int_{\Delta_{\nu}} \frac{\xi_i(\mathbf{r}) \xi_j(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} da da'. \quad (6.36)$$

The summations run over all those triangular elements Δ_{μ} and Δ_{ν} that have the corresponding nodes (i) or (j) in common. Without derivation, the formula for one of these terms will be stated, and the subscripts μ and ν will be dropped for reasons of conciseness. Quantities referring to the second element are marked by a prime. In each element, the *geometric* centroid is again chosen as the relative origin. Then the definitions (see Fig. 6.7)

$$R := |\mathbf{r}'_c - \mathbf{r}_c|, \quad \mathbf{u} := (\mathbf{r}'_c - \mathbf{r}_c)/R \quad (6.37)$$

are in agreement with Eq. (6.33a). It is now advantageous to introduce the dimensionless vectors

$$\mathbf{v}_m := \mathbf{d}_m/4R, \quad \mathbf{v}'_n := \mathbf{d}'_n/4R, \quad (m, n = 1, 2, 3) \quad (6.38a)$$

and their longitudinal components

$$w_m := \mathbf{u} \cdot \mathbf{v}_m, \quad w'_n := -\mathbf{u} \cdot \mathbf{v}'_n, \quad (m, n = 1, 2, 3), \quad (6.38b)$$

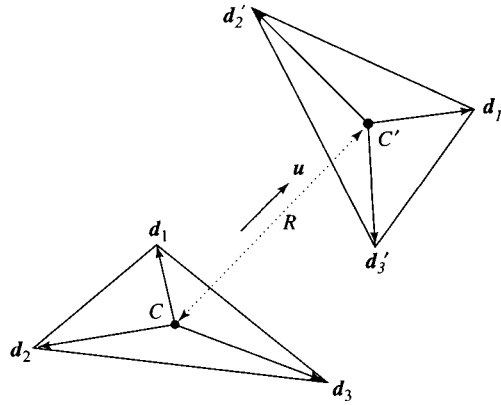


FIGURE 6.7 Notation concerning the Coulomb interaction between two charged triangles in the far-field approximation. The distance R between the two centroids C and C' should be much larger than could be drawn here.

d_m and d'_n again being the vectors leading from the centroid r_c or r'_c to the corresponding edge, respectively. Strictly speaking, we notice that the labels i and j in Eq. (6.36) are the *global* ones that are associated with the corresponding nodes by a very complex kind of memory organization. For reasons of conciseness, we therefore use the natural *local* numbering that provides a 3×3 matrix M . Its elements are now given by

$$M_{i,k} = \frac{aa'}{9R} \left\{ 1 + w_i + w'_k + 3w_i w'_k + 2.4(w_i^2 + w'_k{}^2) + \mathbf{v}_i \cdot \mathbf{v}'_k - 0.8(v_i^2 + v'_k{}^2) + 1.2 \sum_{n=1}^3 (w_n^2 + w'_n{}^2) - 0.4 \sum_{n=1}^3 (v_n^2 + v'_n{}^2) \right\}, \quad (6.39)$$

which include correctly all quadrupole interaction terms, the error hence being of order R^{-4} . This formula appears to be rather complex at first sight but it requires only one square root for all nine matrix elements together, and also the intrinsic summation over the label n is to be carried out only once. With some skill, the algebraic expenditure can be kept fairly small, as the alternative, a double surface Gauss quadrature, requires at least 16 square roots. In the *dipole* approximation, the formula simplifies considerably to

$$M_{i,k}^D = aa' (1 + w_i + w'_k) / 9R + O(R^{-3}), \quad (6.40)$$

so that this should be used whenever it is accurate enough.

TABLE 6.1
OPTIMAL APPROXIMATION

R/L	Approximation
0-2	Analytical
2-4	Gauss 13 P.
4-16	Gauss 7 P.
16-32	Quadrupole
32-64	Dipole
>64	Monopole

In this table, which was found by comparison of different tests [4,6], the variable R is defined by Eq. (6.33a) or Eq. (6.37) where L denotes the longest side-length of the triangle. In the cited investigations, a *hybrid* method was also applied, which consists in a first one-dimensional integration by means of the *rod* formula (6.12), and a subsequent Gauss quadrature in the other direction. This method was not outlined here, as it does not bring a significant gain over the improved analytical method.

In cases of double integration, required in the Galerkin method, the same table can be used for the second triangle, L now denoting its longest side length. In the case $R/L \leq 2$, there is no second analytical formula; the 13-point formula should then be used, as was justified in the earlier section. In summary, we now have a technique for solving this difficult integration problem for every situation.

6.1.4 The Complete Procedure

A complete field calculation program using the BEM is rather complex; it proceeds in the following steps:

- (i) Input, parametrization, and discretization of all boundaries;
- (ii) Definition of boundary conditions and possible external potentials;
- (iii) Calculation of the system matrix and driving terms;
- (iv) Solution of the linear system of equations and storage of the resulting surface sources;
- (v) Repeated field calculations for equipotential surfaces or ray tracings.

The first step is here quite analogous to corresponding techniques in the FEM. The only difference consists in the fact that now the triangulations are to be carried out on bent surfaces instead of meridional sections. Usually the nodes are located on the boundaries themselves, but in case of the Galerkin method some improvement can be obtained if they are shifted in the normal direction in such a manner that, for each element, the averaged displacement nearly vanishes. The discretization supplies coordination tables, as in the FEM.

The following steps must be considered in connection with the memory available for the computation. In contrast to the FEM, the resulting matrix is now *compact*, so that all available memory saving techniques must be employed to obtain a reasonable accuracy. This has the following implications.

- Parts of the configuration that can be approximated reasonably and accurately by an *external* analytical potential $V_e(\mathbf{r})$ should not be discretized but used as a driving term to save surface elements.
- All given geometrical *symmetries* should be used to minimize the surface to be discretized. The corresponding contributions are then to be incorporated in the *kernel* K . If, for example, a system is mirror symmetric with respect to the plane $z = 0$, then only the right half part ($z \geq 0$) of the boundaries should be discretized, and the kernel becomes then

$$K = [(x - x')^2 + (y - y')^2 + (z - z')^2]^{-1/2} + [(x - x')^2 + (y - y')^2 + (z + z')^2]^{-1/2} \quad (6.41)$$

that still satisfies $K(\mathbf{r}, \mathbf{r}') = K(\mathbf{r}', \mathbf{r})$. Each such symmetry reduces the memory by a factor 4 and the computation time by a factor 2 in setting up the system matrix.

- If the kernel is symmetric and positive definite, then the Galerkin method should be chosen as this brings a further gain by a factor of 2 in memory and computation time because the Cholesky algorithm can be applied. Numerical checks have shown that the system matrices resulting from Coulomb kernels are so well-behaved that it makes sense to store only the diagonal elements in double precision and the off-diagonals in single precision, which brings practically another halving of the memory. This is certainly justified because the mere linear approximation of the surface source density implies a larger loss of accuracy than this.

Quite generally the storage of a symmetric matrix $\mathbf{K}$ with diagonal $\mathbf{D}$ and lower trigonal part $\mathbf{L}$ in this reduced form is easily achieved using an integer function

$$i' := (i - 1)(i - 2)/2. \quad (6.42a)$$

The necessary operations are in turn

$$\begin{aligned} D_i &:= K_{i,i} \quad (1 \leq i \leq N) \quad 8 \text{ bytes,} \\ L_{i' + n} &:= K_{i,n} \quad ((1 \leq n \leq i), 2 \leq i \leq N) \quad 4 \text{ bytes.} \end{aligned} \quad (6.42b)$$

The determination of this matrix proceeds as follows: initially all matrix elements are cleared (set equal to zero). Thereafter a double loop runs over all M triangular *elements* (with $M \approx 2N$), and in each configuration of pairs

of elements the set of corresponding labels (i, n) are determined from the table of coordination. The statements in the inner loop are skipped if $i \leq n$ or if this pair has already been considered. Otherwise the double integrations are carried out with full precision before the storage, to minimize rounding errors in the analytical calculations. At the end, their results are accumulated (added) according to Eqs. (6.42). As already demonstrated in Eq. (6.39), it is favorable to calculate all nine contributions from a pair of triangles simultaneously as this saves many operations. The same should be done with the Gauss quadrature and the analytical integrations, where, of course, the parts for the gradient are not necessary here.

The Cholesky decomposition of such a split matrix is concisely be given by

$$\begin{aligned}
 & D_1 \leftarrow D_1^{1/2} \\
 & \text{for } (i = 2, \dots, N) \\
 & \left\{ \text{for } (k = 1, \dots, i-1) \left[L_{i'+k} \leftarrow \left(L_{i'+k} - \sum_{n=1}^{k-1} L_{k'+n} L_{i'+n} \right) / D_k \right] \right. \\
 & \left. D_i \leftarrow \left(D_i - \sum_{n=1}^{i-1} L_{i'+n}^2 \right)^{1/2} \right\}. \tag{6.43}
 \end{aligned}$$

The evaluation of the driving terms is simple and straightforward if a *linear* approximation is sufficient. In the *cyclic local* indexing shown in Fig. 6.3, the result is simply given by

$$\hat{P}_i \leftarrow P_0 = \sum_{k=1}^{N'} a_k (2U_0 + U_k + U_{k+1})/12, \tag{6.44}$$

N' here being the number of triangles with common node 0. The forward and backward substitution loops are to be formulated in agreement with Eq. (6.42), which is a straightforward task. After their execution and the storage of their results, the solution of the integral equation Eq. (6.9) is finished and the applications of the field calculations can now start.

6.1.5 The Normal Derivative

Integral equations involving the normal components of the Green's function, such as Eq. (1.115), can be solved quite generally by using linear functions in triangular meshes; like the formulas of Section 6.1.2, they can be evaluated

at any point in space except at their singularities. In practice, however, the double integrations that are required if Galerkin's method is to be used are too difficult, and the saving of memory achieved by the use of Eq. (6.42) is no longer possible, the system matrix being asymmetric here. The simple *collocation* technique must now be employed, therefore, as follows.

If, for example, Eq. (1.115) is to be solved, it is advantageous to write the latter as

$$\left(\frac{1}{2} + \frac{\mu_1}{\mu_2 - \mu_1}\right) \tau_i + \sum_{j=1}^N \tau_j \mathbf{n}_i \cdot \int_B \xi_j(\mathbf{r}) \frac{(\mathbf{r}_i - \mathbf{r}) d\mathbf{a}}{4\pi |\mathbf{r}_i - \mathbf{r}|^3} = \mathbf{n}_i \cdot \mathbf{H}_0(\mathbf{r}_i),$$

$$(i = 1, \dots, N), \quad (6.45)$$

in which $\tau_i = \tau(\mathbf{r}_i)$ are the function values at the *nodes* and $\mathbf{n}_i = \mathbf{n}(\mathbf{r}_i)$ is the corresponding surface normal. The integration is again to be carried out over all the triangular elements and in all those elements that do not have a node at the reference position $\mathbf{r}_i$; once again, linear form functions can be used for $\xi_j(\mathbf{r})$. The evaluation of the corresponding matrix elements is then a straightforward application of the formulas derived in Sections 6.1.2 and 6.1.3.

However, modifications are necessary when we turn to the diagonal elements as linear form functions can still be used, but the surface elements on curved boundaries must not be approximated by plane triangles: there is no surface normal at the top of a pyramid. We must therefore define a local tangential plane and a corresponding normal $\mathbf{n}_i$ at each node $\mathbf{r}_i$. In the subsequent presentation, we shall consider a configuration such as that of Fig. 6.3, and identify the position $\mathbf{r}_i$ with the centre O of this configuration; now, however, we assume that there exists a common tangential plane at this point. Without loss of generality, a local coordinate system (x, y, z) with origin at the point O and z -axis in the direction of the normal $\mathbf{n}$ can be chosen, as shown in Fig. 6.8. The reference point P is now the origin itself, but the neighboring nodes have nonvanishing vertical coordinates, z_1 and z_2 . An exact representation of such a curved boundary element would become fairly complicated as

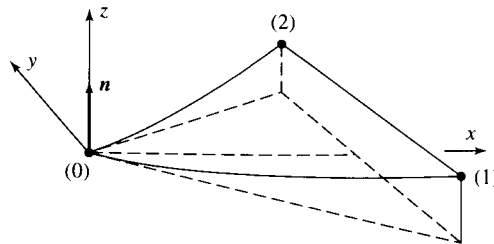


FIGURE 6.8 A bent triangle (0,1,2) and its projection onto the $x - y$ plane, which is the tangential plane at the node (0).

such a formula should match the surface and its normals at every node in the vicinity, but linear form functions would then be too inaccurate. We therefore introduce the following *local* quadratic approximation:

$$z(x, y) = x(px + qy). \quad (6.46)$$

The coefficients p and q are determined from the conditions

$$z(h, -b_1) = z_1, \quad z(h, b_2) = z_2, \quad (6.47)$$

where h_1 , b_1 , and b_2 are defined as in Fig. 6.4; this gives

$$p = \frac{b_1 z_2 + b_2 z_1}{h^2(b_1 + b_2)}, \quad q = \frac{z_2 - z_1}{h(b_1 + b_2)}. \quad (6.48)$$

The form function $\xi(\mathbf{r})$ to be used here is the one that takes the value *one* at the centre:

$$\xi = 1 - x/h, \quad (0 \leq x \leq h). \quad (6.49)$$

The contribution of the triangle considered to the diagonal element in question now becomes

$$J_{i,1} := -\frac{1}{4\pi} \int_{\Delta_1} \frac{\xi(x) z(x, y)}{(x^2 + y^2 + z^2)^{3/2}} dx dy. \quad (6.50)$$

The evaluation of this principal value leads to integrals of the form of Eqs. (6.25), if second order terms in z can be ignored because the curvature of the surface is very small, which will be assumed to be the case here. Analytical evaluation of the integral then yields

$$J_{i,1} = -\frac{ph}{8\pi} \left(\frac{b_1}{d_1} + \frac{b_2}{d_2} \right) + \frac{qh^2}{8\pi} \left(\frac{1}{d_2} - \frac{1}{d_1} \right). \quad (6.51)$$

Strictly speaking, the geometrical quantities appearing in this formula should refer to the projection of the curved triangle on the tangential plane, but this is an unnecessary complication. Because we have already ignored the terms in z^2 in Eq. (6.50), the quantities h , b_1 , b_2 , d_1 , and d_2 of the original triangle can be used within the same approximation, which is a considerable simplification. The components z_1 and z_2 in Eq. (6.48) can simply be calculated as scalar products:

$$z_k = \mathbf{n} \cdot (\mathbf{r}_k - \mathbf{r}_0), \quad (k = 1, 2). \quad (6.52)$$

The diagonal elements $D_{i,i}$ of Eq. (6.45) can now be calculated, giving

$$D_{i,i} = \frac{1}{2} + \frac{\mu_1}{\mu_2 - \mu_1} + \sum_{k=1}^{N'} J_{i,k}, \quad (6.53)$$

in which the summation runs over all triangles with a common node at $\mathbf{r}_i$. In the same manner, the integral Eqs. (1.108) for electric surface sources can be discretized, the essential modification being the replacement of μ_1 by ε_1 and μ_2 by ε_2 , as appropriate.

There remains one important aspect to be discussed if we have $\mu_2 \gg \mu_1$, say $\mu_2/\mu_1 = 10^5$, then the condition of the system matrix resulting from Eq. (6.45) becomes very poor; this integral equation should not then be used. In any case, it is necessary to evaluate the matrix elements with the highest precision available, eight bytes per number, and to perform postiterations. Clearly, the computational effort required is much greater than for Dirichlet problems.

6.2 AXIALLY SYMMETRIC INTEGRAL EQUATIONS

We now resume our considerations of Chapter I, our aim being to solve the integral equations in Sections 1.6–1.8 by means of the techniques outlined in Section 6.1.1. In a more general way than just for the example of Eq. (6.9), this is possible if the boundaries are *rotationally symmetric*, as was discussed in Chapter II. We now describe the positions $\mathbf{r}$ and $\mathbf{r}'$ in *cylindric* coordinates (z, r, φ) and (z', r', φ') respectively. The surface parameters (u, v) are the arc length s' in a meridional cut C through the boundaries and the azimuth φ' . The essential gain will be the reduction to *one* dimension by means of Fourier analysis, which implies that memory limitation becomes far less drastic than in the general case. In fact, the first electron optical application of the BEM by Singer and Braun [1], concerned electrostatic round lenses as the method of Section 6.1 was hardly feasible at that time for reasons of lack of memory.

Apart from the evident restriction to rotationally symmetric boundaries we now encounter the difficulty of evaluating singular nonelementary integrals that requires special techniques. As in the earlier section, the general theory will be outlined first and the most important technical applications will then be studied.

6.2.1 Fourier Analysis of Integral Equations

Quite generally the formulation of Eq. (6.1a) in cylindrical coordinates and for rotationally symmetric boundaries result in

$$\oint_C \int_0^{2\pi} \bar{K}(z, r, \varphi; z', r', \varphi') \bar{Q}(z', r', \varphi') d\varphi' r' ds' = \bar{\lambda}(z, r, \varphi) \bar{Q}(z, r, \varphi) + \bar{P}(z, r, \varphi). \quad (6.54)$$

A Fourier analysis of such an integral equation makes sense only under two assumptions:

- (i) The kernel $\bar{K}$ depends only on the absolute *difference* $\psi = |\varphi - \varphi'|$ of azimuths
- (ii) The coefficient $\bar{\lambda}$ is independent of the azimuth, which means rotationally symmetry.

The former assumption is satisfied for the Green function $G(\mathbf{r}, \mathbf{r}')$ (Eq. (1.88)), its normal derivative $p(\mathbf{r}, \mathbf{r}')$ (Eq. (1.90)), and the Green function of waves (Eq. (1.126)) as kernels; this hence raises no problem. The coefficient $\bar{\lambda}$ results from *material* properties; see for example Eq. (1.115). Condition (ii) implies hence that the geometric configuration of the materials must be rotationally symmetric, which is a stronger restriction. For example, it is *not* satisfied in magnetic deflection yokes that still have round surfaces but angular sectors with different permeabilities. Such configurations must be excluded here, but a nontrivial allowed angular dependence survives in the driving term $\bar{P}(z, r, \varphi)$.

We now introduce the Fourier-series expansions

$$\bar{Q}(z, r, \varphi) = \sum_{n=-\infty}^{\infty} \bar{Q}_n(z, r) \exp(in\varphi), \quad (6.55a)$$

$$\bar{P}(z, r, \varphi) = \sum_{n=-\infty}^{\infty} \bar{P}_n(z, r) \exp(in\varphi), \quad (6.55b)$$

which must be absolutely convergent. Moreover, the reality conditions

$$\bar{Q}_{-n} = \bar{Q}_n^*, \quad \bar{P}_{-n} = \bar{P}_n^*, \quad (6.55c)$$

must be satisfied. The calculation is now straightforward: Eqs. (6.55) are introduced into Eq. (6.54). The latter is then multiplied by $\exp(-im\varphi)$ on both sides and integrated over φ . Then the well-known orthogonality relations of trigonometric functions can be used and a common factor 2π cancelled out on both sides. The result of this calculation can be written concisely as a sequence of integral equations

$$\oint \bar{K}_m(z, r; z', r') \bar{Q}_m(z', r') r' ds' = \bar{\lambda}(z, r) \bar{Q}_m(z, r) + \bar{P}_m(z, r), \quad (m = 0, \pm 1, \pm 2, \dots), \quad (6.56)$$

with the Fourier kernels

$$\overline{K}_m(z, r; z', r') = 2 \int_0^\pi \overline{K}(z, r; z', r'; \psi) \cos m\psi \, d\psi \quad (6.57)$$

It is of great importance that the integral equations (6.56) be uncoupled, so that they can be solved *independently*. Strictly speaking, the set of them is infinite but, for most practical purposes, only the Fourier components up to $|m| = 12$ are needed. This method then brings a significant gain.

The explicitly one-dimensional form is obtained by introducing a suitable *parametrization* $(z(t), r(t))$ of the meridional curve C . If this curve consists of several disjoint parts, these are to be mapped into different nonoverlapping t -intervals with tiny gaps, say 10^{-16} . The whole domain may be $0 \leq t \leq T$. Derivatives with respect to the parameter will be denoted by dots. The choice of the arc length s as such a parameter is possible but not always favorable.

In the light of this parametrization, we are motivated to define transformed functions

$$K_m(t, t') := \overline{K}_m(z(t), r(t); z(t'), r(t')), \quad (6.58a)$$

$$Q_m(t) := \overline{Q}_m(z(t), r(t)), \quad (6.58b)$$

and similarly for $P_m(t)$ and $\lambda(t)$, whereupon Eq. (6.56) now takes now the form

$$\int_0^T K_m(t, t') Q_m(t') r(t') \dot{s}(t') \, dt' = \lambda(t) Q_m(t) + P_m(t). \quad (6.59)$$

This is a Fredholm equation of second kind that includes the first kind as the special case $\lambda(t) \equiv 0$. As the label m is not essential for understanding the following presentation and to avoid double indexing we shall drop it now; it will be restored whenever this is really necessary.

In general, Fredholm equations with transformed Green functions as singular kernels cannot be solved in closed form. The very few cases in which this is fortunately possible, can serve for testing computer programs. The most frequently employed numerical methods are the same as in Section 6.1, namely collocation and Galerkin's method. In analogy to Eq. (6.2), we write now

$$Q(t) = \sum_{j=1}^N S_j N_j(t), \quad (0 \leq t \leq T). \quad (6.60)$$

With respect to the *collocation* we choose now

$$S_j := Q_j = Q(t_j), \quad P_j = P(t_j), \quad (j = 1, \dots, N), \quad (6.61a)$$

$$N_j(t_i) = \delta_{ij}, \quad (i, j = 1, \dots, N). \quad (6.61b)$$

The sampling abscissae t_j must be ordered in strictly monotonic sequence. The trial functions $N_j(t)$ must be chosen in agreement with given symmetry properties of the system in question. In the general case of many disjoint curves, this can lead to some tedious considerations, which we shall ignore here for reasons of conciseness. The condition that the Fredholm equation shall be satisfied at the selected sampling points again leads to a linear system of equations like Eq. (6.6), but now with the matrix elements

$$A_{ij} = \int_0^T K(t_i, t) N_j(t) r \dot{s} dt - \lambda(t_i) \delta_{ij}. \quad (6.62)$$

The formulation of the Galerkin method in analogy to Eqs. (6.7) and (6.8) is straightforward but exhibits one essential disadvantage: the factor $r(t')\dot{s}(t')$ in Eq. (6.59) destroys a possible symmetry of the matrix elements, so that the Cholesky algorithm cannot be used. A suitable transform is therefore necessary to eliminate this asymmetry. This is particularly advantageous if the coefficient λ vanishes, which is the case in the application to Dirichlet problems. We then define new functions of t :

$$U(t) := r^{1/2} P, \quad \phi(t) := r^{1/2} \dot{s} Q \quad (6.63a)$$

and the kernel function

$$H(t, t') = \sqrt{r(t)r(t')} K(t, t'), \quad (6.63b)$$

whereupon Fredholm's equation of the first kind assumes the normalized form

$$\int_0^T H(t, t') \phi(t') dt' = U(t), \quad (6.64)$$

in which the symmetry is conserved. The series expansion (6.60) is now to be replaced by the equivalent expansion

$$\phi(t) = \sum_{j=1}^N \hat{S}_j N_j(t). \quad (6.65)$$

This might possibly require a suitable modification in the set of trial functions, particularly in the vicinity of the optic axis.

The Galerkin approximation is now given by

$$\sum_{j=1}^N \langle N_i | H | N_j \rangle \hat{S}_j = \langle N_i | U \rangle, \quad (i = 1, \dots, N), \quad (6.66)$$

in which all favorable properties of this method are conserved. It will later turn out that the separation of the square root factor in Eq. (6.63a) is quite favorable.

6.2.2 Properties of the Fourier-Green Functions

An important class of kernel functions is obtained from the introduction of the natural Green function (1.88) into Eq. (6.57). These functions — known as the Fourier-Green functions — are explicitly written as

$$G_m(z, r; z', r') = \frac{1}{2\pi} \int_0^\pi \frac{\cos m\psi d\psi}{((z - z')^2 + r^2 + r'^2 - 2rr' \cos \psi)^{1/2}}. \quad (6.67)$$

Because these and their partial derivatives are to be evaluated quite frequently in solutions of integral equations or in the subsequent field calculations, it is necessary to study their analytic properties. Many interesting properties arise in connection with those of the complete elliptic integrals, but for reasons of conciseness we shall keep their presentation as short as possible.

Symmetries

The following properties can already easily be verified by considering the form of the integrand in Eq. (6.67) without representation by an elliptic integral

$$G_m(z, r; z', r') = G_{-m}(z, r; z', r') = G_{|m|}(z, r; z', r'), \quad (6.68a)$$

$$G_m(z, r; z', r') = G_m(z', r'; z, r) =: \hat{G}_m((z - z')^2, r, r'), \quad (6.68b)$$

$$G_m(z, r; z', r') = G_m(z, r'; z', r) = G_m(z', r; z, r'), \quad (6.68c)$$

$$\begin{aligned} G_m(z, r; z', r') &= (-1)^m G_m(z, -r; z', r') = (-1)^m G_m(z, r; z', -r') \\ &= G_m(z, -r; z', -r'). \end{aligned} \quad (6.68d)$$

The latter property allows a definition in the entire coordinate plane.

Differential Equations and Integral Properties

By repeated partial differentiations of the integrand in Eq. (6.67), it is easy to verify that the function G_m satisfies a PDE

$$G_{m|zz} + G_{m|rr} + \frac{1}{r} G_{m|r} - \frac{m^2}{r^2} G_m = \frac{-1}{r} \delta(z - z') \delta(r - r'), \quad (6.69)$$

and, owing to the symmetry (6.68a), a corresponding one with respect to the variables z' and r' . The delta functions on the right-hand side result from

the fact that G_m becomes singular for $z = z'$ and $r = r'$; the exact strength of this singularity will emerge from later considerations. The strength of this singularity is independent of the Fourier order m .

The term in m^2/r^2 can be eliminated by introducing a suitable factor that does not destroy the symmetry. This leads to a new definition

$$P_m(z, r; z', r') := (rr')^{-|m|} G_m(z, r; z', r'), \quad (6.70)$$

from which we obtain the PDE

$$\Delta_\alpha P_m = \left(\partial_{zz} + \partial_{rr} + \frac{\alpha}{r} \partial_r \right) P_m = -r^{-\alpha} \delta(z - z') \delta(r - r') \quad (6.71a)$$

with $\alpha = 2|m| + 1$, or equivalently in self-adjoint form

$$\partial_z(r^\alpha P_{m|z}) + \partial_r(r^\alpha P_{m|r}) = -\delta(z - z') \delta(r - r'). \quad (6.71b)$$

This is nothing but the cylindrical Poisson equation with a point singularity as source, hence the basic properties of its solution, outlined in Section 2.4, are known. In particular the function is finite and even with respect to r and r' at the optic axis. For $m = \pm 1$ even the positive exponent in Eq. (6.70), leading to $\alpha = -1$ makes sense; then the *flux* potential is obtained, which is useful for the calculation of magnetic lenses.

It is now quite easy to find an integral theorem by means of the Gauss-Stokes theorem in the (z, r) plane. For any closed domain D with boundary line C in the upper half plane we obtain in turn

$$\begin{aligned} \oint_C r^\alpha \mathbf{n} \cdot \nabla P_m ds &= \iint_D [\partial_z(r^\alpha P_{m|z}) + \partial_r(r^\alpha P_{m|r})] ds dr \\ &= - \iint_D \delta(z - z') \delta(r - r') dz dr = -\beta(z', r'), \end{aligned} \quad (6.72)$$

$\beta(z', r')$ being the step function of Eqs. (1.91) and (1.92). Thanks to Eq. (6.68a) we can now assume that $m \geq 0$.

Moduli and Elliptic Integrals

If we consider the point (z, r) as reference point for a potential field, then we can consider the function $G_m(z, r, z', r')$ as the static potential of a harmonically charged *ring* located in a plane $z' = \text{const.}$ and having radius r' , as shown in Fig. 6.9. To cast this in a convenient form, it is favorable to introduce the minimum and maximum distance from the ring

$$d_{1,2} := [(z - z')^2 + (r \mp r')^2]^{1/2} \quad (6.73a)$$

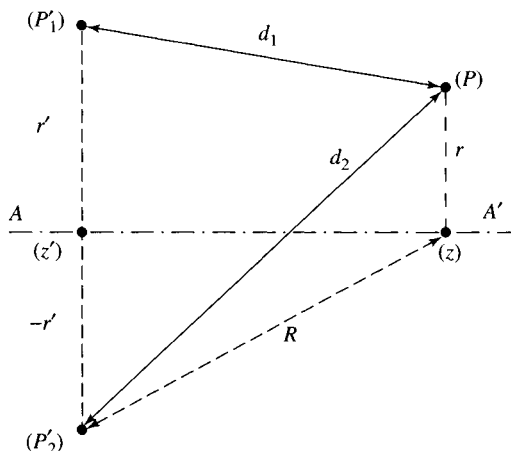


FIGURE 6.9 Geometrical relations between an arbitrary reference point P and the two ring positions P'_1 and P'_2 in the same meridional plane, $\overline{AA'}$ denoting the optic axis. The distances d_1 , d_2 , and R are essential parameters in the theory.

and their frequently appearing arithmetic mean

$$d_a := (d_1 + d_2)/2. \quad (6.73b)$$

These, and hence all regular functions of them, are *symmetric* with respect to the coordinate pairs (z, r) and (z', r') , which may hence exchange their roles.

Using these distances, we can introduce the familiar moduli

$$\bar{k} := d_1/d_2, \quad k := \sqrt{1 - \bar{k}^2} = 2\sqrt{rr'}/d_2. \quad (6.74)$$

Then the substitution $\psi = 2\alpha$ in Eq. (6.67) leads to a representation by elliptic integrals in the form

$$G_m(z, r; z', r') = (\pi d_2)^{-1} I_m(k), \quad (6.75a)$$

$$I_m(k) = \int_0^{\pi/2} \frac{\cos 2m\alpha \, d\alpha}{(1 - k^2 \cos^2 \alpha)^{1/2}}. \quad (6.75b)$$

The representatives of lowest order are the familiar complete elliptic integrals in their Legendre form:

$$I_0(k) = K(k), \quad (6.76a)$$

$$I_1(k) = 2D(k) - K(k) \equiv 2k^{-2}[K(k) - E(k)] - K(k). \quad (6.76b)$$

The functions of higher orders are then related to these by the recurrence relation ($m > 1$)

$$\left(1 + \frac{1}{2m}\right) I_{m+1} - 2(2k^{-2} - 1)I_m + \left(1 - \frac{1}{2m}\right) I_{m-1} = 0, \quad (6.76c)$$

which can be derived by means of the addition theorems and suitable partial integrations. This recurrence relation is numerically instable but can be used as a tridiagonal system of equations, if I_0 and the function I_M of highest order are known.

In the most important domain near the optic axis, this representation is unfavorable, and a better one is obtained by introduction of new moduli

$$\rho := \frac{d_2 - d_1}{d_2 + d_1} \equiv rr'd_a^{-2} \equiv \frac{1 - \bar{k}}{1 + \bar{k}}, \quad (6.77a)$$

$$\bar{\rho} := \sqrt{1 - \rho^2} = d_a^{-1} \sqrt{d_1 d_2}, \quad (6.77b)$$

which are again symmetric in (z, r) and (z', r') .

The kernel function G_m can then be rewritten as

$$G_m = \frac{1}{2\pi d_a} \int_0^\pi \frac{\cos m\psi \, d\psi}{(1 - 2\rho \cos \psi + \rho^2)^{1/2}}. \quad (6.78)$$

This formula is particularly suitable for series expansions, as we shall see later. Another useful form is given by

$$G_m = \frac{\rho^m}{\pi d_a} K_m(\rho) \quad (6.79a)$$

with another kind of complete elliptic integrals:

$$K_m(\rho) = \int_0^{\pi/2} \frac{\sin^{2m} \alpha \, d\alpha}{(1 - \rho^2 \sin^2 \alpha)^{1/2}}. \quad (6.79b)$$

The lowest order is now given by

$$K_0(\rho) = K(\rho), \quad (6.80a)$$

$$K_1(\rho) = D(\rho) = [K(\rho) - E(\rho)]/\rho^2. \quad (6.80b)$$

The relation between this representation and the former one is given by

$$I_m(k) = (1 + \rho)\rho^m K_m(\rho) \equiv (1 + \rho)\hat{K}_m(\rho). \quad (6.81)$$

In context with Eqs. (6.77), this is known as a *Landen transform* (see [20], Chapter 17). The recurrence relation for the functions $\rho^m K_m$ can now easily be derived from Eq. (6.76c) by considering the relation

$$4/k^2 - 2 = \rho + 1/\rho; \quad (6.82)$$

the basic instability of a solution in ascending sequence can, however, not be removed.

Series Expansion

The denominator in Eq. (6.78) can easily be expanded in terms of Legendre polynomials P_n :

$$(1 - 2\rho \cos \psi + \rho^2)^{-1/2} = \sum_{n=0}^{\infty} \rho^n P_n(\cos \psi). \quad (6.83)$$

The integration over ψ becomes easy and results in a series expansion

$$K_m(\rho) = \frac{\pi}{2} \sum_{n=0}^{\infty} F_n F_{n+m} \rho^{2n} \quad (6.84)$$

with the half-integer binomial coefficients

$$F_n := \left| \binom{-1/2}{n} \right| = 1, \frac{1}{2}, \frac{3}{8}, \frac{5}{16}, \frac{35}{128}, \dots \quad (6.85a)$$

These can quite easily be determined up to high orders by means of the recurrence algorithm

$$F_0 = 1, F_n = F_{n-1} \left(1 - \frac{1}{2n} \right), \quad n = 1, 2, 3, \dots \quad (6.85b)$$

This series expansion (6.84) is very favorable for small values of ρ^2 , say $\rho^2 \leq 0.01$. At larger values its convergence becomes slower, so that $\rho^2 = 0.25$ is a practical upper limit for its use, however the convergence can be enhanced significantly by means of *Aitken acceleration*. This procedure is well known in numerical analysis and proceeds as follows.

Let $a_0, a_1, a_2, \dots$, be a sequence of numbers that converges absolutely and asymptotically like a geometric sequence. Then the modified sequence

$$a'_n = a_n - \frac{(a_n - a_{n+1})^2}{a_n - 2a_{n+1} + a_{n+2}} \quad (6.86)$$

converges significantly faster to the same limit. If we apply this twice to the partial sums of Eq. (6.84), an evaluation for $\rho^2 \leq 0.65$ is feasible with reasonable effort.

Axial Derivatives and Paraxial Expansion

From Eq. (6.84) in the context of Eq. (6.79a), the axial potential

$$\phi_m(z) \equiv \phi_m^{(0)}(z) := \lim_{r \rightarrow 0} [r^{-m} G_m(z, r; z', r')] = \frac{1}{2} (r')^m F_m R^{-2m-1}, \quad (m \geq 0) \quad (6.87a)$$

is obtained as a function of z ,

$$R = [(z - z')^2 + r'^2]^{1/2} \quad (6.87b)$$

being the distance of the axial position $(z, 0)$ from the periphery of the ring (see Fig. 6.9). For its derivatives, here scaled by

$$\phi_m^{(n)}(z) := \frac{1}{n!} d^n \phi_m(z) / dz^n, \quad (n \geq 1), \quad (6.88a)$$

a recurrence algorithm can be established, starting with

$$\phi_m^{(1)}(z) = - (2m + 1)(z - z') R^{-2} \phi_m^{(0)}(z), \quad (6.88b)$$

and thereafter for $n \geq 1$ in an ascending sequence:

$$(n + 1) R^2 \phi_m^{(n+1)} = - (2m + 1 + 2n)(z - z') \phi_m^{(n)} - (2m + n) \phi_m^{(n-1)}. \quad (6.88c)$$

This recurrence scheme is numerically stable, so that it is quite easy to determine derivatives of high orders. The corresponding paraxial series expansion (2.72) with $\alpha = 2m + 1$ can then be evaluated. This is of great importance with respect to particle-optical calculations of focusing properties and aberrations, which require explicitly axial derivatives of high orders: in this respect, the BEM is particularly advantageous.

Partial Derivatives

The explicit analytical form of the derivatives $G_{m|z}$ and $G_{m|r}$ for arbitrary values of z and r is required for the solution of integral equations in solving normal derivatives and in field calculations based on their solution. Such field calculations are, for instance, necessary in the determination of equipotentials and in ray tracing.

The easiest way to obtain these derivatives is differentiation under the integral in Eq. (6.67) and the subsequent introduction of suitable moduli. Another way is to differentiate Eqs. (6.79). Both methods must lead to the same final formulas. The results of these longer analytical calculations can be cast in the convenient form

$$G_{m|z} = \frac{(z' - z)\rho^m}{\pi d_a d_1 d_2} J_m(\rho), \quad (6.89a)$$

$$G_{m|r} = \frac{(r' - r)\rho^m}{\pi d_a d_1 d_2} J_m(\rho) + \frac{r'\rho^{m-1}}{2\pi d_a^3} \left(\frac{d_1}{d_2} J_m - K_m \right), \quad (m > 0) \quad (6.89b)$$

$$G_{0|r} = \frac{r}{\pi d_a d_1 d_2} \left[\left(\frac{r'}{d_a} \right)^2 J_1 - J_0 \right], \quad (6.89c)$$

with a new kind of elliptic integrals

$$\begin{aligned} J_m(\rho) &:= \frac{1 - \rho^2}{2\rho^m} \int_0^\pi \frac{\cos m\psi \, d\psi}{(1 - 2\rho \cos \psi + \rho^2)^{3/2}} \\ &\equiv (2m + 1)K_m(\rho) + 2\rho \dot{K}_m(\rho), \end{aligned} \quad (6.90)$$

the dot denoting the derivative with respect to ρ .

From the integral representation it can be seen that the singularity will be stronger than logarithmic, whereas the second form shows that the function is finite for $\rho \rightarrow 0$ and even in ρ .

The two most important representatives of this family are

$$J_0(\rho) = 2E(\rho)/\bar{\rho}^2 - K(\rho), \quad (6.91a)$$

$$J_1(\rho) = 2E(\rho)/\bar{\rho}^2 - D(\rho). \quad (6.91b)$$

Considering Eqs. (6.77), we see that the strength of the singularity is like d_1^{-1} ; this holds also for all numbers of this family. For sufficiently small values of ρ^2 the power series expansion (6.84) can be considered, leading to

$$J_m(\rho) = \frac{\pi}{2} \sum_{n=0}^{\infty} F_n F_{n+m} (2m + 4n + 1) \rho^{2n}, \quad (6.92)$$

which converges more slowly than Eq. (6.84).

The analytical differentiations can be carried on to the second order by considering the self-adjoint differential equation for K_m as a function of $p := \rho^2$

$$\frac{d}{dp} \left(p^{m+1} (1-p) \frac{dK_m}{dp} \right) - \frac{2m+1}{4} p^m K_m = 0. \quad (6.93)$$

The resulting more complicated formulas will not be given here for reasons of space. In the paraxial domain, the differentiation of the paraxial series expansion is much easier.

Singularity Analysis

As is well known in the mathematical literature, the complete elliptic integrals become *logarithmically* singular if the modulus ρ or k approaches unity. The corresponding series expansions for the familiar functions E , K , and D can be found in any comprehensive set of tables and are therefore not stated here. Using these, Ströer [16–18] has determined the corresponding series expansions for the Fourier-Green functions and their normal derivative. Depending on the modulus used, they can take different but equivalent forms.

Here we shall present a series expansion that is better adapted to the kernel (6.63b). With respect to the separated square-root factor, it is favorable to define a new modulus

$$q := d_1^2 / 4rr' = \bar{k}^2 / (1 - \bar{k}^2), \quad (6.94)$$

which can be used only for $rr' > 0$. The transform of the series expansion for G_m then results in

$$\begin{aligned} 2\pi\sqrt{rr'}G_m &=: H_m(q) \\ &= \left[1 + \left(m^2 - \frac{1}{4} \right) q + \left(m^2 - \frac{1}{4} \right) \left(m^2 - \frac{9}{4} \right) \frac{q^2}{4} + O(q^3) \right] L_m(q) \\ &\quad + \left(m^2 + \frac{1}{4} \right) q + \left(m^4 - \frac{7}{6}m^2 - \frac{7}{16} \right) \cdot \frac{3}{8}q^2 + O(q^3), \end{aligned} \quad (6.95a)$$

$$L_m(q) := \frac{1}{2} \ln(16/q) - \sum_{n=1}^m{}' 1/(n - 1/2), \quad (6.95b)$$

the prime indicating that this summation excludes $m = 0$.

From this approximation, it is obvious that the kernel $H(t, t')$ in Eq. (6.63b) becomes singular like

$$H(t, t') = \frac{-1}{4\pi} \ln(t - t')^2 + R(t, t'), \quad (6.96)$$

where $R(t, t')$ is the *finite* remainder. The coefficient of this singularity is obviously independent of the multipole label m . In contrast to this favorable result, the approximation (6.95) causes the difficulty that the series expansions are *semi*convergent, which means that they are useful only for $m^2 q \leq 0.25$, if $m > 0$. This demonstrates the difficulty of achieving high accuracy for $m \gg 1$.

The corresponding approximations for the partial derivatives $G_{m|z}$ and $G_{m|r}$ can now be easily obtained by differentiation of Eqs. (6.95) with respect to z and r . In the lowest order we find:

$$G_{m|z} = \left(\frac{-1}{2\pi\bar{r}d_1^2} + \frac{m^2 - 1/4}{4\pi\bar{r}^3} L_m(q) \right) (z - z'), \quad (6.97a)$$

$$G_{m|r} = \left(\frac{-1}{2\pi\bar{r}d_1^2} + \frac{m^2 - 1/4}{4\pi\bar{r}^3} L_m(q) \right) (r - r') + \frac{1 - L_m(q)}{4\pi r \bar{r}}, \quad (6.97b)$$

where the geometric mean $\bar{r} = (r, r')^{1/2}$. The strongest singularity like d_1^{-2} cancels out by forming the *principal value*, as is required in the integral equations of Sections 1.6 and 1.7, but to find all relevant terms, the parametric form of the boundary curve is now necessary. In the vicinity of the reference point, a Taylor series expansion up to the second order suffices:

$$z(t') \equiv z' = z + \tau\dot{z} + \tau^2\ddot{z}/2, \quad (6.98a)$$

$$r(t') \equiv r' = r + \tau\dot{r} + \tau^2\ddot{r}/2, \quad \tau = t' - t. \quad (6.98b)$$

In this approximation, we have a 'velocity'

$$v = (\dot{z}^2 + \dot{r}^2)^{1/2}, \quad d_1 = v|\tau|, \quad (6.99)$$

and consequently,

$$L_m(q) = L_m^c - \ln|\tau|, \quad L_m^c = \ln(8r/v) - \sum_{n=1}^m \frac{1}{n - 1/2}. \quad (6.100)$$

In the practical calculations, it is favorable to introduce a surface normal N not having unit length, this is then simply given by

$$N_z = \dot{r}, \quad N_r = -\dot{z}. \quad (6.101)$$

By means of the Taylor-series expansion, it can easily verified that the relevant scalar product is

$$P := N \cdot (r - r') = (\dot{z}\ddot{r} - \dot{r}\ddot{z})\tau^2/2 = \kappa v^3 \tau^2/2, \quad (6.102)$$

κ being the *curvature* of the boundary. Using this formula together with Eq. (6.97), we obtain in lowest order the result

$$\bar{r}N \cdot \nabla G_m = -\kappa v/4\pi + \frac{\dot{z}}{4\pi r}(L_m(q) - 1), \quad (6.103)$$

which means that the normal derivative becomes *logarithmically* singular, but now with an amplitude that depends on the shape of the boundary. This has the consequence that the solution of integral equations involving this term is necessarily more complicated than that of Eq. (6.64) with the singularity (6.96).

The Flux Kernel

For reasons of completeness, we present here briefly the most important properties of the Green function for magnetic fluxes, briefly referred to as the flux kernel. This will again be symmetric and satisfy the PDE (6.71) with $\alpha = -1$, fitting Eq. (2.48) with unit source. This function is found to be

$$\Phi(z, r; z', r') = rr'G_1(z, r; z', r'), \quad (6.104)$$

as can easily be verified. In this case, Eq. (6.95) must now be modified to,

$$H_{-1}(q) = 2\pi(rr')^{-1/2}\Phi(z, r; z', r') = H_1(q), \quad (6.105)$$

to arrive at Eq. (6.96); hence the square root factor in Eq. (6.63b) must be modified accordingly. The formula for the normal derivative is essentially modified by the radial factors as follows:

$$\frac{1}{\bar{r}}N \cdot \nabla \Phi = -\frac{\kappa v}{4\pi} - \frac{\dot{z}}{4\pi r}(L_1(q) + 1). \quad (6.106)$$

It is of importance that the amplitude of the logarithmic term has changed its sign here. Another property that is quite useful in the solution of integral equations is the integral theorem

$$\oint_c \frac{1}{r} \mathbf{n} \cdot \nabla \Phi \, ds = \oint_c \frac{1}{r} N \cdot \nabla \Phi \, dt = -1/2 \quad (6.107)$$

valid for any reference point (z', r') on the boundary. This relation can be incorporated into the numerical calculations to achieve conservation of the total lens currents.

Numerical Calculation

There are essentially *four* methods of calculation (apart from numerical integration, which is generally far too slow for practical use):

- (i) The evaluation of the power series expansion (6.84) with double Aitken acceleration according to Eq. (6.86), if $(\rho) \leq 0.8$;
- (ii) The evaluation of a recurrence formula like Eq. (6.76), if $\rho > 0.7$;
- (iii) A logarithmic Chebyshev approximation;
- (iv) The familiar method of iterated arithmetic and geometric means for the basic elliptic integrals $E(\rho)$, $K(\rho)$, and $D(\rho)$.

Method (ii) requires the knowledge of K and D as initial values that must be calculated as accurately as possible. With method (iv) only 4 to 7 iterations are necessary to reach machine accuracy. Thereafter, the recurrence scheme can be used in two different but equivalent versions [19]:

(a) Original Functions K_m

$$K_0 = K(\rho), \quad K_1 = D(\rho), \quad (\text{start}) \quad (6.108a)$$

$$K_{m+1} = \left[(1 + \rho^2)K_m - \left(1 - \frac{1}{2m}\right)K_{m-1} \right] \bigg/ \left[\left(1 + \frac{1}{2m}\right)\rho^2 \right],$$

$$(1 \leq m < M \leq 50). \quad (6.108b)$$

(b) Transformed Functions

$$\hat{K}_m = \rho^m K_m. \quad (6.109)$$

The corresponding scheme is given by

$$\hat{K}_0 = K(\rho), \quad \hat{K}_1 = \rho D(\rho), \quad (\text{start}) \quad (6.110a)$$

$$\hat{K}_{m+1} = \left[(\rho + 1/\rho)\hat{K}_m - \left(1 - \frac{1}{2m}\right)\hat{K}_{m-1} \right] \bigg/ \left(1 + \frac{1}{2m}\right),$$

$$(1 \leq m < M \leq 50). \quad (6.110b)$$

After division by the factor ρ^M , both forms are *equivalent* with respect to accuracy and stability. An important observation that was found empirically by the author, is the fact that their deviations are of nearly equal absolute amount and have opposite signs. This suggests the cancellation by arithmetic averaging according to

$$K_M(\rho) = (K_M + \rho^{-M}\hat{K}_M)/2. \quad (6.111)$$

In this way, the error can be reduced considerably, as is shown in Table 6.2

The presented errors are the differences $K_m - K_m^{(s)}$, $K_m^{(s)}$ being the result of the power-series expansion. This accuracy is sufficient for all practical purposes.

TABLE 6.2
TEST OF ACCURACY FOR SQUARE MODULUS 0.64

Order	Kernel	Recurrence	Averaging
0	1.995303	+8.67639e-12	+8.67639e-12
5	0.606144	+6.40998e-12	+6.40943e-12
10	0.445103	+5.26368e-12	+5.25635e-12
15	0.368737	+4.64223e-12	+4.57973e-12
20	0.321821	+4.64645e-12	+4.13658e-12
25	0.289242	+8.08753e-12	+3.85753e-12
30	0.264916	+3.98114e-11	+4.00041e-12
35	0.245857	+3.15014e-10	+7.20429e-12
40	0.230400	+2.71203e-09	+3.67066e-11
45	0.217537	+2.37462e-08	+2.96940e-10
50	0.206614	+2.09486e-07	+2.59607e-09

The third method, the logarithmic Chebyshev approximation, consists in an approximation of the kind

$$K_m(\rho) = \sum_{n=0}^L (A_{m,n} - B_{m,n} \ln \varepsilon) \varepsilon^n, \\ (\varepsilon := \bar{\rho}^2 = 1 - \rho^2, B_{m,0} = 1/2, m = 0, 1, 2, \dots). \quad (6.112)$$

The coefficients $A_{m,n}$ and $B_{m,n}$ are to be determined in such a way that, for a given value of L , the maximal absolute error is minimized. Formulas of this kind can be found in Abramowitz and Stegun [20], pp. 591–2. Ströer [16] has improved these for $L = 7$ and $m = 0, \dots, 12$. The corresponding coefficients are given in the Appendix.

The author has tested this approximation, the results for one particular value $\rho^2 = 0.5$ are given in Table 6.3.

In this case, the power series expansion $K_m^{(s)}$ is more accurate, which explains that then also the recurrence algorithm gives better results than in Table 6.2. However, even the Chebyshev approximation is accurate enough for practical purposes. Moreover, it has the advantage of being the fastest of all methods.

Graphs of the functions are presented in Fig. 6.10 and Fig. 6.11. They demonstrate the singular behavior at $\rho^2 = 1$, and the very slow decrease for large values of the order m . This is an effect that remains confined to the vicinity of the boundary because towards the optic axis the additional factor ρ^m in Eq. (6.79a) causes a strong damping.

The calculation of the integrals of the *second* kind, $J_m(\rho)$, can be performed in an analogous manner. For $\rho \leq 0.8$ the power series expansion with double Aitken acceleration can be used, which has practically the same speed of

TABLE 6.3
TEST OF ACCURACY FOR SQUARE MODULUS 0.5

Order	Kernel	Recurrence	Chebyshev
0	1.854075	+2.52243e-13	+2.29594e-13
1	1.006862	+2.29372e-13	-1.44329e-14
2	0.777673	+2.10831e-13	-7.73381e-13
3	0.658182	+1.95288e-13	-2.48956e-12
4	0.581507	+1.81521e-13	-5.60063e-12
5	0.526845	+1.68754e-13	-1.04270e-11
6	0.485295	+1.55709e-13	-1.70041e-11
7	0.452308	+1.40388e-13	-2.48752e-11
8	0.425286	+1.19182e-13	-3.31508e-11
9	0.402616	+8.61533e-14	-4.03526e-11
10	0.383239	+2.90323e-14	-4.39041e-11
11	0.366426	-7.45515e-14	-4.08995e-11
12	0.351655	-2.68674e-13	-2.93192e-11

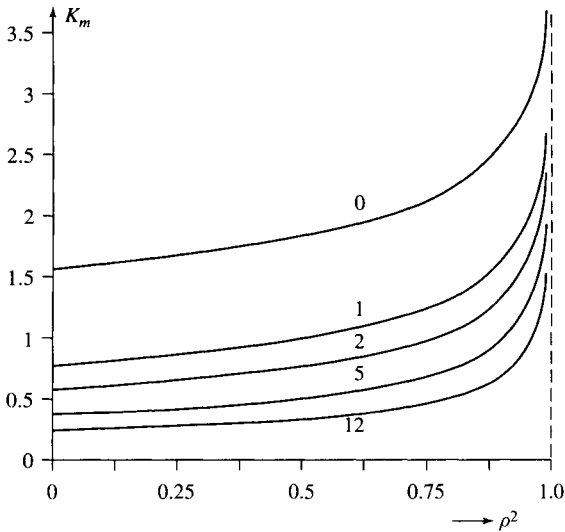


FIGURE 6.10 Elliptic Fourier integrals K_m as functions of the square modulus ρ^2 for $m = 0, 1, 2, 5$, and 12 , respectively.

convergence. For $m \leq 12$ the second form of Eq. (6.90) can also be evaluated in the context of Eq. (6.112), which gives slightly less accurate, but still quite satisfactory results. Moreover—especially for $m > 12$ and $\rho \geq 0.8$ —a strikingly simple recurrence formula can be applied:

$$J_{m+1} - J_m = -(2m + 1)(K_{m+1} - K_m). \tag{6.113}$$

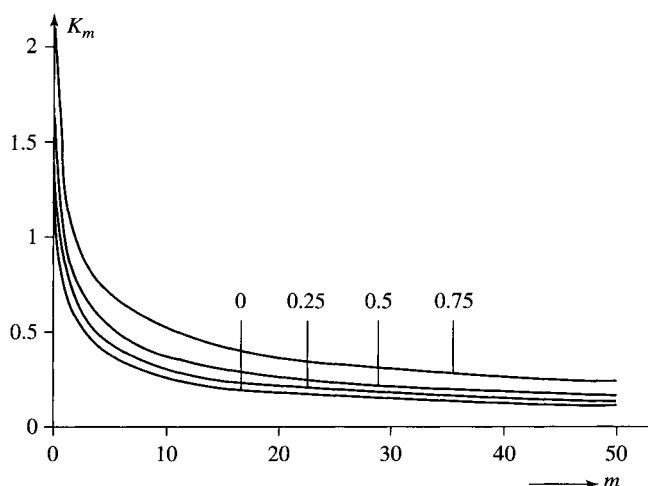


FIGURE 6.11 Elliptic Fourier integrals K_m as functions of the order m for $\rho^2 = 0, 0.25, 0.5$, and 0.75 , respectively.

This can be explicitly verified for a few terms of low order m , especially for $m = 0$ with Eqs. (6.91). For large values $m \leq 50$ it was assumed and then verified numerically, but a rigorous proof should be possible. In summary, all the necessary functions can be evaluated with sufficient precision. A comparison of the different methods of calculating the magnetic field of an current-conducting ring ($m = 1$) can be found in von der Weth [21].

6.3 NUMERICAL SOLUTION OF INTEGRAL EQUATIONS

The techniques for the numerical solution of such integral equations as are derived in Section 6.2.1 are the logical continuation of this topic, but we have devoted a separate section to this task as there will appear now some new aspects. We set out from a general form (6.64), recalled here for conciseness and generalized to

$$\int_0^T H(t, t') \phi(t') dt' - \lambda(t) \phi(t) = U(t), \quad (0 \leq t \leq T). \quad (6.114)$$

At first we disregard any special properties of the kernel H , and assume only that it might become logarithmically singular for $t' \rightarrow t$. Moreover, we assume that Eq. (6.114) has a unique solution, which implies that we are not concerned with an eigenvalue problem.

6.3.1 Basic Collocation Techniques

The simplest approximations are piecewise constant or linear trial functions in each integration element, as shown in Fig. 6.12a,b. The linear approximation is essentially the same as the evaluation of Eqs. (6.60) to (6.62) with the piecewise linear functions $N_j(t)$.

Although the result is a smoother function than the approximation by rectangles, the accuracy will *not* generally be better. This is analogous and related to the difference between the trapezoidal rule and the midpoint rule of integration. In the present context, the use of the latter leads to significant simplifications:

- (i) Because the function values at the interval endpoints are not used as sampling data, it is not necessary to distinguish between the various kinds of boundary conditions at the curve endpoints; in systems with many curves having different properties at their endpoints especially, this brings a considerable simplification of the program structure.
- (ii) The integration over the singularity of $H(t, t')$ is easier as this singularity is always located at the midpoint of an interval and the symmetry properties can hence be exploited. For these reasons we suggest that this method should be used unless better than linear approximations are available.

The linear system to be solved now takes the fairly simple form

$$\sum_{k=1}^N A_{ik} \phi_k = U_i, \quad (i = 1, \dots, N), \quad (6.115)$$

where the sampling data U_i and ϕ_i ($i = 1, \dots, N$) referring to the midpoints t_i . Let h_k be the half-length of the interval with label k , then the matrix elements

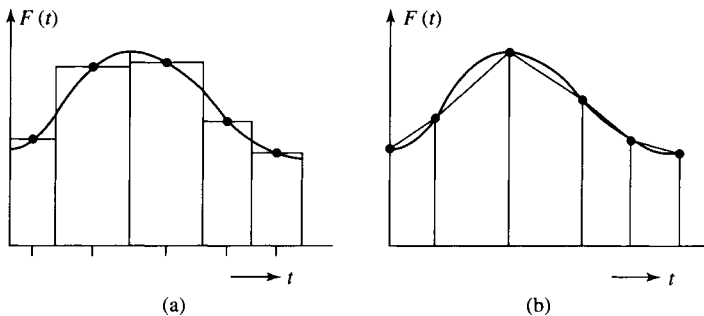


FIGURE 6.12 Elementary methods of integration: (a) midpoint rule with piecewise constant approximation; (b) linear approximation in each interval. (The size of the intervals is exaggerated for clarity.)

are given by

$$A_{ik} = h_k \int_{-1}^1 H(t_i, t_k + h_k \tau) d\tau - \lambda(t_i) \delta_{ik}. \quad (6.116)$$

The integrations required in the off-diagonals can be performed by means of Gauss quadratures. The corresponding order can be lowered with increasing mean distance between the corresponding intervals. In adjacent intervals $|i - k| = 1$, a subdivision is necessary for the reasons outlined in Section 3.6.1. The determination of the diagonal elements requires some special considerations. It is possible to apply adaptative procedures as outlined in Section 3.6.3, provided that there remains a tiny but still finite distance to the singularity. This has the disadvantage that very many function calls are necessary to achieve a sufficient accuracy. A much better way is the development of special quadrature formulas fitting this task.

Let us consider an arbitrary function of the basic form

$$F(x) = f_1(x) \ln |x| + f_2(x)/x + f_3(x) \quad (6.117)$$

with smooth functions which may be approximated by polynomials in x . The quadrature formula in question may have the symmetric form

$$\int_{-h}^h F(x) dx = h \sum_{m=1}^N w_m [F(p_m h) + F(-p_m h)] + O(h^{2N+1}), \quad (6.118)$$

whereupon the x^{-1} singularity is already considered exactly, and, moreover, all other odd contributions vanish exactly for reasons of symmetry. There remains then the task of solving the nonlinear system of equations for the relative positions p_m and weights w_m :

$$\begin{aligned} \int_0^1 t^{2n} dt &= \sum_{m=1}^N w_m p_m^{2n} = (2n+1)^{-1}, \\ - \int_0^1 t^{2n} \ln t dt &= - \sum_{m=1}^N w_m p_m^{2n} \ln p_m = (2n+1)^{-2}, \\ (n &= 0, 1, 2, \dots, N-1). \end{aligned} \quad (6.119)$$

The solution for $N = 4$ is given in Hawkes and Kasper [22]. The still more accurate result for $N = 6$ is presented in Table 6.4.

This is very accurate; for instance, the error of the integration

$$- \int_{\pi/2}^{\pi/2} \ln(|\sin x|) dx = \pi \ln 2 = 2.177586 \dots, \quad (6.120)$$

turns out to be -6.419×10^{-10} , which is very small indeed.

TABLE 6.4
SYMMETRIC INTEGRATIONS

Sampling Positions				Integration Weights			
0.0186	9798	8085	7164	0.0599	5864	4701	3795
0.1377	1897	2966	3167	0.1709	2669	2136	9927
0.3465	1321	7828	6855	0.2377	9187	0666	4534
0.5935	4231	5320	0500	0.2458	2349	4206	6578
0.8177	0525	4328	2321	0.1930	0907	5657	2154
0.9635	0620	1178	6623	0.0924	9022	2631	3011
$N = 6$; Integral = $1/13$:				Error $-5.49094\text{e-}7$			
Integral = $1/169$:				Error $+9.43469\text{e-}7$			

After the system (6.115) has been solved by means of the *LU* algorithm the function $\phi(t)$ is known as a *step function* and is hence constant in each interval. If necessary, it can be smoothed in such a way that in each interval the integral over it remains conserved.

6.3.2 Collocation Techniques Using Splines

In spite of its striking simplicity, the earlier outlined method has the drawback that it requires very many intervals and hence a high rank N to obtain an acceptable accuracy. The approximation can be improved by using polynomial approximations of higher orders in each interval with imposition of continuity conditions on the endpoints; see for example [22,23].

A very favorable method is the use of cubic splines in combination with the integral equation [24,25]. This implies that the interval ends $t_0, \dots, t_N$ are now to be used as sampling positions and that a cubic Hermite polynomial is to be used in each interval. The method becomes more sophisticated because of the fact that now the boundary conditions at the various curve-end points, for example, positive mirror symmetry or cyclic closure, must explicitly be considered. Moreover, great care must be taken when dealing with sharp corners. A naive technique can result in unphysical oscillations (Gibbs phenomenon); this can be avoided only by slightly rounding-off these corners (see Section 3.4.3 and Fig. 3.13) and using sufficiently small interval sizes. For reasons of conciseness we shall not discuss here all these various conditions and confine the presentation to the regular situation.

Since the spline matrices are sparse, it is not necessary to store them. With a view to an easy later elimination, the derivatives at the interval ends are here denoted by:

$$\phi_{N+k} \hat{=} \dot{\phi}_k \equiv d\phi/dt, \quad (t = t_k, k = 1, \dots, N). \quad (6.121)$$

The cubic Hermite polynomial, to be used in each interval $t_{k-1} \leq t \leq t_k$, ($1 \leq k \leq N$), can now be written as

$$\phi(t) = \phi_{k-1} + (\phi_k - \phi_{k-1})F_k(t) + \bar{D}_k(t)\dot{\phi}_{k-1} + D_k(t)\dot{\phi}_k. \quad (6.122)$$

The form functions can easily be determined, as follows: it is favorable to introduce a normalized variable

$$u := (t - t_{k-1})/h_k, \quad h_k := t_k - t_{k-1} \equiv q_k^{-1}, \quad (6.123a)$$

h_k being here the *full* interval length. Then we have to evaluate the expressions

$$\begin{aligned} F_k(t) &= u^2 (3 - u), \\ D_k(t) &= (t - t_{k-1}) (1 - u)^2, \\ \bar{D}_k(t) &= (t - t_{k-1}) (u^2 - u). \end{aligned} \quad (6.123b)$$

Let $\Phi = (\phi_0, \dots, \phi_N)^T$ denote the sampling vector of function values, and similarly for $\dot{\Phi}$ and U . Then the complete system of equations can be cast in the concise form

$$(P - \Lambda)\Phi + Q\dot{\Phi} = U, \quad (6.124a)$$

$$R\Phi + S\dot{\Phi} = 0. \quad (6.124b)$$

The upper equation results from the discretization of Eq. (6.114), whereas the lower one is the necessary additional spline condition. This partitioning is quite favorable, as we shall see later. The matrix elements without special boundary conditions, which form the vast majority, are found to be

$$\begin{aligned} P_{ik} &= \int_{t_{k-1}}^{t_k} H(t_i, t)F_k(t) dt + \int_{t_k}^{t_{k+1}} H(t_i, t)(1 - F_{k+1}(t)) dt, \\ Q_{ik} &= \int_{t_{k-1}}^{t_k} H(t_i, t)D_k(t) dt + \int_{t_k}^{t_{k+1}} H(t_i, t)\bar{D}_{k+1}(t) dt. \end{aligned} \quad (6.125)$$

These are now to be completed by the matrix elements resulting from boundary conditions. Here we shall present only one simple but ever-frequently appearing case: a *cyclic* condition at the endpoints:

$$\phi(0) = \phi(T), \quad \dot{\phi}(0) = \dot{\phi}(T), \quad U(0) = U(T). \quad (6.126)$$

It is then favorable to eliminate ϕ_0 and $\dot{\phi}_0$ so that the numbering starts with unity. In C-programs that require it to begin with zero, it would be favorable to

eliminate ϕ_N and $\dot{\phi}_N$. Equations (6.125) are then to be evaluated for $1 \leq i \leq N$ and $1 \leq k \leq N-1$, and to be completed by

$$\begin{aligned} P_{iN} &= \int_{t_{N-1}}^T H(t_i, t) F_N(t) dt + \int_0^{t_1} H(t_i, t) (1 - F_1(t)) dt, \\ Q_{iN} &= \int_{t_{N-1}}^T H(t_i, t) D_N(t) dt + \int_0^{t_1} H(t_i, t) \bar{D}_1(t) dt. \end{aligned} \quad (6.127)$$

It is favorable to perform all integrations referring to the same interval simultaneously and to accumulate the matrix elements accordingly as the evaluation of the kernel H is the most time-consuming process.

The matrix Λ resulting from the term with $\lambda(t)$ in Eq. (6.114) is a pure diagonal matrix with the elements

$$\Lambda_{ik} = \lambda(t_i) \delta_{ik}, \quad (6.128)$$

which appear already in Eq. (6.116). The spline matrices S and R are *tri-diagonal* ones that result from Eq. (3.99) with appropriate notation:

$$\begin{aligned} S_{i,i-1} &= q_i, \quad S_{i,i} = 2(q_i + q_{i+1}), \quad S_{i,i+1} = q_{i+1}, \\ R_{i,i-1} &= 3q_i^2, \quad R_{i,i} = 3(-q_i^2 + q_{i+1}^2), \quad R_{i,i+1} = -3q_{i+1}^2, \\ (2 \leq i \leq N-1). \end{aligned} \quad (6.129)$$

In the case of a cyclic boundary, these are to be completed by the corresponding elements

$$\begin{aligned} S_{1,1} &= 2(q_1 + q_2), \quad S_{1,2} = q_2, \quad S_{1,N} = q_1, \\ S_{N,1} &= q_1, \quad S_{N,N-1} = q_N, \quad S_{N,N} = 2(q_N + q_1), \end{aligned} \quad (6.130a)$$

and similarly

$$\begin{aligned} R_{1,1} &= 3(q_2^2 - q_1^2), \quad R_{1,2} = -3q_2^2, \quad R_{1,N} = 3q_1^2, \\ R_{N,1} &= -3q_1^2, \quad R_{N,N-1} = 3q_N^2, \quad R_{N,N} = 3(q_1^2 - q_N^2), \end{aligned} \quad (6.130b)$$

which means that R and S are *cyclic* tridiagonal matrices, S being symmetric and positively definite, hence invertible. The complete system matrix, to be built up according to Eqs. (6.124a,b) is now well defined, and we must assume that it is nonsingular and sufficiently well conditioned.

There are now different ways of solving system of Eqs. (6.124). The simplest, though not the best one is the direct implementation and solution with

a matrix of the size $2N \times 2N$, which is a waste of memory and computation time. A better way is *an iterative technique*.

We start with an initial guess for Φ , for instance the one obtained by solution of Eq. (6.124a) with $\dot{\Phi} = 0$. This is then introduced into Eq. (6.124b), which can easily be solved for $\dot{\Phi}$ by means of the well-known spline technique. This vector is then substituted in Eq. (6.124a), which supplies now a better approximation for Φ , and so on, until sufficient convergence is achieved.

The drawback of this method is that very many matrix-vector multiplication may be needed for large rank N . This can be avoided by direct *block-elimination* of $\dot{\Phi}$. Because the matrix S is certainly regular, we can write

$$\dot{\Phi} = -S^{-1}R \cdot \Phi, \quad (6.131)$$

and introduce this into Eq. (6.124a), whereupon we obtain

$$A\Phi \equiv (P - \Lambda - QS^{-1}R) \cdot \Phi = U \quad (6.132)$$

with a matrix A of rank N . This system can be solved by means of the familiar *LU* algorithm (see Section 5.6.1), whereupon the vector Φ is now known.

The inverse S^{-1} loses the favorable property of sparseness, but in Eq. (6.132) this does not increase the memory requirement. However, if Φ is determined, it is better to apply the Cholesky algorithm to Eq. (6.124b) instead of using the inverse again as the former is faster.

In summary, this method can be implemented in an efficient form, and if the λ -values are sufficiently far from the eigenvalues of the system, the obtained results are mostly sufficient. The necessary numerical integrations cannot be performed by means of Eq. (6.118); however, Ströer [26] has derived an analogous summation formula for one-sided singularities. Let

$$C(x) = A(x) + B(x) \ln x, \quad (6.133)$$

be the function to be integrated over the unit interval $0 \leq x \leq 1$, then this integration can be carried out by a quadrature of the form

$$\int_0^1 C(x) dx = \sum_{i=1}^n g_i C(x_i). \quad (6.134)$$

Ströer has determined the corresponding abscissae x_i and weights g_i for $n = 6, 9$, and 15 sampling positions; the corresponding results are given in the appendix. Their derivation is much more complicated than that of Table 6.4. The integration over other intervals than the one assumed above is obtained by appropriate scale transformation. A special solution technique for magnetic lenses has been developed by Becker [27,28].

6.3.3 The Galerkin Method

The principle of this method has already been outlined in Section 6.2.1, and is given by Eqs. (6.66) and (6.65) or more generally in Section 6.1.1. There remains only the task of finding a favorable set of trial functions $N_j(t)$, and of performing the necessary *double* integrations over functions that might become logarithmically singular. Again many cases are to be distinguished in realistic applications with many disjoint boundary curves and different possibilities of geometrical symmetries, and again we shall confine the presentation to the simple case of one closed boundary, as we did in the earlier section.

There are many ways of choosing the trial functions $N_j(t)$. Linear ones are often too inaccurate and cubic Hermite polynomials are then the next alternative. These were tested by Ströer [26,29], who obtained very good results with them. The final solution is then only once continuously differentiable at the nodes, but this is sufficient for most practical applications. On the other hand, these derivatives are now additional degrees of freedom for minimization of the mean square error, so that a better approximation than with cubic splines is to be expected.

Because the derivatives will now *not* be eliminated, it is more advantageous to introduce a *sequential* numbering, in which each node has *two* degrees of freedom, where odd ones refer to function values and even ones to derivatives.

The labellings then need not be altered if further boundaries are added to a configuration already given.

The determination of the system matrix, the so-called matrix compilation is favorably performed element wise, as in the FEM. This means that initially, all matrix elements are set to zero. Then in a loop over all elements, here the t -intervals, all four matrix contributions arising in the same element, are determined simultaneously and added to the corresponding matrix element. Let the interval $t_{k-1} \leq t \leq t_k$ be coordinated with the variables

$$\hat{S}_{n+1} = \phi(t_{k-1}), \quad \hat{S}_{n+2} = \dot{\phi}(t_{k-1}), \quad \hat{S}_{n+3} = \phi(t_k), \quad \hat{S}_{n+4} = \dot{\phi}(t_k) \quad (6.135)$$

as boundary values, n being the last label of the preceding part. Then the trial functions $N_j(t)$ are to be chosen as

$$\begin{aligned} N_{n+1}(t) &= F_k(t), & N_{n+2}(t) &= D_k(t), \\ N_{n+3}(t) &= 1 - F_k(t), & N_{n+4}(t) &= \bar{D}_k(t), \end{aligned} \quad (6.136)$$

with the Hermite functions given by Eqs. (6.123). With these functions as factors, the integrals

$$|H|N_j\rangle = \int_{t_{k-1}}^{t_k} H(t, t') N_j(t') dt', \quad j = n+1, \dots, n+4, \quad (6.137)$$

are to be performed, t being kept fixed. This expression is then a common factor in the four integrals over the variable t , the other factor being a form function determined in an analogous manner. Altogether in each pair of different intervals, there are thus 16 double-integrals to be determined simultaneously and added to the corresponding matrix elements. In adjacent or identical intervals, some of these integrals contribute to the same matrix elements.

In the latter case, there arises the difficulty that a singularity is reached, but this does not remain at an interval endpoint, hence Eq. (6.134) cannot immediately be applied. This difficulty is demonstrated in Figs. 6.13a,b. It is favorable to introduce again the normalized coordinate u from Eq. (6.123), and a corresponding coordinate u' related to t' . Then an integral of the form

$$I = \int_{u=0}^1 \int_{u'=0}^1 [a(u, u') + b(u, u') \ln(u - u')^2] du du' \quad (6.138)$$

is to be determined, whose singularity is located on the diagonal of the hatched square in Fig. 6.13a. Ströer [16] solved this problem by transformation to the coordinates v and v' , oriented in diagonal directions:

$$u = (v + v')/2, \quad u' = (v' - v)/2. \quad (6.139)$$

The transformation of the coefficients become

$$A(v, v') = a\left(\frac{v' + v}{2}, \frac{v' - v}{2}\right), \quad B(v, v') = b\left(\frac{v' + v}{2}, \frac{v' - v}{2}\right), \quad (6.140)$$

and with these functions, the integral can be written as

$$I = \frac{1}{2} \int_{v=0}^1 \left[\int_{v'=v}^{2-v} (A(v, v') + A(-v, v') + (B(v, v') + B(-v, v')) \ln v^2) dv' \right] dv. \quad (6.141)$$

The inner integral over v' is now to be carried out first, which can be achieved by suitable Gauss quadratures. Thereafter there remains an integrand having the basic form of eq. (6.133); the resulting integral can be evaluated by means of the quadrature formula (6.134). In this context it is not necessary to know the coefficient $b(u, u')$ of the logarithm exactly. All that is necessary is to determine the appropriate sampling positions and weights.

The square in the third quadrant of Fig. 6.13a is treated in the same manner, after the coordinates u' and v' have been shifted accordingly. The square-shaped domains in the second and fourth quadrant require a slight modification.

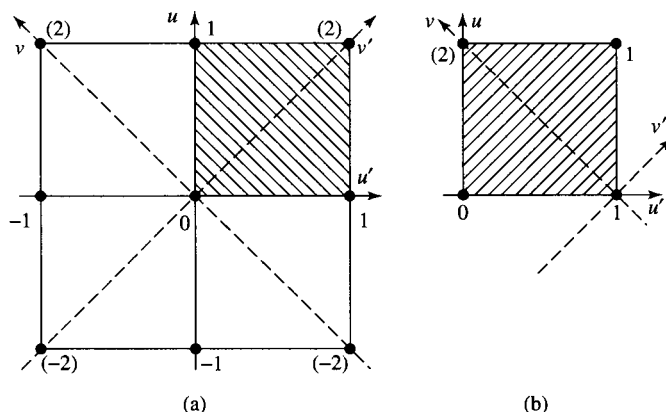


FIGURE 6.13 Choice of local normalized coordinate systems for the double integration over a function having its singularity along the diagonal chosen as the v' -axis ($v = 0$): (a) the hatched area refers to the same interval; (b) the hatched area refers to adjacent intervals; the origins of both coordinates are not the same in this case. The numbers in parentheses are function values of the coordinates v' or v , respectively.

If the coordinate u' is shifted in such a way that $0 \leq u' \leq 1$ is again valid, as is required in Eq. (6.123a), then the situation of Fig. 6.13b is given, in which the line of singularities reaches the domain only at the point $(1,0)$, the common endpoint of the adjacent t -intervals. The appropriate coordinate transform is now

$$u = (v + v')/2, \quad u' = (v' - v)/2 + 1. \quad (6.142)$$

The transformed coefficients hence become

$$\begin{aligned} A(v, v') &= a\left(\frac{v' + v}{2}, \frac{v' - v}{2} + 1\right), \\ B(v, v') &= b\left(\frac{v' + v}{2}, \frac{v' - v}{2} + 1\right), \end{aligned} \quad (6.143)$$

and the integral is then

$$\begin{aligned} I' &= \int_0^1 \int_0^1 (a(u, u') + b(u, u') \ln(u - u' + 1)^2) du' du \\ &= \frac{1}{2} \int_{v=0}^1 \left[\int_{v'=-v}^v \{A(v, v') + B(v, v') \ln v^2\} dv' \right] dv \\ &\quad + \frac{1}{2} \int_{v=1}^2 \left[\int_{v'=v-2}^{2-v} \{A(v, v') + B(v, v') \ln v^2\} dv' \right] dv. \end{aligned} \quad (6.144)$$

The integrand in the second contribution is a regular function; this integral can hence be evaluated by ordinary Gauss quadratures. The first part finally takes the form of Eq. (6.133), which is integrated by means of Eq. (6.134). The integration over the fourth quadrant in Fig. 6.13a, ($u' \geq 0, u \leq 0$) is performed in an analogous manner, hence the complicated task of double integration is altogether soluble and the Galerkin method thus feasible.

An alternative version of the method consists in the use of the modified interpolation kernels as trial functions. In their nonequidistant form, dealt with in Section 3.3.4, these are still fairly general. The local approximation is certainly less accurate, but because the rank of the system matrix is now halved in comparison to the case of Eq. (6.136), the interval size can be halved to obtain the same final size. Then this version might become equivalent or even more accurate.

On the other hand, the use of two degrees of freedom per node implies a greater flexibility of the Galerkin method, so that it is fairly easy to link cubic Hermite elements with other special types, for instance, singular ones near sharp edges (Ströer [26,29]) or others in the vicinity of symmetry planes. In summary the outlined method is a very favorable one.

6.3.4 A Fast Method for Symmetric Integral Equations

The following method was first published by Kasper [30] and later improved gradually by the author. Here we shall present an unpublished but tested version, which also takes into account the second term in the singularity analysis given by Eq. (6.95a).

We set out from the special form of Eq. (6.64), in which the kernel $H(t, t')$ is symmetric and satisfies Eq. (6.96). Because the amplitude of the singularity is independent of the interval size in t or t' , it is always possible to choose equidistant t -intervals of unit length. Moreover, it turns out that it is favorable to locate the sampling positions in the midpoints of these intervals, hence $T = N$ is the rank of the system, no matter how many different symmetry conditions at curve endpoints are given. The sampling data are hence

$$\phi_n := \phi(n - 1/2), \quad U_n := U(n - 1/2), \quad n = 1, 2, \dots, N. \quad (6.145)$$

Subsequently, we again confine the presentation to the case of one closed boundary, as we have done before; the generalization to several boundaries is not a problem, but simply a matter of appropriate indexing. The assumed periodicity implies

$$\phi(t \pm N) = \phi(t), \quad U(t \pm N) = U(t), \quad (6.146a)$$

$$\phi_{n \pm N} = \phi_n, \quad U_{n \pm N} = U_n. \quad (6.146b)$$

It is convenient to introduce a temporary variable $\tau := t + 1/2$. We can use then Eq. (3.113) and write down more concisely

$$\phi(t) = \sum_{n=-\infty}^{\infty} \phi_n F(\tau - n), \quad (6.147a)$$

$$U(\phi) = \sum_{n=-\infty}^{\infty} U_n F(\tau - n). \quad (6.147b)$$

The factor $F(\tau - n)$ will be a *modified* interpolation kernel, as defined in Section 3.3. For conciseness, we shall drop here the order M , which is kept fixed during the whole calculations; the final data refer to $M = 6$. The summations are only virtually extended to infinity; in reality just $2M$ terms do not vanish. The concept of analytical continuation is shown in Fig. 6.14.

On introducing the expansion (6.147a) into Eq. (6.64), now referring to the variable $\tau' = t' + 1/2$, considering that the kernel $H(t, t')$ must have the same periodicity properties in both variables, and converting the formally infinite summation into an infinite integral, we arrive at the integral relation

$$\sum_{n=1}^N \int_{-\infty}^{\infty} H(\tau, \tau') F(\tau' - n) d\tau' \phi_n = U(\tau). \quad (6.148)$$

The evaluation at the sampling positions $\tau = 1, 2, \dots, N$ is already a collocation method that is more accurate than that of Section 6.3.1, but it is not necessary to terminate the development of the method at this point.

According to Galerkin's technique, we now multiply Eq. (6.148) in turn with $F(\tau - j)$, $j = 1, \dots, N$, and integrate also over τ , thus arriving at

$$\sum_{n=1}^N H_{j,n} \phi_n = \bar{U}_j, \quad (j = 1, 2, \dots, N), \quad (6.149)$$

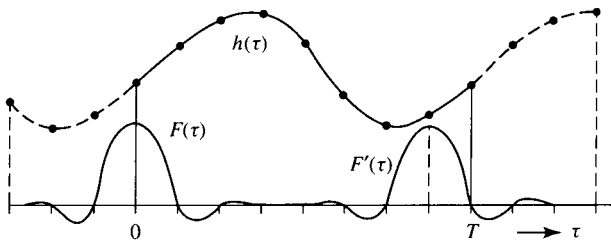


FIGURE 6.14 Analytical continuation of a periodic function $h(\tau)$ in such a way that the scalar products with $F(\tau)$ and $F'(\tau) := F(\tau - T + 1)$ can be evaluated. In this case the cyclic distance D is unity

with the *symmetric* matrix elements

$$H_{j,n} = \iint_{-\infty}^{\infty} F(\tau - j)H(\tau, \tau')F(\tau' - n) d\tau' d\tau \quad (6.150)$$

and the inhomogeneities

$$\bar{U}_j = \int_{-\infty}^{\infty} U(\tau)F(\tau - j) d\tau \approx U_j. \quad (6.151)$$

The latter approximation is justified by Eq. (3.199) and is fairly uncritical. So far no further going simplifications were made.

The kernel H cannot be integrated in this simple manner, because it becomes logarithmically singular. We therefore separate this singularity from the regular remainder and write

$$H(\tau, \tau') = R(\tau, \tau') + S(\tau, \tau'). \quad (6.152)$$

The integration over $R(\tau, \tau')$ is easy owing to the assumed regularity and results in

$$R_{jn} = R(\tau = j, \tau' = n). \quad (6.153)$$

There is thus only one evaluation of the function $R(\tau, \tau')$ at the specified pair of arguments. In the case of *off-diagonal* elements ($j \neq n$), this value is not even required explicitly, as we can evaluate Eq. (6.152) for the arguments $\tau = j, \tau' = n$ and use this equation to eliminate R_{jn} , thus arriving at

$$H_{j,n} = H(j, n) + \iint_{-\infty}^{\infty} F(\tau - j)[S(\tau, \tau') - S(j, n)]F(\tau' - n) d\tau' d\tau, \\ (j \neq n). \quad (6.154a)$$

The determination of the diagonal elements ($j = n$) requires a slight modification: because $H(n, n)$ does not exist, we shift the second argument by a tiny value ε , say $\varepsilon = 10^{-7}$, on both sides and take the average of the two function values. Then we obtain the approximation

$$H_{n,n} = \frac{1}{2}[H(n, n + \varepsilon) + H(n, n - \varepsilon)] + O(\varepsilon^2) \\ + \iint_{-\infty}^{\infty} F(\tau - n) \left[S(\tau, \tau') - \frac{1}{2}S(n, n + \varepsilon) - \frac{1}{2}S(n, n - \varepsilon) \right] \\ \times F(\tau' - n) d\tau' d\tau. \quad (6.154b)$$

The terms of second order are proportional to the partial derivatives of R , and hence so small that they can perfectly well be ignored.

So far, this concept is very general and assumes only that a relation like Eq. (6.152) does exist. This separation is fairly uncritical, because any constant or smooth function, incorporated in $S(\tau, \tau')$, cancels out from Eqs. (6.154). Moreover, from Eq. (6.154a) it can be concluded that $H_{jn} \rightarrow H(j, n)$ for $|n - j| \gg 1$, because then the integrand is far from its singularity, so that the double integral vanishes.

The further evaluations require an explicit specification of the function $S(\tau, \tau')$. If we consider Eq. (6.96) as the approximation of lowest order and recall that this function must be made periodic, we see that it is favorable to define a *cyclic distance* $D_N(p)$ by

$$D_N(p) := \text{Min}(|p|, |p + N|, |p - N|), \quad (6.155)$$

whereupon the singularity function can be approximated by

$$S(\tau, \tau') = -C [1 + a(\tau, \tau') D_N^2(\tau - \tau')] \ln D_N(\tau - \tau'), \quad (6.156)$$

with $C = 1/2\pi$. The coefficient $a(\tau, \tau')$ of the quadratic term must be slowly variable, so that it can be replaced with a suitable mean value. In the context of Eqs. (6.154), there arises then the task of determining integrals of the form

$$I_n(p) := p^{2n} \ln |p| - \iint_{-\infty}^{\infty} F(u)F(v)(p + u - v)^{2n} \ln |p + u - v| du dv, \\ (n = 0, 1, p = \varepsilon, 1, 2, 3, \dots). \quad (6.157)$$

For a fixed value of the order M (here $M = 6$), these integrals are completely defined and can be determined by tedious numerical integrations. This needs to be done only once, and the results are then stored. Thereafter, the off-diagonal elements become simply

$$H_{j,n} = H(j, n) + C [I_0(D_N(j - n)) + a(j, n)I_1(D_N(j - n))] \quad (6.158a)$$

and the diagonal elements become

$$H_{n,n} = \frac{1}{2} [H(n, n + \varepsilon) + H(n, n - \varepsilon)] + C [I_0(\varepsilon) + a(n, n)I_1(\varepsilon)], \quad (6.158b)$$

whereupon the matrix is complete. It satisfies all the conditions imposed on it: symmetry, cyclicity, and appropriate behavior for large distances D_N . From Table 6.5 it is seen that the correction terms decrease rapidly for large value of p , so that $p \leq 10$ is quite sufficient. It is also possible to write down the

TABLE 6.5
VALUES OF THE INTEGRALS IN EQ. (6.157); HERE $M = 6$ AND
 $\varepsilon = 10^{-7}$ WERE CHOSEN. (WITH $\ln 1/\varepsilon = 16.118\dots$, THE
DIAGONAL ELEMENTS ARE STILL POSITIVE.)

p	Integrals $I_0(p)$	Integrals $I_1(p)$
ε	-14.5097 49155 4243	+0.1360 59852 4303
1	+0.1706 11555 9755	-0.0553 48150 0357
2	-0.0821 13912 3582	+0.0242 75580 4422
3	+0.0356 54366 8550	-0.0079 86563 5596
4	-0.0118 40797 1310	+0.0016 67851 5820
5	+0.0029 50750 7671	-0.0002 13477 5978
6	-0.0005 84608 7241	+0.0000 26730 1330
7	+0.0001 01018 3667	-3.3541 76087 e-06
8	-0.0000 14419 0270	-1.1018 49136 e-07
8	+1.4205 33765 e-06	+4.0118 40815 e-08
10	-9.0105 47108 e-08	-1.2673 86551 e-08
11	+8.0639 58390 e-10	-1.7803 92722 e-09
12	-3.1602 40958 e-10	-7.7506 00162 e-10

matrix in the variable t instead of τ ; this requires simply the replacement $H(j, n) \rightarrow H(j - 1/2, n - 1/2)$ and similarly $a(j - 1/2, n - 1/2)$.

6.3.5 The Solution of Dirichlet Problems

The most important practical application is the solution of Dirichlet problems for multipole fields. This is the special case of Eq. (6.59) for $\lambda \equiv 0$ and the Fourier-Green function G_m from Eq. (6.67) taken for $\bar{K}_m$. As there are further labels to be considered, we shall write now the multipole order m as a superscript; there will be no confusion with powers that do not appear in this context. The sampling positions here refer again to *half-integer* values, and in this context it is convenient to introduce two-dimensional vectors

$$\mathbf{u}_n := (z_n, r_n) \equiv (z(n - 1/2), r(n - 1/2)), \quad (n = 1, \dots, N), \quad (6.159)$$

and subscripts on all other functions will refer to these positions.

It is convenient to remove the square root transforms, given in Eqs. (6.63), and to introduce in turn the surface *potentials*

$$P_j^m = U_j r_j^{-1/2}, \quad (j = 1, \dots, N), \quad (6.160a)$$

the surface charge densities σ_n^m and charges q_n^m

$$q_n^m = r_n \dot{s}_n \sigma_n^m = \phi_n r_n^{1/2}, \quad (n = 1, \dots, N), \quad (6.160b)$$

and the Fourier kernels

$$\overline{G}_{j,n}^m = (r_j r_n)^{-1/2} H_{j,n}, \quad (j, n = 1, \dots, N). \quad (6.160c)$$

The linear system of equations to be solved then takes the physically understandable form

$$\sum_{n=1}^N \overline{G}_{j,n}^m q_n^m = P_j^m. \quad (6.161)$$

The off-diagonal elements are now simply given by

$$\overline{G}_{j,n}^m = G_m(\mathbf{u}_j, \mathbf{u}_n) + \frac{1}{2\pi\sqrt{r_j r_n}} [I_0(D_{j,n}) + a_{j,n}^m I_1(D_{j,n})], \quad (6.162a)$$

and the diagonal elements are given by

$$\begin{aligned} \overline{G}_{n,n}^m &= \frac{1}{2} G_m(\mathbf{u}_n, \mathbf{u}_n + \varepsilon \dot{\mathbf{u}}_n) + \frac{1}{2} G_m(\mathbf{u}_n, \mathbf{u}_n - \varepsilon \dot{\mathbf{u}}_n) \\ &+ \frac{1}{2\pi r_n} [I_0(\varepsilon) + a_{n,n}^m I_1(\varepsilon)]. \end{aligned} \quad (6.162b)$$

There still remains the task of determining the coefficients a_{jn}^m ; these are determined by bringing Eqs. (6.155) and (6.156) into agreement with Eqs. (6.94) and (6.95) up to the second term in the logarithmic series expansion. In this context, all contributions without a factor $\ln q$ can be ignored, because these are incorporated in the remainder R .

The distance d_1 appearing in Eq. (6.94) can be approximated by

$$d_1 \equiv |\mathbf{u}(t) - \mathbf{u}(t')| = |t - t'| \cdot \bar{v}. \quad (6.163)$$

If a matrix element with labels j and n is to be calculated, the integrand in Eqs. (6.154a,b) is largest in the vicinity of the points $t = j - 1/2$ and $t' = n - 1/2$; hence a good approximation for the factor $\bar{v}$ is then

$$\bar{v} = v_{jn} = |\dot{\mathbf{u}}_j + \dot{\mathbf{u}}_n|/2. \quad (6.164)$$

Replacing the denominator rr' in Eq. (6.94) with the constant $r_j r_n$, we obtain the approximation

$$q = \frac{v_{jn}^2}{4r_j r_n} (t - t')^2, \quad (6.165)$$

and the agreement of the essential factors in Eqs. (6.95a) and (6.156) after suitable adaptation

$$1 + (m^2 - 1/4)q \equiv 1 + a_{j,n}^m(t - t')^2 \quad (6.166)$$

is achieved with

$$a_{j,n}^m = (m^2 - 1/4)v_{j,n}^2 / (4r_j r_n). \quad (6.167)$$

This approximation can sometimes give too large values, especially for very large values of m or if the radii r_j or r_n are quite small. To prevent this from happening, we *confine* the matrix elements by setting

$$a_{j,n}^m = \frac{(m^2 - 1/4)v_{j,n}^2}{4r_j r_n + (m^2 + 1/2)v_{j,n}^2}. \quad (6.168)$$

This holds also for $j = n$, that is, for diagonal elements.

6.3.6 Generalizations

The method has been demonstrated for one simple case of application—one closed boundary curve remaining sufficiently distant from the optic axis. There are, however, many ways of generalizing it, and these are all shown in Fig. 6.15. The basic form of Eqs. (6.162) remains the same; only some functions are to be altered.

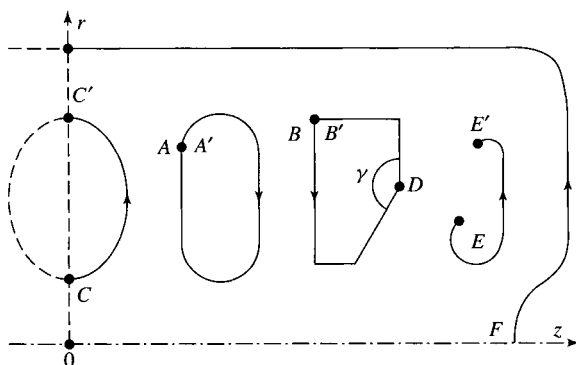


FIGURE 6.15 Different configurations in a complex system of boundaries: A, A' , and B, B' : common endpoints on closed loops; the orientations may be different; C, C' endpoints on a symmetry plane; D sharp edge with inner angle γ ; E, E' endpoints of an open boundary; F axial vertex. In the case of mirror symmetry only one half needs to be defined explicitly

Several Closed Curves (A, A' and B, B'):

Each curve has its own integer period N_A and N_B , the distant function D_N of Eq. (6.155) must then be modified accordingly.

Mirror Symmetry

If the potential has a positive or negative mirror symmetry with respect to the plane $z = 0$, then it is sufficient to discretize one half of the system, and this reduces the necessary memory practically by a factor of 4. Generally, the other half of the system is then considered by a modification of the kernel:

$$G_m^s(z, r; z', r') = G_m(z, r; z', r') \pm G_m(z, r; -z', r'). \quad (6.169)$$

If some curve endpoints C, C' are located at the symmetry plane, then the distance function must be modified accordingly, and the missing parts of the I -integrals must be included. This means the following: let $c < t \leq c'$ be the parameter interval of such a curve, c and c' being integers. Then we have to make the following replacement:

$$\begin{aligned} I_k(D_{j,n}) = & I_k(|j - n|) + I_k(|j + n - 2c + 1|) \\ & + I_k(|2c' + 1 - j - n|), \quad (k = 0, 1). \end{aligned} \quad (6.170)$$

If the length $c' - c$ is sufficiently large, the additional contributions of the two ends do not overlap, so that there then remain only two terms.

Sharp Edges

Generally, this case cannot be dealt with correctly because sharp edges are an unphysical idealization for reasons that have already been explained in Section 2.5.2 and 4.5.2. Within the BEM, a more rigorous method fitting the formalism of Section 6.3.3 has been developed by Ströer [26], who could show that the use of appropriately chosen singular element functions gives very good results. The fast method of Section 6.3.4 seems to be incompatible with the presence of sharp edges, because the necessary assumptions of regular functions are violated. Nevertheless, an attempt can be made to find an acceptable approximation.

We now return to the notation of Fig. 2.8 and Eqs. (2.93a,b) and rename now $r \rightarrow s$. In the close vicinity of a sharp edge, the potential is nearly that of a planar configuration, because the terms produced by the rotational symmetry are corrections of higher orders. Hence, as a result of Eq. (2.93a) the potential is nearly $U - U_0 \sim s^\mu$, and the field strength is proportional to $s^{\mu-1}$. The surface source density must then have the same strength of singularity, and

hence $q = r\dot{s}\sigma \sim s^{\mu-1}\dot{s}$ in agreement with Eq. (6.160b), r being here the finite off-axis distance at the edge. This kind of singularity must be removed by a parametrization $s = \text{const.}t^\lambda$, so that $q(t)$ remains finite, resulting in

$$t^{(\mu-1)\lambda} \cdot t^{\lambda-1} \equiv 1, \quad (6.171)$$

which is satisfied by

$$\lambda = 1/\mu = 2 - \gamma/\pi, \quad (0 \leq \gamma \leq \pi), \quad (6.172)$$

γ being the angle subtended at the edge (Fig. 2.8). If we now locate the position of the edge at an integer value $t_c > 0$, the appropriate discretization becomes the monotonic function

$$s(t) = s_c + a_c |t - t_c|^\lambda \text{ sign } (t - t_c), \quad (6.173)$$

s_c and $a_c > 0$ being free parameters. The edge itself is never a sampling point, because the latter are chosen as $t_n = t_c \pm 1/2, t_c \pm 3/2$, etc.; hence, in fact, singular terms never appear explicitly.

The approximation in Eqs. (6.172) and (6.173) gives correctly $\lambda = 1$ and a linear function $s(t)$ if $\gamma = \pi$, corresponding to a regular boundary. The other limiting case, $\gamma = 0, \lambda = 2$, corresponds to an endpoint of an infinitely thin sheet. This can be considered as an approximation of a very *thin electrode* as shown in Fig. 6.16. This simulation comes close to reality if the charged line is treated as one *open* boundary joining the endpoints A and B and carrying the *sum* of surface charges of both sides. These sums q_n^m are obtained *correctly*,

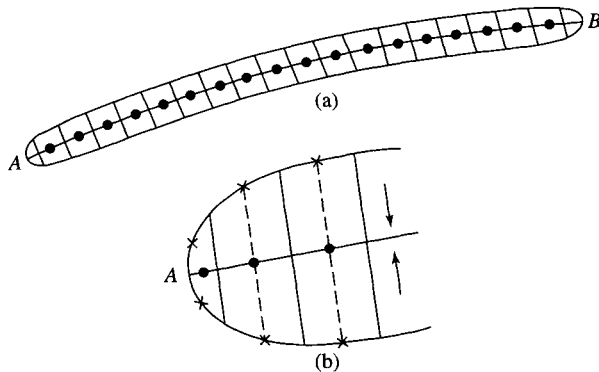


FIGURE 6.16 An open boundary with endpoints A and B as a degenerate case of a closed one, with a thickness that vanishes at the limit. The rings located on the middle surface carry the contributions from both sides. (a) Situation with a still finite, but small thickness; (b) enlarged detail in the vicinity of the endpoint A . The arrows indicate the directions of shrinking

if the discretization law (6.173) with $\lambda = 2$ and the distance function (6.170) are used in the vicinity of both endpoints.

The case of a *rectangular* edge $\gamma = \pi/2$, $\lambda = 3/2$, which cannot be solved correctly, has been tested by the author and has given very good results, so that this simple approximation is, indeed, feasible. The error is strongly reduced if the distance of the two closest neighbour points from the edge is reduced by a factor 0.92, and the slope is increased by a factor 1.05.

Axial Vertices

Sometimes, mainly in electrostatic devices, an electrode may reach the optic axis, thus having there a rotationally symmetric vertex. This is another case in which the method can never be exact for mathematical reasons. However, very good results are obtained if the parametrization of the boundary curve is chosen such that $r(t)$ becomes an *even* function in t : $r(t) = c_2 t^2 + c_4 t^4 + \dots$, and the symmetry rule (6.170) is used in the vicinity of this vertex. The surface charge density σ then remains finite, as it should do. The value in the neighborhood of the vertex ($t = 1/2$) becomes a little too large by a factor of 1.024. This can be corrected. If the function $\sigma(r)$ is then smoothed in the close vicinity of the optic axis in such a way that the total charge remains conserved, very good results are obtained. If $m \geq 2$ and $r(0) = 0$, the lowest radius must be enlarged to $r_1 = r(1/2) + m|\dot{r}(1/2)|/4$ to achieve positive definiteness.

The Simplified Field Calculation

The fast method has some further properties that may be advantageous. Once the linear system in Eq. (6.161) has been solved by means of the Cholesky algorithm, which is always possible in cases of reasonable discretization, because the matrix is then positively definite, the source function $q_m(t)$ is known at half-integer positions $t_n = n - 1/2$, $n = 1, \dots, N$. Then a series expansion like Eq. (6.147a) can be written down for $q_m(t)$. The Coulomb integral for the multipole potential $V_m(z, r)$ at any position $\mathbf{u} := (z, r)$ can be written as

$$\begin{aligned} V_m(\mathbf{u}) &= \int_0^T G_m(\mathbf{u}, \mathbf{u}'(t)) q_m(t) dt + V_m^E(\mathbf{u}) \\ &= \sum_{n=1}^N \int_{-\infty}^{\infty} q_n^m F(\tau - n) G_m(\mathbf{u}, \mathbf{u}'(t)) d\tau + V_m^E(\mathbf{u}), \quad (6.174) \end{aligned}$$

V_m^E being an *external* field, if given.

The arguments for this transformation are quite analogous to those leading to Eq. (6.148). At positions $\mathbf{u}$ that are sufficiently distant from all ring singularities, Eq. (3.199) can be used again, and we then obtain simply

$$V_m(\mathbf{u}) = \sum_{n=1}^N G_m(\mathbf{u}, \mathbf{u}_n) q_n^m + V_m^E(\mathbf{u}). \quad (6.175)$$

With the same reasoning, the formula

$$\nabla V_m(\mathbf{u}) = \sum_{n=1}^N \nabla G_m(\mathbf{u}, \mathbf{u}_n) q_n^m + \nabla V_m^E(\mathbf{u}) \quad (6.176)$$

for the gradient can also be derived.

These formulas can be interpreted as applications of the Euler-Maclaurin formula (3.196) with unit interval size. The boundary conditions are here either periodic or symmetric, so that all terms involving boundary derivatives in Eq. (3.196) cancel out, and this explains the fairly high accuracy.

In the spatial scale, a unit interval size in t corresponds to a local spacing $v = s(t)$ on the surface, and this means that neighboring rings have roughly this distance. In a narrow domain near the surface, the neighboring rings then produce a field like that of a grid with charged wires separated by the distance $h = s$. The periodic field of such a grid decreases as $\exp(-2\pi d/h)$, d being the distance from the grid. This is the physical interpretation of the error term associated with the Euler-Maclaurin formula.

From this result, it emerges that a distance of about $d = 2h$ is necessary to achieve a good accuracy. This implies that the field in the vicinity of a surface cannot be calculated by Eqs. (6.175) and (6.176). Moreover, the method becomes quite unfavorable for systems with very thin plates, covered with rings on *both* sides. If the thickness d is given, the spacing between neighboring rings must not exceed $d/2$. It is then better to use only one layer of rings, as is shown in Fig. 6.16.

The method is particularly well adapted to the requirements of particle optics, as the domain occupied by the rays is usually far distant from the ring singularities.

6.4 SPECIAL TECHNIQUES FOR ASYMMETRIC INTEGRAL EQUATIONS

This section is the logical continuation of the concepts derived in the earlier ones, but for conciseness, we present the material of this section separately because we shall encounter some new aspects once again.

Asymmetric integral equations are always obtained whenever a normal derivative is involved, as this implies the use of a kernel with a singularity like that of Eq. (6.103). These conditions usually appear in cases of applications to *magnetic* fields in systems with rotationally symmetric ferrite yokes. Typical representatives of this are magnetic round lenses and deflection systems with toroidal or saddle coils.

6.4.1 Integral Equation for Round Lenses

In the case of round lenses the use of the flux potential $\psi(z, r)$ (Eq. (2.40)) is particularly advantageous because then the boundary conditions (2.54) are very simple. The most favorable surface variable is the surface current density $\omega(\mathbf{r})$, defined in Section 1.7.2, because a knowledge of ω makes a field calculation possible without solution of further integral equations for other field variables. In the present case this vector ω has only an *azimuthal* component and the integral equation for the latter follows from Eq. (1.124). The transform from the cartesian representation to the cylindrical form leads to a procedure like that of Section 6.2.1, but now the Fourier-series expansion can have only two linearly dependent components with $m = \pm 1$. A corresponding integral equation, even including saturation terms, was published by Ströer [6]. Here we shall omit the saturation terms, as it is extremely difficult to evaluate them.

Using the *flux kernel* Φ from Eq. (6.104), the integral equation can be written as

$$-\frac{1}{r} \oint_Y \mathbf{n} \cdot \nabla \Phi(z, r, z', r') \omega(s') ds' + (\lambda + 1/2) \omega(s) = -H_T(s), \quad (6.177)$$

in which the point $\mathbf{u} = (z, r)$ must be located at the yoke surface Y ; the normal derivative refers to this position and the vector $\mathbf{n}$ directs into the vacuum. The function $H_T(s) = \mathbf{t} \cdot \mathbf{H}_0(z, r)$ is the *tangential* component of the driving field, that is, the field produced by the coils in the absence of the yoke. The constant

$$\lambda = (\mu/\mu_0 - 1)^{-1} \ll 1 \quad (6.178)$$

is the only term containing material properties.

As the arc length s is not the best curve parameter, we introduce a new parameter t and adopt the notation of Section 6.3. Moreover, it is advantageous to use a “velocity” $\mathbf{v} = (\dot{z}, \dot{r}) \equiv \dot{\mathbf{u}}$ with $v = |\mathbf{v}|$ and a new source function

$$J(t) = \omega(s) \dot{s} = v \omega(s), \quad (6.179)$$

which is the *current* per unit interval in t .

The factor r^{-1} in Eq. (6.177) can be removed by a suitable square root transform with $r = r(t)$:

$$\chi(t) := \sqrt{r} J(t) = \sqrt{r} v \omega, \quad (6.180a)$$

$$U(t) := \sqrt{r} v \cdot \mathbf{H}_0 = \sqrt{r} v H_T, \quad (6.180b)$$

$$H_{-1}(t, t') := -(rr')^{-1/2} \mathbf{N} \cdot \nabla \Phi, \quad (6.180c)$$

whereupon Eq. (6.177) can be rewritten as

$$\oint H_{-1}(t, t') \chi(t') dt' + (\lambda + 1/2) \chi(t) = -U(t). \quad (6.181)$$

This integral equation is of Fredholm's second kind and essentially *asymmetric*: owing to Eq. (6.106) this kernel becomes singular like

$$H_{-1}(t, t') = \frac{\kappa v}{4\pi} + \frac{\dot{z}}{4\pi r} (L_1(q) + 1), \quad (6.182)$$

the notation being explained in Section 6.2.2. The amplitude of this singularity depends essentially on the form and parametrization of the boundary curve Y , so that the method of Section 6.3.4 cannot be applied here. However, because Eq. (6.182) is a special case of Eq. (6.114), the general techniques of Section 6.3.1 to 6.3.3 can be used for its solution. It is essential to recall that a singularity like that in Eq. (6.182) is more sensitive with respect to approximation errors than a simple one like $\text{const} \cdot \ln |t - t'|$. On the other hand, the identity (6.107), now becoming

$$\oint \sqrt{r'/r} H_{-1}(t, t') dt = 1/2 \quad (6.183)$$

can be used to improve the accuracy.

6.4.2 Integral Equation for Deflection Systems

The typical structure of magnetic deflection systems is shown in Fig. 6.17. Such a system has a rotationally symmetric ferrite yoke and current-conducting coils with wires that are at least partly oriented in the longitudinal direction. In *toroidal* coils (Fig. 6.17a), these wires form closed loops that are entirely located in *meridional* planes, whereas in saddle coils they are closed by azimuthal arcs along the front faces of the yoke. The figure shows only a strongly simplified form of such systems; in reality they can be much more complicated. In television tubes and monitors—the main field of application

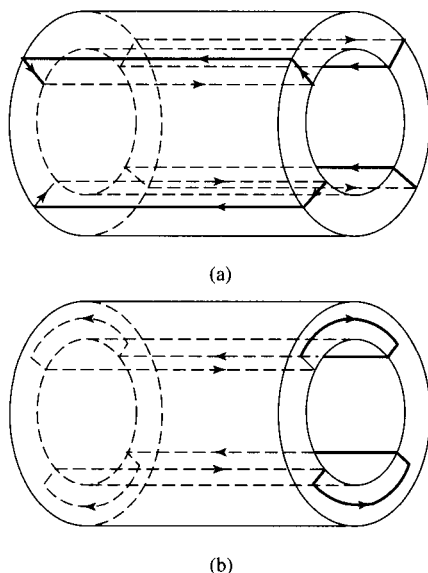


FIGURE 6.17 Schematic presentation of magnetic deflection systems with ferromagnetic cylindrical yoke: (a) system with toroidal coils; (b) system with saddle coils

of such systems — the yoke is hornlike, widened towards the screen; moreover, the wires of the saddle coils do not always have meridional or azimuthal directions, but may be inclined with respect to these. With respect to the following calculations, it is only important that the yoke is rotationally symmetric and that the coils do *not* produce a rotationally symmetric field $\mathbf{H}_0(\mathbf{r})$, like that of a round lens.

There are now two alternatives for describing the field: the solution of the *vector* integral Eq. (1.124) for surface currents or of the *scalar* integral Eq. (1.115) for surface “charges.”

Since the vector equation now has three components, coupled by the continuity condition, the scalar equation is much easier to solve and we therefore consider only the latter. We identify now medium 1 with the yoke and medium 2 with the vacuum, so that the surface normal $\mathbf{n}$ is again directed into the latter. With $\mu_1 =: \mu$, $\mu_2 = \mu_0$, λ from Eq. (6.178), and a suitable exchange of notation, Eq. (1.115) can now be rewritten as

$$\oint \mathbf{n} \cdot \nabla G(\mathbf{r}, \mathbf{r}') \tau(\mathbf{r}') da' + (\lambda + 1/2) \tau(\mathbf{r}) = -H_N(\mathbf{r}), \quad (6.184)$$

$H_N(\mathbf{r}) := \mathbf{n} \cdot \mathbf{H}_0(\mathbf{r})$ being here the *normal* component of the external $\mathbf{H}_0$ field as the driving term. This is again a special case of the general theory, outlined in Section 6.2.1, and we can hence write down the corresponding sequence of

uncoupled reduced integral equations

$$\oint \tau_m(s') \mathbf{n} \cdot \nabla G_m(\mathbf{u}, \mathbf{u}') r' ds' + (\lambda + 1/2) \tau_m(s) = -F_m(s), \quad (6.185)$$

$\tau_m(s)$ denoting the Fourier coefficients of $\tau(\mathbf{r}')$, $F_m(s)$ those of $\mathbf{n} \cdot \mathbf{H}_0$, and G_m the Fourier-Green function, defined by Eq. (6.67).

This representation is easily understandable but is not the best form with respect to numerical solution. Instead, we introduce again a new curve parameter with $s = s(t)$, $v = \dot{s}$, and corresponding new source functions

$$q_m(t) := r v \tau_m(s), \quad (6.186a)$$

$$V_m(t) := r v F_m(s). \quad (6.186b)$$

We define the transformed kernel

$$K_{|m|}(t, t') := r \mathbf{N} \cdot \nabla G_m(\mathbf{u}, \mathbf{u}'), \quad (6.186c)$$

whereupon we obtain a new set of integral equations

$$\oint K_{|m|}(t, t') q_m(t') dt' + \left(\lambda + \frac{1}{2} \right) q_m(t) = -V_m(t). \quad (6.187)$$

There arises now a difficulty: for $m = 0$, that is, for ordinary rotationally symmetric systems, the homogeneous integral equation has an eigenvalue $\lambda_0 = -1/2$. This is a consequence of the general integral theorem (1.116) and now leads to the problem that for $\mu \gg 1$, $\lambda \rightarrow 0$ the inhomogeneous equation becomes ill-conditioned. Hence, we must exclude this case here. For the calculation of magnetic lenses, this difficulty can be circumvented by the use of Eqs. (6.177) or (6.181), in which case this critical eigenvalue is $\lambda_0 = +1/2$. In the analogous case of electrostatic fields in dielectric media, the value $\varepsilon_1/\varepsilon_2$ is fortunately never so large that the corresponding integral equation becomes insoluble, although its condition is not very good.

Finally, it is straightforward to introduce also the square root transform in analogy to Eqs. (6.180), which is favourable for rendering the factor r in Eq. (6.186c) that is symmetric according to (6.103).

The corresponding equations are now given by

$$\phi_m(t) := q_m(t)/\sqrt{r} = v \tau_m \sqrt{r}, \quad (6.188a)$$

$$U_m(t) := V_m(t)/\sqrt{r} = v F_m \sqrt{r}, \quad (6.188b)$$

$$H_m(t, t') := K_{|m|}(t, t') \sqrt{r'/r} = \sqrt{rr'} \mathbf{N} \cdot \nabla G_m, \quad (6.188c)$$

and then

$$\oint H_m(t, t') \phi_m(t') dt' = -(\lambda + 1/2) \phi_m(t) - U_m(t), \quad (6.189)$$

which is again a special case of Eq. (6.114). In all these different versions it is possible to apply the methods of Sections 6.3.1 to 6.3.3. These require careful numerical integrations over the asymmetric logarithmic singularities, which might be tedious.

6.4.3 The Fast Method for Asymmetric Integral Equations

The striking simplicity of the formalism in Eqs. (6.161) and (6.162) and of the subsequent simple field calculation by means of Eqs. (6.175) and (6.176) makes it appropriate to look for an analogous method in the asymmetric case. Such a method has been developed by Kasper and Ströer [18] and works satisfactorily, provided that the discretization function $s(t)$ is chosen properly. As this may be difficult, we shall not outline this method here; instead, we present a simpler version [31]. For conciseness we shall drop again the multipole label m , whenever this is not essential. Again, we assume here a cyclic boundary, dissected into N unit intervals in the variable t .

The first steps in the development of this method are exactly the same as in Section 6.3.4 we can rewrite Eqs. (6.145) to (6.154a) again; the variables having the meanings explained in that section.

A first essential difference arises in the calculation of the diagonal elements: although Eq. (6.152) can always be assumed, the nature of the singularity function now becomes more complicated: considering Eqs. (6.94), (6.95), (6.100), (6.103), and (6.106) in turn, we find that this function is now

$$S(t, t') = C(t) \ln |t - t'| \quad (6.190)$$

with the known amplitude

$$C(t) = -\dot{z}(t) (4\pi r(t))^{-1}. \quad (6.191)$$

The remainder R can be calculated by careful evaluation of the singularity analysis or more easily by four evaluations of the kernel according to

$$\begin{aligned} R(t, t) = & \frac{2}{3} [H(t, t + \varepsilon) + H(t, t - \varepsilon)] - C(t) \left(\ln \varepsilon - \frac{1}{3} \ln 2 \right) \\ & - \frac{1}{6} [H(t, t + 2\varepsilon) + H(t, t - 2\varepsilon)] + O(\varepsilon^4). \end{aligned} \quad (6.192)$$

With $\varepsilon = 5 \cdot 10^{-4}$, a quite sufficient accuracy is achieved, even with curved boundaries ($\kappa \neq 0$).

There remains now only the task of carrying out the double integrals over the singularity. In view of the dependence of the amplitude on the form and parametrization of the boundary, this is feasible only approximately by using a Taylor-series expansion of $C(t)$ about the point $t = j - 1/2$, at which the corresponding matrix element is to be evaluated; this means that we assume a *locally* valid expansion

$$C(t) = \sum_{k=0}^2 C_{j,k}(t - j + 1/2)^k, \quad (6.193a)$$

$$C_{j,k} = \frac{1}{k!} C^{(k)} \left(j - \frac{1}{2} \right), \quad (k = 0, 1, 2). \quad (6.193b)$$

These derivatives of the amplitude function are easily determined, as the boundary curve should be given in locally analytical form. On introducing this series expansion together with Eq. (6.190) into the integrals in Eq. (6.154a) we encounter a new family of integrals

$$L_k(p) = \iint_{-\infty}^{\infty} F(u)F(v) u^k (\ln |p + u - v| - \ln q) du dv, \quad (q = \max(1, |p|)). \quad (6.194)$$

In contrast to the integrals in Eq. (6.157), these integrals are now *asymmetric* with respect to u and v but nevertheless, by means of the symmetry of $F(u)$ and $F(v)$, the relation $L_k(-p) = L_k(p)$ can be established. For a numerical evaluation the order $M = 4$ of the modified interpolation kernels was chosen, because higher orders lead to unfavorable oscillation for large values of p . The corresponding results are given in Table 6.6. These integrals decrease rapidly with increasing distance p , so that the assumptions made in Eq. (6.193) are justified.

The off-diagonal elements can now be written as

$$H_{j,n} = H(j - 1/2, n - 1/2) + \sum_{k=0}^2 C_{jk} L_k(D_N(|j - n|)), \quad (6.195)$$

$D_N(j - n)$ being again the cyclic distance between the two positions j and n . The diagonal elements consist of the remainders R , the integrals over the singularity, and the term $\lambda + 1/2$:

$$H_{n,n} = R(n - 1/2, n - 1/2) + \sum_{k=0}^2 C_{nk} L_k(0) + \lambda + 1/2. \quad (6.196)$$

TABLE 6.6
VALUES OF THE INTEGRALS IN EQ. (6.194) WITH $M = 4$ (KASPER [43])

p	Integrals $L_0(p)$	Integrals $L_1(p)$	Integrals $L_2(p)$
0	-1.5734 07031 3577	+0.0000 00000 0000	-0.2019 72652 5456
1	-0.1896 90014 9739	+0.1529 25257 8614	+0.0816 68229 9485
2	+0.0752 67518 7179	-0.1513 77322 3551	+0.0730 96991 1986
3	-0.0207 80709 3606	+0.0660 55587 7126	-0.0706 64268 3832
4	+0.0032 70353 7714	-0.0139 46931 9470	+0.0186 35313 3107
5	-0.0003 37965 9193	+0.0017 72074 1012	-0.0022 27712 6835
6	+0.0000 34759 5956	-0.0002 21148 8127	+0.0004 15621 3980
7	+4.1162 44366 e-07	-3.7702 85148 e-06	+0.0000 28840 1064
8	+3.7196 32545 e-07	-3.3890 09664 e-06	+0.0000 16092 5619
9	+1.4227 63268 e-07	-1.4556 16951 e-06	+7.6203 31340 e-06
10	+5.9042 10151 e-08	-6.7164 25512 e-07	+3.8994 82524 e-06
11	+2.6737 93960 e-08	-3.3482 86572 e-07	+2.1388 93246 e-06
12	+1.3027 23809 e-08	-1.7807 47054 e-07	+1.2416 30625 e-06

It is now again easy to remove the square root transforms. In the case of multipoles (Eqs. (6.186) and (6.187)) we obtain a linear system of equations

$$\sum_{n=1}^N K_{j,n}^m q_n^m = -V_j^m, \quad (j = 1, \dots, N), \quad (6.197)$$

in which the sampling variables are again the corresponding function values at the half-integer positions $t = j - 1/2$. The matrix elements for $j \neq n$ become simply

$$K_{j,n}^m = r_j N_j \cdot \nabla G_m(\mathbf{u}_j, \mathbf{u}_n) + \sqrt{r_j/r_n} \sum_{k=0}^2 C_{j,k} L_k(D_{jn}). \quad (6.198a)$$

In the diagonal elements, the shifts related to ε must be expanded up to second order when the boundaries are curved; these positions are then

$$\mathbf{u}_{n,\mu} = \mathbf{u}_n + \mu \varepsilon \dot{\mathbf{u}}_n + \mu^2 \varepsilon^2 \ddot{\mathbf{u}}_n / 2, \quad -2 \leq \mu \leq 2. \quad (6.198b)$$

The position $\mu = 0$ must be skipped, and this is simply achieved by confinement and weight zero; hence from Eq. (6.192) it is evident that the weights

$$w_\mu = (-1/6, 2/3, 0, 2/3, -1/6) \quad (6.198c)$$

are to be used. Moreover, it is convenient to combine all contributions referring to the same amplitude $C(n - 1/2) = C_{n0}$ into one:

$$L'_0 = L_0(0) + \frac{1}{3} \ln 2 - \ln \varepsilon \quad (6.198d)$$

The diagonal elements are then concisely given by

$$K_{n,n}^m = r_n N_n \cdot \left\{ \sum_{\mu=-2}^2{}' w_\mu \nabla G_m(\mathbf{u}_n, \mathbf{u}_{n,\mu}) \right\} \\ + \lambda + \frac{1}{2} + C_{n0} L_0' + C_{n2} L_2(0). \quad (6.198e)$$

The linear system of equations for the surface currents J_n in a round lens is obtained in exactly the same manner. Here it is favorable to split off the term with λ ; hence

$$\sum_{n=1}^N (M_{j,n} + \lambda \delta_{j,n}) J_n = -T_j, \quad (j = 1, \dots, N) \quad (6.199)$$

with the matrix elements

$$M_{j,n} = -r_j^{-1} N_j \cdot \nabla \Phi(\mathbf{u}_j, \mathbf{u}_n) + \sqrt{r_u/r_j} \sum_{k=0}^2 C_{j,k} L_k(D_{j,n}) \quad (6.200a)$$

and

$$M_{n,n} = -r_n^{-1} N_n \cdot \left\{ \sum_{\mu=-2}^2{}' w_\mu \nabla \Phi(\mathbf{u}_n, \mathbf{u}_{n,\mu}) \right\} \\ + 1/2 + C_{n,0} L_0' + C_{n,2} L_2(0). \quad (6.200b)$$

The driving terms T_j are then given by

$$T_j = \dot{\mathbf{u}}_j \cdot \mathbf{H}_0(\mathbf{u}_j), \quad (j = 1, \dots, N). \quad (6.201)$$

6.4.4 The Conservation of Total Lens Current

The discretization for round lenses outlined here provides a simple way of controlling the accuracy of the numerical calculation and of using deviations from Ampère's law for purposes of correction.

First, within the validity of the midpoint formula of integration over periodic functions we should have

$$\eta := \sum_{j=1}^N T_j = \sum_{j=1}^N \dot{\mathbf{u}}(j-1/2) \cdot \mathbf{H}_0(\mathbf{u}(j-1/2)) \\ = \int_0^N \dot{\mathbf{u}}(t) \mathbf{H}_0(\mathbf{u}(t)) dt = \oint \mathbf{H}_0(z(s), r(s)) \cdot ds = 0, \quad (6.202a)$$

as there are no *physical* currents flowing in the yoke. Any value $\eta \neq 0$ is an integration error and thus a first measure of accuracy. It is then easy to eliminate this according to

$$T_j \longleftarrow T_j - \eta |T_j| \bigg/ \sum_{n=1}^N |T_n|. \quad (6.202b)$$

However, other correction weights also could be used if the main sources of errors are known.

A second control is given by the discretization of the integral theorem (6.107). If this is carried out in the same manner as in Section 6.4.3, it is finally found that

$$\sum_{j=1}^N M_{j,n} = 1, \quad (n = 1, 2, \dots, N) \quad (6.203)$$

should be valid for all n . These conditions supply a very strong control of accuracy and a powerful means of correction. The main source of errors is the truncation of the series expansion (6.193a) after the second order. As the ignored third order vanishes in the diagonal for reasons of symmetry, the next neighbour elements should be modified. This is easily possible in the following manner (D_{jn} denoting the cyclic distance):

$$\begin{aligned} &\text{for } (n = 1 \text{ to } N, \text{ step } 1) \\ &\{ \delta_n = \sum_{j=1}^N M_{j,n} - 1 \\ &\quad \text{for } (j = 1 \text{ to } N, \text{ step } 1) \\ &\quad \{ \text{if } (D_{jn} = 1) \quad M_{j,n} \longleftarrow M_{j,n} - \delta_n/2 \}. \\ &\} \end{aligned} \quad (6.204)$$

However, other correction methods also could be used if the sources of errors are known better.

From Eqs. (6.202) and (6.203), the conservation of currents can immediately be concluded: summation of Eqs. (6.199) over the label j results in

$$\sum_{n=1}^N (1 + \lambda) J_n = 0 \quad (6.205)$$

for all values of λ , as it should. The discretization is hence self-consistent.

The feasibility of this method is demonstrated in Figs. 5.26 and 5.27 in context with the limitations of the FEM. In this example, the BEM is certainly the better choice. Its inherent limitation lies in the fact that saturation effects must be excluded, as is implicit here in the presumption, $\lambda = (\mu_r - 1)^{-1} = \text{const.}$ A way to control this becomes obvious at the end of the next section.

Another limitation, so far tacitly assumed, is the condition that all sharp corners are slightly rounded off. This is a stronger necessity here than in the case of Dirichlet problems (Section 6.3.5), because otherwise the amplitude $C(t)$ would become discontinuous. An alternative here is Ströer's general Galerkin method with singular trial functions [26], but singularities would cause local saturation effects; hence rounding-off is the simpler way of approximation.

6.4.5 The Complete Field Calculation

The first result of the numerical calculations is a knowledge of a set of *surface sources*, either the Fourier components of electric or magnetic 'ring charges' or a set of coaxial ring currents. The total field is then obtained by superposition of an external field, the coulomb integral over all these sources, and a final Fourier synthesis of all these partial fields. In the far zone, this procedure is simple and—apart from the Fourier synthesis—is given by Eqs. (6.175) and (6.176). In the case of ring currents, the corresponding procedure is given by the *flux* potential

$$\Psi(\mathbf{u}) = \Psi_0(\mathbf{u}) + 2\pi\mu_0 \sum_{n=1}^N J_n \Phi(\mathbf{u}, \mathbf{u}_n), \quad (6.206)$$

Ψ_0 being the flux produced by the coil and Φ defined by Eq. (6.104). However, considering this definition and the general scheme in Section 6.3.5, we notice soon that we have here a special case with $m = 1$; hence we can immediately write

$$\Psi(\mathbf{u}_j) = \Psi_0(\mathbf{u}_j) + 2\pi\mu_0 \sum_{n=1}^N r_j r_n J_n \bar{G}_{j,n}^1 \quad (6.207)$$

for the function values at the sampling points. Hence, by using suitable interpolation techniques, the function $\bar{\Psi}(s)$ on the whole boundary is now known. By suitable differentiation techniques, the *normal* component B_n can be determined

$$B_n = \frac{1}{2\pi r} \frac{d\bar{\Psi}(s)}{ds} = \frac{1}{2\pi r v} \frac{d\bar{\Psi}}{dt}, \quad (6.208)$$

which remains continuous as it must. However, the *tangential* field component can also be determined easily. First of all, the calculated set of currents J_n provides the possibility of determining a function $J(t)$ by means of interpolation. Thereafter, the surface current density is given by Eq. (6.179), or by

$$\omega(s) = J(t)/v(t). \quad (6.209)$$

Now the boundary conditions with $\mu_1 = \mu$, $\mu_2 = \mu_0$ take the form

$$\begin{aligned} H_t &= B_{t,2}/\mu_0 = B_{t,1}/\mu, \\ B_{t,2} - B_{t,1} &= \mu_0 \omega. \end{aligned} \quad (6.210)$$

The solution of this linear system of equations results in Eq. (6.178) and

$$H_t(s) = -\lambda\omega(s), \quad B_{t,1} = -\mu\lambda\omega, \quad B_{t,2} = -\mu_0\lambda\omega. \quad (6.211)$$

Owing to Eqs. (6.209) and (6.205), the closed line integral over this function $H_t(s)$ vanishes as it must. Finally, the normal derivatives of the flux function $\Psi(u)$ are immediately given by these results, as we have

$$\mathbf{n} \cdot \nabla \Psi_{1,2} = -2\pi r B_{t,1,2} = 2\pi r \omega \lambda \mu_{1,2}, \quad (6.212)$$

whereupon the surface field is completely known.

It is now quite easy to control the *linearity* of the field. Because the maximum of the norm $|\mathbf{B}(\mathbf{u})|$ must be assumed at the boundary, it is sufficient to determine the value of this maximum. If this is less than a reasonably chosen threshold for linearity, the calculation can be accepted as accurate enough but if it exceeds this, then a suitable FEM program must be applied.

In analogy to this procedure, the surface Fourier potentials corresponding to the solutions of Eq. (6.185) or (6.187) can be similarly obtained by evaluation of Eqs. (6.161) and (6.162) with suitable adaptation of the notation. Thereafter, the tangential component of $\mathbf{H}(\mathbf{r})$ is obtained by differentiation with respect to the arc length whereas normal components follow from

$$\mathbf{n} \cdot (\mu_0 \mathbf{H}_2 - \mu \mathbf{H}_1) = 0, \quad \mathbf{n} \cdot (\mathbf{H}_2 - \mathbf{H}_1) = \tau. \quad (6.213)$$

These also hold separately for each Fourier component.

Finally, the fast algorithm can be used to determine the normal derivatives of a potential, if the sources result from the solution of a Dirichlet problem (Section 6.3.6). We then start from

$$\nabla(V_m - V_m^E) = \int_0^N \nabla G_m(\mathbf{u}, \mathbf{u}'(t)) q_m(t) dt \quad (6.214)$$

and form the scalar product with $rN(\mathbf{u})$ on both sides. This equation is then in turn written down for the sampling positions $\mathbf{u} = \mathbf{u}_j$, ($j = 1, \dots, N$) and *discretized* according to the scheme of Section 6.4.3. There are now two differences: the term with $\lambda + 1/2$ in the diagonal elements is to be omitted, and the resulting system is now only a matrix vector multiplication. After the appropriate normalization with $r_j^{-1}v_j^{-1}$, we obtain a set of relations

$$\eta_j := (r_j v_j)^{-1} \sum_{n=1}^N K_{jn}^m q_n^m + \mathbf{n}_j \cdot \nabla V_m^E(\mathbf{r}_j), \quad (6.215)$$

with $\mathbf{n}_j = N_j/v_j$. This result η_j is the *principal value* of the normal derivative, which means the arithmetic mean of the values on both sides. However, the *difference* of the two values is also known from the solution of the Dirichlet problem and is just σ_j according to Eq. (6.160b); hence, we have

$$\mathbf{n}_j \cdot \nabla V_m \pm = \eta_j \pm q_j (2r_j v_j)^{-1}, \quad (6.216)$$

whereupon the problem is solved. It is advantageous to store all the calculated values for purposes of interpolation.

With the resulting set of data, an accurate field calculation is now possible in two domains: (i) at a sufficient distance from the boundary, and (ii) on the boundary itself by suitable interpolation with respect to the curve parameter s or t . But there remains a domain near the boundary in which the simple summation formulas (6.175) and (6.176) become too inaccurate or even wrong. There are different ways of solving this problem.

- (a) *Adaptative Integration*. This certainly gives good results, if carried out carefully, but it may become very slow, especially in the close vicinity of the boundary.
- (b) *Condensation of Rings*. This idea is shown in Fig. 6.18. The midpoint form of the Maclaurin formula is compatible with a subdivision of each interval into *three* equal intervals, as then the original midpoints can be further used. To avoid too dramatic an increase of the computation time, this is done only in the vicinity of the foot point on the boundary, as shown in the figure. The field data referring to these positions are determined by interpolation, as would be necessary for any kind of adaptive integration. The errors resulting from the cut on the left-hand side can be reduced strongly with suitable summation weights:

$$\dots \left| \begin{array}{ccc} 0 & 1 & 0 \end{array} \right| \left| \begin{array}{cc} 0 & \frac{26}{27} \quad \frac{2}{9} \end{array} \right| \left| \begin{array}{ccc} \frac{1}{9} & \frac{10}{27} & \frac{1}{3} \end{array} \right| \left| \begin{array}{ccc} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array} \right| \dots$$

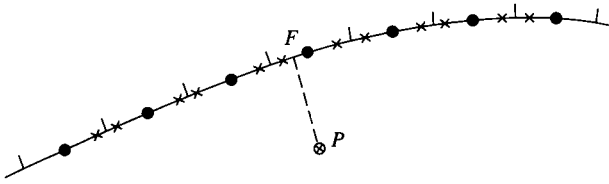


FIGURE 6.18 Dissection of a part of the boundary intervals in the vicinity of a reference point P . The interval endpoints are marked by bars, the original midpoints by dots, and the inserted points by crosses.

and similarly in reversed sequence on the right-hand side. If necessary, this procedure can be repeated a second time, and thus roughly one quarter of the original spacing between the rings can be reached as the limiting distance. It, however, makes little sense to go beyond local divisions into nine subintervals.

(c) Interpolation on the Surface Normal

This procedure is shown in Fig. 6.19 and becomes indispensable in the close vicinity of the boundary. First of all the *footpoint* F on the boundary is determined and then the potential and its gradient at it. Thereafter, the field is calculated at a point P^* on the surface *normal* sufficiently far away for the results to be reliable. The gradient is decomposed into its normal and tangential components. Now the potential and the normal component, both in F and P^* , make it possible to carry out a cubic Hermite interpolation with these data and we obtain then the corresponding values at the reference point P . The tangential components at F and P^* allow only a linear interpolation, which might be insufficient. It is therefore

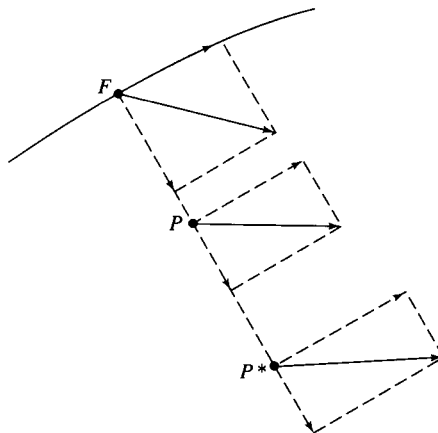


FIGURE 6.19 The potential and its gradient at a reference point P can be obtained by interpolation between the footpoint F on the boundary and a sufficiently distant point P^*

favorable to differentiate the normal derivative at F also with respect to s or t . This mixed derivative of second order can then be used to carry out a quadratic interpolation, so that both components become of about the same accuracy. At the end, the gradient is then transformed into its representation in cylindric coordinates.

This procedure does not require a subdivision of intervals, as outlined here, but it can be combined with it if the original spacings are too large for a quadratic interpolation. With some suitable storage techniques, this kind of field calculation can be made quite efficient. The field is then available in the whole z - r -plane without exceptions.

6.5 THE CALCULATION OF EXTERNAL FIELDS

In the earlier section we had tacitly assumed that the external field, essentially the function $H_0(\mathbf{r})$ originating from current-conducting coils in the absence of ferromagnetic yokes, is known. The determination of such fields is the present topic. At first sight, this seems to be a straightforward task. Provided that the current density $\mathbf{j}(\mathbf{r})$ is known in the domain C of the coils, the integration according to Biot-Savart's law

$$\mathbf{H}_0(\mathbf{r}') = \frac{1}{4\pi} \int_C \frac{\mathbf{j}(\mathbf{r}) \times (\mathbf{r}' - \mathbf{r})}{|\mathbf{r}' - \mathbf{r}|^3} d^3r \quad (6.217)$$

is the required solution.

A more careful consideration soon shows that this straightforward integration is impractical, because a three-dimensional numerical integration is so very slow that it would represent the vast majority of the computation time. We must hence look for more efficient methods of integration.

6.5.1 The Evaluation of Particular Integrals

This concept was worked out by Ströer [16,17], who applied it in the field in extended rotationally symmetric coils, as used in round magnetic lenses. Here we shall demonstrate this method in a more general form, as it might also be very useful in other applications.

We suppose that we have to evaluate a Coulomb integral of the general form

$$V(\mathbf{r}') = \int_{\Gamma} G(\mathbf{r}', \mathbf{r}) \rho(\mathbf{r}) d^3r \quad (6.218)$$

with a *known* source function $\rho(\mathbf{r})$ that vanishes outside the domain Γ of integration. The Green function G is defined by Eq. (1.88) and satisfies Eqs. (1.89)

and (1.90). Hence, the function $V(\mathbf{r})$ satisfies the Poisson equation

$$\Delta V(\mathbf{r}) = \begin{cases} -\rho(\mathbf{r}), & \mathbf{r} \in \Gamma \\ 0, & \mathbf{r} \notin \Gamma. \end{cases} \quad (6.219)$$

The general solution of this PDE, *not* necessarily satisfying any given boundary conditions, consists in a *particular* solution $V_p(\mathbf{r})$ of Eq. (6.219) and the general solution $V_H(\mathbf{r})$ of the associated Laplace equation. Then the evaluation of Eq. (1.93) for the homogeneous function V_H (without ρ -term) gives

$$\begin{aligned} \beta(\mathbf{r}')V_H(\mathbf{r}') &= \int_s G(\mathbf{r}', \mathbf{r}) \partial_n V_H(\mathbf{r}) da \\ &\quad - \int_s p(\mathbf{r}', \mathbf{r}) V_H(\mathbf{r}) da. \end{aligned} \quad (6.220a)$$

Similarly, this integral theorem (1.93) can be applied to the space *outside* the domain Γ . Then we have to change the signs on the right-hand side if the surface normal continues to be directed out of Γ . Moreover, the factor $\beta(\mathbf{r}')$ is to be replaced by its complement $1 - \beta(\mathbf{r}')$. The *total* solution $V(\mathbf{r})$ will satisfy the natural boundary conditions at infinity, as is implied by Eq. (6.218).

We then obtain the integral formula

$$\begin{aligned} (1 - \beta(\mathbf{r}'))V(\mathbf{r}') &= - \int_s G(\mathbf{r}', \mathbf{r}) \partial_n V(\mathbf{r}) da \\ &\quad + \int_s p(\mathbf{r}', \mathbf{r}) V(\mathbf{r}) da. \end{aligned} \quad (6.220b)$$

This equation is now added to Eq. (6.220a) and the homogeneous function V_H eliminated by writing $V_p = V - V_H$, and the result is then

$$\begin{aligned} V(\mathbf{r}') &= \beta(\mathbf{r}')V_p(\mathbf{r}') - \int_s G(\mathbf{r}', \mathbf{r}) \partial_n V_p(\mathbf{r}) da \\ &\quad + \int_s p(\mathbf{r}', \mathbf{r}) V_p(\mathbf{r}) da. \end{aligned} \quad (6.221)$$

The gain achieved with this formula lies in the fact that only a surface integration is necessary instead of a spatial one. It is often not too complicated to find a particular integral V_p if the natural boundary conditions are not imposed on it.

The formalism holds also for *vector potentials* if the vector Poisson equation (1.25) with $\mu = \mu_0$ and given $\mathbf{j}(\mathbf{r})$ is to be satisfied inside the domain

C of a coil. We then have to try to find a suitable particular solution A_p of it. Equation (6.221) can be written down for the three cartesian components, and these are thereafter organized in vectorial form, giving

$$\begin{aligned} A(\mathbf{r}') &= \beta(\mathbf{r}')A_p(\mathbf{r}') - \int_{\partial C} G(\mathbf{r}', \mathbf{r}) \mathbf{n}(\mathbf{r}) \cdot \nabla A_p(\mathbf{r}) d\mathbf{a} \\ &+ \int_{\partial C} p(\mathbf{r}', \mathbf{r}) A_p(\mathbf{r}) d\mathbf{a}. \end{aligned} \quad (6.222)$$

These integrals can now be transformed to cylindrical or other coordinates that are useful for their evaluation.

The integral formulas (6.221) and (6.222) have some favorable properties. Because Eq. (6.222) is essentially only a vector form of Eq. (6.221), we confine the presentation to the scalar case.

- (i) The integral formula holds in the *whole space*, although the particular solution needs to be known only up to the surface. Inside the domain Γ Green's theorem can also be applied to the function $V_p(\mathbf{r}')$ and it then turns out that Eq. (6.218), from which we set out, holds. For points outside the factor $\beta(\mathbf{r}')$ vanishes and a possible analytical continuation of V_p is not necessary. As G and p satisfy Laplace's equation and the natural boundary conditions, the same holds for $V(\mathbf{r}')$.
- (ii) Because the solution of (6.218) is unique and continuously differentiable, it is permissible to differentiate Eq. (6.221) with respect to $\mathbf{r}'$ if $\mathbf{r}' \notin S$. For points $\mathbf{r}' \in S$, a suitable limitation must be performed. Although it is allowed, such a differentiation is *not* very favorable, as it requires mixed second-order derivatives of Green's function.
- (iii) The formalism can be generalized for *several* domains $\Gamma_1, \dots, \Gamma_N$ with disjoint source functions. It is then only necessary to achieve continuity of the potential V_p itself on the inner boundaries. The integrals with factor $p(\mathbf{r}', \mathbf{r})$ vanish on these and survive only on the outer enclosing boundary.

6.5.2 Application to Rotationally Symmetric Fields

We now assume that the domain Γ of sources and all functions associated with it are rotationally symmetric about the z -axis, which is usually also the "optic" axis. This domain may have a meridional section D in the upper half (z, r) plane with a positively oriented and closed boundary line B . The potentials V_p and V are then independent of the azimuth. The integration over the latter results in the Fourier kernel G_0 and its derivative. Hence only a *line* integral

over B remains. With the use of two-dimensional vectors $\mathbf{u} = (z, r)$, we obtain

$$\begin{aligned} V(\mathbf{u}') &= \beta(\mathbf{u}')V_p(\mathbf{u}') - \oint G_0(\mathbf{u}', \mathbf{u}) \mathbf{r}n \cdot \nabla V_p(\mathbf{u}) ds \\ &+ \oint V_p(\mathbf{u}) \mathbf{r}n \cdot \nabla G_0(\mathbf{u}', \mathbf{u}) ds. \end{aligned} \quad (6.223)$$

For domains D_A intersecting the optic axis, the path of integration is closed on the latter, as is shown in Fig. 6.20; however, this part gives no contribution to the integrals because of the factor r in Eq. (6.223). Similarly, a positive mirror symmetry can be used. Then the kernel G_0 must be replaced by its symmetrized form:

$$G_s(\mathbf{u}', \mathbf{u}) = G_0(z', r'; z, r) + G_0(z', r'; -z, r) \quad (6.224)$$

and the path must be closed along the axis of symmetry if a domain D_s (see Fig. 6.20) reaches the latter. Again, this part does not give a contribution.

As a simple example, we consider the source function

$$\rho(z, r) = \rho_0 + Az + Br^2 + Czr^2, \quad (6.225a)$$

in which the coefficients are to be determined from interpolation conditions at the four corners of a quadrangle. A particular potential V_p fitting this source function is readily obtained:

$$V_p(z, r) = -\frac{r^2}{4}(\rho_0 + Az) - \frac{r^4}{16}(B + Cz), \quad (6.225b)$$

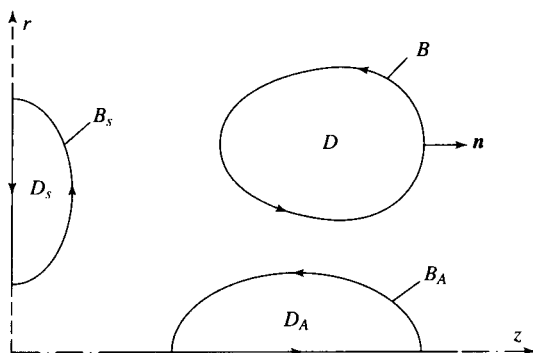


FIGURE 6.20 Different situations of domains and their closed boundaries in the upper half meridional plane

as is verified by differentiation. Of course, this result is not unique, as any linear combination of the polynomials of Eq. (2.66) with $\alpha = 1$ can be superimposed on this function. This has no influence on the final result. For the outer domain, this is immediately evident from Eq. (6.220a) with $\beta(\mathbf{r}') = 0$.

With respect to the field calculation in a rotationally symmetric coil, we set out from the vector potential

$$\mathbf{A}(\mathbf{r}) = \Psi(z, r) \mathbf{e}_\varphi(\varphi) \cdot (2\pi r)^{-1}, \quad (6.226)$$

$\Psi(z, r)$ being the magnetic flux potential. Introduction into Eq. (6.222) and integration over the azimuth φ leads then to the flux kernel Φ from Eq. (6.104), instead of G_0 and to the integral formula

$$\begin{aligned} \Psi(\mathbf{u}') &= \beta(\mathbf{u}') \Psi_p(\mathbf{u}') - \oint r^{-1} \Psi_p(\mathbf{u}) \mathbf{n} \cdot \nabla \Phi(\mathbf{u}', \mathbf{u}) ds \\ &+ \oint r^{-1} \Phi(\mathbf{u}', \mathbf{u}) \mathbf{n} \cdot \nabla \Psi_p(\mathbf{u}) ds. \end{aligned} \quad (6.227)$$

The current density $\mathbf{j}(\mathbf{r})$ with amplitude $j(z, r)$ is frequently approximated by a bilinear function

$$j(z, r) = j_0 + az + br + czt. \quad (6.228a)$$

The corresponding particular flux $\Psi(z, r)$ must then satisfy the PDE (2.48) with $\mu = \mu_0$. A very simple polynomial as solution is here

$$\Psi_p(z, r) = -2\pi\mu_0 r^3(j_0 + az)/3 - \pi\mu_0 r^4(b + cz)/4, \quad (6.228b)$$

as is easily verified. Strictly speaking, a term in r^3 contradicts the conditions of regularity at the optic axis; however, no coil ever has a vanishing lower radius, so that this problem never arises.

The evaluation of Eq. (6.223) or Eq. (6.227) is *not* the solution of a Fredholm equation, but requires only a *line*-integration. It is advantageous to perform this for positions $\mathbf{u}'$ in the boundary B . With $\beta(\mathbf{u}') = 1/2$ the boundary values of the potential are thus obtained, whereby the task of field calculation is now reduced to the solution of a *Dirichlet* problem. This supplies *surface sources* instead of spatial ones, so that the external fields can be calculated in the same manner as those generated by the surface sources on electrodes or yokes. The necessary computation time is merely doubled instead of being increased by a factor of 100 or more.

An alternative method, in which it is not necessary to solve a linear system of equations but which involves an additional gradient, is given in reference [32].

6.5.3 Coils with Rectangular Cross Sections

Quite frequently, the coils in round magnetic lenses have a rectangular cross section in the (z, r) plane with *constant* current density; hence we assume

$$j(z, r) = j_0 \text{ if } (z_1 \leq z \leq z_2 \text{ and } r_1 \leq r \leq r_2) \\ 0 \text{ else,} \quad (6.229)$$

with $z_1 < z_2$ and $0 < r_1 < r_2$. In such a simple case, of the methods other than those outlined earlier, one may be favorable. This is important owing to the fact that the spacing between the coil and the yoke is often so narrow that the integration over surface singularities becomes difficult. Moreover, an analogous method of integration may be useful in other cases of integration over singular functions.

In this section we shall denote the reference point by (z, r) and the source position, being the integration variable, by (z', r') . To simplify the calculations we now introduce relative coordinates

$$u := (z' - z)/2r, \quad v := (r' - r)/2r, \quad (r > 0). \quad (6.230)$$

The method outlined here will be based on series expansions of the integrands in the vicinity of their singularity in terms of these coordinates. We set out from a configuration as shown in Fig. 6.9 and adopt the notation of Eqs. (6.73) to (6.76). Let the ring carry a current J ; then the conventional representation of the magnetic field is given by

$$A(z, r) = \frac{\mu_0 J r'}{\pi d_2} (2D(k) - K(k)), \quad (6.231a)$$

$$B_r(z, r) = (z - z') C(z, r), \quad (6.231b)$$

$$B_z(z, r) = (r' - r) C(z, r) + \frac{2\mu_0 J r'^2}{\pi d_2^3} D(k), \quad (6.231c)$$

$$C(z, r) = \frac{\mu_0 J r'}{\pi d_2 d_1^2} (E(k) - 2k'^2 D(k)). \quad (6.231d)$$

The factor $C(z, r)$ is the amplitude of the circular part of the field; this means that, in the absence of the second term in B_z , the field lines would just become circles round the singularity. This amplitude vanishes on the optic axis. This representation is not always the best one, but in the present context, it is particularly advantageous, as the additional square roots in Eqs. (6.77) and (6.95) would require additional series expansions.

To cast the later expansions in a concise form, it is favorable to introduce a notation for frequently appearing functions

$$L(u, v) := -\ln((u^2 + v^2)/16), \quad (6.232a)$$

$$\overline{C}(u, v) := (1 + v - v^2/2 + v^3/2)/(u^2 + v^2), \quad (6.232b)$$

and with these:

$$\overline{A}(u, v) := \frac{4\pi}{\mu_0 J} A(z, r) = (1 + v + v^2/4 + 3u^2/4) L(u, v) + \delta\overline{A}(u, v), \quad (6.233a)$$

$$\overline{B}_r(u, v) := \frac{4\pi r}{\mu_0 J} B_r(z, r) = -u \overline{C}(u, v) + \frac{3u}{4} (1 - v) L(u, v) + \delta\overline{B}_r(u, v), \quad (6.233b)$$

$$\begin{aligned} \overline{B}_z(u, v) &:= \frac{4\pi r}{\mu_0 J} B_z(u, v) = v \overline{C}(u, v) + \frac{1}{2} \left(1 - \frac{v}{2} + \frac{v^2}{4} - \frac{3u^2}{4} \right) L(u, v) \\ &+ \delta\overline{B}_z(u, v). \end{aligned} \quad (6.233c)$$

The derivation of these series expansions is straightforward. The remainders $\delta\overline{A}$, $\delta\overline{B}_r$, and $\delta\overline{B}_z$ are not necessarily small but are functions that are so smooth that they can be integrated by Gauss quadratures. They are favorably determined by subtraction of the singular functions from the exact ones, so that the truncation of the earlier given expansions is always self-consistent.

We now come to the problem of performing double intergrations over a rectangle. This task can be facilitated considerably by the following theorem: let

$$I = \int_{u_1}^{u_2} \int_{v_1}^{v_2} f(u, v) dv du \quad (6.234a)$$

be the integral required. We first look for an indefinite integral

$$F(u, v) = \iint f(u, v) dv du + g_1(u) + g_2(v), \quad (6.234b)$$

which hence satisfies

$$\frac{\partial^2 F(u, v)}{\partial u \partial v} = f(u, v). \quad (6.234c)$$

If this is found, then the definite integral is

$$I = F(u_1, v_1) + F(u_2, v_2) - F(u_1, v_2) - F(u_2, v_1), \quad (6.234d)$$

from which the additive free functions g_1, g_2 cancel out.

In this sense, the integrations over the functions in Eqs. (6.233) can be carried out analytically. In this context it is favorable to introduce an ancillary function

$$\hat{a}(p, q) := p - q \arctan(p/q), \quad (6.235)$$

with $|\arctan(p/q)| \leq \pi/2$.

This function has hence the symmetry properties

$$\begin{aligned} \hat{a}(-p, q) &= -\hat{a}(p, q), \\ \hat{a}(p, -q) &= +\hat{a}(p, q). \end{aligned} \quad (6.236)$$

The indefinite integrals over the functions given by Eqs. (6.233) are then found to be

$$\begin{aligned} \iint \bar{A} du dv &= u(L(u, v) + 1) \left(v + \frac{v^2}{2} + \frac{v^3}{12} + \frac{u^2}{6} + \frac{u^2 v}{4} \right) \\ &\quad + v \left(1 + \frac{2v}{3} \right) \hat{a}(u, v) + u \left(1 + \frac{u^2}{3} \right) \hat{a}(v, u) \\ &\quad - \frac{uv}{36} (3u^2 + v^2), \end{aligned} \quad (6.237a)$$

$$\begin{aligned} \iint \bar{B}_r du dv &= (L(u, v) + 1) \left[v + \frac{v^2}{2} + \frac{v^3}{12} + \frac{u^2}{2} + \frac{3u^2 v}{4} \right. \\ &\quad \left. - (u^2 + v^2) (5u^2 + v^2)/16 \right] \\ &\quad + (2 + 4u^2/3) \hat{a}(v, u) - 5u^2 v^2/16, \end{aligned} \quad (6.237b)$$

$$\begin{aligned} \iint \bar{B}_z du dv &= -\frac{u}{2} L(u, v) [1 - v + (u^2 + v^2 + u^2 v)/4 - v^3/12] \\ &\quad + u (1 - u^2/3) \hat{a}(v, u) - \hat{a}(u, v) \\ &\quad + uv \left(1 - \frac{3v}{8} + \frac{7v^2}{36} - \frac{u^2}{12} \right). \end{aligned} \quad (6.237c)$$

In these expressions, all terms that depend only on one coordinate have been omitted, as they finally cancel out. The corresponding definite integrals are given by four evaluations, each according to the rule, given in Eq. (6.234d).

The remainders $\delta\bar{A}$, $\delta\bar{B}_z$, and $\delta\bar{B}_r$ must not be disregarded, because at larger values of $u^2 + v^2$ they will give a contribution. These can, however, be perfectly integrated by a 7×7 Gauss quadrature, as all rapidly varying terms are removed from them. If the reference point happens to coincide with a

quadrature position, it is symmetrically shifted by $\pm 10^{-3}$ and the results averaged. At the end, the scale transform to the original (z, r) coordinate system is to be performed, whereupon the field calculation is ready. It gives the correct result, regardless of whether the reference position is located inside the coil or outside, or just on its boundary. In contrast to this, even a Gauss quadrature of orders 24×24 may fail in the interior or in the vicinity of the boundary.

Although the procedure is correct for any position with $r \neq 0$, it should *not* be used at large distances from the surface, say about the double the length of the coil or more, as the separation method implies then the subtraction of numbers of nearly equal value, which gives rise to rounding errors. The 7×7 Gauss quadrature then suffices and is even faster.

The Paraxial Domain

In the paraxial domain, for $r \leq 0.15r_1$, the method described previously becomes unfavorable owing to the normalizations made in Eq. (6.230). It is then possible to employ mere quadratures.

A more efficient method is based on the integration over the axial field strength $B(z)$. It is then possible to make use of a scalar potential $W(z)$, which is of physical interest in electron optics, as it is related to the image rotation in magnetic lenses (ref. [22], Chapter 15).

The axial field strength $b(z)$ corresponding to Eq. (6.231c) is given by

$$b(z) = B_z(z, 0) = \frac{1}{2}\mu_0 J r^2 R^{-3}, \quad (6.238a)$$

in which R denotes the frequently appearing distance

$$R(z - z', r') = [r'^2 + (z - z')^2]^{1/2}, \quad (6.238b)$$

(see Fig. 6.9). The scalar potential $w(z)$ becomes

$$w(z) = \int b(z) dz = \frac{1}{2}\mu_0 J (z - z')/R + C. \quad (6.239)$$

The choice $C = \mu_0 J/2$ implies $w(-\infty) = 0$ and is hence particularly favorable. Apart from this, a corresponding *indefinite* double integral is found to be

$$\tilde{W}(z) = \iint w dz' dr' = -\frac{\mu_0 j}{4} [r'R + (z - z')^2 \ln(r' + R)]. \quad (6.240)$$

This is to be evaluated four times according to the rule of Eq. (6.234d) to obtain $W(z)$. The differentiation with respect to z can be exchanged with the

integration over z' and r' ; we obtain thus

$$\tilde{B}(z) = \iint b dz' dr' = -\frac{\mu_0 J}{2} (z - z') \ln(r' + R), \quad (6.241)$$

$$\tilde{B}'(z) = \frac{-\mu_0 J}{2} [\ln(r' + R) + 1 - r'/R], \quad (6.242)$$

and so on. In this way, repeated differentiations are possible and with these the paraxial series expansions given in Section 2.4.

6.5.4 Magnetic Fields of Deflection Systems

We are now concerned with the magnetic field in systems such as those shown in Fig. 6.17, but only with that contribution that is generated by the coils in the absence of the yoke. The cylindrical shape is often too special; hence, we assume now the more general form shown in Fig. 6.21. The wires are not located directly on the surface of the yoke but are at a finite distance from it for reasons of electrical insulation. It is now necessary to assume that the layer of winding is very thin so that the approximation by *surface currents* can be made. For simplification, it is necessary that this surface be *rotationally* symmetric, as otherwise a Fourier analysis is impossible, and a rigorous evaluation of Biot-Savart's law would then be necessary. The following calculations refer to this surface that may be open or closed. We represent it in parametric form $z(s)$, $r(s)$ in cylindric coordinates, the second parameter then being the azimuth φ .

The concept of surface current density $\mathbf{J}(\mathbf{r})$ was already introduced in Section 1.5.2, and we employ this concept here. According to the assumptions introduced earlier, this current density is here a vector field

$$\mathbf{J}(s, \varphi) = J_t(s, \varphi) \boldsymbol{\tau}(s, \varphi) + J_\varphi(s, \varphi) \mathbf{e}_\varphi(\varphi), \quad (6.243)$$

$\boldsymbol{\tau}(s, \varphi)$ being the normalized tangential vector in the meridional plane $\varphi = \text{const.}$; this vector has the cartesian components

$$\boldsymbol{\tau} = (r'(s) \cos \varphi, r'(s) \sin \varphi, z'(s)). \quad (6.244)$$

The two components J_t and J_φ cannot be chosen independently but are related to each other by the law of conservation for currents

$$\frac{\partial}{\partial s}(r(s)J_t(s, \varphi)) + \frac{\partial}{\partial \varphi}J_\varphi(s, \varphi) = 0. \quad (6.245)$$

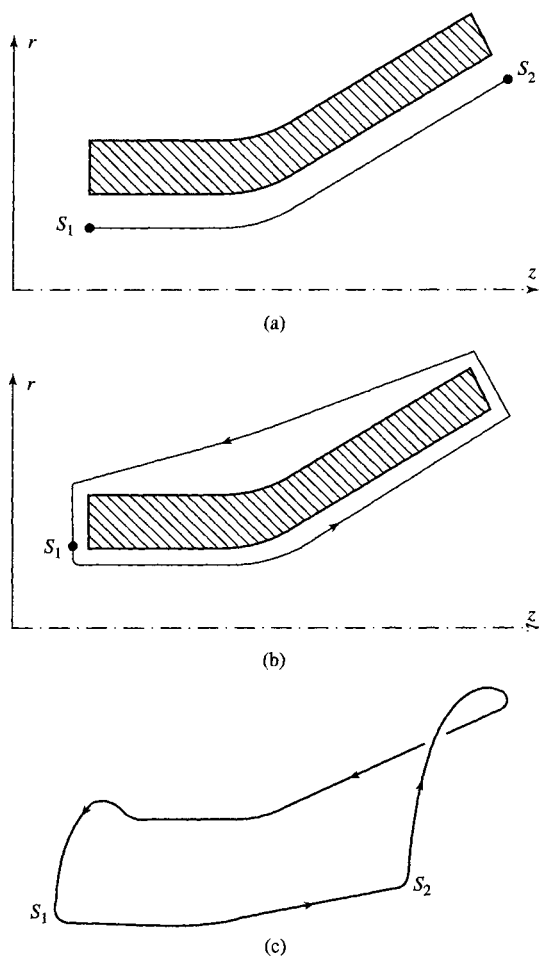


FIGURE 6.21 Structure of deflection systems: (a) yoke (hatched area) and saddle coils; (b) yoke and toroidal coils; (c) perspective view of one winding of the saddle coils. The figure shows only one half part of the system

In practice this is automatically satisfied by the fact that the current distribution is produced by layers of thin wires with constant current in each of them.

For an efficient field calculation and also for recalling the technical application, we introduce Fourier-series expansions:

$$r(s)J_t(s, \varphi) = \sum_m w_m(s) \cos m(\varphi - \alpha_0), \quad (6.246a)$$

$$r(s)J_\varphi(s, \varphi) = \sum_m a_m(s) \sin m(\varphi - \alpha_0). \quad (6.246b)$$

From the continuity condition (6.245), the relations

$$m a_m(s) = -r(s) w'_m(s) \quad (6.246c)$$

are immediately obtained, and this shows that the azimuthal components can be eliminated (the case $m = 0$ is unreasonable and excluded). In principle, these relations hold for all orders m , but only *odd* orders $m = 1, 3, 5, \dots$, are useful in deflection systems; this selection is implicit in the antisymmetric structure of the field. Without loss of generality we shall assume now that $\alpha_0 = 0$, and this means that the coordinate system is adapted to the symmetry planes of the coils.

The following calculations are a generalization of those given by Ströer [16,17], as here two nonvanishing components J_t and J_φ at the same position are possible, whereas Ströer and most other authors assumed either meridional or azimuthal directions of $\mathbf{J}$. We set out from

$$\mathbf{A}(\mathbf{r}_0) = \frac{\mu_0}{4\pi} \iint \mathbf{J}(\mathbf{r}) |\mathbf{r} - \mathbf{r}_0|^{-1} d\mathbf{a}, \quad (6.247)$$

where we have denoted the reference position by $\mathbf{r}_0$ instead of $\mathbf{r}'$ to avoid confusion with the derivatives with respect to the arc length s . More explicitly this coulomb integral becomes

$$\mathbf{A}(\mathbf{r}_0) = \frac{\mu_0}{4\pi} \int_s \int_{\varphi=0}^{2\pi} r(s) [\tau(s, \varphi) J_t(s, \varphi) + \mathbf{e}_\varphi(\varphi) J_\varphi(s, \varphi)] R^{-1} d\varphi ds \quad (6.248)$$

with the denominator

$$R = |\mathbf{r}(s, \varphi) - \mathbf{r}_0| = [(z - z_0)^2 + r^2 + r_0^2 - 2rr_0 \cos \psi]^{1/2}, \quad (6.249)$$

and $\psi := \varphi - \varphi_0$. This vector integral is most favorably evaluated in cylindrical coordinates taking Eqs. (6.246) into consideration. Since $\tau_z = z'(s)$, the *longitudinal* coordinate turns out to be

$$A_z = \frac{\mu_0}{4\pi} \sum_m \int_s \left[z'(s) w_m(s) \int_0^{2\pi} R^{-1} \cos m\varphi d\varphi \right] ds. \quad (6.250)$$

With $\varphi = \varphi_0 + \psi$, and by making use of the addition theorems, we recover the Fourier-Green function of Eq. (6.67), as the integral over $\sin \psi$ vanishes.

A first result is therefore

$$A_z(\mathbf{r}_0) = \sum_m \cos m\varphi_0 \int_s z'(s) w_m(s) G_m(z_0, r_0; z, r) ds. \quad (6.251)$$

The *radial* component is obtained in an analogous manner, the only difference being that now the scalar product

$$\mathbf{e}_r(\varphi) \cdot \mathbf{e}_r(\varphi_0) = \cos(\varphi - \varphi_0) = \cos \psi \quad (6.252)$$

appears as an additional factor. We hence have to evaluate the integral

$$A_{r1} = \frac{\mu_0}{4\pi} \sum_m \int_s \left(r'(s) w_m(s) \int_0^{2\pi} R^{-1} \cos \Psi \cos m(\varphi_0 + \varphi) d\psi \right) ds. \quad (6.253)$$

The Fourier analysis of this expression now results in *two* components:

$$\begin{aligned} & \cos \psi \cos m(\varphi_0 + \psi) \\ &= \frac{1}{2} \cos m\varphi_0 \{ \cos(m-1)\psi + \cos(m+1)\psi \} + \dots, \end{aligned} \quad (6.254)$$

where again the sine term vanish by integration; we hence obtain

$$A_{r1}(\mathbf{r}_0) = \frac{1}{2} \mu_0 \sum_m \cos m\varphi_0 \int_s r'(s) w_m(s) (G_{m-1} + G_{m+1}) ds. \quad (6.255a)$$

This is, however, *not* the entire result, if $J_\varphi \neq 0$, as this component also contributes to the complete component via a factor

$$\mathbf{e}_\varphi(\varphi) \cdot \mathbf{e}_r(\varphi_0) = \sin(\varphi_0 - \varphi) = -\sin \psi. \quad (6.255b)$$

We hence obtain

$$A_{r2}(\mathbf{r}_0) = -\frac{\mu_0}{4\pi} \sum_m \int_s \left(a_m(s) \int_0^{2\pi} R^{-1} \sin \psi \sin m(\varphi_0 - \psi) d\psi \right) ds, \quad (6.255c)$$

and by means of the addition theorems and of Eq. (6.246c), we finally have (with $m \neq 0$)

$$A_{r2}(\mathbf{r}_0) = \frac{1}{2} \mu_0 \sum_m \frac{1}{m} \cos m\varphi_0 \int_s r(s) w'_m(s) (G_{m-1} - G_{m+1}) ds, \quad (6.255d)$$

$$A_r(\mathbf{r}_0) = A_{r1}(\mathbf{r}_0) + A_{r2}(\mathbf{r}_0). \quad (6.255e)$$

In the same manner, we also find

$$A_{\varphi 1}(\mathbf{r}_0) = \frac{\mu_0}{2} \sum_m \sin m\varphi_0 \int_s r'(s) w_m(s) (G_{m-1} - G_{m+1}) ds, \quad (6.256a)$$

$$A_{\varphi 2}(\mathbf{r}_0) = -\frac{\mu_0}{2} \sum_m \frac{1}{m} \sin m\varphi_0 \int_s r(s) w'(s) (G_{m-1} + G_{m+1}) ds, \quad (6.256b)$$

$$A_{\varphi}(\mathbf{r}_0) = A_{\varphi 1}(\mathbf{r}_0) + A_{\varphi 2}(\mathbf{r}_0), \quad (6.256c)$$

whereupon the A -field is completely determined. There are some obvious rules:

- (i) Each component is obtained as a line integral over Fourier kernels, and the function $w_m(s)$ or its derivative, the longitudinal component A_z is formed with the odd order m , whereas the transversal components are built up from the even orders $m \pm 1$.
- (ii) The dependence on the azimuth φ_0 is found to be

$$\begin{aligned} A_z &= \sum_m A_z^m(\mathbf{u}_0) \cos m\varphi_0, \\ A_r &= \sum_m A_r^m(\mathbf{u}_0) \cos m\varphi_0, \\ A_{\varphi} &= \sum_m A_{\varphi}^m(\mathbf{u}_0) \sin m\varphi_0, \end{aligned} \quad (6.257)$$

where again the abbreviation $\mathbf{u}_0 = (z_0, r_0)$ has been introduced.

The field strength $\mathbf{B}(\mathbf{r}_0)$ is now obtained by evaluation of $\mathbf{B} = \text{curl} \mathbf{A}$ in cylindric coordinates

$$\begin{aligned} B_z &= \sum_m (\partial A_{\varphi}^m / \partial r_0 + A_{\varphi}^m / r_0 + m A_r^m / r_0) \sin m\varphi_0, \\ B_r &= \sum_m (-\partial A_{\varphi}^m / \partial z_0 - m A_z^m / r_0) \sin m\varphi_0, \\ B_{\varphi} &= \sum_m (\partial A_r^m / \partial z_0 - \partial A_z^m / \partial r_0) \cos m\varphi_0. \end{aligned} \quad (6.258)$$

These differentiations are to be carried out under the integrals in Eqs. (6.251) to (6.256) and refer then to the first pair of variables in $G_m(z_0, r_0; z, r)$ etc. In this context, the formulas (6.89) and (6.90) can be used. Moreover, if the arc length s is an unfavorable curve parameter, the transformation to another parameter t is quite easy by substituting

$$z' ds = \dot{z} dt, \quad r' ds = \dot{r} dt, \quad w'_m ds = \dot{w}_m dt. \quad (6.259)$$

Then, with suitable choice of this parameter t , the advantages of the Euler-Maclaurin formula can be used here.

With respect to the field calculation for a ferrite yoke, it is of importance that the components B_z and B_r have the *same* angular behaviour as $\sin m\varphi_0$. This has the consequence that H_N in Eq. (6.184) also has this behavior. Consequently, only *sine terms* of this form contribute to Eqs. (6.185) and (6.186). This implies that the simulation with a *scalar* surface source $\tau(\mathbf{r})$ is compatible with this kind of external field and has the advantage of leading to *uncoupled* integral equations, once the set of driving terms $F_m(s)$ has been determined.

6.5.5 Special Cases of Deflection Systems

The frequent assumption that the windings in deflection coils have either the meridional or the azimuthal direction emerges as a simplifying specialization of the general theory.

In the meridional parts, we have $a_m(s) = 0$ and hence $w_m(s) = \text{const.}$ because of Eq. (6.246c). We can then take this factor in front of the corresponding integrals, and there remain the purely *geometrical* dimensionless deflection coefficients

$$M_z^m = \int_{s_1}^{s_2} z'(s) G_m(\mathbf{u}_0, \mathbf{u}(s)) ds, \quad (6.260a)$$

$$M_{r,\varphi}^m = \frac{1}{2} \int_{s_1}^{s_2} r'(s) (G_{m-1} \pm G_{m+1}) ds. \quad (6.260b)$$

In toroidal systems, these are the only nonvanishing coefficients, as the loop from s_1 to s_2 is closed (see Fig. 6.21b).

In saddle coil systems, additional contributions result from the azimuthally directed windings at the two front planes. The discontinuity of $w'_m(s)$ at $s = s_1$ or s_2 cannot be ignored but must be considered as

$$w'_m(s) = w_m(\delta(s - s_1) - \delta(s - s_2)). \quad (6.260c)$$

On introducing this into Eqs. (6.255d) and (6.256b), the corresponding integrals can be evaluated completely and they give rise to the coefficients

$$\begin{aligned} T_{r,\varphi}^m &= \frac{1}{2m} (G_{m-1}(\mathbf{u}_0, \mathbf{u}_1) - G_{m-1}(\mathbf{u}_0, \mathbf{u}_2)) \\ &\quad \pm \frac{1}{2m} (G_{m+1}(\mathbf{u}_0, \mathbf{u}_1) - G_{m+1}(\mathbf{u}_0, \mathbf{u}_2)). \end{aligned} \quad (6.260d)$$

Hence, Eqs. (6.257) can now be rewritten more explicitly as

$$A_z = \mu_0 \sum_m M_z^m w_m \cos m\varphi_0,$$

$$\begin{aligned}
 A_r &= \mu_0 \sum_m (M_r^m + T_r^m) w_m \cos m\varphi_0, \\
 A_\varphi &= \mu_0 \sum_m (M_\varphi^m + T_\varphi^m) w_m \sin m\varphi_0,
 \end{aligned}
 \tag{6.261}$$

and corresponding equations can be written down for the $\mathbf{B}$ -field. Further evaluations now require a detailed specification of the geometric shape. Apart from the notation, the above formulas agree with those derived by Ströer [16,17].

The numerical evaluation of the integrals occurring in Eqs. (6.257), (6.258), and (6.261) is straightforward, if the coils are sufficiently distant from the yoke, but often this is not the case. A useful technique is then the separation of the almost singular terms from the regular remainders, as in Eqs. (6.233), but now for multipole fields. The remainders can then again be integrated numerically, where as the singular terms are to be integrated analytically — here only in a one-dimensional form.

The outlined general method becomes unnecessarily complicated if saddle coils have only a few windings and a shielding yoke is absent. A method of integration in coils built up by straight wires and circular arcs has been developed by Munack [33]. Other information is found in Hawkes and Kasper [34], Plies [35] and Ding Shouqian et al. [36]. Methods for systems with inductance effects are dealt with in references [37–39], for example.

6.6 OTHER APPLICATIONS OF INTEGRAL EQUATIONS

In this section, we shall discuss very briefly some cases in which the general formalism of integral equations with singular kernels can be used. Here we consider only the modifications that are necessary to ensure that the methods work successfully.

6.6.1 Planar Fields

We assume now that the whole configuration does not depend on one cartesian coordinate, which may be the coordinate z . It is then possible to use the methods of complex analysis for purpose of field calculation. Another approach is the solution of integral equations. The surface charge density σ is now to be replaced with a *line charge density* η , the charge per unit arc on the boundary line in a cross section through the system.

The boundary B must now be parametrized as functions $s(t)$ and $\bar{x}(t)$, $\bar{y}(t)$, whichever is appropriate. Then the *Dirichlet problem* is defined by

$$\bar{V}(t) = \int_B \eta(s') Z(\bar{x}(t), \bar{y}(t); \bar{x}(t'), \bar{y}(t')) s'(t') dt'. \tag{6.262}$$

The kernel Z is here given by

$$Z(x, y; x', y') = Z_0 - \frac{1}{4\pi} \ln[(x - x')^2 + (y - y')^2]. \quad (6.263)$$

It has the same strength of singularity as the Fourier-Green function but is distinctly simpler.

If this integral equation is discretized in such a manner that the resulting linear system of equations becomes symmetric, it will be positive definite only when the constant Z_0 is chosen large enough. The *same* constant must be used again if the integral representation

$$V(x, y) = \int_B \eta(s) \dot{s}(t) Z(x, y; \bar{x}(t), \bar{y}(t)) dt, \quad (6.264)$$

and its derivatives for x and y are used for purposes of field calculation; it has no influence on the results.

One important difference from the former cases is the fact that the *natural* boundary conditions do *not* exist here. Hence the domain of solution must always be *closed*, and only its *interior* can be used for field calculation. If these conditions are respected, the previously outlined general techniques can be applied to planar fields.

The integral equations for magnetic fields in systems with iron yokes can also be transferred and solved accordingly. This is much simpler than for those dealt with in Section 6.4, as there is *no singularity*

$$N \cdot \nabla Z = \frac{-N_x(x - x') - N_y(y - y')}{2\pi[(x - x')^2 + (y - y')^2]} \longrightarrow -\frac{\kappa \dot{s}}{4\pi}, \quad (6.265)$$

in contrast to Eqs. (6.103).

6.6.2 Wave Fields

The calculation of high-frequency fields in resonant cavities is a task of great importance in electrical engineering, and consequently, the literature in this topic is huge. Simple examples can already be found in any textbook for graduate students. It is not the object of the present volume to deal with this topic in detail; we give just one example showing how the family of Fourier-Green functions can be generalized to wave fields. This implies that the subsequent considerations can hold only for *rotationally symmetric* cavities.

As a simple example, we consider the *Dirichlet* problem (with $w = 2\pi/\lambda$ instead of k).

$$\int_B G_w(\mathbf{r}, \mathbf{r}') \sigma(\mathbf{r}') d\mathbf{a}' = 0, \quad \mathbf{r} \in B, \quad (6.266)$$

in which $\sigma(\mathbf{r})$ is again a surface source function and G_w is now the Green function given by Eq. (1.126). Equation (6.266) is the specialization of Eq. (1.125) for boundary points. A nontrivial solution is then only possible if w becomes an *eigenvalue* of the integral equation. Although the kernel should have only real values in the preceding defined case, we generalize it to

$$G_w(\mathbf{r}, \mathbf{r}') := \frac{\exp(iwD)}{4\pi D}, \quad (6.267a)$$

$$D := |\mathbf{r} - \mathbf{r}'| \quad (6.267b)$$

being the distance between the two positions.

We can again write down a Fourier-series expansion for $\sigma(\mathbf{r})$, whereupon we obtain a generalization of Eq. (6.67) in the form

$$G_w^m(z, r; z', r') = \frac{1}{4\pi} \int_0^{2\pi} \frac{\exp[im\psi + iwD(\psi)] d\psi}{D(\psi)}, \quad (6.268a)$$

with

$$D(\psi) = [(z - z')^2 + r^2 + r'^2 - 2rr' \cos \psi]^{1/2}. \quad (6.268b)$$

However, as $D(\psi) = D(-\psi)$, the contribution in $\sin m\psi$ must vanish for reasons of symmetry and we arrive at

$$G_w^m(z, r; z', r') = \frac{1}{2\pi} \int_0^\pi D^{-1}(\psi) \cos m\psi e^{iwD} d\psi, \quad (6.269)$$

which comprises Eq. (6.67) as the special case $w = 0$. It is again possible to introduce the moduli, defined in Section 6.2 and we then obtain

$$G_w^m = (\pi d_2)^{-1} \int_0^{\pi/2} \frac{(-1)^m \cos 2m\alpha}{(1 - k^2 \sin^2 \alpha)^{1/2}} e^{i\Phi(\alpha)} d\alpha, \quad (6.270a)$$

$$\Phi(\alpha) := w d_2 (1 - k^2 \sin^2 \alpha)^{1/2} \quad (6.270b)$$

as a generalization of Eqs. (6.75) on replacement of α with $\pi/2 - \alpha$. The Landen transform to the modulus ρ is now not advantageous, as this complicates the form of the exponent. Apart from this Green function, we need its partial derivatives with respect to z and r , which lead to similar expressions, but with the exponent $3/2$ in the denominator.

Since analytical methods for the evaluation of such integrals are not known, we shall again discuss numerical techniques, which must, of course, require more effort than those in Section 6.2. One way is based on Taylor-series expansions.

The Taylor-series expansion of the exponential function is absolutely convergent for all arguments but requires an increasing number of terms with increasing argument to reach a given threshold. Hence this technique is feasible only for sufficiently small values of Φ . Since the factor $\cos 2m\alpha$ can be written as a polynomial of degree m in terms of $\sin^2 \alpha$, we encounter integrals of the form.

$$J_{m,n}(k) := \int_0^{\pi/2} \sin^{2m} \alpha (1 - k^2 \sin^2 \alpha)^{n/2} d\alpha, \\ (m = 0, 1, 2, \dots, n = -3, -2, \dots). \quad (6.271)$$

The integrals of lowest orders are well-known, ($\bar{k}$ being the complementary modulus):

$$J_{0,-3} = E(k)/\bar{k}^2, \quad J_{0,-2} = \frac{\pi}{2\bar{k}}, \quad J_{0,-1} = K(k), \\ J_{0,0} = \pi/2, \quad J_{0,1} = E(k), \quad J_{0,2} = \frac{\pi}{2} (1 - k^2/2). \quad (6.272)$$

Generally, the integrals of higher orders can easily be obtained from the recurrence relation [40]

$$J_{0,n} = \frac{n-1}{n} (2 - k^2) J_{0,n-2} - \frac{n-2}{n} J_{0,n-4}, \quad (n \geq 1), \quad (6.273)$$

which is numerically stable in ascending direction. The integrals with $m > 0$ can then be determined by linear combination of these functions:

$$J_{m,n} = (J_{m-1,n} - J_{m-1,n+2})/k^2, \quad (m > 0), \quad (6.274)$$

the most familiar example being

$$J_{1,-1} = D(k) = (K(k) - E(k))/k^2, \quad (n = -1). \quad (6.275)$$

The Green function for $m = 0$ is now, for example, given by evaluation of

$$G_w^0 = (\pi d_2)^{-1} \sum_{n=0}^N \frac{(iwd_2)^n}{n!} J_{0,n-1}(k). \quad (6.276)$$

However, the power and the factorial should not be evaluated in this simple manner but combined in such a way that very large numbers do not appear

in the numerator. Röhlm [40], who has tested this method, found that it works successfully for $wd_2 \leq 34$ and $N \leq 120$, the threshold being 10^{-6} .

Another method of evaluating integrals of the form of Eqs. (6.270) consists in the separation of a *constant* exponential factor:

$$\exp(i\Phi(\alpha)) = \exp(i\Phi_M) \exp[i(\Phi(\alpha) - \Phi_M)]. \quad (6.277)$$

It is favorable to choose the *mean* value

$$\Phi_M := wd_2 \sqrt{1 - k^2/2} = \Phi(\pi/4), \quad (6.278a)$$

whereupon the second factor can be written as

$$\exp[i(\Phi - \Phi_M)] = \exp \left[\frac{0.5 i w d_2 k^2 \cos 2\alpha}{(1 - k^2/2)^{1/2} + (1 - k^2 \sin^2 \alpha)^{1/2}} \right]. \quad (6.278b)$$

For $k \ll 1$, especially that is near the optic axis or at very large distance from the ring, this function oscillates so slowly that a numerical Romberg quadrature makes sense.

A detailed analysis shows that the function G_w^m has the same strength of logarithmic singularity as G_m , the factor being again $(2\pi)^{-1}$. The normal derivative has the same singularity amplitude. Only the regular remainder contains additional terms. Hence, the general methods for solving integral equations are again applicable. An important new task is now, the calculation of eigenvalues from the condition that the determinant of the system matrix must vanish. This is a standard problem of linear algebra and is not discussed here.

The BEM for high-frequency fields offers the possibility of calculating configurations in which the standard FEM techniques become unfeasible as result of extreme differences in the geometric proportions. Röhlm [40], for instance, applied the BEM to a resonant cavity with a tip cathode for the extraction of electron beams. It is beyond the scope of this volume to discuss in detail such examples, and other questions such as initial value problems and wake field effects, for which we refer to the corresponding literature [41,42]. The intention of this section was merely to show that the generalizations outlined are possible.

REFERENCES

1. Singer, B. and Braun, M. (1970). Integral equation method for computer evaluation of electron optics, *IEEE Trans. Electron Dev.* **ED-17**, 10, 926–934.

2. Harting, E. and Read, F. H. (1976). *Electrostatic Lenses*, Amsterdam: Elsevier.
3. Eupper, M. (1982). Eine Methode zur Lösung des Dirichlet-Problems in drei Dimensionen und ihre Anwendung auf einen neuartigen Elektronenstrahlerzeuger, *Optik* **62**: 299–307.
4. Eupper, M. (1985). Eine verbesserte Integralgleichungsmethode zur numerischen Lösung dreidimensionaler Dirichletprobleme und ihre Anwendung in der Elektronenoptik, Dissertation, Universität, Tübingen, Fakultät für Physik, Germany.
5. Munro, E. and Chu, H. C. (1982). Numerical analysis of electron beam lithography systems, Part II: Computation of fields in electrostatic deflectors, *Optik* **61**, 1: 1–16.
6. Ströer, M. (1988). The integral equation method for field calculations in three dimensions and its reduction to a sequence of two-dimensional problems, *Optik* **81**: 12–20.
7. Bowring, N. J. and Read, F. H. (1994). *Charged Particle Optics—3 Dimensional*, presented at a computer workshop in Delft.
8. Read, F. H. (1999). Edgeways electrostatic deflector with reduced aberrations, *Nucl. Instrum. Meth. A* **427**: 177–181.
9. Cubric, D., Lencova, B., Read, F. H. and Zlamal, J. (1999). Comparison of FDM, FEM, and BEM for electrostatic charged particle optics, *Nucl. Instrum. Meth. A* **427**: 357–362.
10. Baker, C. (1977). *The Numerical Treatment of Integral Equations*. Oxford: Clarendon Press.
11. Jaswon, M. A. and Symm, G. T. (1977). *Integral Equation Methods in Potential Theory and Electrostatics*. London & New York: Academic Press.
12. Stakgold, I. (1979). *Green's Functions and Boundary Value Problems*. New York: Wiley.
13. Kress, R. (1989). Linear integral equations, *Applied Mathematical Sciences*, Volume 82, London, New York, Heidelberg, Berlin: Springer.
14. Pipkin, A. C. (1981). A course on integral equations, *Applied Mathematical Sciences*, Volume 9.
15. Polyanin, A. D. and Manzhirov, A. V. (1998). *Handbook of Integral Equations*. Boston, London, New York: CRC Press.
16. Ströer, M. (1990). Formulierung und Lösung magnetostatischer Probleme für rotationssymmetrische Anordnungen mittels Integralgleichungen, Dissertation, Universität Tübingen, Fakultät für Physik, Germany,
17. Kasper, E. and Ströer, M. (1989). A new method for the calculation of magnetic fields in systems with unsaturated yokes and its application in electron optics, *Optik* **83**: 93–100.
18. Kasper, E. and Ströer, M. (1990). A new method for the calculation of magnetic lenses, *Nucl. Instrum. Meth. A* **298**: 1–9.
19. Kasper, E. and Scherle, W. (1982). On the analytical calculation of fields in electron optical devices, *Optik* **60**: 339–352.
20. Abramowitz, M. and Stegun, I. A. (1970). *Handbook of Mathematical Functions*, New York: Dover Publications.
21. Weth, A. von der. (1997). Berechnung statischer axialsymmetrischer Magnetfelder mit der Methode der Randlelemente (BEM) unter Berücksichtigung nichtlinearer Materialeigenschaften, Dissertation, University Frankfurt (Main), Fachbereich Physik, Germany.
22. Hawkes, P. W. and Kasper, E. (1989). *Principle of Electron Optics*, Volume 1, Chapter 10. London & New York: Academic Press.
23. Kasper, E. (1983). On the solution of integral equations arising in electron optical field computations, *Optik* **64**: 157–169.
24. Scherle, W. (1983a). Berechnung von magnetischen Ablenssystemen, Dissertation, Universität Tübingen, Fakultät für Physik, Germany.
25. Scherle, W. (1983b). Eine Integralgleichungsmethode zur Berechnung magnetischer Felder von Anordnungen mit Medien unterschiedlicher Permeabilität, *Optik* **63**, 3: 217–226.
26. Ströer, M. (1987). Eine Galerkin-Methode mit singulären Formfunktionen und ihre Anwendung auf die Berechnung magnetostatischer Felder, *Optik* **77**: 15–25.

27. Becker, R. (1989). INTMAG: a program for the calculation of magnetic fields by integration, *Nucl. Instr. Meth. B* **42**: 303–306.
28. Becker, R. (1990). Magnetic fields calculated by INTMAG compared with analytical solutions and precision measurements, *Nucl. Instrum. Meth. A* **298**: 13–21.
29. Ströer, M. (1986). Treatment of corner singularities in integral equations, *In Proceedings of Eleventh International Congress on Electron Microscopy*, Kyoto, Japan, 289–290.
30. Kasper, E. (1987). An advanced method of field calculation in electron optical systems with axisymmetric boundaries, *Optik* **77**: 3–12.
31. Kasper, E. (2000). An advanced boundary element method for calculation of magnetic lenses, *Nucl. Instrum. Meth. A* **450**: 173–178.
32. Kasper, E. (1995). On the calculation of fields generated by coils, *J. Microscopy* **179**: 158–160.
33. Munack, H. (1990). Field calculations for saddle coils of large-angle deflection systems, *Optik* **85**: 161–166.
34. Hawkes, P. W. and Kasper, E. (1990). *Principles of Electron Optics*, Volume 2, Chapter 40. London & New York: Academic Press.
35. Plies, E. (1994). Electron optics of low-voltage electron beam testing and inspection. Part I: simulation tools. *Advances in Optical and Electron Microscopy*, Volume 13, pp. 123–242, London & New York: Academic Press.
36. Ding Shou-qian, Sheng Xia, Sun Jie (1995). The relations between the field parameters and magnetic potential harmonic coefficients, *Nucl. Instrum. Meth. A* **363**: 337–340.
37. Mayergoyz, I. D. (1984). Boundary Galerkin approach to the calculation of eddy currents in homogeneous conductors, *J. Appl. Phys.* **55**(6): 2192–2194.
38. Miyazawa, H., Koizumi, M., Mizuta, T., Oku, K., Shirai, S. and Yoshimi, I. (1995). Development of an electron optics simulator to consider the 3-D magnetic deflection fields in CRTs, *Nucl. Instrum. Meth. A* **363**: 341–346.
39. Miyazawa, H., Koizumi, M., Fukumoto, H., Hisada, T., Okuyama, N., Oku, K. and Shirai, S. (1995). Development of inductance simulation system for deflection yoke, *Nucl. Instrum. Meth. A* **427**: 214–218.
40. Röhm, E. (1986). Berechnung zeitabhängiger Felder mittels der Integralgleichungsmethode und die Anwendung auf ein Hochfrequenzstrahlerzeugungssystem, *Optik* **74**: 99–113.
41. Leis, R. (1985). *Initial Boundary Value Problems in Mathematical Physics*, New York: Wiley.
42. Kawaguchi, H. and Honma, T. (1995). On the wake fields reaction force which acts on electrons in an accelerator cavity, *Nucl. Instrum. Meth. A* **363**: 145–152.
43. Kasper, E. (2000). An advanced boundary element method for calculation of magnetic lenses, *Nucl. Instrum. Meth. A* **450**: 173–178.

CHAPTER VII

Hybrid Methods

The individual methods of field calculation, outlined in earlier chapters, are not always sufficient in every respect. As we have seen, they all have some advantageous properties but also disadvantages or even limitations of applicability. Good results can then be expected with hybrid methods, by which we mean with *combinations* of the pure methods, whenever it is possible to utilize the favorable properties and yet to circumvent the difficulties. Moreover, a general hybrid procedure can also include the use of suitable series expansions and field models. There is thus a great variety of choices. We cannot describe all of these in detail, and we hence confine this account to some relevant examples.

7.1 COMBINATION OF THE FEM WITH THE BEM

This is appropriate if the magnetic field in a round lens has to be calculated under two complicating conditions:

- (i) The yoke has an open structure, so that the fringe field is extensive.
- (ii) The lens is so strongly excited that saturation effects in the yoke have to be taken into account.

Case (i) in the absence of condition (ii) can be dealt with successfully by means of the BEM, as is demonstrated in Figs. 5.26 and 5.27 whereas the reverse case, nonlinear media in closed yokes, can be solved efficiently by using the FEM, see Figs. 5.24 and 5.25.

We consider now the example of an open lens, as shown in Fig. 7.1. The dissection of the pole piece into triangular elements is here simplified for reasons of demonstration; of course, more sophisticated meshes could be used. Provided that the *boundary values* of the flux function $\psi(z, r)$ on the yoke surface are *known*, the nonlinear Dirichlet problem could be solved in the interior by means of the FEM, and the inhomogeneously linear one in the exterior by means of the BEM. The problem is hence reduced to the task of finding these boundary values.

In the previous presentations of the BEM, we have used integral equations mainly to determine surface *sources* such as surface currents, see for example Eq. (6.177) as the field calculation can then be performed directly without

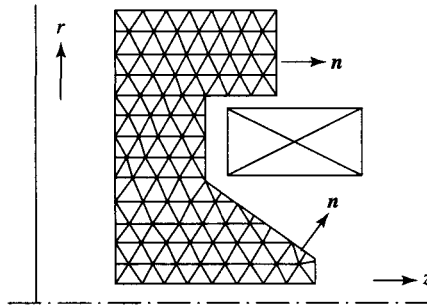


FIGURE 7.1 Half axial section through an open round magnetic lens with a ferromagnetic pole piece and a rectangular distribution of windings. Only the interior of the pole piece is discretized by a triangular mesh.

solving additional systems of equations—sufficiently far from the surfaces at least. This is now possible, as the assumptions on which their derivation was based are not valid. But Green's theorem Eq. (1.87) can also be used to determine linear relations between surface potentials and normal derivatives on one side of the boundary, here the *vacuum* side. There are only a few points to be noted:

- (i) The surface normal $\mathbf{n}$ is now directed towards the *interior* of the vacuum domain of integration, which implies that some signs must be changed.
- (ii) The spherical sector round the singularity is now located in the outside, and hence the complementary sector to the one shown in Fig. 1.7, is to be cut out. This requires a factor $\bar{\beta} = 1 - \beta(\mathbf{r}')$, β defined in Eq. (1.91).
- (iii) The vector potential and the vectorial Poisson equation (1.25) with $\mu = \mu_0$ must first be used in cartesian form; thereafter the transformation to cylindrical coordinates can be made.
- (iv) The spatial integration is confined to the coil (C) and the surface integration to the boundary (Y) of the yoke as contributions by an infinitely large enclosing sphere vanish as a result of the natural conditions.

On considering all this, we obtain

$$\begin{aligned}
 \mu_0 \int_C G(\mathbf{r}, \mathbf{r}') \mathbf{j}(\mathbf{r}) d^3r &=: \mathbf{A}_0(\mathbf{r}') \\
 &= \int_Y G(\mathbf{r}, \mathbf{r}') \mathbf{n}(\mathbf{r}) \cdot \nabla \mathbf{A}(\mathbf{r}) da + \bar{\beta}(\mathbf{r}') \mathbf{A}(\mathbf{r}') \\
 &\quad - \oint_Y \mathbf{A}(\mathbf{r}) \mathbf{n}(\mathbf{r}) \cdot \nabla G(\mathbf{r}, \mathbf{r}') da.
 \end{aligned} \tag{7.1}$$

We now introduce the function

$$\mathbf{A}(\mathbf{r}) = (2\pi r)^{-1} \Psi(z, r) \mathbf{e}_\varphi(\varphi) \quad (7.2)$$

into this integral equation, and multiply by $2\pi r' \mathbf{e}_{\varphi'}(\varphi')$ throughout, whereupon it becomes a scalar equation. The operator $\mathbf{n} \cdot \nabla = \partial_n = n_z \partial_z + n_r \partial_r$ does not depend on the azimuth. The integration over the latter then leads to the kernel function of Eq. (6.104) and its normal derivative. With the abbreviations $\mathbf{u} = (z, r)$, $\mathbf{u}' = (z', r')$ we finally arrive at

$$\begin{aligned} \Psi_0(\mathbf{u}') &= \bar{\beta}(\mathbf{u}') \Psi(\mathbf{u}') + \oint_Y \frac{1}{r} \Phi(\mathbf{u}, \mathbf{u}') \partial_n \Psi_v(\mathbf{u}) ds \\ &\quad - \oint_Y \frac{1}{r} \Psi(\mathbf{u}) \partial_n \Phi(\mathbf{u}, \mathbf{u}') ds, \end{aligned} \quad (7.3)$$

which is valid only for positions $\mathbf{u}'$ on the boundary Y . The subscript v indicates that this derivative must be evaluated on the vacuum side, while $\partial_n \Phi$ is the principal value, as is usual.

The term $\Psi_0(\mathbf{u}')$ is the flux produced by the coils in the absence of any yoke and can be evaluated by means of the methods outlined in Section 6.5. In this context it should be mentioned that Eq. (7.3) is *not* the same as Eq. (6.227), though the formal structure is quite similar; Eq. (6.227) refers to the surface of the coil, and the result of it must now be identified with $\Psi_0(\mathbf{u}')$, ($\mathbf{u}' \in Y$).

The integral equation (7.3) is still underdetermined because both the potential Ψ and its normal derivative are initially not known. However, the latter can be related to the normal derivative on the other side of the boundary by means of the condition

$$\partial_n \Psi_v(\mathbf{u}) = \mu_0 \mu_F^{-1} \partial_n \Psi_F(\mathbf{u}), \quad (7.4)$$

the label F denoting the ferromagnetic yoke. This condition represents the necessary coupling between the two domains. Anticipating an iterative technique, we can rewrite Eq. (7.3) in operator form

$$\begin{aligned} \mathbf{H} \Psi(\mathbf{u}') &:= \bar{\beta}(\mathbf{u}') \Psi(\mathbf{u}') - \oint_Y \Psi(\mathbf{u}) r^{-1} \partial_n \Phi(\mathbf{u}, \mathbf{u}') ds, \\ &= \Psi_0(\mathbf{u}') - \int_Y r^{-1} \mu_0 \mu_F^{-1} \Phi(\mathbf{u}, \mathbf{u}') \partial_n \Psi_F(\mathbf{u}) ds. \end{aligned} \quad (7.5)$$

With $\bar{\beta} = 1/2$ on smooth boundary parts this integral operator $\mathbf{H}$ is essentially *adjoint* to the one in Eq. (6.177). In a discretized version this becomes a

system of equations

$$\sum_{k=1}^N H_{ik} \psi_k = \psi_i^{(0)} - \sum_{k=1}^N L_{ik} D_k, \quad (i = 1, \dots, N) \quad (7.6)$$

with the derivatives

$$D_k = \mu_0 \mu_k^{-1} \partial_n \psi_F(\mathbf{u}_k), \quad (k = 1, \dots, N). \quad (7.7)$$

The permeabilities $\mu_k = \mu_F(\beta_k)$ may now become nonlinear owing to the hysteresis effect. The matrix elements H_{ik} and L_{ik} depend only on the geometrical shape of the yoke and the chosen discretization. They need to be calculated only *once* and stored. The matrix $\mathbf{H}$ can then be decomposed by means of the *LU* algorithm (see Section 5.6.1), the decomposed matrix occupying the same memory locations. Also the vector Ψ_0 of inhomogeneities needs to be calculated only *once* by integration over the coil and stored.

All preparations are now ready to start an iterative procedure. Initially some reasonable guess of the boundary values ψ_i is made, and with this the Dirichlet problem is solved for the interior by means of the FEM. Then the vector $\mathbf{D}$ of derivatives can be determined by suitable differentiation techniques, to be applied in the corresponding finite elements. This is then substituted on the right-hand side of Eq. (7.6), whereupon an improved set of boundary values ψ_i is found by means of *LU* substitutions. A further iteration cycle of the same kind can now start. This procedure is repeated until sufficient convergence is achieved.

In the linear case (μ independent of $|\mathbf{B}|$), a constant rate of convergence of the relative error is to be expected for theoretical reasons, whereas the nonlinear case may be more complicated. There is little information available about this problem, as this method has not been tested in this form, in spite of an early proposal [1]. A major difficulty is certainly the increased programming effort, because the specific techniques of the FEM and the BEM must now be combined. Nevertheless, the initially stated difficulties can hardly be overcome by using simpler techniques. Other techniques are given in references [2–5].

At the end, when the boundary values $\Psi(\mathbf{u}')$ are known, the corresponding integral equation for surface current densities has to be solved:

$$\Psi(\mathbf{u}') = \Psi_0(\mathbf{u}') + 2\pi\mu_0 \oint \Phi(\mathbf{u}, \mathbf{u}') \omega(\mathbf{u}) ds. \quad (7.8)$$

After solving this, the integral relation and those obtained by differentiation can be used to calculate the field in the whole vacuum domain.

With the transformation (6.179) or (6.209), the discretized form of Eq. (7.8) becomes essentially the same as Eq. (6.207) and hence (6.206) is valid.

Because the material properties appear only in the parameter λ given in Eq. (6.178), the following approximation is reasonable: if the exact form of the field distribution in the ferromagnetic yoke is *not* required (it is certainly unnecessary for ray tracing), then the field in the whole *vacuum* domain can be simulated correctly by replacement of the constant λ with a suitable function $\lambda(t)$. Its sampling values must be chosen such that the relations

$$\sum_{i=1}^N \lambda_i J_i = 0, \quad \sum_{i=1}^N J_i = 0 \quad (7.9)$$

are simultaneously satisfied and that the value of $\mu = \mu_i$, to be used in Eq. (6.178) comes close to that obtained from the magnetization curve. The use of the FEM is then unnecessary, which leads to a considerable simplification. This is indeed possible, as the author has shown [6].

7.2 COMBINATION OF THE FDM WITH THE BEM

In Section 6.5.1, we have derived a method for the reduction of three-dimensional Coulomb integrals to two-dimensional ones, which already leads to a significant reduction of the computational effort. However, this may not be sufficient, especially if the space charge distribution is strongly inhomogeneous, so that very many elements are necessary to achieve sufficient accuracy. This situation usually arises in electron guns with space charge-limited emission and is most pronounced in the vicinity of the cathode, where the emitted electrons are still very slow. To overcome this difficulty, Kasper [1] has proposed a combination of the FDM with the BEM, which has been successfully implemented by Killes [7,8] who succeeded in calculating electron gun properties with very good accuracy. The determination of the space source distribution from emission characteristics of the gun and from ray tracings is beyond the scope of this volume. Here we consider this as given and refer to the corresponding publications [7,8]. We shall, however, outline the hybrid method in a fairly general form, which would be valid even in three dimensions, but for reasons of conciseness we shall present only a version valid for *rotationally symmetric* boundaries.

7.2.1 The General Procedure

We set out from Poisson's equation

$$\Delta V(\mathbf{r}) = -\rho(\mathbf{r}), \quad (7.10)$$

in which we have set $\varepsilon \rightarrow 1$ for convenience.

This is to be solved in a domain S with a rotationally symmetric boundary ∂S and prescribed boundary values on it:

$$V(\mathbf{r}) = \bar{V}(\mathbf{r}), \quad \mathbf{r} \in \partial S. \quad (7.11)$$

This Dirichlet problem can be split into a sequence of uncoupled two-dimensional ones by means of Fourier series expansion with respect to the azimuth, preferably in cylindric coordinates, resulting in

$$\begin{aligned} (\partial_{zz}^2 + \partial_{rr}^2 + r^{-1}\partial_r - m^2 r^{-2}) V_m(z, r) &= -\rho_m(z, r), \\ (m = 0, \pm 1, \pm 2, \dots), \end{aligned} \quad (7.12)$$

with given boundary values $\bar{V}_m(z, r)$ on the boundary lines ∂D of the two-dimensional domain D in the z, r -plane.

We now introduce a decomposition

$$V_m(z, r) = V_m^p(z, r) + U_m(z, r) \quad (7.13)$$

with boundary values

$$V_m^p = \bar{V}_m^p, \quad U_m = \bar{U}_m = \bar{V}_m - \bar{V}_m^p, \quad (z, r) \in \partial D, \quad (7.14)$$

$V_m^p(z, r)$ being a *particular* solution of Eq. (7.12) and $U_m(z, r)$ a solution of the associated *homogeneous* PDE. The latter can be determined using the BEM, if the boundary values $\bar{U}_m$ on ∂D are known because the necessary conditions for this are satisfied. These solutions can be cast in the integral form

$$U_m(z, r) = \oint_{\partial D} G_m(z, r; \bar{z}(s), \bar{r}(s)) \sigma_m(s) \bar{r}(s) ds, \quad (7.15)$$

G_m being the Fourier-Green function (6.67).

The PDE (7.12) for V_m^p does not have a favorable form if $m \neq 0$; we hence introduce modified functions thus:

$$V_m^p =: r^{|m|} W_m(z, r), \quad \rho_m =: r^{|m|} g_m(z, r), \quad (7.16)$$

whereupon we obtain the standard form

$$\Delta_\alpha W_m(z, r) = -g_m(z, r), \quad \alpha = 2|m| + 1, \quad (7.17)$$

which is the same as Eq. (2.58). This can be solved accurately by means of the FDM, provided that the mesh fits the boundary, which will generally *not* be

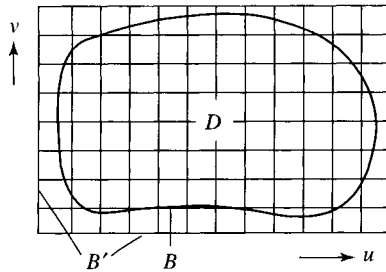


FIGURE 7.2 Extension of a domain D with curved boundary B to a rectangular domain with boundary B' .

the case. Even the transformation into curvilinear coordinates u and v will not always solve this problem. Instead of using irregular meshes, which is always feasible but depends on the circumstances and is less accurate, we propose the following method, shown in Fig. 7.2.

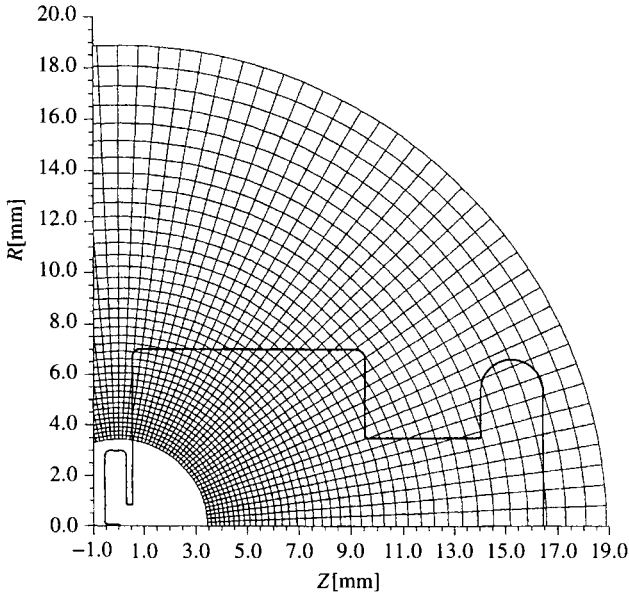
The mesh is *extended* beyond the physical boundary so that only *regular* elements are obtained and nine-point formulas can be applied without exception. The boundary values of the transformed potential $\Phi_m(u, v) = W_m(z, r)$ at the outer boundary B' can be set at will, for instance zero; they must remain unaltered. If the source function does not vanish on the inner boundary B , it is extrapolated beyond this; this can be done arbitrarily provided that this function remains continuous or has at least very small discontinuities. These extrapolated values will not contribute to the final solution in the interior. The Dirichlet problem for the extended domain is now well-defined and easily soluble by means of the FDM. This can be performed completely independently of the BEM contribution and must be done as the first step.

The interpolation techniques outlined in Sections 5.5.3 and 5.5.4 can now be used to determine the function values on the inner boundary $B = \partial D$.

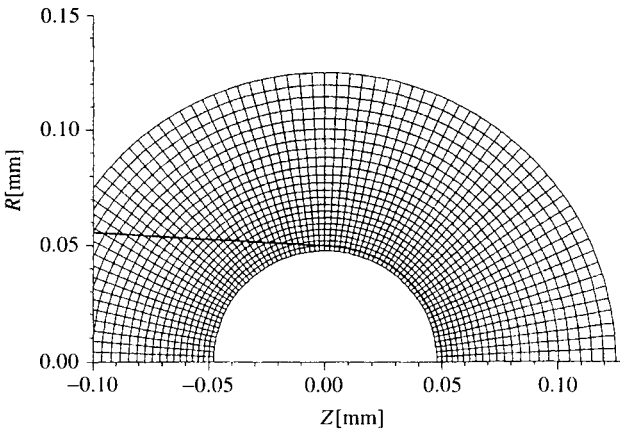
$$V_m^p(\bar{z}, \bar{r}) = (\bar{r})^{|m|} W_m(\bar{z}, \bar{r}) = (\bar{r})^{|m|} \Phi_m(\bar{u}, \bar{v}), \quad (7.18)$$

which are then introduced on the left-hand side of Eq. (7.15) to solve this Fredholm equation for $\sigma_m(s)$. Once this is complete, a field calculation in the whole domain D is possible, and by Fourier synthesis in the three-dimensional domain S , whereupon the problem is solved. We emphasize that the different steps are *uncoupled*. An iterative procedure arises only from the need to calculate the space charge from ray tracings.

A typical application of this method is the field calculation in an electron gun, with hemispherical cathode on the tip of a conical shank as is shown in Fig. 7.3a,b. In this case the spherical mesh, defined by Eqs. (4.167) is most favorable; its parameters should be chosen such that the mesh fits the cathode



(a)



(b)

FIGURE 7.3 A triode gun with hemispherical cathode on a conical shank in a spherical mesh: (a) coarse view; (b) vicinity of the cathode; Killes [8].

surface in the vicinity of its apex. The nine-point formulas, derived in Section 4.4.7, can then be employed. As the charge density has its maximum just at the cathode surface, it should be extrapolated linearly up to the margin of the mesh to prevent strong inhomogeneities.

7.2.2 The Modified Galerkin Method

A complete field calculation within the frame of the BEM is possible by means of the techniques described in Section 6.4.5. This is even accurate enough for ray tracing, if the particles have a high velocity, for instance, if a beam is to be focused on a target. However, in the case of electron emission from a cathode, the particles are initially so slow that interpolation errors may become significant. In this case, it is appropriate to shift the charge-carrying surface back into the interior of the cathode, so that there remains always a finite distance between them. All necessary integrations can then be carried out accurately by means of Gauss quadratures. Such a modification is always possible in the case of Dirichlet problems, if the field in the *exterior* is not needed and may hence be altered: see Fig. 7.4.

Let $\mathbf{u}_n = (\bar{z}_n, \bar{r}_n)$ ($n = 0, \dots, L$) be the sampling points and $\dot{\mathbf{u}}_n = (\dot{\bar{z}}_n, \dot{\bar{r}}_n)$ be the derivatives on the physical boundary B ; its curve parameter being denoted by t . The corresponding sampling positions on the shifted boundary $\tilde{B}$ with parameter τ are denoted by $\mathbf{v}_n := (\tilde{z}_n, \tilde{r}_n)$. A simple definition is then given by

$$\tilde{z}_n = \bar{z}_n - a_n \dot{\bar{r}}_n, \quad \tilde{r}_n = \bar{r}_n + a_n \dot{\bar{z}}_n, \quad (n = 0, \dots, L). \quad (7.19)$$

The shift parameter a_n should be proportional to the mean interval length:

$$a_n = C(t_{n+1} - t_{n-1})/2, \quad (7.20)$$

with cyclic or symmetric conditions at the end points. Killes [8] suggested the choice $C = 0.5$. The curve $\tilde{B}$ joining these points can then be constructed by suitable splines, as is outlined in Section 3.4. Some care is necessary in the vicinity of sharp edges: these must necessarily be rounded off with a sufficiently large radius of curvature! Otherwise, the above algorithm will break down.

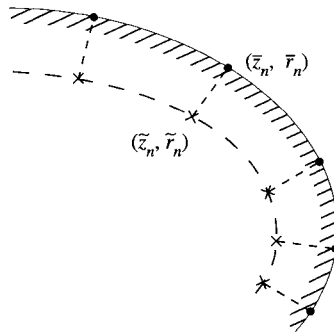


FIGURE 7.4 Construction of a charge carrying surface in the interior of an electrode.

The modified Galerkin method is now outlined briefly in a more general form. We start from a Fredholm equation

$$\bar{U}(t) \equiv U(\mathbf{u}(t)) = \int_B K(\mathbf{u}(t), \mathbf{v}(\tau)) Q(\tau) d\tau, \quad (\mathbf{u}(t) \in B). \quad (7.21)$$

A set of M linearly independent trial functions $N_j(\tau)$ for interpolation on $\tilde{B}$ must now be chosen and an *equal* number of trial functions $\bar{N}_i(t)$ referring to the curve B . As usual, the unknown function $Q(\tau)$ is expanded as

$$Q(\tau) = \sum_{j=1}^M S_j N_j(\tau) \quad (7.22)$$

to be introduced into Eq. (7.21), and again the linear system of equations is obtained by forming projections, now with the functions $\bar{N}_i(t)$. The result is then given by

$$\sum_{j=1}^M K_{ij} S_j = T_i \equiv \langle \bar{N}_i | \bar{U} \rangle, \quad (i = 1, \dots, M) \quad (7.23)$$

with the matrix elements

$$K_{ij} = \int_B \int_{\tilde{B}} \bar{N}_i(t) K(\mathbf{u}(t), \mathbf{v}(\tau)) N_j(\tau) d\tau dt, \quad (7.24)$$

which are now *asymmetric*. After the determination of the coefficients S_j , a general field calculation is possible in the form

$$U(\mathbf{u}) = \sum_{j=1}^M S_j \int_{\tilde{B}} K(\mathbf{u}, \mathbf{v}(\tau)) N_j(\tau) d\tau, \quad (7.25)$$

$\mathbf{u} = (z, r)$ here being any position *inside* the domain D or on the boundary B .

Suitable sets of trial functions are again cubic Hermite polynomials (Eqs. (6.123), (6.135), and (6.136)) with adaptation of the notation. Other choices are the modified interpolation kernels for nonequidistant intervals (see Section 3.3.4).

The accuracy gained with this hybrid method depends on the particular geometrical configuration and can be quite high, provided that the space charge density can be calculated correctly. Usually this latter part, the determination of the space charge from ray tracings causes the dominant error, but this topic goes beyond the scope of this volume. The field calculation

method in its general form offers the possibility of determining also nonrotationally symmetric configurations, in which the deviations from the rotational symmetry may similarly be caused by perturbations of the boundary conditions or asymmetry of the space charge cloud. Its application is then limited by the highest necessary multipole order, which can still be computed with acceptable accuracy.

7.3 THE CHARGE SIMULATION METHOD (CSM)

This method is a generalization of the technique for shifting the singularities of the charge distribution into the interior of electrodes.

In the preceding section, these were still surface charges on boundaries close to the real material surfaces; now, however, these sources can be suitable functions for which a reasonably simple analytical potential can be found. The distance from the electrode surface may be arbitrarily large, it is only important to adjust the free parameters of the chosen field model in such a way that the physically prescribed boundary conditions—most frequently a Dirichlet problem—are approximated as accurately as possible.

The most familiar representative of such a simulation is the potential of a spherically symmetric charge distribution: in the charge-free domain outside the sphere this is *exactly* the same as the Coulomb potential produced by the concentration of the total charge in the center. Furthermore, any other spherically symmetric distribution of the same total charge would produce the same field in the outer domain. Hence extrapolation back from a given potential field to the source distribution producing it is *not unique*. This rule, which is evident in this simple well-known example, must, of course, hold generally.

7.3.1 The General Procedure

We set out from a very general series expansion of the form

$$V(\mathbf{r}) = \sum_{j=1}^N p_j V_j(\mathbf{r}), \quad (7.26a)$$

with *free* parameters $p_1, \dots, p_N$ and simulation potentials $V_j(\mathbf{r})$, which should be well-known *analytical* functions. This latter condition implies here that these functions must be continuously differentiable arbitrarily often except at their singularities, and that such singularities, if there are any, must be situated a *finite* distance from the boundary B of the physical domain D . This implies

that we can form

$$\nabla V(\mathbf{r}) = \sum_{j=1}^N p_j \nabla V_j(\mathbf{r}), \quad (7.26b)$$

$$\Delta V(\mathbf{r}) = \sum_{j=1}^N p_j \Delta V_j(\mathbf{r}), \quad (\Delta \equiv \nabla^2) \quad (7.26c)$$

at least in D and in the vicinity of B outside D . Most frequently a Dirichlet problem of Laplace equation is to be solved, in the form

$$\Delta V(\mathbf{r}) = 0, \quad (\mathbf{r} \in B \text{ or } \mathbf{r} \in D). \quad (7.26d)$$

However, this latter condition is not necessary. Thus, it may be possible to define a potential, which simulates approximately a space charge distribution in front of a cathode. This must then be a *regular* function, which implies that caustics of the electron beam must not be present. As such applications are rather limited, we only point out this possibility and do not pursue it further.

The CSM is favorable for the solution of Dirichlet problems. It becomes more complicated but still feasible in case of Neumann conditions and is practically unusable for the solution of boundary value problems involving material conditions. Hence, we consider here a Dirichlet problem that can be cast in the form

$$V(\mathbf{r}) = \bar{V}(\mathbf{r}) \quad (\mathbf{r} \in B). \quad (7.27)$$

A general but not necessarily the best technique for solving it is *collocation*. We have to define exactly N suitable sampling positions $\mathbf{r}_1, \dots, \mathbf{r}_N$ on the boundary B , and to solve the system of equations.

$$\sum_{j=1}^N V_j(\mathbf{r}_i) p_j = \bar{V}(\mathbf{r}_i), \quad (i = 1, \dots, N). \quad (7.28)$$

Provided that the system matrix is well conditioned, we obtain a reliable set of parameters p_j , which can be introduced into Eqs. (7.26), whereupon the field can be calculated at any position $\mathbf{r} \in D$. This does *not* automatically give high accuracy as in the midpoints between adjacent sampling points the deviations from (7.27) may be quite large for inappropriate choices.

Another, more accurate method is the *least squares fit* (LSQ) approximation that can be cast in the form

$$\sum_{k=1}^M w_k \left(\bar{V}(\mathbf{r}_k) - \sum_{j=1}^N p_j V_j(\mathbf{r}_k) \right)^2 =: \varepsilon = \min, \quad (7.29)$$

in which the number M of surface points should be *larger* than N and the associated summation weights $w_k > 0$. The minimization conditions yields the linear system of equations

$$\sum_{j=1}^N p_j \left(\sum_{k=1}^M w_k (V_i(\mathbf{r}_k) V_j(\mathbf{r}_k)) \right) = \sum_{k=1}^M w_k V_i(\mathbf{r}_k) \bar{V}(\mathbf{r}_k), \quad (i = 1, \dots, N). \quad (7.30)$$

If the configuration is chosen reasonably, as is assumed here, the symmetric system matrix will also be positive definite, so that the system (7.30) can then be solved by means of the Cholesky algorithm. A matrix with poor condition indicates that the simulation model was chosen inadequately and that it should be modified in such a way that this situation does not arise again. With suitable matrix allocation techniques, analogous to those in the FEM the computational effort can be kept reasonably small: in fact, only $N \cdot M$ evaluations of the analytical trial functions are really necessary.

The resulting potential satisfies the prescribed boundary conditions only on average, however the oscillatory errors decrease rapidly (exponentially) with increasing distance from the boundaries.

There is considerable freedom in the choice of a suitable type of charge distribution and the corresponding Coulomb potential: from among many possible examples, we mention systems of point charges, dipoles [9], line elements [10], grids of wires [11,12], coaxial rings [13,14], and additional sets of charged plates [15]. Finding the most favorable type of charge distribution is a matter of intuition and experience; the appropriate answer also depends on the nature of the device in question with the result that there can be no general recipes. The example presented below is hence only one among the many possible demonstrations of an application of the CSM.

7.3.2 *Pointed Cathode Models*

The application of suitable field models is appropriate in all cases, in which the BEM causes difficulties owing to extreme differences in the geometric proportions, the FDM and the FEM being still more unfavorable. An example of such a configuration is an electron gun with a pointed cathode for field-electron emission. The emission process itself cannot be dealt with here, and we refer to the corresponding literature [16]. The actual problem in this section arises from the fact that the hemispherical cathode, welded on a conical shank, has a radius of about 10^{-5} to 10^{-6} mm, whereas the other geometric parameters are of the order 1 mm. Even the FDM with spherical mesh (Section 4.4.7) runs into numerical problems. There are different ways of overcoming these. One

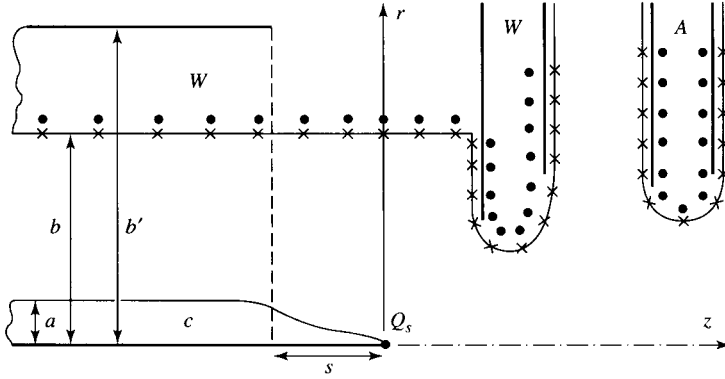


FIGURE 7.5 Simulation of an electron gun with pointed cathods by a point charge, a system of charged axial wires, a sequence of rings close to the surface, two pairs of aperture plates, and a Wehnelt tube in a half-meridional section. Kasper [17]. (The notation has been adapted to conform with the used in this chapter.)

of them is shown in Fig. 7.5 [16–18]. In these papers, the use of combined electrostatic-magnetic fields is also discussed.

The cathode is here simulated by the superposition of the potentials caused by a *point charge* and two *axial wires* (line charges). The anodes are simulated by a cylindrical tube of radius $b' > b$, a set of *ring charges* and two pairs of *aperture plates*. The latter will be the topic of Section 7.3.3. This cathode model is too limited, as it has only two free parameters. A model with more degrees of freedom is given by the following set of functions:

$$\begin{aligned}
 V_1(z, r) &= \ln(z + R) - \ln[z + s + ((z + s)^2 + r^2)^{1/2}] + 2sG_0(z, r; -s, b) \\
 V_2(z, r) &= R^{-1} - 2G_0(z, r; -s, b) \\
 V_3(z, r) &= -zR^{-3} \\
 V_4(z, r) &= \frac{1}{2} (3z^2 - R^2)R^{-5}
 \end{aligned} \tag{7.31}$$

with $R := (z^2 + r^2)^{1/2}$.

Here the origin of the cylindrical coordinate system is located at the point charge, which is also the end point of the line charge. The parameter s is the length of the shank and b the radius of the tube. The charge on the far distant ring with position $(-s, b)$ neutralizes the total charge of this cathode model, see Fig. 7.6. In the close vicinity of the tip, that means for $R \ll s$, the second logarithm can be approximated by $\ln(2s)$ and the ring potential by

$$2G_0(z, r; -s, b) = (b^2 + s^2)^{-1/2}. \tag{7.32}$$

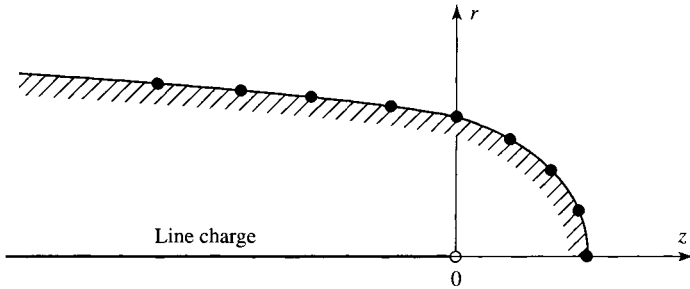


FIGURE 7.6 Definition of a cathode surface in the vicinity of its apex by some control points. The multipole sources are located in the origin 0, which is also the endpoint of the finite line charge.

It is then also permissible to assume that the external potential V_E , that is the potential in the absence of the above-defined model potential, is locally constant and may be evaluated at the origin. If the cathode surface has the constant potential V_c , we have now to solve Eqs. (7.30) with $\bar{V} = V_c - V_E$ and $N = 4$. For this purpose some few sampling points on the tip and on the transition to the shank are to be defined, as shown in Fig. 7.6. After obtaining the solution the cathode surface is defined as the equipotential $V(z, r) = V_c$, V now being the total potential. Outside the singularities this is an exact solution of Laplace equation. If necessary, the number of trial functions can be increased beyond four, but we shall not pursue this here.

The Cathode Tube

A major problem is the calculation of the field produced by the *cathode support* as this has the geometrical form of a hairpin. Such a field calculation is feasible (see Eupper [10]), but much more complicated than the whole of the remaining calculations. Therefore, we choose here a simplified model that does not produce a completely wrong fringe field. The *anode tube* of radius b has the axial extent $-L \leq z \leq -s$ with $L \gg b$, a constant potential U_T and constant surface charge density. The cathode support has the same axial extent, the potential U_c and asymptotically a cylinder radius $a < b$. It is simulated by an *axial line charge* distribution the density of which is to be determined by these conditions. The length L is chosen so large that the fringe field from the remote side $z = -L$ can be ignored. In the vicinity of the end at $z = -s$, the whole cathode surface is then again defined as the equipotential $V(z, r) = U_c$ resulting from the superposition of all partial potentials. Although it is not difficult to calculate with two values U_c and U_T , there is no loss of generality to set $U_T = 0$, and this is now assumed for reasons of conciseness.

The asymptotic potential in the tube is then given by

$$V_{as}(r) = U_c \ln(b/r) / \ln(b/a), \quad (U_c < 0), \quad (7.33)$$

which evidently satisfies the above-defined boundary conditions. The surface charge density on the tube wall becomes

$$\sigma = -\varepsilon_0 U_c (b \ln(b/a))^{-1} > 0, \quad (7.34a)$$

and the line charge density on the axis is then

$$dQ/dz = -2\pi b\sigma =: 4\pi\varepsilon_0 q < 0. \quad (7.34b)$$

The evaluation of the Coulomb integral over this charge distribution with subsequent limitation for $L \rightarrow \infty$ results in the integral

$$\begin{aligned} V_T(w, r) &= \frac{q}{\pi} \int_0^\pi \ln(w + (w^2 + r^2 + b^2 - 2br \cos \varphi)^{1/2}) d\varphi \\ &\quad - q \ln(w + (w^2 + r^2)^{1/2}) \end{aligned} \quad (7.35)$$

with $w := z + s$. Quite generally, the identity

$$w + \sqrt{w^2 + d^2} = d^2 / (\sqrt{w^2 + d^2} - w) \quad (7.36)$$

can be used and is favorable for $w < 0$. We obtain then the symmetry relation

$$\begin{aligned} V_T(w, r) + V_T(-w, r) &= \frac{q}{\pi} \int_0^\pi \ln \left(1 - \frac{2b}{r} \cos \varphi + \frac{b^2}{r^2} \right) d\varphi \\ &= 2q \ln(b/r) = V_{as}(r) \end{aligned} \quad (7.37)$$

for $r < b$. Hence in the front plane $w = 0, z = -s$ the value $V_T = V_{as}/2$ is obtained, as it must be.

An expression for the integral in Eq. (7.35) in closed form is not known, but as it is a solution of Laplace's equation, an approximate evaluation by means of the paraxial series expansion Eq. (2.72) with $\alpha = 1$ is easily possible and results in

$$\begin{aligned} V_T(w, r) &= q \ln \left[\frac{w + \sqrt{w^2 + b^2}}{w + \sqrt{w^2 + r^2}} \right] \\ &\quad - q \sum_{n=1}^{\infty} \frac{1}{2n} \binom{-1/2}{n} \left(\frac{r^2}{w^2 + b^2} \right)^n P_{2n-1} \left(\frac{w}{\sqrt{w^2 + b^2}} \right), \end{aligned} \quad (7.38)$$

where P_{2n-1} denotes Legendre polynomials of their argument. With this series expansion, Eq. (7.37) can again be confirmed, hence $V_T(0, r) = V_{as}/2$, $V_T(-\infty, r) = V_{as}$, $V_T(+\infty, r) = 0$. This paraxial series expansion is most favorable in the vicinity of the cathode tip and the electron beam, where very few terms are sufficient. Its convergence becomes slow in the vicinity of the edge ($w = 0, r = b$), but with not more than 20 terms a reasonable accuracy can always be obtained for $r^2 \leq w^2 + b^2$.

The derivatives of this potential can be calculated by summation of the differentiated series expansions, or if these converge too slowly, in analytical form. These consist each of a contribution (1) from the axial wire and (2) from the shielding wall. The former are elementary functions: with

$$R_0 := \sqrt{w^2 + r^2} \equiv [(z + s)^2 + r^2]^{1/2} \quad (7.39)$$

being the distance from the end point of the wire, we obtain then the elementary functions

$$V_{T|w}^{(1)} = -q/R_0, \quad (7.40a)$$

$$V_{T|r}^{(1)} = \begin{cases} -qr/(R_0(w + R_0)) & (w \geq 0) \\ -qr/(R_0(w - R_0)) - 2q/r & (w < 0). \end{cases} \quad (7.40b)$$

Moreover, we obtain exactly

$$V_{T|w}^{(2)} = 2qG_0(w, r; 0, b) \equiv 2qG_0(z, r; -s, b). \quad (7.40c)$$

This is the same kernel function as is already required in Eq. (7.31). The evaluated of $V_{T|r}^{(2)}$ by differentiation under the integral leads to an elliptic integral of the third kind. Its evaluation is rather complicated; hence we shall not present here this expression. A fast method is the differentiation of the series expansion resulting in

$$V_{T|r}^{(2)} = -\frac{qr}{w^2 + b^2} \sum_{n=1}^{\infty} \binom{-1/2}{n} \left(\frac{r^2}{w^2 + b^2} \right)^{n-1} P_{2n-1} \left(\frac{w}{\sqrt{w^2 + b^2}} \right). \quad (7.40d)$$

This has a similar form to the second part of Eq. (7.38), the essential difference being the fact that now, the factor $(2n)^{-1}$ is missing. This has the consequence that the rate of convergence is substantially slower; this series expansion is therefore sufficiently fast only for $4r^2 \leq w^2 + b^2$, which is perfectly justified in the domain of the electron beam. If this derivative is ever needed in the far zone, it can be determined by numerical differentiation of Eq. (7.38).

The Complete Field

It remains to incorporate such a field model into a general program, which allows also the calculation to be performed for other anode profiles. Quite generally the potential $V(z, r)$ is built up from four contributions:

$$V(z, r) = V_T(z + s, r) + V_A(z, r) + V_R(z, r) + V_C(z, r) \quad (7.41)$$

with the following meanings:

- (i) $V_T(z + s, r)$ is the tube potential, as was derived earlier. Once the parameters s , a , and b are chosen, this function is determined uniquely.
- (ii) $V_A(z, r)$ is the part produced by aperture plates. The need to introduce this will become obvious in the next section. This potential can be determined by matching asymptotic boundary conditions and does not contain unknowns in the sense described later.
- (iii) $V_R(z, r)$ is the potential, originating from *surface charges* on the *real* anode surfaces; it is coupled with the last part V_C . There is no need to shift the rings away from the real surfaces, as was shown in Fig. 7.5. In fact, the “fast” method for solving Fredholm equations (Sections 6.3.4 to 6.3.6) is here much more favorable, and we write hence

$$V_R(z, r) = \sum_{n=1}^N q_n G_0(\mathbf{u}, \mathbf{u}_n) \quad (7.42)$$

with $\mathbf{u} := (z, r)$.

- (iv) $V_C(z, r)$ is the contribution from the *cathode* tip and its shank; it contains four additional unknowns $p_1 \cdots p_4$. The simplest way of setting up and solving the complete system of equations for *all* unknowns, is to *append* these new ones to $q_1, \dots, q_N$, hence

$$q_{N+j} := p_j, \quad j = 1, \dots, 4. \quad (7.43)$$

Similarly, the sampling points for the evaluation of Eq. (7.30) are appended to the vectors $\mathbf{u}_n (n = 1, \dots, N)$ defining the anode surfaces. The total length of this array is now $N + M$.

The system matrix has now the partitions

$$S = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \tilde{\mathbf{B}} & \mathbf{C} \end{pmatrix}, \quad (7.44)$$

with $\mathbf{A}$ having rank N and $\mathbf{C}$ rank 4. The submatrix $\mathbf{A}$ can be calculated simply by means of Eqs. (6.162):

$$A_{jn} = S_{jn} = \bar{G}_{j,n}^0(\mathbf{u}_j, \mathbf{u}_n), \quad (j, n = 1, \dots, N). \quad (7.45a)$$

The submatrix C , arising from Eq. (7.30), becomes, with an appropriate shift of the indices,

$$C_{jn} \equiv S_{j+N, n+N} = \sum_{k=1}^M w_k V_j(\mathbf{u}_{k+N}) V_n(\mathbf{u}_{k+N}). \quad (j, n = 1, \dots, 4) \quad (7.45b)$$

Lastly, the coupling partition $\tilde{B}$ is found to be

$$\begin{aligned} \tilde{B}_{jn} = S_{j+N, n} &= \sum_{k=1}^M w_k V_j(\mathbf{u}_{k+N}) G_0(\mathbf{u}_{k+N}, \mathbf{u}_n), \\ (j &= 1, \dots, 4; n = 1, \dots, N), \end{aligned} \quad (7.45c)$$

whereas the upper diagonal part is simply

$$B_{jn} = S_{j, n+N} = V_n(\mathbf{u}_j), \quad (j = 1, \dots, N; n = 1, \dots, 4). \quad (7.45d)$$

The vector P of inhomogeneities must now be extended similarly to a length of $N + M$. On the anode surfaces, the physical boundary potentials U_i are prescribed. Following Eq. (7.41), the contributions from V_T and V_A must be subtracted from these, giving

$$P_j = U_j - V_T(z_j + s, r_j) - V_A(\mathbf{u}_j), \quad (j = 1, \dots, N). \quad (7.46a)$$

In the extended part, however, the right-hand side of Eq. (7.30) must be evaluated, which leads to

$$\begin{aligned} \tilde{P}_j = P_{j+N} &= \sum_{k=1}^N w_k V_j(\mathbf{u}_{k+N}) (U_c - V_T(z_{k+N} + s, r_{k+N}) - V_A(\mathbf{u}_{k+N})), \\ (j &= 1, \dots, 4), \end{aligned} \quad (7.46b)$$

U_c being the prescribed cathode potential.

The total system of linear equations for the unknowns $q_1, \dots, q_{N+4}$ can, of course, be solved directly, but almost half of the memory requirements for the large system matrix S can be saved by exploiting the fact that the block matrices A and C are symmetric and positive definite; Cholesky decomposition is therefore permissible. It is then advisable to express the system of equations in the coupled form

$$\begin{aligned} A \cdot q &= P - B \cdot p \\ C \cdot p &= \tilde{P} - \tilde{B} \cdot q, \end{aligned} \quad (7.47)$$

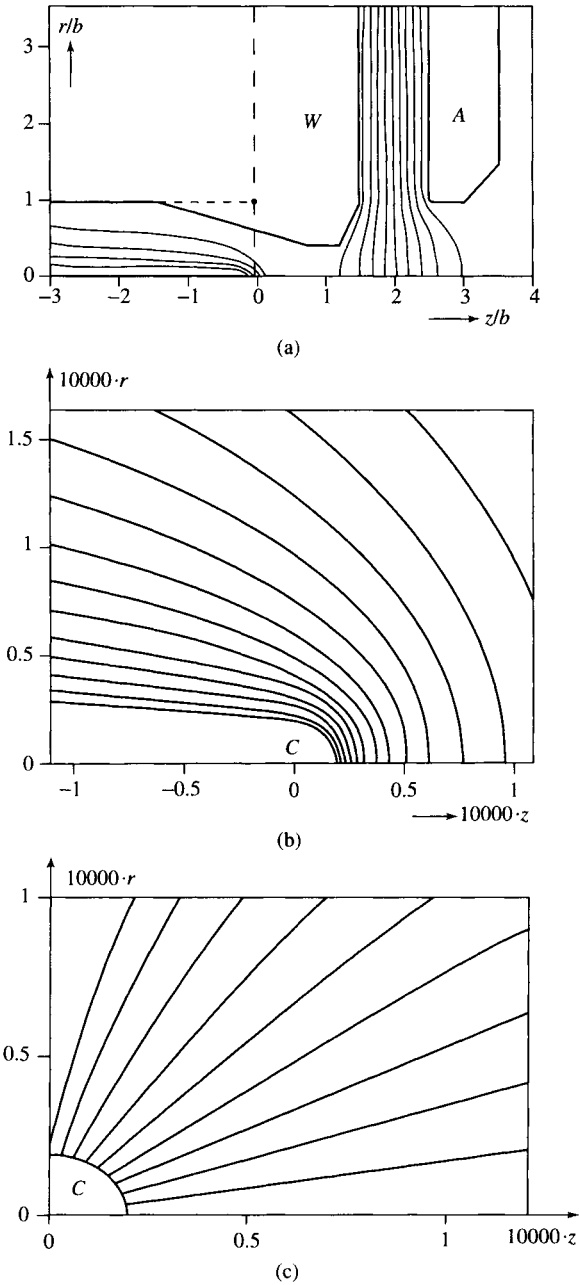


FIGURE 7.7 Equipotentials in the domain between the cathode and the first anode (W) of an electron gun: (a) global picture; (b) vicinity of the cathode; (c) some electron trajectories.

for the submatrices and vectors that are not appended and begin with unit label. This system of equations can easily be solved by an iterative combination of Cholesky algorithms in which the most time-consuming part—the decomposition of the matrices A and C —needs to be carried out only once at the beginning. The mutual Cholesky substitutions, starting from $\mathbf{p} = 0$, are fast matrix–vector multiplications and converge rapidly until the limit imposed by rounding errors is reached. Altogether, the necessary computational effort has increased only moderately, compared with a system without a cathode.

After solving the corresponding system of linear equations, Eq. (7.41) can be evaluated at any position (z, r) outside the cathode surface. The latter is now defined by the *equipotential* surface $V(\mathbf{u}) = U_c$. The deviation from the original one can be ignored in practice. Some results are shown in Figs. 7.7a,b and some trajectories starting from the cathode in Fig. 7.7c.

7.3.3 Charged Aperture Plates

All fields, produced by *finite* charge distributions, like rings and rods of finite length, satisfy the *natural* boundary conditions at infinity. The potential may be shifted at will by an additive constant, but this is unimportant. It is, however, impossible to built up an *acceleration step* by such elements, that means potential with

$$V(z \rightarrow -\infty, r) \neq V(z \rightarrow +\infty, r). \quad (7.48)$$

In principle, this is unnecessary and even unrealistic as such fields cannot be produced by any technical device. However, without them the simulation of accelerator systems would become very expensive, as then the domain of solution must be closed by locating charge elements on the *whole* surface, even on very remote parts that are unimportant for particle optics. The introduction of aperture plates serves the purpose of simulating the correct asymptotic conditions, so that the discretization can be confined to the essential parts of the boundary.

It is favorable to introduce prolate spheroidal coordinates (u, v, φ) , φ being the familiar azimuth. Because only rotationally symmetric functions are required here, the azimuth is not needed and we start from a transformation between (u, v) and cylindric coordinates (z, r) :

$$r = a\sqrt{(1+u^2)(1-v^2)} \quad 0 \leq v \leq 1, \quad (7.49a)$$

$$z = a u v \quad -\infty < u < \infty, \quad (7.49b)$$

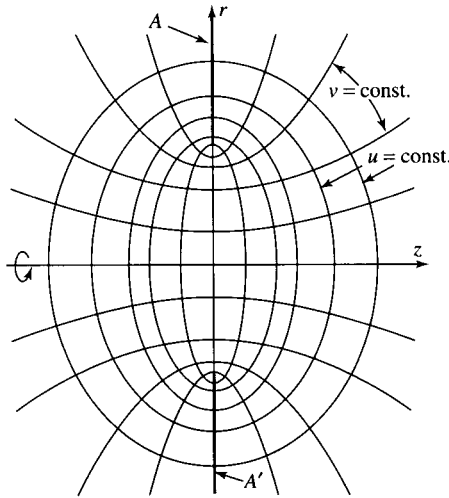


FIGURE 7.8 Oblate spheroidal coordinates (u, v) and their relation to cylindric coordinates (z, r) in a meridional plane. The position of the aperture plate ($v = 0$) is marked by A, A' .

a being the bore radius of the plate. This coordinate system is shown on Fig. 7.8; in the half meridional section the lines $u = \text{const.}$ are semiellipses, whereas the lines $v = \text{const.}$ are semihyperbolas. Among the latter the optic axis ($r = 0$) is the curve $v = 1$, whereas the charged plate becomes the line $v = 0$.

The inverse transformation, also required, leads to a system of biquadratic equations. Its solution is unique only with the choice of a suitable cut surface. Because the bore of the plate must belong to the regular domain, this surface is then the plate itself, $v = 0$. A suitable algorithm is now given by

$$\begin{aligned} s &:= (z^2 + r^2 - a^2)/2, \\ t &:= (s^2 + z^2 a^2)^{1/2}, \end{aligned} \quad (7.50a)$$

$$\begin{aligned} v &= a^{-1}(t - s)^{1/2}, \quad (s \leq 0) \\ v &= |z|(t + s)^{-1/2}, \quad (s > 0), \end{aligned} \quad (7.50b)$$

$$u = z/(av). \quad (7.50c)$$

On the charged plate, $v = 0$, the coordinate u becomes discontinuous according to

$$u = \pm(r^2/a^2 - 1)^{1/2}, \quad (v = 0, r \geq a). \quad (7.51)$$

The matrix T of derivatives, necessary for the calculation of Lamé coefficients, and for field calculation, is found to be

$$T = \begin{pmatrix} v_{1z} & u_{1z} \\ v_{1r} & u_{1r} \end{pmatrix} = \frac{1}{a(u^2 + v^2)} \begin{pmatrix} u(1 - v^2), & v(1 + u^2) \\ -rv/a, & ru/a \end{pmatrix}. \quad (7.52)$$

This matrix becomes singular for $u = v = 0$. From the general relation

$$u^2 + v^2 = 2t/a. \quad (7.53)$$

We see that this happens for $t = 0$, hence $s = 0$, and $r = a$; this is just the *bore-circle* of the plate.

According to the algorithm, given in Section 3.1.1, the Lamé-coefficients become here

$$L_u = a \left(\frac{u^2 + v^2}{1 + u^2} \right)^{1/2}, \quad L_v = a \left(\frac{u^2 + v^2}{1 - v^2} \right)^{1/2}, \quad L_\varphi = r, \quad (7.54)$$

and from Eq. (7.49a) the Jacobian is

$$J = L_u L_v L_\varphi = a^3(u^2 + v^2), \quad (7.55)$$

whereupon the Laplace equation according to Eq. (3.44) has the form

$$\begin{aligned} \Delta \Phi(u, v, \varphi) &= \frac{1}{a^2(u^2 + v^2)} \{ \partial_u[(1 + u^2)\Phi|_u] + \partial_v[(1 - v^2)\Phi|_v] \} \\ &+ r^{-2} \partial^2 \Phi / \partial \varphi^2 = 0. \end{aligned} \quad (7.56)$$

We are here only interested in solutions that are rotationally symmetric, that is independent of φ . These can be found by the familiar method of separation of variables

$$\Phi(u, v) = \sum_{m=0}^{\infty} C_m Q_m(u) \cdot P_m(v), \quad (7.57)$$

where the functions $P_n(v)$ turning out to be the Legendre polynomials. The remaining differential equation.

$$\frac{d}{du}((1 + u^2)Q'_m(u)) = m(m + 1)Q_m(u) \quad (7.58)$$

is solved by modified Legendre functions. These are obtained from Legendre functions of the second kind by introducing an imaginary argument; with

suitable amplitudes, the function values can be made real. The solutions of lowest orders $m = 0$ and 1 are given explicitly by

$$\begin{aligned} Q_0(u) &= \arctan u, & P_0(v) &= 1, \\ Q_1(u) &= u \arctan u + 1, & P_1(v) &= v, \end{aligned} \quad (7.59)$$

as can be easily verified. These are already sufficient for the purpose of this section as the functions of higher orders increase stronger than linearly at infinity and are hence unsuitable to represent an asymptotically homogeneous field. The function $Q_0(u)$ is here not appropriate as it is related to the potential of a charged disc filling the bore. Hence there remains only one solution with *three* free parameters F_1 , F_2 , and Φ_0 :

$$\Phi(u, v) = \Phi_0 + \frac{1}{2}(F_1 + F_2)auv + \frac{1}{\pi}(F_2 - F_1)av(1 + u \arctan u). \quad (7.60)$$

The first term Φ_0 remaining for $v = 0$, is the potential on the plate itself. With Eq. (7.49b) and $\bar{F} := (F_1 + F_2)/2$, the second term can be written as $\bar{F} \cdot z$ and represents hence a *homogeneous* field of strength $\bar{F}$. The third one is the essentially new result: for $z \rightarrow \pm\infty$ this behaves like $(F_2 - F_1)|z|/2$; altogether, the asymptotic field strengths are thus

$$\lim_{z \rightarrow -\infty} \Phi_{1z} = F_1, \quad \lim_{z \rightarrow +\infty} \Phi_{1z} = F_2. \quad (7.61)$$

This can be concluded easily from the general formulas

$$\Phi_{1z} = \frac{1}{2}(F_1 + F_2) + \frac{1}{\pi}(F_2 - F_1) \left(\arctan u + \frac{u}{u^2 + v^2} \right), \quad (7.62a)$$

$$\Phi_{1r} = \frac{1}{\pi a}(F_2 - F_1) \frac{rv}{(1 + u^2)(u^2 + v^2)}. \quad (7.62b)$$

The surface charge density $\sigma(r)$ is obtained from Eq. (7.62a) with $v = 0$:

$$\varepsilon_0 \sigma = \Phi_{1z}^{(-)} - \Phi_{1z}^{(+)} = \frac{2}{\pi}(F_2 - F_1) (\arctan |u| + |u|^{-1}), \quad (7.63)$$

and with Eq. (7.51) for $r \geq a$,

$$\varepsilon_0 \sigma = \frac{2}{\pi}(F_1 - F_2) \left(\arccos(a/r) + \frac{a}{\sqrt{r^2 - a^2}} \right). \quad (7.64a)$$

Asymptotically this becomes the constant $F_1 - F_2$, whereas near the bore it is singular like $|r - a|^{-1/2}$. The integrated surface charge $Q(r)$ becomes

$$\varepsilon_0 Q(r) = 2(F_1 - F_2) (r^2 \arccos(a/r) + a\sqrt{r^2 - a^2}). \quad (7.64b)$$

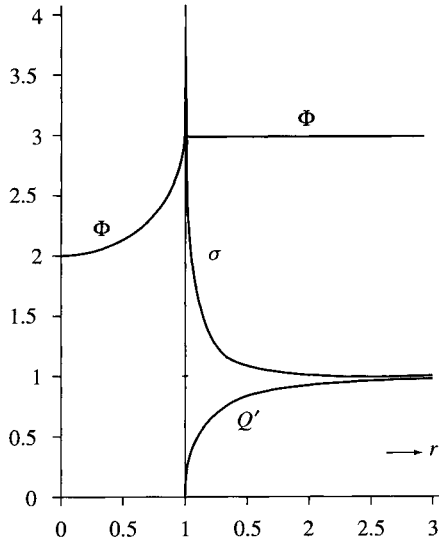


FIGURE 7.9 Normalized radial functions referring to the potential field of a charged aperture plate: Φ : potential in the plane of the plate; σ : surface charge density; Q' : integrated surface charge per area, according to Eq. (7.64).

The function $Q' = \varepsilon_0 Q / (\pi r^2)$ becomes asymptotically the same constant. These and the potential in the aperture plane:

$$\Phi(0, v) = \Phi_0 - av(F_1 - F_2)/\pi = \begin{cases} \Phi_0 & (r \geq a) \\ \Phi_0 - \frac{1}{\pi}(F_1 - F_2)\sqrt{a^2 - r^2} & (r \leq a) \end{cases} \quad (7.65)$$

are shown in Fig. 7.9.

The elliptic coordinates u and v can be related to the moduli ρ and $\bar{\rho}$, defined by Eqs. (6.77). The distances d_1 and d_2 are here defined by

$$d_{1,2} = (a^2 + z^2 + r^2 \mp 2ar)^{1/2}, \quad (7.66)$$

and with these the simple relations

$$\begin{aligned} \sqrt{1 + u^2} &= (d_2 + d_1) / 2a, \\ \sqrt{1 - v^2} &= (d_2 - d_1) / 2a = 2r / (d_1 + d_2), \\ u^2 + v^2 &= d_1 d_2 / a^2 \end{aligned} \quad (7.67)$$

can be derived, leading to

$$\rho = \sqrt{\frac{1-v^2}{1+u^2}}, \quad \bar{\rho} = \sqrt{\frac{u^2+v^2}{1+u^2}}. \quad (7.68)$$

Hence the potential of a *ring charge* of value Q , located just on the circular opening of the plate, is given by:

$$\Phi_R(u, v) = \frac{Q}{\varepsilon_0 \pi a \sqrt{1+u^2}} K(\rho). \quad (7.69)$$

7.3.4 Systems of Charged Aperture Plates

We now consider a system of N coaxial plates as shown in Fig. 7.10. Each plate is specified by the position of its aperture circle, (z_i, a_i) and its potential P_i ($i = 1, \dots, N$). The asymptotic field strengths for $r \rightarrow \infty$ in the gaps are then already well defined, and only the field strength F_0 for $z \rightarrow -\infty$ and F_N for $z \rightarrow +\infty$ can be chosen freely. The present choice of parameters [19] is different from that in ref. [15] but is equivalent to it.

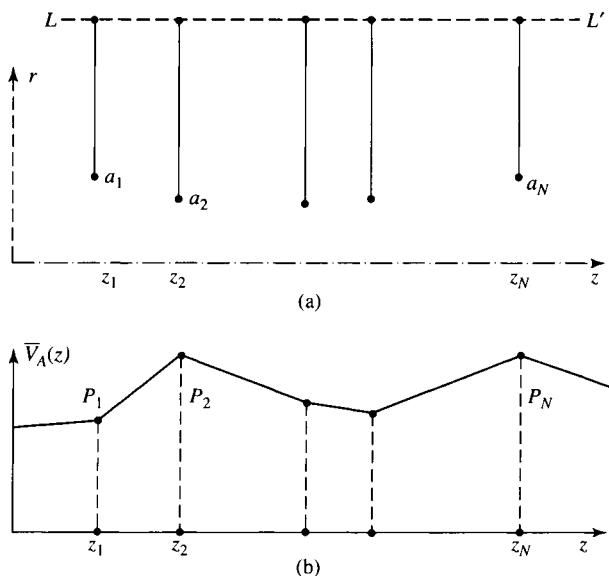


FIGURE 7.10 Example for a system of $N = 5$ coaxial charged aperture plates with associated asymptotic potential: (a) the positions of the plates in the upper half part of a meridional plane; (b) the asymptotic potential $\bar{V}_A(z)$ as a piecewise linear function; this is approximately reached along a line LL' at large off-axis distance.

It is now necessary to introduce elliptic coordinates (u_i, v_i) for each plate (i) separately. This can easily be done in a subprogram using a_i and $z - z_i$ instead of z . The total potential can then be written down as the superposition of N -shifted *symmetrical* contributions and a suitable homogeneous field:

$$V_A(z, r) = \sum_{i=1}^N C_i v_i (1 + u_i \arctan u_i) + \bar{V} + \bar{F}_z. \quad (7.70)$$

The constants C_i are determined by the differences of field strengths F_i . F_0 and F_N are to be defined explicitly, whereas the other values are obtained from

$$F_i = (P_{i+1} - P_i)/(z_{i+1} - z_i), \quad (i = 1, \dots, N-1), \quad (7.71)$$

whereupon we obtain the result

$$C_i = \frac{1}{\pi} a_i (F_i - F_{i-1}), \quad (i = 1, \dots, N) \quad (7.72)$$

according to the form of the third term in Eq. (7.60). For $r \gg a_i$ the approximations

$$\arctan u_i \rightarrow \frac{\pi}{2} \text{sign}(z - z_i), \quad v_i \ll 1 \quad (7.73)$$

can be made, hence

$$\frac{1}{\pi} a_i v_i (1 + u_i \arctan u_i) \rightarrow \frac{1}{2} |z - z_i| \quad (7.74)$$

and thus

$$\bar{V}_A(z) = V_A(z, \infty) = \bar{V} + \bar{F}_z + \frac{1}{2} \sum_{i=1}^N (F_i - F_{i-1}) |z - z_i|. \quad (7.75)$$

This is a piecewise linear function, see Fig. 7.10b

For all positions apart from the discontinuities, the differentiation with respect to z results in

$$\bar{V}'_A(z) = \bar{F} + \frac{1}{2} \sum_{i=1}^N (F_i - F_{i-1}) \text{sign}(z - z_i), \quad (7.76)$$

which can be rewritten as

$$\bar{V}'_A = \begin{cases} \bar{F} + (F_0 - F_N)/2 & z < z_1 \\ \bar{F} + F_k - (F_0 + F_N)/2 & z_k < z < z_{k+1} \leq z_N \\ \bar{F} + (F_N - F_0)/2 & z > z_N \end{cases} \quad (7.77)$$

From this it is obvious that all field strengths F_i are simultaneously correct with

$$\bar{F} = (F_0 + F_N)/2. \quad (7.78)$$

Finally, the constant $\bar{V}$ is determined by matching one of the plate potentials to the prescribed value. An elementary calculation results in

$$\bar{V} = (P_1 + P_N - F_0 z_1 - F_N z_N)/2, \quad (7.79)$$

whereupon the configuration is completely determined. Most frequently, the asymptotic half spaces are *field-free*. This is comprised in the above given formulas as the special case

$$F_0 = F_N = 0, \quad \bar{F} = 0, \quad \bar{V} = (P_1 + P_N)/2. \quad (7.80)$$

With suitable labelling, the single plate, $N = 1$, is also a special use, as it must be. The expressions of Eq. (7.71) now cannot be evaluated and are hence to be omitted.

Any function of the form of Eq. (7.70) is an exact solution of Laplace equation (apart from positions on the plates themselves) but does *not* satisfy a Dirichlet condition. This is demonstrated in Fig. 7.11, showing an attempt to simulate a triode-electron source with flat cathode and Wehnelt-electrode at negative potential. Such systems are used, for instance, as electron emission microscopes for analyzing metal surfaces [20]. The details of this application are here unimportant; we are only concerned with the task of field calculation.

An equipotential surface $V(0, r) = 0$, as in Fig. 7.11 is exactly simulated by imposing the antisymmetry condition $V(-z, r) = V(z, r)$ and using only the positive half part. As is obvious from Fig. 7.11, the shape of the Wehnelt electrode, which is very important, can be simulated only inaccurately, and many trials by variation of parameters are necessary to match one of the equipotential surfaces to the prescribed form as closely as possible.

This tedious procedure can be avoided by using additional ring charges on the surfaces, as discussed in Section 6.3. These need to cover only areas of limited extend in the radial direction ($r \leq 2$ in Fig. 7.12) as the partial potential V_A already satisfies all the asymptotic conditions. The Dirichlet problem

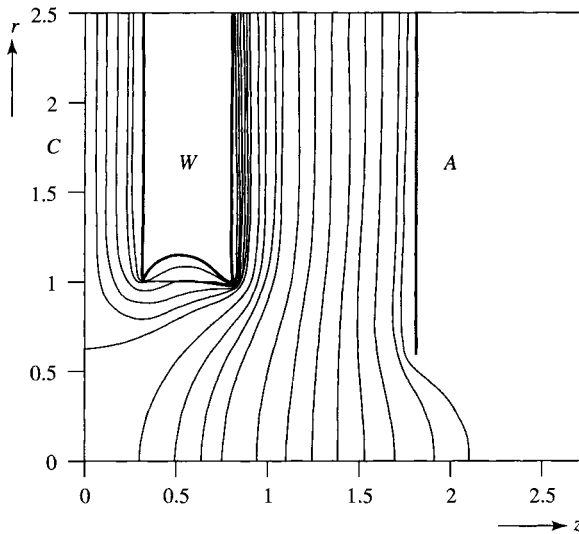


FIGURE 7.11 Attempt to simulate the electrostatic field in an electron gun, consisting of a flat cathode C , a Wehnelt electrode W , and a thin anode A , having the potentials 0, -1 , and 10 units, respectively. Here only six aperture plates are chosen, of which the three in the region $z > 0$ are shown. The shape of the Wehnelt electrode is wrong.

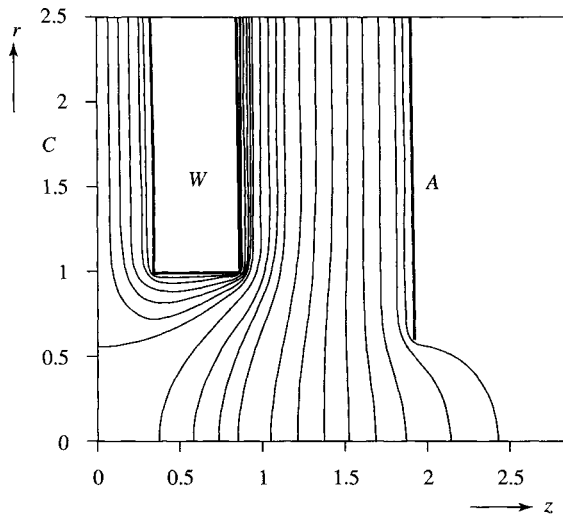


FIGURE 7.12 Recalculation of the configuration of Fig. 7.11 by additional consideration of surface-charge rings covering the Wehnelt and the anode up to $r = 2$. Now, the surface of the Wehnelt is fairly accurately an equipotential.

to be solved for the ring charges, consists in the condition, that the derivation $V_B(z, r) - V_A(z, r)$ from the prescribed boundary potentials $V_B(z, r)$ must vanish. The rings should be located in such a way that the singularity of the plate-charge density does not fall directly on one ring, but in the gap between two adjacent ones. If the edges are slightly rounded off, it is also possible to locate the aperture circles in the original edges, as is done in Figs. 7.12 and 7.13a. Moreover, the term $u^2 + v^2$ in the denominators in Eqs. (7.62) must be greater than a small threshold value, for example $u^2 + v^2 \geq 10^{-4}$. Then the inevitable errors, caused by the singularities, remain confined to a very small vicinity of the corresponding edge.

The result of this hybrid technique is demonstrated in Fig. 7.12. The deficiencies of the approximation in Fig. 7.11 have clearly been remedied. The most important correction is the decrease of the "zero Volt radius," the radius of the off-axis saddle point—from 6.5 units to 5.5 units.

The absolute necessity to use superpositions of ring fields and plate fields in *open* structures is demonstrated in Figs. 7.13a and 7.13b, where an attempt

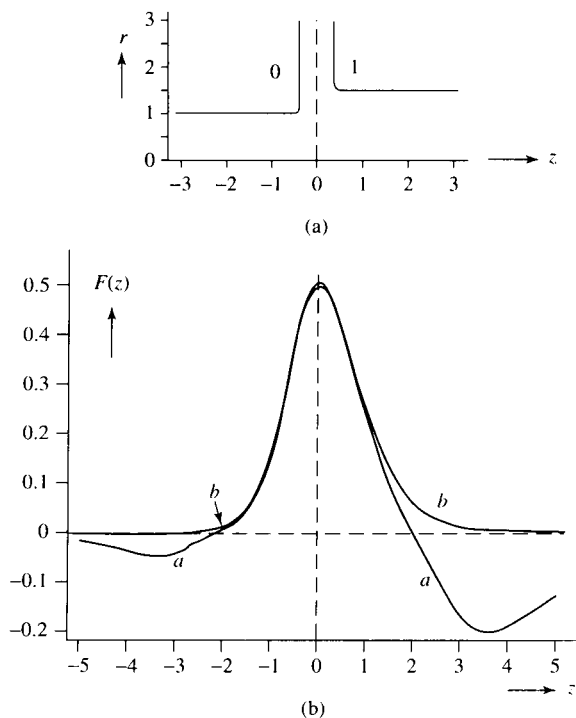


FIGURE 7.13 Field calculation in an open lens with simple structure of the pole pieces: (a) half axial section with normalized potential step; (b) axial field strength $F(z)$: *a*: for an open structure; *b*: for the same structure with aperture plates on the plane faces of the gap.

is made to simulate a simple round lens with two pole pieces or electrodes analogous to the configuration shown in Fig. 2.4.

In such a case the axial field strength $F(z)$ must not change sign. Curve *a* in Fig. 7.13b shows that this requirement is strongly violated near the open ends of the cylindrical tubes. The effect is more pronounced on the right-hand side and is an inevitable consequence of the fact that the potential due to any configuration of charged rings satisfies the natural boundary condition at infinity. This error can be eliminated in two ways:

- (i) closure of the boundary sufficiently far from the midplane;
- (ii) open boundary, but additional use of two charges plates, held at the corresponding potentials.

In the present example these plate circles can be located at the positions $(-0.4, 1)$ and $(0.4, 1.5)$, with potentials 0 and 1, respectively. The result is the curve (*b*) for $F(z)$ in Fig. 7.13b, which satisfies perfectly all the requirements. This alternative has the additional advantages that the cylinders, carrying ring charges, may be shorter and that the transition to the field-free space is an analytical function.

In view of these results, it is clearly necessary to incorporate charged plates in the model for pointed-filament guns, as was already mentioned in Section 7.3.2; these can be located directly on the plane anode surfaces.

7.4 THE CURRENT SIMULATION MODEL

The correct method for the field calculation in magnetic deflection systems has been worked out in Sections 6.2 to 6.5 and most specifically in Sections 6.5.4 and 6.5.5. The practical evaluation of these formulas requires careful and lengthy calculations, especially if the current-conducting windings are very close to the shielding yoke. If the relative permeability of the latter is very large, $\mu \geq 10^3$, then the following method may be advantageous.

This method is based on the assumption that the $\mathbf{H}$ -field in the ferromagnetic yoke is very weak, so that the boundary conditions on its surface can be simplified. As this assumption is not self-evident, we study first some simple cases in which an accurate solution can be obtained analytically.

7.4.1 Magnetic Mirror Properties

As in electrostatics, the case of an infinitely extended plane mirror can be solved exactly, as follows.

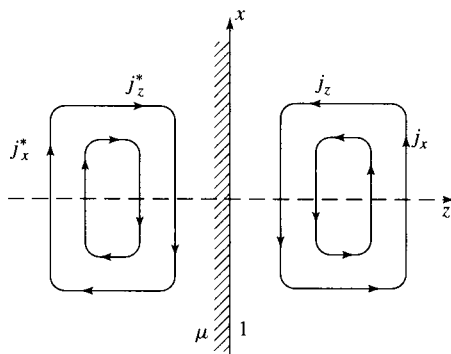


FIGURE 7.14 A system of current-conducting coils and its magnetic "mirror-image"; the components of $\mathbf{j}$ parallel to the surface keep their direction, whereas the normal component changes direction.

Let the coordinate plane $z = 0$ be the surface of a homogeneous ferromagnetic material; the *relative* permeability is thus unity for $z > 0$ (vacuum) and $\mu > 1$ for $z \leq 0$ (see Fig. 7.14).

The current density $\mathbf{j}(\mathbf{r})$ for $z > 0$ is regarded as known and defined by a system of current conducting wires. At infinity the natural boundary conditions hold and on the surface $z = 0$ the familiar electromagnetic ones. Then this configuration is specified uniquely.

To find the solution, we introduce a mirror current density $\mathbf{j}^*(\mathbf{r})$ thus:

$$\left. \begin{aligned} j_x^*(x, y, -z) &= j_x(x, y, z) \\ j_y^*(x, y, -z) &= j_y(x, y, z) \\ j_z^*(x, y, -z) &= -j_z(x, y, z) \end{aligned} \right\} \quad (z > 0), \quad (7.81)$$

so that $\mathbf{j}(\mathbf{r})$ is nonzero only for $z > 0$ and $\mathbf{j}^*(\mathbf{r})$ only for $z < 0$. By means of these, two kinds of vector Coulomb integrals can be written down:

$$\mathbf{A}_1(\mathbf{r}) = \mu_0 \int_{(z' > 0)} \frac{\mathbf{j}(\mathbf{r}') d^3 r'}{4\pi |\mathbf{r} - \mathbf{r}'|}, \quad (7.82a)$$

$$\mathbf{A}_2(\mathbf{r}) = \mu_0 \int_{(z' < 0)} \frac{\mathbf{j}^*(\mathbf{r}') d^3 r'}{4\pi |\mathbf{r} - \mathbf{r}'|}. \quad (7.82b)$$

They can be evaluated at any position $\mathbf{r}$ and refer formally to a medium with unit relative permeability.

The total solution is now built up by forming suitable linear combinations of these integrals. The resulting vector potential must have the source $\mathbf{j}(\mathbf{r})$ for

$z > 0$ and *none* for $z < 0$. This implies the form

$$\begin{aligned} \mathbf{A}(\mathbf{r}) &= C\mathbf{A}_2(\mathbf{r}) + \mathbf{A}_1(\mathbf{r}) \quad (z \geq 0), \\ \mathbf{A}(\mathbf{r}) &= C^*\mathbf{A}_1(\mathbf{r}) \quad (z \leq 0). \end{aligned} \quad (7.83)$$

The coefficients are determined by the boundary conditions in the plane $z = 0$. By means of symmetry considerations it turns out that $A_{1x}(x, y, 0) = A_{2x}(x, y, 0)$, with analogous relations for A_y and all derivatives with respect to x and y . Then the continuity of A_x , A_y , and B_z requires that

$$C^* = C + 1. \quad (7.84)$$

With $A_{1z|z} = A_{2z|z}$ the continuity of $\text{div} \mathbf{A}$ is also ensured. The remaining derivatives have an *antisymmetric* property, for example $A_{1x|z} = -A_{2x|z}$, $A_{1z|x} = -A_{2z|x}$, and so on.

The continuity of the tangential components H_x and H_y of the vector $\mu_0 \mathbf{H} = \mu^{-1} \nabla \times \mathbf{A}$ requires that

$$\mu^{-1} C^* = 1 - C. \quad (7.85)$$

The solution of these linear equations is given by

$$C = \frac{\mu - 1}{\mu + 1}, \quad C^* = \frac{2\mu}{\mu + 1}. \quad (7.86)$$

The component A_z itself is then not continuous, but this is not necessary here; the above-defined boundary value problem is hence now solved *exactly*.

7.4.2 Local Properties

This simple method may sometimes be helpful with respect to field calculations in systems that are shielded on one side by a very thick ferromagnetic wall, but apart from this special situation there are unfortunately very few such simple analytically soluble cases. We therefore try to cast the relations obtained in such a form that they can serve as a *general* model for the *local* behavior of the magnetic field in the vicinity of a ferromagnetic wall.

In the following considerations, the vicinity of sharp edges is excluded. It is always possible to choose a sufficiently small domain on the surface of a ferromagnetic yoke, in which this part of the surface can be replaced by its tangential plane, and the cartesian coordinate system then be adapted to the latter. Such an approximation makes sense only if the current-conducting

wires are located quite near to the surface. This implies that $j_z(x, y, z) \equiv 0$ and consequently $A_z \equiv 0$.

It is now advantageous to introduce a surface current density $\mathbf{J}(\mathbf{r})$ by writing

$$\mathbf{J}(x, y) := \int_0^\infty \mathbf{j}(x, y, z) dz, \quad (7.87)$$

with $\mathbf{J}^* = \mathbf{J}$ owing to the symmetry. Equations (7.83) with (7.86) can now be cast in a more concise form:

$$\mathbf{A}(\mathbf{r}) = \frac{\mu\mu_0}{2\pi(\mu+1)} \iint_{z'=0} \frac{\mathbf{J}(x', y') dx' dy'}{|\mathbf{r} - \mathbf{r}'|} + \mathbf{A}_0(\mathbf{r}), \quad (7.88)$$

which is valid on both sides of the surface. The term $\mathbf{A}_0(\mathbf{r})$ results from more distant parts of the surface and has no sources in the domain in question. In a local approximation such contributions cannot be excluded.

By differentiation under the integral for $|z| > 0$ and subsequent by proceeding to the limit $|z| \rightarrow 0$, it can be shown that the component B_z remains continuous, as it must. The *tangential* components H_x and H_y are related to the vector $\mathbf{J}$, as was already stated in Eq. (1.84). Beyond this general relation we obtain here more specifically:

$$\begin{aligned} H_x^{(+)} &= H_x^{(0)} - \frac{\mu}{\mu+1} J_y, & H_x^{(-)} &= H_x^{(0)} + \frac{1}{\mu+1} J_y, \\ H_y^{(+)} &= H_y^{(0)} + \frac{\mu}{\mu+1} J_x, & H_y^{(-)} &= H_y^{(0)} - \frac{1}{\mu+1} J_x, \end{aligned} \quad (7.89)$$

in which the positive label refers to the vacuum side, and the terms with label (0) contain *all* nonlocal continuous contributions.

Because the latter are not yet known, these boundary relations are not yet useful for practical purpose. A reasonably simple and integrable relation is to be expected only for $\mu \rightarrow \infty$, say $\mu \geq 10^3$. It is then often permissible to assume that $|\mathbf{H}_0| \sim \mu^{-1}$, so that this field can be ignored, Eqs. (7.89) now simplifying to

$$\mathbf{H}_v(\mathbf{r}) \times \mathbf{n}(\mathbf{r}) = \mathbf{J}(\mathbf{r}). \quad (7.90)$$

This simple law can now be generalized for any smoothly bent yoke surface. Then the vector $\mathbf{n}$, (being $-\mathbf{e}_z$ in the special case) is directed into the ferromagnetic material, and the label v indicates that the field $\mathbf{H}(\mathbf{r})$ is to be evaluated on the vacuum side, where it is considerable. This simplification is not at all self-evident, and we shall therefore study a simplified model of toroidal and saddle coils, from which it becomes obvious that $|\mathbf{H}_0| \sim \mu^{-1}$ makes sense.

In contrast to this very strong simplification, the error introduced by ignoring the finite distance d of the windings from the surface is negligible. If D denotes the distance of the reference point from the surface, then this error decreases at least as $(d/D)^2$. In the case of curved surfaces, this error is minimized if the mirror currents are located according to the law of reciprocal radii:

$$(R - d)(R + d^*) = R^2, \quad (7.91)$$

R being the radius of curvature.

7.4.3 A Simple Model for Cylindrical Coils

The field calculation in systems of toroidal or saddle coils of finite extent in the axial direction is so complicated that it can be performed only by numerical tools. If, however, this extent is so large that the influence of the front faces on the field in the middle plane can in practice be ignored, then a fairly simple analytical solution can be obtained. This case is certainly unrealistic with respect to technical applications, but the results offer a way of studying the screening effect of the yoke as a function of the various parameters of the model.

Let us consider an infinity long cylindrical tube with radii r_1 and $r_2 > r_1$; the cross section $z = \text{const.}$ is shown in Fig. 7.15. The interior $r < r_1$ and the exterior $r > r_2$ may be vacuum, whereas the wall ($r_1 < r < r_2$) has the relative permeability $\mu \neq 1$.

The inner and outer surfaces may both carry a surface current flowing in the longitudinal direction. The corresponding densities, being functions of the azimuth φ , can be approximated by Fourier-series expansions. Because the

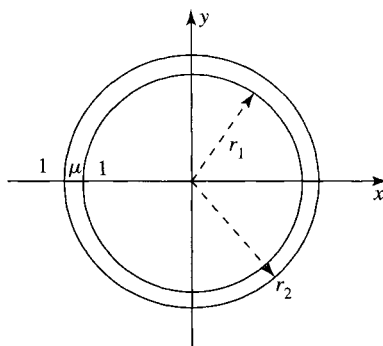


FIGURE 7.15 Cross section through a ferromagnetic cylindric wall. (The current distribution on its surfaces is not shown here.)

fields, produced by the different components of these, superimpose independently, it is sufficient to consider then separately in turn. Moreover, there is no loss of generality in assuming only cosine terms as the coordinate system can be rotated accordingly. Hence we now assume that the surface current densities are given by

$$\begin{aligned} J_1(\varphi) &= r_1^{-1} C_1 \cos m\varphi \quad (r = r_1), \\ J_2(\varphi) &= r_2^{-1} C_2 \cos m\varphi \quad (r = r_2), \end{aligned} \quad (7.92)$$

with *positive* integer orders m . Because there are no spatial sources, it must be possible to introduce *scalar* potentials $U_n(r, \varphi)$ such that $\mathbf{H} = -\text{grad } U_n$ is piecewise valid for $r \neq r_1$ and $r \neq r_2$. These potentials can favorably be written as

$$U_1(r, \varphi) = \frac{a_1}{m} \left(\frac{r}{r_1} \right)^m \sin m\varphi, \quad (0 \leq r < r_1), \quad (7.93a)$$

$$U_2(r, \varphi) = \left[\frac{a_2}{m} \left(\frac{r}{r_1} \right)^m - \frac{b_2}{m} \left(\frac{r_1}{r} \right)^m \right] \sin m\varphi, \quad (r_1 < r < r_2), \quad (7.93b)$$

$$U_3(r, \varphi) = \frac{-b_3}{m} \left(\frac{r_2}{r} \right)^m \sin m\varphi, \quad (r > r_2). \quad (7.93c)$$

Strictly speaking, all amplitudes should have an additional label m , but for reasons of conciseness this is dropped here. These amplitudes are determined uniquely by the boundary conditions, namely, that $B_r = \mu H_r$ must remain continuous, and that the difference between the tangential components H_φ must be equal to the corresponding surface current density J . These conditions can be cast in a concise form; with $\rho := r_1/r_2 < 1$:

$$\begin{aligned} \mu^{-1} a_1 &= a_2 + b_2, \quad \mu^{-1} b_3 = \rho^{-m} a_2 + \rho^m b_2, \\ a_1 &= a_2 - b_2 + C_1, \quad b_3 = -\rho^{-m} a_2 + \rho^m b_2 - C_2. \end{aligned} \quad (7.94)$$

The solution of this system of linear equations is given by

$$\begin{aligned} a_1 &= \mu D^{-1} \{ [\mu + 1 - (\mu - 1)\rho^{2m}] C_1 - 2\rho^m C_2 \} \\ a_2 &= -D^{-1} \{ (\mu - 1)\rho^{2m} C_1 + (\mu + 1)\rho^m C_2 \} \\ b_2 &= D^{-1} \{ (\mu + 1) C_1 + (\mu - 1)\rho^m C_2 \} \\ b_3 &= \mu D^{-1} \{ 2\rho^m C_1 - [\mu + 1 - (\mu - 1)\rho^{2m}] C_2 \} \end{aligned} \quad (7.95)$$

with the common denominator

$$D = (\mu + 1)^2 - \rho^{2m}(\mu - 1)^2 > 0. \quad (7.96)$$

These formulas are correct for all values of the model parameters. Saddle coils are obtained with $C_2 = 0$ and toroidal coils with $C_2 = C_1$. There are two simple special cases:

- (i) Tube with *large outer radius* : $\rho^m \ll 1$

$$a_1 = \frac{\mu C_1}{\mu + 1}, \quad b_2 = \frac{C_1}{\mu + 1} = a_1/\mu, \quad (7.97)$$

whereas the other coefficients can be ignored. These are the limits of shielding in a system of saddle coils at high multipole orders m .

- (ii) Tube wall with high permeability: $\mu \gg 1$

It then becomes obvious that $|a_2|$ and $|b_2|$ become very small like μ^{-1} , as $D \sim \mu^2$. The remaining coefficients become now

$$a_1 = C_1, \quad b_3 = -C_2, \quad (7.98)$$

as a factor $1 - \rho^{2m}$ cancels out from the numerators and the denominator. These are just the results that would be obtained by evaluation of Eq. (7.90), which is thus justified for these kinds of coil systems.

The calculation also demonstrates that it is permissible to use local scalar potentials; however, these are *not* continuous at the yoke surfaces; the latter are here necessarily cut surfaces with respect to the definition.

7.4.4 Generalization of the Method

We now assume that Eq. (7.90) is valid on the vacuum side of a ferromagnetic yoke that implies $\mu \gg 1$. As was already mentioned above, the surface must be considered to be a discontinuity of the scalar potential, and the function $U(\mathbf{r})$ introduced below always refers to the vacuum domain on introducing $\mathbf{H} = -\text{grad } U$ into Eq. (7.90), and after a second vector multiplication with $\mathbf{n}$ this can be rewritten as

$$\nabla U_t := (\mathbf{n} \times \nabla U) \times \mathbf{n} \equiv \nabla U - \mathbf{n} \cdot \nabla U \mathbf{n} = \mathbf{J} \times \mathbf{n}, \quad (7.99)$$

which means that the *tangential* component ∇U_t on the surface S is prescribed.

There arises now the problem of the integrability of this boundary condition. For arbitrary vector functions $\mathbf{J}(\mathbf{r})$ this is certainly not valid, but we shall see that the law of continuity is necessary and sufficient in practice.

As in Section 6.5.4, the boundary S , being rotationally symmetric, will be parameterized by the azimuth φ , and the arc length s along a meridional cut, but in contrast to the former considerations this surface S , containing currents, will be identical with the yoke surface, as is necessary here.

We recall now Eq. (6.243):

$$\mathbf{J}(s, \varphi) = J_t \boldsymbol{\tau} + J_\varphi \mathbf{e}_\varphi \quad (7.100)$$

and write down $\text{grad } U$ in this coordinate system. According to the formulas in Section 3.1.5 this becomes

$$\nabla U = \boldsymbol{\tau} \frac{\partial U}{\partial s} + \frac{1}{r} \mathbf{e}_\varphi \frac{\partial U}{\partial \varphi} + \mathbf{n} \cdot \nabla U \mathbf{n}. \quad (7.101)$$

After introducing these Equations into Eq. (7.99) and considering the orthonormality of the basic vectors, we arrive at

$$\partial \bar{U}(s, \varphi) / \partial \varphi = r(s) J_t(s, \varphi), \quad (7.102a)$$

$$\partial \bar{U}(s, \varphi) / \partial s = -J_\varphi(s, \varphi), \quad (7.102b)$$

in which the function $\bar{U}(s, \varphi)$ denotes the boundary values of the potential $U(\mathbf{r})$ on S . A necessary condition of integrability is obviously the uniqueness of the mixed derivative of second order:

$$\frac{\partial^2 \bar{U}}{\partial s \partial \varphi} = \frac{\partial}{\partial s} (r J_t) = -\frac{\partial}{\partial \varphi} J_\varphi, \quad (7.102c)$$

which is the law of continuity according to Eq. (6.245). This is also sufficient, as then Eqs. (7.102) can be integrated by means of Fourier-series expansions.

In fact, if we substitute Eqs. (6.246) into the right-hand side of Eq. (7.102a) and integrate over the azimuth φ , it becomes obvious that the Fourier-series expansion of $\bar{U}$ must have the form

$$\bar{U}(s, \varphi) = \sum_m \bar{U}_m(s) \sin m(\varphi - \alpha_0), \quad (7.103)$$

which is then compatible with Eq. (7.102b). By comparison of coefficients for all orders $m \neq 0$ the relations

$$\bar{U}_m(s) = m^{-1} w_m(s), \quad (7.104a)$$

$$\bar{U}'_m(s) = a_m(s) / (m r(s)) \quad (7.104b)$$

are obtained which are compatible, thanks to Eq. (6.246c).

The ultimate simplification is reached in systems of toroidal coils. With $a_m(s) \equiv 0$ we obtain $\bar{U}_m = w_m/m = \text{const}$. This leads to a considerable simplification, as the highly complicated field calculation problem that would be

solved with rigorous methods, is now reduced to a sequence of uncoupled Dirichlet problems with constant boundary potentials. Though the Fourier analysis is quite favorable, Eq. (7.102a) can also be integrated without its use, and results in

$$\bar{U}(\varphi) = \int_0^\varphi r(s) J_t(s, \varphi') d\varphi' \equiv \int_0^\varphi w(\varphi') d\varphi', \quad (7.105)$$

which is Schwertfeger's formula [21]. Moreover, Eqs. (7.98) become now a special case with $\alpha_0 = 0$, $a_1 = C_1 = w_m$, and $b_3 = -C_2 = -w_m$. It should be mentioned that Eqs. (7.104) and (7.105) hold also for horn-shaped yokes, as in Fig. 6.21.

7.4.5 Comparison with Correct Calculations

The validity of the earlier-introduced simplifications can be examined by comparison of the field with that given by more rigorous computations. Such calculations have been carried out by Scherle [22], who used a collocation technique with splines (Section 6.3.2) in the context of the approximations (6.260). The geometric form of the yoke is shown in Fig. 6.17 with cylinder length 2 cm and radii 1 cm and 1.5 cm. The results for the normalized transversal axial field strengths are shown in Fig. 7.16. Scherle confirmed that for $\mu \rightarrow \infty$, these results agree very accurately with those of the simplified approximation so that the latter is then certainly justified.

These calculations are also useful for studying the limitations of the current simulation method. As is to be expected, the transversal field strength on the optic axis decreases with decreasing permeability μ . Although Eqs. (7.95) refer to an infinitely long system of coils, a comparison of the results of Fig. 7.16 with those of Eqs. (7.95) reveals only modest deviations. In the notation of Eqs. (7.95) the relative axial field strength is the ratio a_1/C_1 that becomes unity for $\mu \rightarrow \infty$. With $m = 1$ and $\rho = r_1/r_2 = 2/3$, we obtain in the worst case

$$\begin{aligned} a_1/C_1 &= 1/6 \quad \text{for } \mu = 1 \text{ and } C_2 = C_1, \\ a_1/C_1 &= 1/2 \quad \text{for } \mu = 1 \text{ and } C_2 = 0. \end{aligned} \quad (7.106)$$

These are in fair agreement with the ratios that can be determined from the graphs; the conclusion of Eq. (7.98) obtained from Eqs. (7.95) for $\mu \rightarrow \infty$, is hence justified in practice.

The bell-shaped curves for the transversal axial field strength with $\mu \rightarrow \infty$ are quite similar, but do not have exactly the same form. In Fig. 7.17 the curves for a saddle coil and a toroidal coil having the same geometrical parameters,

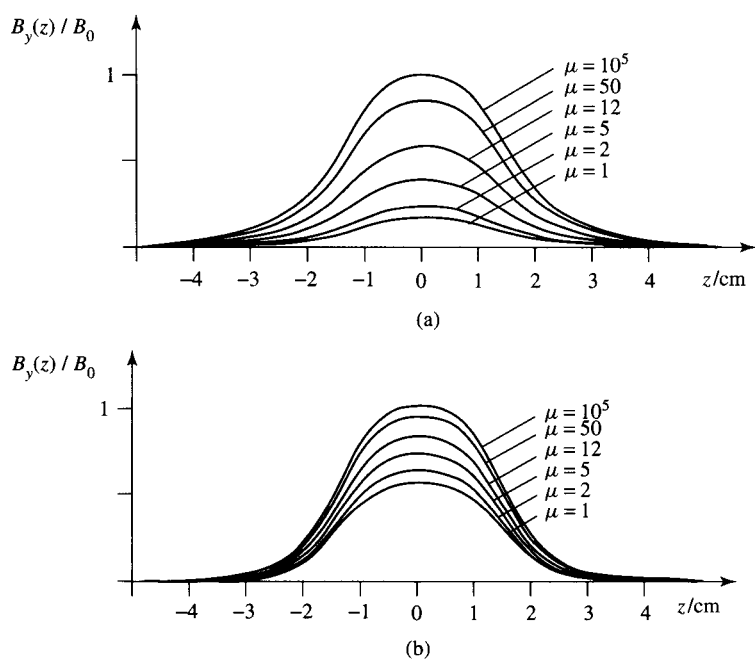


FIGURE 7.16 Normalized transversal axial field strength: (a) toroidal coil; (b) saddle coil (after W. Scherle [22]).

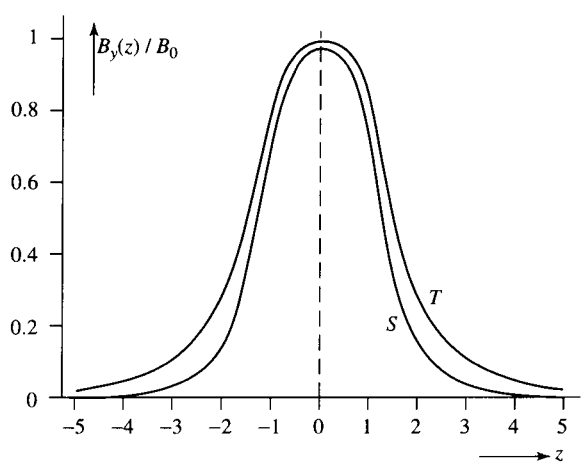


FIGURE 7.17 Normalized transversal axial field strength for saddle coils (S) and toroidal coils (T), with equal geometry and equal current distribution on the inner surface; B_0 is the maximum of the function $B_y(z)$ in the toroidal coil system for $\mu \rightarrow \infty$.

and the same distribution of windings on the inner surface, are plotted together for comparison. The larger width for the toroidal coil results from the fact that the windings on the outer surface part also contribute to the axial field, as is to be expected.

7.5 THE GENERAL ALTERNATION METHOD

Quite often the solution of a Dirichlet problem in the interior of a domain D with complicated structure requires a very large effort, if we attempt to find it in one step; conversely, it can be calculated fairly easily if this domain is built up from partly overlapping subdomains with surfaces of simple form. An example is shown in Fig. 7.18. The domain D consists of the union of a rectangular box and an intersecting sphere, and the boundary values of a potential $V(r)$ may be prescribed on the outer shell of this domain. There are efficient techniques of field calculation in the pure box and in the pure sphere, but not in the union of them, owing to the very different geometrical symmetries.

This problem can be solved if there is a systematic method of determining the boundary values on the inner surfaces, which are needed to complete the geometrical symmetry, and this is just the task of the general alternation method (GAM). This method was introduced as long ago as 1870 by H. A. Schwarz for purposes of mathematical proofs, but gained some practical importance with the development of powerful computers; see, for example, the work of Schaefer [23,24].

7.5.1 Formulation of the Method

The GAM can be applied to a system of several partly overlapping domains, but for reasons of conciseness, we shall consider here only two of them, as shown in Fig. 7.19.

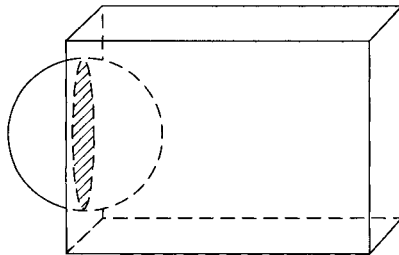


FIGURE 7.18 Union of a rectangular box and an intersecting sphere.

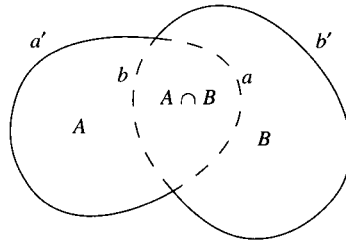


FIGURE 7.19 Illustration of the notation used in the theory of the general alternating method.

Let $D = A \cup B$ be the union of two domains A and B and $a' \cup b'$ its *outer* boundary, whereas a and b are the inner boundaries of the *overlapping* $A \cap B$. We assume that in D a linear PDE of the general form

$$LV(\mathbf{r}) = S(\mathbf{r}) \quad (7.107)$$

is to be solved with prescribed boundary values

$$V(\mathbf{r}) = \bar{V}(\mathbf{r}), \quad \mathbf{r} \in a' \cup b' \quad (7.108)$$

and that this Dirichlet problem has a unique finite solution. This is certainly true if Eq. (7.107) is some form of Poisson's equation, and the source function $S(\mathbf{r})$ is finite.

The GAM consists now in the following iteration technique:

Start

Choose some reasonable, not too crude guess, for the boundary values $V_0(\mathbf{r}), \mathbf{r} \in a$.

Alternation

For $n = 0, 1, 2, \dots$

- (1) Solve the Dirichlet problem for $V_n(\mathbf{r})$ in A .
- (2) Determine the boundary values $V'_n(\mathbf{r})$ on b by means of interpolation.
- (3) Solve the Dirichlet problem for $V'_n(\mathbf{r})$ in B .
- (4) Determine the boundary values $V_n(\mathbf{r})$ on a by means of interpolation.
- (5) Determine the maximum absolute difference δ_n of boundary values on $a \cup b$.
- (6) Stop for $\delta_n < \varepsilon$, else continue, (ε being a prescribed error threshold).

This procedure converges with the speed of a geometric sequence, at least in the case of Poisson's equation with invariant source function $S(\mathbf{r})$. In the form presented earlier it is easily understandable, but not yet very efficient, as the

whole information about the field must be available at every step. This can be avoided by means of the following modifications: owing to the linearity of Eq. (7.107) we can introduce series expansions of the form

$$V(\mathbf{r}) = V_0(\mathbf{r}) + \sum_{n=1}^{\infty} U_n(\mathbf{r}), \quad (\mathbf{r} \in A), \quad (7.109a)$$

$$V(\mathbf{r}) = V'_0(\mathbf{r}) + \sum_{n=1}^{\infty} U'_n(\mathbf{r}), \quad (\mathbf{r} \in B), \quad (7.109b)$$

in which only the start terms must be solutions of the inhomogeneous PDE; we now assume

$$LV_0(\mathbf{r}) = S(\mathbf{r}) \quad \text{in } A, \quad V_0(\mathbf{r}) = \bar{V}(\mathbf{r}), \quad \mathbf{r} \in a', \quad (7.110a)$$

and analogously,

$$LV'_0(\mathbf{r}) = S(\mathbf{r}) \quad \text{in } B, \quad V'_0(\mathbf{r}) = \bar{V}(\mathbf{r}), \quad \mathbf{r} \in b', \quad (7.110b)$$

whereas the other terms must then necessarily satisfy the conditions

$$LU_n(\mathbf{r}) = 0 \quad \text{in } A, \quad U_n(\mathbf{r}) = 0 \quad \mathbf{r} \in a', \quad (7.110c)$$

and similarly,

$$LU'_n(\mathbf{r}) = 0 \quad \text{in } B, \quad U'_n(\mathbf{r}) = 0 \quad \mathbf{r} \in b'. \quad (7.110d)$$

The procedure can now be rewritten concisely if we use the notation $\mathbf{r}_a$ for any position on the internal boundary a , and similarly $\mathbf{r}_b$ for position on b , and assume that the corresponding function values are to be determined by interpolation, so that the corresponding Dirichlet problem is completely defined. We may write then

Start

Define $V_0(\mathbf{r}_a)$ by guess;

Solve Eqs. (7.110a);

Determine $V'_0(\mathbf{r}_b) = V_0(\mathbf{r}_b)$;

Solve Eqs. (7.110b);

Determine $U_1(\mathbf{r}_a) = V'_0(\mathbf{r}_a) - V_0(\mathbf{r}_a)$.

Thereafter the inhomogeneous terms are no longer needed explicitly. The iteration now takes the following form:

For $n = 1, 2, 3, \dots$,

- (1) Solve Eqs. (7.110c).
- (2) Determine $U'_n(\mathbf{r}_b) = U_n(\mathbf{r}_b)$.
- (3) Solve Eqs. (7.110d).
- (4) Determine $U_{n+1}(\mathbf{r}_a) = U'_n(\mathbf{r}_a)$.
- (5) Determine $\delta_n = \max(|U_n(\mathbf{r}_b)|, |U_{n+1}(\mathbf{r}_a)|)$.
- (6) Stop for $\delta_n < \varepsilon$, else continue.

Though not explicitly necessary for this procedure, it is quite favorable to calculate also the partial sums according to Eqs. (7.109) as then the complete sets of boundary values are known and the field calculations in A or B without coupling are possible.

The important advantage lies in the fact that *different* techniques of field calculations may be used in the different subdomains. Hence, it is possible to combine the FEM with the BEM, or with the FDM, or even with entirely different methods like *series expansions* techniques. The choice of the appropriate combination of procedures is not prescribed by the GAM but by the geometrical form of the domains in question. Therefore, there is now a great flexibility of choices.

7.5.2 Practical Examples

One typical application of the GAM has already been dealt with in Section 4.6: the subdivision of meshes within the frame of the FDM. As already obvious from Fig. 4.18, the two coupled meshes must overlap by at least one row along the *inner* boundary, and Fig. 4.19 demonstrates the necessary interpolation.

Another example is the combination of the FDM with BEM, as outlined in Section 7.2. The special property of this method, shown in Fig. 7.2, is that the domain of the BEM is *completely* included in that of the FDM, which means that it has then *no outer* boundary in the sense of Fig. 7.19.

It is possible to incorporate entirely different methods, *series expansions*, for instance, with coefficients to be determined as the calculations proceed. A quite simple example, calculated by R. Weysser [25] is shown in Figs. 7.20 and 7.21; also seen in Fig. 4.3. These show the calculation of the electrostatic field in a quarter of a deflector, the missing parts to be completed by symmetry operations. For reasons of simplification, only a two-dimensional Dirichlet problem was solved, which means that the fringe fields were ignored. In practice, this is not yet sufficient, but for purpose of demonstration it may be assumed here. The shape of the electrodes make it appropriate to introduce exponentially expanding cylindrical coordinates (u, v) according to Eq. (4.30).

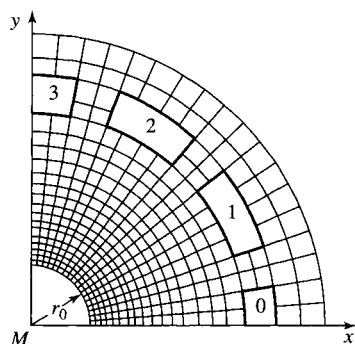


FIGURE 7.20 Polar mesh for the solution of Laplace equation in a quarter of an electrostatic deflector (after Weysser [25]). The labels of the electrodes are the factors j in $\sin(\pi j/6)$. (A realistic mesh must be finer by a factor of at least 2.)

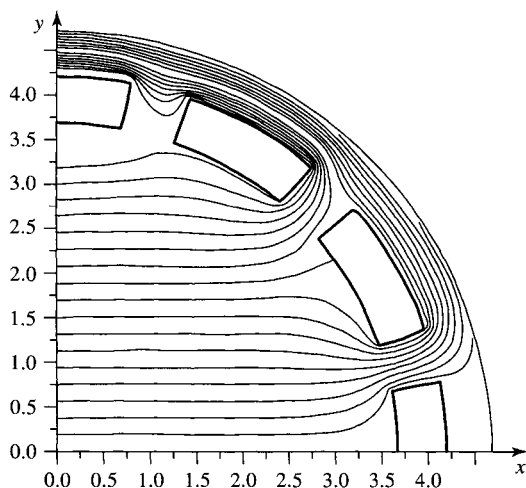


FIGURE 7.21 Equipotentials of the electrostatic potential in the configuration of Fig. 7.20 (Weysser [25]).

The nine-point formula in its simplest and most accurate version can then be used:

$$V_{i,k} = \frac{1}{5}(V_{i,k+1} + V_{i,k-1} + V_{i+1,k} + V_{i-1,k}) + \frac{1}{20}(V_{i+1,k+1} + V_{i+1,k-1} + V_{i-1,k+1} + V_{i-1,k-1}) + O(h^8) \quad (7.111)$$

with a corresponding symmetric six-point formula for points on the y axis. (These follow from the general scheme in Section 4.4 as the special case

$\alpha = 0$). In the vicinity of the edges Eq. (4.210) can be applied if the mesh is made fine enough, so that a node is not the nearest neighbor of two edges.

However, this kind of mesh can never reach the cylinder axis, and there necessarily remains an *open* inner boundary. In this situation, it is appropriate to introduce a series expansion, generally of the form

$$V(\mathbf{r}) = \sum_{m=0}^M C_m \Phi_m(\mathbf{r}), \quad (7.112)$$

in terms of partial solutions of the PDE

$$L\Phi_m(\mathbf{r}) = S(\mathbf{r})\delta_{m,0}. \quad (7.113)$$

In the present example, this is the Laplace equation in polar coordinates; hence these functions are of the form $r^m \cos m\varphi$ or $r^m \sin m\varphi$.

It is now important that the number of coefficients C_m can be strongly reduced by means of symmetry considerations. If the electrode potentials are chosen to be

$$\bar{V}_j = V_D \sin(j\pi/6), \quad (j = 0, 1, 2, 3), \quad (7.114)$$

the screening electrode having vanishing potential, then only *sine* terms of *odd* orders can appear in Eq. (7.112), hence

$$V(\mathbf{r}) = C_1 r \sin \varphi + C_3 r^3 \sin 3\varphi + \dots \quad (7.115)$$

The highest order to be considered depends on the maximum of r , for which Eq. (7.115) will be evaluated; in the domain of linearity, the first one is already sufficient.

The coupling in the sense of the GAM proceeds now in the following manner, as shown in Fig. 7.22. The inner boundary a is a quarter of a circle with radius r_0 to be chosen reasonably. The boundary b is then similarly a quarter of a circle with radius

$$r_b = r_0 \exp(ph), \quad (p = 1, 2 \text{ or } 3) \quad (7.116)$$

fitting the mesh. Let $N = \pi/2h$ be the number of meshes in the azimuthal direction, then the boundary values for the FDM on a are obtained from Eq. (7.115) by

$$V_{0,K} = C_1 r_0 \sin(kh), \quad (k = 0, \dots, N), \quad (7.117a)$$

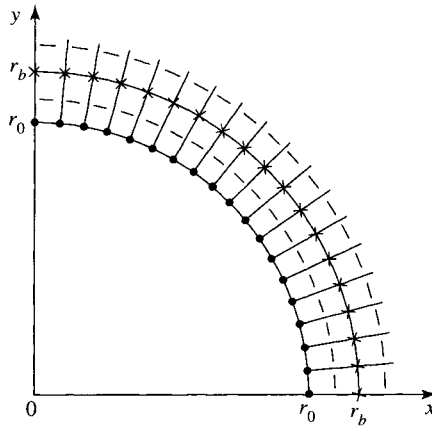


FIGURE 7.22 Coupling of algorithms: the boundary values for the FDM refer to the points with dots on the circle $r = r_0$, whereas those of the series expansion refer to the points with crosses on the circle $r = r_b$. Here $p = 2$ is assumed.

whereas the coefficient C_1 is obtained by Fourier analysis of the potentials on the meshline b :

$$C_1 = \frac{2}{Nr_b} \left\{ \frac{1}{2} V_{p,N} + \sum_{k=1}^{N-1} V_{p,k} \sin(kh) \right\}. \quad (7.117b)$$

It is of course reasonable to calculate these sine factors only once at the beginning and to store them, instead of recalculating them after each iteration.

This example demonstrates two favorable properties of the GAM:

- (i) It is not necessary to choose $M = N$ in Eq. (7.112); this would be quite unpractical if N is large. Only as many terms are actually used as are necessary to approximate $V(\mathbf{r})$ with sufficient accuracy.
- (ii) The coupling between the two domains can already be made after each FDM iteration over the mesh, and the SOR or SLOR (Section 5.6.3) can then be employed to speedup the convergence.

The results of this procedure are the equipotential lines, shown in Fig. 7.21. They demonstrate that a homogeneous field can be obtained within a fairly large circle and few electrodes. The same result could also be obtained with other methods, such as the BEM with infinity long wires as sources, but it is questionable whether the computational effort would then be smaller.

For this example, the FDM in spherical coordinates (Section 4.4.7) can also be combined with a series expansion in terms of spherical functions. In the most frequent case of rotational symmetry, this essentially means replacement of the trigonometric functions by Legendre polynomials. Depending on

the nature of the application, the powers of the radial coordinate may be positive or negative or even both. The coupling algorithm is quite similar to the one outlined earlier. A further generalization is the “sphere on orthogonal cone” model using Legendre functions of fractional orders [26]. The BEM can similarly be coupled with series expansions in spherical coordinates [27].

7.5.3 Systems with Several Different Materials

An entirely different class of application of the GAM is the solution of integral equations for configurations with more than two different materials. In the derivation of Eqs. (1.102), (1.108), (1.113), (1.115), and (1.124), it was assumed that there are only two different media distinguished by the labels 1 and 2. This is often too restrictive, and in this situation, the GAM provides a tool for solving more general boundary value problems. In fact, Scherle [28] applied it to Eq. (1.113) to calculate the magnetic field yoke and screening plate, the result being shown in Fig. 7.23. Here we shall start from a generalization of Eq. (1.115), because a knowledge of the scalar sources $\tau(r)$ makes it

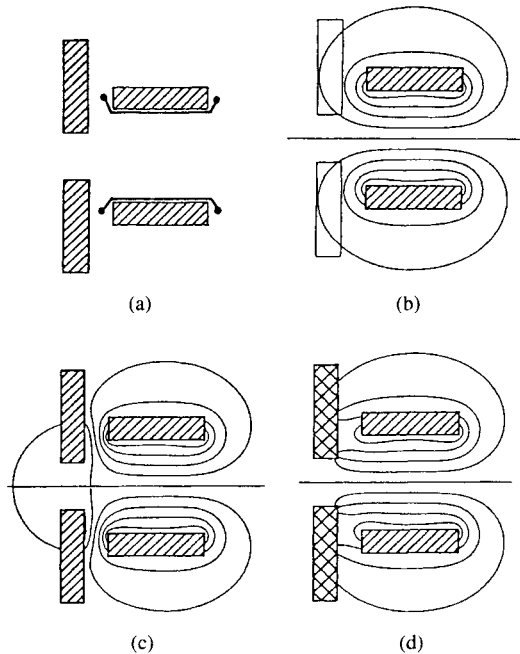


FIGURE 7.23 Saddle coils with ferromagnetic yoke and separate screening plate: (a) geometric configuration; (b) screening plate ignored (start); (c) ferromagnetic screening plate ($\mu \gg 1$); (d) superconducting screening plate ($\mu = 0$); after W. Scherle [28], [22].

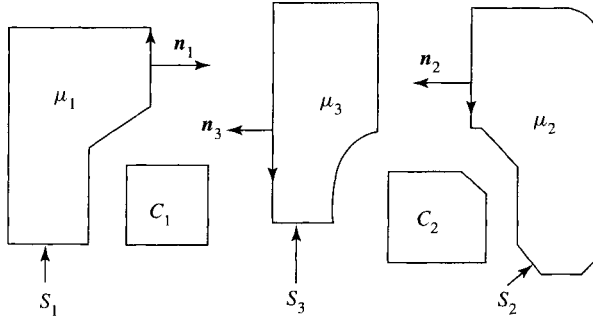


FIGURE 7.24 A system of three pole pieces with different material properties and of two coils. This configuration has no particular technical purpose, but merely illustrates the notation used.

possible to perform direct field calculations according to Eq. (1.114) without having to solve an additional Dirichlet problem. The formulation of the GAM is now slightly different from that presented in Section 7.5.1 (see Fig. 7.24).

Here we assume a system of N disjoint nonoverlapping and closed domains D_k with surfaces S_k in space. Their normals are directed outwards into the vacuum, and their relative permeabilities μ_k may be different from 1. The coils producing the driving field $\mathbf{H}_0(\mathbf{r})$ may be entirely located in the vacuum domain. The common outer boundary enclosing the whole system can be shifted to infinity, and there the natural boundary conditions must hold.

With a suitable choice of the notation, Eq. (1.114) can now be rewritten as

$$\mathbf{H}(\mathbf{r}) = \mathbf{H}_0(\mathbf{r}) + \sum_{k=1}^N \oint_{S_k} \nabla G(\mathbf{r}, \mathbf{r}') \tau_k(\mathbf{r}') da'. \quad (7.118)$$

With the definition

$$\lambda_i = (\mu_i - 1)^{-1}, \quad (i = 1, \dots, N) \quad (7.119)$$

and the replacement $\mu_1 \rightarrow \mu_i, \mu_2 = 1$, the application of Eq. (1.115) to any surface S_i can be cast in the form

$$\begin{aligned} & \left(\lambda_i + \frac{1}{2} \right) \tau_i(\mathbf{r}) + \oint_{S_i} \mathbf{n}_i(\mathbf{r}) \cdot \nabla G(\mathbf{r}, \mathbf{r}') \tau_i(\mathbf{r}') da' \\ &= -\mathbf{n}_i(\mathbf{r}) \cdot \left\{ \mathbf{H}_0(\mathbf{r}) + \sum_{k=1}^N \oint_{S_k} \nabla G(\mathbf{r}, \mathbf{r}') \tau_k(\mathbf{r}') da' \right\}, \\ & \quad (\mathbf{r} \in S_i), \quad (i = 1, \dots, N), \end{aligned} \quad (7.120)$$

in which the primed summation denotes the condition that the term for which $k = i$ is excluded, as it already appears on the left-hand side. This is now a set of *coupled* Fredholm equations. Its justification is easily understandable. Because $\mathbf{H}_0$ in Eq. (1.115) represents the external field, and in Eq. (7.118) this consists of several terms, all contributions with the exception of the self-interaction must now be included in it.

Of course, this system of integral equations can be transformed to a large system of linear equations by means of discretization and then solved directly. An iterative method is also possible and may become indispensable if the matrix is too large. The terms on the right-hand side thus have to be treated as inhomogeneities. Their evaluation is much simpler than the calculation of the principal values of improper integrals, appearing on the left-hand side. With suitable numerical techniques, these calculations can be performed efficiently.

To show briefly the method of iteration, we rewrite Eq. (7.120) as a system of linear equations, resulting from some kind of discretization, the details of the latter not being discussed here.

Let $\mathbf{T}_i$ be the vector (one-dimensional array) of sample data representing the function $\tau_i(\mathbf{r})$ on the surface S_i and similarly $\mathbf{F}_i$ the vector representing the function $\mathbf{n}_i \cdot \mathbf{H}_0$. These have rank L_i that will usually be different for different boundaries. Moreover, the integral operations are then to be approximated by matrix-vector products, the matrices being $\mathbf{J}_{ik}$ defined as follows: the first label i refers to the surface S_i with normal $\mathbf{n}_i$, the second label k to the surface S_k , over which the integration is carried out. This matrix $\mathbf{J}_{ik}$ consequently has rank $L_i \times L_k$. It is favorable to incorporate the diagonal elements $\lambda_i + 1/2$ in the matrix $\mathbf{J}_{ii}$; we must assume here that all these “diagonal” matrices are nonsingular, so that they can be decomposed by means of the *LU* algorithm (Section 5.6.1). Equations (7.120) now takes the concise form:

$$\mathbf{J}_{ii} \cdot \mathbf{T}_i = -\mathbf{F}_i - \sum_{k=1}^{N'} \mathbf{J}_{ik} \mathbf{T}_k, \quad (i = 1, \dots, N) \quad (7.121)$$

(remember that the labels refer here to entire boundaries, not to particular points!)

The repeated evaluation of the driving terms $\mathbf{F}_i$ can be avoided by representation of the unknown vectors $\mathbf{T}_i$ in terms of the iteration number n as in Eqs. (7.109):

$$\mathbf{T}_i = \sum_{n=0}^{\infty} \mathbf{T}_i^{(n)}. \quad (7.122)$$

As in Section (5.6.3), these are now different strategies for carrying out these iterations. Here we confine the presentation to the simplest case, that of the

Gauss-Seidel algorithm for $N = 2$:

$$\mathbf{J}_{11}\mathbf{T}_1^{(0)} = -\mathbf{F}_1, \quad \mathbf{J}_{22}\mathbf{T}_2^{(0)} = -\mathbf{F}_2 - \mathbf{J}_{21}\mathbf{T}_1^{(0)}; \quad (7.123a)$$

thereafter for $n = 1, 2, \dots$,

$$\mathbf{J}_{11}\mathbf{T}_1^{(n)} = -\mathbf{J}_{12}\mathbf{T}_2^{(n-1)}, \quad \mathbf{J}_{22}\mathbf{T}_2^{(n)} = -\mathbf{J}_{21}\mathbf{T}_1^{(n)} \quad (7.123b)$$

until sufficient convergence has been achieved. The rate of convergence is again the same as that of a geometric-series expansion and increases strongly with increasing distance between the coupled boundaries. The tedious numerical surface integrations, leading to the matrices J_{ik} and the LU decompositions need to be carried out only once at the beginning and the results are then stored. There remains only a sequence of matrix-vector multiplications to be performed which are fairly fast.

We have taken here Eq. (1.115) as an example to demonstrate the application of the GAM, but Eq. (1.108) for electrostatic fields has the same basic structure and can therefore be solved quite analogously. Moreover, the discretization of complicated Dirichlet problems for static fields leads to a system of linear equations, having the structure of Eq. (7.121) and can hence be solved in the outlined way; the matrices and vectors have a slightly different form and meaning.

7.5.4 Nonoverlapping Domains

The cases treated above have the special property that the domain of overlap is here the *entire space* as the surfaces S_k appearing there, are only carriers of *sources*, not boundaries of validity of Eq. (7.118) and analogous ones. We now return to the situation, dealt with in Section 7.5.1.

The overlap of domains, shown in Fig. 7.19 may sometimes be unpractical, for example when the physical surface of a material with discontinuous properties across it is simultaneously the boundary for the applicability of a chosen calculation technique. A standard example of this situation is the coupling of the FEM with the BEM, at the surface of a ferromagnetic yoke, as is outlined in details in Section 7.1.

Quite generally two coupling conditions are always necessary to join two domains together. In the conventional form of the GAM these are the consistency of boundary potentials on two *disjoint* surfaces (see Section 7.5.1). Here it is now the continuity of the flux Ψ , and the material equation for its normal derivative, which must finally be achieved by iteration.

Formally this case could be reduced to a structure similar to that shown in Fig. 7.19 as we could define two surfaces a and b in the vicinity of the real

yoke surface and then determine the potentials on these by linear extrapolation. This is, however, quite unfavorable because for small shifts the coupling becomes very weak and for larger shifts the accuracy obtained very poor. Hence the method of Section 7.1 is the most advantageous.

In summary, the GAM is a very powerful and flexible method of field calculation. Here we could present only a few characteristic examples, and many more can be found in the corresponding literature.

7.6 FAST FIELD CALCULATION

After the numerical solution of a boundary value problem for the electric or magnetic field, there remains the task of calculating the field strength at any position $\mathbf{r}$ inside the given domain, as this is necessary for ray tracing. For field data obtained by applications of the FDM or the FEM, this task has already been outlined in Section 5.5, and we refer here to these methods.

If the BEM were applied, and a set of known surface sources is given, a special interpolation technique is not in principle necessary because the analytic functions presented in Section 6.2 can be evaluated at any position apart from their singularities. There is, however, one disadvantage: since the number of discrete ring sources is fairly large, typically some hundreds, direct field calculation by evaluation of these analytical expressions becomes fairly *slow*. This disadvantage becomes obvious if many trajectories are to be calculated to determine the particle-optical properties of the device in question.

There are various strategies for overcoming this deficiency. A simple one, though not the best, is the calculation and storage of potentials in a suitable rectangular grid, whereupon the methods of Section 5.5 can be applied. The method outlined below [29] considers the special properties of systems having a straight optic axis and of fields obtained by multipole series expansions with respect to the latter. An important aspect is that, for particle optical reasons, the beam rarely occupies more than a quarter of the bore diameter, so that the field calculation can be confined to a narrow domain near the optic axis. This will bring an enormous simplification of the otherwise tedious field computations.

The development of this method proceeds in three natural steps:

- (i) A purely radial interpolation technique;
- (ii) A two-dimensional interpolation for solutions of the multipole-Laplace equation;
- (iii) The efficient evaluation of the Fourier-series expansion.

The resulting method will be more flexible than those described in Section 5.5, as it is *not* necessarily confined to the use of regular rectangular meshes.

7.6.1 Radial Interpolation

It is favorable to consider explicitly the fact that the functions to be calculated, the amplitudes of the Fourier-series expansions, are *even* with respect to the radial coordinate r .

We hence introduce a variable

$$s := r^2 = x^2 + y^2 \quad (7.124)$$

and consider *regular* functions of the form

$$F(s) = f(r), \quad (7.125)$$

which means that these can be expanded as a power series in terms of s :

$$F(s) = \sum_{n=0}^{\infty} a_n s^n. \quad (7.126)$$

Of course, the summation must be truncated after a finite number N , and the next task is to determine this limit and the corresponding coefficients as economically as possible. Subsequently we shall denote derivatives with respect to s by dots and those with respect to r by primes. We find then that the function

$$\dot{F}(s) = \frac{1}{2r} f'(r) = \sum_{n=1}^{\infty} n a_n s^{n-1} \quad (7.127)$$

must remain *finite* for $r \rightarrow 0$ and obtain in particular

$$\dot{F}(0) = \frac{1}{2} f''(0) = a_1. \quad (7.128)$$

Because the function values on the optic axis are the most important ones in particle optics, we now assume that $a_0 = F(0)$ and a_1 must be considered *exactly* in any kind of approximation, which means that Chebyshev polynomials cannot be used.

A very accurate interpolation with only few references to the original function $f(r)$ and its derivative is obtained by matching these at two additional positions

$$s_1 = r_1^2 = h^2, \quad s_2 = r_2^2 = 2h^2, \quad (7.129a)$$

which means that the values

$$f_j := f(r_j), \quad f'_j := f'(r_j), \quad (j = 1, 2), \quad (7.129b)$$

are reproduced exactly; this implies that $N = 5$. The coefficients $a_2, \dots, a_5$ are then determined uniquely by these requirements. The results can be written concisely in terms of reduced coefficients:

$$\bar{a}_n := a_n h^{2n}, \quad (n = 1, \dots, 5), \quad (7.130a)$$

$$\begin{pmatrix} \bar{a}_2 \\ \bar{a}_3 \\ \bar{a}_4 \\ \bar{a}_5 \end{pmatrix} = \begin{pmatrix} 4 & -4 & 7 & -2 \\ -4 & 8 & -17 & 5 \\ 1 & -5 & 13 & -4 \\ 0 & 1 & -3 & 1 \end{pmatrix} \begin{pmatrix} f_1 - a_0 - \bar{a}_1 \\ h f'_1/2 - \bar{a}_1 \\ (f_2 - a_0 - 2\bar{a}_1)/4 \\ (h f'_2/\sqrt{8} - \bar{a}_1)/4 \end{pmatrix}. \quad (7.130b)$$

After evaluation of this matrix-vector product, it is favorable to solve Eq. (7.130a) for the original coefficients a_n , because h may alter with the axial coordinate z . The accuracy of this technique will be discussed at the end of this section.

The main drawback of this approximation is the condition that f'_1 and f'_2 must be available, which is not always the case; we hence look for a technique that does not require these. It is then necessary to introduce additional sampling positions. These should be compatible with those of Eqs. (7.129a), and minimize the oscillations of the error polynomial. With a renumbering of the sequence these are favorably chosen as

$$s_j \equiv r_j^2 = c_j h^2, \quad (j = 1, \dots, 4), \quad (7.131a)$$

with the fixed numbers

$$c_j = (1/2, 1, 5/3, 2), \quad (7.131b)$$

and the requirement is now that

$$f_j = f(r_j), \quad (j = 1, \dots, 4) \quad (7.131c)$$

shall be interpolated exactly. The corresponding matrix-vector product now takes the form

$$\bar{a}_{i+1} = \sum_{k=1}^4 B_{ik} (f_k - a_0 - c_k \bar{a}_1) \quad (7.132a)$$

with the matrix

$$B = \begin{pmatrix} 320/21 & -5 & 9.72/7 & -1.25/3 \\ -32 & 15.5 & -4.86 & 1.5 \\ 64/3 & -12.5 & 4.86 & -4.75/3 \\ -32/7 & 3 & -9.72/7 & 0.5 \end{pmatrix} \quad (7.132b)$$

Apart from the fact that here these fractional numbers are obtained, this version is no more complicated than the former one. Both kinds of approximation are schematically shown in Fig. 7.25; their accuracy can be determined from the corresponding *error polynomials*, shown in Fig. 7.26.

Because all powers of r up to the tenth are considered correctly, this error polynomial must have the general form

$$\varepsilon_{1,2}(r) = \frac{h^{12}}{12!} f^{(12)}(\bar{r}) \eta_{1,2}(r^2/h^2), \quad (7.133)$$

$\bar{r}$ being some value for which $0 \leq \bar{r} \leq h$. The function $\eta(x)$ is a polynomial of order 6; its form is determined by the interpolation conditions, as it must have zeros at all prescribed sampling positions. We obtain hence the functions

$$\eta_1(x) = x^2(x-1)^2(x-2)^2, \quad (7.134a)$$

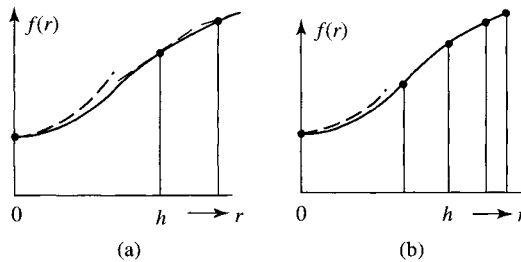


FIGURE 7.25 Approximation of an even function $f(r)$ by a polynomial of fifth degree in r^2 in parabolic order at the origin: (a) with two additional function values and derivatives; (b) with four additional function values.

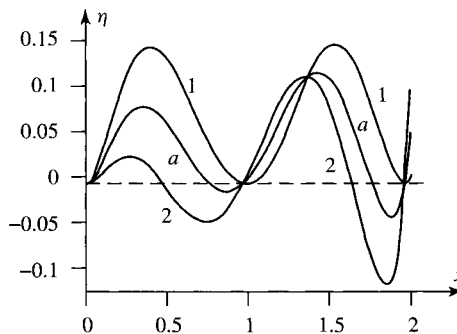


FIGURE 7.26 Reduced error polynomials $\eta(x)$: 1: $\eta_1(x)$ according to Eq. (7.134a); 2: $\eta_2(x)$ according to Eq. (7.134b); a : average $(\eta_1 + \eta_2)/2$.

$$\eta_2(x) = x^2 \prod_{n=1}^4 (x - c_n) \quad (7.134b)$$

with $x = r^2/h^2$. From their graphs, shown in Fig. 7.26, it becomes obvious that both approximations are practically equivalent with respect to accuracy. A slight improvement, essentially for $x > 1.5$, is obtained by using the arithmetic means $(\eta_1 + \eta_2)/2$, which implies introducing the arithmetic means of the coefficients a_i of Eqs. (7.130a) and (7.132a). In contrast to these approximations, the Taylor-series expansion with respect to the origin has the corresponding polynomial $\eta_T = x^6$ with maximum 64 at $x = 2$. This error is more than 400 times larger than the maxima of η_1 and η_2 .

The accuracy obtained can be quite high: for instance with $h = 0.5$, the function $\cos r$ is approximated with an error of about 10^{-14} in the interval $0 \leq r \leq 0.7$. In the worst case the potential of a charged ring is approximated with a relative error of about 10^{-10} with $h = 0.35 R$, R being the ring radius. This covers half the bore radius, which is perfectly sufficient in practice. This method is significantly different from the slice method, introduced by Barth *et al.* [30]; in the latter, equidistant spacing is assumed.

7.6.2 Two-Dimensional Interpolation

We now consider regular solutions $P(z, r)$ of the multipole-Laplace equation

$$\Delta_\alpha P = P_{|zz} + P_{|rr} + \alpha r^{-1} P_{|r} = 0 \quad (7.135)$$

as are obtained with the BEM for rotationally symmetric boundaries. For $z = \text{const.}$ these satisfy the basic assumptions (7.125) with (7.124), and so the outlined interpolation techniques can be applied to them. Now, however, we generalize the series expansion (7.126) to

$$P(z, r) = \sum_{n=0}^N a_n(z) s^n, \quad (N = 5). \quad (7.136)$$

Strictly, the coefficients should now have an additional label α , but for reasons of conciseness we shall drop this here as long as this causes no confusion.

A novel aspect is now that the derivatives $a_n''(z)$ are defined by the PDE (7.135). In fact, on introducing Eq. (7.136) into Eq. (7.135) and comparing equal powers of s , we obtain in turn

$$\begin{aligned} a_n''(z) &= -2(n+1)(2n+1+\alpha)a_{n+1}(z), \quad (0 \leq n < N), \\ a_N''(z) &= 0. \end{aligned} \quad (7.137)$$

For the application of Hermite-interpolation techniques with respect to the coordinate z , it is necessary to know the derivatives of first order as well, but these are not furnished by the above scheme. This is, however, no problem if the BEM program supplies an analytic function for $P_{|z}(z, r)$, which is usually the case. As the differentiation with respect to z commutes with the operator Δ_α , we have

$$\partial_z \Delta_\alpha P = \Delta_\alpha P_{|z} = 0; \quad (7.138)$$

the radial-series expansion technique can also be applied to this function, and we hence obtain

$$P_{|z}(z, r) = \sum_{n=0}^N a'_n(z) s^n, \quad (7.139)$$

$$\begin{aligned} a''_n(z) &= -2(n+1)(2n+1+\alpha) a'_{n+1}(z), \quad (0 \leq n < N), \\ a'''_N(z) &= 0. \end{aligned} \quad (7.140)$$

The sampling data must now be determined from the potential field. For selected values of z and h , the functions $P(z, r_j)$, $P_{|z}(z, r_j)$, and $P_{|r}(z, r_j)$ are to be evaluated at $r_j = hc_j^{1/2}$ ($j = 1, \dots, 4$). Similarly the *axial* values

$$a_0(z) = P(z, 0), \quad a'_0(z) = P_{|z}(z, 0) \quad (7.141)$$

are immediately furnished by the BEM program. However, the coefficients a_1 and a'_1 are not directly given. It is not difficult to determine the derivatives a''_0 and a'''_0 by evaluation of Eqs. (6.87) and (6.88) for each charged ring and linear superposition of the results. Then with $n = 0$ in Eqs. (7.137) and (7.140) the relations

$$a_1(z) = -a''_0(z) / (2 + 2\alpha), \quad (7.142a)$$

$$a'_1(z) = -a'''_0(z) / (2 + 2\alpha) \quad (7.142b)$$

are obtained, which can now be used to evaluate the series expansions.

This kind of approximation has some properties in common with the familiar paraxial series expansion, given by Eqs. (2.69) and (2.72). In fact, it will give us exactly the same results, provided that $P(z, r)$ is *exactly* a polynomial of tenth degree in r . But as this is usually not true, it is then better to determine $a_2, \dots, a_5$ and their derivatives by the method described here than by repeated differentiations of $a_0(z)$.

The practical realization of this technique is shown in Fig. 7.27. A strictly monotonic sequence of sampling planes $z = z_i$ ($0 \leq i \leq M_z$) is chosen, and in each of these a suitable radial interval size $h = h_i$, so that the particle beam

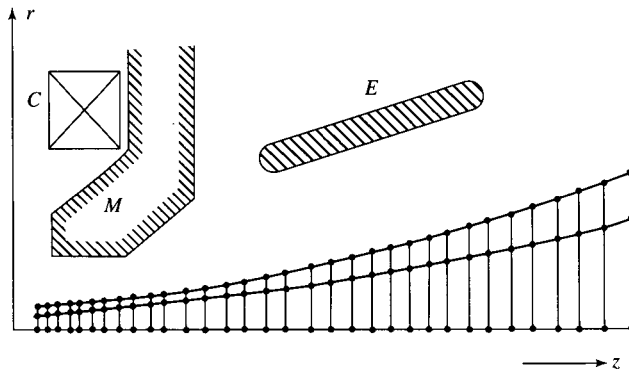


FIGURE 7.27 Example of the use of a grid for fast interpolation in a system composed of an open magnetic lens M with coil C and an electrostatic deflector electrode E .

stays entirely in the interval $0 \leq r \leq 1.4h_i$, which must not exceed half of the local bore radius. It is not necessary either to choose equidistant planes or to assume equal values of h_i , giving a wide flexibility in matching this grid to the given geometrical configuration.

The algorithm outlined above is now carried out in turn for the sampling points on the planes $z = z_i$, and the results stored as two-dimensional arrays $a_{n,i}$, $a'_{n,i}$ for $n = 0, \dots, 5$ and $i = 0, \dots, M_z$. The storage of the derivatives of higher orders is not necessary as their evaluation according to Eqs. (7.137) and (7.140) is quite fast.

If now the potential and its derivatives are required in a certain plane with coordinate $z_{i-1} \leq z \leq z_i$, then all marginal values $a_n^{(j)}(z_{i-1})$ and $a_n^{(j)}(z_i)$, $j = 0, \dots, 3$, which are needed to carry out a sequence of Hermite interpolations of seventh degrees ($M = 3$) for the calculation of $a_n^{(j)}(z)$, are available. This procedure can be made fairly fast as the corresponding form functions, the coefficients A and B in Eq. (3.74), need to be calculated only *once* for all six sets of coefficients. The results are splines $a_n(z)$, which are three times continuously differentiable, which is perfectly sufficient for practical purposes.

The conditions (7.137) and (7.140) will generally not be satisfied by Hermite polynomials, as a seventh degree in z conflicts with a tenth degree in r . As it is well known that differentiation of interpolation polynomials necessarily leads to a loss of accuracy owing to the lowering of the degree, it is favorable to evaluate Eqs. (7.137) and (7.140), which are even faster than the calculation of differentiated Hermite polynomials.

The differentiation schemes (7.137) and (7.140) run into difficulties if the factor in front of $-a_{n+1}$ and $-a'_{n+1}$ becomes too large, which may happen mainly for large values of α . This effect can be noticed if the graphs of the corresponding functions show unreasonable oscillations. In such a case, it is

then necessary to determine these derivatives numerically, for instance, by means of a quintic spline Eq. (3.103). Another possibility is the application of the finite-difference approximation Eq. (5.104), the notation being adapted correspondingly. This proceeds as follows. Let $a_{n,i}$ denote the function value of $a_n(z)$ at the position $z = ih$, and analogous abbreviations be introduced for the derivatives. Then we have

$$a''_{n,i} = 2h^{-2}(a_{n,i-1} - 2a_{n,i} + a_{n,i+1}) - 0.5h^{-1}(a'_{n,i+1} - a'_{n,i-1}) \quad (7.143a)$$

and

$$a'''_{n,i} = 7.5h^{-3}(a_{n,i+1} - a_{n,i-1}) - 1.5h^{-2}(a'_{n,i+1} + 8a'_{n,i} + a'_{n,i-1}), \quad (7.143b)$$

with a discretization error of fourth order in both cases. These formulas can be generalized for unequal interval lengths and then, of course, become more complicated. The corresponding coefficients are uniquely defined by the requirement that the function $a_n(z)$ and its derivative assume prescribed values at three subsequent positions $z_{i-1} < z_i < z_{i+1}$, which leads to a polynomial of fifth degree.

After performing the Hermite interpolations with respect to the coordinate z , we have a set of series expansion coefficients $a_0(z), \dots, a_5(z)$ and their derivatives, and these can be used directly in particle optics programs that require these explicitly, for instance, to analyze aberrations. However, it is no problem to go further and to evaluate the potential and its partial derivatives at any position (z, r) within the domain of approximation by summation of the corresponding series expansions.

7.6.3 Three-Dimensional Interpolation

We now have developed the necessary tools for interpolations and differentiations in three dimensions. These must be carried out in such a way that they become most accurate in the vicinity of the optic axis. Because we have to superimpose multipole fields of different orders m , it is necessary to introduce the latter as an additional label, and to consider the relation $\alpha = 2|m| + 1$ according to Eq. (2.9). With an upper limit M for m , the Fourier-series expansion of the potential in combination with radial power-series expansions of the general form of Eq. (7.136), the total potential can be written as

$$\tilde{V}(z, r, \varphi) = \sum_{m=0}^M \sum_{n=0}^N r^m s^n [a_{mn}(z) \cos m\varphi + b_{mn}(z) \sin m\varphi]. \quad (7.144)$$

This form is, however, not advantageous because it requires the repeated evaluations of square roots and trigonometric functions. Moreover, special cases

are to be distinguished on calculating the partial derivatives in the vicinity of the optic axis. Furthermore, a complex representation that is possible in modern programming languages does not bring a significant gain.

A favorable computer-adapted algorithm is obtained by introducing the two-dimensional harmonic functions

$$\begin{aligned} c_m(x, y) &:= r^m \cos m\varphi = \Re(x + iy)^m, \\ s_m(x, y) &:= r^m \sin m\varphi = \Im(x + iy)^m. \end{aligned} \quad (7.145)$$

Without using complex algebra these can be calculated efficiently by means of the familiar addition theorems:

$$c_0 = 1, \quad s_0 = 0$$

$$\text{for } (m = 1, 2, \dots, M) \quad \{c_m = xc_{m-1} - ys_{m-1}, \quad s_m = xs_{m-1} + yc_{m-1}\}. \quad (7.146)$$

Partial differentiation of these functions is quite easy:

$$c_{m|x} = mc_{m-1}, \quad s_{m|x} = ms_{m-1}, \quad c_{m|y} = -ms_{m-1}, \quad s_{m|y} = mc_{m-1} \quad (7.147)$$

as can be verified by induction. There is no problem in carrying this scheme on to higher orders.

In terms of these functions the potential $\tilde{V} = V$ of Eq. (7.144) now takes the concise *cartesian* form

$$V(x, y, z) = \sum_{m=0}^M \sum_{n=0}^N (x^2 + y^2)^n [a_{mn}(z)c_m(x, y) + b_{mn}(z)s_m(x, y)]. \quad (7.148)$$

Because the coefficient functions $a_{mn}(z)$ and $b_{mn}(z)$ are splines of sevenths order, the differentiation of $V(x, y, z)$ with respect to all three coordinates is now a simple straight forward procedure. This is a considerable advantage for application in ray-tracing programs. A particular advantage of this representation is the fact that the linear (focusing) term can be separated from the remainder, which facilitates aberration analysis.

The author has tested this method and observed that the replacement of Eqs. (7.137) and (7.140) by Eq. (7.143) has only very little effect on the final result in Eq. (7.148), because the other factors reduce these contributions very significantly especially in the vicinity of the optic axis.

Finally, it should be mentioned that the flux field in a round magnetic lens also fits the general interpolation scheme. It is then favorable to use the potential $\Pi(z, r)$, defined by Eq. (2.41) and satisfying Eq. (7.135) with $\alpha = 3$. Using $s = r^2 = x^2 + y^2$, we can rewrite Eqs. (2.42) in the cartesian form

$$B_z = \Pi + s \partial \Pi / \partial s, \quad (7.149a)$$

$$B_x = -\frac{x}{2} \partial \Pi / \partial z, \quad (7.149b)$$

$$B_y = -\frac{y}{2} \partial \Pi / \partial z, \quad (7.149c)$$

which are well behaved in the vicinity of the optic axis.

7.6.4 Variation of Parameters and Perturbations

Quite often, a charged-particle optical device consists of many components, and it can be assumed that the electric or magnetic fields produced by these superimpose linearly. It then makes sense to alter some of the parameters characterizing these fields to study the sensitivity of the device to them.

Another task of importance is analysis of perturbations caused by constructional imperfections, for instance, by shift, tilt, or elliptic deformation of its surfaces [31,32], or by deviation of the electrode potentials. We must assume here that these effects are quite *small*. It is then possible to transform a geometric shift $s(\mathbf{r})$ of the surface into an equivalent potential deviation by writing

$$\delta \bar{V}(\mathbf{r}) = s(\mathbf{r}) \cdot \nabla V(\mathbf{r}) + \delta U(\mathbf{r}), \quad (7.150)$$

with the gradient to be evaluated at the corresponding surface point. The additional term $\delta U(\mathbf{r})$ describes the deviation of the potential on the ideal surface from the described boundary value, for instance, by electric instability.

This function $\delta \bar{V}(\mathbf{r})$ is now to be expanded as a Fourier series with respect to the azimuth, and the set of corresponding Dirichlet problems is then solved numerically. It is of importance that this procedure is to be carried out only *twice* for each component: the first time for the nominal boundary values and the second time for a variation or perturbation with *unit* amplitude; the same holds for the calculation and storage of the interpolation coefficients.

Thereafter—within the range of linearity—the coefficients in Eq. (7.148) can be written as

$$\begin{aligned} a_{mn}(z) &= a_{mn}^{(o)}(z) + \lambda_m \delta a_{mn}(z), \\ b_{mn}(z) &= b_{mn}^{(o)}(z) + \mu_m \delta b_{mn}(z). \end{aligned} \quad (7.151)$$

The nominal components vanish and consequently their calculation and storage can be omitted if the corresponding Fourier potential is entirely caused by perturbations. Ray tracing with various choices of the parameters λ_m and μ_m is now quite fast as only the superpositions and interpolations are to be repeated, not entire field calculations.

7.7 CALCULATION OF EQUIPOTENTIALS

It is customary and favorable to illustrate potential fields by systems of equipotentials, that is, by lines or surfaces with prescribed *constant* values of the potential. As three-dimensional perspective plots require a very large effort, we confine here our consideration to the two-dimensional case. For reasons of conciseness the potential in question will be denoted by $V(x, y)$ and its gradient by $\mathbf{g}(x, y)$, though the potential itself and the coordinates may have very different mathematical meanings. Most frequently a rectangular frame $x_a \leq x \leq x_b, y_a \leq y \leq y_b$ is prescribed and a set of selected values $U_1, U_2, \dots, U_N$, and there then arises the task of determining all those lines $V(x, y) = U_n, (n = 1, \dots, N)$ that are located in this frame. Examples for such graphs have already been presented at various places in this volume. Here we are concerned with algorithms for determining such graphs as efficiently as possible. As this technique is always the same for the different lines, it is sufficient to outline it for a representative one with a constant potential $V(x, y) = U$.

7.7.1 Equipotentials in FEM Grids

The calculation of equipotentials in triangular grids is most easy in the first order of the FEM as shown in Fig. 7.28: for $V_1 \leq U \leq V_2$ an equipotential must intersect a mesh line between the corresponding node values V_1 and V_2 ; the intersection point is then determined by linear interpolation. As the next intersection point must certainly be located on one of the adjacent sides of the triangle, it can be determined analogously, and the two points then joined by a straight line, whereupon the search is continued in the next triangle.

The whole procedure is shown in Fig. 7.29. We have to find a first pair of values V_1 and V_2 with $U \in (V_1, V_2)$, preferably at the margin if possible.

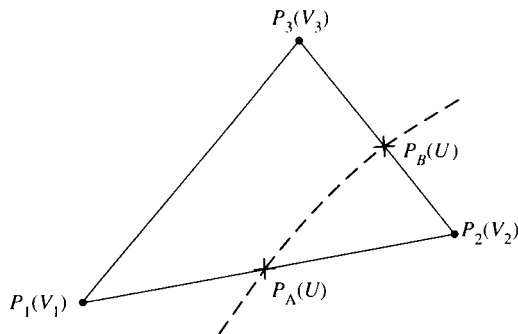


FIGURE 7.28 Linear interpolation of the potential in a triangle (P_1, P_2, P_3) with edge potentials V_1, V_2, V_3 .

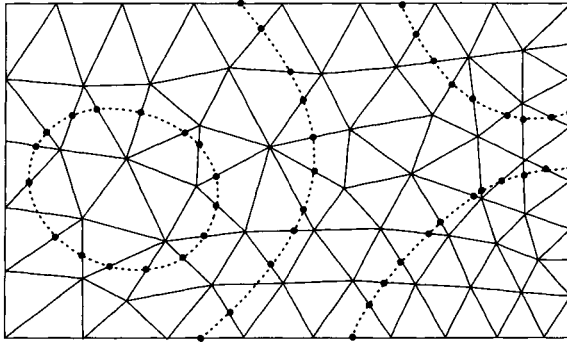


FIGURE 7.29 Determination of equipotentials in triangular grids (dotted lines): a closed line (l), a single line (m), and a line consisting of two disjoint parts (r).

Then the sequence of searches and interpolations outlined above is carried out until it again reaches a margin or the starting point. Special considerations are necessary whenever an equipotential consists of several disjoint lines. Then this procedure has to be repeated with suitably chosen starting points, until all the parts have been found.

The first order of the FEM and the polygons resulting from it are often not accurate enough; approximations of higher orders must then be used, such as those described in Section 5.3. The calculation of intersection points between equipotentials and mesh lines must also be carried out more accurately to avoid artefacts in the plots. A suitable algorithm is presented below. Apart from this modification, the general algorithm is quite analogous to the one outlined above: the result is always a sequence of discrete sampling points representing the equipotential. Here, however, it makes sense to trace a *curve spline* through these points, as the approximation is better than linear.

7.7.2 Determination of Intersection Points

The following algorithm is fairly general and may be applied in various situations, for instance, to determine intersection points between equipotentials and mesh lines, or parts of a boundary, or in context with the method of the next section.

We consider now a part of a curve, given in *parametric* form (see Fig. 7.30)

$$\mathbf{r}(s) = (x(s), y(s)), \quad s_2 \leq s \leq s_2 \quad (7.152a)$$

specified by its two end points and tangents in

$$\mathbf{r}_{1,2} = \mathbf{r}(s_{1,2}), \quad \mathbf{t}_{1,2} = \mathbf{r}'(s_{1,2}). \quad (7.152b)$$

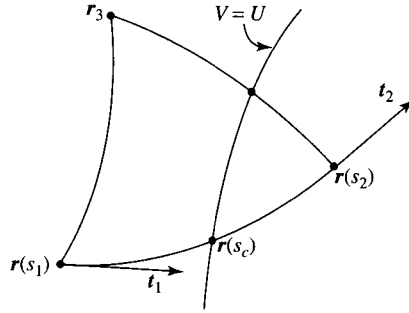


FIGURE 7.30 Intersection of an equipotential $V(\mathbf{r}) = U$ with a bent line element.

This function can then simply be approximated by a vectorial cubic Hermite polynomial, and for graphical purposes this is quite sufficient.

We now assume that the potential and its tangential derivative are known at the two end points

$$V_{1,2} = V(\mathbf{r}_{1,2}), \quad g_{1,2} = \mathbf{t}_{1,2} \cdot \mathbf{g}(\mathbf{r}_{1,2}). \quad (7.153)$$

By means of these, a cubic Hermite polynomial for $V(s)$ can again be defined. In the worst case, the cubic equation $V(s) = U$ must be solved for s to find the intersection point. However, this procedure is rather unsatisfactory and time-consuming, as it is to be carried out quite often.

An essential gain is obtained if we can ensure that the function $V(s)$ is *strictly monotonic* in the whole interval $s_1 \leq s \leq s_2$. A sufficient criterion for this is given by

$$|V_2 - V_1| > \frac{1}{6}(s_1 - s_2)|g_1 + g_2|, \quad (g_1 \cdot g_2 > 0), \quad (7.154)$$

where $V_2 - V_1$, g_1 , and g_2 must all have the *same sign*. This results from the condition that the derivative at the midpoint must also have this sign. A necessary criterion for the existence of an intersection point is then

$$(U - V_1)(U - V_2) \leq 0. \quad (7.155)$$

If Eq. (7.154) is valid and Eq. (7.155) is not, then the corresponding line element can be skipped. In the majority of cases, the line elements are so short that $V(s)$ is nearly a linear function, which implies that both $|g_1|$ and $|g_2|$ are large enough for

$$|V_2 - V_1|/(s_2 - s_1) \leq 1.5 \text{ Min}(|g_1|, |g_2|) \quad (7.156)$$

to be satisfied. It then makes sense to approximate the *inverse* function $s(V)$, defined in the interval $V \in [V_1, V_2]$ by a cubic Hermite polynomial. This

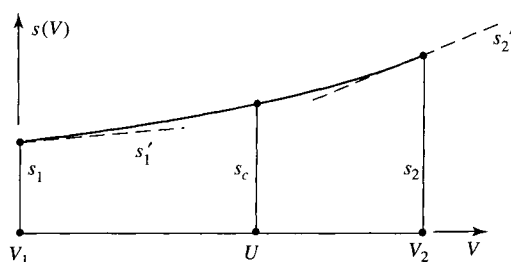


FIGURE 7.31 Hermite interpolation of the inverse function, $s'_1 = 1/g_1$ and $s'_2 = 1/g_2$ being the slopes at the endpoints.

has the known marginal values s_1, s_2 , and g_1^{-1}, g_2^{-1} , and the calculation of $s_c = s(U)$ is now a single evaluation of this polynomial instead of an iterative procedure (see Fig. 7.31). The evaluation of the polynomial of Eqs. (7.152) at $s = s_c$ is then straightforward. In this context, it is not necessary to assume normalized tangents, so that the parameter s is not necessarily the arc length, though this choice is not forbidden.

If one of the conditions for the use of the inverse polynomial is not satisfied, then the more tedious iterative solution of the equation $V(s) = U$ must be carried out. This will be necessary in the vicinity of saddle points, in transitions to practically field-free domains, or even if an equipotential intersects a line element twice or more.

7.7.3 The General Search Algorithm

We now consider the more general case, in which a function $V(x, y)$ and its gradient $\mathbf{g}(x, y)$ are defined in a rectangle $x_a \leq x \leq x_b, y_a \leq y \leq y_b$ without any meshes, for instance, as result of a computation using the BEM.

A first task is then the determination of a starting point (x_0, y_0) for the line $V(x, y) = U$. A simple strategy to find all those lines that intersect the rectangular margin consists in dissecting the latter into a number of sufficiently small intervals and storing the values of V and g referring to their end-points. If these intervals are counted sequentially from 1 to M , then the interval with label s and with

$$(U - V_{s-1})(U - V_s) \leq 0 \quad (7.157)$$

includes the starting point, which can then be determined more accurately using the algorithm outlined above. If no such interval exists, then the equipotential in question lies entirely outside or inside the given frame, which can be decided by the direction of the gradient. In the latter case the marginal position with smallest absolute deviation is chosen to start inwards in the direction of the gradient (see Fig. 7.32), until a step is found that satisfies Eq. (7.155), so

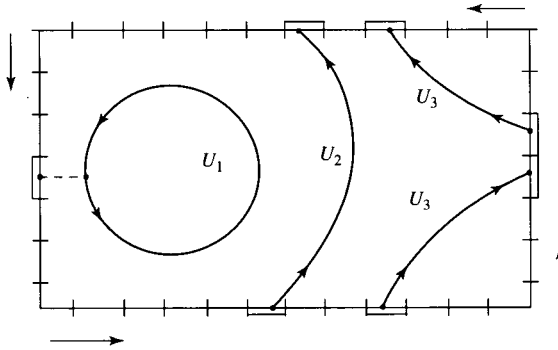


FIGURE 7.32 Different cases of searching for a start point. The arrows indicate the direction of search and the continuation.

that the correct starting point can be calculated. In sophisticated configurations this procedure may not always be successful; it is then necessary to introduce suitable starting guesses “by hand.”

We assume now that the position (x_0, y_0) with gradient $\mathbf{g}_0$ is known. Immediately after that, only a step in the tangential direction is possible, that is normal to $\mathbf{g}_0$; subsequently however, the second last data set can be considered too to determine a new step as shown in Fig. 7.33a,b. We assume here that the curvature is so small that the difference between arc length and chord length can be ignored. It can then be shown that the gradient at the *midpoint* of each chord must be perpendicular to the chord itself to cancel out quadratic error terms. With two given vectors $\mathbf{g}_n$ and $\mathbf{g}_{n-1}$, referring to the parameters s_n and $s_n - h_{n-1}$ a linear function

$$\mathbf{g}(s) = \mathbf{g}_n + (s - s_n)(\mathbf{g}_n - \mathbf{g}_{n-1})/h_{n-1} \quad (7.158a)$$

is given. Its value at the next midpoint $s_n + h_n/2$ is consequently

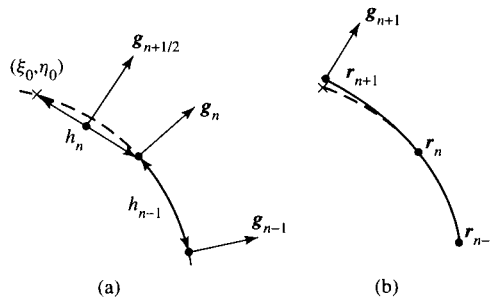


FIGURE 7.33 Search for a new point: (a) predictor step according to Eqs. (7.158), (7.159); (b) corrector step according to Eq. (7.161).

$$\mathbf{g}_{n+1/2} = \mathbf{g}_n + \frac{h_n}{2h_{n-1}}(\mathbf{g}_n - \mathbf{g}_{n-1}). \quad (7.158b)$$

The next position becomes

$$\begin{aligned} x_{n+1} &\approx \zeta_0 = x_n - \sigma h_n N_y, \\ y_{n+1} &\approx \eta_0 = y_n + \sigma h_n N_x, \end{aligned} \quad (7.159)$$

with the normalized vector

$$N := \mathbf{g}_{n+1/2} / |\mathbf{g}_{n+1/2}|, \quad (7.160)$$

and $\sigma = +1$ or -1 depending on the chosen orientation.

This new position is not located exactly at the equipotential line, as can be determined by a field evaluation at this position. This deviation can be eliminated by means of Newton iterations:

$$\begin{aligned} \xi_{v+1} &= \xi_v - \frac{(V(\xi_v, \eta_v) - U) g_x(\xi_v, \eta_v)}{g_x^2 + g_y^2 + \varepsilon^2}, \\ \eta_{v+1} &= \eta_v - \frac{(V(\xi_v, \eta_v) - U) g_y(\xi_v, \eta_v)}{g_x^2 + g_y^2 + \varepsilon^2}, \end{aligned} \quad (7.161)$$

for $v = 0, 1, 2, \dots$ until the deviation $|V - U|$ has dropped below a reasonably chosen error limit. The tiny constant ε^2 in the denominators prevents divisions by zero. The final values of (ξ, η) are stored as (x_{n+1}, y_{n+1}) and the associated gradient as $\mathbf{g}_{n+1}$, after which the search is continued.

Newton's method, although quite familiar, is not the fastest one. As soon as the equipotential is included in an interval, inverse cubic Hermit interpolation converges much faster, provided it is coded in such a way that it is insensitive to rounding errors. Then even slight extrapolations (not more than 5% of the interval length) may speed up the convergence considerably.

The whole procedure is terminated if one of the following conditions is satisfied:

- (i) Crossing the rectangular margin; then the last point should be located on the latter.
- (ii) Crossing an internal boundary; this is most easily discovered if the sign of the distance from the boundary changes.
- (iii) Cyclic closure: the line runs again into the interval between points 0 and 1.

Some experience is necessary to choose the accuracy parameters and the controls of the step size h in such a way that the curve plots have a reasonable and smooth shape without too many evaluations of the field functions.

7.7.4 Magnetic Flux Lines

The electrostatic potential $V(\mathbf{r})$ has a certain physical meaning as the potential energy of a particle with unit charge, located at the position $\mathbf{r}$. In contrast to this, the different potentials, used in magnetostatics, have only a formal meaning, and it is therefore usual to illustrate magnetic fields in the form of flux lines. There are lines in two or three-dimensional graphs along which the vector $\mathbf{B}(\mathbf{r})$ has the direction of the local *tangent*. This definition leads immediately to the differential equation

$$\frac{d\mathbf{r}(\tau)}{d\tau} = \lambda(\mathbf{r}, \tau) \mathbf{B}(\mathbf{r}), \quad (7.162)$$

τ being any suitable curve parameter and $\lambda(\mathbf{r}, \tau)$ an arbitrary but regular factor. The *geometric* form of the line is already uniquely given by the choice of the starting point $\mathbf{r}(0)$; different parameters τ or factors λ lead only to different formal representations of the same curve. The numerical solution of Eq. (7.162) can be carried out by such well-known techniques as, for example, the Runge Kutta method; these are not outlined here.

The result is a graphical illustration of the spatial distribution of *directions* of the magnetic field. By tracing many such lines their local *density* should also provide a measure for the strength $|\mathbf{B}|$ of the field. However, both requirements cannot be combined in the same plot. This will be demonstrated for the special case of rotationally symmetric fields.

We now recall Eqs. (2.42) and introduce them on the right-hand side of Eq. (7.162), after the latter has been written down in cylindrical coordinates:

$$\begin{aligned} \dot{z}(\tau) &= \lambda (2\pi r)^{-1} \partial \Psi / \partial r, \\ \dot{r}(\tau) &= -\lambda (2\pi r)^{-1} \partial \Psi / \partial z. \end{aligned} \quad (7.163)$$

From these we can conclude that

$$\begin{aligned} \dot{r} \partial \Psi / \partial r + \dot{z} \partial \Psi / \partial z &= 0, \\ \Psi(z, r) &= \text{const.}; \end{aligned} \quad (7.164)$$

hence the equipotentials of the flux potential are solutions of Eq. (7.162).

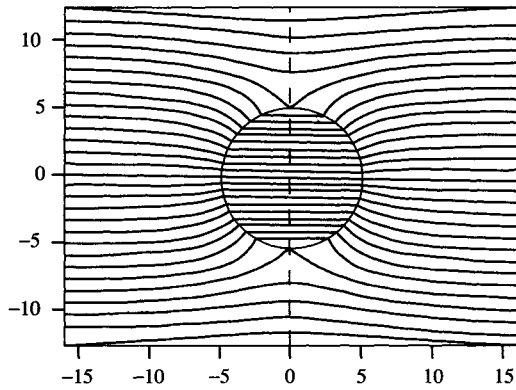
On the other hand, the amplitude $A(z, r)$ of the vector potential is also frequently used essentially in FEM programs for magnetic lenses, and it is then appropriate to produce equipotential plots of this amplitude. However, owing to the different radial factors in Eqs. (2.41 and 2.42), these must have another shape.

To show the difference, we consider here a simple analytical model, a sphere with radius a and relative permeability μ in a homogeneous external field of

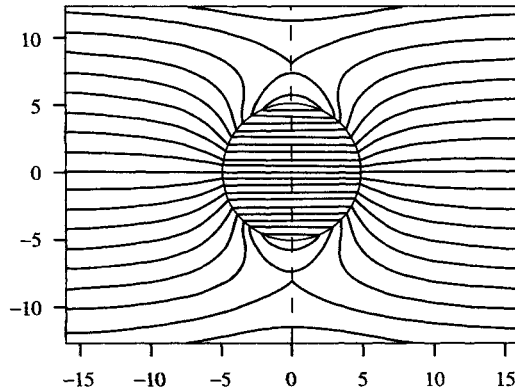
strength B_∞ . The well-known analytical solution is given by

$$\Psi(z, r) = \begin{cases} \frac{3\mu}{\mu + 2} \pi B_\infty r^2, & (r \leq a) \\ \pi B_\infty r^2 \left(1 + \frac{2(\mu - 1)a^3}{(\mu + 2)(z^2 + r^2)^{3/2}} \right), & (r \geq a). \end{cases} \quad (7.165)$$

This expression remains finite as $\mu \rightarrow \infty$. Asymptotically we always have $B_z \rightarrow B_\infty$ and for $\mu \rightarrow \infty$ the constant field $B_0 = 3B_\infty$ in the interior; moreover, the flux lines must be orthogonal to the surface of the sphere on its outer side.



(a)



(b)

FIGURE 7.34 A sphere of radius $r = 5$ units and very large permeability ($\mu = 10^5$) in an asymptotically homogeneous external $\mathbf{B}$ -field: (a) equipotentials of the flux potential ψ ; (b) equipotentials of the vector potential amplitude A .

The result of such a plot is shown in Fig. 7.34a. To simulate homogeneous fields by lines of constant density, it is necessary to choose constants $\Psi_n = \text{const} \cdot n^{1/2}$ with $n = 0, 1, 2, \dots$. The orthogonality on the surface is then respected, two singular points being an exception; however, the ratio of densities is $\sqrt{3}$ instead of 3: the plot shows the correct directions but *not* the real strengths.

In contrast to this graph, the second one shows equipotentials of the amplitude $A(z, r)$. This is related to Eq. (7.165) by

$$A(z, r) = \Psi(z, r) (2\pi r)^{-1}. \quad (7.166)$$

An asymptotically homogeneous field is now obtained by choosing $A_n = \text{const} \cdot n$, n being any positive or negative integer. The ratio of line densities is here $B_0/B_\infty = 3$, as is expected, but the lines are not at all orthogonal to the surface on the outer side, and this means that these lines do *not* have the direction of the $\mathbf{B}$ -field as is obvious from Fig. 7.34b. The most interesting features are the two saddle points at $r = \pm a \cdot 4^{1/3}$, which are clearly extrema of the function

$$A(0, r) = \pm B_\infty (r/2 + a^3 r^{-2}) \quad (r \geq a). \quad (7.167)$$

This simple example demonstrates very convincingly that the correct interpretation of graphs showing magnetic field distributions is not at all a trivial task and must be made with great care.

REFERENCES

1. Kasper, E. (1984). Improvements of methods for electron optical field calculations, *Optik* **68**: 341–362.
2. Lencova, B. and Lenc, M. (1984). The computation of open electron lenses by coupled finite element and boundary integral methods, *Optik* **68**: 37–60.
3. Jung, H., Lee, G. and Hahn, S. (1984). 3-D magnetic field computations by infinite elements, *J. Appl. Phys.* **55**: 2201–2203.
4. Weth, A.v.d. and Becker, R. (1999). A hybrid method using BEM and FDM for the calculation of 2-D magnetic fields including materials with field-dependent permeability, *Nucl. Instrum. Meth. A* **427**: 399–403.
5. Weth, A.v.d. (1997). Berechnung statischer axialsymmetrischer Magnetfelder mit der Methode der Randlelemente (BEM) unter Berücksichtigung nicht-linearer Materialeigenschaften, Dissertation, Universität Frankfurt (Main), Fachbereich Physik, Germany.
6. Kasper, E. 2000 An advanced boundary element method for calculation of magnetic lenses, *Nucl. Instrum. Meth. A* **450**: 113–178.
7. Killes, P. (1985). Solution of Dirichlet problems using a hybrid finite differences and integral-equation method applied to electron guns, *Optik* **70**: 64–71.

8. Killes, P. (1988). Neuartige Verfahren zur Berechnung rotationssymmetrischer Elektronenstrahlerzeugungssysteme unter Berücksichtigung der Raumladung, Dissertation, Universität Tübingen, Fakultät für Physik, Germany.
9. Degenhardt, R. and Berz, M. (1999). High-accuracy field description of particle spectrographs, *Nucl. Instrum. Meth. A* **427**: 151–156.
10. Eupper, M. (1982). Eine Methode zur Lösung des Dirichlet-Problems in drei Dimensionen und ihre Anwendung auf einen neuartigen Elektronenstrahlerzeuger, *Optik* **62**: 299–307.
11. Schönecker, G., Spehr, R. and Rose, H. (1990). Fast charge-simulation procedure for planar and simple three-dimensional electrostatic fields, *Nucl. Instrum. Meth. A* **298**: 360–376.
12. Read, F. H., Bowring, N. J., Bullivant, P. D. and Ward, R. R. A. (1999). Short- and long-range penetration of fields through meshes, grids or gauzes, *Nucl. Instrum. Meth. A* **427**: 363–367.
13. Takaoka, A., Shin-ya, Y. and Ura, K. (1985). An improvement of the charge simulation method in electron optical systems, *Optik* **69**: 166–171.
14. Kasper, E. (1986). A simple method of field calculation in electron optical computer simulations, In *Proceedings of Eleventh International Congress on Electron Microscopy*, Kyoto, Japan, 285–286.
15. Hoch, H., Kasper, E. and Kern, D. (1978). Darstellung stationärer rotationssymmetrischer elektromagnetischer Felder durch Superposition von Lochblenden- und Kreisringfeldern, *Optik* **50**: 413–425.
16. Kasper, E. (1982). Field electron emission sources, *Adv. Opt. Electron Microsc.* **4**: 207–260.
17. Kasper, E. (1979). On the numerical field calculation in field emission devices, *Optik* **54**: 135–147.
18. Kasper, E. (1981). Numerical design of electron lenses, *Nucl. Instrum. Meth.* **187**: 175–180.
19. Hawkes, P. W. and Kasper, E. (1989). *Principles of Electron Optics*, Volume 1, Chapter 10.3. London & New York: Academic Press.
20. Möllenstedt, G. and Lenz, F. (1963). *Adv. Electron. Electron Phys.* **18**: 251–329.
21. Schwertfeger, W. and Kasper, E. (1974). Zur numerischen Berechnung elektromagnetischer Multipolfelder, *Optik* **41**: 160–173.
22. Scherle, W. (1983). Eine Integralgleichungsmethode zur Berechnung magnetischer Felder von Anordnungen mit Medien unterschiedlicher Permeabilität, *Optik* **63**: 217–226.
23. Schaefer, C. H. (1982). Methoden zur numerischen Lösung der Laplacegleichung bei komplizierten Randwertaufgaben in drei Dimensionen und ihre Anwendung auf Probleme der Elektronenoptik, Dissertation, Universität Tübingen, Fakultät für Physik, Germany.
24. Schaefer, C. H. (1983). The application of the alternating procedure by H. A. Schwarz for computing three-dimensional electrostatic fields in electron-optical devices with complicated boundaries, *Optik* **65**: 347–359, and in *Electron Optical Systems*, J. J. Hren et al. (ed.), Scanning Electron Microscopy, Chicago. (1984) pp. 85–89.
25. Weysser, R. (1981). Numerische Berechnung elektronenoptischer Eigenschaften elektrostatischer Zwölfpolablenksysteme, unpublished work, Tübingen.
26. Kern, D. (1978). Theoretische Untersuchungen an rotationssymmetrischen Strahlerzeugungssystemen mit Feldemissionsquelle, Dissertation, Universität Tübingen, Fakultät für Physik, Germany.
27. Weysser, R. (1983). Feldberechnung in rotationssymmetrischen Elektronenstrahlerzeugern mit Spitzenkathode und Raumladungen, *Optik* **64**: 143–156.
28. Scherle, W. (1983). Berechnung von magnetischen Ablenssystemen, Dissertation, Universität Tübingen, Fakultät für Physik, Germany.
29. Kasper, E. (1991). An advanced method for the direct calculation of electron optical aberration discs, *Optik* **89**: 23–30.

30. Barth, J. E., Lencova, B. and Wisselink, G. (1990). Field evaluation from potentials calculated by the finite element method: the slice method, *Nucl. Instrum. Meth. A* **298**: 263–268.
31. Janse, J. (1971). Numerical treatment of electron lenses with perturbed axial symmetry, *Optik* **33**: 270–281.
32. Lei Wei and Tu Yan (1999). Determination of the acceptable tolerance in the manufacture of electron optical system, *Nucl. Instrum. Meth. A* **427**: 393–398.

Appendix

Tables for Numerical Calculations

TABLE A1
POSITIONS (P_K, Q_K) AND WEIGHTS W_K FOR NUMERICAL INTEGRATIONS OVER THE UNIT
TRIANGLE, TO BE APPLIED IN EQ. (5.43)

k	p_k			q_k			w_k	$M = 7$
1	0.3333	33333	3333	0.3333	33333	3333	0.1125	00000 0000
2	0.4701	42064	1051	0.4701	42064	1051	0.0661	97076 3943
3	0.0597	15871	7898	0.4701	42064	1051	0.0661	97076 3943
4	0.4701	42064	1051	0.0597	15871	7898	0.0661	97076 3943
5	0.1012	86507	3235	0.1012	86507	3235	0.0629	69590 2724
6	0.7974	26985	3531	0.1012	86507	3235	0.0629	69590 2724
7	0.1012	86507	3235	0.7974	26985	3531	0.0629	69590 2724
k	p_k			q_k			w_k	$M = 13$
1	0.0651	30102	9022	0.0651	30102	9022	0.0266	73617 8044
2	0.8697	39794	1956	0.0651	30102	9022	0.0266	73617 8044
3	0.0651	30102	9022	0.8697	39794	1956	0.0266	73617 8044
4	0.3128	65496	0049	0.0486	90315	4253	0.0385	56880 4452
5	0.6384	44188	5698	0.3128	65496	0049	0.0385	56880 4452
6	0.0486	90315	4253	0.6384	44188	5698	0.0385	56880 4452
7	0.6384	44188	5698	0.0486	90315	4253	0.0385	56880 4452
8	0.3128	65496	0049	0.6384	44188	5698	0.0385	56880 4452
9	0.0486	90315	4253	0.3128	65496	0049	0.0385	56880 4452
10	0.2603	45966	0790	0.2603	45966	0790	0.0878	07628 7166
11	0.4793	08067	8419	0.2603	45966	0790	0.0878	07628 7166
12	0.2603	45966	0790	0.4793	08067	8419	0.0878	07628 7166
13	0.3333	33333	3333	0.3333	33333	3333	-0.0747	85022 2339

TABLE A2

LOGARITHMIC CHEBYSHEV APPROXIMATION OF ELLIPTIC FOURIER INTEGRALS: SERIES EXPANSION COEFFICIENTS CORRESPONDING TO EQ. (6.112) FOR $0 \leq M \leq 12$ (COMPILED BY STRÖER, MARTIN (1988). THE INTEGRAL EQUATION METHOD FOR FIELD CALCULATIONS IN THREE DIMENSIONS AND ITS REDUCTION TO A SEQUENCE OF TWO-DIMENSIONAL PROBLEMS, OPTIK 81: 12–20.). FOR REASONS OF SPACE THESE COEFFICIENTS ARE ORDERED SEQUENTIALLY IN ROWS FOR EACH ORDER M IN TURN: FIRST AND SECOND ROW: COEFFICIENTS A_{M1} TO A_{M8} , THIRD AND FOURTH ROW: COEFFICIENTS B_{M1} TO B_{M8} .

$m = 0$			
1.386294361120072	0.096573604792707	0.030891446746570	0.015272858721217
0.012600888645968	0.016872000426597	0.010893399208991	0.001397767132554
0.500000000000000	0.124999998195220	0.070311302560537	0.048734007055494
0.035690124742644	0.020920433004995	0.005786112004658	0.000340500527793
$m = 1$			
0.386294361122069	0.039720946480868	0.013877685631877	0.011051646365682
0.052565071743731	0.150617095815989	0.115640062562669	0.015631293671909
0.500000000000000	0.374999978179412	0.351547901669162	0.340637386379290
0.315252185659005	0.212531071594955	0.062962910500374	0.003829331800319
$m = 2$			
0.052961027796581	-0.183797901720478	-0.201813988807319	-0.187570123935092
0.024309673970935	0.538258930762281	0.479144934497365	0.067556069971613
0.500000000000000	0.624999899011878	0.820244448203696	1.019932801281646
1.132336086011537	0.851250941576949	0.266205238713845	0.016626709046236
$m = 3$			
-0.147038972182182	-0.507315647358413	-0.690900325335614	-0.803273734343274
-0.275792140972901	1.348924611090542	1.366708788706388	0.199561272480832
0.500000000000000	0.874999683207112	1.476347550514833	2.238459272740019
2.879723996890914	2.363371470508255	0.772246403271758	0.049307570883882
$m = 4$			
-0.289896114992850	-0.902259867856210	1.506496699765093	-2.093130081595788
-1.162456809967748	2.765774782166597	3.144950839339827	0.473028573184726
0.500000000000000	1.124999207857373	2.319771468450815	4.145264988907702
6.028988728111818	5.319554694253942	1.802743365181324	0.117267352515358
$m = 5$			
-0.401007226015542	-1.352756129577286	-2.689610986487678	-4.336554546425374
-3.053884466587842	4.973834567232638	6.277898235007967	0.968643711198342
0.500000000000000	1.374998297880751	3.350392651263854	6.886055889146606
11.11637782354093	10.42482886709553	3.643576936056174	0.240832972738042

(continued overleaf)

TABLE A2
(CONTINUED)

$m = 6$			
-0.491916316772399	-1.848701545084303	-4.273450550739532	-7.827241663109319
-6.460884110703678	8.151147089185454	11.31957885613281	1.785817802705674
0.500000000000000	1.624996721748286	4.568045801620706	10.60275327481664
18.73011479996919	18.50505117774181	6.641484486762556	0.445142376916233
$m = 7$			
-0.568839393452216	-2.383100747157549	-6.285737549784954	-12.86738248589327
-11.97479783291880	12.46268934328316	18.90469923510796	3.041508309134890
0.500000000000000	1.874994188722359	5.972521117090509	15.43314765865546
29.49914937964538	30.49180487288699	11.19904051612891	0.759865032179401
$m = 8$			
-0.635506059752076	-2.950822680163688	-8.750098524338769	-19.76339459646274
-20.25558625762304	18.05633699701123	29.73878985216147	4.868755217104028
0.500000000000000	2.124990348090137	7.563562807279788	21.51065108198583
44.08345089242326	47.40735787301771	17.76897266948520	1.218844874759970
$m = 9$			
-0.694329588635809	-3.547942439087919	-11.68696610517039	-28.82271853889121
-32.02067149099778	25.06040894640033	44.58847982112482	7.415075900939003
0.500000000000000	2.374984789828158	9.340868687646134	28.96413704579285
63.16574965235668	70.35040248006382	26.84827936816972	1.859701903827639
$m = 10$			
-0.746961166851191	-4.171360229447272	-15.11419562791189	-40.35126209211272
-48.03409103425827	33.58258005695129	64.27143231036217	10.84062746268862
0.500000000000000	2.624977045072872	11.30408985047890	37.91781073056277
87.44413239284796	100.4818554916639	38.97184498304095	2.723369141851764
$m = 11$			
-0.794580213486634	-4.818565502465630	-19.04751265584654	-54.65158527049681
-69.09761195943462	43.70929441703581	89.64806725417400	15.31668316865398
0.500000000000000	2.874966588405372	13.45283182848557	48.49118723716045
117.6262957751210	139.0133005731470	54.70716416360714	3.853708603554693
$m = 12$			
-0.838058473068955	-5.487483125103609	-23.50083751051596	-72.02112535654775
-96.04035919596173	55.50711770343709	121.6107078133956	21.02321949585577
0.500000000000000	3.124952839203688	15.78665428359067	60.79899033410681
154.4232918354280	187.1925859238283	74.64676427142708	5.296884259848489

TABLE A3
 POSITIONS AND WEIGHTS FOR THE GAUSS QUADRATURE OVER
 LOGARITHMICALLY SINGULAR FUNCTIONS ACCORDING TO
 EQS. (6.133) AND (6.134) (COMPILED BY STRÖER, MARTIN (1987).
 EINE INTEGRALGLEICHUNGSMETHODE ZUR BERECHNUNG
 MAGNETISCHER FELDER VON ANORDNUNGEN MIT MEDIEN
 UNTERSCHIEDLICHER PERMEABILITÄT, OPTIK 77: 15–25.)

i	x_i	g_i
$n = 6$		
1	0.003025 80213 75463	0.011351 33881 72726
2	0.040978 25415 59506	0.075241 06995 49165
3	0.170863 29552 68773	0.188790 04161 54163
4	0.413255 70884 47932	0.285820 72182 72273
5	0.709095 14679 06286	0.284486 42789 14088
6	0.938239 59037 71671	0.154310 39989 37584
$n = 9$		
1	0.000711 07328 87084	0.002694 89114 90210
2	0.010194 53302 62549	0.019498 06475 26352
3	0.046833 86722 11245	0.057273 68794 91312
4	0.130367 83136 51315	0.111551 01434 87582
5	0.270447 25718 89118	0.167174 87686 32406
6	0.458319 45709 51280	0.203697 11869 47111
7	0.665444 63307 03512	0.203382 45316 41999
8	0.850116 72984 92690	0.158655 27983 01063
9	0.969977 04487 05807	0.076072 61324 81966
$n = 15$		
1	0.000418 79439 15881	0.001582 16996 94974
2	0.005903 80719 67481	0.011165 09016 49139
3	0.026460 76443 02601	0.031641 43000 00962
4	0.071654 63925 22795	0.059392 70284 48534
5	0.145256 11889 91253	0.087235 36748 54213
6	0.244193 49573 89433	0.109400 04049 06654
7	0.360887 45060 90450	0.121724 17387 97666
8	0.479916 57660 99632	0.110649 95004 32181
9	0.577544 59034 88119	0.090780 31814 89713
10	0.667020 85592 77383	0.071523 66777 17527
11	0.719305 90870 54335	0.067327 14853 49943
12	0.815643 92531 46289	0.102903 67809 82828
13	0.895722 54422 78316	0.047896 59493 30462
14	0.937176 91967 75713	0.052228 39893 54531
15	0.986253 44087 18497	0.034549 26869 90674

INDEX

Note: Boldface numbers indicate illustrations; italic *t* indicates a table.

- Acceleration step in charge simulation
 - method (CSM), 377
- Adaptive integration,, 110–111, 333
- Additive contribution, nine-point
 - configurations and FDM, 140
- Affine deformations, 122–124, 222
- Aitken acceleration, 292–293, 298, 299–300
- Alternation method (*see* General alternation method)
- Ampère's law, 18, 329
- Analytic continuation, 50–51, 55–57
- Analytic functions, 53, 54–55, 99
- Anisotropic materials, 3, 11
- Anode tube and charge simulation method (CSM), 371
- Antisymmetric potentials, 56, 389
- Aperture plates, 370, 377–387, **378**, **381**, **382**
- Approximation error in interpolation, 81–82
- Approximation in one mesh, nine-point
 - configurations and FDM, 135–137, **135**
- Area coordinates, *2t*, 204–205, **206**, **207**
- Area elements in orthogonal coordinate system, 61
- Asymmetric integral equations using special BEM techniques, 321–335
 - azimuthal component in magnetic lenses, 322, 332
 - complete field calculation in, 331–335
 - condensation of rings in, 333–334, **334**
 - conservation of total lens current and, 329–331
 - deflection systems for, 323–326, **324**
 - Dirichlet problems and, 331–333
 - discretization in, 333
 - fast method for, 326–329, 328*t*, 332–333
 - flux potential for, 331
 - footpoints in, 334–335, **335**
 - Fourier components in, 332
 - Fourier-Green function in, 325
 - Fourier-series expansions in, 322
 - Fredholm's equation for, 323
 - Galerkin method for, 331
 - interpolation on the surface normal and, 334–335, **335**
 - Maclaurin formula in, 333–334
 - principle value in, 333
 - surface sources in, 331
 - tangential components in, 322, 332
- Axial derivatives, 293
- Axial line charges, charge simulation method (CSM), 371
- Axial potential, 48, 50
- Axial vertices, 320
- Axial wires, 370
- Axially symmetric integral equations and BEM, 284–301
 - Aitken acceleration in, 292–293, 298–300
 - axial derivatives and paraxial expansion in, 293
 - Chebyshev approximation in, 298, 299, 300*t*
 - Cholesky algorithm in, 287
 - collocation in, 286–287
 - cylindric coordinates in, 284
 - differential equations and integral properties in, 288–289
 - flux kernel in, 297
 - Fourier analysis of, 284–288
 - Fourier-Green function properties in, 288–301
 - Fourier-series expansions in, 285–286
 - Fredholm's equation in, 286, 287
 - Galerkin's method in, 286, 287–288
 - Gauss-Stokes theorem in, 289
 - Green's function in, 285, 286, 288
 - iterated arithmetic in, 298
 - Landen transform in, 292
 - Legendre form in, 290
 - moduli and elliptic integrals in, 289–292
 - numerical calculations in, 297–301
 - parametrization in, 286
 - partial derivatives in, 293–295
 - Poisson equation in, 289
 - power series expansion in, 298
 - recurrence formula in, 298, 300–301
 - rotationally symmetric, 284
 - series expansion in, 292–293

- Axially symmetric integral equations and BEM, (*cont.*)
 - singularity analysis in, 295–297
 - symmetries in, 288
 - Taylor-series expansion in, 296–297
 - trial functions in, 287
- Azimuthal component in round lenses, 322, 332
- Azimuthal Fourier-series expansions, 31–35
- Barycentric coordinates (*see* Area coordinates)
- Basic mathematical tools, 59–114
- Bernoulli numbers, 112
- Bessel function, 50
- Bessel-Hermite quadratures, 109–110
- Bicubic splines, 106–107
- Bilinear trial functions, first-order finite element method (FOFEM), 223
- Biot-Savart's law, 25, 335
- Bisection method, in field interpolation, 231–232
- Bivariate Hermite interpolation, 104–106, 105, 213
- Block-elimination, in solving integral solutions using BEM, 307
- Bode's rule, 111
- Boundary conditions, 13, 15–19, 21, 22
 - in interpolation, 77–78, 237
 - in magnetic round lenses, 43–45
- Boundary element method (BEM), xii, 86, 263–356, 403
 - adaptive integration and, 333
 - advantages and disadvantages of, 263
 - Aitken acceleration in, 292–293, 298–300
 - asymmetric integral equations, special techniques for, 321–335
 - axial derivatives and paraxial expansion in, 293
 - axial vertices and, 320
 - axially symmetric integral equations, 284–301
 - Biot-Savart's law and, 335
 - Chebyshev approximation in, 298, 299, 300*t*
 - Cholesky algorithm in, 280, 281, 287, 307, 320–321
 - in coils with rectangular cross sections, 340–344
 - collocation in, 265, 282, 286–287, 302–307
 - condensation of rings in, 333–334, 334
 - conservation of total lens current and, 329–331
 - Coulomb integrals for rods and triangles, 267–276, 268, 269, 270, 274, 275, 278, 280, 335
 - cyclical notation in, 273–274, 281
 - deflection coefficients and, 349, 350
 - in deflection systems, integral equation for, 323–326, 324
 - in deflection systems, magnetic fields of, 344–349, 345
 - in deflection systems, special cases, 349–350
 - dipole approximation in, 278–279, 279*t*
 - Dirichlet problem and, 267, 315–317, 331–333, 339, 350–352
 - discretization in, 263, 264–284, 279–281, 333
 - Euler-Maclaurin formula in, 321, 348–349
 - evaluation of particular integrals in external fields, 335–337
 - external field calculations, 335–350
 - far-field approximation in, 276–279
 - fast field calculation and, 408, 413
 - fast method for asymmetric integral equations using, 326–329, 328*t*, 332–333
 - fast method for symmetric integral equations using, 311–315, 312, 315*t*
 - field interpolation vs., 263
 - finite-difference method (FDM) and, 361–367
 - finite-element method (FEM) and, 279, 280, 357–361
 - flux kernel in, 297
 - footpoint in, 273–274, 270, 274, 334–335, 334
 - Fourier analysis of integral equations using, 284–288
 - Fourier-Green function in, 288–301, 315, 325, 346–347, 351
 - Fourier-series expansions in, 285–286, 322, 345–346, 352
 - Fredholm's equation in, 264–265, 286, 287, 323

- Galerkin method and, 265–267, 277, 279, 280, 282, 286–288, 312, 308–311, 331
- Gauss quadratures and, 275–276, 303, 309, 311, 342–343
- Gauss-Stokes theorem in, 289
- general alternation method (GAM) and, 400, 407
- global vs. local coordinates in, 272–273, 278
- Green's function in, 281–285, 286, 288, 335–337, 352–353
- Gauss quadrature in, 278
- Hermite polynomials in, 304, 308
- integral equations in, 263
- interpolation on the surface normal and, 334–335, 334
- iterative methods, 307
- Landen transform in, 292, 352
- Laplace equation and, 267, 336
- line charge density in, 350
- Maclaurin formula in, 333–334
- mirror symmetry and, 318
- natural boundary conditions and, 264
- nonlinear material properties and, 263
- normal derivative in, 281–284
- numerical calculations in, 297–301
- numerical solutions of integral equations using, 301–321
- parametrization in, 286
- paraxial domains and, 343–344
- partial derivatives in, 293–295
- planar fields and, 350–351
- Poisson equation in, 289, 336–337
- power series expansion in, 298
- principle value in, 333
- recurrence formula in, 298, 300–301
- rotationally symmetric boundaries and, 264, 337–339, 338
- round lenses and, 322–323
- several closed curves and, 318
- sharp edges and, 318–320, 319
- simplified field calculation in, 320–321
- singularity analysis in, 295–297, 351
- step functions in, 304
- surface sources and, 339
- symmetric integration in, 303–304, 304r
- symmetries in, 280
- symmetrized kernel in, 280, 318
- Taylor-series expansion and, 353
- three-dimensional configurations and, 263
- trial functions in, 287
- triangular element analysis and, 204
- uncoupled integral equations in, 349
- vector potentials and, 336–337
- wave fields and, 351–354
- Boundary potentials, 25–26
- Cartesian coordinates, 2r, 4, 10, 13, 51
- Cathode tube and charge simulation method (CSM), 371–373
- Cauchy-Riemann equations, 52–54, 121–122
- Centroids, discretization of integral equations using BEM, 276–278, 278
- Charge simulation method (CSM), 367–387
 - acceleration step in, 377
 - anode tube and, 371
 - axial line charges, 371
 - cathode tube and, 371–373
 - charged aperture plates and, 377–387, 378, 381, 382
 - Cholesky algorithm and, 369, 375, 377
 - collocation in, 368
 - complete field in, 374–377, 376
 - coupling in, 375
 - Dirichlet problems in, 367, 368, 384, 385
 - equipotential surfaces and, 377
 - field-free spaces in, 384
 - Lamé coefficients in, 379
 - Laplace equation in, 368, 372, 384
 - least squares fit (LSQ) approximation, 368–369
 - Legendre functions in, 373, 379–380
 - Neumann conditions and, 368
 - normalized radial functions in, 381
 - paraxial series expansion in, 372
 - pointed cathode models in, 369–377, 370
 - ring and plate fields in, 386–387, 386
 - ring charges, 382
 - surface charges, 374
 - Wehnelt electrodes and, 384, 385
- Charged aperture plates, CSM and, 382–387, 378, 381, 382
- Charged particle optics, xii
- Chebyshev approximations, 298, 299, 300r, 409, 428–430r

- Cholesky algorithm in axially symmetric integral equations and BEM, 287
- in charge simulation method (CSM), 369, 375, 377
- in discretization of integral equations using BEM, 266, 280, 281
- in large systems of equations, using FEM, 243–244, 249, 259
- in numerical solution of integral solutions using BEM, 307, 320–321
- Chord lengths, 99
- Closed domain with two inner exclusions, rotationally symmetric boundaries, **38**
- Closed outer surfaces, 15
- Coefficient transformation, 7–8
- Coils, 17, 18, 36
 - current simulation model and, 391–393, **391**
 - deflection systems and, integral equation for, 323–326, **324**
 - FEM and BEM combination for, 357–361, **358**
 - with rectangular cross sections, 340–344
- Collocation
 - in axially symmetric integral equations and BEM, 286–287
 - in charge simulation method (CSM), 368
 - in discretization of integral equations using BEM, 265, 282
 - in numerical solution of integral solutions using BEM, 302–307, **302**
- Complete mesh formula, 137–141
- Complex logarithm and exponential, planar fields, 55
- Complex powers, 54
- Condensation of rings, special BEM techniques, 333–334, **334**
- Conductivity (κ), 2*r*, 3, 12–13
- Confined derivatives, 48
- Conformal mapping
 - of nine-point configurations using FDM, 141
 - of orthogonal meshes, 121–122
 - of planar fields, 52–54, **53**
- Conical shanks, FDM and BEM combinations, 363–364, **364**
- Conjugate gradient method, 120, 247–249, **248**
- Conservation of total lens current, 329–331
- Converging series expansion, 47
- Coordinate systems, 2*r*, 59–74
- Corner singularities, 170–172, **171**
- Corners, rounding off, 101–102, **101**
- Coulomb integrals, Coulomb potentials, xii, 20–22
 - in charge simulation method (CSM), 369
 - in discretization of integral equations using BEM, 267–276, **268**, **269**, **270**, **275**, 280
 - in external field calculations, 335
 - in FDM and BEM combinations, 361
 - in spherical meshes and, 165
- Coupling in charge simulation method (CSM), 375
 - in Fourier-series expansions, 35
 - in general alternation method (GAM) and, 402–403, **403**
- Cubic Hermite interpolation, 80, **81**, 82, 95, 100, 223, 420–421, **421**
- Curl operator, discretization of, 63–64, **64**, 67
- Current density (*j*), 3, 9, 10, 12, 34
- Current simulation model, 387–397
 - antisymmetric properties and, 389
 - comparison with correct calculations and, 395–397
 - cylindrical coils and, 391–393, **391**
 - ferromagnetic walls and, 389–391, **391**
 - Fourier-series expansion and, 394
 - general method, 393–395
 - local properties, 389–391
 - magnetic mirror properties, 387–389, **388**
 - orthonormality of vectors in, 394
 - permeability in, 388, 391
 - saddle coils and, 393, 395, **396**, 397
 - surface current density, 390, 392
 - toroidal coils and, 394–395, **396**, 397
- Current-free domains, 6
- Curve parameter, rotationally symmetric boundaries, 37–38
- Curves, mathematical representation of, 96–102
 - analytical functions, 99
 - chord lengths, 99
 - cubic Hermite interpolation, 100
 - differential geometrical functions and, 97–98

- double points and corners in, 97, 101–102
- Frenet's equation and, 98
- Hermite interpolation and, 97, 98, 101–102, **101**
- normal, binormal and torsion in, 98
- parametric form of, 97 planar, 98
- quintic spline in, 100
- rounding off corners in, 101–102, **101**
- sampling array determination for, 98–101
- smoothing, 97, 101–102, **101**
- splines in, improved, 100
- velocity in, 97
- Curvilinear coordinates, vectors in, 61–62
- Curvilinear triangles, FEM and, 210
- Cyclic distance, 314
- Cyclic local indexing, using BEM, 281
- Cyclical modulo, 133
- Cyclical notation, 217, 219, 273–274
- Cylindric coordinates, $2r$, 31–35, 284
- Cylindric Laplace operator, $2r$, 45
- Cylindrical coils, 391–393, **391**
- Cylindrical Poisson equation
 - in FDM, 145–167
 - correction of functional in, 157–158
 - discretization of, 149–156
 - flux potential, 157
 - implicit algorithm of, 158–159
 - ordinary differential equation (ODE)
 - in, 147, 149, 150, 156
 - radial discretization in, 147–149
 - radial power transform in, 156–157
 - spherical meshes and, 159–167
 - Taylor-series expansion and, 146, 147
- D'Alembert's wave equation, 15
- Damped electromagnetic waves, 13
- Deflection coefficients, 349, 350
- Deflection systems
 - external field calculations for, magnetic fields of, 344–349, **345**
 - integral equation for, 323–326, **324**
 - special cases for BEM in, 349–350
- Deformation of meshes, 143–145
- Dielectric coefficient (ϵ), 2, $2r$, 4, 7, 12, 15, 17, 22
- Dielectric constants, 23
- Differential equations, x
 - Fourier-Green function properties in, 288
 - homogenous, 47
 - magnetic round lenses and, 43
 - three-dimensional, 32
 - two-dimensional, 32, 35
- Differential forms, 62–69
- Differential geometrical functions, 97–98
- Differential operators, 33
- Differentiation in systems of triangles, 213–216, **214**
- Dimensionless coefficients, geometrical, 349
- Displacement vector (D), electric, 1, $2r$, 3
- Dipole approximation, 278–279, $279t$
- Dipoles, 20, 44
- Direct solution methods of large equation systems, 242–247
- Dirichlet problems, xiii, 17, 22–26
 - asymmetric integral equations using special BEM techniques, 331–333
 - charge simulation method (CSM), 367, 368, 384, 385
- Discrete Fourier analysis, 75
- Discretization of
 - integral equations using BEM, 267, 284
 - external field calculations, 339
 - fast field calculation, 417
 - FDM and BEM combinations, 362, 365
 - FEM and BEM combination for, 357
 - general alternation method (GAM), 397–400, 407
 - integral equations, 267, 284, 350–352
 - magnetic round lenses and, 45
 - numerical solution for, 315–317
 - planar fields and, 54
 - rotationally symmetric boundaries and, 39
 - spherical meshes and, 165
 - wave fields and, 28–29
- Discretization, discretization errors, xi, 8, 74
 - in asymmetric integral equations using special BEM techniques, 333
 - in boundary element method (BEM), 263
 - cylindrical Poisson equation in FDM and, 149–156
 - in integral equations using BEM, 279–281
 - in fast field calculation, 415
 - in FEM and BEM combination, 360–361
 - in finite-element method (FEM), 197

Discretization, discretization errors (*cont.*)

- in first-order finite element method (FOFEM), 219–222
 - in five-point configurations and FDM, 133
 - in general alternation method (GAM), 406
 - interpolation and, 87–88, 96
 - in irregular configurations and FDM, 178, 181–182, 184–185
 - in Maxwell's equations, 72–74, **73**
 - in orthogonal coordinate system, 63–67, **63**, **64**, **66**
 - radial, 147–149
 - in spherical meshes, 161–165, 161, 166†
 - Taylor series method for, 127–129
 - of variational principle, in FEM, 200–204
- Discretization of integral equations using BEM, 264–284
- centroids in, 276, 277–278, **278**
 - Cholesky algorithm and, 266, 280, 281
 - collocation in, 265, 282
 - complete procedure for, 279–281
 - Coulomb integrals in, 267–276, **268**, **269**, **270**, **275**, 280
 - cyclical notation in, 273–274, 281
 - dipole approximation in, 278–279, 279†
 - Dirichlet problem and, 267, 284
 - far-field approximation in, 276–279
 - finite-element method (FEM) and, 279, 280
 - footpoint in, 273–274, **274**
 - Fredholm's equation in, 264–265
 - Galerkin method and, 265–267, 277, 279, 280, 282
 - Gauss quadratures in, 275–276, 278
 - general methods for, 264–267
 - global vs. local coordinates in, 272–273, 278
 - Green's function in, 281–284
 - kernel in, 280
 - Laplace equation and, 267
 - multipole expansions, 276
 - natural boundary conditions and, 264
 - normal derivative of, 281–284
 - rotationally symmetric boundaries and, 264

singularity in, 267

symmetries in, 280

Displacement vector, electric (D), 4, 13, 16

Divergence operator, orthogonal coordinate system, 65–66

Divergence relations, 13

Double points and corners, 97, 101–102

Driving fields, 6

Edge or corner singularities, FDM and, 170–172, **171**

Electric material conditions, 16–17

Electrodes, 22

Electromagnetic field vectors, x

Electromagnetic potentials, 3–8

Electron guns, 236, 361

Electrostatic deflectors, general alternation method (GAM) and, 400, **401**

Electrostatic fields, 4, 53

Electrostatic potential (V), 4, 15, 22, 24, 32

Elliptic Fourier integrals, 289, 301, **300**, **301**

Elliptical partial differential equation, 15

Equipotential surfaces, charge simulation method (CSM), 377

Error polynomials, in fast field calculation, 411–412, **411**

Euler equations, 8, 120

Euler-Maclaurin formula, 111–113, 321, 348–349

Exploration vs. interpolation, 230–232

External field calculations, 335–350

Biot-Savart's law and, 335

coils with rectangular cross sections and, 340–344

Coulomb integrals in, 335

deflection coefficients, 349–350

in deflection systems, 344–350, **345**

Dirichlet problems and, 339

Euler-Maclaurin formula in, 348–349

evaluation of particular integrals in, 335–337

Fourier-Green function in, 346–347

Fourier-series expansions in, 345–346

Gauss quadratures and, 342–343

Green's function in, 335–337

Laplace equation and, 336

paraxial domains and, 343–344

Poisson equation in, 336–337

rotationally symmetric fields and, 337–339, **338**

surface current density in, 344

- surface sources and, 339
 - uncoupled integral equations in, 349
 - vector potentials and, 336–337
- Extrapolation, 92–94, **93**
- Far-field approximation, 276–279
- Fast BEM for symmetric integral equations, 311–315, **312**, 315_t
- Fast field calculation, 408–418
 - boundary element method (BEM) and, 408, 413
 - cubic Hermite polynomials, 420–421, **421**
 - Dirichlet problems and, 417
 - discretization in, 415
 - equipotential in FEM grids, 418–419, **418**, **419**
 - error polynomials in, 411–412, **411**
 - finite-difference method (FDM), 408
 - finite-element method (FEM), 408
 - Fourier-series expansions and, 408, 415, 417
 - general search algorithm for, 421–423, **422**
 - harmonic functions and, 416
 - Hermite interpolation, 413, 414, 415, 420–421, **421**
- interpolation and, 408, 412–417
 - intersection point determination using, 418–421, **419**
 - magnetic lenses and, 413–415, **414**
 - multipole-Laplace equations and, 412
 - Newton iterations in, 423
 - parameter variation in, 417
 - perturbations in, 417
 - radial interpolation, 409–412
 - radial-series expansion in, 413
 - ray tracing and, 417
 - Taylor-series expansions in, 412
 - three-dimensional interpolation, 415–417
 - two-dimensional interpolation in, 412–415
- Fast method for asymmetric integral equations using BEM, 326–329, 328_t, 332–333
- Ferromagnetic walls, current simulation model and, 389–391, **391**
- Ferromagnetic yokes, general alternation method (GAM) and, 407–408
- Field calculations, x, 1–3
- Field interpolation, 229–242
 - bisection method in, 231–232
 - boundary conditions in, 237
 - boundary element method (BEM) vs., 263
 - exploration vs., 230–232
 - form functions in, 239–240, **240**
 - Hermite interpolation in, 232–237, **235**, **240**
 - mesh position determination in, 230–232
 - paraxial interpolation in, 237–240, **238**
 - rectangular meshes, 232–233
 - Taylor-series expansion, 235
 - trigonal meshes and, 241–242, **241**
 - two-dimensional techniques in, 230
- Field strength, electric (E), 1, 2_t, 4, 13, 14, 15, 16, 17
- Field-free spaces, 384
- Finite difference approximations, 75
- Finite-difference method (FDM), xi, 115–191
 - affine distortions in 122–124
 - boundary element method (BEM) and, 361–367
 - Cauchy-Riemann equations in, 121
 - classification of configurations in, 126
 - complete mesh formula for, 137–141
 - conformal mapping in, 121–122, 141
 - correction of functional in, 157–158
 - cylindrical Poisson equation in, 145–167
 - deformation of meshes in, 143–145
 - discretization in, 149–156
 - edge or corner singularities, 170–172, **171**
 - exponentially expanding meshes, 122, **123**
 - finite-element method (FEM) and, 115, 227–228
 - five-point configurations, 127–134, **127**
 - general alternation method (GAM) and, 400, 402, 403
 - harmonic functions and, 177–181
 - hexagonal meshes in, 143, **143**
 - inner mesh points and, 167–170
 - irregular configurations, 167–185, **167**
 - Lagrange function in, 136
 - Laplace equation, 170–172, 177–182
 - Liebmann's method, 115

Finite-difference method (FDM) (*cont.*)

mesh points on boundaries of irregular materials in, 172–173, **173**

nine-point configurations, 134–145, **134**

Numerov formula and, 188

ordinary differential equation (ODE) in, 147, 149, 150, 156

orthogonal meshes and, 121–124, 141–142

partial differential equation (PDE) in, 115–126, 131, 135, 140, 144–145, 182–184

radial discretization, 147–149

radial power transform in, 156–157

regularization of meshes in, 143–145, **144**

rhombic meshes in, 142–143, **143**

ring-integral method for, 129–131, **130**

series expansions evaluations, 173–177

seven-point configurations and, 186–187, **186**, 190

Simpson's rule in, 136

spherical meshes, Poisson equation and, 159–167

subdivision of meshes in, 185–189, **186**, **188**

Taylor-series expansion and, 127–129, 136, 140, 146, 147, 155

trial functions in, 174, 177

two-dimensional meshes in, 115–126, **116**, **117**

variational principles in, 118–120, 118

Finite-element method (FEM), xi–xii,

193–261

bilinear Hermite interpolation in, 213

boundary element method (BEM) and, 204, 228, 357–361

Cholesky algorithm in, 243–244, 249, 259

completely external point triangulation in, 197

conjugate gradient method in, 247–249, **248**

differentiation in systems of triangles and, 213–216, **214**

direct solution methods of large equation systems, 242–247

discretization in, 197, 200–204, 279, 280

equipotential grids and FEM, 418–419, **418**, **419**

field interpolation and, 229–242

finite-difference method (FDM) and, 115, 227–228

first-order finite element method (FOFEM) in, 212, 216–228

form functions for paraxial interpolation in, 239–240, **240**

Galerkin method in, 200

Gauss-Cholesky algorithm in, 246

Gauss-elimination technique in, 242, 243

Gauss-Seidel method in, 249, 251

general alternation method (GAM) and, 400, 407

Hermite interpolation in, 232–237, **235**, **240**

Hesse matrix and, 259

hexagonal mesh generation in, 194–195, **195**

inner point triangulation, 197

integration over triangular domains in, 207–209

Jacobian determinants in, 249, 258

large systems of equations, solutions of, 242–259

limitation function for triangulation in, 198–199

magnetic lenses and, 223–228

mesh generation in, 193–200

Newton's multidimensional iteration method and, 258

nonlinear systems of equations and, 258–259

paraxial interpolation in, 237–240, **238**

partly external point triangulation in, 197–198, **198**

quadrature formulas/parameters for, 208–209, 208*r*, **208**

quadrilateral elements in, 213, **213**

quadrilateral meshes in, 194, **194**, 223

Rayleigh quotients in, 252–253

rectangular meshes, 232–233

regularization of triangulation in, 199–200, **200**

relaxation methods in, 249–253

second-order finite element method (SOFEM) in, 212, 216

- self-adjoint partial differential equations in, 216–219
- sparse matrices and, 244, **244**
- successive line overrelaxation (SLOR) in, 253–258, **251, 255**
- table of elements for triangulation in, 200
- table of nodes for triangulation in, 200
- Taylor-series expansion and, 258
- triangular (trigonal) mesh generation in, 194–195, **195**
- triangular (trigonal) mesh interpolation in, 241–242, **241**
- triangular element analysis in, 204–216, **204**
- triangulation of cylindrical surface for, 196, **196**
- two-dimensional indexing in, 195–196, **196**
- First-order finite element method (FOFEM), 212, 216–228
 - affine deformations and, 222
 - bilinear trial functions in, 223
 - closed vs. open lenses in, 226, **226, 228**
 - cyclicity of notation in, 217, 219
 - discretization coefficients in, 218–222, **219**
 - discretization errors in, 219–222
 - error analysis and improvements to, 220–222, **220**
 - five-point formula for, 220
 - Gauss quadratures and, 223
 - homogenous fields and, 219
 - homogenous vs. inhomogenous PDE in, 220–221
 - linear vs. nonlinear trial functions in, 219, 225
 - magnetic lenses and, 223–228
 - orthogonal structures in, 219
 - quadrilateral meshes and, 223
 - self-adjoint partial differential equations in, 216–219
 - source coefficients in, 221–222
 - tilted meshes and, 222
- Five-point configuration/formulas and FDM, 127–134, **127**
 - cyclical modulo and, 133
 - discretization error in, 133
 - first-order finite element method (FOFEM), 220
 - Gauss's theorem in, 129
 - generalization of method in, 133–134
 - irregular configurations and, 133
 - isotropic PDE in, 133
 - partial differential equation (PDE) in, 131
 - ring-integral method for, 129–131, **130**
 - rotational symmetry in, 128
 - Taylor series method for, 127–129
 - triangular cells and, 133
- Flux kernel, 297
- Flux lines, 424–426, **425**
- Flux potential
 - in asymmetric integral equations using BEM, 331
 - in cylindrical Poisson equation using FDM, 157
 - in magnetic round lenses, 40–42, **41, 42**
 - in series expansions, 49–50
- Footpoints, 72, 273–274, **274**, 334–335, **335**
- Form functions for paraxial interpolation, 239–240, **240**
- Form invariant PDE, 120
- Fourier analysis, 284–288, 332, 403
- Fourier integral, 56–57
- Fourier-Bessel expansions, 50–51
- Fourier-Green function
 - in asymmetric integral equations using BEM, 325
 - in axially symmetric integral equations using BEM, 288–301
 - in external field calculations, 346–347
 - in integral equations, 351
 - in numerical solution of integral solutions using BEM, 315
- Fourier-series expansion
 - in axially symmetric integral equations and BEM, 285–286
 - azimuthal, 31–35
 - coupling and, 35
 - current simulation model and, 394
 - in deflection systems, 345–346
 - in external field calculations, 345–346
 - in fast field calculation, 408, 415, 417
 - in integral equations, 352
 - in rotationally symmetric boundaries, 38–39
 - in vector fields, 34–35
- Four-point interpolation, 94

- Fredholm's equation, 23–25, 264–265, 286–287, 323
 Frenet's equations, 98
 Fringing fields, 51
 Galerkin method, 200
 in asymmetric integral equations using BEM, 331
 in axially symmetric integral equations and BEM, 286–288
 in discretization of integral equations using BEM, 265–267, 277, 279, 280, 282
 in FDM and BEM combinations, 365–367
 in numerical solution of integral solutions using BEM, 308–312
 Gauge, 5, 10
 Gauss-Choleksy algorithm, 246
 Gauss-elimination technique, 242, 243
 Gauss-Legendre quadrature, 108–109
 Gauss quadratures, 208, 223, 278, 430–431
 in discretization of integral equations using BEM, 278
 in discretization of integral equations using BEM, 275–276
 in external field calculations, 342–343
 in FDM and BEM combinations, 365
 in first-order finite element method (FOFEM), 223
 in numerical solution of integral solutions using BEM, 303, 309, 311
 Gauss-Seidel algorithm, 249, 251, 407
 Gauss-Stokes theorem, 289
 Gauss's theorem, 65, 120, 129
 General alternation method (GAM), 397–408, **398**
 boundary element method (BEM) and, 400, 403, 407
 coupling and, 402–403, **403**
 Dirichlet problems and, 397–400, 407
 discretization in, 406
 electrostatic deflectors and, 400, **401**
 FDM and BEM combination with, 400
 FEM and BEM combination with, 407
 ferromagnetic yokes and, 407–408
 finite-difference method (FDM) and, 400, 402, 403
 finite-element method (FEM) and, 400, 407
 Fourier analysis and, 403
 Gauss-Seidel algorithm and, 407
 intersecting sphere and box, 397, **397**
 iterative arithmetic and, 406
 Laplace equation and, 402
 Legendre functions and, 404
 LU algorithm, 406, 407
 nine-point formula and, 401
 nonoverlapping domains and, 407–408
 overlapping domains and, 397–398, **397, 398**
 permeability, 405
 Poisson's equation and, 398–400
 practical examples of, 400–404
 saddle coils and, 404–407, **404**
 series expansions and, 400–401
 six-point formula and, 401–402
 successive line overrelaxation (SLOR), 403
 successive-overrelaxation (SOR), 403
 systems with several different materials and, 404–407, **405**
 General coordinates, 21
 General search algorithm for interpolation, 421–423, **422**
 Gibbs phenomenon, 304
 Gradients, 22
 Green's function, 19, 20, 25–26
 in axially symmetric integral equations and BEM, 285, 286
 in discretization of integral equations using BEM, 281–284
 in external field calculations, 335–337
 in FEM and BEM combination, 358
 in integral equations, 352, 353
 in wave fields, 28, 351–354
 Half-integer values, 112–113, 311, 315
 Half-spheres in integral equations, 20–24
 Harmonic axial potential, 48–50, 52–56
 Harmonic functions, 177–181, 416
 Harmonic time dependence, 14
 Heap in interpolation, 83, 93
 Helmholtz equation, 15, 28–29
 Hermite interpolation, 77–82, **81**, 240
 Bessel-Hermite quadratures and, 109–110
 bicubic splines and, 106–107
 bivariate, 104–106, **105**, 213
 curves and, 97, 98
 fast field calculation and, 413–415, 420–421, **421**

- numerical solution of integral solutions
 - using BEM and, 304, 308
 - polynomials, 88
 - rectangular meshes and, 232–237, **235**, 232
 - in triangular element analysis in FEM, 209, 211–212, **211**, 216
- Hermite splines, 82–86, 101–102, **101**
- Hertz vector potential (Z), 12–14
- Hesse matrix, 259
- Hexagonal meshes, 143, **143**, 194–195, **195**
- Homogenous differential equation, 47
- Homogenous fields, first-order finite element method (FOFEM), 219
- Homogenous materials, 23
- Hybrid methods, xiii, 357–431
- Hysteresis curve, inverted $H(B)$, 11, **12**
- Ideal ferromagnetic materials, 44–45
- Implicit algorithm, 158–159
- Inflection points or natural boundary conditions, 83
- Inhomogenous equation (*see* Cylindric Poisson equation)
- Inhomogenous materials, 5
- Inner boundaries, 15–16
- Integral equations, x , 15, 22
 - asymmetric, special techniques for, 321–335
 - axially symmetric, using BEM, 284–301
 - boundary element method (BEM) and, 263
 - Dirichlet problems and, 350–352
 - discretization of, using BEM, 264–284
 - eigenvalues of, 354
 - for electrostatic fields, 19–25
 - Fourier analysis of, 284–288
 - Fourier-Green function and, 351
 - Fourier-series expansion and, 352
 - Green function and, 352, 353
 - Landen transform and, 352
 - line charge density in, 350
 - for magnetic fields, 25–28
 - mean values and, 354
 - numerical solution of (*see* Numerical solution of integral equations using BEM)
 - planar fields and, 350–351
 - scalar, 25–26
 - singularity and, 351
 - for surface sources, 24–25
 - symmetric, fast BEM for, 311–315, **312**, 315†
 - Taylor-series expansion and, 353
 - vector, 27–28
 - wave fields and, 28–29, 351–354
- Integration interval, 108
- Integration over triangular domains in FEM, 207–209
- Interpolation, xii, 74–86
 - accuracies, table of, 92†
 - approximation error in, 81–82
 - asymmetric integral equations using BEM, on the surface normal and, 334–335, **334**
 - basic relations in, 86–92
 - basic rules for, 74–76
 - bivariate Hermite, 104–106, **105**, 213
 - boundary conditions and, 77–78, 237
 - cubic interpolation, Hermite, 80, 81, 82, 95
 - discrete Fourier analysis and, 75
 - discretization and, 74, 87–88, 96
 - exploration vs., 230–232
 - extrapolation vs., 92–94, **93**
 - fast field calculation, 408
 - field, 229–242
 - finite-element method (FEM) and, 229–242
 - form functions in paraxial, 239–240, **240**
 - four-point, 94
 - heap in, 83
 - Hermite, 77–88, **81**, 211–212, **211**, 216, 232–237, **235**, 240, 413–415, 420–421, **421**, 420
 - inflection points or natural boundary conditions in, 83
 - kernel functions, 74, 86–96, **89**, **96**, 312, 327
 - Lagrange polynomials and, 75, 86, 88, **90**, 106, 212
 - modified kernels for (*see* Kernel functions)
 - Neville algorithm of, 75–76
 - Newton algorithm of, 75
 - nonequidistant intervals and, 94–96
 - normalized form functions for, 77
 - paraxial 237–240, **238**
 - periodic splines in, 83–86, **85**
 - polynomials and, 75

Interpolation, (*cont.*)

- radial, 409–412, **411**
 - rectangular meshes and, 232–233
 - recurrence algorithm in, 89–92
 - sampling for, 74, 86
 - splines for, 80, 82–86
 - symmetric functions and, 77, **238**
 - three-dimensional, 415–417
 - trial or test functions in, 74
 - trigonal meshes and, 241–242, **241**
 - two-dimensional, 230, 412–415
- Intersecting sphere, general alternation method (GAM) and, 397–398, **397**
- Intersection point determination, 418–421, **419**
- Inverse transformations, 72
- Inverted hysteresis curve $H(B)$, 11, **12**
- Irregular configurations and FDM, 167–185, **167**
- arbitrary configuration of nodes and space vectors in, 175–176, **175**
 - degenerate case of triangle in, 168–169, **168**
 - discretization in, 178, 181–182, 184–185
 - edge or corner singularities in, 170–172, **171**
 - five-point configurations in, 133
 - general method for, applications of, 181–184
 - harmonic functions and, 177–181
 - inner mesh points and, 167–170
- Laplace equation in, 170–172, 177–182
- material coefficients in, 172–173, 184
 - mesh points on boundaries of materials in, 172–173, **173**
 - nine-point configuration in, 169–170, **169**
 - partial differential equation (PDE) in, 182–184
 - Poisson equation in, 179–180, 183, 185
 - series expansions evaluations, 173–177
 - Taylor-series expansion in, 184–185
 - totally pivoting algorithm in, 178
 - trial functions in, 174, 177
- Isoparametric functions, 210–211, **210**
- Isotropic material coefficients, 11, 121
- Isotropic media, 11

Isotropic nonlinear media, Maxwell's equations and, 3, 5

Isotropic PDE, 122, 133, 141–142

Iterative techniques, 249–259, 307, 397–408

Jacobian determinants

- in large systems of equations, 249, 258
- in orthogonal coordinate system, 61, 65–66
- in orthogonal meshes, 121
- in triangular element analysis in FEM, 208, 211
- in two-dimensional meshes, 118

Kernel functions, 74, 88, **89**, 95–96, **96**, 264, 280, 285–286

Kronecker symbol, 20, 60

Lagrange function or density (*see* Lagrangian)

Lagrangian (L), 8–12, 119–120, 202, 224

Lagrange density (L), 8, 9, 10

Lagrange interpolation, 75, 86, 88, **90**, 106

Lamé coefficients, 69–61, 68, 70, 379

Landen transform, 292, 352

Laplace equation, 7, 15, 20, 27

- in charge simulation method (CSM), 368, 372, 384

- in discretization of integral equations using BEM, 267

- in external field calculations, 336

- in fast field calculation, 412

- in general alternation method (GAM) and, 402

- in irregular configurations and FDM, 170–172, 177–182

- in orthogonal coordinate system, 69

- in orthogonal meshes, 122

- in planar fields, 52–55, 57

Laplace operator (Δ), 2t, 4, 7, 13, 15, 17, 18, 19, 23, 29, 33, 34, 69

Large systems of equations, 242–259

- Cholesky algorithm in, 243–244, 249, 259

- conjugate gradient method in, 247–249, **248**

- direct solution methods in, 242–247

- Gauss-Cholesky algorithm in, 246

- Gauss-elimination technique in, 242, 243

- Gauss-Seidel method in, 249, 251

- Hesse matrix and, 259
- Jacobian determinants in, 249, 258
- Newton's multidimensional iteration method and, 258
- nonlinear systems of equations and, 258–259
- preconditioning in, 249
- Rayleigh quotients in, 252–253
- relaxation methods in, 249–253
- residuals in, 248
- sparse matrices and, 244, **244**
- successive line overrelaxation (SLOR) in, 253–258, **251, 255**
- successive-overrelaxation (SOR), 249–251, **251**
- Taylor-series expansion and, 258
- tridiagonal systems of splines and, 246
- well-conditioned matrices in, 242
- Least squares fit (LSQ) approximation, 368–36
- Legendre functions, 290, 373, 379–380, 404
- Lenses (*see also* Magnetic round lenses), xi, 27, 31, 36, 39–45, 223–230, 322–323, 329–332, 351–361
- Liebmann's method (*see* Finite-difference method)
- Limitation function for triangulation, 198–199
- Limitation processes (for Green's integral theorem), 20–21, 24–25
- Line charge density, 350
- Line elements, in orthogonal coordinate system, 59–61
- Linear trial or form functions, 209–210, **217**, 219, 225, 267–274
- Linear material equations, 23–24
- Linear media, 3, 11
- Local properties, current simulation model and, 389–391
- LU algorithm, 243, 406, 407
- Maclaurin formula, 111–113, 333–334, **334**
- Magnetic deflectors, xi, 36, 51, **51**
- Magnetic energy density, 11
- Magnetic excitation vector (H), 1, 2*t*, 3, 5, 6, 10, 11, 13, 17–19, **19**, 25, 26
- Magnetic field strength (B), 2*t*, 3, 5, 6, 9, 10, 11, 14, 15, 17, 27, 41–44
- Magnetic fields, 25–28, 322, 344–349, **345**
- Magnetic flux lines, 424–426, **425**
- Magnetic flux potential (ψ), 7
- Magnetic lenses, 39–45, **40**
 - asymmetric integral equations using special BEM techniques and, 322–323
 - boundary conditions in, 43–45
 - boundary element method (BEM) and, 228, **228**
 - closed vs. open lenses in, 226, **226**, **227**, **228**
 - conservation of total lens current and, 329–331
 - differential equations for, 43
 - Dirichlet problem and, 45
 - fast field calculation and, 413–415, **414**
 - FDM vs. FEM in, 227–228, **226**, **227**, **228**
 - FEM and BEM combination for, 357–361, **358**
 - first-order finite element method (FOFEM) for, 223–228
 - flux potential in, 40–42, **42**
 - ideal ferromagnetic materials in, 44–45
 - magnetic field strength (B), 41–42, 44
 - Maxwell equations for, 43
 - Neumann problem and, 44–45
 - Stokes's integral theorem and, 41–42
 - superconductors and, 45
 - variational principles in, 45
 - vector potential (A) in, 41
- Magnetic material conditions, 17–19
- Magnetic mirror properties, 387–389, **388**
- Magnetic refraction, 17, **18**
- Magnetic scalar potentials, 6–7, **6**, 41
- Magnetization (M), 1, 2*t*, 3, 9
- Material coefficients, discontinuity, 172–175, 220–222
- Material condition, 16–19, 23–24
- Maxwell equations, x, 1–4, 8, 13, 14, 43, 72–74, **73**
- Meissner-Ochsenfeld effect, 19
- Meridional section, rotationally symmetric boundaries, 37–38
- Mesh generation in FEM, 193–200
- Metallic electrodes, 22
- Method of meshes (*see* Finite-difference method)
- Mirror properties, 318, 387–389, **388**
- Modified interpolation kernels (*see* Interpolation kernel functions)

- Moduli and elliptic integrals, Fourier-Green function and, 289–292
- Multipole correctors, xi, 36
- Multipole expansions for charged triangles, 276
- Multipole-Laplace equations, 412 (*see also* Cylindric Poisson equation)
- Nabla operator transform, 62
- Natural boundary conditions, 20, 264
- Negative symmetry and extrapolation, 93–94
- Neumann problem, 28–29, 39, 44–45, 368
- Neville algorithm of interpolation, 75–76
- Newton algorithm of interpolation, 75
- Newton iterations, in fast field calculation, 423
- Newton-Cotes formulas and adaptive procedures, 110–111
- Newton's multidimensional iteration method, 258
- Nine-point configuration/formulas and FDM, xi, 134–145, **134**
 - approximation in one mesh for, 135–137, **135**
 - complete mesh formula for, 137–141
 - conformal mapping of, 141
 - deformation of meshes in, 143–145
 - general alternation method (GAM) and, 401
 - hexagonal meshes in, 143, **143**
 - irregular configurations and FDM and, 169–170, **169**
 - Lagrange function in, 136 (*see also* Lagrangian)
 - orthogonality in, 141–142
 - partial differential equation (PDE) in, 135, 140, 144–145
 - regularization of meshes in, 143–145, **144**
 - rhombic meshes in, 142–143, **143**
 - Simpson's rule in, 136
 - special cases of, 141–143
 - Taylor-series expansion in, 136, 140
- Nonequidistant intervals, 94–96
- Nonlinear material properties, 3, 11–12, **12**, 125, 224–225, 263, 360
- Nonlinear systems of equations, using FEM, 258–259
- Normal derivatives, 17
- Normal, binormal and torsion in curves, 98
- Normalized coefficients, in FOFEM, 218–219
- Normalized coordinate system, using BEM, 309–311, **310**
- Normalized form functions for Hermite interpolation, 77, 80, **81**
- Normalized radial functions for aperture plates, **381**
- Normalized vectors, 61–64
- Numerical calculations of elliptic Fourier integrals, 297–301, 428*t*
- Numerical differentiation, 74–86
- Numerical integration, 107–114
 - Bessel-Hermite quadratures in, 109–110
 - Euler Maclaurin formulas for, 111–113
 - Gauss-Legendre quadrature in, 108–109
 - integration interval in, 108
 - Maclaurin formulas for, 111–113
 - Newton-Cotes formulas and adaptive procedures in, 110–111
 - Simpson rule, extended, 112
- Numerical solution of integral equations
 - using BEM, 301–321
 - axial vertices and, 320
 - block-elimination in, 307
 - Cholesky algorithm in, 307, 320–321
 - collocation in, 302–307, **302**
 - cyclic distance in, 314
 - Dirichlet problems in, 315–317
 - Euler-Maclaurin formula in, 321
 - fast method for symmetric integral equations, 311–315, **312**, 315*t*
 - Fourier-Green functions in, 315
 - Galerkin method in, 308–312
 - Gauss quadratures in, 303, 309, 311
 - generalizations of boundaries in, 317–321, **317**
 - Gibbs phenomenon, 304
 - half-integer values in, 315
 - Hermite polynomials in, 304, 308
 - iterative arithmetic in, 307
 - mirror symmetry in, 318
 - modified interpolation kernel in, 312
 - normalized coordinate system for, 309–311, **310**
 - off-diagonal elements in, 313
 - several closed curves and, 318
 - sharp edges and, 318–320, **319**
 - simplified field calculation in, 320–321

- step functions in, **302**, **304**
 - symmetric integration in, **303–304**, **304t**
- Numerov formula, **1**, **149**, **188**
- Off-diagonal elements, using fast BEM, **313–317**, **326–329**
- Optic axis, **34**
- Optimization, ix–x
- Ordinary differential equation (ODE), in FDM, **147**, **149**, **150**, **156**
- Orthogonal coordinate systems, **59–74**
 - area elements in, **61**
 - curvilinear coordinates for, vectors in, **61–62**
 - differential forms of, **62–69**
 - discretization of curl operator in, **63–64**, **64**, **67**
 - discretization of divergence in, **65–66**, **66**
 - discretization of gradient in, **63**
 - discretization of Maxwell's equations, **72–74**, **73**
 - footpoint of surface in, **72**
 - Gauss's integral theorem and, **65**
 - inverse transformations in, **72**
 - Jacobian determinants in, **61**, **65–66**
 - Lamé coefficients in, **59–61**, **68**, **70**
 - Laplace operator in, **69**
 - line elements in, **59–61**, **59**
 - Nabla operator transform in, **62**
 - normalized vectors in, **61–64**
 - rotationally symmetric boundaries and, **71**
 - self-consistent approximations in, **67**
 - separation method in, **59**
 - spatial seven-point configuration of, **68**
 - surface-adapted coordinate system and, **69–72**, **70**
- Orthogonal meshes, **121–124**, **141–142**
- Orthogonal structures in FEM, **219**, **220**
- Overlapping domains, **285–287**, **286**, **397–398**, **397**, **398**
- Parameter variation, **417**
- Parametric form
 - of curves, **97**
 - of surfaces, **102**
- Parametrization, in BEM, **286**
- Paraxial domains, **343–344**
- Paraxial interpolation, **237–240**, **238**
- Paraxial series expansion, **48–49**, **57**, **293**, **372**
- Partial derivatives of the Fourier Green function, **293–295**
- Partial differential equation (PDE), **4–7**, **32–34**, **115–126**
 - in FDM and BEM combinations, **362**
 - in five-point configurations and FDM, **131**
 - in irregular configurations and FDM, **182–184**
 - in nine-point configurations and FDM, **135**, **140**, **144–145**
- Particle optics, xii, **34**
- Periodic splines, **83**
- Periodicity and extrapolation, **93–94**
- Permanent magnets, Maxwell's equations and, **3**
- Permeability, magnetic (μ), **2t**, **7**, **12**, **15**, **25**, **27**, **34**, **360**, **388**, **391**, **405**
- Perturbation theory, x
- Perturbations, in fast field calculation, **417**
- Planar curves, **98**
- Planar fields, **51–57**, **51**
 - analytic continuation in, **55–57**
 - analytic functions for, **52–55**
 - antisymmetric potentials in, **56**
 - boundary element method (BEM) and, **350–351**
 - Cartesian and polar coordinates of point in, **54**
 - Cauchy-Riemann equations in, **52–54**
 - complex logarithm and exponential in, **55**
 - complex powers and, **54**
 - conformal mapping of, **52–54**, **53**
 - Dirichlet problems and, **54**
 - Fourier integral in, **56–57**
 - Laplace equation in, **52–57**
 - paraxial series expansion in, **57**
- Planar structures of triangles, **205**
- Point charges (in CSM), **370**
- Pointed cathode models and CSM, **369–377**, **370**
- Poisson's equation, xiii, **4**, **19–20**, **23**
 - in axially symmetric integral equations and BEM, **289**
 - cylindrical, in FDM, **145–167**
 - in external field calculations, **336–337**
 - in FDM and BEM combinations, **361–362**

- Poisson's equation, (*cont.*)
 in FEM and BEM combination, 358
 in general alternation method (GAM), 398–400
 in irregular configurations and FDM and, 179–180, 183, 185
 in spherical meshes, 159–167
 in vector fields, 34–35
- Polarization (P), electric, 1, 2, 2*t*, 21, 22
- Polygons, 99
- Polynomial in paraxial series expansion, 47
- Positive symmetry and extrapolation, 93–94
- Potential density (p), 9
- Power series expansion for elliptic Fourier integrals, 292, 294, 298
- Preconditioning of large systems of equations, using FEM, 249
- Prescribed spline conditions, 83
- Principle values, 21, 333
- Prisms, magnetic deflection, 51, **51**
- Quadratic trial functions in triangles, 210–211
- Quadrature formulas for triangles, 208–209, 208*t*, **208**, 429*t*
- Quadrilateral meshes, 194, **194**, 213, **213**, 223
- Quintic spline, 84, 100
- Radial discretization, 147–149
- Radial interpolation, 409–412
- Radial power transform, 156–157
- Radial-series expansion, 413
- Radio frequency fields, x, 1
- Ray tracing, x, xii–xiii, 236, 276, 417, 361, 366
- Rayleigh quotients, 252–253
- Rectangular meshes, 103–104, **103**, 232–233
- Recurrence algorithm, 89–92, 298, 300–301
- Reduced scalar potential, 6, 25
- Reducible systems, 31–57
- Refraction, magnetic, 17, **18**
- Regularization of triangulation, 143–145, **144**, 199–200, **200**
- Relaxation methods, 249–253
- Reluctivity (ν), 3, 5, 12
- Repeated z-differentiations, 46–47
- Residuals, in large systems of equations, using FEM, 248
- Resonant cavities, 14, 28, 354
- Rhombic meshes, 142–143, **143**
- Ring and plate fields, 386–387, **386**
- Ring charges, 370, 382
- Ring-integral method, 129–131, **130**
- Rotated normal vectors in triangles, 205
- Rotational symmetry, xi, 31, 36–39, **36**
 in axially symmetric integral equations, 284
 of boundaries, 36–39, **36**
 of cavities, 351
 of closed domain with two inner exclusions, **38**
 curve parameter for, 37–38
 Dirichlet problem and, 39
 in discretization of integral equations using BEM, 264
 in FDM and BEM combinations, 361–367
 of fields, 337–339, **338**, 357–361, 369–387
 of five-point configurations and FDM, 128
 Fourier analysis of, 38–39
 mathematical form for, 37–38
 meridional section of, 37–38
 Neumann problem and, 39
 in orthogonal coordinate system, 71
 in sequences of boundary values, 39
- Round lenses (*see* Magnetic lenses)
- Rounding off corners curves, 101–102, **101**
- Runge Kutta method, 424
- Saddle coils, 349, 393, 395, **396**, 397, 404–407, **404**
- Sampling, 74, 86, 98–101
- Scalar integral equation, 25
- Scalar potentials in general, 2*t*, 6–7, 9
- Schwarz, H.A., general alternation method, 397 Second-order finite element method (SOFEM), 212, 216
- Self-adjoint equations, 4, 6, 7, 33, 43, 67, 216–219
- Self-consistent approximations, 67
- Separation method, 59
- Series expansions (*see also* Taylor-series expansions), x, 45–57
 axial potential and, 48, 50
 in axially symmetric integral equations and BEM, 292–293, 292
 Bessel function and, 50
 converging, 47

- flux potential and, 49–50
- Fourier-Bessel, 50–51
- general alternation method (GAM) and, 400–401
- homogenous differential equation and, 47
- inhomogenous equation and, 49
- in irregular configurations and FDM, evaluations of, 173–177
- paraxial, 48–49, 57
- polynomial solutions for paraxial, 47
- repeated z-differentiations in, 46–47
- symmetry conditions and, 46
- Seven-point configurations/formulas, **68**, 186–187, **186**, 190
- Sharp edges, boundary element method (BEM), 318–320, **319**
- Simplified field calculation, 320–321
- Simpson's rule, 112, 136
- Singularity, 20, 267, 295–297, 307, 313, 323, 326, 351
- Slit electrodes, 55–56, 56
- Smooth curves, 97
- Source coefficients in FOFEM, 221–222
- Source-free wave propagation, 13
- Space charge density (ρ), $2t$, 6–7, 9, 13
- Sparse matrices, 244, **244**, 304
- Speed of light (c), 13
- Spherical coordinates, $2t$, 159–161, 364
- Spherical meshes
 - Coulomb potentials in, 165
 - Dirichlet problems and, 165
 - discretization in, 161–165, 166*t*
 - FDM and BEM combinations and, 363–364, **364**
 - Poisson equation and, 159–167
 - source coefficients in, 161, 164
 - Taylor-series expansion in, 161, 163
- Spherical sectors, 20–21
- Splines, xi, 80
 - bicubic, 106–107
 - collocation using, in integral equations and BEM, 304–307
 - curves, 100
 - Hermite, 82–86, 101–102, **101**
- Standard notations of electrodynamics, $2t$
- Stationary (time-independent) fields, 3
- Step functions, 304
- Stokes's integral theorem for flux potential, 41–42
- Subdivision of meshes and FDM, 185–189, **186**, **188**
- Successive line overrelaxation (SLOR), 253–258, **251**, **255**, 403
- Successive-overrelaxation (SOR), 249–251, **251**, 403
- Superconductors, 18–19, 45
- Surface charge density (σ), 16, 22, 23, 374
- Surface current density (J), $2t$, 17–18, **18**, 344, 390, 392
- Surface sources, 15, 24–25, 331, 339, 357
- Surface-adapted coordinate system, 69–72, **70**
- Surfaces
 - mathematical representation of, 102–107, **102**
 - bicubic splines on, 106–107
 - bivariate Hermite interpolation in, 104–106, **105**
 - kernel functions on, 103
 - Lagrange interpolation on, 106
 - rectangular meshes for, 103–104, **103**
- Symmetric integration over singularities, 303–304, 304*t*
- Symmetry conditions
 - paraxial, 46
 - of Fourier Green function, 288
- Table of elements for triangulation, 200
- Table of nodes for triangulation, 200
- Tangential components of H-field, 17–19, 27–28, 322, 332
- Taylor-series expansion
 - in axially symmetric integral equations and BEM, 296–297
 - cylindrical Poisson equation in FDM and, 146, 147, 155
 - in fast field calculation, 412
 - field interpolation and, 235
 - in five-point configurations and FDM, 127–129
 - integral equations and, 353
 - in irregular configurations and FDM, 184–185
 - in large systems of equations, using FEM, 258
 - in nine-point configurations and FDM, 136, 140
 - spherical meshes and, 161, 163
- Tensor law, 119, 125
- Three-dimensional configurations, 263

- Three-dimensional differential equations, 32
- Three-dimensional interpolation, 415–417
- Three-slit lenses, 55–56, **56**
- Tilted meshes, 222
- Toriodal coils, 36, 323–326, **324**, 394–395, **396**, 397
- Totally pivoting algorithm, 178
- Trial functions, xii, 74, 86
 - in axially symmetric integral equations and BEM, 287
 - in FDM and BEM combinations, 366
 - in irregular configurations and FDM, 174, 177
 - in triangular element analysis in FEM, 209–212, 216, **217**
- Triangular (trigonal) meshes, 194–195, **195**, 241–242, **241**
- Triangular cells, five-point configurations and FDM, 133
- Triangular element
 - analysis in FEM, 204–216, **204**
 - area coordinates in, 204–205, **206**, **207**
 - bivariate Hermite interpolation in, 213
 - boundary element method (BEM) and, 204
 - centroid data in, 212
 - curvilinear triangles and, 210
 - differentiation in systems of triangles and, 213–216, **214**
 - first-order finite element method (FOFEM) in, 212
 - general relations and area coordinates for, 204–207
 - global vs. local numbering in, 204
 - Hermite interpolation in, 209, 211–212, **211**, 216
 - integration over triangular domains in, 207–209
 - isoparametric functions in, 210–211, **210**
 - Jacobian determinants, 208, 211
 - Lagrange interpolation in, 209, 212
 - linear functions for, 209–210
 - planar structures in, 205
 - quadratic functions in, 210–211
 - quadrature formulas/parameters for, 208–209, 208*t*, **208**
 - quadrilateral elements in, 213, **213**
 - rotated normal vectors in, 205
 - second-order finite element method (SOFEM) in, 212, 216
 - side midpoints in, 206
 - trial functions for, 209–212, 216, **217**
- Triangulation, xii
 - completely external point, 197
 - cylindrical surface for BEM, 196, **196**
 - inner point, 197
 - limitation function for, 198–199
 - partly external point, 197–198, **198**
 - regularization of, 199–200, **200**
 - table of elements for, 200
 - table of nodes for, 200
- Tridiagonal systems of splines, 83–85, 246, 304–307
- Trigonometric functions, 75, 97
 - fast method, 416
- Triode-electron sources, 384
- Two-dimensional differential equations,
 - reduction to, 32
- Two-dimensional indexing in FEM, 195–196, **196**
- Two-dimensional interpolation, 103–107, 412–415 (*see also* Field interpolation)
- Two-dimensional meshes, 115–126, **116**, **117**
 - affine distortions in, 122–124
 - Cauchy-Riemann equations in orthogonal meshes and, 121
 - classification of configurations in, 126
 - conformal mappings in orthogonal meshes and, 121
 - conjugate gradient in, 120
 - eigenvalues in, 124
 - Euler equations in, 120
 - exponentially expanding meshes, 122, **123**
 - form invariant PDE in, 120
 - Gauss's theorem in, 120
 - general coordinate transforms in, 117–118
 - internal boundaries in, 126
 - isotropic material coefficients in orthogonal meshes and, 121
 - Jacobian determinant in, 118, 121
 - Lagrange equation in, 119, 120, 124
 - Laplace equation in orthogonal meshes and, 122
 - orthogonal meshes as, 121–124
 - partial differential equation (PDE) in, 115–126
 - regular vs. irregular points in, 126

- scalar invariants and, 117
 - sources and nonlinearities in, 124–126
 - space charge in, 124
 - symmetry axes in, 126
 - tensor law in, 119, 125, **125**
 - variables in, 126
 - variational principles in, 118–120
- Uncoupled integral equations, 349
- Unsaturated media, Maxwell's equations
and, 3 (*see also* Linear material
equations; Linear media)
- Vacuum domains, 22
- Variational principles, x, 8–12
discretization of, in FEM, 200–204
- in magnetic round lenses, 45
in two-dimensional meshes, 118–120
- Vector differentiation (∇), 2*t*, 6, 7, 9, 13, 19,
24, 29, 31
- Vector fields and Fourier-series
expansions, 34
- Vector integral equations, 27–28
- Vector Poisson equation, 5
- Vector potential (A), 2*t*, 5, 9–10, 25, 27,
34, 41, 336–337
- Velocity, in curve calculation, 97
- Wave equations, 12–15
- Wave fields, 28–29, 351–354
- Wehnelt electrodes, 384, **385**

This Page Intentionally Left Blank