

# HANDBOOK OF Charged Particle Optics

Second Edition



EDITED BY

Jon Orloff



CRC Press  
Taylor & Francis Group

A CHAPMAN & HALL BOOK

**HANDBOOK OF**

**Charged**

**Particle Optics**

**Second Edition**



# **HANDBOOK OF**

# **Charged**

# **Particle Optics**

## **Second Edition**

**EDITED BY**

**Jon Orloff**



**CRC Press**

Taylor & Francis Group

Boca Raton London New York

---

CRC Press is an imprint of the  
Taylor & Francis Group, an **informa** business

Cover: Focused ion beam (FIB) image of a FIB cross-sectioned NiAl thermal spray splat on a stainless steel (SS) substrate. The channeling contrast shows columnar grain growth of the NiAl. The large grains of the SS are visible. The slight contrast changes in the SS grain under the splat are due to slight orientation changes due to the mechanical polishing of the SS prior to the splat deposition. Sample prepared and micrograph taken by Dr. Lucille Giannuzzi, FEI Company. The image was produced on a FEI Company Quanta 200 3D DualBeam. The sample is courtesy of Prof. Sanjay Sampath at SUNY Stony Brook.

CRC Press  
Taylor & Francis Group  
6000 Broken Sound Parkway NW, Suite 300  
Boca Raton, FL 33487-2742

© 2009 by Taylor & Francis Group, LLC  
CRC Press is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works  
Printed in the United States of America on acid-free paper  
10 9 8 7 6 5 4 3 2 1

International Standard Book Number-13: 978-1-4200-4554-3 (Hardcover)

This book contains information obtained from authentic and highly regarded sources. Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access [www.copyright.com](http://www.copyright.com) (<http://www.copyright.com/>) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

**Trademark Notice:** Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

---

**Library of Congress Cataloging-in-Publication Data**

---

Handbook of charged particle optics / edited by Jon Orloff. -- 2nd ed.

p. cm.

Includes bibliographical references and index.

ISBN-13: 978-1-4200-4554-3

ISBN-10: 1-4200-4554-7

1. Optical instruments--Design and construction--Handbooks, manuals, etc. 2. Electron optics--Handbooks, manuals, etc. I. Orloff, Jon. II. Title.

QC372.2.D4H36 2009

681'.4--dc22

2008013026

---

Visit the Taylor & Francis Web site at  
<http://www.taylorandfrancis.com>

and the CRC Press Web site at  
<http://www.crcpress.com>

---

# Contents

Preface to the Second Edition .....	vii
Editor .....	ix
Contributors .....	xi
<b>1 Review of ZrO/W Schottky Cathode .....</b>	<b>1</b>
<i>Lyn W. Swanson and Gregory A. Schwind</i>	
<b>2 Liquid Metal Ion Sources .....</b>	<b>29</b>
<i>Richard G. Forbes and (the late) Graeme L. R. Mair</i>	
<b>3 Gas Field Ionization Sources .....</b>	<b>87</b>
<i>Richard G. Forbes</i>	
<b>4 Magnetic Lenses for Electron Microscopy.....</b>	<b>129</b>
<i>Katsushige Tsuno</i>	
<b>5 Electrostatic Lenses .....</b>	<b>161</b>
<i>Bohumila Lencová</i>	
<b>6 Aberrations .....</b>	<b>209</b>
<i>Peter W. Hawkes</i>	
<b>7 Space Charge and Statistical Coulomb Effects .....</b>	<b>341</b>
<i>Pieter Kruit and Guus H. Jansen</i>	
<b>8 Resolution .....</b>	<b>391</b>
<i>Mitsugu Sato</i>	
<b>9 The Scanning Electron Microscope .....</b>	<b>437</b>
<i>András E. Vladár and Michael T. Postek</i>	
<b>10 The Scanning Transmission Electron Microscope.....</b>	<b>497</b>
<i>Albert V. Crewe (updated by Peter D. Nellist)</i>	
<b>11 Focused Ion Beams.....</b>	<b>523</b>
<i>M. Utlaut</i>	

<b>12</b>	Aberration Correction in Electron Microscopy.....	601
<i>Ondrej L. Krivanek, Niklas Dellby, and Matthew F. Murfitt</i>		
<b>Appendix: Computational Resources for Electron Microscopy.....</b>		<b>641</b>
<i>J. Orloff (with valuable information from Peter W. Hawkes and Bohumila Lencová)</i>		
<b>Index.....</b>	<b>645</b>	

---

# Preface to the Second Edition

The purpose of the second edition of this handbook, as with the first, is to provide a convenient place to find answers to many, if not all, questions pertaining to the basic physics of how and why high-resolution focused beam instruments work the way they do. This is a reference book for the users and designers of instrumentation.

By high-resolution systems we mean those that are designed to produce a focused beam whose dimensions are in the range 0.1–1000 nm, or so. The chapters on the high-brightness Schottky electron and liquid metal ion sources most widely used for probe-forming instruments today provide a thorough coverage of these critical elements. In recognition of a new microscopy based on the use of light ions, there is a new chapter on the physics and optics of the gas field ionization source that puts this source technology on a firm footing.

The chapters on the electrostatic and magnetic lenses used for focused beam systems provide an up-to-date account of these technologies; the expanded chapter on nonrelativistic aberrations provides a rather complete accounting of the subject. There is an expanded chapter on space-charge and coulomb effects, which place a limitation on high resolution, especially with focused ion beam (FIB) systems. These four chapters along with the chapters on electron and ion sources provide a basis for the design of any nonrelativistic high-resolution focused beam system (which is virtually all of them).

The properties and applications of the scanning electron microscope (SEM), the scanning transmission electron microscope (STEM), and the FIB system are covered in chapters providing the latest (2008) information on these subjects. Since the first edition of this handbook there have been dramatic developments in aberration correction for electron beam instruments, both theoretically and practically (due in great part to the tremendous increase in power and decrease in cost of small computers). A new chapter on this subject has therefore been added.

With the increase in the capability of focused electron and ion beam systems has come the difficult (but wonderful) problem of defining what the resolving power of a system capable of a 0.1 nm focused beam size really means—in terms of the relation between the beam and individual atoms in the target. Or, for that matter, what is meant by *beam size*. This is a subject of intense study (ca. 2008) and ideas are evolving rapidly as to how to define resolution and beam size and their relationship in this regime. The introduction to the chapter on resolution attempts to lay out some of the issues.

Computational tools for system optical design have reached the point where it no longer seems necessary to devote a chapter to them. A brief mention is made in the Appendix, along with references and locations (URLs) where some of the most used software tools can be found.

I wish to acknowledge here the great effort put into the creation of this book by its numerous contributors, who supplied 99% of the energy that went into it. The reader will notice that different chapters have different styles; this is appropriate for a group of authors from around the globe. He or she may also notice that this extends to the way references are handled. Your editor considered making the reference style uniform, and then decided that the potential gain of uniformity (minuscule) was not worth the potential cost of introducing errors by attempting uniformity (significant). Speaking of errors, we are confident they are only of the nature of typos and there will doubtless be some not caught in the editorial process. That is the fault only of the editor.

Finally, it is with great sadness that we must report the death of Dr. Graeme Mair, one of the trail-blazing scientists in the development and understanding of liquid metal ion sources and author of the chapter on liquid metal ions sources in the first edition of this handbook. He brought great

physical insight to the field and his presence is greatly missed. Dr. Richard Forbes was kind enough to revise and expand Dr. Mair's chapter, as well as to provide the new chapter on gas field ionization sources. I also wish to thank Dr. Peter Nellist for generously agreeing to update the chapter on STEM. The original chapter was written by Dr. Albert Crewe who, unfortunately, is in ill health and was unable to work on its update.

---

# Editor

**Jon Orloff** is emeritus professor in the Department of Electrical and Computer Engineering at the University of Maryland at College Park, where he was in the faculty from 1993–2006. From 1978–1993 he was professor of applied physics at the Oregon Graduate Institute. He has worked on the development and application of high resolution focused ion beam technology, and on applications of high brightness electron and ion sources. He is author or coauthor of 85 papers, most having to do with focused ion beam technology, as well as a monograph, *High Resolution Focused Ion Beams: FIB and Its Applications* (with L. Swanson and M. Utlaut). After serving as associate chair for undergraduate education in his department for 5 years, Professor Orloff retired from the University of Maryland in 2006, and currently lives in a minuscule town on the Oregon coast where it is very quiet and charged particle optics is almost completely unknown. He occasionally consults on matters related to focused ion beam technology.



---

# Contributors

**Albert V. Crewe**

Department of Physics and Enrico Fermi  
Institute  
University of Chicago  
Chicago, Illinois

**Niklas Dellby**

Nion Company  
Kirkland, Washington

**Richard G. Forbes**

Advanced Technology Institute  
University of Surrey  
Guildford, United Kingdom

**Peter W. Hawkes**

CEMES-CNRS  
Toulouse, France

**Guus H. Jansen**

Caneval Ventures  
Haarlem, The Netherlands

**Ondrej L. Krivanek**

Nion Company  
Kirkland, Washington

**Pieter Kruit**

Faculty of Applied Physics  
Delft University of Technology  
Delft, The Netherlands

**Bohumila Lencová**

Institute of Scientific Instruments  
Brno, Czech Republic

**Graeme L. R. Mair** (deceased)

Department of Physics  
University of Athens  
Zografos, Greece

**Matthew F. Murfitt**

Nion Company  
Kirkland, Washington

**Peter D. Nellist**

Department of Materials  
Oxford University  
Oxford, United Kingdom

**Michael T. Postek**

National Institute of Standards  
and Technology  
Gaithersburg, Maryland

**Mitsugu Sato**

Hitachi, Ltd.  
Ibaraki, Japan

**Gregory A. Schwind**

FEI Company  
Hillsboro, Oregon

**Lyn W. Swanson**

FEI Company  
Hillsboro, Oregon

**Katsushige Tsuno**

JEOL Ltd.  
Tokyo, Japan

**M. Utlaat**

Department of Physics  
University of Portland  
Portland, Oregon

**András E. Vladár**

National Institute of Standards  
and Technology  
Gaithersburg, Maryland



---

# 1 Review of ZrO/W Schottky Cathode

*Lyn W. Swanson and Gregory A. Schwind*

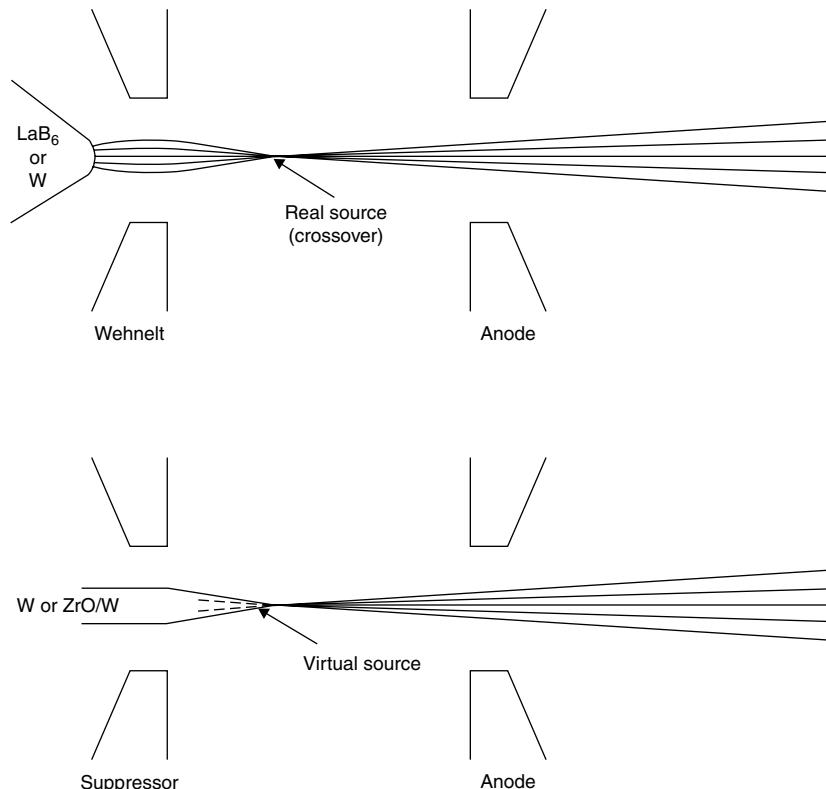
## CONTENTS

1.1	Introduction.....	1
1.2	ZrO/W Cathode Background.....	2
1.3	Schottky Emission .....	5
1.4	Extended Schottky Emission .....	6
1.5	Relationship among $\beta$ , Emitter Radius, and Work Function .....	8
1.6	Angular Intensity/Extraction Voltage Relationships .....	10
1.7	Emitter Shape Stability .....	12
1.8	Total Energy Distribution .....	19
1.9	Emitter Brightness .....	22
1.10	Current Fluctuations .....	23
1.11	Emitter Environmental Requirements .....	24
1.12	Emitter Life Considerations .....	26
1.13	Summary.....	27
	Acknowledgment .....	28
	References .....	28

## 1.1 INTRODUCTION

Point cathodes are used in electron optical systems to produce high-brightness, submicron, focused electron beams. Initially, the point cathode most often used in commercial applications was the room-temperature field emission source. In the past 20–30 years several high-temperature point cathodes have been developed into commercially viable electron sources. A number of early investigators, including Hibi,<sup>1</sup> Maruse and Sakaki,<sup>2</sup> and Everhart,<sup>3</sup> pioneered the study of high-temperature Schottky point cathodes. Today, the high-temperature field emission source is the most commonly used electron source for a variety of commercial electron beam instruments requiring a high-brightness cathode, including electron beam lithography, scanning and transmission microscopes, critical dimension measurement tools, etc.

Virtual source point cathodes are distinguished from the more conventional “crossover mode” thermionic cathodes, which operate near the space charge limit, by the high electric field at the cathode surface. Figure 1.1 illustrates the difference between the virtual and crossover modes of cathode operation. In the latter case the electron optical system uses the crossover as the object in the subsequent electron optical system, whereas in the former the virtual crossover, located a short distance behind the physical cathode, is the object. The difference in the source electrode configuration between the virtual and real crossover modes of operation is rather trivial, as shown in Figure 1.1, and simply consists of a difference in protrusion length from the suppressor electrode (usually referred to as the Wehnelt or Schottky suppressor electrode). This small change in geometry dramatically alters the electric field at the cathode surface and the electron trajectories.



**FIGURE 1.1** Electron trajectories are illustrated for the real (upper diagram) and virtual crossover (lower diagram) source optics.

Poly-crystalline and single-crystal LaB<sub>6</sub> and CeB<sub>6</sub> point cathodes have been used over the past 15 years in a variety of microprobe instruments, but almost exclusively in the crossover mode. One purpose of the suppressor electrode, which operates several hundred volts negative with respect to the cathode, is to reduce the magnitude of the total emitted current by reducing extraneous thermal emission from the electrode shaft. This is particularly important for low-work-function cathodes, for example, LaB<sub>6</sub> thermionic and the ZrO/W Schottky cathodes, where the total current at normal operating temperatures of  $\sim$ 1800 K would exceed several hundred microamperes without the suppressor electrode.

Although the crossover versus noncrossover mode of operation is the major dividing line, several other important differences in cathode performance result as summarized in Table 1.1 where the high-temperature ZrO/W Schottky<sup>4–7</sup> and LaB<sub>6</sub> thermionic<sup>8–10</sup> cathodes are used as representative examples. In this review, the focus will be on the properties and emission characteristics of the ZrO/W Schottky cathode.

## 1.2 ZrO/W CATHODE BACKGROUND

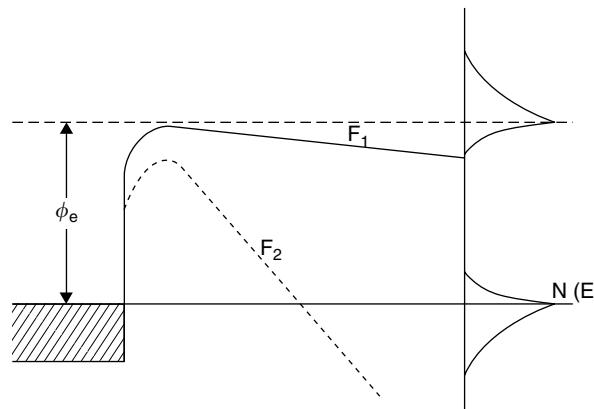
The noncrossover-type cathode, with the emitter protruding through the suppressor, is a high-field cathode and typically operates in the emission regime known as the “extended Schottky” regime.<sup>11,12</sup> In contrast, the crossover mode cathode (e.g., LaB<sub>6</sub> cathode) typically operates near the space charge limit with a low applied electric field. These two emission regimes are illustrated in Figure 1.2 which shows the extremes in terms of the average energy level and distribution of the emitted electrons relative to the Fermi level. Emission at or near the Fermi level occurs at low temperature and high field and is commonly referred to as cold field emission (CFE) and when operated at higher temperatures is

**TABLE 1.1**

**Comparison of Various Properties of the Crossover and Noncrossover Cathode Modes Using the ZrO/W and LaB<sub>6</sub> Cathodes as Examples**

Property	Noncrossover (ZrO/W)	Crossover (LaB <sub>6</sub> )
Electric field (V/m)	$>5 \times 10^8$ V/m	Space charge limited
Operating temperature (K)	1700–1800	1700–1800
Virtual source (nm)	20–40	10,000–20,000
Emitter radius ( $\mu\text{m}$ )	0.3–1.0	Several micrometers
Reduced source brightness at 5 kV (A/m <sup>2</sup> sr V)	$5 \times 10^7$ to $3 \times 10^8$	$2 \times 10^4$
Energy spread at cathode at 1800 K (eV)	0.4	0.15 <sup>a</sup>
Work function (eV)	2.95	2.75
Emission regime	Extended Schottky	Schottky
Cathode life (h)	>18,000 (2 years)	2000
Vacuum required (torr)	<1 $\times 10^{-8}$ (total) and $\leq 1 \times 10^{-9}$ for H <sub>2</sub> O and O <sub>2</sub>	<1 $\times 10^{-7}$

<sup>a</sup> Energy spread increases at subsequent crossover positions due to coulomb interactions.

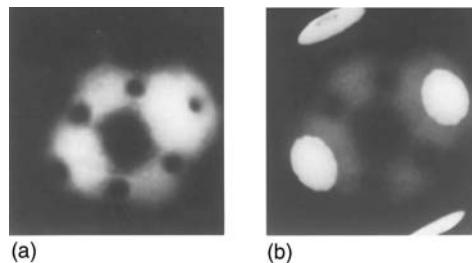


**FIGURE 1.2** Diagram of the electron potential energy at the solid–vacuum interface for the low-temperature cold field emission ( $F_2$ ) and high-temperature Schottky ( $F_1$ ) emission modes. Energy distributions for the two emission modes are indicated.

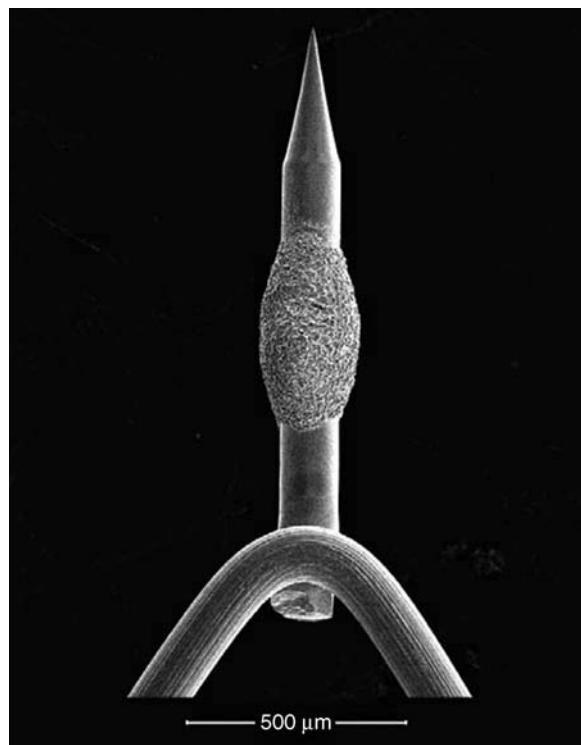
referred to as extended Schottky emission (SE).<sup>12</sup> In the crossover mode of operation a relatively low applied electric field exists at the cathode because of electrode geometry and space charge effects.

The ZrO/W point cathode has been studied extensively by several investigators.<sup>4–7,13–20</sup> From their studies it has been determined that the ZrO/W point cathode can operate in the temperature range from 300 to 1800 K, exhibits a low work function ( $2.9 \pm 0.1$  eV), and is convenient to operate with an emitter radius of 0.3–1.0  $\mu\text{m}$ . The low-work-function region of the emitter is highly localized to the (100) crystal face, as shown in Figure 1.3, by the selective adsorption of a ZrO complex supplied by surface diffusion from a reservoir of ZrO<sub>2</sub> on the emitter shank as shown in Figure 1.4.<sup>21</sup> By use of a ⟨100⟩ oriented single-crystal wire W for the emitter substrate, the selective work function lowering of the (100) plane allows for a narrow ( $\pm 7^\circ$  angular divergence) electron beam centered on the cathode axis as shown in Figure 1.5b.

The surface physics and chemistry processes which allow a co-adsorbed layer of Zr and O<sub>2</sub> to selectively adsorb and lower the work function of the W(100) crystal face has been the subject of several studies.<sup>15,19,22,23</sup> It has been determined that both the Zr coverage and the Zr/O ratio are critical



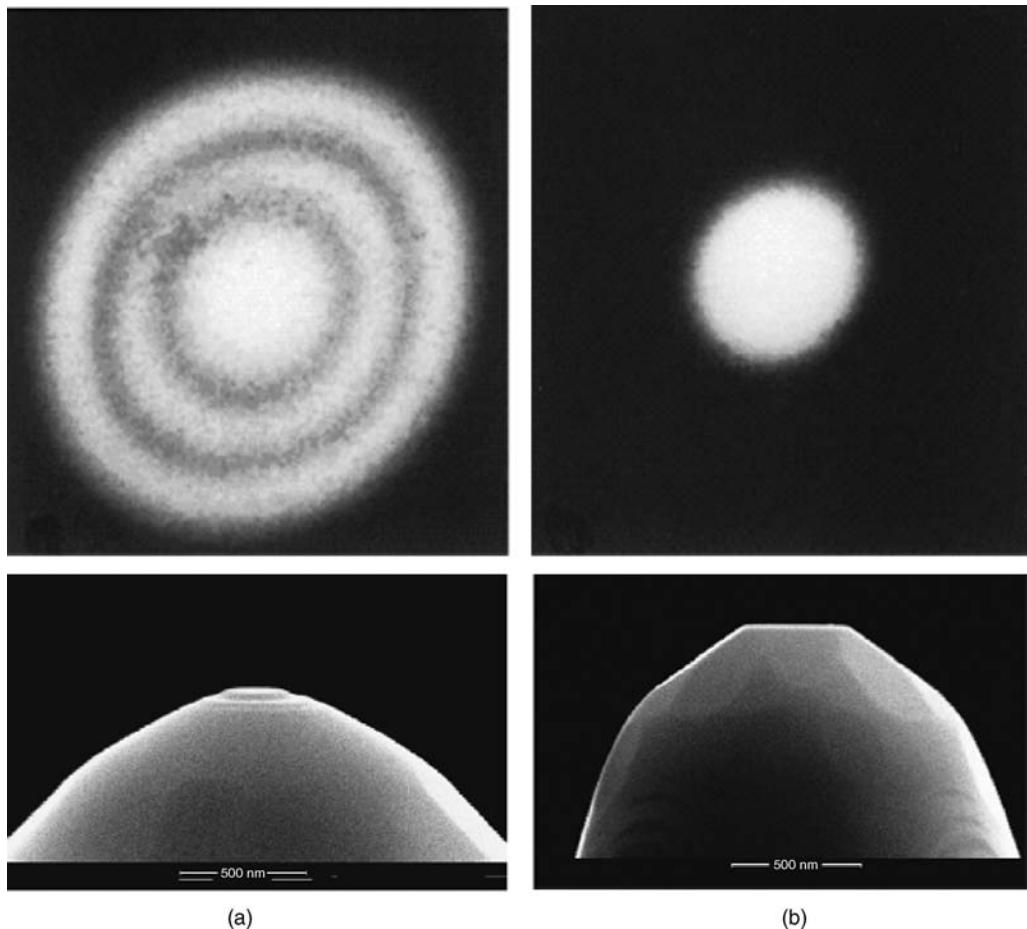
**FIGURE 1.3** Field electron microscope emission patterns of (a) clean and (b) ZrO<sub>2</sub>-coated tungsten ⟨110⟩ oriented emitters. Image (b) shows preferential emission from the (100) crystal faces.



**FIGURE 1.4** A scanning electron microscope image of the ZrO/W ⟨100⟩ Schottky emitter showing the ZrO<sub>2</sub> reservoir on the emitter shaft.

in establishing the minimum work function. Excess O<sub>2</sub> pressure can lead to an increase in the work function and concomitant reduction in emission; however, the work function change is completely reversible so long as the temperature does not exceed 1850 K. It is this remarkable thermal stability of the adsorbed ZrO complex that gives the cathode its ability to function as a viable, high-temperature cathode. A study<sup>24</sup> using a combination of X-ray photoelectron spectroscopy (XPS), low-energy electron diffraction (LEED), and work function measurement techniques determined the surface composition to be a Zr–O complex at the minimum work function.

The thermally annealed and field-stabilized end forms of the ZrO/W SE lead to different emission patterns as shown in Figure 1.5. The thermally annealed end form consists of terraced steps of the ⟨100⟩ crystal plane whose step height is enhanced by the adsorbed ZrO. This latter end form is transformed into a stable, single-facet end form by application of temperature and electric field as shown in



**FIGURE 1.5** Field electron microscope emission patterns (upper) and corresponding scanning electron microscope images of the emitter shapes (lower) for the ZrO/W SE cathode: (a) immediately after thermal processing at 1800 K in a low electric field and (b) after long-term heating at 1800 K in a high electric field.

Figure 1.5b. This (100) facet remains stable indefinitely as long as the field strength exceeds 8 MV/cm at an operating temperature of  $\leq 1800$  K which corresponds to a current/solid angle (angular intensity) of 0.2 mA/sr for a 0.5- $\mu\text{m}$  emitter. Similar results can be obtained by substituting Sc,<sup>25</sup> Y,<sup>26</sup> Hf,<sup>27</sup> Ti,<sup>28</sup> or others<sup>29</sup> for Zr. The low work function permits a relatively large radius field emitter, operated at an elevated temperature in the extended SE mode, to compete effectively in terms of brightness with a small radius, high-work function CFE emitter operating in the pure field emission mode. Despite as much as a factor of 5–10 increase in emitter radius between the CFE and SE emitter types, the virtual source size of the SE is still relatively small, for example, 15–20 nm. In combination with the high angular intensity capable from the SE cathode, this results in a very high-brightness electron source.

### 1.3 SCHOTTKY EMISSION

The cathode current density  $J_S$  for the SE regime is related to the work function  $\phi$ , temperature  $T$ , and electric field  $F$  by the Schottky equation<sup>30</sup>

$$J_S = \frac{4\pi m_e (kT)^2}{h^3} \exp\left(\frac{e^{3/2} F^{1/2}}{(4\pi\epsilon_0)^{1/2} kT} - \frac{\phi}{kT}\right) \quad (1.1)$$

where  $m_e$ ,  $e$ ,  $k$ , and  $h$  are the usual physical constants. For very low electric fields, Equation 1.1 reduces to the Richardson–Dushman equation.<sup>31</sup> As the field increases, the Schottky equation describes the effect of the lowering of the potential barrier in Figure 1.1 by the Schottky term  $e^{3/2}F^{1/2}/(4\pi\varepsilon_0)^{1/2}$ .

Experimentally accessible parameters  $V_E$  (extraction voltage) and  $I'$  (current angular intensity) are related to the Equation 1.1 variables as follows:

$$F = \beta V_E \quad (1.2)$$

where the field factor  $\beta$  is a function of emitter radius and electrode spacing and  $I'$  is related to  $J_S$  by

$$I' = J_S \left( \frac{r}{m} \right)^2 \quad (1.3)$$

where  $r$  is the emitter apex radius and  $m$  the angular magnification (between the cathode surface and virtual source position). The slopes of experimental plots of  $\ln(I')$  versus  $V_E^{1/2}$  (Schottky plots) will yield the value of  $\beta$  via Equation 1.1. In Section 1.5 an empirical relationship among  $\beta$ ,  $r$ , and work function will be developed.

## 1.4 EXTENDED SCHOTTKY EMISSION

The Schottky Equation 1.1 is valid for fields sufficiently low so that negligible tunneling current contributes to the value of  $J_S$ . A measure of the tunneling current is given by the dimensionless parameter  $q$ ,<sup>11,12</sup> where

$$q = \frac{h(4\pi\varepsilon_0)^{1/4}F^{3/4}}{2\pi^2 m_e^{1/2} k T} = 1.656 \times 10^{-4} \frac{F^{3/4}}{T} \quad (1.4)$$

and  $F$  is in V/m. For example, when  $q = 0.5$ , approximately half of the current is due to tunneling. When  $q$  is small the emission characteristics are adequately described by Equation 1.1 with only a small fraction of current due to tunneling. For higher values of  $q$  the emission is in the extended Schottky regime and the current density  $J_{ES}$  can be approximately described by the following equation:<sup>11,12</sup>

$$J_{ES} = J_S \frac{\pi q}{\sin(\pi q)} \quad (1.5)$$

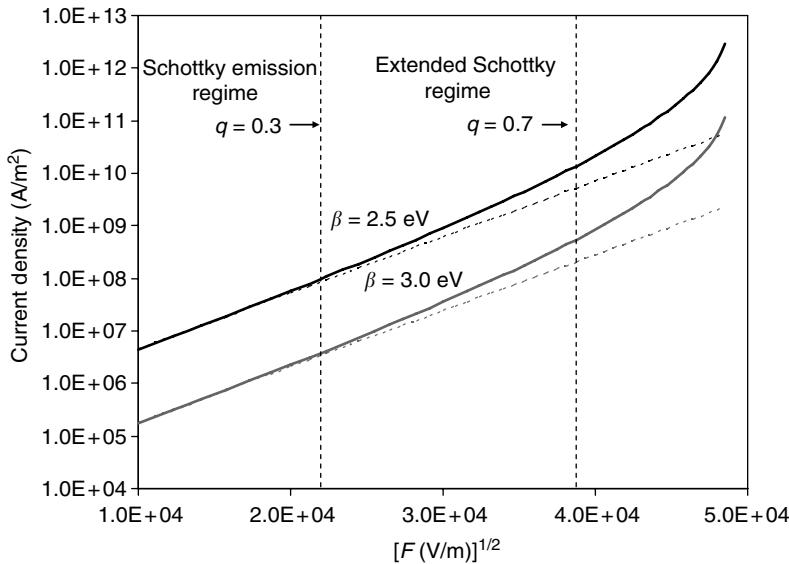
This equation has been believed to extend the usefulness of the Schottky equation to  $0 < q < 0.7$ , beyond which Equation 1.5 becomes inaccurate. A Schottky plot of Equation 1.5 is shown in Figure 1.6. If the upper boundary of the extended Schottky regime is considered to be at  $q = 0.7$ , this occurs at a field of

$$F_1(T) = \left[ \frac{1.4\pi^2 m_e^{1/2} k T}{h(4\pi\varepsilon_0)^{1/4}} \right]^{4/3} = 6.82 \times 10^4 T^{4/3} \quad (1.6)$$

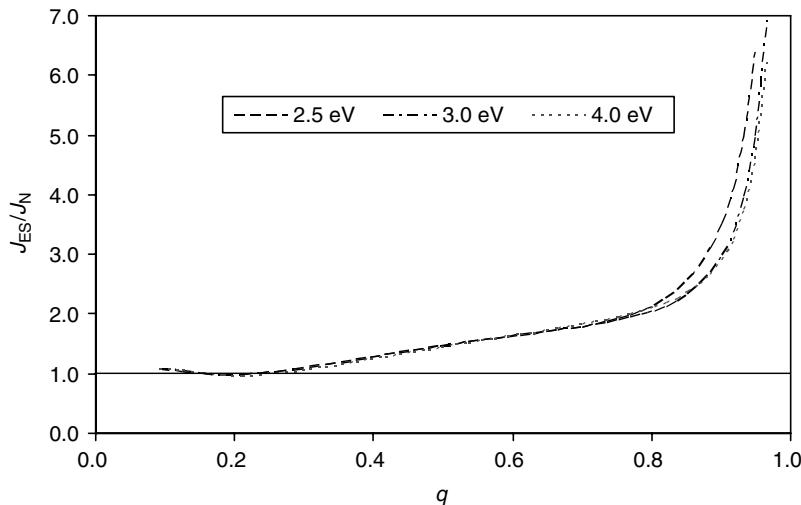
The boundary between SE and extended SE emission is at approximately  $q = 0.3$ .

Experimental values of  $\beta$  can be obtained in the Schottky regime from Schottky plots ( $\ln(I')$  versus  $V_E^{1/2}$ ); however in the extended Schottky regime, these  $\beta$  values will be too large as the factor  $\pi q / \sin(\pi q)$  introduces a stronger field dependence than is assumed in the Schottky equation 1.1. To obtain the correct  $\beta$  value, it is necessary to examine the slope of the Schottky plot using a correction term. It can be shown that the Schottky slope for Equation 1.5 is given by

$$\frac{d \{ \ln J_{ES} \}}{d \{ V_E^{1/2} \}} = \frac{e^{3/2} \beta^{1/2}}{(4\pi\varepsilon_0)^{1/2} k T} \left[ 1 + \frac{3kT(4\pi\varepsilon_0)^{1/2}}{2e^{3/2} F^{1/2}} [1 - \pi q \cot(\pi q)] \right] \quad (1.7)$$



**FIGURE 1.6** Theoretical Schottky plots at  $T = 1800$  K for the indicated values of work function using Equations 1.1 and 1.5. Regions where the Schottky and extended Schottky emission models are valid are indicated.



**FIGURE 1.7** Graph shows ratio of  $J_{ES}$  to the numerically calculated current density  $J_N$  for  $T = 1800$  K and the indicated work function values. A ratio of unity indicates the  $J_{ES}$  model is valid.

The term in brackets can be viewed as a correction factor to the true Schottky slope  $e^{3/2}\beta^{1/2}/((4\pi\epsilon_0)^{1/2}kT)$ . By measuring the experimental Schottky slope, Equation 1.7 can be solved iteratively to give the true  $\beta$  value. The current density over the full range of  $F$  and  $T$  can be obtained using a numerical method<sup>32,33</sup> that incorporates a fourth order Runge–Kutta–Fehlberg numerical integration method.<sup>34</sup> A comparison between the numerically calculated current density  $J_N$  and  $J_{ES}$  (obtained from the approximate analytical Equation 1.5) reveals that for  $q > 0.3$  the difference between  $J_N$  and  $J_{ES}$  is sufficient to question the accuracy of using Equation 1.7 for determination of  $\beta$ . In Figure 1.7 the ratio  $J_{ES}/J_N$  plotted versus  $q$  shows a serious deviation between  $J_{ES}$  and  $J_N$  for  $q > 0.3$  instead of the generally accepted view that  $J_{ES}$  is valid for  $0 > q > 0.7$ . This casts a doubt

on the validity of using Equations 1.5 and 1.7 to calculate  $\beta$  from experimental Schottky plots. In Section 1.8 a similar comparison between the total energy distribution (TED) curves predicted by ES theory and numerical calculations will be given.

## 1.5 RELATIONSHIP AMONG $\beta$ , Emitter RADIUS, AND WORK FUNCTION

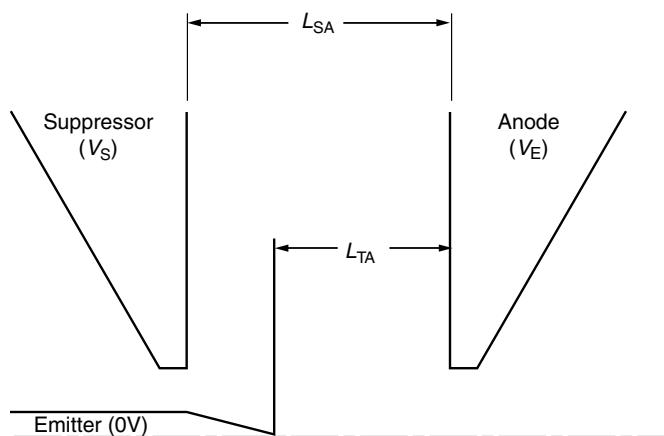
To relate the field factor  $\beta$  and the angular magnification  $m$  to the emitter radius, a model of the emitter and the electron gun must be chosen. Numerous approximate model shapes have been used for point emitters, such as hyperbolic,<sup>35</sup> parabolic,<sup>36</sup> sphere-on-cone,<sup>37</sup> and so on, but none of these can accurately model the end facet on the ZrO/W emitter as shown in Figure 1.5b. A finite difference computer program using the spherical coordinates with an increasing mesh (SCWIM) model<sup>38</sup> was used to determine  $\beta$  and  $m$  for several emitter radii. The geometry used for these calculations is shown in Figure 1.8. The suppressor electrode potential  $V_s$  was typically  $-300$  V with respect to the emitter potential and served to reduce thermal emission from the low-work-function surface along the emitter's cylindrical shaft. Computer modeling with the SCWIM program resulted in the following empirical relationship between  $m$ ,  $\beta$ , and  $r$ :

$$m = 8.713 \times 10^{-5} (\beta r)^{0.42} \quad (1.8)$$

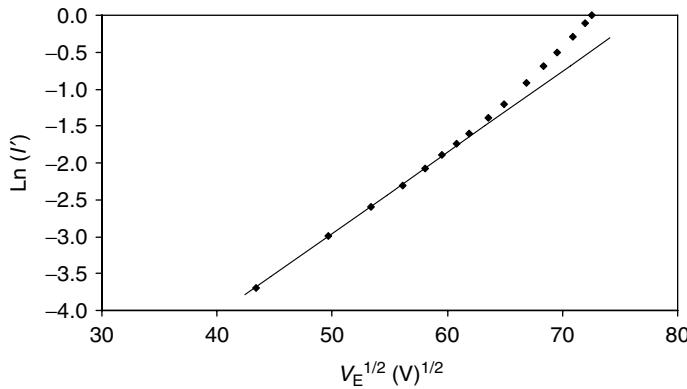
where  $\beta$  and  $r$  are in units of meter<sup>-1</sup> and nanometers, respectively. Equation 1.8 is accurate to within  $\pm 1\%$  for  $\beta$  values between 20,000 and 300,000 m<sup>-1</sup>.

A typical Schottky plot of  $\ln(I')$  versus  $V_E^{1/2}$  is shown in Figure 1.9 for the ZrO/W SE source. The experimental data deviate from the straight-line Schottky plot at the upper end of the data range as the emission mechanism crosses the boundary separating the Schottky and extended Schottky regimes described in Section 1.4.

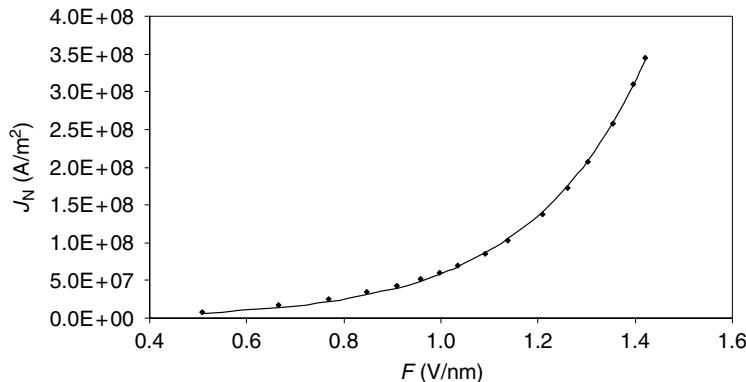
In view of the unexpected variance between  $J_{ES}$  and  $J_N$  and the difficulty of ascribing an initial Schottky slope where the data becomes nonlinear as shown in Figure 1.9, a two-parameter curve fit method for evaluating  $\beta$  and  $\phi$  was developed. The curve fit method follows the Nelder–Mead<sup>39</sup> approach where an iterative, best fit of the experimental data to a  $J_N$  versus  $F$  curve is achieved using Equations 1.2 and 1.3 to convert experimental  $I'(V_E)$  data to  $J_{ES}(F)$  data and where  $\beta$  and  $\phi$  are the fitting parameters. The empirical equation 1.8 is used to determine  $m$  from the  $\beta$  and  $r$  values. The  $J_N$  values are calculated numerically from the basic field emission equations as described in Section 1.4. The resulting curve fit for the Figure 1.9 data is shown in Figure 1.10.



**FIGURE 1.8** Details of the emitter geometry used for the ZrO/W emission studies and computer modeling. For the experimental studies  $L_{SA} = 760$   $\mu\text{m}$ ,  $L_{TA} = 508$   $\mu\text{m}$ , and  $V_s = -300$  V unless otherwise noted.



**FIGURE 1.9** Experimental Schottky plot showing  $\ln(I')$  versus  $V_E^{1/2}$ , where  $r = 320$  nm and  $T = 1800$  K.



**FIGURE 1.10** Best fit of the experimental data (data points) to a  $J_N$  versus  $F$  curve (solid line) where  $r = 320$  nm and  $T = 1800$  K. The fitting parameters  $\phi$  and  $\beta$  are  $2.94$  eV and  $2.6 \times 10^5 \text{ m}^{-1}$ , respectively.

From experimental  $I'(V_E)$  data for several emitters at various values of  $r$ , a set of  $\phi$ ,  $m$ , and  $\beta$  values were obtained as shown in Table 1.2. From the Table 1.2 data it can be shown that an empirical relationship between  $r$  and  $\beta$  of the form

$$\beta = 6.738 \times 10^7 r^{-0.96} (\text{m}^{-1}) \quad (1.9)$$

can be obtained over the range  $r = 300\text{--}1400$  nm. In addition, from experimental results similar to Table 1.1 where  $L_{\text{TA}}$  was varied from 550 to 1400  $\mu\text{m}$  it was found that

$$\beta = 1.12 \times 10^7 L_{\text{TA}}^{-0.632} (\text{m}^{-1}) \quad (1.10)$$

Equations 1.9 and 1.10 can be combined to form the following relationship accurate to  $\pm 5\%$  over the aforementioned ranges of  $r$  and  $L_{\text{TA}}$ :

$$\beta = 3.50 \times 10^9 L_{\text{TA}}^{-0.632} r^{-0.96} (\text{m}^{-1}) \quad (1.11)$$

For the Equations 1.9 through 1.11 empirical relationships the emitter protrusion  $L_{\text{SA}} - L_{\text{TA}}$  was constant at 268  $\mu\text{m}$ .

**TABLE 1.2**  
**Results of Analysis of  $I'$  (V) Data for Various ZrO/W Schottky Emission Sources  
 Obtained Using the Curve Fit Method**

$r$ (nm) (Observed)	$r$ (nm) (Calculated from Equation 1.9)	$\beta$ ( $m^{-1}$ )	$\varphi$ (eV)	$m$
270	318	267,200	2.96	0.186
320	322	263,500	2.94	0.186
400	392	218,400	2.92	0.187
500	480	179,838	2.90	0.188
670	521	166,078	2.88	0.188
830	832	106,000	2.85	0.189
890	988	89,900	2.79	0.182
1220	1365	65,876	2.83	0.191
1440	1473	61,260	2.85	0.191

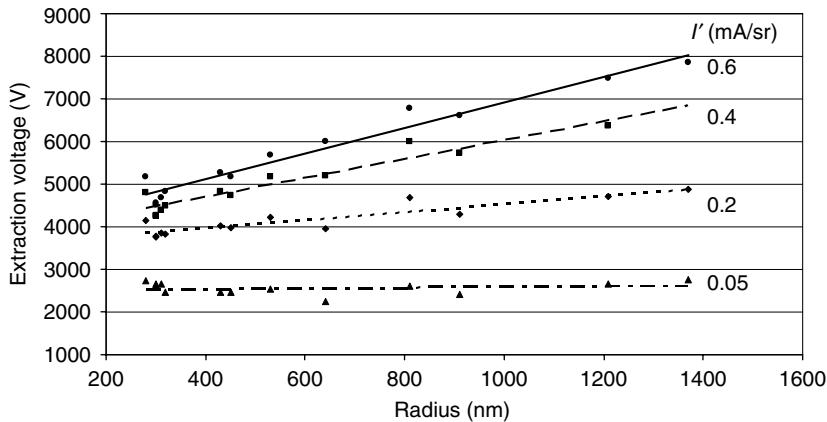
Note: The values in the table are for the electrode geometry given in Figure 1.8.

## 1.6 ANGULAR INTENSITY/EXTRACTION VOLTAGE RELATIONSHIPS

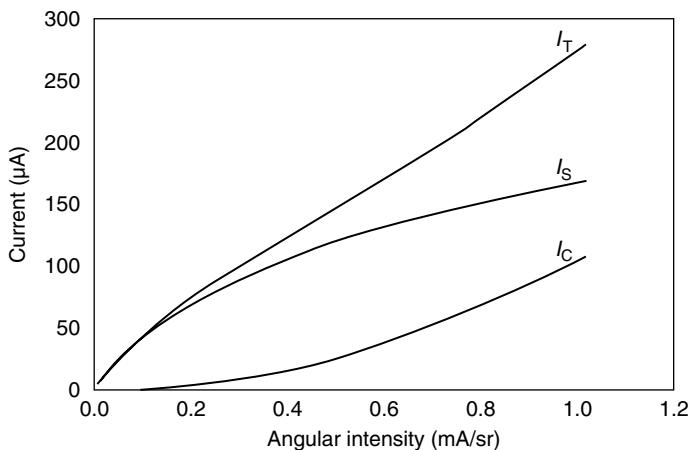
Figure 1.11 shows the experimentally measured relationship between  $I'$ ,  $V_E$ , and  $r$  where  $V_E$  is plotted versus  $r$  for the indicated values of  $I'$ . The electrode geometry is as shown in Figure 1.8 where  $L_{SA} = 760 \mu m$ ,  $L_{TA} = 508 \mu m$  and the suppressor and extractor bore diameters are both  $380 \mu m$ . It is clear from the Figure 1.11 plots that the variation of  $V_E$  with  $r$  is significantly reduced for low values of  $I'$ . This is due to the increasing contribution of pure SE at the lower values of electric field strength, that is, electrons predominately escaping over the potential barrier (see Figure 1.2). The total emission current ( $I_T$ ) consists of two components—(1) emission from the central, low-work-function (100) crystal plane ( $I_c$ ) and (2) emission from the four (100) planes located  $90^\circ$  from the central (100) plane and along the emitter shank ( $I_s$ ). Figure 1.12 shows an experimental plot of  $I_c$  and  $I_s$  versus  $I'$ . The relative contribution of  $I_c$  to the total current increases with  $I'$ . This results from the fact that for a given  $I'$  (or  $V_E$ ) the electric field at the central (100) plane is much higher than the electric field along the emitter shank. This means that the current comprising  $I_c$  transitions into the more field-dependent, extended SE regime at a lower value of  $V_E$  than does the emission from the shank region. The smaller the value of  $r$ , the lower the value of  $I'$  (or  $V_E$ ) where this transition occurs.

The electron trajectories from the emitter region contributing to  $I_c$  and  $I_s$  are shown in Figure 1.13 for the Figure 1.8 electrode geometry. For the indicated dimensions, one can observe that the emission from the central (100) plane ( $I_c$ ) is transmitted through the extractor electrode and the emission from all the other regions of the emitter is collected on the extractor electrode. If the extraction electrode aperture diameter is increased to 1.5 mm, the central (100) current, along with most of the shank emission, is transmitted. The latter situation is shown in Figure 1.14 where a fluorescent screen placed downstream from the extraction aperture shows the emission distribution. The emission from the central (100) plane is contained in a  $7^\circ$  half angle. The emission from the shank region is separated from the central (100) emission by  $18^\circ$ .

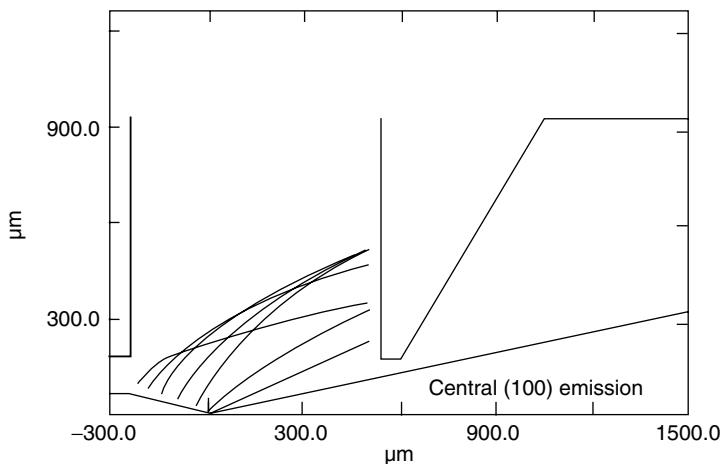
The emission distribution from the central (100) plane, obtained by scanning the beam across a small probe hole, is shown in Figure 1.15 for various values of  $V_E$ .<sup>17</sup> As  $V_E$  increases, the emission distribution changes from a uniform flat to a “ring-shaped” distribution. In the operating range of  $I' = 0.1\text{--}0.7 \text{ mA/sr}$ , the angular emission distribution is relatively flat between  $-6^\circ$  and  $+6^\circ$ . From the SCWIM program the emitter apex field distribution was computed and is shown in Figure 1.16 for various emitter radii. The higher field at the edge of the central (100) facet accounts for the Figure 1.15 emission distribution. At higher values of  $V_E$  (or  $F$ ) the transition at the facet edge to the more field-dependent extended SE regime becomes more pronounced, thereby accounting for the onset of the ring-shaped emission distribution.



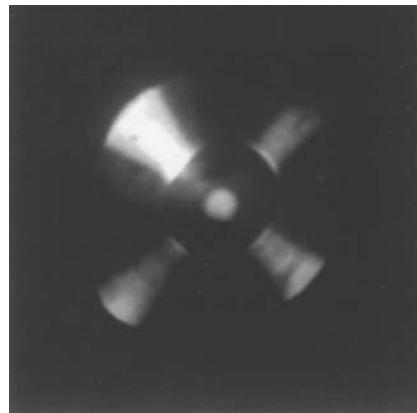
**FIGURE 1.11** Plots of measured  $V_E$  versus  $r$  for the indicated values of  $I'$  at  $T = 1800$  K.



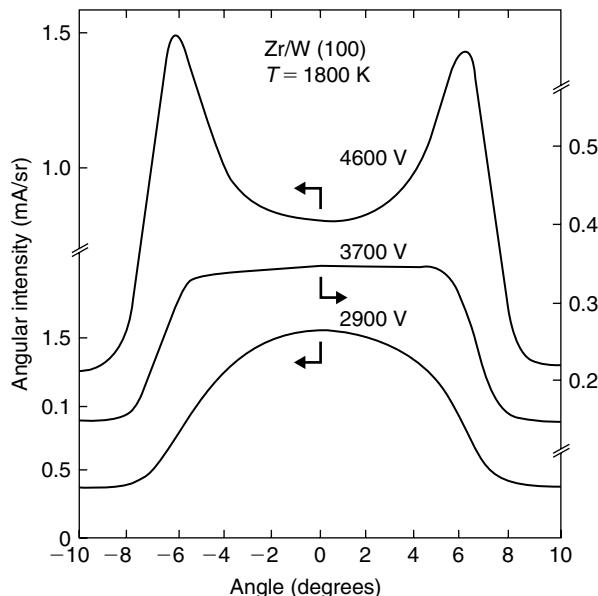
**FIGURE 1.12** Experimental plots of the total emission  $I_T$ , the shank  $I_S$ , and the central (100) plane currents  $I_C$  versus  $I'$  for a ZrO/W emitter using the electrode geometry shown in Figure 1.8.



**FIGURE 1.13** Computer-calculated trajectories for electron emission from the shank and central (100) plane regions of the ZrO/W emitter (values in  $\mu$ m).



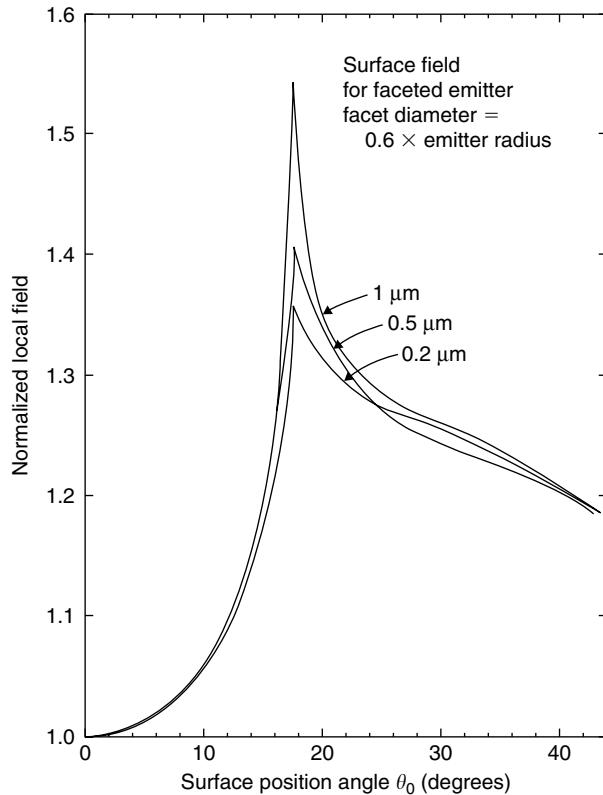
**FIGURE 1.14** Emission pattern of the shank (outer four emission lobes) and central (100) plane emission for the ZrO/W emitter.



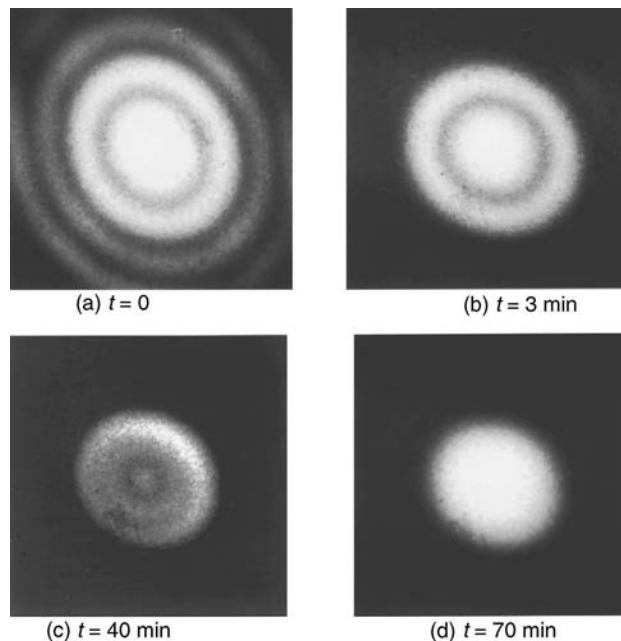
**FIGURE 1.15** Experimentally measured current angular intensity distribution at the indicated extractor voltages for the ZrO/W emitter with  $r = 0.8 \mu\text{m}$ .

## 1.7 Emitter Shape Stability

In the presence of an applied electric field with temperature sufficient to allow for surface self-diffusion, all field emitters undergo significant macroscopic shape changes. This so-called field buildup process has been well studied and understood,<sup>40,41</sup> and in the case of the tungsten body-centered cubic (bcc) structure it can lead to several end forms. With ZrO present on the surface, the (100), (110), and (112) crystal planes grow at the expense of lower-index crystal faces and the resulting end form is shown in Figure 1.5b. The process by which this occurs, as shown in Figure 1.17, is where the (100) net planes sequentially shrink in size and eventually vanish as surface W atoms and adsorbed  $\text{ZrO}_x$  entities migrate away from the terrace edges, which are visible in Figure 1.5a, and diffuse toward the edge of the next lower net plane. After several hours the stable end form, shown in Figure 1.5b, with the corresponding emission distribution, shown in Figure 1.17d, is achieved.



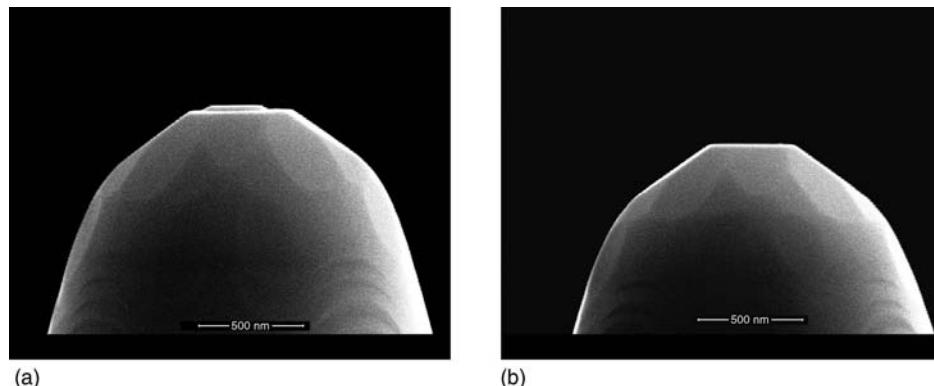
**FIGURE 1.16** Computer-calculated normalized surface field versus emission angle for the indicated emitter radii. The ratio of facet diameter to emitter radius was 0.6.



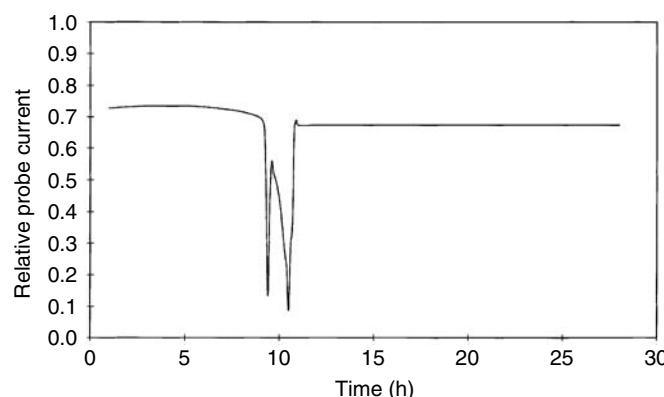
**FIGURE 1.17** Sequence of emission patterns showing (100) plane terrace collapse during field buildup of a ZrO/W Schottky emission cathode at  $T = 1900 \text{ K}$ : (a)  $t = 0$ ; (b)  $t = 3 \text{ min}$ ; (c)  $t = 40 \text{ min}$ ; and (d)  $t = 70 \text{ min}$ .

This sequence of events is illustrated more dramatically in Figure 1.18a where the collapsing (100) net plane was frozen in place by reducing the temperature and inserting the emitter into a scanning electron microscope (SEM) for a side profile view. From this view it can be determined that the height of the collapsing net plane is  $\sim 30$  nm. Figure 1.18b shows the emitter after further operation at 1800 K with the electric field applied. In the latter case the net plane has fully collapsed and the Figure 1.5b emission distribution is obtained. The ratio of the central (100) plane diameter to the overall emitter radius is  $\sim 0.6$  for a fully faceted emitter stabilized at an angular intensity of 0.5–1.0 mA/sr.

During this faceting process the central emission current (e.g., current within a half angle of  $\pm 6^\circ$ ) undergoes a cyclic change as the retreating (100) net plane vanishes as shown in Figures 1.17c and 1.17d. The change in the central emission current during the final stages of the net plane collapse is shown in Figure 1.19. The time period between the current oscillations as shown in Figure 1.19, varies from one to several hours depending on the emitter temperature and O<sub>2</sub> partial pressure (see Section 1.11). During the time between the current oscillations (Figure 1.19), the total emission current is relatively stable. However, until the faceting or field buildup process is completed and a stable end form has been achieved, the probe emission current will undergo several of these cycles of significant current change followed by a period of stability.

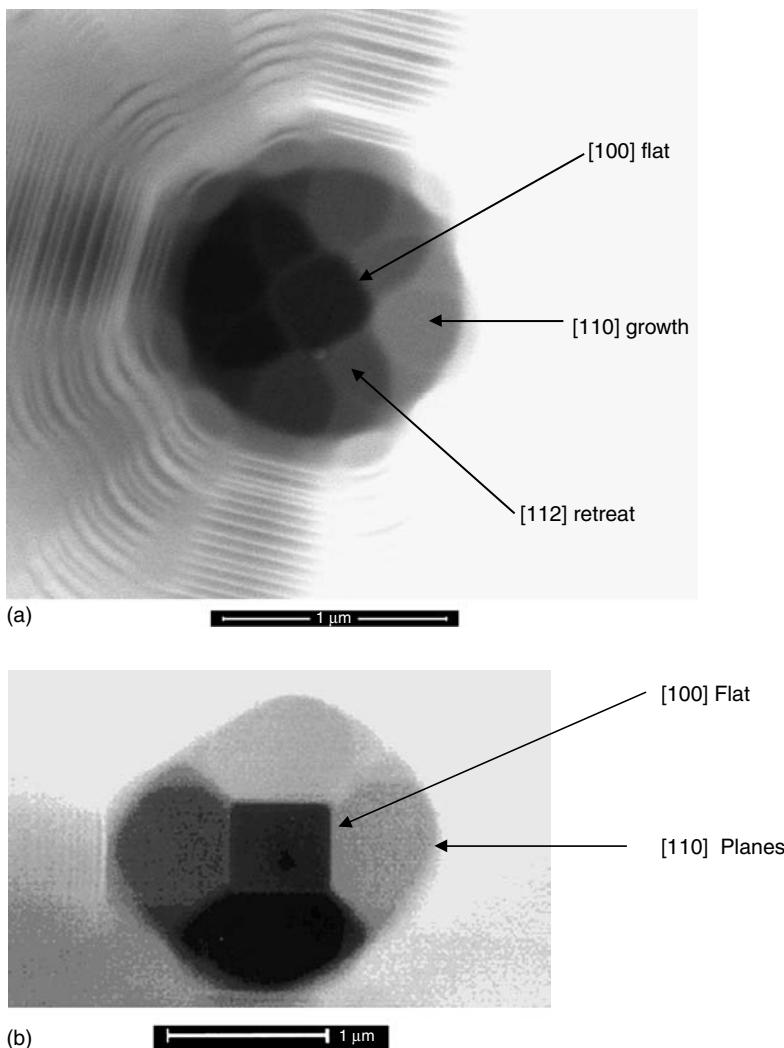


**FIGURE 1.18** Scanning electron microscope micrographs of a ZrO/W cathode for  $r = 0.90 \mu\text{m}$  that was undergoing thermal field buildup at 1800 K. Photos (a) and (b) show a time sequence of a (100) plane terrace undergoing collapse by surface diffusion.

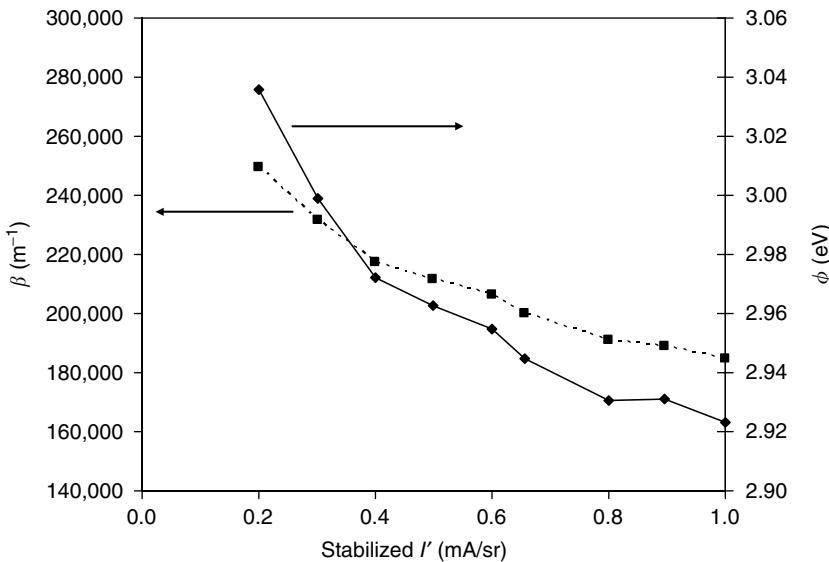


**FIGURE 1.19** A plot is shown of the emission current accepted from a 1.6 mrad semiangle probe centered on the axial (100) crystal face of the ZrO/W emitter. The large instability is due to a (100) plane terrace undergoing collapse (as shown in Figures 1.17 and 1.18) across the probe acceptance aperture.

Although the emission pattern from the central (100) plane often exhibits a round shape, the actual stable shape of the faceted emitter is a truncated pyramid with a (100) facet at the top of the pyramid and four (110) planes on the sides. Figure 1.20 shows a top-down SEM photo of a typical stable shape of the ZrO/W Schottky cathode. The figure 1.20a photo was taken during the early stages of the facet formation where both (112) and (110) planes make up the sides of the truncated pseudopyramid. Figure 1.20b shows the final stable shape that will remain unchanged during the several thousand hour life of the emitter so long as the electric field remains relatively constant and within a range of angular current densities described in more detail below. It is interesting to note the  $45^\circ$  rotation of the (100) flat shown in Figures 1.20a and 1.20b as the (110) side facets grow in size at the expense of the higher index and less thermodynamically stable (112) facets. This end form is generally observed for all emitters with radii in the range 200–1000 nm. For larger radii emitters the (112) and some higher index plane facets do not become extinguished and remain indefinite and the central (100) facet retains a more rounded shape.



**FIGURE 1.20** (a) Top-down view of Schottky source after 48 h of operation at 0.5 mA/sr angular current density. (b) Same as (a) but after several hours of additional operation at 0.5 mA/sr.



**FIGURE 1.21** Curves show the variation of the geometric factor  $\beta$  and work function  $\phi$  as the Schottky emitter with  $r = 0.60 \mu\text{m}$  is heated at 1800 K for 24 h at the indicated value of  $I'$ .

For the Figure 1.20b shape the ratio of the flat side  $f$  to the radius  $r$  of the circle inscribed tangent to the (100) and side (110) planes remains constant at 0.828 as dictated by the cubic geometry of the W unit cell. However, if the profile view of the Figure 1.20b shape is such that the (112) ridges are normal to the viewer, the  $f/r$  ratio is 0.634. Thus, the profile view for most emitters examined show an  $f/r$  ratio between the former values. The emitter radius defined in this matter can vary by 30% for a given flat size depending on the profile view; thus, a better definition of emitter size would be the size of the central (100) flat from a top-down view as shown in Figure 1.20.

Studies have shown that the size of the (100) facet (and thus  $r$ ) increases with applied electric field after which a stable shape is again achieved. Figure 1.21 shows that equilibrating the emitter at increasing values of  $I'$  (or  $F$ ) at 1800 K decreases the  $\beta$  factor (measured from the  $I'$  (V) characteristics from the probe current from the central region of the (100) flat) ~35% over the range of  $I'$  investigated due to the increasing size of the (100) flat. A small reduction of ~4% in the work function is also noted as the flat size increases.

In the absence of an electric field or if the field is reduced below a critical value  $F_0$ , the net plane collapse of a fully faceted emitter restarts and overall emitter dulling occurs; that is, emitter radius increases. The rate of increase ( $dr/dt$ ) of the emitter radius for a spherical shape is given by<sup>42</sup>

$$\frac{dr}{dt} = \frac{1.25\gamma\Omega^2 D_0 \alpha}{AkTr^3} \exp\left(\frac{-E_d}{kT}\right) \quad (1.12)$$

where  $\Omega$  is the volume per atom ( $\Omega = 1.57 \times 10^{-29} \text{ m}^3/\text{atom}$  for W),  $A$  is the surface area per atom ( $A = 1 \times 10^{-19} \text{ m}^2/\text{atom}$ ),  $\gamma$  is the surface tension ( $\gamma = 2.9 \text{ N/m}$ ),  $\alpha$  is the emitter cone half angle,  $D_0$  is the surface diffusivity constant ( $D_0 = 4 \times 10^{-4} \text{ m}^2/\text{s}$  for clean W), and  $E_d$  is the activation energy for surface diffusion. Thus, for a clean W emitter, Equation 1.12 becomes<sup>41</sup>

$$\frac{dr}{dt} = \frac{2.6 \times 10^5 \alpha}{Tr^3} \exp\left(\frac{-E_d}{kT}\right) (\mu\text{m/s}) \quad (1.13)$$

where  $\alpha$  is in radians and  $r$  is in micrometers and  $E_d = 3.14$  eV. Integrating Equation 1.13 gives the relationship between the initial  $r_i$  and final  $r_f$  emitter radius as follows:

$$r_f^4 - r_i^4 = 1.04 \times 10^6 \frac{\alpha t}{T} \exp\left(-\frac{E_d}{kT}\right) (\mu\text{m}^4) \quad (1.14)$$

For the clean W(100) emitter an experimental study of the change in radius in the temperature range of 1850–2300 K gave the following form of Equation 1.14:

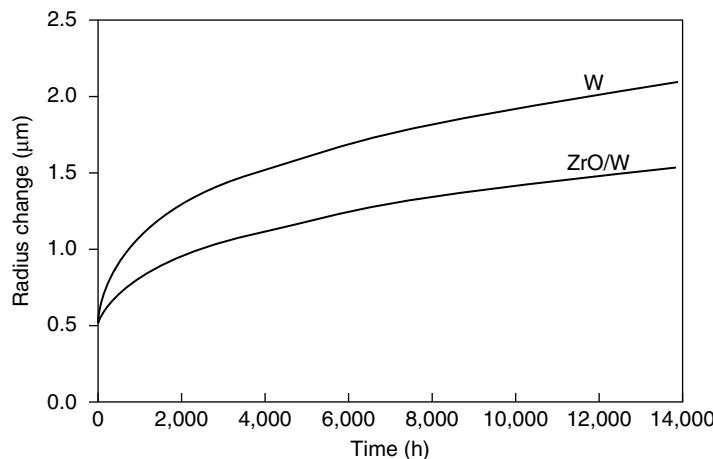
$$r_f^4 - r_i^4 = 2.18 \times 10^5 \frac{\alpha t}{T} \exp\left(-\frac{2.78}{kT}\right) (\mu\text{m}^4) \quad (1.15)$$

where  $E_d = 2.78$  eV. The latter value of  $E_d$  for the W(100)-oriented emitter is slightly lower than the value of 3.14 eV for a W(110)-oriented emitter and the pre-exponential term of Equation 1.15 compares reasonably well with the calculated value in Equation 1.14. With the ZrO layer present the measured dulling rate has been determined to be given by the following Equation 1.15 parameters:

$$r_f^4 - r_i^4 = 1.41 \times 10^{15} \frac{\alpha t}{T} \exp\left(-\frac{6.53}{kT}\right) (\mu\text{m}^4) \quad (1.16)$$

The ZrO layer not only increased the activation energy from 2.78 to 6.53 eV, but also dramatically increased the pre-exponential factor. These two factors partially compensate each other and lead to a modest lowering of the overall dulling rate.

Figure 1.22 shows the time variation of  $r$  at 1800 K based on the Equations 1.15 and 1.16 parameters for the W(100) and ZrO/W(100) emitters with an initial starting value of  $r = 0.5$   $\mu\text{m}$ . Because of the reciprocal dependence of  $dr/dt$  on  $r^3$ , the rate of dulling decreases rapidly with increasing  $r$ . Although the ZrO/W cathode has a lower dulling rate than the clean W(100), the zero field dulling rate of both the clean and ZrO-coated W(100) emitter is still unacceptably high for practical use at or above the normal operating temperature of 1800 K.



**FIGURE 1.22** Plots of the zero field radius change with time for a clean W(100) and ZrO/W(100) emitter using the Equations 1.15 and 1.16 parameters with  $T = 1800$  K and an emitter shank cone semiangle of 0.27 rad. Initial radii were 0.50  $\mu\text{m}$ .

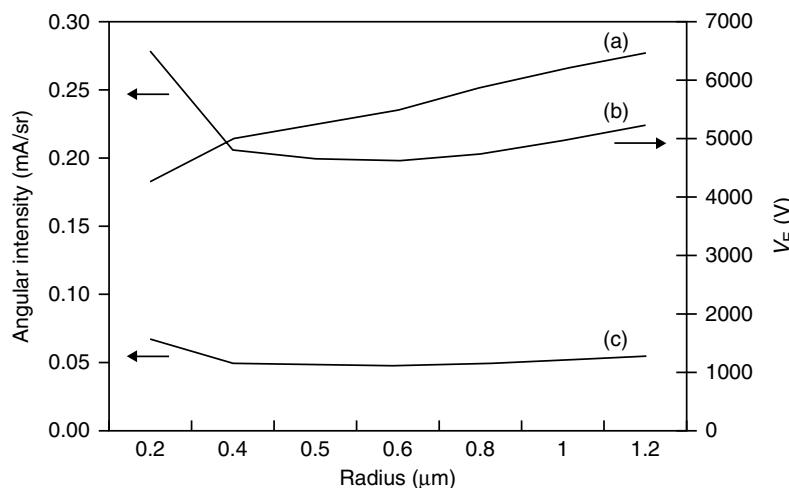
As mentioned previously, with a constant applied electric field of sufficient strength, the emitter dulling at 1800 K via the net plane collapse process leads to a faceting of the (100) plane and ultimate cessation of emitter dulling and net plane collapse. The electric field strength must exceed a minimum value of  $F_0$ ; that is, the field where the macroscopic surface tension stress and electric field stresses reach a balance. An analysis of the effect of an external applied electric field using an idealized model consisting of a hemispherical emitter apex has shown that the rate of dulling given by Equation 1.12 is modified as follows:<sup>34</sup>

$$\left(\frac{dr}{dt}\right)_F = \left(1 - \frac{F^2 r}{8\pi\gamma}\right) \frac{dr}{dt} \quad (1.17)$$

where  $\gamma$  is the surface tension. If the applied field is such that the term in brackets is equal to zero (i.e.,  $F_0 = (8\pi\gamma/r)^{1/2}$ ), the rate of dulling approaches zero.

In the case that  $F > F_0$ , emitter buildup occurs until a stable emitter shape (e.g., the end form shown in Figure 1.5b) is achieved. It has been determined that the end form shown in Figure 1.5b is achieved and stabilized if the apex electric field is  $\geq 0.8$  V/nm for an emitter radius of 500 nm. This corresponds to a measured angular current intensity  $I'$  of  $\sim 0.2$  mA/sr. Setting  $F = F_0 = 0.8$  V/nm in Equation 1.17, a value of  $\gamma = 1.41$  N/m is calculated for the ZrO/W emitter. This compares with a value of 2.9 N/m obtained for a clean W emitter.<sup>41</sup>

From Equations 1.1 through 1.5 and Equation 1.8 in combination with the condition  $F_0 = (8\pi\gamma/r)^{1/2}$ , one can determine the variation of  $I'$  and  $V_0$  (where  $V_0 = F_0/\beta$ ) with  $r$ . The result is shown in Figure 1.23 where the variation of both  $I'$  and  $V_0$  with  $r$  is given for two values of  $\phi$ . Although the values of the extraction voltage  $V_0$  are independent of  $\Phi$ , the corresponding values of  $I'$  decrease with increasing  $\phi$ . For a specific value of  $r$ , if the value of the extraction voltage  $V_0$  and, hence, the corresponding value of  $I'$  are below the relevant curves shown in Figure 1.23, ring collapse and concomitant emitter dulling will occur. This, in turn, will lead to the beam current instability noted in Figure 1.19. In contrast, if the extraction voltage and the corresponding  $I'$  values exceed the values shown in Figure 1.23 for a specific value of  $r$ , the facet size increases slightly over time but ring collapse and emitter dulling will not commence.



**FIGURE 1.23** The extractor voltage  $V_0$  (a) and corresponding values of  $I'$  (for which emitter dulling and concomitant ring collapse ceases) are shown as a function of  $r$ . Two  $I'$  curves are shown at work function values of 2.95 eV (b) and 3.20 eV (c). The Figure 1.8 electrode geometry was used.

In summary, it is concluded from this analysis that a stable emitter geometry for the SE cathode, and hence stable emission, can be realized if the emitter shape shown in Figure 1.5b is formed and  $F \geq F_0$ .

## 1.8 TOTAL ENERGY DISTRIBUTION

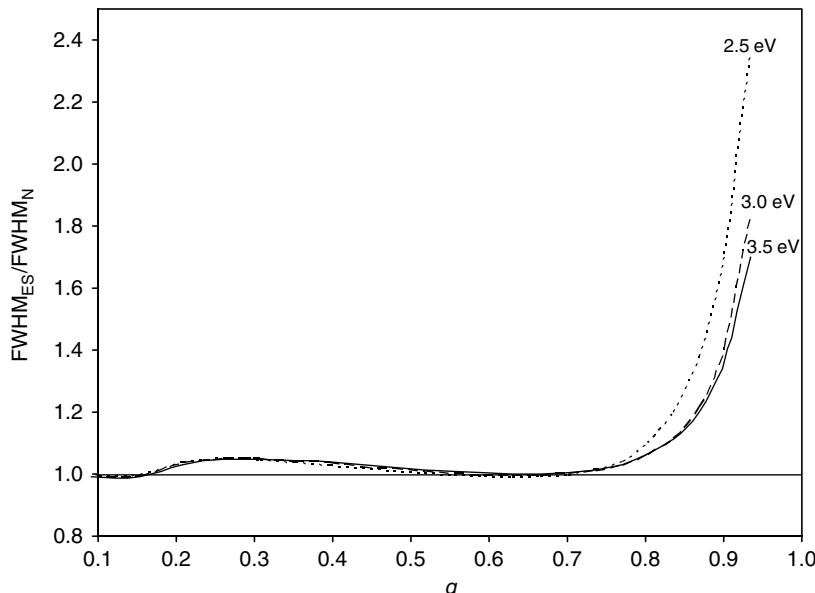
A significant contributor to the total aberration load in most electron microprobe systems is the chromatic aberration. Since the magnitude of the chromatic aberration is directly proportional to the width of the energy distribution of the electron beam, it is of considerable importance to understand the energy distribution of the emitted electrons in the extended Schottky regime. The following analytical expression for the TED for the extended Schottky regime can be derived:<sup>12</sup>

$$J_{\text{ES}}(\varepsilon) = \frac{J_S q}{kT} \ln \left[ 1 + \exp \left( \frac{\varepsilon}{qkT} \right) \right] \left[ 1 + \exp \left( \frac{\varepsilon + \varphi - E_0}{kT} \right) \right]^{-1} \quad (1.18)$$

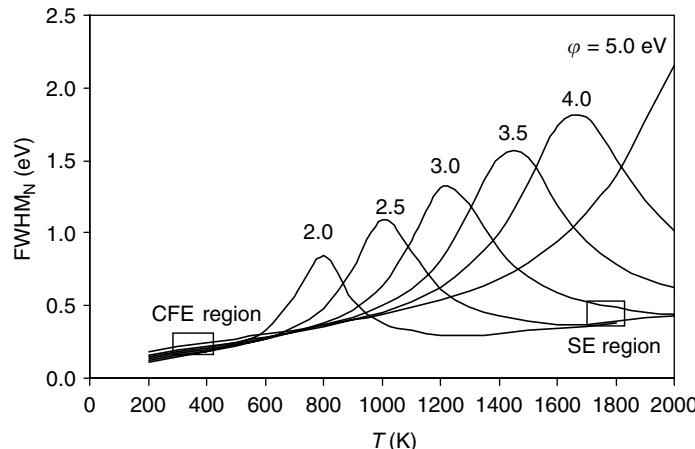
where  $\varepsilon$  is the electron energy with respect to the barrier maximum and  $E_0$  is the Schottky reduction of the work function barrier given by

$$E_0 = \frac{e^{3/2} F^{1/2}}{(4\pi\varepsilon_0)^{1/2}} \quad (1.19)$$

In Figure 1.24 the full width at half maximum ( $\text{FWHM}_N$ ) values of the TED obtained by the numerical method described in Section 1.4 are compared with the  $\text{FWHM}_{\text{ES}}$  values obtained from Equation 1.18. Unlike the current density ratio  $J_{\text{ES}}/J_N$  in Figure 1.7, the ratio  $\text{FWHM}_{\text{ES}}/\text{FWHM}_N$  is relatively close to 1 for  $q < 0.7$ , thereby giving confidence that the Equation 1.18 for the TED is accurate for  $0 < q < 0.7$ .



**FIGURE 1.24** Graph shows the ratio of the full width at half maximum ( $\text{FWHM}_{\text{ES}}$ ) to the numerically calculated  $\text{FWHM}_N$  for  $T = 1800$  K and the indicated work function values.



**FIGURE 1.25** Curves show the variation of the full width at half maximum ( $\text{FWHM}_N$ ) of the total energy distribution (TED) with temperature at the indicated values of work function and for  $J = 2 \times 10^7 \text{ A/m}^2$ . Boxes show the typical operating range of the cold field emission (CFE) and Schottky emission (SE) sources.

In Figure 1.25 the  $\text{FWHM}_N$  values versus  $T$  at constant  $J$  show that a maximum in the  $\text{FWHM}$  occurs for each value of  $\phi$  and it occurs at decreasing values of  $T$  as  $\phi$  decreases. In the SE regime (i.e.,  $\text{FWHM}$  values to the right side of the peaks in Figure 1.25), a minimum  $\text{FWHM}$  value occurs as a function of temperature for each work function value.

It becomes clear from this graph that a crucial factor for realizing a small  $\text{FWHM}$  of the TED for a high-temperature Schottky source is a low value of  $\phi$ . What this means is that in the extended Schottky regime an increasing number of the emitted electrons escape over the Schottky barrier and thereby more closely approximate the pure Schottky regime where the  $\text{FWHM}$  of the TED approaches  $\sim 2 \text{ kT}$ . The boxes in Figure 1.25 show that the  $\text{FWHM}$  of the TED in the operating range of the ZrO/W SE cathode is not greatly different from that of a CFE cathode.

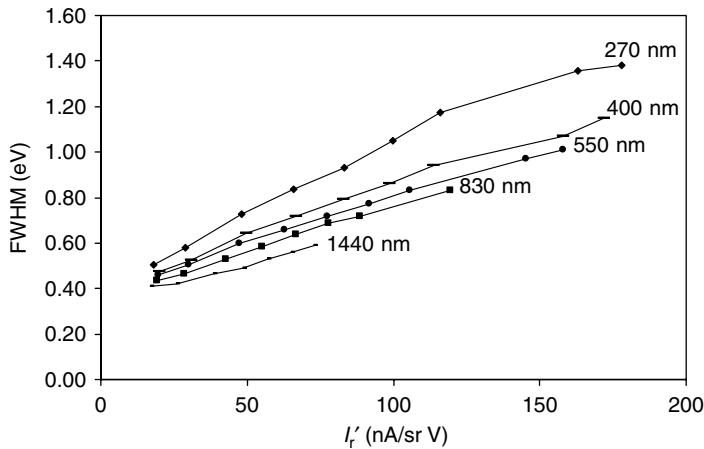
It can be shown<sup>43</sup> that the beam current  $I$  in an electron-focusing column is proportional to the reduced current angular intensity  $I'_r = I'/V_E$  and beam voltage  $V_i$  as follows:

$$I = I'_r \pi (M\alpha_i)^2 V_i \quad (1.20)$$

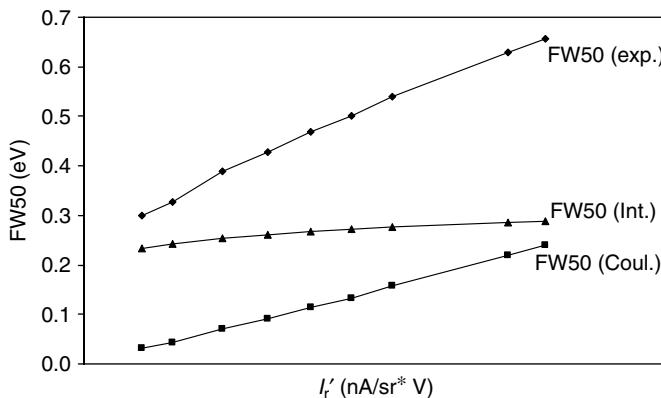
where  $M$  is the column magnification and  $\alpha_i$  is the beam convergence semiangle at the target. Thus, for electron optical applications where chromatic aberration dominates, a desirable electron source is one which minimizes the  $\text{FWHM}$  of the TED while maximizing  $I'_r$ .

Measurements of the  $\text{FWHM}$  versus  $I'_r$  for most high-brightness, point electron sources generally show an increase in  $\text{FWHM}$  values beyond that predicted by the theoretical expectations due to stochastic coulomb interactions external to the emitting interface (see Chapter 7). The ZrO/W SE cathode is no exception as shown in Figure 1.26, the energy spread increases with  $I'_r$  and decreasing  $r$  far beyond theoretical expectations.<sup>43</sup> A detailed study of the coulomb interactions carried out by Jansen<sup>44</sup> indicates that the energy spread  $\Delta E_c$  due to external coulomb interactions has the form:  $\Delta E_c \propto I^n/V_E^m$ , where  $n = 0.5-1$  and  $m \sim 1$ . Thus, by increasing  $r$ , the value of  $V_E$  for a specific value of  $I'$  (or  $I'_r$ ) increases, thereby causing a decrease in the energy spread. The ZrO/W SE source with  $\phi = 2.9 \text{ eV}$  has the desirable property of a low energy spread at 1800 K provided  $I'_r$  is not excessive.

It is of interest to extract the contribution of the coulomb interaction (referred to as the Boersch effect) from the experimental  $\text{FWHM}$  values of the TED. It can be shown<sup>44,45</sup> that using the energy



**FIGURE 1.26** Experimental values of the full width at half maximum (FWHM) of the total energy distribution versus  $I_r'$  are given for the indicated radii of the Schottky emission source at 1800 K.

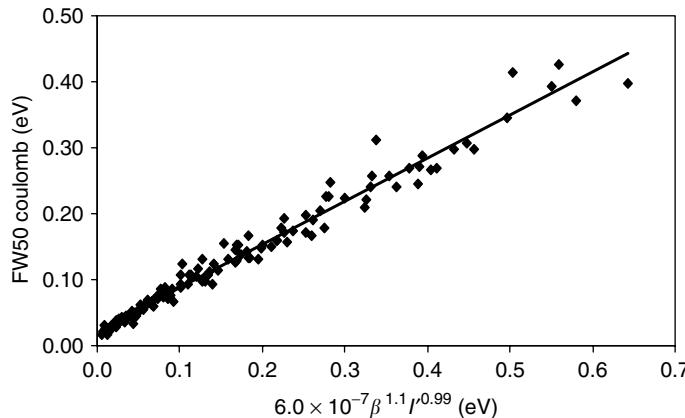


**FIGURE 1.27** Coulomb (Coul.) and intrinsic (Int.) contributions to the experimental (exp.) FW50 values are shown for a 550 nm radius Schottky emission source.

spread containing 50% of the current (FW50), an expression relating the FW50 due to the coulomb interactions to the respective intrinsic (i.e., theoretical) and experimental FW50 values can be obtained as follows:

$$\text{FW50(exp.)}^\gamma = \text{FW50(int.)}^\gamma + \text{FW50(coul.)}^\gamma \quad (1.21)$$

where an effective  $\gamma$  can be found independent of the shape of the respective distributions if the FW50 values are used. By use of two fit parameters that allow the convolution of the coulomb and intrinsic TED curves to match the experimental TED curve,<sup>43</sup> a value of  $\gamma = 1.56$  was obtained for the SE regime. In Figure 1.27 the intrinsic and coulomb contributions to the total FW50 values are shown as a function of  $I_r'$  for the SE source with  $r = 550$  nm. These results clearly show that the stochastic coulomb interaction is the main contribution to the increase in the FW50 (exp.) values with increasing  $I_r'$ .



**FIGURE 1.28** FW50 values due to coulomb interactions for the Schottky emission source are shown to obey a simple power law with respect to  $\beta$  and  $I'$ .

It was determined that the FW50 (Coul.) variation with  $I'$  and  $\beta$  could be fitted reasonably well to a simple power law function of  $I'$  (in mA/sr) and  $\beta$  (in m<sup>-1</sup>) as shown in Figure 1.28. Inserting the Equation 1.11 dependence of  $\beta$  on  $r$  and  $L_{TA}$  into the Figure 1.28 power law results in FW50 (Coul.)  $\propto I'^{0.99} r^{-1.1} L_{TA}^{-0.695}$ . To reduce the energy broadening, one must operate the SE source at the lowest value of  $I'$  and largest value of  $r$  consistent with beam current and source brightness requirements for a particular optical application.

## 1.9 Emitter Brightness

The reduced brightness  $B_r$  of a point source is related to the virtual source  $d_v50$  and  $I'_r$  as follows:

$$B_r = \frac{4I'_r}{\pi d_v 50^2 V_E} \quad (1.22)$$

If one assumes a Gaussian emission distribution from the virtual source, the diameter  $d_v50$  containing 50% of the current is given by<sup>46,47</sup>

$$d_v50 = 1.67 \frac{r}{m} \left( \frac{\langle E_t \rangle}{V_E} \right)^{1/2} \quad (1.23)$$

where the average initial transverse energy  $\langle E_t \rangle$  of the emitted electrons is  $kT$  for the SE regime.

From Equations 1.22 and 1.23 and Table 1.2 the reduced brightness can be calculated. However, it is emphasized that the expected increase in  $d_v50$  and resulting decrease in  $B_r$  due to stochastic Coulomb interactions<sup>44</sup> is not included in the calculation of  $d_v50$ . The results, given in Table 1.3 for  $I' = 0.20$  mA/sr, indicate that brightness levels in the range of low  $1 \times 10^8$  A/m<sup>2</sup> sr V can be expected for small values of  $r$ . The source figure of merit as defined by  $B_r/\text{FW50}$  improves as  $r$  decreases. Increasing  $B_r/\text{FW50}$  by increasing  $I'$  much above 0.2 mA/sr is unlikely since FW50 increases with  $I'$ , as shown in Figure 1.28, and values of  $d_v$  obtained from Equation 1.23 will be overly optimistic due to an increasing contribution of coulomb interaction.

**TABLE 1.3**

**Calculated Values for  $d_v$  and  $B_r$  Using Equations 1.22 and 1.23 and Experimental Data for  $r$ ,  $V_E$ ,  $m$ , and FW50**

$r$ (nm)	$V_E$ (V)	$m$	$d_v50$ (nm)	$B_r$ ( $\text{A}/\text{m}^2 \text{ sr V}$ )	FW50 (eV)	$B_r/\text{FW50}$ ( $\text{A}/\text{m}^2 \text{ sr V}^2$ )
300	3900	0.19	17	$2.4 \times 10^8$	0.42	$5.6 \times 10^8$
500	4000	0.20	26	$9.4 \times 10^7$	0.36	$2.6 \times 10^8$
1000	4300	0.18	56	$1.9 \times 10^7$	0.33	$5.8 \times 10^7$

## 1.10 CURRENT FLUCTUATIONS

Current stability considerations typically fall into two frequency ranges: short-term ( $f > 0.1$  Hz) and longer-term ( $f < 0.01$  Hz) current drift. The physical processes involved in both of these cases are work function and localized electric field fluctuations. Work function fluctuations come about via adsorption/desorption and local surface concentration fluctuations of certain gases. Local field fluctuations come about via macroscopic geometry variations as described in Section 1.7 or atomic size displacement due to ion bombardment. If work function fluctuations due to the adsorption and desorption of gases and geometric fluctuations are eliminated, high-frequency work function fluctuations due to self-diffusion or diffusion-induced concentration fluctuations of adsorbed substrate atoms in the probed area still exist. At elevated temperature (i.e.,  $\sim 1800$  K), substrate atom concentration fluctuations due to self-diffusion are the primary mechanisms of beam current noise generation since residual gas adsorption will be negligible at pressures  $< 1 \times 10^{-8}$  torr.

Figure 1.29 shows a spectral density function  $\omega(f)$  of the current fluctuations (normalized to the probe current  $I_p$ ) as a function of frequency<sup>18</sup> for various aperture acceptance solid angles  $\Omega$ , where

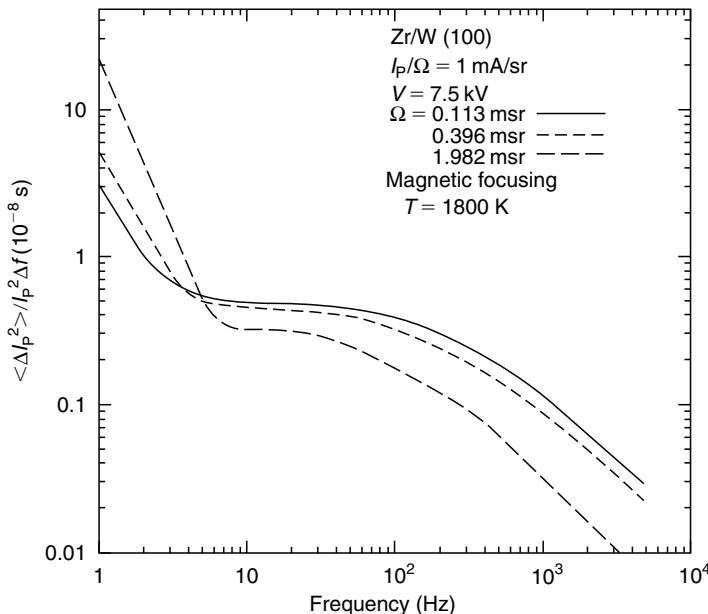
$$\langle \Delta I_p^2 \rangle = \int_f^{f+\Delta f} \omega(f) df \quad (1.24)$$

The integration of Equation 1.24, which is given by the area under the Figure 1.29 curves, yields the total mean square, fractional noise  $\langle \Delta I_p^2 \rangle / I_p^2$  in the frequency interval 0 to  $f$ . In conjunction with a similar measurement of the current fluctuation versus emitter radius, the following empirical relationship valid for  $f > 10$  Hz and  $T = 1800$  K was derived:

$$\frac{\langle I_p^2 \rangle}{I_p^2} = \frac{20}{(1 \times 10^4 + \alpha r^2)} + \frac{3.2 \times 10^{-10} f}{\pi I' \alpha^2} \quad (1.25)$$

where  $\alpha$  (the acceptance semiangle),  $r$ ,  $I'$ , and  $f$  are in mrad, nm, mA/sr, and Hz, respectively. The first term in Equation 1.25 is the dominant term in the low frequency range and is due to surface diffusion-based flicker noise; the last term is the statistical shot noise contribution. In the usual operating range of values for  $f$ ,  $\alpha$ , and  $r$  the total noise contribution is less than 1%. In the absence of shot noise and as  $\alpha r^2$  approaches zero, that is, a very small emitting area seen by the acceptance aperture, the percent total noise can reach 4–5%.

To minimize long-term current drift, the partial pressures of certain residual gases (see Section 1.11) must be  $< 1 \times 10^{-8}$  torr and remain constant. In addition, the electric field must be sufficiently high to prevent (100) plane collapse (see Figure 1.19). The long-term drift of the current from the central region of the (100) plane has been observed to be  $\leq 0.2\%/\text{h}$  in commercial focusing columns when the preceding environmental conditions are met.



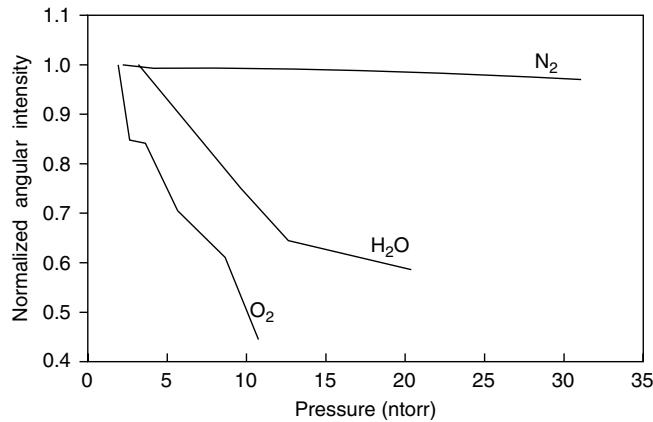
**FIGURE 1.29** Graph shows the experimental normalized spectral density function of the beam current versus  $f$  for the indicated acceptance solid angles at  $T = 1800 \text{ K}$  for the ZrO/W Schottky emission. The current angular intensity was 1 mA/sr.

## 1.11 Emitter Environmental Requirements

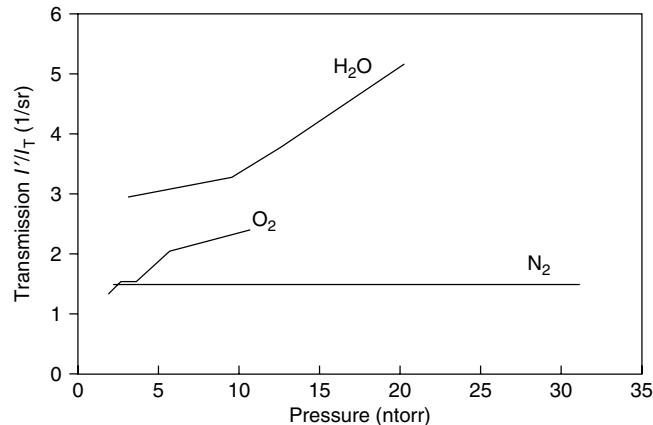
The ZrO/W cathode operated at elevated temperature is considerably less sensitive to residual gas pressures than field emission cathodes operating at room temperature. Electronegative gases, for example, O<sub>2</sub>, H<sub>2</sub>O, N<sub>2</sub>, CO, CH<sub>4</sub>, CO<sub>2</sub>, H<sub>2</sub>, and so on, adsorbed on a room-temperature cathode cause a substantial increase in work function. Ion and neutral bombardment of the cathode due to electron-stimulated desorption (ESD) from nearby anode surfaces can be as important as base residual gas pressure in causing long-term current instability. The elevated operating temperature of the SE cathode not only maintains low equilibrium coverage of adsorbed gases, but also rapidly anneals atomic displacements caused by ion bombardment. Nevertheless, the ZrO/W cathode operated at  $\sim 1800 \text{ K}$  has been found to be sensitive to partial pressures of two oxygen-bearing gases, H<sub>2</sub>O and O<sub>2</sub>.

Figures 1.30 and 1.31 show the variation of  $I'$  and the axial transmission  $I'/I_T$  with the partial pressure of three gases. Fortunately, neither O<sub>2</sub> nor H<sub>2</sub>O is typically present at significant partial pressures in well-baked, stainless steel vacuum enclosures. However, if a significant portion of the electron beam is allowed to impinge on a nearby oxide-coated anode surface, a local flux of O<sub>2</sub> formed by ESD can adsorb on the cathode surface. Figure 1.30 shows that partial pressures of both O<sub>2</sub> and H<sub>2</sub>O reduce the emission current, whereas N<sub>2</sub> has little effect. The reduction of the emission current due to these gases may take several hours to equilibrate after a pressure change; however, it is completely reversible throughout the pressure range. The emission level will ultimately stabilize after a pressure change if the gas partial pressure remains constant. Figure 1.31 shows that the overall beam transmission (defined as  $I'$  divided by the total emission current) increases with increasing O<sub>2</sub> and H<sub>2</sub>O pressure.

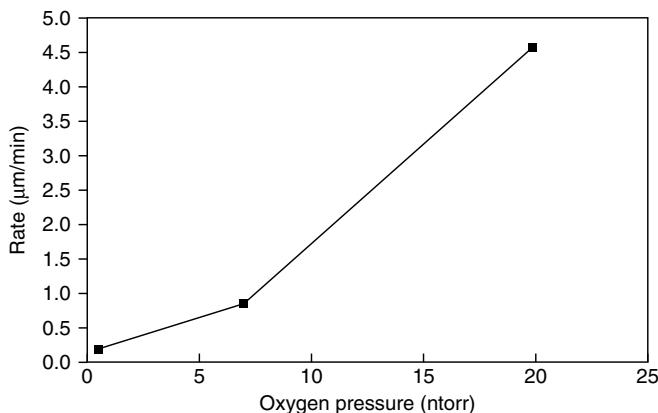
If one is operating the ZrO/W cathode under conditions where emitter dulling is occurring, the measured rate of net plane collapse (see Figure 1.17) is also affected by O<sub>2</sub> pressure and emitter temperature. As shown in Figure 1.32, the rate of net plane collapse at 1800 K increases with oxygen pressure. This, in turn, changes the frequency of the central emission instability event shown in Figure 1.17.



**FIGURE 1.30** Experimental measured variation of the equilibrium, normalized angular intensity of the ZrO/W Schottky emission with the partial pressure of the indicated gases. Emitter temperature was maintained at 1800 K.



**FIGURE 1.31** Same as Figure 1.25, but showing the variation of the beam transmission with partial pressure of the indicated gases.

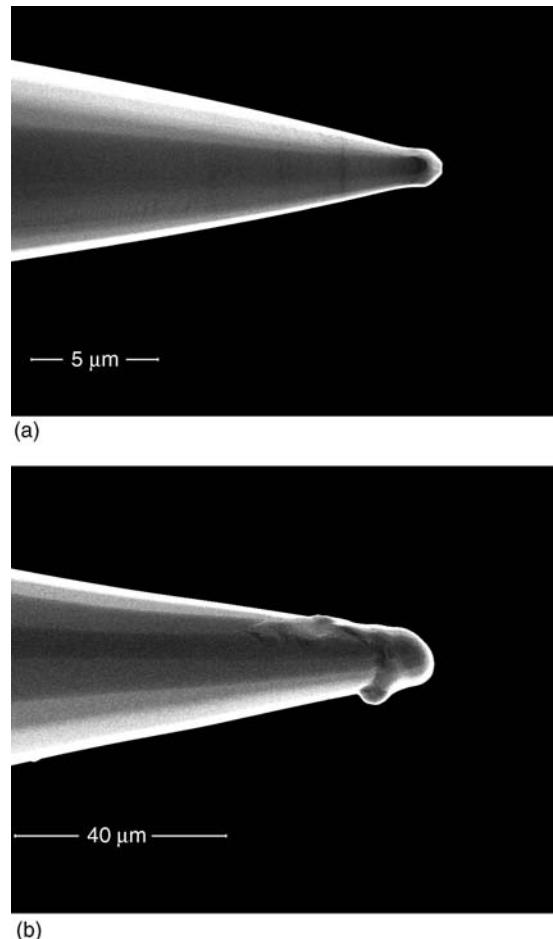


**FIGURE 1.32** Graph shows the effect of  $O_2$  pressure on the rate of (100) plane collapse (see Figure 1.5) for the ZrO emitter at  $T = 1800$  K. The emitter radius was  $0.8 \mu\text{m}$  and  $F \leq F_0$ .

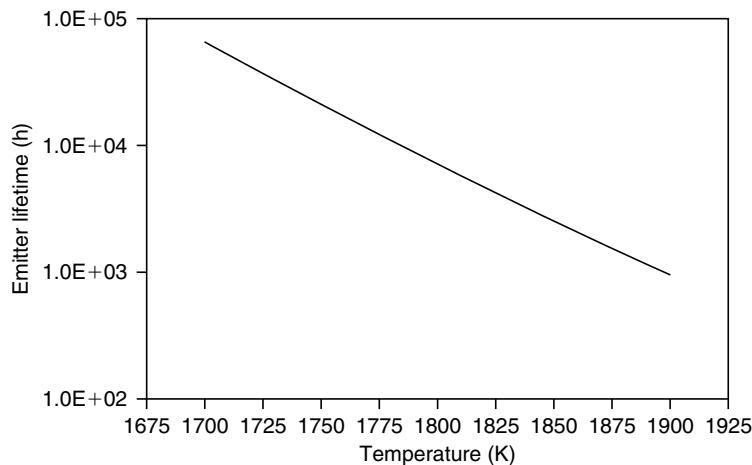
## 1.12 Emitter Life Considerations

Barring catastrophic failure, the usual life termination mechanism for a field emitter is gradual dulling of the emitter. Emitter life can be terminated by catastrophic dulling caused by an arc between the emitter and extractor electrode induced by inadvertent extractor electrode over-voltage or pressure burst in the emitter region. When compared with the clean W cathode operating at room temperature, the emitter radius of the ZrO/W cathode is a factor of several times larger. Experience has shown that larger radius emitters are better able to withstand unexpected arcs without significant dulling. Typically, a catastrophic emitter-dulling event for a ZrO/W emitter more or less leaves the emitter intact, except for a larger radius, as shown in Figure 1.33b, and functioning as a low-field Schottky emitter. The low-work-function (100) plane remains on axis after such an event and the axial angular intensity may not undergo a dramatic change at the normal extraction voltage. In fact, it is sometimes difficult to determine from the emission characteristics alone that a catastrophic dulling event has occurred. Clearly, however, the virtual source size of the emitter will greatly increase.

The natural event which terminates emitter life for the ZrO/W emitter is the disappearance by evaporation of the ZrO reservoir shown in Figure 1.4. The rate of evaporation of a 350- $\mu\text{m}$ -long by 36- $\mu\text{m}$ -thick Zr layer is given by the graph in Figure 1.34 as a function of emitter temperature.



**FIGURE 1.33** Scanning electron microscope photos of (a) a normal ZrO/W emitter with  $r = 0.30 \mu\text{m}$  and (b) an arced emitter with  $r = 3.2 \mu\text{m}$ .



**FIGURE 1.34** The curve shows the calculated emitter lifetime due to evaporation of the Zr reservoir (see Figure 1.4) versus emitter temperature. Calculation of Zr reservoir depletion time was made according to Equations 1.26 and 1.27 with a Zr reservoir assumed to be 36  $\mu\text{m}$  thick by 350  $\mu\text{m}$  in length.

The relationship between evaporation time  $t$ , reservoir volume  $V$ , and area  $A$ , and pressure  $P_{\text{zr}}$  and is given by

$$t = \frac{\rho_{\text{zr}} V}{A P_{\text{zr}}} \left( \frac{2\pi k T}{m_{\text{zr}}} \right)^{1/2} \quad (1.26)$$

where  $T$ ,  $m_{\text{zr}}$ , and  $\rho_{\text{zr}}$  are the temperature, mass, and density of the evaporating species, respectively. The vapor pressure of the evaporating material, assumed to be pure Zr, is given by

$$P_{\text{zr}} = 1.96 \times 10^{11} \exp\left(-\frac{68986.5}{T}\right) (\text{N/m}^2) \quad (1.27)$$

The Figure 1.34 graph shows that at 1800 K the cathode life with respect to evaporation of the Zr coating is several thousands of hours. From the results of a dozen ZrO/W cathodes operating in a commercial instrument environment, an average life with respect to evaporation of the ZrO/W layer of  $6400 \pm 1650$  h was obtained. In one case, a cathode life of 17,000 h was obtained with no appreciable change in radius.

## 1.13 SUMMARY

The results of numerous past investigations of the ZrO/W Schottky cathode have led to a substantial understanding of its fundamental mechanisms of operation and emission properties. Because of its robustness with respect to environmental factors and high-performance parameters, this cathode has found increasing use in commercial electron beam systems to satisfy the need for both higher resolution and higher current density. Although LaB<sub>6</sub> and CFE cathodes still dominate many of the SEM applications, both SEM and transmission electron microscope (TEM) instrument makers are joining with e-beam lithography manufacturers as substantial users of the Schottky point cathode.

When used at the proper value of  $I'$  (i.e., 0.2–1.0 mA/sr) and partial pressures of H<sub>2</sub>O and O<sub>2</sub> (i.e.,  $\leq 1 \times 10^{-9}$  torr), the critical emission characteristics, energy spread, source brightness, and short- and long-term current fluctuations are all very competitive with the CFE cathode and superior to all other known high-brightness cathodes.

## ACKNOWLEDGMENT

The authors wish to acknowledge considerable experimental contributions to the investigations of the ZrO/W cathode from D. W. Tuggle and N. A. Martin.

## REFERENCES

1. T. Hibi, 1955. *J. Electron Microsc.*, 3, 15.
2. S. Maruse and Y. Sakaki, 1958. *Optik*, 15, 485.
3. T. Everhart, 1967. *J. Appl. Phys.*, 38, 223.
4. L. W. Swanson and L. C. Crouser, 1969. *J. Appl. Phys.*, 40, 4741.
5. L. W. Swanson and N. A. Martin, 1975. *J. Appl. Phys.*, 46, 2029.
6. J. E. Wolf, 1975. *J. Vac. Sci. Technol.*, 12, 1169.
7. P. B. Mee, 1976. *J. Appl. Phys.*, 47, 3904.
8. J. M. Lafferty, 1951. *J. Appl. Phys.*, 22, 299.
9. A. N. Broers, 1967. *J. Appl. Phys.*, 38, 1991, 3040.
10. M. Futamoto, M. Nakazawa, K. Usami, S. Hosoki, and U. Kawabe, 1980. *J. Appl. Phys.*, 51, 3869.
11. S. G. Christov, 1966. *Phys. Status Solidi*, 17, 11.
12. L. W. Swanson and A. E. Bell, 1973. *Adv. Electron. Electron Phys.*, 32, 193.
13. V. N. Sherdnik, 1961. *Sov. Phys.-Solid State*, 3, 1268.
14. G. N. Fursei and S. A. Shakirova, 1966. *Sov. Phys.- Tech. Phys.*, 11, 827.
15. R. A. Collins and B. H. Blott, 1968. *Surf. Sci.*, 10, 349.
16. L. W. Swanson, 1975. *J. Vac. Sci. Technol.*, 12, 1228.
17. D. W. Tuggle and L. W. Swanson, 1985. *J. Vac. Sci. Technol.*, B3, 220.
18. L. W. Swanson and D. W. Tuggle, 1981. *J. Appl. Surf. Sci.*, 8, 185.
19. S. C. Lee, Y. Irokawa, M. Inoue, and R. Shimizu, 1995. *Surf. Sci.*, 330, 289.
20. S. Yamamoto, S. Hosoki, S. Kawase, and Y. Hirai, 1981. *J. Appl. Phys.*, 52, 1011.
21. J. E. Wolfe, 1979. *J. Vac. Sci. Technol.*, 16, 1704.
22. S. C. Le, H. Tamura, M. Inoue, and R. Shimizu, 1994. *Microbeam Anal.*, 3, 311.
23. L. R. Danielson and L. W. Swanson, 1979. *Surf. Sci.*, 88, 14.
24. H. Satoh, H. Nakane, and H. Adachi, 1995. *Proc. Eighth Int. Vac. Microelectron Conf.*, 425.
25. T. Kawano, Y. Takai, and R. Shimizu, 2000. *Jpn. J. Appl. Phys.*, 39, 577.
26. T. Kawakubo, Y. Saito, N. Miyamoto, H. Nakane, and H. Adachi, 2004. *J. Vac. Sci. Technol. B*, 22, 1258.
27. M. Troyon, A. K. Pandy, and J. P. Orlowski, 1989. *J. Vac. Sci. Technol.*, A7, 3112.
28. S. Hosoki, K. Takata, and H. Kaga, 1986. *Proc. Eleventh Int. Congr. Electron Microsc.*, Kyoto, 219.
29. H. Nishiyama, T. Ohshima, and H. Shinada, 1999. *Appl. Surf. Sci.*, 146, 382.
30. W. Schottky, 1914. *Z. Phys.*, 15, 872.
31. S. Dushman, 1930. *Rev. Mod. Phys.*, 2, 381.
32. A. Modinos, 1984. *Field, Thermionic and Secondary Electron Emission Spectroscopy*, Plenum Press, NY.
33. A. Bahm, G. A. Schwind, L. W. Swanson (to be published *JVST B* Nov/Dec 2008 issue).
34. R. L. Burden and J. D. Faires, 2001. *Numerical Analysis*, 7th ed., Brooks/Kole, Pacific Grove, CA.
35. C. F. Erying, S. Mackeown, and R. A. Millikan, 1928. *Phys. Rev.*, 31, 900.
36. J. A. Becker, 1951. *Bell Syst. Tech. J.*, 30, 907.
37. W. P. Dyke, J. K. Trolan, W. V. Dolan, and G. Barnes, 1953. *J. Appl. Phys.*, 24, 570.
38. N. K. Kang, D. W. Tuggle, and L. W. Swanson, 1983. *Optik*, 3, 313.
39. J. Nelder and R. Mead, 1965. *Comput. J.*, 7, 308.
40. W. P. Dyke, F. M. Charbonnier, R. W. Strayer, R. L. Floyd, J. P. Barbour, and J. K. Trolan, 1960. *J. Appl. Phys.*, 31, 790.
41. P. C. Bettler and F. M. Charbonnier, 1960. *Phys. Rev.*, 119, 85.
42. J. P. Barbour, F. M. Charbonnier, W. W. Dolan, W. P. Dyke, E. E. Martin, and J. K. Trolan, 1960. *Phys. Rev.*, 117, 1452.
43. G. A. Schwind, G. Magera, and L. W. Swanson, 2006. *J. Vac. Sci. Technol. B*, 24(6), 2897.
44. G. H. Janson, 1990. *Coulomb Interactions in Particle Beams*, Suppl. 21, Academic Press, NY.
45. J. E. Barth and M. D. Nykerk, 1999. *Nucl. Instrum. Methods Phys. Res. A*, 427, 86.
46. M. S. Bronsgeest, J. Barth, L. W. Swanson and P. Kruit, 2008. *JVST B*, 26, 949.
47. J. Worcester, 1969. *Br. J. Appl. Phys. (J. Phys. D)*, 2, 4457 (new old 38).

---

# 2 Liquid Metal Ion Sources

*Richard G. Forbes and (the late) Graeme L. R. Mair*

## CONTENTS

2.1	General Introduction .....	31
2.1.1	Field Ion Emission Sources .....	31
2.1.2	Liquid Metal Ion Sources—General Background .....	32
2.1.3	Conventions .....	33
2.2	Introduction to LMIS Behavior.....	34
2.2.1	The Shape of the Liquid Emitter .....	34
2.2.2	Electrohydrodynamic Spraying.....	35
2.2.3	Time-Dependent Behavior .....	35
2.2.3.1	The Most-Steady Regime.....	35
2.2.3.2	The Pulsation Regime .....	35
2.2.3.3	The Upper Unsteady Regime.....	36
2.3	Basic Theory—Ion Emission Related.....	36
2.3.1	Charged-Surface Models and Maxwell Stress .....	36
2.3.2	Field Evaporation.....	37
2.3.2.1	Introduction.....	37
2.3.2.2	Relevance of FEV Theory.....	37
2.3.2.3	FEV Rate Constant and Time Constant.....	37
2.3.2.4	Field Evaporation in the Equilibrium and Supply Limits.....	38
2.3.2.5	Prediction of Evaporation-Field Values .....	39
2.3.2.6	Escape Mechanism.....	40
2.3.2.7	Field Dependence of Activation Energy .....	40
2.3.3	Field-Emitted Vacuum Space Charge .....	42
2.4	Basic Theory—Electrohydrodynamics.....	44
2.4.1	The Formula for Surface Pressure Jump .....	44
2.4.2	Electrohydrostatic Equilibrium and Stability Conditions .....	45
2.4.3	Taylor's Mathematical Cone .....	46
2.4.4	The Steady High-Electrical-Conductivity Gilbert–Gray Cone-Jet.....	47
2.4.4.1	Pressure Relationships .....	47
2.4.4.2	Limiting Parameter Values .....	48
2.4.4.3	Nonturbulent Flow .....	48
2.4.4.4	Viscous-Loss Terms .....	49
2.4.4.5	Quasi-Ellipsoidal Model for the Liquid Cap.....	49
2.4.4.6	Pressure Drop in the Cone-Jet .....	50
2.4.4.7	The Zero-Base-Pressure Approximation .....	50
2.4.5	Apex Boundary Conditions .....	50
2.4.5.1	The Kingham and Swanson Model.....	50
2.4.5.2	Steady-State Boundary Condition Based on Current Densities .....	51

2.5	Steady-State Current-Related Characteristics.....	51
2.5.1	Basic Theoretical Formulation .....	52
2.5.2	Extinction and Collapse Voltages.....	54
2.5.3	Current–Voltage Characteristic above Extinction.....	54
2.5.3.1	Mair’s Equation.....	54
2.5.3.2	Mahoney’s Equation.....	55
2.5.3.3	The Effect of Flow Impedance .....	56
2.5.3.4	Effects of Temperature.....	57
2.5.4	Dependence of Cusp Length on Emission Current.....	57
2.5.5	Extinction Current.....	57
2.5.5.1	The Practical LMIS as a Physical Chaotic Attractor.....	58
2.6	Energy Distribution and Ion-Optical Characteristics .....	59
2.6.1	Introduction .....	59
2.6.2	Ion Energy Distributions .....	60
2.6.2.1	Onset Energy Deficits .....	60
2.6.2.2	Energy Spreads .....	61
2.6.2.3	Distribution Shape and the Low-Energy Tail .....	62
2.6.2.4	Effects of Temperature.....	62
2.6.3	Ion-Optical Effects .....	63
2.6.3.1	Angular Intensities .....	63
2.6.3.2	Optical Source Size.....	64
2.7	Ionization Mechanisms and Emitted Species .....	65
2.7.1	Elemental Ion Sources .....	65
2.7.1.1	Atomic Ions.....	65
2.7.1.2	Other Atomic Ion Generation Mechanisms .....	65
2.7.1.3	Cluster Ions .....	65
2.7.2	Alloy Ion Sources .....	66
2.8	Time-Dependent Phenomena .....	67
2.8.1	Pulsation .....	67
2.8.2	Droplet Emission in the Upper Unsteady Regime .....	68
2.8.3	Globule Emission from the Needle and Cone .....	69
2.8.4	Source Turn-On .....	70
2.8.5	Numerical Modeling of Liquid-Shape Development .....	71
2.9	Technological Issues .....	71
2.9.1	Fabrication of the Normal LMIS.....	71
2.9.1.1	Needle Fabrication .....	71
2.9.1.2	Needle Wetting.....	71
2.9.1.3	Reservoir and Heating Arrangements.....	73
2.9.2	Source Operation.....	73
2.9.2.1	General Conditions .....	73
2.9.2.2	Secondary Electrons .....	74
2.9.3	Alternative LMIS Geometries.....	74
2.10	Concluding Remarks.....	74
	Acknowledgment .....	76
A.1	Appendix: Numerical Data .....	76
	References .....	80

---

This chapter is dedicated to my friend and close colleague, Graeme Mair, the sole author of this chapter in the first edition. By his untimely death, the subject lost a colorful character and a deep thinker about difficult technical physics—RGF.

## 2.1 GENERAL INTRODUCTION

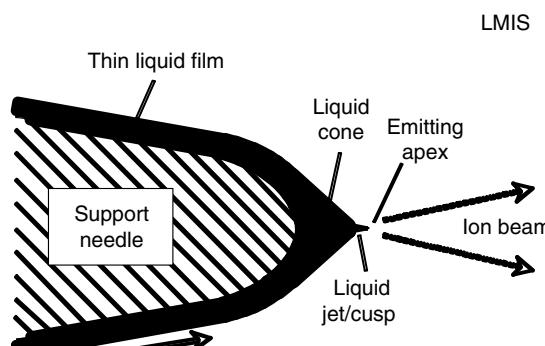
### 2.1.1 FIELD ION EMISSION SOURCES

This chapter provides an introduction to the *liquid metal ion source* (LMIS) and the physics of how it works. An LMIS is a field ion emission source. Such sources generate high-brightness positive ion beams from neutral atoms or molecules, by field-induced ion formation at the tip of a needle-like emitter. The very high electric fields needed, typically 10 V/nm or more, can be produced at convenient applied voltages only if the emitter is sharply pointed. Systems using a few or many such emitters operating in parallel have sometimes been investigated, but emission uniformity is a major problem. Emphasis here is on single-point emitters.

Of these emitters, the best known and most highly developed one is the low-drag blunt-needle LMIS, called here the *normal LMIS*. This source is commonly used in focused ion beam (FIB) machines.<sup>1-3</sup> As shown in Figure 2.1, ion formation occurs at the tip of a liquid metal jet drawn out by electrical forces from a roughly conical body of liquid metal, itself formed on top of a refractory metal needle (usually tungsten) with an apex radius of a few micrometers. The conical liquid body acts as a local reservoir and is supplied by thin-film flow along the needle surface from a main reservoir associated with the supporting structure. At the emission currents ( $\sim 2 \mu\text{A}$ ) normally used in FIB machines, the jet apex has a radius of  $\sim 1$  to 2 nm. The resulting high current densities (of order  $10^{11} \text{ A/m}^2$ ) immediately above the jet apex cause space-charge effects, which play an important role in LMIS physics. In particular, they are the main reason why, notwithstanding the small physical size of the emitting area, the apparent optical size is much larger (of the order of 50 nm in diameter). More generally, the behavior of a LMIS depends on a complex interplay between (1) electrohydrodynamic (EHD) effects on the liquid ionizant supply and (2) ion emission and space-charge effects.

The other established single-point ion emitter is the *gas field ionization source* (GFIS), based on surface field ionization (FI) of gas atoms at the tip of a refractory metal needle. Recently, a specialized GFIS—the so-called atomic-level ion source (ALIS<sup>TM</sup>)<sup>4</sup>—has been developed as a working helium ion source for commercial scanning ion microscopes. The critical feature of an ALIS is that the metal substrate is atomically sharp. The physics and ion optics of the GFIS are described in Chapter 3.

At the time of writing, a third type of point ion emitter—the *ionic liquid ion source* (ILIS)<sup>5</sup>—seems to be emerging. These sources are based on partially organic molten salts, which are liquid at room temperature, and have relatively high electrical conductivity. Although the emission mechanism will presumably be some form of field desorption, rather than field evaporation (FEV), aspects of its theory will be similar to LMIS theory. However, it seems premature to attempt detailed discussion here.



**FIGURE 2.1** Schematic diagram of an operating blunt-needle liquid metal ion source.

### 2.1.2 LIQUID METAL ION SOURCES—GENERAL BACKGROUND

The normal LMIS can use pure metals or alloys—our term LMIS covers both. The alloy variant is sometimes called a *liquid alloy ion source* (LAIS) or sometimes a *liquid metal alloy ion source* (LMAIS). The gallium LMIS is widely used because gallium has a low melting point and low vapor pressure, is relatively unreactive, and produces a beam that is nearly all Ga<sup>+</sup>, and because the Ga<sup>+</sup> ion is heavy enough to be effective for ion milling. The LAIS, and for some elements the LMIS, produce mixtures of ions; unwanted ions have to be removed by ion-optical filtering.

The normal LMIS rests on several historical developments. Although the scientific phenomena involved are much older, the LMIS had its technical origin in attempts to develop EHD colloid thrusters for spacecraft propulsion.<sup>6</sup> Research originally aimed at developing charged-metal-droplet thrusters,<sup>7,8</sup> from 1961 onward, showed that liquid metals generated large numbers of ions. In 1974, Krohn<sup>9</sup> suggested that EHD ion sources might provide very bright sources for ion microprobes, and by 1976 he tested a gallium source in a prototype microprobe.<sup>10</sup>

All these early devices used a liquid supply via a capillary tube, which was the standard EHD source procedure. This is less steady than the normal LMIS and more difficult to control, particularly at low emission currents. However, it gives higher emission currents and is still under development in spacecraft applications, as part of the technology of field emission electric propulsion (FEEP). Following Taylor's theory,<sup>11</sup> it was assumed that the emission took place at the top of a liquid cone formed at the tube end.

At UKAEA Culham in the early 1970s, a FEEP research project aimed to develop higher-thrust devices by operating many sources in parallel, using arrays of needles. As a follow-up, Clampitt and colleagues investigated the current–voltage characteristics of a single-point emitter where liquid metal is supplied to the apex of a single sharp needle by thin-film flow on its surface. The emission current was relatively small.

The blunt-needle LMIS was discovered accidentally when S. Ventakesh (an undergraduate student on a professional placement year) raised the voltage too high and blew the end off the sharp needle. Her supervisor, Clampitt, then discovered that the emission currents were much higher. In 1975, high emission currents had been obtained for several metals,<sup>12</sup> including gallium. It was soon established<sup>13</sup> (and later confirmed<sup>14</sup>) that this was due to the formation of a liquid cone-jet at the emitter apex, and it was assumed<sup>13</sup> that liquid flow was assisted by grooves or roughness on the needle surface. Later, liquid supply was improved by deliberately creating grooves or roughness on the needle surface:<sup>15–17</sup> this reduces certain electrical and viscous flow problems, allowing *low-drag* liquid flow.

At present, the LMIS is used mainly for ion milling or gas-assisted nanoscale materials processing,<sup>3</sup> but there is continued interest in its use as a source of specific atomic ions<sup>18</sup> or cluster ions, as well as the continuing FEEP interest,<sup>19,20</sup> particularly for missions that require low, continuously adjustable thrust. There is also renewed interest in its possible use as an electrosprayer for depositing thin films of controlled geometry,<sup>21</sup> and there is interest in relating the behaviors of the ILIS<sup>5,22</sup> and LMIS.

For FIB machines, desirable source properties include

- Ability to provide desired ion species with minimum of unwanted species
- High brightness, which needs adequate emission current, small optical source size, and small ion energy spread
- High efficiency of ionization
- High general stability
- Long lifetime, which means robustness in normal operating conditions
- Need for easy and cheap manufacture, and easy installation

For other uses modified criteria may apply; in particular, high emission current may be more important than high brightness.

The normal LMIS is an efficient ion source, but its performance is constrained by the physics of how it works. Knowledge of source physics aids assessment of the merits and limits of existing

designs, the possibilities for improvement, and the effects on machine performance, and may aid the interpretation of scientific and technical results. The most recent overviews of source physics are in the original version of this handbook, in Ref. 2, and in the articles by Niedrig<sup>23</sup> and by Forbes.<sup>24</sup>

When this handbook was first published in 1997, there was only one LMIS textbook,<sup>1</sup> and this concentrated more on fundamentals and source technology than applications. Since then, the use of FIB machines in materials research has much increased, and two further textbooks have appeared on FIB and its applications.<sup>2,3</sup> There has been useful progress on LMIS fundamentals, and the focused helium ion beam machines<sup>4</sup> have appeared. As technology and applications are adequately discussed elsewhere, and this handbook contains a separate chapter (Chapter 3) on GFIS, this updated chapter concentrates on LMIS fundamentals, including the (often unwanted) secondary LMIS role as an EHD sprayer of charged droplets.

An aim of this chapter is to show how recent developments in relevant basic science (often not well recognized in the literature) improve understanding of LMIS behavior and limitations. Rightly, the main focus on LMIS work moved to applications many years ago, but hopefully this chapter may be useful to those who want to know better how the machines they use work, and to those who are involved in the development of new generations of sources.

The plan of this chapter, after an introduction, is to deal first with relevant basic science, then develop the theory of the so-called metallic Gilbert–Gray (GG) cone-jet, and then show how relevant LMIS operational and ion-optical properties relate to those of the cone-jet. This is followed by brief comments on technological issues and concluding remarks. We do not aim to be comprehensive in our discussion or complete in our referencing. The articles cited are those thought to be most relevant. A comprehensive LMIS/FIB bibliography, up to about 1990, was compiled by McKenzie and Smith.<sup>25</sup>

### 2.1.3 CONVENTIONS

This chapter uses the term *steady* for situations in which liquid flows smoothly through shapes that do not change with time. Steady flow can be quantified by the *flux*  $R$  (amount of substance transported per unit time);  $R$  is most conveniently measured in atoms per second, treating the *atom* as an atomic-level unit of amount of substance. For dimensional consistency,  $n_1$  is used to denote a *count of one atom*. The *flux density*  $Z$  (flux per unit area) relates to the flow of neutral atoms, each of mass  $m$  and moving with speed  $v$ , by

$$Z = (n_1/m)\rho v = n_1 v / \omega \quad (2.1)$$

where  $\rho$  is the material density and  $\omega$  ( $= m/\rho$ ) its atomic volume.

After emission, each ion has an average charge  $ze$ , where  $e$  is the elementary positive charge. If ions are emitted in two or more charge states, then  $z$  is nonintegral and may vary with applied voltage. The emission current density  $J$  and emission current  $i$  corresponding to flux density  $Z$  and flux  $R$  are

$$J = (ze/n_1)Z, \quad i = (ze/n_1)R \quad (2.2)$$

The ratio  $(e/n_1)$  is the *Faraday constant*, but in LMIS theory it is often more convenient to write

$$J/Z = i/R = ze/n_1 \cong (1.602177 \times 10^{-13} z) \mu\text{A}/(\text{atom/s}) \quad (2.3)$$

Equation 2.3 shows that, for given ionization conditions, a liquid flux has an *equivalent emission current*. For an LMIS, it is convenient to set up hydrodynamic flow equations in terms of the equivalent electrical currents and current densities, using Equation 2.3, and this is done here. The subscripts F and S distinguish between quantities applicable to FEV and liquid supply, respectively.

More generally, this chapter uses the conventions of the international system of measurement, including the rationalized meter–kilogram–second (rmks) equation system, but does not adhere

rigidly to SI units. The 2006 values of the fundamental physical constants<sup>26</sup> are used, and (as a convention to identify their status) values of universal constants are given to seven significant figures.

Table A.2 lists the most important quantities defined and gives their values for the metals of most interest as possible elemental ion sources.

## 2.2 INTRODUCTION TO LMIS BEHAVIOR

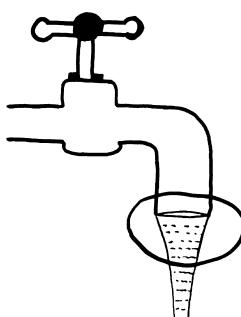
### 2.2.1 THE SHAPE OF THE LIQUID Emitter

Formation of a liquid cone is essential to the normal LMIS but is a more general liquid phenomenon. Electrically induced liquid cones were first reported by Gilbert<sup>27</sup> in 1600, who placed rubbed amber above water drops. Electrically induced liquid spraying was first reported by Gray<sup>28</sup> in 1732, who noted the following: (1) pull-out of a liquid jet from a water-cone apex, (2) droplet formation (more plentiful when using hot water), (3) formation of a cone with mercury, and (4) a brief electrical discharge at the cone apex. This discharge involved the emission of both light and sound, occurred with both water and mercury (but more strongly with mercury), and was immediately followed by the collapse of the liquid cone. In 1741, Desaguliers<sup>29</sup> showed, in demonstrations before the Royal Society of London, that the emitted droplets were charged and hypothesized that rain clouds might contain charged water droplets. Bose<sup>30</sup> and Nollet (see Ref. 31, p.5) also contributed to early work.

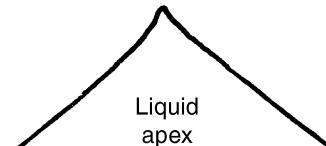
However, these liquid cones are usually called *Taylor cones*, because Taylor<sup>11</sup> (in 1964) was the first person to find a conical solution to the classical condition for the electrohydrostatic (EHS) equilibrium of an electrically conducting liquid (see Section 2.4.3).

At higher applied voltages there is a greater electrostatic *pull*, and liquid flow occurs. In this case, experimental liquid shapes do not correspond to Taylor's static-equilibrium mathematics. The cone pulls out into a jet. This is a well-known hydrodynamic effect, which is also seen, for example, when gravity pulls a thin stream of water from a tap or faucet (Figure 2.2). We call these experimental electrically induced liquid shapes *GG cone-jets*, reserving the name *Taylor cone* for Taylor's ideal, static, and classical mathematical shape. *Cone-jet* is a name used in recent EHD articles.<sup>32</sup>

At the low emission currents (of order 2  $\mu\text{A}$ ) used in FIB machines, the so-called jet is extremely short and probably more in the nature of a 1–2 nm bump at the top of a slightly distorted Taylor cone. Exact shapes are not known but are most likely a smaller version of the shape reported by Benassayag<sup>14,33</sup> for a 5  $\mu\text{A}$  emission current (see Figure 2.3).



**FIGURE 2.2** A well-known hydrodynamic analogy for the liquid metal ion source (LMIS) driving force. Here, gravity pulls water from the pipe, forming a *vena contracta*. With an LMIS, the driving force is the electrostatic Maxwell stress.



**FIGURE 2.3** Shape of the apex of an operating Ga liquid metal ion source at an emission current of 5  $\mu\text{A}$ , as observed in a high-voltage electron microscope. (Redrawn from Figure 7 in Benassayag, G., Etude du mécanismes d'émission des sources d'ions à métal liquide. Cas des métaux nobles, Docteur thesis, University Paul Sabatier, Toulouse, 1984; see also Benassayag, G. et al., *Ultramicroscopy* 16, 1–8, 1985.)

### 2.2.2 ELECTROHYDRODYNAMIC SPRAYING

With nonmetallic liquids, GG cone-jets mainly emit charged droplets. This is called *EHD spraying*, *electrospraying*, or *electrostatic atomization*. The term *EHD spraying* is favored here. Strictly, spraying from a single GG cone-jet is only one of the several spray modes;<sup>32</sup> others involve alternate or simultaneous spraying from two or more cone-jets.

The liquid metal GG cone-jets are a special case. The normal LMIS is really an EHD-driven charged-particle generator that can operate in two distinctive ways. A charged droplet may be generated by some EHD instability (several modes exist). But atomic ions can also be generated by a *local* electron-transfer process that occurs where the field is very high, at the jet apex. The term LMIS does not satisfactorily recognize this dual role.

For long jets, differences in behavior between metals and liquids with low electrical conductivity probably relate to the voltage drop that occurs along the jet for the latter: this may be hundreds of volts.<sup>34</sup> With a metallic jet, in which the voltage drop is typically a fraction of a volt,<sup>35</sup> it is easier (electrostatically) to get the high local apex field needed for fast atomic-ion formation; this in turn allows a quasi-steady situation whereby the jet is continuously consumed by (and ejects momentum via) the ion emission at its apex. With liquids of limited electrical conductivity, this does not happen: the jet continues to grow until it becomes hydrodynamically unstable, and the mass and momentum are ejected via droplets.

A practical difference is that nonmetallic liquids are almost always supplied via a capillary tube, whereas the normal LMIS uses liquid flow along a needle surface. A needle of radius 2–10 μm encourages the formation of a single cone-jet at its apex, although double cone-jet formation has sometimes been observed.<sup>36</sup> A capillary tube readily allows multiple cone-jet formations around the rim; this occurs in some conditions, particularly for nonmetallic liquids.<sup>32</sup>

### 2.2.3 TIME-DEPENDENT BEHAVIOR

The time-dependent behavior of an LMIS is important. Strictly, what is measured is an average emission current. For moderate-to-high time-average emission currents (or, for some materials, possibly at all currents), an LMIS cone-jet does not emit atomic ions steadily. Periods of ion emission, typically of order 1 μs or less, appear to be terminated by EHD-driven droplet emission events. The jet then returns rapidly to its original configuration and restarts atomic-ion emission. If the on-periods are sufficiently long, this behavior is described as quasi-steady and the jet is described as *self-healing*.

#### 2.2.3.1 The Most-Steady Regime

Every LMIS has a material-specific extinction current  $i_x$ , ~0.45 μA for gallium,<sup>37,38</sup> below which it either turns off or pulsates. Between this and a slightly higher current, a low-drag source has an intermediate regime (up to ~2 μA for gallium). Above this current, some materials (notably gallium) have a steady or quasi-steady regime that is relatively noise-free. In this *most-steady regime*, the *full-width at half-maximum* (FWHM) ( $\Delta E$ ) of the ion energy distribution increases according to the power law  $\Delta E \sim i^\beta$ , where  $\beta \approx 0.7$ .<sup>39,40</sup> Occasional droplet emission from the jet may occur at the low end of the regime; droplet emission becomes more probable as current and jet length increase.

#### 2.2.3.2 The Pulsation Regime

For an LMIS that is not low drag, there exists below the most-steady regime a regime where a phenomenon called *pulsation* occurs, associated with the collapse and reformation of the jet and maybe the cone. In this regime  $\Delta E$  becomes independent of current<sup>37,41–43</sup> as the applied voltage is lowered. The normal (low-drag) LMIS does not exhibit this regime. Section 2.8.1 discusses this further.

### 2.2.3.3 The Upper Unsteady Regime

For materials with a most-steady regime, there is a critical current<sup>44</sup>  $i_u$  ( $\sim 25 \mu\text{A}$  for gallium) at which  $\beta$  changes from  $\sim 0.7$  to a value in the range of 0.3–0.4. This is accompanied by rapidly increasing current noise, usually interpreted<sup>44</sup> as due to droplet emission. These features mark a change to an *upper unsteady regime*, discussed further in Section 2.8. Some materials (e.g., gold and tin) appear to have no significant most-steady regime, but go directly, or almost directly, from a pulsation regime to an upper unsteady regime.

In summary, emission of charged droplets is intrinsic to LMIS behavior. For standard FIB applications, part of the art of LMIS design is to facilitate LMIS operation in a regime where droplet emission is minimized and other requirements can be satisfied. In practice, because unwanted space-charge effects get worse at higher currents, this usually means operating the LMIS at as low an emission current as is compatible with the avoidance of pulsation or extinction.

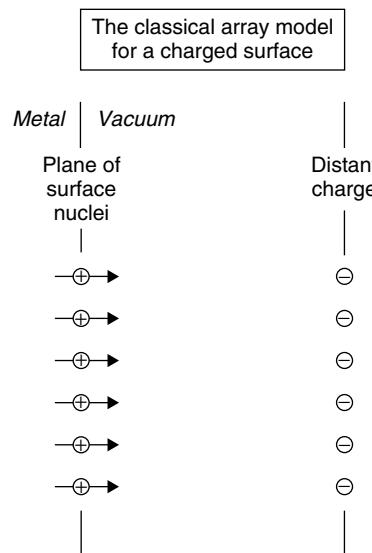
## 2.3 BASIC THEORY—ION EMISSION RELATED

### 2.3.1 CHARGED-SURFACE MODELS AND MAXWELL STRESS

In Maxwell's electrostatics,<sup>45</sup> metal surface charge is treated as a thin uniform layer. But, in a simple atomic-level approach,<sup>46</sup> each charged surface atom is modeled by a charge and a dipole, both located at the position of the atomic nucleus. Figure 2.4 represents a positively charged flat surface in this way; an opposing (negatively charged) capacitor plate is also shown.

Let the negative charge array have average surface charge density  $s$ . This creates a field, in both directions, of magnitude  $F^s = s/2\epsilon_0$ , where  $\epsilon_0$  is the electric constant. This field, acting on the charge density  $s$  on the positive plate, causes an outward stress (force per unit area)  $M = s^2/2\epsilon_0 = (1/2)\epsilon_0 F^2$ , where  $F (= 2F^s)$  is the total field in space between the charge layers. This is the well-known *Maxwell stress* ( $M$ ).<sup>45</sup>

In an atomic model, a second effect occurs. The field, due to the distant negative charges, pulls the surface-atom nuclei of the positive plate outward until each nucleus is sufficiently displaced



**FIGURE 2.4** The classical array model for a positively charged, atomically flat surface. Each surface atom is represented by a superimposed dipole and positive charge at the position of the atomic nucleus. The distant negative charge is needed for electrostatic self-consistency.

from the *electrical center* of its surrounding electrons that a balancing force is created. Thus, the surface atoms become polarized. The resulting dipole layer is a universal property of charged metal objects, irrespective of their shape.

There are two significant consequences. First, the metal's *electrical surface*<sup>47,48</sup> (where the field appears to start) is repelled outward from the plane of the surface-atom nuclei to a position near the outer edge of these atoms by the so-called<sup>48</sup> *repulsion distance d*. Second, the induced dipole layer creates a correction to the formula for the pressure difference across a charged conducting liquid surface<sup>49</sup> (see Section 2.4.1).

### 2.3.2 FIELD EVAPORATION

#### 2.3.2.1 Introduction

The main LMIS ion emission mechanism is field evaporation (FEV).<sup>50</sup> This surface process occurs in high electric fields of order 10 V/nm or more, the precise field value depending on the metal involved. It generates ions directly at the LMIS apex by simultaneously breaking the bonds holding an atom to the surface and ionizing the atom.

Field ionization (FI) is a different mechanism by which neutral atoms approaching the emitter from the surrounding space are ionized by field-induced electron tunneling. This can occur with an LMIS if a source of neutral ionizant atoms exists. However, at the relatively high fields above an LMIS apex, this occurs well away from the surface<sup>51</sup> and is known as *free-space FI* (FSFI).

FEV can sometimes be a multistage process in which the atom first *escapes* as a lower charge state ion and is then immediately *post-field ionized* by surface FI to a higher charge state.

At one stage, uncertainty existed about the main LMIS emission mechanism. Energy measurements showed<sup>52</sup> ions were formed in a surface process. For pure metals, the measured<sup>52</sup> relative abundances of the singly and doubly charged ions derived by post-field ionization (PFI), coupled with PFI theory,<sup>53</sup> enabled empirical estimates of apex field strength.<sup>52</sup> FEV field strengths can be predicted with adequate accuracy and agreed with the empirical strengths. This identified the main emission mechanism as FEV.<sup>52</sup>

#### 2.3.2.2 Relevance of FEV Theory

FEV theory is relevant to LMIS design and operation because it facilitates prediction of

- The field strength at which FEV occurs
- The charge state(s) of emitted ions
- The dependence of FEV activation energy on field

The FEV activation energy determines ionization times and rate-constants. FEV of elements is adequately understood, but alloy FEV theory needs more development.

Our theoretical knowledge comes mainly from research into FEV from solids, from 1960 to 1985, when there was prolonged dispute and debate about detailed mechanisms.<sup>46,50</sup> LMIS literature often reflects the state of knowledge early in this period, rather than the better knowledge achieved later; thus, FEV theory is discussed in detail here.

#### 2.3.2.3 FEV Rate Constant and Time Constant

The normal FEV escape process is thermally activated and governed by an Arrhenius-type equation. The *FEV rate-constant*  $k_F$  determines how fast escape occurs for an individual atom situated in a field  $F$  and is given by

$$k_F = A_F \exp[-Q(F)/k_B T] \quad (2.4)$$

where  $Q(F)$  is the (real) *FEV activation energy*,  $k_B T$  the usual Boltzmann factor, and  $A_F$  the *FEV rate-constant pre-exponential*. In principle,  $A_F$  can depend weakly on field and the local environment of the atom in question; values for  $A_F$  are not well known but are often taken as  $\sim 10^{13} \text{ s}^{-1}$ , or somewhat less. The reciprocal of  $k_F$ , denoted by  $\tau_k$ , is the *FEV time-constant*, that is, the mean time taken for a particular atom exposed to a field  $F$  to field evaporate.  $k_F$  is measured in  $\text{second}^{-1}$  and  $\tau_k$  in seconds.

### 2.3.2.4 Field Evaporation in the Equilibrium and Supply Limits

The FEV of solids occurs in what is called the *thermodynamic equilibrium limit*. The solid usually has a sharply curved surface where kink-site atoms protrude more than others and are at higher risk of FEV. It is assumed that a constant number of atoms per unit area  $\sigma_F$  are at high risk of FEV. The emitted *FEV current density*  $J_F$  is given by

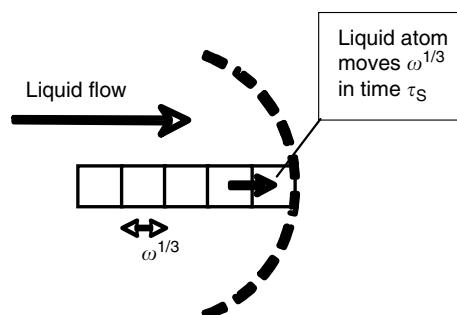
$$J_F = ze\sigma_F k_F = J_{FP} \exp[-Q(F)/k_B T] \quad \text{with } J_{FP} = ze\sigma_F A_F \quad (2.5)$$

The field that goes into Equations 2.4 and 2.5 is the field acting on the high-risk atoms, and the role of Equation 2.5 is to determine the value of  $J_F$ , given the value of  $F$ . It is convenient to measure  $J_F$  in the units microampere per square nanometer.

In the LMIS context, however, FEV takes place *in the supply limit*. In the steady state, neutral liquid atoms approaching the liquid to vacuum boundary at the LMIS apex are ionized with 100% efficiency. So, in a steady-state situation, the apex FEV flux density is given by the liquid-atom-supply flux density. The role of an FEV equation is then to derive an FEV time-constant  $\tau_k$  from the current density and obtain the apex field.

Suppose that, near the emitting apex, the LMIS can be modeled as a set of linear files of atoms, parallel to the emitter axis, all moving parallel to the axis with speed  $v_a$ . Consider the behavior of one such file, illustrated as a file of cubes in Figure 2.5. Each cube has volume  $\omega$  and side-length  $\omega^{1/3}$  ( $= 0.267 \text{ nm}$  for gallium). The time  $\tau_S$  taken for the file to move forward exactly one place is  $\tau_S = \omega^{1/3}/v_a$ . If each of the top atoms in each parallel file were subject to the same field  $F$  for all the time that they move forward, then all parallel files would be subject to simultaneous FEV with time constant  $\tau_k$ , and the steady-state condition would be  $\tau_k = \tau_S$ .

The top atoms on the most protruding files may shield neighbors electrostatically. Assume  $N$  neighbors are shielded: this changes the condition to  $\tau_k = \tau_S/N$ . A similar effect occurs in solid FEV, where it is assumed that between 1 and 10% of atoms are at high risk of FEV at any one time. For a small liquid apex, one probably expects the percentage of atoms at high risk ( $=100\%/N$ ) to be greater, so a first guess might be  $3 < N < 10$ .



**FIGURE 2.5** Schematic diagram to illustrate the process of field evaporation in the supply limit. The square boxes represent a file of atoms approaching the emitting liquid apex. In the steady state, each atom moves forward a distance  $\omega^{1/3}$  in the (average) time  $\tau_S$  taken for the topmost atom to field evaporate.

The previous situation assumes the field acting on the topmost atom does not depend on its position as the atom moves forward by a distance  $\omega^{1/3}$ . If this is not so, a further correction factor  $\alpha_r (>1)$  is needed, and the condition becomes  $\tau_k = \tau_s/N\alpha_r$ . Hence, we obtain

$$J_{Fa} = zev_a/\omega = zew^{-2/3}/\tau_s = zew^{-2/3}(N\alpha_r\tau_k)^{-1} = zew^{-2/3}(A_F/N\alpha_r)\exp[-Q_n(F)/k_B T] \quad (2.6)$$

Comparison with Equation 2.5 shows that in the supply limit we need to put the FEV current-density prefactor

$$J_{FP} = zew^{-2/3}(A_F/N\alpha_r) \quad (2.7)$$

For gallium, the surface number density of atoms ( $\omega^{-2/3}$ ) is  $\sim 14 \text{ nm}^{-2}$ . So  $3 < N < 10$  corresponds to the range  $1.4 \text{ nm}^{-2} < \sigma_F < 5 \text{ nm}^{-2}$ , which seems physically plausible.

Obviously, this theory makes the assumption that in LMIS emission the FEV is thermally activated. If physical circumstances at the LMIS apex are such that the field at the LMIS apex is very close to the field  $F_E$  at which the classical activation energy  $Q$  becomes zero, then other (faster) mechanisms involving ion and electron tunneling may come into operation. Theory relevant to this situation has never been properly explored, and the best that one can currently do is to approximate  $Q$  as equal to zero and the apex field as equal to  $F_E$ .

### 2.3.2.5 Prediction of Evaporation-Field Values

Several evaporation-field definitions exist. The most useful theoretical parameter is the field  $F_E$  (the so-called *zero-Q evaporation field* (ZQEF)), which was just discussed. The theoretically predicted activation energies  $Q_n(F)$  for ion escape into different *escape charge states*  $n$  are all different. To predict the ZQEF, one calculates for each  $n$ -value the *escape field*  $F_n^e$  at which  $Q_n(F_n^e) = 0$ ; the lowest escape field is taken as the estimate of the actual ZQEF. This is *Brandon's criterion*.<sup>54</sup> The corresponding  $n$ -value defines the *dominant escape charge state*  $n_d$ .

Escape fields are now estimated by methods<sup>50</sup> based directly on conservation of energy. In the simplest useful model,  $F_n^e$  is taken as the *Müller escape field*  $F_n^M$ , given by

$$F_n^M = b_n\{K_n^0\}^2 \quad (2.8a)$$

with

$$b_n = 4\pi\varepsilon_0/n^3e^3 \cong (0.6944616/n^3) \text{ V nm}^{-1} \text{ eV}^{-2} \quad (2.8b)$$

$$K_n^0 = \Lambda^0 + H_n - n\phi \quad (2.8c)$$

The *thermodynamic term*  $K_n^0$  is given in terms of the zero-field bonding energy  $\Lambda^0$  for the atom in question, the sum  $H_n$  of its first  $n$  (free-space) ionization energies, and the local work-function  $\phi$  of the surface left after the removal of the atom. For elemental FEV, tabulations of calculated escape fields  $F_n^M$  are given in Appendix A.1. For each material, the lowest escape field given by an equation of form 2.8 can be called its *Müller ZQEF*.

For many years, Equation 2.8 was derived from a simple (but physically incomplete) mathematical model for a particular supposed kinetic mechanism of FEV, the *image-hump* (IH) mechanism. As a result, Equation 2.8 has been known as the IH formula, and it was believed that its success in predicting experimental onset evaporation fields was evidence for this mechanism. It is now known that Equation 2.8 can be derived directly from thermodynamic arguments and cannot validly be used as evidence for any specific FEV kinetic mechanism. It seems better to call Equation 2.8 *Müller's formula*.

In reality, other effects, including repulsive energy terms, charge-sharing terms, and field dependence of bonding energy must be taken into account. Final predicted fields are higher than the Müller ZQEF by a factor typically in the order of 1.2 (but 1.13 for Ga, making its predicted ZQEF 17 V/nm). In a few cases where experimental ZQEF values can be obtained, the revised predictions are still somewhat less than the experimental values (typically by a factor of 1.4), but there is uncertainty over field calibration. At present, it seems best to use the Müller ZQEF as a working estimate for  $F_E$ .

For metallic elements, all parameters in Equation 2.4 are those for the element concerned. For alloy FEV, the ionization energies will be those for the species concerned, but the values of bonding energy and the local work-function may depend on the precise chemical environment of the bound atom in question. Currently, the best discussion of alloy FEV is in Ref. 55. When the ZQEF for a surface solute atom is significantly different from that for the underlying matrix, various problems can arise, including temporary preferential retention of the solute atom if its ZQEF is much higher than the matrix ZQEF.

### 2.3.2.6 Escape Mechanism

Historically, FEV is the name not of a single well-defined mechanism but of a class of related possibilities. It took many years to gain adequate understanding of the mechanism of FEV from solid metals. As already noted, FEV is often a two-stage process: escape, followed by PFI. The FEV field strength and activation-energy variation are determined by the escape stage alone, but the final ionic charge state is determined by both stages.

For LMIS theory, an important question is whether the escape mechanism is the IH mechanism (as widely stated in LMIS literature) or a charge-exchange (CE)-type mechanism. In the IH mechanism an integrally charged ion is first formed and then escapes over a Schottky IH; in the CE-type mechanisms bond-breaking and ionization occur together. CE-type mechanisms fall into two main categories: *charge hopping* and *charge draining*.

There is now broad consensus in the old mainstream field ion emission community that, certainly for most materials, escape takes place by a charge-draining mechanism, basically as illustrated in Figure 2.6. At the already highly charged surface, an atom at high risk of FEV is already in a partially ionized state. As its nucleus moves away from the surface, electron charge drains out of the ion until it is in a pure integral state. During this process the ion passes over an activation-energy hump. Unresolved questions remain about the precise path in space taken by the departing nucleus and about precisely how the charge associated with the atom varies along this path.

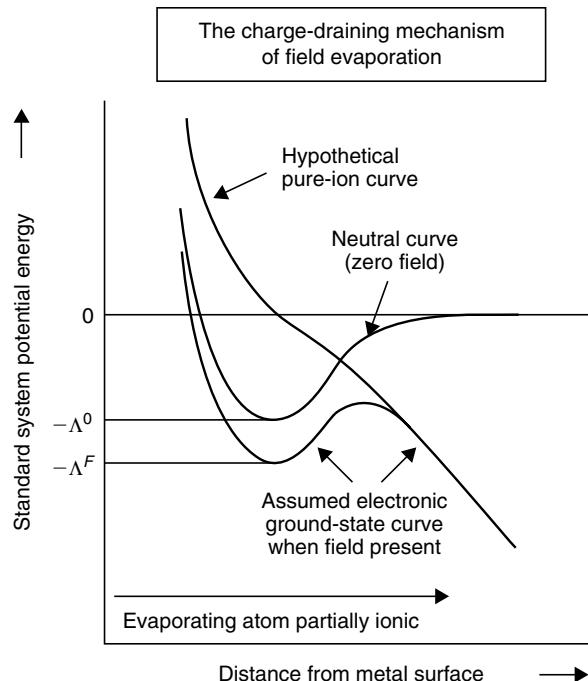
For most elements of interest to LMIS design, it can be shown convincingly that the IH mechanism is not a physical possibility because, at the field strengths necessary for FEV, there is no IH in the ionic potential.<sup>56</sup> Also, in all cases tested, the experimental field sensitivity of FEV flux or rate constant is different from that predicted by the simple IH model.<sup>57</sup> Unfortunately, gallium is one of the few materials for which the numerics are not convincing, and there have been no experimental tests. Nevertheless, there seems no good reason to doubt that all metallic elements escape via the charge-draining mechanism.

### 2.3.2.7 Field Dependence of Activation Energy

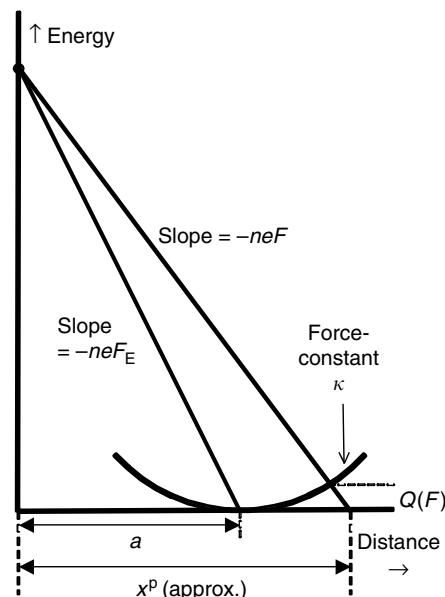
To link emission current density to the LMIS apex field, an expression for  $Q(F)$  is needed. For charge draining no convincing *general* calculation has been carried out (although a good quantum-mechanical theory exists for aluminum<sup>58</sup>). For the charge-hopping mechanism a simple model exists,<sup>59</sup> based on taking the atomic nucleus and moving it in a parabolic bonding well of force-constant  $\kappa$ , as illustrated in Figure 2.7. This model, replaced an older treatment, that had been found to contain a mathematical error.

Figure 2.7 shows that, in a first approximation, the newer model yields

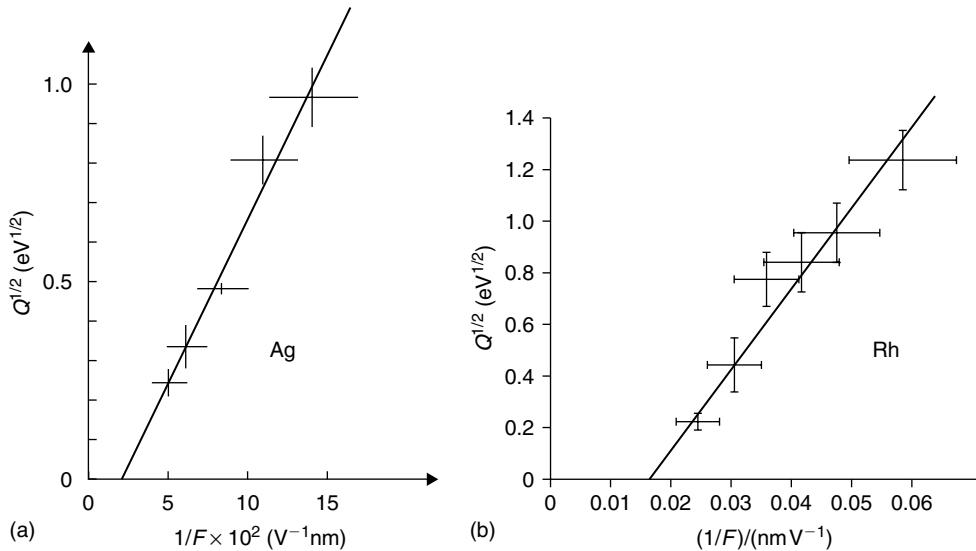
$$Q(F) \approx \frac{1}{2} \kappa \{x^p - a\}^2 = \frac{1}{2} \kappa a^2 \{(x^p/a) - 1\}^2 = \frac{1}{2} \kappa a^2 \{(F_E/F) - 1\}^2 \quad (2.9a)$$



**FIGURE 2.6** Schematic diagram to illustrate the charge-draining mechanism of field evaporation, as currently understood.  $\Lambda^0$  and  $\Lambda^F$  denote atomic bonding energies with and without the external field  $F$  present. If a field is present, then the atom is partially charged when initially bound on the surface, as required by Gauss' theorem. As the atomic nucleus leaves the surface, the charge drains out (to leave an integrally charged ion), and the ion core escapes over a small activation energy hump (or tunnels through it).



**FIGURE 2.7** Schematic diagram to illustrate the derivation of the Forbes 1982 formula for field evaporation activation energy in the context of a charge-hopping model.



**FIGURE 2.8** To illustrate the experimental linear relationship between  $Q^{1/2}$  and  $1/F$ , which is predicted for field evaporation by the Forbes 1982 formula: (a) for rhodium<sup>60</sup>; and (b) for silver (data provided by the late Professor J. Block.)

where  $a$  and  $x^p$  are the distances shown in Figure 2.7. A better approximation<sup>59</sup> is written here as

$$Q(F) \approx C_Q (F_E/F - 1)^2 \equiv C_Q (\mu^{-1/2} - 1)^2 \quad (2.9b)$$

$$J_F = J_{FP} \exp[-(C_Q/k_B T)(\mu^{-1/2} - 1)^2] \quad (2.9c)$$

where  $C_Q$  is a semiconstant, and  $\mu \equiv F^2/F_E^2 = M/M_E$ , where  $M_E = (1/2)\epsilon_0 F_E^2$ ;  $\mu$  is the *normalized apex Maxwell stress*, and (it is assumed) must have a value in the range  $0 < \mu < 1$ . The  $\mu$ -based form, with  $J_{FP}$  given by Equation 2.7, is useful in LMIS theory.  $C_Q$  is difficult to predict reliably, but form 2.9b fits experimental results, in the few cases where the field dependence of  $Q$  has been directly measured. (The plots of  $Q^{1/2}$  versus  $1/F$  shown in Figure 2.8 are linear down to about  $F/F_E = 0.3$  and yield the empirical  $C_Q$  values: 0.040 eV for silver and 0.25 eV for rhodium.<sup>60</sup>) Forbes and Chibane<sup>61</sup> argue that form 2.9 should also model the charge-draining mechanism. The model is not valid at low fields, where a Morse potential (or equivalent) is needed to describe bonding-well shape.

After escape, if the field is high enough, a departing ion may be post-field ionized (in one or more steps), from its escape charge state  $n$  into a charge state  $\xi$ . In some field ranges, mixtures of different charge states may be generated, and the *final charge state*  $z$  should be set equal to the mean value. Kingham<sup>53</sup> has published diagrams showing the field ranges over which particular ionic species are predicted; these ranges cover most metals of interest to LMIS design. Field values derived from these diagrams are given in Table A.1.

### 2.3.3 FIELD-EMITTED VACUUM SPACE CHARGE

Results from the theory of field-emitted vacuum space charge (FEVSC) are relevant. From equations in Ref. 62 one can derive the *dimensionless planar FEVSC equation*

$$9\zeta^2\theta^2 - 3\theta + 3 - 4\zeta = 0 \quad (2.10)$$

where  $\zeta$  describes the *strength* of the space-charge effects and  $\theta$  the resulting reduction in surface field.  $\theta$  is defined by  $\theta = F_p/F_L$ , where  $F_p$  is the *Poisson surface field*, that is, the emitter surface

field that exists in the presence of space-charge effects, and  $F_L$  is the *Laplace surface field*, that is, the field that would exist in their absence, for the same applied voltage. The parameter  $\zeta$  is given by

$$\zeta = \kappa_{sc} J F_p^{-2} V^{1/2} \quad \text{with } \kappa_{sc} = \epsilon_0^{-1} (m/2ze)^{1/2} \quad (2.11)$$

Equation 2.10 has the solution

$$\theta(\zeta) = (1/6\zeta^2)[1 \pm \sqrt{S_q}] \quad \text{with } S_q \equiv 1 + 4\zeta^2(4\zeta - 3) \quad (2.12)$$

Figure 2.9a shows  $\theta(\zeta)$ . For low and high  $\zeta$ , there are approximate solutions:

$$\theta \approx 1 - \frac{4}{3}\zeta \quad (\zeta \text{ low}) \quad \theta \approx \frac{2}{3}\zeta^{-1/2} \quad (\zeta \text{ high}) \quad (2.13)$$

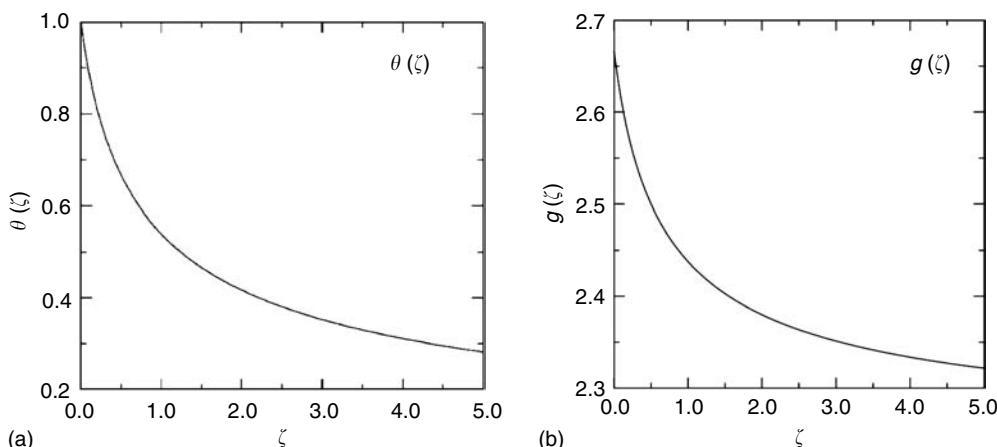
For gallium,  $\kappa_{sc} = 6.7888 \times 10^7 \text{ A}^{-1} \text{ V}^{3/2}$ . Assuming apex radius  $r = 1.5 \text{ nm}$  gives a rough estimate of emission area  $A = \pi r^2 = 7.1 \text{ nm}^2$ . Assuming  $i = 2 \mu\text{A}$  yields a current density  $J = i/A = 2.8 \times 10^{11} \text{ A/m}^2$ . Taking  $F_p = 15 \text{ V/nm}$  and  $V = 3000 \text{ V}$ , then yields  $\zeta = 4.7$ . Equation 2.12 then gives  $S_q = 1370$  and  $\theta = 0.29$ . This result suggests that, even at this low current level, space-charge has caused the Poisson field ( $F_p$ ) at the apex to be much less than the Laplace field ( $F_L$ ) that would have been present for the same geometry in the absence of space-charge. This calculation is too simple, not least because it applies planar space-charge theory to the very sharp LMIS emitter, but the result is qualitatively correct. Space-charge affects LMIS behavior, even at the low emission currents used in FIB machines.

The space-charge also reduces the Maxwell stress ( $M_p$ ) on the emitter surface below what it would have been: ( $M_L$ ) for the same geometry in the absence of space-charge, by an amount  $\Delta_M$  given by

$$\Delta_M = M_L - M_p = M_p[(M_L/M_p) - 1] = M_p(\theta^{-2} - 1) \quad (2.14)$$

If we define  $g(\zeta) \equiv (\theta^{-2} - 1)/\zeta$  (giving  $\Delta_M = g\zeta M_p$ ), then, in the limits of low and high  $\zeta$ , Equation 2.13 yields

$$g(\zeta) \approx \frac{8}{3} \cong 2.67 \quad (\zeta \text{ low}) \quad g(\zeta) \approx \frac{9}{4} = 2.25 \quad (\zeta \text{ high}) \quad (2.15)$$



**FIGURE 2.9** Two results from the theory of planar field-emitted vacuum space-charge, showing: (a) how the ratio  $\theta$  of the Poisson field  $F_p$  to the Laplace field  $F_L$  depends on the space-charge strength  $\zeta$ ; and (b) that the parameter  $g$  varies slowly with  $\zeta$ .

Figure 2.9b shows the variation of  $g(\zeta)$  over the range  $0 < \zeta < 5$ . This slow variation (assumed to apply to sharply curved emitters, too) plays a key role in generating the simple form of Mair's equation (see Section 2.5.3.).

## 2.4 BASIC THEORY—ELECTROHYDRODYNAMICS

The main EHD issues relating to LMIS behavior are (1) what liquid shapes represent situations of EHS or EHD equilibrium, (2) under what conditions are such shapes stable or steady against small displacements, (3) how do we calculate the liquid flow through a steady shape, and (4) what is the correct form for the LMIS apex boundary condition. A preliminary issue is the pressure jump across a charged, curved, and conducting liquid surface. For simplicity, discussion here is presented in a form that applies only to good conductors into which external fields do not penetrate.

### 2.4.1 THE FORMULA FOR SURFACE PRESSURE JUMP

The *surface pressure jump* is the difference  $\Delta p$  between the pressure  $p_{\text{in}}$  inside a liquid surface and the pressure  $p_{\text{out}}$  outside. According to a classical ansatz (see Ref. 45, p.190),  $\Delta p$  is given for a charged, conducting, and curved liquid surface by

$$\Delta p = p_{\text{in}} - p_{\text{out}} = \gamma c - \frac{1}{2} \epsilon_0 F^2 \quad (2.16)$$

where  $\gamma$  is the surface free energy per unit area in the absence of electric fields, and  $c$  the total surface curvature, given by  $c = 1/r_1 + 1/r_2$ , where  $r_1$  and  $r_2$  are the principal radii of curvature. When  $F$  is sufficiently large,  $\Delta p$  becomes negative.

The well-known simple surface-tension formula, without the Maxwell stress term, was derived empirically by Laplace<sup>63,64</sup> in 1806–1807 but proved by Gibbs<sup>65</sup> using the laws of thermodynamics around 1875. It became clear in the 1980s that Equation 2.16 had never been formally derived from thermodynamics. In 1995, Ljepojevic and Forbes (LF) presented a formal proof.<sup>66</sup> The best approach uses a new variant of thermodynamic potential, best considered as an electrical Gibbs function.

Let there be a potential difference  $V$  between the liquid body and its surroundings, a charge  $q$  on the body, and a complementary charge  $-q$  on the surroundings, and consider an infinitesimally small change in liquid shape. Under conditions where both temperature and  $V$  are constant, the condition for thermodynamic and EHS equilibrium is

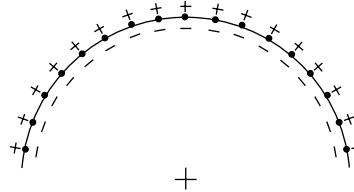
$$\Delta^1 G = \Delta^1 \mathcal{F} - V \Delta^1 q = 0 \quad (2.17)$$

where  $\mathcal{F}$  and  $G$  are the *Helmholtz free energy* and *electrical Gibbs function* for the system (excluding the *external* constant-potential electrical source). Infinitesimally small changes  $\Delta^1 G$ ,  $\Delta^1 \mathcal{F}$ , and  $\Delta^1 q$  are caused by the small change in liquid shape.

For a classical conductor, assuming that elements  $ds$  of the original liquid surface  $\Sigma$  move forward by a set of distances  $\delta x$ , LF proved that condition 2.17 becomes

$$\Delta^1 G = \int_{\Sigma} \{-\Delta p + \gamma c - \frac{1}{2} \epsilon_0 F^2\} \delta x \, ds = 0 \quad (2.18)$$

This result is also obtained if the external electrical source is removed, and the changes are considered to take place at constant temperature and charge  $q$ . At equilibrium, Equation 2.18 must be true for all possible infinitesimally small smooth changes in liquid-surface position, so the integrand must be identically zero (except, possibly, at a finite number of singular points). Equation 2.16 follows.



**FIGURE 2.10** Schematic diagram to illustrate how a layer of polarizable dipoles at a charged liquid metal surface gives rise to a pressure increase in the liquid (see text).

This classical-conductor result ignores the existence of atoms. But important parts of an LMIS have nanoscale dimensions, so—unlike the situation in many hydrodynamic applications—the existence of atoms must be taken into account. The energies associated with the surface dipoles in the model described in Section 2.3.1 contribute additional terms to  $\mathcal{F}$  and hence to  $G$ . Analysis of the general case is difficult, but for a sphere of large radius  $r$  the outcome is

$$\Delta p = c(\gamma - \frac{1}{4} \varepsilon_0 F^2 d) - \frac{1}{2} \varepsilon_0 F^2 (1 - cd) = 2\gamma/r - \frac{1}{2} \varepsilon_0 F^2 + \frac{1}{2} \varepsilon_0 F^2 d/r \quad (2.19)$$

The repulsion distance  $d$ , defined earlier, has a value comparable with atomic radii.<sup>48</sup>

This formula can be interpreted as a field-dependent reduction in surface tension, combined with a curvature-dependent reduction in the magnitude of the Maxwell stress, the overall effect<sup>49</sup> being to add a small positive term to  $\Delta p$ . This formula has been proved for conditions when the sphere radius is “not too small,” so the form of the additional term may not be correct for small  $r$ -values.

Figure 2.10 shows a simple spherical model for understanding the physical origin of these effects. The positive charge that actually resides on the sphere surface can be modeled as a single large positive charge at the sphere center. The dipoles are in an inhomogeneous field due to this charge and are pulled weakly toward it; this causes a small increase in pressure inside the liquid. An effect also occurs, due to mutual repulsion between the dipoles, that is equivalent to a reduction in surface tension; this causes a small reduction in pressure. The overall result is a small increase in pressure inside the liquid.

In absolute terms, the correction term in Equation 2.19 is relatively small and can usually be disregarded. But its existence is physically important, because it destroys the validity of Taylor’s mathematical proof that a Taylor cone is a situation of EHS equilibrium.

#### 2.4.2 ELECTROHYDROSTATIC EQUILIBRIUM AND STABILITY CONDITIONS

For a liquid body to be in EHS equilibrium the internal pressure must be uniform, and the body must have no excess kinetic energy other than thermal energy. If the external pressure  $p_{\text{out}}$  is uniform, then the pressure jump  $\Delta p$  must be the same at all nonsingular points on its surface.

In respect of stability conditions, the situation in the literature is confused (and has been for well over 100 years). Consider a small, finite, isolated body of liquid. In these circumstances  $G = \mathcal{F}$ . Thermodynamic arguments can be applied. Minimization of  $\mathcal{F}$  requires minimization of surface free energy: the equilibrium shape is a sphere (of radius  $r$ ). Now consider the body to be in its ground thermodynamic state after losing all excess (nonthermal) kinetic energy, and consider a charge  $q$  to be added to the body. This adds to  $\mathcal{F}$  an energy  $q^2/2C$ , where  $C$  is the body’s electrical capacitance. The condition for stability is that the liquid shape corresponds to a sufficiently deep local minimum in  $G$  (and  $\mathcal{F}$ ) with respect to all possible distortions of the charged liquid body. This requires that the first derivatives of  $G$  be zero with respect to all possible distortions (which leads to the EHS equilibrium requirements already stated) and that the second derivatives be positive with respect to all possible distortions.

In 1882, Raleigh<sup>67</sup> made a well-known analysis of this type, assuming a particular form of distortion, and reached conclusions about the maximum charge that the droplet could hold before

becoming unstable. In terms of the equivalent surface field  $F$  on the original spherical droplet, his result can be written as

$$\left(\frac{1}{2}\epsilon_0 F^2\right)/(2\gamma/r) < 1 \quad (2.20)$$

The droplet is either spherical and stable, or unstable. Theoretically, the charged droplet cannot maintain its stability by adjusting its shape. Raleigh's original derivation was very difficult to follow, even for Raleigh himself,<sup>68</sup> and alternative derivations have been given.<sup>68-70</sup>

Before this, Maxwell's 1873 textbook had discussed the closely related case of an electrified soap bubble,<sup>45</sup> but note his  $T_0 = 2\gamma$ . Though worded carefully, and strictly correct, his remarks are easily misinterpreted to imply that the condition for stability is pressure based, namely,

$$\Delta p > 0 \quad (2.21)$$

With Equation 2.16, this condition also leads to Equation 2.20.

Much subsequent confusion has arisen from these two approaches to the derivation of condition 2.20. Condition 2.21 has been used incorrectly for around 100 years and continues to be used incorrectly as a general result applicable to other geometries, both in LMIS theory and in EHD theory more generally.

As Taylor has pointed out<sup>11</sup> (though not in these words), the pressure-based condition 2.21 has not been deduced as a general mathematical conclusion from the thermodynamic requirement that, for stability, the second derivatives of  $G$  need to be positive with respect to all possible distortions. The charged spherical droplet is a very special case, and condition 2.21 has been demonstrated only for this special case.

Taylor<sup>11</sup> shows this convincingly by analyzing the case of a neutral droplet in a uniform field  $F_M$ . He finds that as  $F_M$  increases, the originally spherical drop can initially maintain its stability by elongating (parallel to the field) into a shape closely resembling an ellipsoid of revolution (a prolate spheroid) but that this shape becomes unstable when the ratio  $\lambda$  of the semimajor axis to the semiminor axis approaches the value 1.9. This is in good agreement with an experimental value of 2.2. In terms of the local field  $F$  and the local radius of curvature  $r$  at the apex of prolate spheroid with  $\lambda = 1.9$ , we calculate that Taylor's criterion corresponds to the stability condition:  $(1/2)\epsilon_0 F^2/(2\gamma/r) < 0.19$ .

A conclusion is that it is unsafe to use Equation 2.21 as an LMIS onset condition, as sometimes done, although it may work adequately in some situations (including the metallic GG cone-jet) where, for other reasons, the onset condition requires the liquid hydrostatic pressure to be relatively close to zero.

### 2.4.3 TAYLOR'S MATHEMATICAL CONE

EHS equilibrium shapes are important because they are configurations into which a body of liquid in motion may tend to settle as it loses its excess energy. In analogy with chaos theory, we call such configurations *attractors*. In 1964, Taylor<sup>11</sup> argued that existing analyses based on liquid spheres and prolate spheroids were not helpful in understanding the disintegration mechanism of spherical droplets, which had long been known to deform in practice by developing pointed conical ends and throwing out fine jets. As a basis for further thought, he found a classical EHS equilibrium shape, based on a conical solution of Laplace's equation, satisfying the requirement that condition 2.16 with  $\Delta p = 0$  holds at all points on its surface. This is the well-known *Taylor cone*, with half-angle approximately equal to 49.3°. The solution has other requirements, not of immediate interest. This half-angle prediction has been confirmed by experiments on water by Taylor, and, later, by experiments<sup>71</sup> (at low LMIS emission currents) on a specially prepared gallium LMIS.

Taylor's argument has sometimes been misunderstood. In particular, he made no claim about cone stability. In fact, an energetically favorable ( $\Delta G < 0$ ) classical distortion mode seems to exist<sup>72</sup> whereby the Taylor cone elongates into a cusp, particularly if liquid supply can occur.

The existence of the correction term in Equation 2.19 creates a new situation. Taylor's mathematical cone was derived to satisfy the requirement that  $\Delta p$  is exactly zero at all points on its surface and used the classical pressure-jump formula 2.16. But the correction term in Equation 2.19 gets increasingly large and significant as the Taylor-cone apex is approached. The existence of this term (or of any more carefully derived correction terms) destroys the validity of Taylor's derivation. It seems that the classical-physics proposition that "the mathematical Taylor cone is an exact EHS equilibrium shape" is incompatible with the known physical existence of electrically polarizable atoms and is now a failed hypothesis.

Equation 2.19 raises the possibility that, near the cone apex, a rounded cone may be an attractor shape. But, as yet, there is no proof, mathematical or experimental, that equations can be formulated that have a solution. So it is unclear whether any rounded-cone-like EHS equilibrium configuration exists.

Since the correction term is large only near the cone apex, where the curvature  $c$  is high (within a distance of, perhaps, 100 nm from the apex), the mathematical Taylor cone is an attractor near its base. So the base of a steady metallic GG cone-jet should resemble a Taylor cone, particularly at low emission currents. This conclusion appears consistent with the experiments.<sup>71</sup>

#### 2.4.4 THE STEADY HIGH-ELECTRICAL-CONDUCTIVITY GILBERT–GRAY CONE-JET

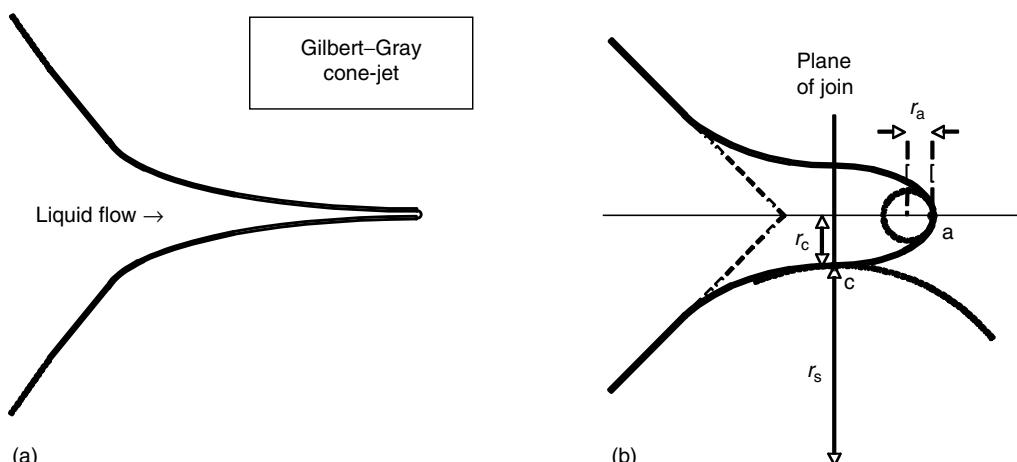
The main LMIS model now used is the *steady high-electrical-conductivity GG cone-jet*, illustrated in Figure 2.11a. The cone base is anchored by surface-tension forces around its rim (not shown), and pressure in the base-plane is assumed uniform (with value  $p_b$ ). Liquid enters the base and is assumed to flow nonturbulently through a steady time-independent shape to the jet cap, where it is ionized with 100% efficiency by FEV. The flow is driven by the pressure difference ( $p_b - p_a$ ), where  $p_a$  is the apex pressure. As with the Taylor cone, this model is to be treated as a "basis for further thought."

##### 2.4.4.1 Pressure Relationships

The pressure differences affecting cone-jet behavior *in vacuum* are

$$\Delta p = p = (p - p_b) + p_b \quad (2.22)$$

$$p_b = (p_b - p_r) + p_r = (p_b - p_r) + \Delta p_r \quad (2.23)$$



**FIGURE 2.11** The Gilbert–Gray cone-jet. To illustrate (a) the general shape of the cone-jet and (b) a revised model for its emitting apex (drawn for a very short jet).

where  $\Delta p$  and  $p$  are the surface pressure jump and the pressure just inside the surface at any point on the cone-jet surface, and  $\Delta p_r$  and  $p_r$  are the values for the main liquid reservoir. These equations inform the hydrodynamics of, respectively, the cone-jet itself and its supply from the main reservoir. If the classical form 2.16 is used for the surface pressure jump, Equation 2.22 becomes

$$\frac{1}{2}\varepsilon_0 F^2 - c\gamma - (p_b - p) + p_b = 0 \quad (2.24)$$

#### 2.4.4.2 Limiting Parameter Values

A simple treatment applies Equation 2.24 to an emitter apex that is assumed to be cylindrically symmetric. Take  $p_b$  and  $(p_b - p_a)$  as zero, and put  $c = 2/r_a$ , where  $r_a$  is the apex radius. This yields

$$2\gamma/r_a - \frac{1}{2}\varepsilon_0 F_a^2 \approx 0 \quad (2.25)$$

where  $F_a$  is the apex field. An estimate of evaporation field thus provides an estimate of apex radius.

The ZQEF  $F_E$  applies when the barrier to ion detachment has just become zero and is the maximum value that normal FEV theory allows. We can define material-specific parameters  $M_E$  and  $r_E$  by

$$M_E \equiv \frac{1}{2}\varepsilon_0 F_E^2 \equiv 2\gamma/r_E \quad (2.26)$$

In normal theory, the Maxwell stress  $M_E$  is an upper limit to the outward stress on the emitter surface, because, if there is no barrier to ion detachment, the applied field has difficulty transmitting additional stress to the underlying metal: it accelerates the ions instead. It follows that  $r_E$  is a lower bound for the apex radius during *steady-state* LMIS operation (provided  $p_b \leq 0$ ). (The apex radius can become smaller than this and may do so during emitter-shape fluctuations.)

For gallium,  $F_E \approx 15$  V/nm, the stresses  $M_E$  and  $2\gamma/r_E$  are each of magnitude  $\sim 1$  GPa ( $\sim 10000$  atm, where 1 atm = 101325 Pa), so  $r_E \approx 1.4$  nm. There is claimed experimental support<sup>14</sup> from high-resolution transmission electron micrographs for jet diameters\* of  $\sim 3$  nm, but there has also been concern as to the reliability of measurements of such small diameters.

#### 2.4.4.3 Nonturbulent Flow

If flow is assumed to be nonturbulent, Bernoulli's equation provides that along a streamline,

$$\frac{1}{2}\rho v^2 = p_b - p \quad (2.27)$$

This relation yields several useful formulas. Inserting it in Equation 2.24 yields the condition

$$W \equiv \frac{1}{2}\varepsilon_0 F^2 - \gamma c - \frac{1}{2}\rho v^2 + p_b = 0 \quad (2.28)$$

where  $W$  is a mathematical function defined by this equation. For nonturbulent flow, this condition applies at all points along a streamline that lies just inside the liquid surface. One of the old numerical shape-modeling programs<sup>73</sup> works by varying parameters to minimize the integral of  $W$  over the liquid surface.

For a streamline that lies along the axis of the cone-jet, Equations 2.1, 2.2 and 2.27 yield a formula for the equivalent supply current density  $J_{Sa}$  at the emitter apex (i.e., the  $J$ -value equivalent to the atom flux density  $Z$ ):

$$J_{Sa} = \pm C_p |p_b - p_a|^{1/2} \quad (2.29a)$$

$$C_p \equiv z e \omega^{-1} (2/\rho)^{1/2} \quad (2.29b)$$

---

\* With reference to Figure 2.11b, the jet diameter near the apex is  $2r_c$ ;  $r_c$  provides an upper limit for the apex radius  $r_a$ .

where  $C_p$  is a material-specific constant defined by Equation 2.29b. In Equation 2.29a, the positive sign is taken if  $p_b > p_a$ , as is the case with an operating LMIS. (The condition  $p_b < p_a$  corresponds to the unsteady situation of a collapsing cone-jet, when the negative sign is taken.) Values of  $C_p$  and derived constants are given in Appendix A.1.

It is useful to put Equation 2.29 into a different form by defining the dimensionless quantities

$$\mu \equiv \frac{1}{2} \varepsilon_0 F_a^2 / M_E = (F_a / F_E)^2 \quad (2.30a)$$

$$\chi \equiv (2\gamma / r_a) / M_E = r_E / r_a \quad (2.30b)$$

The parameters  $\mu$  and  $\chi$  are (the magnitudes of) the *normalized apex Maxwell stress* and the *normalized apex surface-tension stress*, respectively. When the LMIS is in vacuum, the classical pressure-jump formula 2.16 yields

$$p_a / M_E = \Delta p_a / M_E = \chi - \mu \quad (2.31)$$

So Equation 2.29 becomes

$$J_{Sa} = \pm J_{SP} |(p_b / M_E) - (\chi - \mu)|^{1/2} \cong \pm J_{SP} |\mu - \chi|^{1/2} \quad (2.32a)$$

$$J_{SP} = C_p M_E^{1/2} = z e \omega^{-1} F_E (\varepsilon_0 / \rho)^{1/2} \quad (2.32b)$$

where  $J_{SP}$  is a material-specific constant (the *supply-current-density prefactor*); for gallium,  $J_{SP} \approx 5 \mu\text{A}/\text{nm}^2$ . Since  $\mu$  cannot exceed 1,  $J_{SP}$  is an upper bound on the supply current density  $J_{Sa}$ .

#### 2.4.4.4 Viscous-Loss Terms

The Bernoulli equation assumes there is no significant loss of flow kinetic energy to internal heat as the liquid moves. Analyses of LMIS behavior usually assume this. If viscous effects do occur, then Equation 2.27 has to be replaced by

$$\frac{1}{2} \rho v^2 = (p_b - p) - p_{vis} \quad (2.33)$$

where  $p_{vis}$  is a viscous-loss term. Beckman<sup>37</sup> investigated the possible size of (the apex value of) this term when calculating LMIS extinction current. He found corrections of order 5% or less. This chapter disregards the numerics of viscous-loss terms.

#### 2.4.4.5 Quasi-Ellipsoidal Model for the Liquid Cap

For reasons made clear in Section 2.4.5.1, we now model the liquid cap as approximately an ellipsoid of revolution about the jet axis. The basic version uses a hemiellipsoid as shown in Figure 2.11b. A circular cross-sectional plane C of radius  $r_c$  separates the cap from the main jet, which is arbitrarily assumed to be cylindrical close to the plane. At the edge of the plane just inside the liquid surface and just on the main-jet side of plane C, the liquid hydrostatic pressure is denoted by  $p_c$  and speed parallel to the axis by  $v_c$ . As compared with the hemispherical cap model, this new model allows a distinction in equations between  $r_c$  and the apex radius  $r_a$ .

In reality, at the point where the cap joins the cusp, the cusp may be conical rather than cylindrical, particularly at short cusp lengths. This introduces some minor mathematical complications, similar to those discussed in Ref. 74. We can afford to ignore these here, but they may become important in future work.

#### 2.4.4.6 Pressure Drop in the Cone-Jet

A first estimate of the pressure drop ( $p_b - p_c$ ) can be obtained as follows. Assume that the liquid flow is normal to plane C everywhere and that pressure, speed, and supply current density  $J_S$  are uniform in plane C. The equivalent supply current crossing this plane is  $i_S = \pi r_c^2 J_S$ . The assumption that  $r_c \approx 1.5$  nm when  $i_S = 2 \mu\text{A}$  implies  $J_S \approx 0.28 \mu\text{A}/\text{nm}^2$ . For gallium, Equation 2.29, and the assumption that  $p_c \approx p_a$ , imply  $(p_b - p_c) \approx 3.5 \text{ MPa} \approx 35 \text{ atm}$ ; Equation 2.27 then implies  $v_c \approx 35 \text{ m/s}$ . The pressure drop is small in comparison with the sizes ( $\sim 1 \text{ GPa}$ ) of the large terms in Equation 2.25, and appears to justify the approximation used to derive Equation 2.25. It needs to be added that this simple estimation method is not robust and that actual pressure drops may be somewhat higher.

#### 2.4.4.7 The Zero-Base-Pressure Approximation

By definition, a *low-drag* LMIS is taken to have  $(p_r - p_b) \ll (p_b - p_a)$ , thereby justifying the approximation  $(p_b - p_r) \approx 0$  in Equation 2.23. Further, when the whole source assembly is in vacuum, we can take  $p_b \approx p_r$ , where  $p_r$  is determined by a curved liquid surface associated with the main reservoir. This surface is unlikely to have a radius less than  $5 \mu\text{m}$ , which puts  $|p_r|$  (and hence  $|p_b|$ ) in the range  $0\text{--}0.3 \text{ MPa}$ . This is much smaller than the estimate of  $(p_b - p_c)$  as  $\sim 3.5 \text{ MPa}$  and completes justification of the approximation that for a low-drag LMIS we can put  $p_b \approx 0$ . These estimates are rough approximations, but the picture emerges that the normal LMIS is a device driven by the negative hydrostatic pressure at its apex.

### 2.4.5 APEX BOUNDARY CONDITIONS

In the steady state, continuity at the liquid/vacuum boundary needs the FEV flux  $R_F$  (or equivalent current  $i_F$ ) equal to the supply flux  $R_S$  (or equivalent current  $i_S$ ). But theoretical analysis needs a second boundary condition. In 1989, Kovalenko and Shabalin<sup>75</sup> were the first to require that the apex FEV current density  $J_{Fa}$  be equal to the apex supply current density  $J_{Sa}$ . This seems to be a good steady-state boundary condition, provided that one treats FEV in the supply limit. Older approaches (used in numerical modeling) need re-examination, as shown now.

#### 2.4.5.1 The Kingham and Swanson Model

The older approaches<sup>73,74,76</sup> assume a low-drag LMIS ( $p_b \approx 0$ ) and a hemispherical cap. Applying Equation 2.24 at the apex with ( $p_b \approx 0$ ) yields

$$2\gamma/r_a - \frac{1}{2}\epsilon_0 F_a^2 + (p_b - p_a) = 0 \quad (2.34)$$

Multiplying by  $(2/\epsilon_0)$  and rearranging, and then doing similar operations for point c, and then inserting the apex boundary condition  $F_a = F_E$  yields

$$F_E^2 = (2/\epsilon_0)(p_b - p_a) + (2/\epsilon_0)(2\gamma/r_a) \quad (\text{at } a) \quad (2.35)$$

$$F_c^2 = (2/\epsilon_0)(p_b - p_c) + (2/\epsilon_0)(\gamma/r_c) \quad (\text{at } c) \quad (2.36)$$

Note that  $(2\gamma/r_a)$  appears in Equation 2.35 because the shape above plane C is assumed hemispherical but that  $(\gamma/r_c)$  appears in Equation 2.36 because the shape below plane C is assumed cylindrical.

Equation 2.36 appears as equation 2.15 in Kingham and Swanson's (KS) article<sup>74</sup> on LMIS shape, except that KS have evaluated  $(2/\epsilon_0)$  as  $2.259 \times 10^{11}$  SI units and have put the pressure term in a specific form. The remaining issue for KS was to choose  $F_c$ . They argued that the evaporation field must be uniform across the hemisphere and put  $F_c = F_E$ . Unfortunately, their assumption of a hemispherical cap (i.e.,  $r_c = r_a$ ) then implies an internal pressure difference of  $\gamma/r_a$  ( $\approx 0.5 \text{ GPa} \approx 5000 \text{ atm}$ ) across plane C. This is clearly unphysical.

In the limit of low supply current (i.e., low  $p_b - p_c$ ), the KS field choice yields a cap-radius estimate about four times lower than is predicted by Equation 2.25. But one expects Equation 2.25 to yield a lower boundary for apex radius, so the KS choice is physically inconsistent. This choice also causes Equation 2.16 in KS to be incorrect and invalidates part of the remarks made by KS in relation to their Equations 2.15 and 2.16.

Unfortunately, KS equation 2.15 has played a key role in LMIS shape-modeling programs (including those of the present author and colleagues), where it is used to find the apex radius for a given input emission current. The inconsistency in KS equation 2.15 causes all these programs to yield incorrect results at low emission currents, though trends are correct, and the results themselves are better at high emission currents. Also, this problem indirectly caused an anomaly in recent work of Radlicka and Lencová<sup>77</sup> on modeling of LMIS optical characteristics. On replacing KS equation 2.15 by the equivalent of Equation 2.35, this anomaly was removed.<sup>78</sup> This provides independent evidence that correction was necessary.

Alternatives to the KS condition are using: (1) Equation 2.35; (2) Equation 2.36, either with  $F_c = F_E/\sqrt{2}$  (which gets things approximately right) or with some more complex condition; or (3) a completely different form of boundary condition, such as  $J_{Fa} = J_{Sa}$ . Using Equation 2.35 is significantly better than using Equation 2.36, because Equation 2.35 relates to the conditions at the emitter apex. But Equation 2.35 does use the assumption  $F_a = F_E$ . This also may need re-examination, as shown in Section 2.4.5.2.

#### 2.4.5.2 Steady-State Boundary Condition Based on Current Densities

The condition  $J_{Fa}^2 = J_{Sa}^2$ , with Equations 2.5 through 2.7 and 2.32, yields (when  $\mu > \chi$ )

$$J_{Sa}^2 = J_{SP}^2[\mu - \chi] = J_{FP}^2 \exp[-2Q(\mu)/k_B T] = J_{Fa}^2 \quad (2.37)$$

Since  $\chi = \mu - (J_{Fa}^2/J_{SP}^2)$  and there is a physical limitation (in normal models) that  $\mu \leq 1$ , it follows that in *steady-state flow*,  $\chi$  has a maximum value,  $\chi_{max}$ , less than 1. So the apex radius has a minimum value  $r_{min} (= r_E/\chi_{max})$  greater than  $r_E$ . Kovalenko and Shabalin<sup>75</sup> made this point, but in fact it is a general result, applicable whichever of the existing FEV models is used (though the predicted  $r_{min}$  value may be model dependent). It reflects the fact that to drive the liquid flow the pressure difference ( $p_b - p_a$ ) must be greater than zero.

Equations 2.30b and 2.37 yield a formal expression for apex radius  $r_a$  in terms of apex field  $F_a$ :

$$r_a = r_E/\chi = r_E / \{\mu(F_a) - (J_{FP}^2/J_{SP}^2) \exp[-2Q(F_a)/k_B T]\} \quad (2.38)$$

A formal expression for emission current can then be obtained from

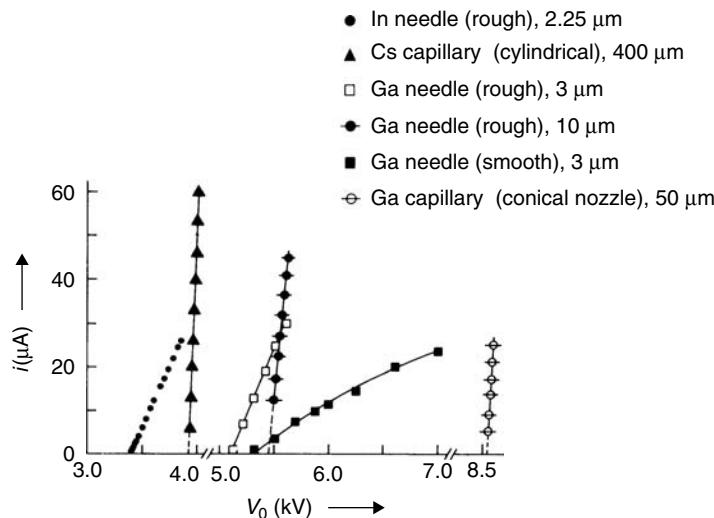
$$i = AJ_{Fa} = \alpha r_a^2 J_{Fa} \quad (2.39)$$

where  $A$  is the notional area of emission, and  $\alpha$  is a parameter that is almost certainly field dependent.

Clearly, the numerical values of  $r_a$  and  $i$  obtained from these equations will depend on the model assumed for  $Q(F)$  and on the values assumed for  $F_a$ ,  $J_{FP}$ , and  $\alpha$ . Also, there is nothing in the general mathematics set out here that requires  $F_a = F_E$ . So, in principle, it looks like  $J_{Fa} = J_{Sa}$  is a better (more general) apex boundary condition than  $F_a = F_E$ .

## 2.5 STEADY-STATE CURRENT-RELATED CHARACTERISTICS

A normal (low-drag) LMIS has the following experimental steady-state characteristics. (1) It turns on at an *onset voltage*  $V_{on}$ , and turns off at an *extinction voltage*  $V_x$ , usually slightly lower than  $V_{on}$ . (2) There is an associated ion emission current  $i_x$ , usually called the *minimum steady emission*



**FIGURE 2.12** Current–voltage characteristics for gallium and indium liquid metal ion sources (LMISs), together with one for a cesium capillary-type LMIS<sup>74,79–83</sup> (needle apex radius or capillary outer radius is indicated).

current but here called the *extinction current*. (3) Above  $V_x$ , the emission current is a nearly linear function of voltage (though with some curvature at higher voltages) and is usually well described by Mair’s equation (see Section 2.5.3). (4) The length of the liquid jet is proportional to emission current. Figure 2.12 shows experimental current–voltage characteristics for a variety of sources<sup>74,79–83</sup> (though without details near the extinction voltage).

Note that measured anode currents sometimes include components other than the ion emission current, possibly a large component due to secondary electrons and a small component due to other charged entities (see Section 2.8). FIB machines normally include a suppressor electrode, which minimizes the secondary-electron back flux. If no precautions have been taken, then true emission currents may be less than the measured currents by a factor as much as 2. With some older experimental articles, it is difficult to determine the contribution of secondary electrons.

### 2.5.1 BASIC THEORETICAL FORMULATION

Following Ref. 84, a convenient high-level approach starts by applying Newton’s second law to the motion of the GG cone-jet as a whole. It is assumed that liquid entering the cone base is emitted from the jet apex as ions with mean velocity equal to  $v_c$ , the bulk liquid velocity in plane C. This yields

$$f_I = d\mathcal{P}/dt = v_c d\mathcal{M}/dt = (mv_c/ze)i \equiv w_I i \quad (2.40)$$

where  $f_I$  is the effective steady-state force acting to accelerate the liquid bulk in the cone-jet,  $\mathcal{P}$  the total momentum of the cone-jet,  $d\mathcal{M}/dt$  the rate at which mass is ejected from the jet apex as ions, and  $w_I$  ( $\equiv mv_c/ze$ ) a parameter defined by Equation 2.40.

Writing  $i = f_I/w_I$ , we see that  $w_I$  has the nature of a resistance to force-induced motion. Parameters with the dimensions *force/electric current* have no well-recognized general name, so we refer to them as *force–current ratios (FCRs)*.

Writing down the various forces acting on the cone-jet, we obtain

$$f_L + f_{sc} + f_{st} + f_{hb} + f_{vis} - w_I i = 0 \quad (2.41)$$

where the terms represent, respectively, the Maxwell force that would act on the cone-jet in the absence of space charge (called, here, the *Laplace force*  $f_L$ ), the force  $f_{sc}$  due to the space-charge,

the force  $f_{\text{st}}$  due to surface tension that acts on the perimeter of the base of the liquid cone, the force  $f_{\text{hb}}$  due to the hydrostatic pressure that acts on the cone base, and a term  $f_{\text{vis}}$  associated with internal viscous forces. This last force is not really an external force but a way of including within Equation 2.41 the fact that the real external forces generate heat within the liquid due to atomic-scale viscous interactions.

Now consider a reference situation in which a quasi-conical liquid body is taken to be in hydrostatic equilibrium with zero emission current. Labeling forces that apply to this situation with the superscript “0,” and then subtracting Equation 2.42 from Equation 2.41, we have (using an obvious notation)

$$f_{\text{L}}^0 + f_{\text{st}}^0 + f_{\text{hb}}^0 = 0 \quad (2.42)$$

$$\Delta f_{\text{L}} + f_{\text{sc}} + \Delta f_{\text{st}} + \Delta f_{\text{hb}} + f_{\text{vis}} - w_{\text{I}}i = 0 \quad (2.43)$$

Take the direction of fluid flow along the cone-jet as the positive force direction. Then the Laplace term in Equation 2.43 is positive, is the driving term, is a function of applied voltage  $V$ , and may be written as  $\Delta f_{\text{L}}(V)$ . The next four terms are negative; each may be written formally in terms of an FCR in the forms  $f_{\text{sc}} = -w_{\text{sc}}i$ ,  $\Delta f_{\text{st}} = -w_{\text{st}}i$ , etc. Rearrangement then yields the formal current-voltage relationship

$$i = \Delta f_{\text{L}}(V) / [w_{\text{sc}} + w_{\text{st}} + w_{\text{hb}} + w_{\text{vis}} + w_{\text{I}}] \equiv \Delta f_{\text{L}}(V) / (\sum w) \quad (2.44)$$

where  $\sum w$  is defined by Equation 2.44. Obviously, Equation 2.44 has a superficial resemblance to Ohm’s law.

Further general formulas can be obtained via Van Dyke’s *slender-body approximation*,<sup>85,86</sup> as used by Taylor.<sup>87,88</sup> This states that for a slender body subject to applied voltage  $V$ , the Laplace force  $f_{\text{L}}$  does not depend on the details of apex shape and is given by

$$f_{\text{L}} = (4\pi\epsilon_0/k_{\text{v}})V^2 \quad (2.45)$$

where  $k_{\text{v}}$  is a constant dependent on apparatus geometry. The full rationalization factor  $4\pi\epsilon_0$  is included here to ensure that  $k_{\text{v}}$  is given by the dimensionless formulas in Taylor’s articles (where  $V$  denotes Gaussian potential rather than rmks potential). In the reference situation, define  $V_0$  by

$$f_{\text{L}}^0 = (4\pi\epsilon_0/k_{\text{v}})V_0^2 \quad (2.46)$$

so

$$\Delta f_{\text{L}}(V) = (4\pi\epsilon_0/k_{\text{v}})[V^2 - V_0^2] = f_{\text{L}}^0[(V/V_0)^2 - 1] \quad (2.47)$$

If the reference situation is assumed chosen such that the internal hydrostatic pressure is zero, then (1)  $f_{\text{hb}}^0 = 0$  and (2) we may consider the liquid shape to approximate to a Taylor cone (certainly near the cone base). In this case Equation 2.42 yields

$$f_{\text{L}}^0 = -f_{\text{st}}^0 = 2\pi r_{\text{b}}\gamma \cos\phi_{\text{T}} \quad (2.48)$$

where  $r_{\text{b}}$  is the cone base radius. Alternatively, we can assume (arbitrarily) that the liquid shape is an exact Taylor cone, subject to Taylor’s field distribution<sup>11</sup> over its surface, and integrate the Maxwell-stress component parallel to the cone axis over the surface. This yields

$$f_{\text{L}}^0 \sim \epsilon_0 F_{\text{b}}^2 \cdot \pi r_{\text{b}}^2 \quad (2.49)$$

where  $F_b$  is the field acting on the conical surface at the base of the liquid cone, *as given by Taylor's formula*.<sup>11</sup> Equation 2.49 should not be considered an exact physical result, because the conditions for exact validity previous of the assumptions it involves rarely, if ever, apply.

From the above analysis, two useful results follow. First, a relatively general expression for the LMIS  $i$ - $V$  relationship is

$$i = [f_L^0 / (\sum_w)] [(V/V_0)^2 - 1], \quad (2.50)$$

where  $f_L^0$  is given by Equation 2.48 or 2.49. Equation 2.48 is preferable, because it is physically more secure, and because it contains parameters that can be easily measured. By making detailed assumptions and approximations in Equation 2.50, specific  $i$ - $V$  formulae are obtained.

Second, if in the reference situation the internal hydrostatic pressure in the liquid cone is zero ( $f_{hb}^0 = 0$ ), then we can identify  $V_0$  as the static collapse voltage  $V_c$  for a low-drag emitter. Equations 2.46 and 2.48 then show that  $V_c$  is given by

$$V_c = (k_V r_b \gamma \cos \phi_T / 2 \epsilon_0)^{1/2}. \quad (2.51)$$

### 2.5.2 EXTINCTION AND COLLAPSE VOLTAGES

LMIS extinction has usually been thought to have a total-force-based explanation, as just set out. This approach is generally similar to Taylor's<sup>88</sup> explanations of the behavior of water cones and soap films. However, discussions<sup>37,38,75,89</sup> on the origin of extinction current imply the possibility of an alternative, *steady-state dynamic*, explanation based on the relative sizes of  $J_{Sa}$  and  $J_{Fa}$ . Possibly a *steady-state-dynamic collapse voltage*  $V_d$ , slightly different from  $V_c$ , would be predicted. Further, due to hydrodynamic fluctuations, one might expect observed extinction voltages  $V_x$  to be slightly higher than whichever of  $V_c$  and  $V_d$  corresponds to the correct theoretical explanation of steady-state extinction. However, all these voltages are expected to be close, and predictions of  $V_c$  have been successfully used<sup>16,90</sup> to predict  $V_x$ . This justifies applying total-force arguments and the slender-body approximation to the LMIS and supports the idea of using  $V_c$  to predict  $V_x$ .

While passing, we note that successful predictions of LMIS onset voltage have been made<sup>15,90</sup> using the presume-jump criterion  $\Delta p = 0$  and needle models that give the apex field as a function of voltage. Given our remarks in Section 2.4.2, this is not necessarily expected *a priori*. It presumably happens because the true criterion is that the internal pressure within the film adhering to the needle has some value  $p_{on}$  that happens to be small in magnitude, in comparison with the absolute magnitudes of the electrical and surfaces tension stresses acting. In such circumstances, the approximation  $p_{on} \approx 0$  (hence  $\Delta p \approx 0$ ) may give small error in the estimation of onset field and voltage.

### 2.5.3 CURRENT-VOLTAGE CHARACTERISTIC ABOVE EXTINCTION

#### 2.5.3.1 Mair's Equation

Mair's equation<sup>91,92</sup> describes the  $i$ - $V$  characteristics of a low-drag LMIS and applies to both the capillary LMIS and the normal LMIS. The derivation assumes that space-charge effects<sup>93</sup> dominate, and uses only the first two terms in Equation 2.41, that is, only the FCR  $w_{sc}$  in Equation 2.44.

The force  $f_{sc}$  relates to space-charge. In planar theory, the force  $f_{psc}$  acting on an area  $S$  is given by Equations 2.11, 2.14, and 2.15 as

$$f_{psc} = -g(\zeta) \zeta M_p S = -g \kappa J F_p^{-2} V^{1/2} \cdot \frac{1}{2} \epsilon_0 F_p^2 \cdot S = -\frac{1}{2} g(m/2ze)^{1/2} V^{1/2} i \quad (2.52)$$

Although LMIS geometry is not planar, Mair assumes we can approximate  $f_{sc} \approx f_{psc}$ , which implies

$$w_{sc} \approx \frac{1}{2} g(m/2ze)^{1/2} V^{1/2} \quad (2.53)$$

For gallium, for  $V = 5$  kV,  $w_{sc} \approx 0.05$  N/A.

In Equation 2.50, Mair uses (in  $\Sigma w$ ) only the term  $w_{sc}$ . From Equations 2.48, 2.50, and 2.53, assuming  $V_x \equiv V_0$ , we get

$$i = (2g^{-1})(2\pi r_b \gamma \cos \phi_T)(2ze/mV_x)^{1/2}[(V/V_x)^{3/2} - (V/V_x)^{-1/2}] \quad (2.54)$$

For large  $V$ ,  $i$  becomes proportional to  $V^{3/2}$ , which is the signature of space-charge-limited current.

The weak space-charge (low  $\zeta$ ) result (see Equation 2.15) can be used to put  $g^{-1} \approx 3/8$ . Then, on writing  $V = V_x + \delta V$ , binomial expansions give *Mair's equation* in its usual cited form:

$$i \approx 3\pi r_b \gamma \cos \phi_T (2ze/mV_x)^{1/2}[(V/V_x) - 1] \quad (2.55)$$

It could be argued that a slightly better result would, in principle, be obtained by using the strong space-charge result  $g^{-1} \approx 4/9$ , but the difference is unimportant and it is convenient to have the simple factor “3” in Equation 2.55. An abbreviated form is also convenient:

$$i = C_M r_b V_x^{-1/2} (V/V_x - 1) \quad (2.56a)$$

$$C_M \equiv 3\pi(2ze/m)^{1/2} \gamma \cos \phi_T \quad (2.56b)$$

where  $C_M$  is a material-specific constant. Appendix A.1 lists values for some elements of interest.

From Equation 2.56a, it can be seen that  $di/dV$  near extinction is predicted to depend only on known constants and the measurable parameters  $r_b$  and  $V_x$ . This prediction was tested by Mair,<sup>92</sup> by using experimental  $di/dV$  values to plot values of  $[(di/dV)/(C_M V_x^{-3/2})]$  against  $r_b$  for gallium, indium, and cesium low-drag sources. The outcome<sup>92</sup> is a convincing straight line over the  $r_b$  range from 1 to 300  $\mu\text{m}$ . This theory certainly works.

Most of the approximations made are easily justified. Using Equation 2.40 and the value  $v_c \approx 35$  m/s derived in Section 2.4.4,  $w_I$  may be estimated as  $2.5 \times 10^{-5}$  N/A; this is much smaller than  $w_{sc}$ . Beckman<sup>37</sup> has shown that viscous effects are small, which implies that  $w_{vis}$  is small. By definition, a low-drag source has  $|f_{hb}|$  and  $|\Delta f_{hb}|$  very small and  $w_{hb} \ll w_{sc}$ , and this implies that  $|\Delta f_{st}|$  and  $w_{st}$  are very small. So, for a low-drag source,  $w_{sc}$  is the only significant term in  $\Sigma w$ .

The only real difficulties are in the assumptions that planar space-charge theory may be applied and that the force due to the space-charge acts only on the liquid area (near the emitter apex) from which emission occurs. While neither of these things is plausible for a general situation, they may be reasonable for the cusp-like LMIS geometry, where much of the space-charge is close to the apex, and the cusp partially shields at least the nearer parts of the cone from fields due to space-charge.

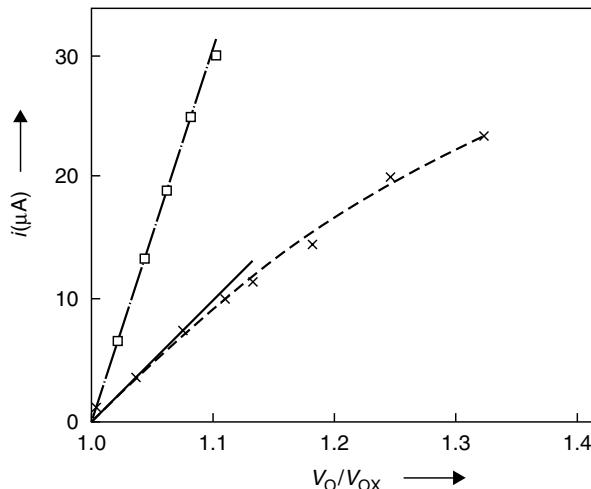
We also note that if  $(w_I + w_{vis})$  is only a very small component in  $\Sigma w$ , then in effect, the operating LMIS is subject to the constraint that the total force on the GG cone-jet be zero. This is perhaps the most fundamental way of looking at Mair's theory.

### 2.5.3.2 Mahoney's Equation

Many years ago, Mahoney et al.<sup>94</sup> found good agreement between  $i$ - $V$  characteristics for cesium, up to currents of 100  $\mu\text{A}$ , and the empirical equation

$$i = K[(V/V_x)^{3/2} - 1] \quad (2.57)$$

where  $K$  is a constant of value  $\sim 50$   $\mu\text{A}$ . Clearly, Equation 2.54 is close in form to Equation 2.57, particularly at higher voltages. A theoretical equation of form 2.57 has also been derived by Weinstein,<sup>95</sup> by analyzing the space-charge effects of a cylindrical ion beam. (This was an early attempt to develop what is now called a two-dimensional Child-Langmuir Law.<sup>96</sup>) If Equation 2.49, rather than Equation 2.48, is used in Equation 2.50, then the result for  $K$  is similar to the expression he derived.



**FIGURE 2.13** Current–voltage characteristics for a *rough* (low-drag) (□) and a *smooth* (×) gallium liquid metal ion source, each with apex radius 3 μm, showing the effect of the flow impedance of the smooth needle. (Experimental points are taken from Kingham, D.R., and Swanson, L.W., *Appl. Phys.*, A 34, 123–132, 1984.) The fitted curves illustrate that the experimental results can be fitted by the theory presented in general form here, more explicitly in Mair, G.L.R., *J. Phys. D: Appl. Phys.*, 17, 1945–1950, 1997, using physically reasonable choices of parameter values.

### 2.5.3.3 The Effect of Flow Impedance

Experimentally, it was found that flow impedance between the reservoir and the cone base reduces the slope of the  $i$ – $V$  characteristic.<sup>15</sup> This is to be expected if  $w_{hb}$  in Equations 2.44 and 2.50 is no longer small. Figure 2.13 shows this effect for Ga sources with apex radius 3 μm; the smooth needle has higher impedance than the rough one.

For flow along the needle, the hydrodynamic *volume flow impedance*  $Z_{vol}$  is defined as  $(p_r - p_b)/R_{vol}$ , where  $R_{vol}$  is the *volume flow rate*, and  $(p_r - p_b)$  is the pressure drop from reservoir to cone base. If we can assume  $p_r \approx 0$ , then  $w_{hb}$  is related to  $Z_{vol}$  by

$$w_{hb} = \pi r_b^2 (\omega/ze) Z_{vol} \quad (2.58)$$

where  $\omega$ , as before, is the atomic volume.

If  $w_{hb}$  is not small, then one expects  $w_{st}$  to be nonzero. Mair<sup>97</sup> has investigated several flow models for  $Z_{vol}$  and also the effect of variation in the angle  $\phi$  with current (this determines  $w_{st}$ ); he has shown that the smooth-needle data in Figure 2.13 can be fitted using physically plausible parameter values.

If  $(w_{hb} + w_{st}) \gg w_{sc}$ , then the LMIS can be described as *high drag*, the  $i$ – $V$  relationship is determined by the hydrodynamics of liquid flow along the needle, and the LMIS operates in the regime identified by Wagner<sup>98</sup> (but the theory given in this chapter is more satisfactory). In the intermediate-drag regime, where  $(w_{hb} + w_{st}) \sim w_{sc}$ , the  $i$ – $V$  relationship can be determined either directly from Equation 2.50 or by multiplying Mair's formula by the correction factor  $w_{sc}/(w_{sc} + w_{hb} + w_{st})$ . Mair<sup>97</sup> gives some explicit formulas.

More generally, the analysis here brings out that for a low-drag source the *opposing force* (that counteracts the Laplace force) is due to space-charge above the liquid apex, whereas the opposing force for a high-drag source originates at the needle shank and is transmitted through the liquid. Physically, it is probably no surprise that high-drag sources pulsate and low-drag sources do not.

### 2.5.3.4 Effects of Temperature

At the high apex flux densities involved in LMIS operation, changes in temperature ( $T$ ) are expected to have limited effect on FEV rate-constants, so the main effect is via the change in the surface free energy  $\gamma$ . Jousten et al.<sup>99</sup> used this effect to modulate emission current by directing a chopped laser beam at an LMIS. In other experiments,<sup>100,101</sup> anomalies in the  $T$ -dependence of  $\gamma$  have been found. Theory has been developed by Mair.<sup>102</sup>

In summary, the substantial agreement between experimental results and the theory presented here shows that steady-state LMIS  $i$ - $V$  behavior can be explained by a theory based on total-force arguments, the slender-body approximation, and the presence of space-charge, without any need to resort to detailed arguments concerning apex field and current density. Although this may seem counterintuitive, it is characteristic of space-charge-controlled situations that emission details are of limited importance.

### 2.5.4 DEPENDENCE OF CUSP LENGTH ON EMISSION CURRENT

The LMIS jet length  $\lambda$  can be defined as the distance between the jet apex and the tip of the cone defined by the limiting shape of the GG cone-jet. Electron microscope (EM) observations<sup>14,23,33,103,104</sup> show that jet length usually increases linearly with emission current. For gallium and many other metals, lengths of tens to hundreds of nanometers are observed at moderate-to-high currents; for cesium, lengths at high current can reach 1  $\mu\text{m}$  or more. Analyzing this effect helps establish consistent overall theory.

The physics is reasonably clear, but details are not fully established. When the applied voltage is increased, the Laplace force  $f_L$  on the emitter increases. The resultant force  $f$  temporarily increases and additional liquid is pulled into the cone-jet, which grows. The growth process stops when current and space-charge have increased sufficiently to counteract the increase in  $f_L$ . As noted earlier, LMISs operate under conditions where change in apex field has limited effect on the FEV rate-constant and apex current density. So, current increase is presumably associated mainly with increase in the jet radius  $r_c$  at the jet/cap join. The relationship between apex field and applied voltage depends on emitter shape, even when space-charge effects are taken into account; so, it may be assumed that increase in  $r_c$  would (for a given jet length) produce an emitter that is blunter and has a lower apex field. However, FEV cannot occur if the apex field gets too low. So, the new apex equilibrium position needs to involve both greater jet radius and greater jet length to ensure that apex field is roughly the same as before.

Initial attempts to model jet-length changes used LMIS shape-modeling programs,<sup>73,74</sup> designed to treat space-charge self-consistently in a real experimental configuration and an assumed emitter geometry. Later, analytical treatments<sup>105,106</sup> bypassed detailed space-charge considerations by using the current–voltage characteristic; this gave the equilibrium jet length ( $\lambda$ ) as

$$\lambda = C_\lambda V_x^{1/2} i, \quad C_\lambda \equiv (2m/ze)^{1/2} / 3\pi\gamma \quad (2.59)$$

Predicted values of  $C_\lambda$  given in Appendix A.1 are in fair agreement with measured values (Forbes et al., unpublished work).

### 2.5.5 EXTINCTION CURRENT

As already noted, a normal (low-drag) LMIS has an observed, material-specific, temperature-dependent extinction current  $i_x$  associated with relatively sharp turnoff at the observed extinction voltage  $V_x$ . Beckman<sup>37</sup> examined many sources and concluded that  $i_x$  relates only to the physical processes that take place near the tip of a metallic GG cone-jet. For gallium at 30°C, he found  $i_x \equiv 0.45 \mu\text{A}$ , which is close to the value ( $\sim 0.5 \mu\text{A}$ ) Mair et al. found earlier.<sup>39,40</sup> Lower currents,

observed for emitters that are not low drag,<sup>37,42</sup> were shown to be average currents in a pulsation regime,<sup>37</sup> as suggested earlier by Shabalin.<sup>89</sup> For gallium, between 30 and 890°C, Beckman found<sup>37</sup>

$$i_x = (1.19 \mu\text{A}) \exp[-(0.025 \text{ eV})/k_B T] \quad (2.60)$$

At least in principle, a normal LMIS probably has the following jet-length ( $l$ ) regimes. (1) For  $l < l_1$  the pressure gradient is reversed, and the jet actively collapses. (2) For  $l_1 < l < l_2$ , the flow is forward but so slow that FEV is dominant and the jet shrinks. (3) For  $l_2 < l < l_3$ , an equilibrium situation (and length  $\lambda$ ) can exist where liquid flow and FEV can balance (but the jet can be knocked into regime 1 or regime 2 by external events or statistical fluctuations). (4) For  $l_3 < l < l_4$ , the jet is quasi-steady. (5) Longer jets ( $l > l_4$ ) are subject to hydrodynamic instability and break-up.

When an equilibrium length exists, there exist associated values of other parameters. Theory seeks the values ( $F_x, J_x, r_x, A_x, i_x$ ) of apex field, apex current density, apex radius, effective emission area, and of current that correspond to  $l_2$ . It is easy to argue that at the point at which the extinction process (jet collapse) begins the LMIS is still emitting with a current given by

$$i_x = J_x A_x = J_x \alpha_x r_x^2 \quad (2.61)$$

where the area  $A_x$  has been written in the form  $\alpha_x r_x^2$ . But, so far, it has proven impossible to derive reliable estimates of any of  $J_x$ ,  $\alpha_x$ , or  $r_x$ .

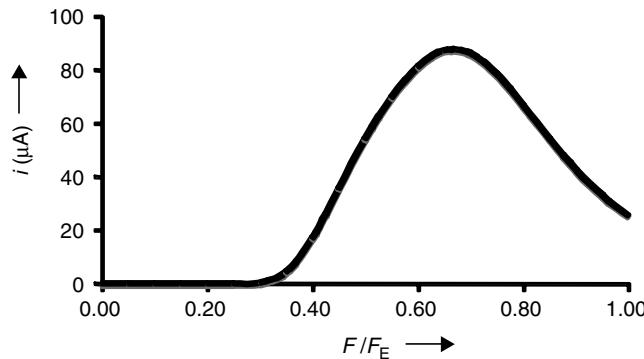
Treatments have been put forward by Kovalenko and Shabalin,<sup>75</sup> Beckman and colleagues,<sup>37,38</sup> and Suvorov and Forbes.<sup>107</sup> All three yield roughly similar values for  $r_x$ , all roughly equal to the parameter  $r_E$  defined earlier. The first two were empirically successful in predicting  $i_x$  for gallium, apparently mainly because they used models in which  $\alpha_x = 0.46$  and 0.474, respectively. (By contrast, Ref. 107 used  $\alpha_x = \pi$ .)

However, there are conceptual difficulties with all these treatments; none of which is fully consistent with all established aspects of LMIS physics. Briefly, the first two treatments use models of FEV that are known to be flawed. The third treatment does not have this difficulty, but its mathematical behavior is qualitatively different from the older models. It identifies an equilibrium ( $J_{Fa} = J_{Sa}$ ) situation, but this appears to be unstable. It would seem that either FEV at the LMIS apex takes place in a high-field regime for which existing FEV theory is not adequately developed, or that the FEV equation used here needs to be replaced by an equation that includes the effects of field-emitted vacuum space-charge, or that the control mechanisms that keep the LMIS steady have essential three-dimensional aspects that are not adequately captured by the existing one-dimensional analyses. Thus, substantial uncertainty exists, and (at the time of writing) the precise origin of the LMIS extinction is not physically understood. The theoretical work of Higuera,<sup>108</sup> and possibly that of Boltachev and Zubarev,<sup>109</sup> will also need to be taken into account.

One provisional result from this confused situation deserves mention. If Equation 2.38 is evaluated using the 1982 CE formula 2.9, and Equation 2.39 is used with  $\alpha = \pi$ , then the dependence of emission current on field has the form shown in Figure 2.14. Near extinction the absolute ion-current values predicted are much too high. But, counterintuitively, it is predicted that the LMIS surface field *decreases* slowly as the current increases. Since the probabilities of both PFI and direct escape into higher charge states increase with field, this prediction can and should be checked by carefully directed experiments.

### 2.5.5.1 The Practical LMIS as a Physical Chaotic Attractor

Consider an LMIS operating at a length just above  $l_2$  and subject to a source or sources of repeated external perturbation, for example, instability in the high-voltage supply, mechanical vibration, or impact events. If any of these can increase the theoretical value of  $l_2$  sufficiently and sufficiently quickly (so the jet cannot follow) or can knock the jet length below  $l_2$  (perhaps by inducing the



**FIGURE 2.14** To illustrate the relationship between emission current  $i$  and apex field  $F$  predicted for a gallium liquid metal ion source by the Forbes 1982 mathematical model of field evaporation, using Equations 2.9, 2.38, and 2.39 with  $\alpha = \pi$ , and the estimated value  $C_Q = 0.2$  eV. The  $x$ -axis is normalized by dividing  $F$  by the maximum evaporation field  $F_E$ . Note that, according to this treatment, *increase* in emission current is associated with *decrease* in evaporation field.

ejection of a charged nanodroplet), the jet may be put into a situation where it shrinks or collapses. Thus, minimum length ( $l_3$ ) and current ( $i_3$ ) values for practical steadiness (i.e., low chance of being knocked unsteady) are likely to be greater than those predicted by disregarding perturbations.

In this case, one can think of a theoretical equilibrium length  $\lambda$  ( $>l_3$ ) as a physical chaotic attractor about which the actual jet length is quasi-randomly fluctuating. For quasi-steady LMIS operation,  $l_3$  has to be sufficiently greater than  $l_2$  so that the chance of the jet getting into a region of its configuration space where it shrinks or collapses is acceptably small. The corresponding practical minimum steady current  $i_3$  determines the lowest current that can be reliably used in FIB applications.

Given this physical picture, practical ways of minimizing  $i_3$  might include<sup>107</sup> using highly stable high-voltage power supplies, reducing any other propensity of the system to initiate hydrodynamic excursions, slowing down the rate at which these excursions develop (perhaps by designing the LMIS such that viscous-drag effects operate near the support needle tip but not elsewhere, if possible), applying a fast-feedback control system to the high-voltage supply, and operating at a temperature as low as possible.

## 2.6 ENERGY DISTRIBUTION AND ION-OPTICAL CHARACTERISTICS

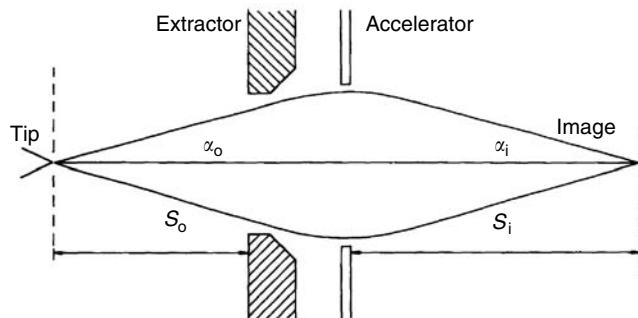
### 2.6.1 INTRODUCTION

These properties are considered together because ion-optical characteristics are affected by the ion energy distribution, and both are affected by strong space-charge. Ion-optical column design is discussed in the literature, but is beyond the scope of this chapter. The following two equations show, in a simple way, how source characteristics relate to column design. An implicit assumption is that the beam profile and the energy distribution can both be regarded as Gaussian.

Figure 2.15 shows an idealized beam-forming system. For low values of the angle of beam convergence ( $\alpha_i$ ) on the image side of the optics, the two most important terms in the usual approximate expression for the spot diameter  $d$  at the specimen relate to (1) the virtual diameter  $d_v$  of the source and (2) the *chromatic aberration* associated with the full-width-at-half-maximum (FWHM)  $\Delta E$  of the ion energy distribution. ( $\Delta E$  is often called the *energy spread*.) Thus

$$d^2 = M_o^2 d_v^2 + \alpha_i^2 C_c^2 (\Delta E/E)^2 \quad (2.62)$$

where  $M_o$  is the optical magnification,  $E$  the (final) ion energy in the focused beam, and  $C_c$  the chromatic aberration coefficient referred to the image side. In the limit of very low  $\alpha_i$ , the term in  $d_v$



**FIGURE 2.15** Schematic arrangement of a simple two-element ion probe-forming system using a liquid metal ion source.

defines the size of the central part of the spot, but any low-energy ions present may cause the spot to have a weak outer *penumbra*. In calculations on a specific system, Cleaver and Ahmed<sup>110</sup> found that the term in  $d_v$  became limiting at spot diameters below  $\sim 50$  nm.

According to Prewett and Mair,<sup>1</sup> if  $\alpha_i$  is chosen by design to provide the required current  $i_B$  in the focused beam (typically well below 1 nA and often below 10 pA), then Equation 2.62 may be rewritten as

$$d^2 = M_o^2 d_v^2 + \pi^{-1} (i_B M_o C_c / E)^2 / \Phi \quad (2.63)$$

where the *chromatic angular intensity*  $\Phi$  is defined in terms of the source angular intensity  $di/d\Omega$  by

$$\Phi = (di/d\Omega) / (\Delta E)^2 \quad (2.64)$$

In general terms, therefore, a source needs low  $d_v$  to minimize the first term and high  $\Phi$  to minimize the second term in Equation 2.64. Modern FIB systems tend to be source-size limited, but for completeness this chapter provides basic angular-intensity information.

## 2.6.2 ION ENERGY DISTRIBUTIONS

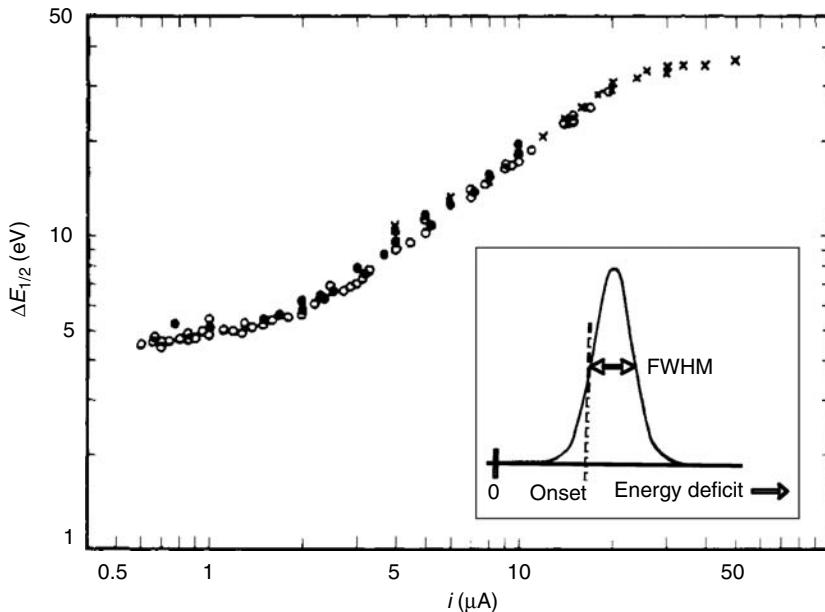
The inset to Figure 2.16 shows a typical ion energy distribution for the  $\text{Ga}^+$  ions emitted by a gallium LMIS. This distribution is plotted as a function of the measured *ion energy deficit*  $D$  defined by  $D = \xi e V_s$ , where  $V_s$  is the voltage difference ( $V_{\text{emitter}} - V_{\text{collector}}$ ) needed to bring an emitted ion (in final charge state  $\xi$ ) to a halt just outside the surface of a collector of work-function  $\phi_c$ .

### 2.6.2.1 Onset Energy Deficits

If an ion in final charge state  $\xi$  is created at the emitter surface by transferring its electrons directly to the emitter Fermi level, then its initial  $D$ -value is the *onset energy deficit*  $D^{\text{on}}$  given by<sup>111</sup>

$$D^{\text{on}} = \Lambda^F + H_\xi - \xi \phi_c - Q \quad (2.65)$$

where  $\Lambda^F$  is the bonding energy of the atom in the presence of the field, and  $H_\xi$  the sum of its first  $\xi$  ionization energies. For  $\text{Ga}^+$  ions the predicted value is  $D^{\text{on}} \approx 5$  eV. Space-charge broadening, as the ion moves, will randomly change the ion energy; the measured peak energy deficit may then be a better estimate of the theoretical onset energy deficit. If, as with  $\text{Ga}^+$ , the distribution peak has a measured deficit close to the predicted  $D^{\text{on}}$  value, then ion formation was a surface process. Significantly higher energy deficits indicate either ion formation in space away from the emitter or energy loss in ion-ion interactions during motion.



**FIGURE 2.16** To illustrate how the energy spread (full width at half maximum) of a gallium liquid metal ion source (LMIS) varies with emission current. The inset shows a typical measured ion energy distribution for an LMIS, plotted as a function of energy deficit. Onset as defined by a *half-maximum* criterion is shown. Because space-charge broadening is taking place, the theoretical onset energy deficit is expected to correspond to some energy between the half-maximum deficit as shown and the energy deficit of the distribution peak.

### 2.6.2.2 Energy Spreads

Perfect focusing in the ion optics requires *monochromatic* beams with all ions having the same energy deficit. This cannot be achieved, so chromatic aberration occurs. The main effect is associated with the width of the ion energy distribution, which is assessed by its FWHM. The FEV natural width, at low current density, is of order 1 eV or less.<sup>40,112</sup> The FWHM for a normal LMIS is always significantly greater than this, typically  $\sim 5$  eV or more, due to space-charge broadening.

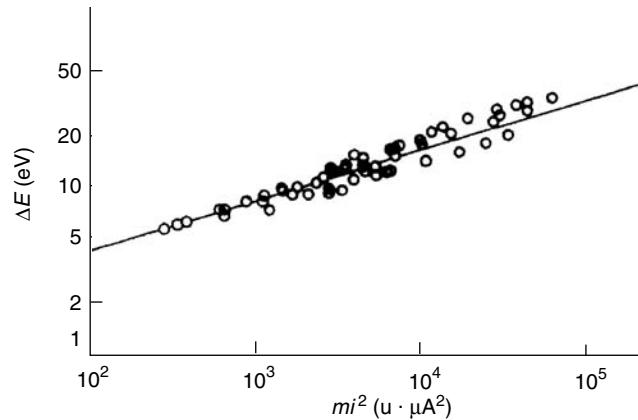
For many materials, the variation of FWHM with emission current, when plotted logarithmically, has the S-shaped form as shown in Figure 2.16. The linear middle region of this logarithmic plot is associated with the *most-steady* emission regime and has received considerable attention. General theoretical discussions include Refs 113–118; computer simulations include Refs 77 and 119. For gallium, this regime starts at  $\sim 2 \mu\text{A}$ . Accepted thinking is that the FWHM here is determined by many-pair, *collision free*, stochastic coulomb interactions in the emitted beam,\* close to the emitter and should theoretically have the approximate functional dependences:

$$\Delta E \sim (mi^2)^{1/3} T^{1/2} \quad (2.66)$$

Dependence on some power of the product  $(mi^2)$  is expected on dimensional grounds;<sup>117</sup> dependence on the value  $1/3$  is a particular interpretation<sup>†</sup> of Knauer's work.<sup>114,115</sup> For the experimental dependence, Mair's graphical summary<sup>44</sup> is convincing, with a value just over 0.3 (see Figure 2.17).

\* This effect is sometimes called the “Boersch effect,” but there is no real crossover associated with a field emitter, so it is debatable whether the name is applicable. The theory is similar, but not identical, to that of the original Boersch effect. See Ref. 118 for a brief review.

† Knauer's calculations are performed in the context of spherical symmetry. In principle, more than one way exists of adapting these calculations to the context of the conical beam from a field emitter.



**FIGURE 2.17** Energy spread (full width at half maximum) plotted against the product  $mi^2$ , for several metals (gallium, indium, cesium, aluminum, and lithium) that generate mainly singly charged ions. Data correspond to the middle (linear) part of S-curves similar to that shown in Figure 2.16.

As already noted, low-drag sources turn off below an extinction current ( $\sim 0.45 \mu\text{A}$  for Ga). Below the most-steady regime down to extinction, there is an *intermediate regime* where the  $\Delta E$  versus  $i$  curve flattens, and the measured peak energy deficit exhibits anomalous behavior.<sup>120</sup> It is not surprising that this regime exists, but its details are not properly understood.

The upper part of the S-curve is characterized by having a different power dependence of  $\Delta E$  on current. Mair<sup>44</sup> has argued that the transition is associated with the onset (at higher applied voltages) of gross jet instability that causes violation of a *no ion path crossing* condition used in the derivation of Equation 2.55. He also argues that some materials (in particular  $\text{Au}^+$ ,  $\text{Sn}^+$ , and  $\text{Bi}^+$ , which follow an  $i^\beta$  power law with  $\beta$  in the range 0.35–0.39) do not have a most-steady regime but only an upper regime (discussed further in Section 2.8).

### 2.6.2.3 Distribution Shape and the Low-Energy Tail

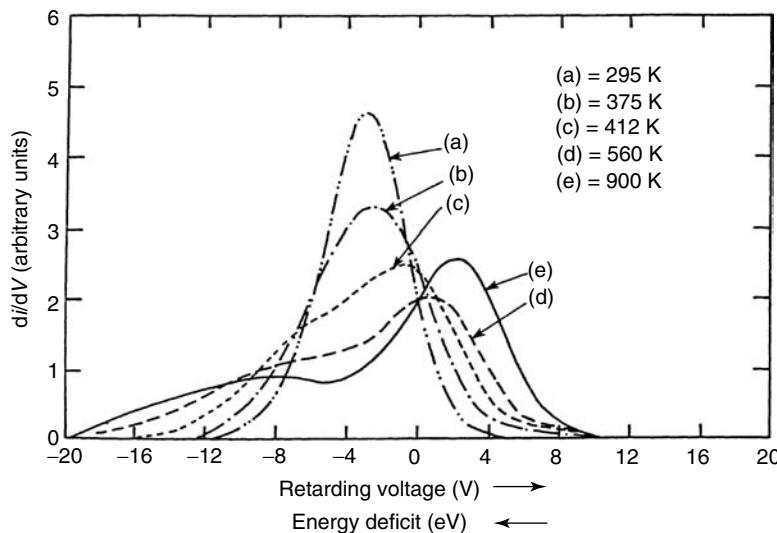
The shape of the main peak is often approximated as Gaussian, but Ward et al.<sup>121</sup> suggested that it is better represented as a Holtsmark distribution, because a tail of low-energy (high energy deficit) ions exists<sup>122</sup> that can create a *penumbra of light milling* around the main artifact under creation. Ward et al. attributed these ions to *transverse thermal velocity broadening*. Kubena and Ward showed later that they were not due to scattering between primary beam ions and residual gas molecules.<sup>123</sup> However, it is perhaps more likely that many of these ions, particularly those with high energy deficit, are created by free-space FI of neutrals\* or by charge-exchange (CE) between an energetic primary ion and a slow neutral in the vacuum space.<sup>124</sup>

### 2.6.2.4 Effects of Temperature

The effect of temperature on the FWHM has been investigated theoretically by Kim et al.<sup>125</sup> Their work suggested that small thermal-energy variations (of order 0.01 eV) existing at emission could be magnified by stochastic coulomb effects into a  $T^{1/2}$  broadening of the FWHM by as much as several electron volts at moderate to high emission currents. This is the observed experimental behavior.<sup>126,127</sup>

A second experimental effect was the development of a double-hump distribution<sup>128</sup> at higher temperatures as shown in Figure 2.18. Later work by Komuro et al.<sup>129</sup> suggested this might really be a triple-peak distribution; they associated the peaks with FEV, free-space FI of incoming neutrals,

\* See discussion of this effect in Chapter 3.



**FIGURE 2.18** Measured energy distributions, at various temperatures, for a Ga liquid metal ion source operating at a current of  $2 \mu\text{A}$ . Note that energy deficit increases toward the left. (From Swanson, L.W. et al., *J. Appl. Phys.*, 51, 3453–3455, 1980.)

and slow ions created by CE, respectively. The calculations of Kosuge et al.<sup>130</sup> provide some support to the three-peak hypothesis.

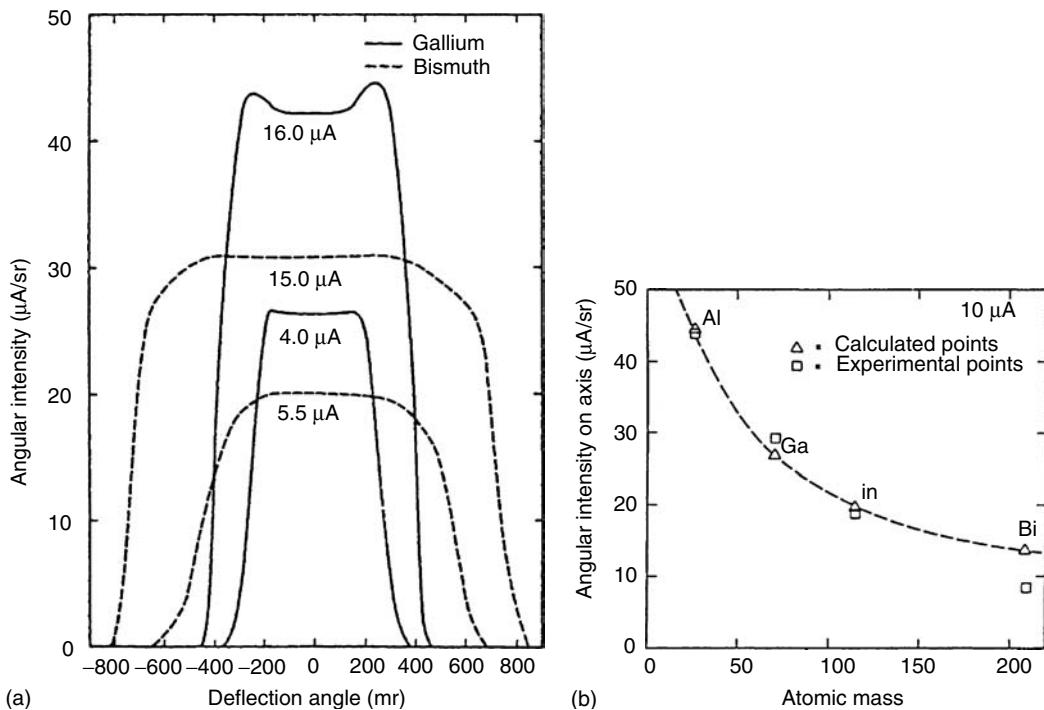
An unresolved problem is that the ions in the lowest-energy-deficit peak (due to FEV) have deficits less than the predicted onset energy deficit. This shows that a mechanism exists by which these ions receive additional kinetic energy after emission. (It is not statistically possible for this to happen during normal FEV.) Hornsey<sup>131</sup> has attributed the effect to oscillation of the jet behind the departing ions: when its (constant potential) surface is moving forward, this pushes additional energy into the departing ions; when it is moving backward this extracts energy. Computer simulations are stated<sup>131</sup> to support this idea. An alternative is that the effect is associated with the development of space-charge oscillations; a similar double-hump effect has been observed in electron emission from solids.<sup>132</sup> We suggest that a third possibility is a curious space-charge effect, arising because the ions in the second Komuro et al. peak are created in space (by FSFI) at distances from the emitter greater than those associated with the ions in the first peak (created by FEV). The two groups of ions would tend to “push each other apart,” as they begin to move: this would reduce the energy deficit for the first group (as observed) and increase it for the second.

### 2.6.3 ION-OPTICAL EFFECTS

#### 2.6.3.1 Angular Intensities

As illustrated in Figure 2.19a, early measurements<sup>128</sup> showed that, for emission currents of interest, there is a considerable range of axial angle over which angular intensity  $di/d\Omega$  is constant. This range increases with emission current, presumably due mainly to the related jet-length increase and associated electrostatic effects. For gallium, at  $i = 4 \mu\text{A}$ , the range is  $\sim 200$  mrad; this is much greater than the source-side acceptance angles employed in column design, which are typically less than 10 mrad.

The value of this constant angular intensity depends on the ion mass and the shape of the emitting structure, as shown in Figure 2.19b. Cleaver and Ahmed<sup>110</sup> assumed a value  $20 \mu\text{A}/\text{sr}$  in their design calculations. Extensive information on experimental angular intensities exists in the literature.



**FIGURE 2.19** Angular intensity ( $di/d\Omega$ ) data (a) as a function of angle for gallium and bismuth at different emission currents (using apertures subtending angles of 108 and 146  $\mu\text{sr}$ , respectively) (From Swanson et al., L.W., *J. Vac. Sci. Technol.*, 16, 1864–1867, 1979.) and (b) comparison for an emission current of 10  $\mu\text{A}$  between on-axis experimental values and numerical modeling (assuming a low-drag source). (From Kingham, D.R. and Swanson, L.W., *Appl. Phys. A*, 34, 123–132, 1984.)

In the most-steady regime  $(\Delta E)^2$  increases with emission current more rapidly than  $di/d\Omega$ , so the chromatic angular intensity  $\Phi$  decreases with emission current. The second RHS term in Equation 2.63 is thus minimized by operating at low emission current.

### 2.6.3.2 Optical Source Size

For modern FIB-machine designs, a more important limitation is that the ion-optical diameter of a normal LMIS is deduced from experiments<sup>133</sup> to be around 40–50 nm. This is much greater than the apex size as predicted theoretically or as observed by electron microscopy.

Suggested causes have been lateral movement of the liquid jet (*jet wobble*) and a space-charge effect called *lateral broadening* or *trajectory displacement* in which stochastic coulomb interactions between particle pairs “push the particles sideways.” This modifies their trajectories in such a fashion that backward ray tracing, outside the region of strong coulomb interaction near the emitter, predicts that the ions come from a virtual source of diameter much greater than that of the physical emitter.

Jansen<sup>118,134</sup> has published general theoretical discussions of trajectory displacement. For the LMIS, Monte Carlo calculations were made by Ward,<sup>135</sup> using the sphere-on-orthogonal-cone (SOC) model for emitter geometry and an apex radius of 10 nm and later by Georgieva et al.<sup>136</sup> Ward derived virtual source diameters of 50–100 nm and Georgieva et al. diameters of order 100 nm. More recent calculations by Radlicka and Lencová,<sup>77,78</sup> based on the shape model used in the LMIS modeling programs<sup>73,74</sup> but with  $r_a = r_c = 1.86 \text{ nm}$ ,  $\lambda = 6.12 \text{ nm}$ , predict a virtual source diameter of 50 nm, which is comparable with experimental values. In consequence, there seems no need to consider that jet wobble contributes significantly.

## 2.7 IONIZATION MECHANISMS AND EMITTED SPECIES

### 2.7.1 ELEMENTAL ION SOURCES

#### 2.7.1.1 Atomic Ions

As already noted, an LMIS can field evaporate atomic ions in a mixture of charge states. This will happen if the field necessary for PFI is comparable with the field necessary for escape. For elements, Table A.1 shows Müller escape fields  $F_n^M$  for atomic ions in charge states 1–3 and also the Kingham-predicted PFI fields  $F_{12}^K$  and  $F_{23}^K$  for the  $1 \rightarrow 2$  and  $2 \rightarrow 3$  transitions, respectively (defined as the fields at which the beam composition is 50% of each charge state).

A simple test for elements that generate singly charged atomic ions, but few in higher charge states, is to evaluate the parameter

$$t_1 = (F_2^M / F_1^M) - 1 \quad (2.67)$$

If  $t_1 > 0.5$ , then ions generated should be mainly singly charged. (If  $t_1 < 0$ , then doubly charged ions should be the majority component.) As shown in Table A.1, this test picks out the Group 1 elements Li, Na, K, Cs, and Rb (mostly too reactive for ordinary LMIS use) and the Group 13 elements Al, Ga, In, and Tl (which includes all elements currently recognized as good emitters of singly charged atomic ions). Obviously, this is mainly a periodic system effect, as the elements selected have a closed subshell and one extra electron. It also picks out Ag, Bi, and Se.

#### 2.7.1.2 Other Atomic Ion Generation Mechanisms

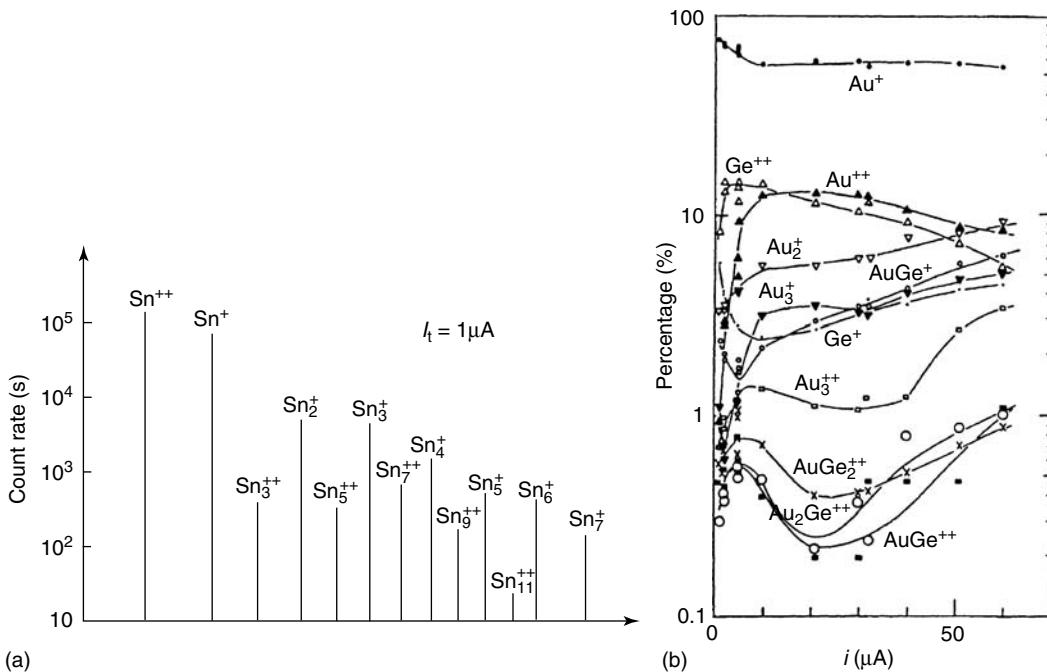
Other mechanisms that can generate atomic ions are the free-space FI of neutral atoms, collision processes involving a primary beam ion (which may create a new ion or slow down the primary ion), and fragmentation of cluster ions or charged droplets. The resulting ion may enter the ion optics if formed close to the axis between the emitter and the beam defining aperture, but would have energy significantly less than the primary beam ions.

The first two mechanisms occur (with the imaging gas) in the field ion microscope (FIM) and are well understood (see Section 3.2.2). If there are neutral ionizant atoms in space above the emitter, they will be attracted toward it by polarization forces. Some might hit low-field regions of the emitter and be reabsorbed into the liquid ionizant film; those approaching the emitting apex directly would be field ionized before reaching it.<sup>51</sup> Also, from FIM work, there is direct evidence that formation of slow ions by CE can occur all along the path of the primary beam (see Section 3.2.2). That neutral atoms do exist near an LMIS apex is demonstrated by spectroscopic analysis of light emitted from the apex region: it comes from excited neutrals.<sup>95,137</sup> One reason for requiring that LMIS liquid ionizants have low vapor pressure is to keep down the space concentration of neutrals and thereby minimize the number of slow ions that get through the beam defining aperture.

#### 2.7.1.3 Cluster Ions

Some metals also emit cluster ions. Often, when this happens, many different forms of cluster ion are emitted, as shown in Figure 2.20a for the case of tin.<sup>138</sup> Obviously, in this case there is a singly charged series and a doubly charged series; possibly, clusters such as  $\text{Sn}_4^{++}$  are also present but have not been distinguished experimentally from clusters, such as  $\text{Sn}_2^+$ , with the same charge-to-mass ratio. Broadly similar phenomena are sometimes observed in field desorption mass spectrometry.<sup>139</sup>

In general, there have been only a few investigations into the mechanisms involved in cluster ion formation.<sup>140,141</sup> One scenario is that the spectrum of observed cluster ions results from statistical nanoscale surface EHD processes in which atom clusters detach themselves and are charged because the LMIS apex is charged. But ion energy measurements show that in some cases cluster ions are generated in space above the emitter;<sup>142,143</sup> this suggests that these ions are generated by a gas-phase



**FIGURE 2.20** (a) Ion mass spectrum generated by tin needle-type liquid metal ion source (LMIS). (From Dixon, A. et al., *Phys. Rev. Lett.*, 46, 865–868, 1981.) (b) Relative abundance, as a function of source current, of species emitted by a Au–Ge alloy LMIS. (From Waugh, A.R., 28th Intern. Field Emission Symp., Portland, July 1981 (unpublished abstracts) 80–82, 1981.)

mechanism of some kind and by fragmentation of larger clusters or liquid droplets (possibly induced by ion impact). Usually, it is not clearly known what pathway is operating in any particular case.

Elements that generate a rich menagerie of ion products can give rise to fascinating high-field chemistry but cause problems for those interested in producing a well-defined beam of specific ionic composition. Filtering is required, and the generated flux of specific ions may have a complex dependence on applied voltage and measured current<sup>144</sup> (see Figure 2.20b).

### 2.7.2 ALLOY ION SOURCES

There seems to be a renewed interest in using the LMIS or LAIS as a source of specific ions.<sup>145</sup> The corresponding elemental sources are often impracticable due to high melting point, high vapor pressure, or other unsuitable behavior, and the solution has been to look for an appropriate alloy with low melting point and vapor pressure. Solutions have long been known for the main semiconductor dopants,<sup>146–148</sup> for example, the alloy B<sub>28</sub>Pt<sub>72</sub> (where the subscripts indicate the percentage atomic composition), but there has continued to be steady progress in exploring alloy-source properties and finding ways to generate additional species, for example, dysprosium.<sup>148</sup>

The simplest way of finding what has been tried (and related references) is to type “liquid metal ion source LMIS ‘element’ ‘chemical symbol’” into a Web search engine such as Google. At the time of writing, positive results were returned for the following 39 elements: Ag, Al, As, Au, B, Be, Bi, Ce, Co, Cr, Cs, Cu, Dy, Er, Ga, Ge, Hg, In, K, Li, Mg, Mn, Na, Nb, Nd, Ni, P, Pb, Pd, Pr, Pt, Rb, Sb, Si, Sm, Sn, U, Y, and Zn.

For dilute alloys, estimates of escape fields, ZQEF, and escape charge state can be made from Equation 2.8 by inserting the values of  $\Lambda^0$  and  $H_n$  for the escaping ion and the work-function for the major alloy component. Assessments as to whether PFI will occur can be made by comparing the

resulting ZQEF with the Kingham-predicted PFI fields (see Appendix A.1) for the escaping ion.<sup>149</sup> However, ordinary FEV theory (including PFI theory) was not designed to cope with complex alloy situations, so these estimates should be treated with caution.

In some cases, complex ions can be formed between different alloy components. If the major alloy component is an element prone to cluster-ion generation, then the ion menagerie generated (and the associated flux dependences on applied voltage and total emission current) can be exceedingly complicated, as illustrated in Figure 2.20b for an Au–Ge alloy.

Alloy sources seem more prone than elemental sources to needle-wetting problems (see Section 2.9). In some cases complicated metallurgical difficulties, due to needle–alloy interactions, have been found.<sup>150</sup>

## 2.8 TIME-DEPENDENT PHENOMENA

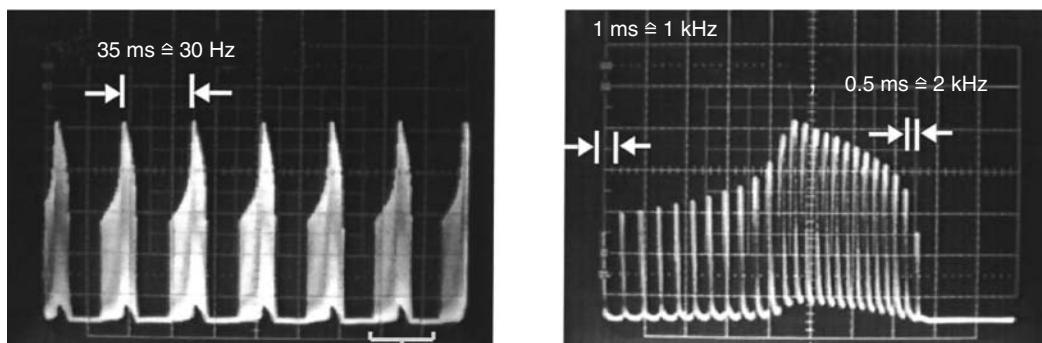
LMIS operation is affected by various time-dependent phenomena, notably droplet emission. Emission during pulsation has been noted. Droplets may also be emitted from the liquid cone and the supporting needle, and, in the high-current regime, from the jet. Droplet emission is a nuisance for a FIB machine but is of interest for metal electrospray applications, such as thin-film deposition.<sup>21</sup>

### 2.8.1 PULSATION

Pulsation involves collapse and reformation of the jet and maybe the cone, and is often associated with emission of a charged droplet: this droplet shields the liquid behind it and thus temporarily reduces the Maxwell stress acting. Oscillation of the liquid cone without droplet emission also seems to be possible.

With nonmetallic liquids, pulsation is a near-universal phenomenon, and is recognized as a mode/regime of EHD spraying.<sup>32</sup> Pulsation can occur with a capillary-based LMIS.<sup>94,151</sup> With needle-based LMISs, Wagner and Hall<sup>15</sup> established, from oscilloscope measurements on gold sources, that pulsation occurs for sources with significant flow impedance but not for a low-drag LMIS. Beckman<sup>37</sup> established the same result for gallium sources.

Combined oscilloscope and stroboscopic measurements on the pulsation of water cone-jets driven by a syringe pump, by Juraschek and Röllgen,<sup>152</sup> provide useful insight. They identified two pulsation regimes. In the lower-voltage regime, the current exhibited two frequencies (see Figure 2.21). The higher frequency corresponds to tear off and reformation of the jet, accompanied by oscillation of the cone. After a finite number of events of this kind (perhaps 10), the reservoir of available liquid in the cone is depleted. It then re-forms as water that is driven into it by the syringe



**FIGURE 2.21** Capillary current versus time oscillograms obtained from an H<sub>2</sub>O/MeOH mixture. Each of the complex pulses in the left-hand diagram is a group containing 25 higher-frequency pulses as shown in the right-hand diagram. (From Juraschek, R. and Röllgen, F.W., *Int. J. Mass Spectrom. Ion Process.*, 177, 1–15, 1998.)

pump: the lower pulse frequency corresponds to the cycle of cone diminution and reformation. Emission current is generated only while both cone and jet are present (though not by FEV); the observed current is an average derived from the on-periods of each cycle. In the higher-voltage pulsation regime, only the higher-frequency current component is present, and it was found that only the jet and the top part of the cone were collapsing and reforming. The partial similarity with the Beckman observations on liquid metals is obvious, and the remaining observations provide material for thought.

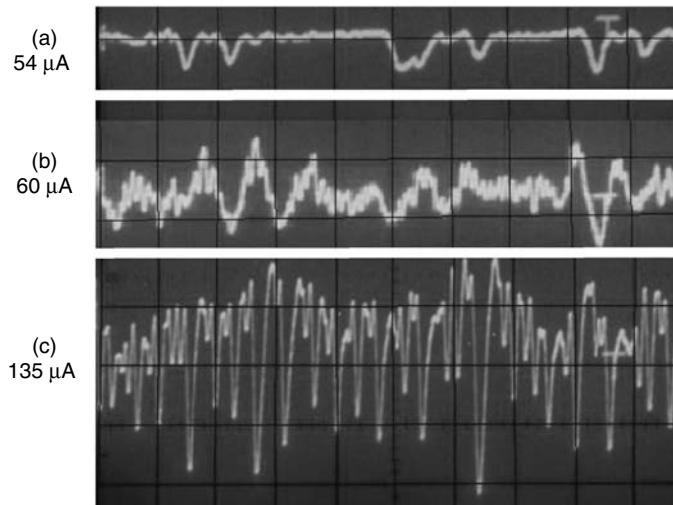
### 2.8.2 DROPLET EMISSION IN THE UPPER UNSTEADY REGIME

Significant levels of noise occur at higher LMIS emission currents. Figure 2.22 illustrates this for gallium.<sup>153–155</sup> noise onset occurs at a critical current  $i_u$  of  $\sim 25 \mu\text{A}$ . This coincides with the upper bend in the S-type logarithmic  $\Delta E$  versus  $i$  characteristic. While noise could be partly associated with jet vibration, droplet emission due to EHD instability is the more likely cause.<sup>156</sup> As with pulsation, the droplet would electrically shield the residual jet and cause ion emission to cease temporarily. Vladimirov and colleagues<sup>157</sup> attempted detailed theoretical analysis of jet breakup, assuming that Raleigh and Faraday instabilities may occur.

Mair<sup>44</sup> argued that noise onset is determined by the timescales for different EHD processes in the jet. Using a simple model, he found a formula for the lower boundary  $i_u$  of the upper unsteady regime:

$$i_u \sim 13.4\pi\rho^{1/2}\gamma^2(ze/m)/(\epsilon_0^{3/2}/F_a^3) \quad (2.68)$$

Values for selected elements, assuming  $z = 1$  and putting  $F_a = F_1^M$ , are shown in Table A.2. The predicted  $i_u$  values are very sensitive to choice of  $F_a$  but show an important qualitative trend: materials that do not emit cluster ions have predicted  $i_u$  values well above observed minimum emission currents; materials known to emit cluster ions, in particular gold, tin, and bismuth, have predicted  $i_u$  values comparable with minimum emission currents.



**FIGURE 2.22** Sequence of emission-current oscillograms for a gallium liquid metal ion source operated at room temperature at the emission currents shown. The horizontal scale is  $1 \mu\text{s}/\text{division}$ . At low currents the emission is relatively stable; low-frequency transients begin to appear at  $\sim 25 \mu\text{A}$ ; the high-current behavior is as shown here, with the onset of higher-frequency transients at  $\sim 60 \mu\text{A}$ , and then gradual diminution of the lower-frequency transients as emission current increases.

These materials (Au, Sn, and Bi) have no intermediate (power law 0.7) region in their  $\Delta E$  versus  $i$  characteristics. They seem to go directly from a pulsation regime to the upper unsteady regime where the  $\Delta E$  versus  $i$  power-law dependence is  $\sim 0.3$ . Both regimes involve EHD unsteadiness. It is plausible that such materials should exhibit EHD-driven *cluster shake-off* at low emission currents. Thus, formula 2.68 may be a helpful criterion for deciding whether a material will emit cluster ions.

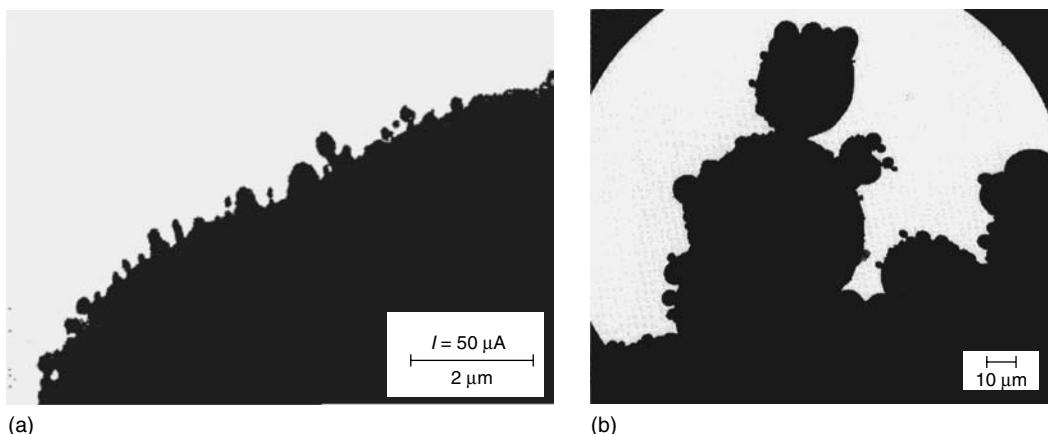
Figure 2.22 shows that for gallium the upper unsteady regime seems to comprise two sub-regimes associated with the predominance of low-frequency (up to  $\sim 60 \mu\text{A}$ ) and high-frequency (above  $\sim 60 \mu\text{A}$ ) pulses, respectively. Comparison with Juraschek and Rollgen's<sup>152</sup> results makes it tempting (but possibly premature) to interpret the  $60 \mu\text{A}$  current level as the boundary between two different droplet-formation subregimes involving the tearoff of different amounts of the cone-jet.

### 2.8.3 GLOBULE EMISSION FROM THE NEEDLE AND CONE

In some LMIS situations, globules are detached from the liquid film on the supporting needle and the liquid cone.<sup>103,158</sup> Video recordings of EM images of operating sources<sup>159</sup> show that, when it exists, the phenomenon is pervasive. Hundreds of sites may be emitting globules at any one time, and the rate of emission increases sturdily with emission current (see Figure 2.23a). The globule size before detachment appears to diminish as one moves toward the higher-field regions near the apex of the needle/cone. Still images from such sequences show that self-similar (quasi-fractal) growth processes occur<sup>103</sup> as shown in Figure 2.23b.

The origin of this quasi-fractal effect has been mysterious. We now think that, on each observed scale, protrusion growth is initiated by a nanoscale (near-atomic level) disruptive event. After this, a high field at the protrusion apex (due to field enhancement) creates low hydrostatic pressure that pulls liquid in, and the usual *balling up* influence of surface tension then operates. The neck joining the globule to the emitter can remain or become very small (see Figure 2.23).

As regards initiation, if a liquid nanoproltrusion on a charged surface is sufficiently high and pointed then it will grow. However, an activation-energy barrier ( $\Delta G$  barrier) exists for its formation. Rough calculations suggest that single-atom displacements will not be sufficient but that displacements that generate a three-atom group sitting on a seven-atom group on top of the normal liquid surface (or something similar) might be sufficient. Analogies are that near-atomically-sharp field emitters can be produced by *in situ* electroformation,<sup>160</sup> and that small nanocones form on very hot solid metallic emitters.<sup>161,162</sup>



**FIGURE 2.23** Liquid indium globules on the sides of the needle apex, observed by high-voltage electron microscopy. (a) Profile related to video recording<sup>159</sup> that shows very active formation and detachment of globules (contrast artificially enhanced). (b) Liquid shape assumed to result from dynamic self-similar processes. (From Praprotnik, B. et al., *Surf. Sci.*, 314, 353–364, 1994.)

The problem is to identify the nature of the disruptive events. There seem to be three possibilities: ordinary statistical fluctuations in a hot liquid, protrusions thrown up by the flow of liquid over a rough surface<sup>163</sup> (though this seems less applicable to the self-similar structures), and the impact of negatively charged droplets or cluster ions.

The relevance of globule emission is as follows. Subsequent fragmentation (e.g., on impact with a counterelectrode) may release neutral atoms that eventually become low-energy ions in the beam. Conceivably, the impact may create a rebounding negatively charged droplet that causes a disruptive event at the emitter. Also, this emission wastes ionizant. It is known to occur with indium,<sup>103</sup> a favored FEEP ionizant,<sup>19</sup> so implications exist for satellite thruster lifetime.

More generally, the increasing incidence of droplet and globule emission as  $i$  increases probably accounts for the reducing mass fraction of atoms emitted as ions.<sup>164</sup> For a more complete discussion, see Ref. 1.

#### 2.8.4 SOURCE TURN-ON

When the applied voltage  $V$  is increased beyond the onset value  $V_{\text{on}}$ , a GG cone-jet begins to form. Thompson and Prewett<sup>165</sup> developed a simple classical model. If viscous-drag effects in the liquid supply are neglected, this predicts turn-on times (delay between application of voltage  $V_{\text{app}}$  and emission-current turn-on) of order 1  $\mu\text{s}$  with the exact value depending on the degree of overvoltage  $O_V$  ( $\equiv V_{\text{app}}/V_{\text{on}}$ ). Their sources had practical turn-on times between 50 ms ( $O_V = 1.1$ ) and 500 ms ( $O_V = 1.0$ ). A revised model,<sup>165</sup> incorporating viscous-drag effects, predicted times of order milliseconds.

Experimental current-time characteristics depend on the degree of overvoltage: there is both an initial *overshoot and decay-back* effect, as observed by Wagner and Hall,<sup>15</sup> and (for higher final current) a slower *growth to a steady-level* effect. This, and the repeated pulses seen by Wagner and Hall, show that complicated EHD effects take place during source turn-on, particularly if  $O_V > 1$  and viscous drag exists in the liquid supply. Broadly similar effects occur with ILISs.<sup>22</sup>

Putting all things together<sup>11,15,152,165–168</sup> yields the following hypothesis about source turn-on. We start from a stable situation with a thin liquid film on the rounded needle apex. Initially, as voltage is increased, the liquid may be able to remain stable by a small shape adjustment that leads to a thicker film at the needle apex (and smaller liquid-apex radius). When the total *outward* force on the body of liquid becomes so great that it cannot be balanced by opposing forces (and when pressure differences develop within it),<sup>\*</sup> a liquid protrusion begins to grow. Details of the liquid-apex motion are not clear (and may vary depending on the strength of viscous-drag effects in the liquid supply), but the back-end of the liquid body grows to *fill out* a Taylor-cone-like shape. As this happens, the liquid near the apex gains increasing momentum and eventually *overshoots* the Taylor-cone shape. Initially, the protrusion continues to grow. However, a new form of instability rapidly develops at its top, and a droplet is emitted that carries away mass and (more important) momentum. The liquid protrusion then reorganizes itself toward the shape of the *jet* part of a GG cone-jet (possibly with the emission of one or more smaller droplets). For small degrees of overvoltage at growth onset, the length of the resulting jet is greater than the steady-state length for the applied voltage, and the jet shrinks back as a result of ion emission. At the same time, the *cone* part continues to grow toward the shape of the steady-state GG cone-jet (if this has not already been achieved). The shape and length of the jet part may adjust further as this happens. Finally, the steady-state shape is achieved.

The physical thinking here is that the quasi-steady-state GG cone-jet is the final attractor but that in the initial unsteady motion the EHD of the upper part of the liquid body (jet plus top of the cone) can partially decouple from the EHD of the back-end of the cone (which is more affected by viscous drag in the liquid supply). In the unsteady motion, issues of liquid supply to the cone and from the cone to the jet need to be considered separately.

---

\* Contrary to what is often stated, the onset condition is *not* exactly  $\Delta p < 0$ —see Section 2.4.2.

Details of behavior may vary with the degree of overvolting, the time dependence of the applied voltage increase, and the extent of viscous-drag effects. There seems to be scope for a new round of experiments, using modern electronic equipment and based on better knowledge of source preparation. Greater understanding of the onset process would probably add to our understanding of pulsation and source extinction and might help source development.

### 2.8.5 NUMERICAL MODELING OF LIQUID-SHAPE DEVELOPMENT

Numerical modeling might aim to explain the effects just described and establish the steady low-current shape, but progress has been limited. Early modeling<sup>166</sup> confirmed that in an electric field droplets are expected to first elongate into a roughly ellipsoidal shape and then develop pointed ends as seen experimentally. However, LMIS modeling met a technical problem: shape must be specified precisely after each hydrodynamic time step; if the method used generates an unphysical kink or protrusion, then subsequent time-steps amplify this unphysical feature.<sup>24</sup>

Cui and Tong<sup>167</sup> had overcome this problem by repeated manual interventions (Cui, private communication to RGF). It can also be overcome by a method<sup>168,169</sup> involving repeated special transformations of coordinates (but this is very computer intensive) and by the so-called level-set method recently used<sup>170</sup> in the context of FEEP. Alternatively, the simplified approach of Higuera<sup>108</sup> can be adopted.

If source turn-on involves liquid overshoot and droplet emission, it seems unlikely that dynamic modeling will help to establish LMIS low-current shape. It may be best to return to steady-state modeling with a wider range of apex-shape models.

In principle, dynamic modeling could be applied to droplet emission in LMIS operation, as it has been in FEEP.<sup>170</sup> But, if sometimes the critical initiating feature is some statistical atomic effect or is due to external impact or some other non-EHD factor, then it may be more important to first understand the initiation process.

## 2.9 TECHNOLOGICAL ISSUES

### 2.9.1 FABRICATION OF THE NORMAL LMIS

Fabrication details and general technological requirements for the normal LMIS are described in Refs 1 and 2, and Ref. 25 often identifies useful original articles. This section draws attention to the main issues and briefly describes some typical processes (but not all variants).

#### 2.9.1.1 Needle Fabrication

Tungsten needles are usually prepared and roughened by electrolytic etching in KOH or NaOH solution,<sup>15</sup> although mechanical grinding and FIB-type milling are options in specialist source work. Etching with AC current was found to produce longitudinal grooves along the axis of the needle (see Figure 2.24). Clever grinding can produce spiral grooves near the needle apex. The groove size controls the impedance the needle presents to flow. After etching, the source assembly is ultrasonically cleaned using deionized water and hydrocarbon solvents. Alternatives to tungsten are needed when the liquid ionizant interacts with tungsten. For example, a titanium-coated graphite needle was used for an Al source<sup>171</sup> and a silicon-coated graphite needle for a Pd<sub>72</sub>B<sub>28</sub> source.<sup>172</sup>

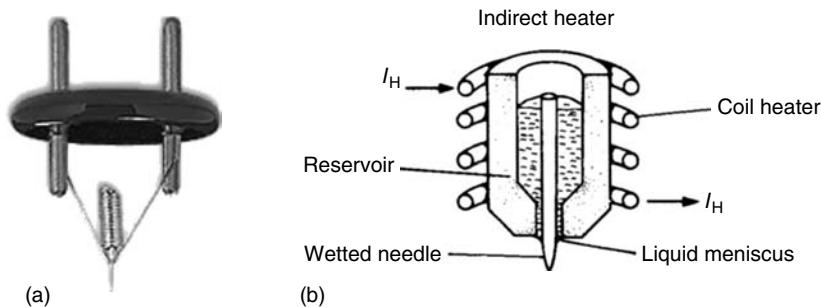
#### 2.9.1.2 Needle Wetting

Vacuum wetting is the most critical stage of source fabrication. A poorly wetted source with a tenuous liquid metal coating may work for a few hours, but its performance soon becomes erratic, leading to eventual failure. Under clean ultrahigh vacuum conditions, a properly wetted, correctly



**FIGURE 2.24** Electron micrograph showing grooves on the surface of an electrochemically etched tungsten needle.

grooved needle will emit steadily for many hundreds of hours (though its performance may need to be *renovated* by heating every 40 h or so). There has to be chemical compatibility between the ionizant and the needle that leads to wetting but not to electrochemical attack on the needle. Prior to wetting, the needle and reservoir assembly are vacuum outgassed to remove residual surface contaminants. Electron bombardment to white heat is one approach. The assembly is then dipped into a precleaned boat of molten metal. Other details are needle or ionizant specific.



**FIGURE 2.25** Two forms of LMIS design: (a) the directly heated FEI Co. design (illustration reproduced with permission) and (b) indirectly heated Culham-type design. (From Prewett, P.D. and Kellogg, E.M., *Nucl. Instrum. Meth. Phys. Res. B*, 6, 135–142, 1985.)

### 2.9.1.3 Reservoir and Heating Arrangements

Heating for the normal LMIS can be direct or indirect. For direct heating, current is passed through the wetted structure. Indirect heating involves a separate (unwetted) heating filament. The most popular design is the tungsten *hairpin*, originally used in field electron microscopy and then applied for LMIS work at the Oregon Graduate Center.<sup>126</sup> The needle is spot-welded to a heating loop. The FEI-type source\* (Figure 2.25a) is a development from this. This design is widely used for gallium and other ionizants compatible with tungsten. The simplicity and ease of manufacture of directly heated sources make them widely used with considerable success in many applications.

Indirectly heated LMIS designs have physical separation between needle and heater. Figure 2.25b illustrates the original Culham design,<sup>173</sup> which allowed a wide choice of needle, reservoir, and heater materials but was significantly more expensive to manufacture. Special designs have to be employed for air-reactive metals such as cesium.<sup>173</sup>

## 2.9.2 SOURCE OPERATION

### 2.9.2.1 General Conditions

Reliable LMIS operation requires good vacuum conditions and care to avoid source contamination. For example, gallium tends to oxidize in air. It is best to operate an LMIS in ultrahigh vacuum ( $<\sim 10^{-8}$  torr at the source). Most sources do not require fully bakeable systems, but these are essential when dealing with highly reactive metals (such as Cs and Li) that combine with water vapor.

The operating environment should be designed with attention to the LMIS metallurgical properties. Since some sputtering inevitably occurs when the extracted beam hits surrounding electrodes such as aperture-defining plates,<sup>174</sup> materials should (ideally) be chosen such that side-effects due to the resulting source contamination are kept to a minimum.<sup>175</sup> As shown clearly by Galovich,<sup>174</sup> source contamination can lead to various difficulties, including increase in melting temperature, increase in hydrodynamic flow resistance, and (especially with alloys) chemical and metallurgical effects that may, for example, change the composition of the emitted beam and alter the physical structure of the support needle. All these effects tend to reduce the source working lifetime.

In some cases where film contamination has occurred, normal operation can be restored by heating or by overvolting the source for a period of several seconds: in the latter case the source emits a spray of droplets, thereby removing the contaminated film and replacing it with clean liquid from the reservoir.

---

\* As described in Web page, <http://www.fei.com/products/components/electron-ion-sources.aspx>.

**TABLE 2.1****Alternative Liquid Metal Ion Source Geometries Considered and Investigated**

Capillary tube	Higher currents (up to 100 $\mu\text{A}$ or more) but difficult to control at low currents. Has two extra high-voltage emission regimes, an <i>exponential regime</i> , and a <i>saturation regime</i> . May operate in multicusp mode at high currents. Under exploration for field emission electric propulsion. <sup>19,20,94,151,164,179</sup>
Sharp needle	Lower currents than normal LMIS.
Atomic size ion source (ASIS)	Feasibility of achieving emission from an atomically sharp emitter supplied by thin-film flow along needle surface has been demonstrated. No liquid cone involved. <sup>180</sup>
Impregnated needle	Liquid supply is through porous needle. Did not succeed commercially. <sup>181</sup>
Linear wire array	Investigated for spacecraft applications. Less effective than linear slit emitter. <sup>79</sup>
Linear slit	Investigated for spacecraft applications. Emission takes place from a regularly spaced series of cusps along the slit length. Gives current of 10 mA or more. <sup>182</sup>
Flat liquid plane	Investigated as possible <i>repeated one-shot</i> high-current-density emitter. Aim was to generate emission simultaneously from large 2-D array of cusps. Did not work in easily controllable fashion. <sup>183–186</sup>
Miniaturized needle array/volcano array	Based on arrays of miniaturized volcanos or needles, built using the technology originally developed for Spindt-type electron emitter arrays. Structures have been fabricated and tested with nonmetallic liquids that yield ion emission; liquid metals have not been tested. <sup>187–189</sup>

In general, an LMIS should be operated at a temperature as close above its melting point as practicable. Using higher temperatures increases the energy spread and increases the thermal evaporation rate for neutral atoms, both of which tend to degrade machine performance.

### 2.9.2.2 Secondary Electrons

The impact of the primary ion beam onto the defining aperture or other counterelectrode generates secondary electrons that may travel back toward the emitter. On striking the supporting structure, they will cause heating and excite surface atoms that subsequently emit light and X-rays. Also, the presence of a secondary-electron component causes error in the measurement of ion emission current. Other than this, there seems no evidence that secondary electrons affect normal LMIS operation at relatively low emission currents, although this has been disputed.<sup>176–178</sup>

### 2.9.3 ALTERNATIVE LMIS GEOMETRIES

During LMIS development, various geometrical options were explored for various applications. Table 2.1 lists some geometrical options, with brief comments and selected references. The development aim was usually (1) high-brightness single-point emitters for FIB-type applications, (2) high-current emitters (e.g., for FEEP), or (3) large-area high current-density sources. There are also potential applications (e.g., parallel-write lithography) that need an array of bright single-point emitters. For all except the single-point applications, a central problem is uniformity of emission from the individual emission sites; our view is that most multisite applications will require external control for each site. It is no surprise that the most successful application (the FIB machine) involves only one single-point emitter and is in a context where (in comparison with most alternative technologies) the high optical brightness of the LMIS gives it special advantages.

## 2.10 CONCLUDING REMARKS

As this chapter shows, the so-called LMIS is a complex EHD device. Recent years have seen slow but useful progress in establishing its physics, but overall understanding is still incomplete. Part of the aim here has been to set down the underlying science that is understood outside the LMIS community but not well understood within it.

An LMIS goes through a sequence of different operating regimes as the applied voltage increases, and the regimes exhibited may depend both on the metal being used and on the state of the support needle. Table 2.2 exhibits the known regimes for needle-based gallium LMISs. (A capillary LMIS may have a different behavior pattern with extra high-current regimes.<sup>151</sup>) There has been useful progress toward developing this *regime map*, but we need greater understanding of how the boundaries arise.

Not surprisingly, the physics (and ion optics) of the most-steady regime (where the LMIS is operated when in a FIB) has been most studied and is best understood. For this regime we begin to have a coherent scientific picture. The role of space-charge immediately above the LMIS apex is crucial, and we understand how this affects current–voltage and length–current characteristics, the energy spread of the beam, and virtual source diameter. The most important questions remaining are the related issues of the detailed nature of the stabilizing mechanism that allows an LMIS to operate quasi-steadily, the shape of the emitter at low emission currents, and an exact theory of extinction current. Better detailed understanding of why and how some metals generate clusters is also needed.

For other applications, for example, metal electrospraying and FEEP, it would be useful to develop improved characterization and understanding of the higher-current emission regimes, where droplet emission becomes a dominant process.

Notwithstanding past gaps in basic understanding, the Ga LMIS and the related FIB machine have been successfully developed technologically. Table 2.3 illustrates current FIB properties. The chief source-derived limitation is the LMIS large virtual source size.<sup>190</sup> This can be reduced only if we can reduce the incidence of space-charge effects, and there has been some thinking that it might be possible to get the LMIS to operate quasi-steadily at slightly lower currents. The existence of the

---

**TABLE 2.2**  
**Regime Map: Emission Regimes Exhibited by Needle-Based Gallium Liquid Metal Ion Sources (LMISs)**

Current Range ( $\mu\text{A}$ )	Low Drag	Not Low Drag
>25	<i>Upper unsteady regime</i> (May comprise two subregimes with higher subregime >60 $\mu\text{A}$ )	
2–25	<i>Most-steady regime</i>	
0.45–2	<i>Intermediate regime</i>	<i>Pulsation regime</i>
<0.45	<i>Unstable—collapses</i>	(Might comprise two subregimes?)

---

**TABLE 2.3**  
**Typical Optical Properties of a Gallium Liquid Metal Ion Source and Focused Ion Beam System**

Property	Value	Property	Value
Virtual source radius	25 nm	Angular current density	10 $\mu\text{A}/\text{sr}$
Emission current	2 $\mu\text{A}$	Brightness	$10^{10} \text{ A m}^{-2} \text{ sr}^{-1}$
Energy spread	5 eV	Reduced brightness	$10^6 \text{ A m}^{-2} \text{ sr}^{-1} \text{ V}^{-1}$
Extraction voltage	10 kV	Best spot size	6 nm

*Source:* Illustrative data are from Orloff, J. et al. in *High Resolution Focused Ion Beams: FIB and Its Applications*, Kluwer, New York, 2003. Predicted spot size is from Hawkes, P.W. and Lençova, B. *E-nano-newsletter*, Issue 6, 2006 <[www.phantomsnet.net/Foundation/newsletters.php](http://www.phantomsnet.net/Foundation/newsletters.php)>.

---

extinction phenomenon will prevent major progress, but small improvements might be possible, particularly if we understood extinction in more detail.

For the longer-term future of point-ion-source development, it would seem that for FIB applications, we should explore the merits of other types of point ion emitter, and for the LMIS we should look to its other possible applications, in particular, its uses as a source of specific ions and clusters in controlled metal electrospraying and for FEEP. Also, ideas used in LMIS theory might contribute to better understanding of ILISs and (possibly) other forms of EHD ion source. There seems to be much scope for further attention to the underlying high-field physics. The scientific work started by Gray in 1732 remains unfinished.

## ACKNOWLEDGMENT

This review has used various published illustrations and diagrams. We want to record our thanks to the Institute of Physics, the American Institute of Physics, the American Vacuum Society, Elsevier Scientific Publishers, FEI Co., North-Holland Publishing Co., and Springer-Verlag GmbH.

## A.1 APPENDIX: NUMERICAL DATA

This appendix provides numerical data relating to theories of FEV and LMIS operation. Table A.1 contains data relating to ZQEF and PFI. The first eight columns include the following chemical thermodynamic data for the elements shown:  $Z_p$ —proton number; Gp—chemical group (modern classification);  $\phi$ —typical local work-function;  $\Lambda^0$ —zero-field atomic bonding energy, taken as given by the enthalpy of atomization;  $I_s$ —the  $s$ th ionization energy. Data are provided for all the metal elements normally considered in solid FEV and for all specialist species used so far in LMIS development. For the former, the data are those used by Tsong<sup>191</sup> and Miller et al.;<sup>55</sup> for the added elements the data are mostly taken from standard references—Ref. 192 for  $\Lambda^0$ ; Ref. 193 for  $\phi$  and the  $I_s$  values. Exceptions: the  $\phi$  values for Pr, Dy, and Er are taken from data in Ref. 194;  $\phi$  for P has been guessed; and  $I_2$  and  $I_3$  for U are taken from a Web site.\*

Columns 10, 12, and 14 show the Müller escape fields for escape charge states  $n = 1, 2, 3$ , calculated from Equation 2.3. Columns 11, 13, and 15 show the “50% PFI fields” as calculated by Kingham:<sup>53</sup> thus  $F_{12}^K$  is the field at which there is 50% chance that an ion initially evaporated as a single-charged ion will be post-field ionized to a doubly charged ion, etc. Columns 16 and 17 show the predicted escape charge state  $n_d$ , and the parameter  $t_1$  defined by Equation 2.67; negative  $t_1$  values have been suppressed. Large positive  $t_1$  values indicate species likely to evaporate as mainly singly charged ions.

The PFI field data can be used as follows. From Brandon’s criterion, the predicted escape field and charge state are determined by the lowest of the alternative escape fields. If the PFI field to the immediate right of this is close to or lower than the escape field, then significant PFI into the next higher charge state is expected. Lead (Pb) is an example of this.

As noted earlier, estimates based on Equation 2.8 are of limited accuracy and may be systematically low. Kingham’s calculations are better but are also limited by the theoretical approximations he uses. Conclusions drawn from this table should be regarded as indicative rather than definitive, and anomalies may exist.

Table A.2 shows quantities, relevant to LMIS operation, calculated from equations given earlier. Rows 2–10 provide basic chemical information:<sup>195</sup>  $m_r$ —relative atomic mass;  $m$ —atomic mass;  $\rho$ —density;  $\omega$ —atomic volume;  $\gamma$ —surface free energy per unit area in zero electric field;  $T$ —assumed operating temperature (taken as the melting point for all species except mercury for which  $T$  is put equal to 300 K). Row 11 shows the assumed (mean) charge state for the emitted ions. The remaining rows show the quantities defined earlier, by the equations shown.

---

\* From <http://www.gordonengland.co.uk/elements/u.htm>, January 2008.

**TABLE A.1**  
**Thermodynamic Properties of Selected Elements and Predicted Field Evaporation Behavior**

<b>Z<sub>p</sub></b>	<b>E<sub>I</sub></b>	<b>G<sub>p</sub></b>	<b>φ (eV)</b>	<b>Λ<sup>0</sup> (eV)</b>	<b>I<sub>1</sub> (eV)</b>	<b>I<sub>2</sub> (eV)</b>	<b>I<sub>3</sub> (eV)</b>	<b>F<sub>12</sub><sup>M</sup> (V/nm)</b>	<b>F<sub>23</sub><sup>K</sup> (V/nm)</b>	<b>F<sub>3</sub><sup>M</sup> (V/nm)</b>	<b>F<sub>34</sub><sup>K</sup> (V/nm)</b>	<b>n<sub>d</sub></b>	<b>t<sub>i</sub></b>	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	17
3	Li	1	2.50	1.65	5.39	75.64	122.45	Li	14	24	524	>50	1005	—
4	Be	2	3.90	3.33	9.32	18.21	153.89	Be	53	—	46	—	770	2
5	B	13	4.60	5.92	8.30	25.15	37.90	B	64	—	79	—	104	1
6	C	14	4.80	7.40	11.26	24.38	47.89	C	133	—	97	—	151	0.2
11	Na	1	2.30	1.13	5.14	47.29	71.64	Na	11	—	208	—	360	—
12	Mg	2	3.70	1.53	7.65	15.04	80.14	Mg	21	17	25	>40	224	—
13	Al	13	4.10	3.34	5.99	18.83	28.45	Al	19	26	35	42	50	1
14	Si	14	4.20	4.63	8.15	16.34	33.46	Si	51	19	37	>40	64	0.8
15	P	15	4.50	3.92	10.49	19.77	30.29	P	68	—	55	—	67	—
19	K	1	2.20	0.94	4.34	31.63	45.72	K	7	75	92	108	149	18
20	Ca	2	2.70	1.83	6.11	11.87	50.91	Ca	19	11	18	132	101	—
22	Ti	4	4.00	4.86	6.82	13.58	27.49	Ti	41	13	26	40	43	—
23	V	5	4.10	5.30	6.74	14.65	29.31	V	44	18	30	42	49	—
24	Cr	6	4.60	4.10	6.77	16.50	30.96	Cr	27	20	29	51	51	—
25	Mn	7	3.80	2.98	7.44	15.64	33.67	Mn	30	19	30	60	60	1
26	Fe	8	4.40	4.29	7.90	16.16	30.65	Fe	42	19	33	49	54	0.0
27	Co	9	4.40	4.39	7.86	17.06	33.50	Co	43	22	37	58	63	—
28	Ni	10	5.00	4.44	7.64	18.17	35.17	Ni	35	24	36	64	65	0.0
29	Cu	11	4.60	3.50	7.73	20.29	36.83	Cu	30	30	43	71	77	—
30	Zn	12	3.80	1.35	9.39	17.96	39.72	Zn	33	22	39	81	84	—
31	Ga	13	4.10	2.78	6.00	20.51	30.71	Ga	15	31	39	49	59	0.2
32	Ge	14	4.80	3.98	7.88	15.93	34.22	Ge	35	—	29	—	58	—
33	As	15	4.70	3.00	9.81	18.63	28.35	As	46	26	42	42	54	—
34	Se	16	5.90	2.35	9.75	21.19	30.82	Se	27	—	40	—	55	2
37	Rb	1	2.10	0.86	4.18	27.28	40	Rb	6	59	69	87	112	1
38	Sr	2	2.59	1.70	5.69	11.03	42.89	Sr	16	—	15	—	74	—
39	Y	3	3.10	4.40	6.38	12.24	20.50	Y	41	—	25	—	30	—
40	Zr	4	4.20	6.32	6.84	13.13	22.99	Zr	56	17	28	29	35	—
41	Nb	5	4.00	7.47	6.88	14.32	25.04	Nb	74	17	37	34	45	—
42	Mo	6	4.20	6.81	7.10	16.15	27.16	Mo	65	20	41	40	51	—

(continued)

TABLE A.1 (Continued)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Z <sub>p</sub>	E <sub>1</sub>	G <sub>p</sub>	ϕ (eV)	Λ <sup>0</sup> (eV)	I <sub>1</sub> (eV)	I <sub>2</sub> (eV)	I <sub>3</sub> (eV)	E <sub>1</sub>	F <sub>1<sup>M</sup></sub> (V/nm)	F <sub>12<sup>K</sup></sub> (V/nm)	F <sub>2<sup>M</sup></sub> (V/nm)	F <sub>23<sup>K</sup></sub> (V/nm)	F <sub>3<sup>M</sup></sub> (V/nm)	F <sub>34<sup>K</sup></sub> (V/nm)	n <sub>d</sub>	t <sub>1</sub>
44	Ru	8	4.50	6.62	7.37	16.76	28.47	Ru	62	22	41	43	54	—	2	—
45	Rh	9	4.60	5.75	7.46	18.08	31.06	Rh	52	25	42	44	61	—	2	—
46	Pd	10	5.00	3.94	8.34	19.43	32.92	Pd	37	28	41	58	63	—	1	0.1
47	Ag	11	4.60	2.96	7.58	21.49	34.83	Ag	24	35	45	65	72	—	1	0.8
48	Cd	12	4.10	1.16	8.99	16.91	37.48	Cd	25	22	31	75	70	—	1	0.2
49	In	13	4.00	2.60	5.79	18.87	28.03	In	13	28	32	42	48	—	1	1.4
50	Sn	14	4.40	3.12	7.34	14.63	30.50	Sn	26	17	23	50	46	—	2	—
51	Sb	15	4.60	2.70	8.64	16.53	25.30	Sb	32	21	30	35	40	—	2	0.0
55	Cs	1	2.10	0.83	3.89	25.10	35	Cs	5	55	57	71	88	—	1	11
56	Ba	2	2.50	1.86	5.21	10.00	35.13	Ba	15	8	13	72	51	—	2	—
57	La	3	3.30	4.49	5.58	11.06	19.18	La	32	10	18	21	24	—	2	—
58	Ce	Lan	2.90	4.38	5.54	10.85	20.20	Ce	34	—	19	—	27	—	2	—
59	Pr	Lan	2.96	3.69	5.46	10.55	21.62	Pr	27	—	16	—	27	—	2	—
60	Nd	Lan	3.20	3.40	5.53	10.73	22.10	Nd	23	—	15	—	27	—	2	—
62	Sm	Lan	2.70	2.14	5.64	11.07	23.40	Sm	18	—	16	—	30	—	2	—
66	Dy	Lan	3.25	3.01	5.94	11.67	22.80	Dy	23	—	17	—	29	—	2	—
68	Er	Lan	3.25	3.29	6.11	11.93	22.74	Er	26	—	19	—	30	—	2	—
72	Hf	4	3.50	6.35	7.00	14.90	23.30	Hf	67	19	39	31	43	48	2	—
73	Ta	5	4.20	8.09	7.89	16	22	Ta	96	21	48	27	44	47	2	—
74	W	6	4.50	8.66	7.98	18	24	W	102	28	57	32	52	53	3	—
75	Re	7	5.10	8.10	7.88	17	26	Re	82	22	45	38	49	62	2	—
76	Os	8	4.60	7.00	8.70	17	25	Os	86	23	48	35	50	75	2	—
77	Ir	9	5.30	6.93	9.10	17	27	Ir	80	22	44	41	50	64	2	—
78	Pt	10	5.30	5.85	9.00	18.56	28	Pt	63	27	45	44	53	—	2	—
79	Au	11	4.30	3.78	9.23	20.50	30	Au	53	34	54	51	66	—	1	0.0
80	Hg	12	4.50	0.69	10.44	18.76	34	Hg	31	28	38	66	66	—	1	0.2
81	Tl	13	3.70	1.87	6.11	20.43	29.83	Tl	13	34	38	50	57	—	1	2.0
82	Pb	14	4.10	2.04	7.42	15.03	31.94	Pb	20	18	23	57	50	—	1	0.2
83	Bi	15	4.30	2.15	7.29	16.69	25.56	Bi	18	23	27	37	39	—	1	0.5
90	Th	Act	3.40	6.20	6.31	11.50	20	Th	58	—	26	—	29	—	2	—
92	U	Act	3.63	5.55	6.19	14.72	19.10	U	46	—	32	—	31	—	3	—

See text for definitions of symbols.

**TABLE A.2**  
**Liquid Metal Ion Source Properties for the Liquid Metals of Most Interest**

1	Equation	Ok?	Quantity	Units	Al <sup>+</sup>	Ga <sup>+</sup>	Ag <sup>+</sup>	In <sup>+</sup>	Sn <sup>++</sup>	Cs <sup>+</sup>	Au <sup>+</sup>	Hg <sup>+</sup>	Pb <sup>+</sup>	Bi <sup>+</sup>	
2		2	Proton number	—	13	31	47	49	50	55	79	80	82	83	
3		3	Group	—	13	13	11	13	14	1	11	12	14	15	
4		4	$m_i$	—	26.98	69.72	107.87	114.82	118.71	132.90	196.97	200.59	207.19	208.98	
5		5	$M$	$\times 10^{-25} \text{ kg}$	0.448	1.158	1.791	1.907	1.971	2.207	3.271	3.331	3.440	3.470	
6		6	$\rho$	$\text{kg/m}^3$	2385	6100	9330	7030	6980	1840	17360	13550	10678	10050	
7		7	$\omega$	$\times 10^{-2} \text{ nm}^3$	1.878	1.898	1.920	2.712	2.824	11.994	1.884	2.458	3.222	3.453	
8		8	$\omega^{1/3}$	nm	0.266	0.267	0.268	0.300	0.305	0.493	0.266	0.291	0.318	0.326	
9		9	$\gamma$	mN/m	914	718	966	556	560	70	1169	486	458	378	
10		10	T	K	933	303	1234	430	505	302	1336	300	600	544	
11		11	$z$	—	1	1	1	1	2	1	1	1	1	1	
12		12	$ze/m$	MC/kg	3.576	1.384	0.894	0.840	1.626	0.726	0.490	0.481	0.466	0.462	
13		13	Fair	V/mm	19.0	15.2	24.5	13.4	23.1	5.1	52.6	30.5	19.9	18.3	
14		14	Fair	$M_E$	1.60	1.02	2.66	0.79	2.36	0.115	12.2	4.12	1.75	1.48	
15		15	Fair	$r_E$	mm	1.14	1.40	0.73	1.40	0.47	1.22	0.19	0.24	0.52	0.51
16		16	Good	$C_p$	$\mu\text{A nm}^{-2} \text{ MPa}^{-1/2}$	0.247	0.153	0.122	0.100	0.192	0.044	0.091	0.079	0.068	0.065
17		17	Good	$J_{sp}$	$\mu\text{A}/\text{nm}^2$	9.87	4.89	6.30	2.81	9.34	0.47	10.1	5.08	2.85	2.52
18		18	Poor	$J_{fp}$	$\mu\text{A}/\text{nm}^2$	2.27	2.25	2.23	1.77	3.46	0.66	2.26	1.90	1.58	1.51
19		19	Poor	$(J_{fp}J_{sp})^2$	—	0.053	0.212	0.126	0.399	0.137	1.94	0.050	0.139	0.308	0.359
20		20	Poor	$\chi_{max}$	—	0.947	0.788	0.874	0.601	0.863	na	0.950	0.861	0.692	0.641
21		21	Poor	$r_a(Q=0)$	nm	1.21	1.78	0.83	2.33	0.55	na	0.20	0.27	0.76	0.80
22		22	Good	$C_M$	$\mu\text{A } \mu\text{m}^{-1} \text{ kV}^{1/2}$	475	232	251	140	196	16.4	225	92.6	85.9	70.6
23		23	Good	$C_A$	$\text{nm } \mu\text{A}^{-1} \text{ kV}^{-1/2}$	2.75	5.62	5.19	9.31	6.65	79.6	5.80	14.1	15.2	18.5
24		24	Fair	$t_1$	—	0.8	1.5	0.4	1.4	-0.1	10	0.02	0.2	0.5	0.5
25		25	Fair	$i_u$	$\mu\text{A}$	34	25	9	14	6	1.8	1.0	0.7	2.0	1.7

See text for important remarks on accuracy.

Comments on accuracy are necessary for Table A.2.

1. The parameters on rows 2–10 are well defined.
2. Results are shown for  $z = 1$  (except Sn, for which we put  $z = 2$ ). Where parameters (e.g., current density) depend on the mean (nonintegral) charge state of the ions in the beam, then results need correction.
3. For consistency of calculation, some tabulated parameters are shown to three significant figures, although their intrinsic accuracy does not justify this; the comments in the column headed “ok?” provide a qualitative indication of likely accuracy.
4. The ZQEF  $F^M$  derives from Müller’s formula, which may underestimate true values by 20% or more. The derived quantity  $M_E$  could be 40% too low, and  $r_E$  and  $r_a$  40% too high (or more), for this reason.
5.  $J_{FP}$  has been derived from Equation 2.7 using the arbitrary values  $A_F = 5 \times 10^{12} \text{ s}^{-1}$ ,  $N = 5$ ,  $\alpha_r = 1$ . The true value of  $J_{FP}$  could be significantly less, and this uncertainty makes the values on rows 19 and 20 very uncertain.
6. Estimates of the coefficients  $C_M$  and  $C_\lambda$  involve only well-defined constants and should be good.
7. Estimates of  $t_1$  should be fair to good.
8. Estimates of  $i_u$  involve  $(F^M)^{-3}$  and could, in principle, be too high (though there is little evidence of this in the comparisons Mair made with experiment<sup>44</sup>).

## REFERENCES

1. P.D. Prewett and G.L.R. Mair, 1991. *Focused Ion Beams from Liquid Metal Ion Sources* (Research Studies Press, Taunton, UK).
2. J. Orloff, M. Utlaut and L.W. Swanson, 2003. *High Resolution Focused Ion Beams: FIB and Its Applications* (Kluwer, New York).
3. L.A. Giannuzzi and F.A. Stevie (eds), 2005. *Introduction to Focused Ion Beams* (Springer, New York).
4. B.W. Ward, J.A. Notte and N.P. Economou, 2006. Helium ion microscope: A new tool for nanoscale microscopy and metrology. *J. Vac. Sci. Technol.* **24**, 2871–2874.
5. P. Lozano and M. Martínez-Sánchez, 2005. Ionic liquid ion sources: Characterization of externally wetted emitters. *J. Colloid Interface Sci.* **282**, 415–421.
6. For a useful brief review, see: C.D. Hendricks, 1966. Physics of charged colloidal particles and the technology of their production, in: F.E. Marble and J. Suruge (eds) *Physics and Technology of Ion Motors* (Gordon and Breach, New York), pp. 139–152.
7. V.E. Krohn, 1961. Liquid metal droplets for heavy particle propulsion. *Prog. Astronaut. Rocketry* **5**, 73–80.
8. J. Perel, 1968. Alkali metal ion sources. *J. Electrochem. Soc.* **115**, 343C–350C.
9. V.E. Krohn, 1974. Electrohydrodynamic capillary sources of ions and charged droplets. *J. Appl. Phys.* **45**, 1144–1146.
10. V.E. Krohn and G.R. Ringo, 1976. Microprobe design using a liquid gallium ion source. *Int. J. Mass Spectrom. Ion Phys.* **22**, 307–308.
11. G.I. Taylor, 1964. Disintegration of water drops in an electric field. *Proc. R. Soc. Lond. A* **280**, 383–397.
12. R. Clampitt, K.L. Aitken and D.K. Jefferies, 1975. Intense field emission ion source of liquid metals (abstract). *J. Vac. Sci. Technol.* **12**, 1208.
13. K.L. Aitken, 1977. Mechanisms of ion emission from liquid caesium. *Proc. Field Emission Day at ESTEC, Nordwijk, 9th April 1976* (European Space Agency, Paris, Report SP-119), 232–239.
14. G. Benassayag, P. Sudraud and B. Jouffrey, 1985. *In-situ* high-voltage TEM observation of an electrohydrodynamic (EHD) ion source. *Ultramicroscopy* **16**, 1–8.
15. A. Wagner and T.M. Hall, 1979. Liquid gold ion source. *J. Vac. Sci. Technol.* **16**, 1871–1874.
16. A.E. Bell and L.W. Swanson, 1986. The influence of substrate geometry on the emission properties of a liquid metal ion source. *Appl. Phys. A* **41**, 335–346.
17. L.W. Swanson and J.Z. Li, 1988. The influence of electrode geometry on liquid metal ion source performance. *J. Vac. Sci. Technol. B* **6**, 1062–1065.

18. R. Mühle, M. Döbeli and A. Melnikov, 2007. Investigations on a Co–Dy alloy ion source. *J. Phys. D: Appl. Phys.* **40**, 2594–2602.
19. M. Tajmar, A. Genovese and W. Steiger, 2004. Indium FEEP microthruster experimental characterization. *AIAA J. Propulsion Power* **20**, 211–218.
20. F.G. Rudenauer, 2007. Field emission devices for space applications, *Surf. Interface Anal.* **39**, 116–122.
21. A. Jaworek, 2007. Electrospray droplet sources for thin film deposition. *J. Mater. Sci.* **42**, 266–297.
22. P. Lozano and M. Martínez-Sánchez, 2005. On the dynamic response of externally wetted ion-liquid ion sources. *J. Phys. D: Appl. Phys.* **38**, 2371–2377.
23. H. Niedrig, 1996. Ion and electron emission from liquid metal ion sources. *Scanning Microscopy* **10**, 919–935.
24. R.G. Forbes, 1997. Understanding how the liquid metal ion source works. *Vacuum* **48**, 85–97.
25. R.A.D. Mackenzie and G.D.W. Smith, 1990. Focused ion beam technology: A bibliography. *Nanotechnology* **1**, 163–201.
26. P.J. Mohr, B.N. Taylor and D.B. Newell, 2007. CODATA recommended values of the fundamental physical constants. See <<http://physics.nist.gov/constants>>.
27. W. Gilbert, 1600. *De Magnete* (Petrus Short, London) (in Latin). Translation by P.F. Mottelay, 1958 (Dover, New York). See p. 89.
28. S. Gray, 1731/2. A letter concerning the electricity of water. *Phil. Trans. (1683–1775)* (of the Royal Society of London) **37**, 227–230 and 260. (Available on-line via JSTOR.)
29. J.T. Desaguliers, 1741. Several electrical experiments, made at various times, before the Royal Society. *Phil. Trans. (1683–1775)* (of the Royal Society of London) **41**, 661–667. (Available on-line via JSTOR.)
30. G.M. Bose, 1745. *Recherches sur la cause et sur la véritable théorie de l'électricité* (Wittenberg).
31. A.G. Bailey, 1988. *Electrostatic Atomisation* (Research Studies Press, Taunton, UK).
32. M. Cloupeau and B. Prunet-Foch, 1994. Electrohydrodynamic spraying functioning modes: A critical review. *J. Aerosol. Sci.* **25**, 1021–1036.
33. G. Benassayag, 1984. Etude du mécanismes d'émission des sources d'ions à métal liquide. Cas des métaux nobles. Docteur thesis, University Paul Sabatier, Toulouse.
34. A.M. Gañan-Calvo, 1997. On the theory of electrohydrodynamically driven jets. *J. Fluid Mech.* **335**, 165–188.
35. R.G. Forbes, 2000. Liquid-metal ion sources and electrosprays operating in cone-jet mode: Some theoretical comparisons and comments. *J. Aerosol. Sci.* **31**, 97–120.
36. W. Diesel and Ch. Dietzsch, 1996. *In situ* HVTEM observation of the tip shape of tin liquid metal ion sources. *Appl. Surf. Sci.* **93**, 179–190.
37. J.C. Beckman, 1997. Low-current emission characteristics of liquid metal ion sources. PhD thesis, Stanford University.
38. J.C. Beckman, T.H.P. Chang, A. Wagner and R.F.W. Pease, 1997. Minimum emission current of liquid metal ion sources. *J. Vac. Sci. Technol. B* **15**, 2332–2336.
39. G.L.R. Mair, D.C. Grindrod, M.S. Mousa and R.V. Latham, 1983. Beam energy distribution measurements of liquid gallium field-ion sources. *J. Phys. D: Appl. Phys.* **16**, L209–L213.
40. G.L.R. Mair, R.G. Forbes, R.V. Latham and T. Mulvey, 1983. Energy spread measurements on a liquid metal ion source, in: *Microcircuit Engineering '83* (ISBN 0.12.044980.3) (Academic, London).
41. P. Marriott, 1986. A liquid metal ion source analysis system. *J. Phys. D: Appl. Phys.* **19**, L115–L119.
42. A.E. Bell, K. Rao, G.A. Schwind and L.W. Swanson, 1988. A low-current liquid metal ion source. *J. Vac. Sci. Technol. B* **6**, 927–930.
43. J.C. Beckman, T.H.P. Chang, A. Wagner and R.F.W. Pease, 1996. Energy spread in liquid metal ion sources at low currents. *J. Vac. Sci. Technol. B* **15**, 3911–3915.
44. G.L.R. Mair, 1996. Electrohydrodynamic instabilities and the energy spread of ions drawn from liquid metals. *J. Phys. D: Appl. Phys.* **29**, 2186–2192.
45. J.C. Maxwell, 1873. *A Treatise on Electricity and Magnetism*, Vol. I, (1st edition, 1873, 3rd edition 1891; 3rd edition reprinted as an Oxford Classic Text, 1998, by Clarendon Press, Oxford).
46. R.G. Forbes, 2003. Field ion and electron emission: A strategic historical review of theoretical concepts. *Ultramicroscopy* **95**, 1–18.
47. N.D. Lang and W. Kohn, 1973. Theory of metal surfaces: Induced surface charge and image potential. *Phys. Rev. B* **7**, 3541–3550.
48. R.G. Forbes, 1999. The electrical surface as the centroid of the surface-induced charge. *Ultramicroscopy* **79**, 25–34.
49. R.G. Forbes, 2001. Fundamental new results in the energetics and thermodynamics of charged metal surfaces. *Ultramicroscopy* **89**, 1–5.

50. R.G. Forbes, 1995. Field evaporation theory: A review of basic ideas. *Appl. Surf. Sci.* **87/88**, 1–11.
51. C.M.C. de Castilho and D.R. Kingham, 1986. Field ion energy deficit calculations for liquid metal ion sources. *J. Phys. D: Appl. Phys.* **19**, 147–156.
52. L.W. Swanson, 1983. Liquid metal ion sources: Mechanism and applications. *Nucl. Instrum. Meth.* **218**, 347–353.
53. D.R. Kingham, 1982. The post-ionization of field evaporated ions: A theoretical explanation of multiple charge states. *Surf. Sci.* **116**, 273–301.
54. D.G. Brandon, 1964. The structure of field evaporated surfaces. *Surf. Sci.* **3**, 1–18.
55. M.K. Miller, A. Cerezo, M.G. Heatherington and G.D.W. Smith, 1996. *Atom Probe Field Ion Microscopy* (Clarendon Press, Oxford).
56. R.K. Biswas and R.G. Forbes, 1982. Theoretical arguments against the Müller-Schottky mechanism of field evaporation. *J. Phys. D: Appl. Phys.* **15**, 1323–1338.
57. R.G. Forbes, R.K. Biswas and K. Chibane, 1982. Field evaporation theory: A re-analysis of published field sensitivity data. *Surf. Sci.* **114**, 498–514.
58. C.G. Sanchez, A.Y. Lozovoi and A. Alavi, 2004. Field evaporation from first principles. *Mol. Phys.* **102**, 1045–1055.
59. R.G. Forbes, 1982. New activation-energy formulae for charge-exchange-type mechanisms of field evaporation. *Surf. Sci.* **116**, L195–L201.
60. R.G. Forbes, K. Chibane and N. Ernst, 1984. Derivation of bonding distance and vibration frequency from field evaporation experiments. *Surf. Sci.* **141**, 319–340.
61. R.G. Forbes and K. Chibane, 1986. Derivation of an activation energy formula in the context of charge draining. *J. de Phys.* **47**, Colloque C7, 65–70.
62. T.E. Stern, B.S. Gossling and R.H. Fowler, 1929. Further studies in the emission of electrons from cold metals. *Proc. R. Soc. Lond. A* **280**, 383–397.
63. P.S. Laplace, 1806. *Traité de mécanique céleste; supplément au dixième live, “Sur l'action capillaire”* (Courvier, Paris).
64. P.S. Laplace, 1807. *Supplement à la théorie de l'action capillaire* (Courvier, Paris).
65. J.W. Gibbs, 1875/76. On the equilibrium of heterogeneous substances. *Trans. Connecticut Acad.* **3**, 108–248 and 343–524. Reprinted in *The Scientific Papers of J.W. Gibbs, Vol. I Thermodynamics* (Dover, New York, 1961). See pp. 219–229 and eq. (500).
66. N.N. Ljepojevic and R.G. Forbes, 1995. Variational thermodynamic derivation of the formula for pressure difference across a charged conducting liquid surface and its relation to the thermodynamics of electrical capacitance. *Proc. R. Soc. Lond. A* **450**, 177–192.
67. Raleigh Lord, 1882. On the equilibrium of liquid conducting masses charged with electricity. *Philos. Mag.* **14**, 184–186.
68. Raleigh Lord, 1916. On the electrical capacity of approximate spheres and cylinders. *Philos. Mag.* **31**, 177–186.
69. C.D. Hendricks and J.M. Schneider, 1962. Stability of a conducting droplet under the influence of surface tension and electric forces. *Am. J. Phys.* **31**, 450–453.
70. J.M.H. Peters, 1980. Raleigh's electrified water drops. *Eur. J. Phys.* **1**, 143–146.
71. D.R. Kingham and A.E. Bell, 1985. Comment on “Variational formulation for the equilibrium condition of a conducting fluid in an electric field.” *Appl. Phys. A* **36**, 67–70.
72. M.S. Chung, P.H. Cutler, J. He and N.M. Miskovsky, 1991. A first-order electrohydrodynamic treatment of the shape and instability of liquid metal ion sources. *Surf. Sci.* **146**, 118–124.
73. R.G. Forbes and N.N. Ljepojevic, 1991. Calculation of the shape of the liquid cone in a liquid metal ion source. *Surf. Sci.* **246**, 113–117.
74. D.R. Kingham and L.W. Swanson, 1984. Shape of a liquid-metal ion source: A dynamic model including liquid flow and space-charge effects. *Appl. Phys. A* **34**, 123–132.
75. V.P. Kovalenko and A.L. Shabalin, 1989. Lower limit on the current of an electrohydrodynamic emitter. *Pis'ma Zh. Tekh. Fiz.* **15**(3), 62–65. (Translated: *Sov. Tech. Phys. Lett.* **15**, 232–234.)
76. W. Liu and R.G. Forbes, 1995. Modelling the link between emission current and LMIS cusp length. *Appl. Surf. Sci.* **87**, 122–126.
77. T. Radlicka and B. Lencová, 2008. Coulomb interactions in Ga LMIS. *Ultramicroscopy* **108**, 445–454.
78. R.G. Forbes, 2008. Appendix to “Coulomb interactions in Ga LMIS” by Radlicka and Lencová. *Ultramicroscopy* **108**, 455–457.
79. K.L. Aitken, D. K. Jefferies and R. Clampitt, 1975. *Culham Lab. Rept.* CLM/RR/E1/21.
80. G.L.R. Mair, 1979. Physics of anode emissions. DPhil thesis, Oxford University.

81. G.L.R. Mair and T. Mulvey, 1986. Some aspects of liquid metal ion sources. *J. Microsc.* **142**, 191–200.
82. S. Papadopoulos, 1986. Ion emission from liquid alloy ion sources. DPhil thesis, Oxford University.
83. G.L.R. Mair and S. Thoms, 1989. Ions field emitted from a liquid-gallium capillary. *J. Phys. D: Appl. Phys.* **22**, 975–976.
84. R.G. Forbes, G.L.R. Mair, N.N. Ljepojevic and W. Liu, 1995. New understandings in the theory of liquid metal ion sources. *Appl. Surf. Sci.* **87/88**, 99–105.
85. M.D. van Dyke, 1965. An asymptotic approximation for slender bodies. Appendix in Ref. 87.
86. M.D. van Dyke, 1969. The electrostatics of a needle normal to a plane. Appendix in Ref. 88.
87. G.I. Taylor, 1965. The force exerted by an electric field on a long cylindrical conductor. *Proc. R. Soc. Lond. A* **291**, 145–158.
88. G.I. Taylor, 1969. Electrically driven jets. *Proc. R. Soc. Lond. A* **313**, 453–475.
89. A.L. Shabalin, 1989. Can an electrohydrodynamic emitter operate with weak currents? *Pis'ma Zh.Tekh. Fiz.* **15**(12), 27–29. (Translated: *Sov. Tech. Phys. Lett.* **15**, 924.)
90. G.L.R. Mair, 1989. On the onset voltage of liquid metal ion sources. *Nucl. Instrum. Meth. Phys. Res. B* **43**, 240–242.
91. G.L.R. Mair, 1984. Theoretical determination of current–voltage curves for liquid-metal ion sources. *J. Phys. D: Appl. Phys.* **17**, 2323–2330.
92. G.L.R. Mair, 1986. Current–voltage curves in liquid metal ion sources. *Vacuum* **36**, 847–850.
93. G.L.R. Mair, 1982. Space-charge effects in liquid metal ion sources. *J. Phys. D: Appl. Phys.* **15**, 2523–2530.
94. J.F. Mahoney, A.Y. Yakihu, H.L. Daley, R.D. Moore and J. Perel, 1969. Electrohydrodynamic ion source. *J. Appl. Phys.* **40**, 5101–5106.
95. B.W. Weinstein, 1975. High current ion emission in electrohydrodynamic spraying. PhD thesis, University of Illinois at Urbana-Champaign.
96. Y.Y. Lau, 2001. Simple theory for the two-dimensional Child-Langmuir Law. *Phys. Rev. Lett.* **87**, 278301.
97. G.L.R. Mair, 1997. The effects of flow impedance on the current–voltage characteristics of liquid-metal ion sources. *J. Phys. D: Appl. Phys.* **17**, 1945–1950.
98. A. Wagner, 1982. The hydrodynamics of liquid metal ion sources. *Appl. Phys. Lett.* **40**, 440–442.
99. K. Jousten, J.F. Holmes and J. Orloff, 1991. High frequency modulation of a gallium liquid metal ion source using a laser beam and thermal effects. *J. Phys. D: Appl. Phys.* **24**, 458–468.
100. L. Bischoff, J. Teichert, Th. Ganetsos and G.L.R. Mair, 2000. Temperature dependence of the electric characteristics of liquid metal alloy ion sources. *J. Phys. D: Appl. Phys.* **33**, 692–695.
101. Th. Ganetsos, G.L.R. Mair and L. Bischoff, 2003. On the temperature dependence of the electric characteristics and mass spectra of liquid metal alloy ion sources. *Ultramicroscopy* **95**, 171–181.
102. G.L.R. Mair, 1992. Theory of current modulation in a liquid metal ion source using thermal effects. *J. Phys. D: Appl. Phys.* **25**, 1284–1287.
103. B. Praprotnik, W. Diesel, Ch. Dietzsch and H. Niedrig, 1994. HV-TEM *in-situ* investigations of the tip shape of indium liquid metal ion emitter. *Surf. Sci.* **314**, 353–364.
104. W. Diesel, Ch. Dietzsch and R. Mühlé, 1995. Tip-shape investigation of a Au–Si alloy liquid metal ion source using a high-voltage transmission electron microscope. *J. Phys. D: Appl. Phys.* **28**, 787–793.
105. G.L.R. Mair and R.G. Forbes, 1991. Analytical determination of the dimensions and evolution with current of the ion emitting jet in liquid-metal ion sources. *J. Phys. D: Appl. Phys.* **28**, 2217–2221.
106. G.L.R. Mair and R.G. Forbes, 1992. An analytical calculation of LMIS cusp length. *Surf. Sci.* **266**, 180–184.
107. V.G. Suvorov and R.G. Forbes, 2004. Theory of minimum emission current for a non-turbulent liquid-metal ion source. *Microelectron. Eng.* **73–74**, 126–131.
108. F.J. Higuera, 2004. Liquid flow induced by ion evaporation in an electrified meniscus. *Phys. Rev. E* **69**, 066301.
109. G.Sh. Boltachev and N.M. Zubarev, 2006. Influence of the space charge on the configuration of conical spikes on a liquid-metal surface. *Europhys. Lett.* **76**, 36–41.
110. J.R.A. Cleaver and H. Ahmed, 1981. A 100-kV ion probe microfabrication system with a tetrode gun. *J. Vac. Sci. Technol.* **19**, 1145–1148.
111. R.G. Forbes, 1976. A generalised theory of standard field-ion appearance energies. *Surf. Sci.* **61**, 221–240.
112. R.J. Culbertson, G.H. Robertson, Y. Kuk and T. Sakurai, 1980. Atom-probe field ion microscopy of a high intensity gallium ion source. *J. Vac. Sci. Technol.* **17**, 203–206.
113. B. Zimmerman, 1970. *Adv. Electron. Electron Phys.* **29**, 257.

114. W. Knauer, 1979. Analysis of energy broadening in electron and ion beams. *Optik* **54**, 211–234.
115. W. Knauer, 1981. Energy broadening in field emitted electron and ion beams. *Optik* **59**, 335–354.
116. M.A. Gesley and L.W. Swanson, 1984. Analysis of energy broadening in charged particle beams. *J. de Phys.* **45**, Colloque C9, 167–172.
117. J. Puretz, 1986. Dimensional analysis and energy broadening in particle beams. *J. Phys. D: Appl. Phys.* **19**, L237–L240.
118. G.H. Jansen, 1990. *Coulomb Interactions in Particle Beams* (Academic, Boston) (*Adv. Electron. Electron Phys., Supplement 21*).
119. J.W. Ward and R.L. Kubena, 1990. Stochastic effects occurring after emission from liquid metal ion sources. *J. Vac. Sci. Technol. B* **8**, 1923–1926.
120. G.L.R. Mair, 1987. Energy deficit measurements with a liquid gallium ion source. *J. Phys. D: Appl. Phys.* **20**, 1657–1660.
121. J.W. Ward, R.L. Kubena and M.W. Utlaut, 1988. Transverse thermal velocity broadening of focused beams from liquid metal ion sources. *J. Vac. Sci. Technol. B* **6**, 2090–2094.
122. R.L. Kubena and J.W. Ward, 1987. Current-density profiles for a  $\text{Ga}^+$  ion microprobe and their lithographic implications. *Appl. Phys. Lett.* **51**, 1960–1962.
123. R.L. Kubena and J.W. Ward, 1988. Comment on “Current-density profiles for  $\text{Ga}^+$  ion microscope and their lithographic implications.” *Appl. Phys. Lett.* **52**, 2089.
124. L.W. Swanson, A.E. Bell and G.A. Schwind, 1988. A comparison of boron emission characteristics for liquid metal ion sources of PtB, PdB and NiB. *J. Vac. Sci. Technol.* **6**, 491–495.
125. Y-G Kim, Y-S Kim, E-H Choi, S-O Kang, G. Cho and H.S. Uhm, 1998. Temperature effects on the energy spread in liquid metal ion sources. *J. Phys. D: Appl. Phys.* **31**, 3463–3468.
126. L.W. Swanson, G.A. Schwind, A.E. Bell and J.E. Brady, 1979. Emission characteristics of gallium and bismuth liquid metal ion sources. *J. Vac. Sci. Technol.* **16**, 1864–1867.
127. C.J. Aidinis, G.L.R. Mair, L. Bischoff and I. Papadopoulos, 2001. A study of the temperature dependence of the energy spread and energy deficit of a  $\text{Ge}^{++}$  ion beam produced by a liquid alloy ion source. *J. Phys. D: Appl. Phys.* **34**, L14–L16.
128. L.W. Swanson, G.A. Schwind and A.E. Bell, 1980. Measurement of the energy distribution of a gallium liquid metal ion source. *J. Appl. Phys.* **51**, 3453–3455.
129. M. Komuro, H. Arimoto and T. Kato, 1988. On the mechanism of energy distribution in liquid metal ion sources. *J. Vac. Sci. Technol. B* **6**, 923–926.
130. T. Kosuge, H. Tanitsu and T. Makabe, 1989. Computer simulation of ion energy distribution in liquid metal ion sources, in: T. Takagi (ed.), *Proc. 12th Symp. on Ion Sources and Ion-Assisted Technology (ISIAT '89)*, 43–46.
131. R. Hornsey, 1991. Simulations of the current and temperature dependence of liquid metal ion source energy distributions. *Jpn. J. Appl. Phys.* **30**, 366–375.
132. E.H. Hirsch, 2001. Anomalous energy distribution in electron beams. *J. Phys. D: Appl. Phys.* **34**, 3229–3233.
133. M. Komuro, T. Kanayama, H. Hiroshima and H. Tanoue, 1982. Measurement of virtual cross-over in liquid gallium ion source. *Appl. Phys. Lett.* **42**, 908–910.
134. G.H. Jansen, 1998. Trajectory displacement effect in particle projection lithography systems: modifications to the extended two-particle theory and Monte-Carlo simulation technique. *J. Appl. Phys.* **84**, 4549–4567.
135. J.W. Ward, 1985. A Monte Carlo calculation of the virtual source size for a liquid metal ion source. *J. Vac. Sci. Technol. B* **3**, 207–213.
136. S. Georgieva, R.G. Vichev and N. Drandarov, 1993. Some estimates of the virtual source size of a liquid metal ion source. *Vacuum* **44**, 1109–1111.
137. G.L.R. Mair, 1980. Emission from liquid metal ion sources. *Nucl. Instrum. Meth.* **172**, 567–576.
138. A. Dixon, C. Colliex, R. Ohana, P. Sudraud and J. van de Walle, 1981. Field-ion emission from liquid tin. *Phys. Rev. Lett.* **46**, 865–868.
139. D.L. Cocke, G. Abend and J.H. Block, 1976. Mass spectrometric observation of large sulfur molecules from condensed sulfur. *J. Phys. Chem.* **80**, 524–528.
140. P. Joyes and J. Van de Walle, 1985. Liquid-metal ion source study of critical sizes of multiply-charged positive ions. *J. Phys. B: At. Mol. Phys.* **18**, 3805–3810.
141. P. Joyes, J. Van de Walle and R.-J. Tarento, 2006. Nanotubes stretched in liquid-metal-ion-sources and their influence on isotopic anomalies. *Phys. Rev. B* **73**, 115436.
142. P. Sudraud, C. Colliex and J. van de Walle, 1978. Energy distribution of EHD emitted gold ions. *J. de Phys.—Lett.* **40**, L207–L211.

143. A.J. Dixon, 1979. The energy spread of metal ions drawn from their liquid. *J. Phys. D: Appl. Phys.* **12**, L77–L80.
144. A.R. Waugh, 1981. Current-dependence of mass and energy spectrum of ions from liquid gold and gold alloy ion sources, in: *28th Intern. Field Emission Symp.*, Portland, July 1981 (unpublished abstracts) 80–82.
145. L. Bischoff and G.L.R. Mair, 2003. Alloy liquid metal ion sources for mass separated focused ion beams. *Recent Res. Devel. Appl. Phys.* **6**, 123–133.
146. T. Ishitani, S. Umemura, S. Hosoki, S. Tayama and H. Tamura, 1984. Development of boron liquid-metal ion source. *J. Vac. Sci. Technol. A* **2**, 1365–1369.
147. T. Ishitani, K. Umemura, and H. Tamura, 1984. Development of phosphorus liquid-metal ion source. *Jpn. J. Appl. Phys.* **23**, L330–L332.
148. W.M. Clark, R.L. Seliger, M.W. Utlauf, A.E. Bell, L.W. Swanson, G.A. Schwind, and J.B. Jergenson, 1987. Long-lifetime reliable liquid metal ion sources for boron, arsenic and phosphorus. *J. Vac. Sci. Technol. B* **5**, 197–202.
149. Th. Ganetsos, A.W.R. Mair, G.L.R. Mair, L. Bischoff, C. Akhamadliev and C.J. Aidinis, 2007. Can field evaporation of doubly-charged ions and post-ionisation from the singly-charged state co-exist? *Surf. Interface Anal.* **39**, 128–131.
150. R.H. Higuchi-Rusli, K.C. Cardien, J.C. Corelli and A.J. Steckl, 1987. Development of boron liquid metal ion source for focused ion beam system. *J. Vac. Sci. Technol. B* **5**, 190–194.
151. D.S. Swatik, 1969. Production of high current density ion beams by electrohydrodynamic spraying techniques. PhD thesis, University of Illinois at Urbana-Campaign.
152. R. Juraschek and F.W. Röllgen, 1998. Pulsation phenomena during electrospray ionization, *Int. J. Mass Spectrom. Ion Process* **177**, 1–15.
153. E. Hesse, G.L.R. Mair, L. Bischoff and J. Teichert, 1996. Parametric investigation of current pulses in a liquid metal ion source. *J. Phys. D: Appl. Phys.* **29**, 2193–2197.
154. C. Akhamadliev, G.L.R. Mair, C.J. Aidinis and L. Bischoff, 2002. Frequency spectra and electrohydrodynamic phenomena in a liquid gallium field ion emission source. *J. Phys. D: Appl. Phys.* **35**, L91–L93.
155. C. Akhamadliev, L. Bischoff, G.L.R. Mair, C.J. Aidinis and Th. Ganetsos, 2006. Investigation of emission stabilities of liquid metal ion sources. *Microelectron. Eng.* **73–74**, 120–125.
156. D.L. Barr and W.L. Brown, 1989. Droplet emission from a gallium liquid metal ion source as observed with an ion streak camera. *J. Vac. Sci. Technol. B* **7**, 1806–1809.
157. V.V. Vladimirov, V.E. Badan, V.N. Gorshov and I.A. Soloshenko, 1993. Liquid metal microdroplet source for deposition purposes. *Appl. Surf. Sci.* **65/66**, 1–12.
158. A. Wagner, T. Venkatesan, P.M. Petroff and D. Barr, 1981. Droplet emission in liquid metal ion sources. *J. Vac. Sci. Technol.* **19**, 1186–1189.
159. W. Diesel, H. Niedrig and B. Praprotnik. TEM *in-situ* investigations of the emission behaviour of liquid metal ion emitter. Video recording provided privately to RGF by H. Niedrig.
160. V.T. Binh and N. Garcia, 1992. On the electron and metallic ion emission from nanotips fabricated by field-surface melting technique: Experiments on W and Au tips. *Ultramicroscopy* **42–44**, 80–90.
161. W. Polanschutz and E. Krautz, 1974. FIM studies at higher temperature of f.c.c. metals Cu, Ag, Ni, Pd, Pt, in: *21st Intern. Field Emission Symp.*, Marseille, July 1974, 60. (The oral presentation included a convincing video.)
162. K. Ishimoto, H.M. Pak, T. Nishida and M. Doyama, 1974. Field ion microscopy of W and Pt-Ti at about 1000 degrees C. *Surf. Sci.* **41**, 102–112.
163. V.G. Suvorov, 2004. Numerical analysis of liquid metal flow in the presence of an electric field: Application to liquid metal ion source. *Surf. Interface Anal.* **36**, 421–425.
164. S.P. Thompson, 1984. Neutral emissions from liquid metal ion sources. *Vacuum* **34**, 223–238.
165. S.P. Thompson and P.D. Prewett, 1984. The dynamics of liquid metal ion sources. *J. Phys. D: Appl. Phys.* **17**, 2305–2321.
166. P.R. Brazier-Smith, S.G. Jennings and J. Latham, 1971. An investigation into the behaviour of drops and drop-pairs subjected to strong electrical forces. *Proc. R. Soc. Lond. A* **325**, 363–376.
167. Z. Cui and L. Tong, 1988. A new approach to simulating the operation of liquid metal ion sources. *J. Vac. Sci. Technol. B* **6**, 2104–2107.
168. V.G. Suvorov and E.A. Litvinov, 2000. Dynamic Taylor cone formation on liquid metal surface: Numerical modelling. *J. Phys. D: Appl. Phys.* **33**, 1245–1251.
169. O.M. Belozerovskii, 1994. *Numerical Simulations in Continuous Media Mechanics* (Physics and Mathematics Literature, Moscow), p. 442.

170. A. VanderWyst, A. Christlieb, M. Sussman and I.D. Boyd, 2005. Simulation of charge and mass distributions of indium droplets created by field emission, in: *37th AIAA Plasmadynamics and Lasers Conf.*, San Francisco, June 2006 (AIAA, Reston, VA), pp. 1–20.
171. Y. Torii and H. Yamada, 1983. Field emission assisted liquid aluminum ion source, in: T. Takagi (ed.) *Proc. Intern. Ion Eng. Congress*, Kyoto, September 1983, 363–368.
172. M. J. Bozack, L. W. Swanson, and A. E. Bell, 1987. Wettability of transition metal boride eutectic alloys to graphite. *J. Mater. Sci.* **22**, 2421–2430.
173. P.D. Prewett and E.M. Kellogg, 1985. Liquid metal ion sources for FIB microfabrication systems—Recent advances. *Nucl. Instrum. Meth. Phys. Res. B* **6**, 135–142.
174. C.S. Galovich, 1987. Effects of back-sputtered material on gallium liquid metal ion source behavior. *J. Appl. Phys.* **63**, 4811–4816.
175. C.S. Galovich and A. Wagner, 1988. A new method for improving gallium liquid metal ion source stability. *J. Vac. Sci. Technol. B* **6**, 1186–1189.
176. W. Czarczynski, 1995. Secondary electrons in liquid metal ion sources. *J. Vac. Sci. Technol.* **13**, 113–116.
177. G.L.R. Mair, 1997. On the role of secondary electrons in liquid metal ion sources. *J. Phys. D: Appl. Phys.* **30**, 921–924.
178. W. Czarczynski, 1997. Author's response to the comment on the role of secondary electrons in liquid metal ion sources. *J. Phys. D: Appl. Phys.* **30**, 925–926.
179. C.D. Hendricks and D.S. Swatik, 1973. Field emitted ion beams from liquids for electric thrusters. *Astronautica Acta* **18**, 295–300.
180. S.T. Purcell, V.T. Binh and P. Thevenard, 2001. Atomic-size metal ion sources: Principles and use. *Nanotechnology* **12**, 168–172.
181. J. Ishikawa, H. Tsuji, T. Kashiwagi and T. Takagi, 1989. Intense metal ion beam formation by an impregnated-electrode-type liquid-metal ion source. *Nucl. Instrum. Meth. Phys. Res. B* **37/38**, 155–158.
182. C. Bartoli, H. von Rhoden, S.P. Thompson and J. Blommers, 1984. A liquid caesium field ion source for space propulsion. *J. Phys. D: Appl. Phys.* **17**, 2473–2483.
183. L. Tonks, 1935. A theory of liquid rupture by a uniform electric field. *Phys. Rev.* **48**, 562–568.
184. Ya. Frenkel, 1936. On the Tonks theory of the rupture of a liquid surface by an uniform electric field in vacuum, *Zh. Tekh. Fiz.* **6**, 347–350.
185. M.D. Gabovich, 1983. Liquid metal ion emitters. *Usp. Fiz. Nauk.* **140**, 137–152. (Translated: *Sov. Phys. Uspekhi* **26**, 447–455.)
186. A.L. Pregenzer, 1985. Electrohydrodynamically driven large-area liquid-metal ion sources. *J. Appl. Phys.* **58**, 4509–4511.
187. S.E. Buttrill Jr. and C.A. Spindt, 1978. Evaluation of volcano-style field ionization source and field emitting cathodes for mass spectrometry and applications. NASA Technical Reports Server, Report No. NASA-CR-156986.
188. J. Mitterauer, 1995. Microstructured liquid metal ion and electron sources. *Appl. Surf. Sci.* **87/88**, 79–90.
189. M. Martinez-Sanchez and A.I. Akinwande, 2005. Colloid thrusters, physics, fabrication and performance. US Defense Technical Information Centre, Report No. ADA442444.
190. P.W. Hawkes and B. Lençova, 2006. Charged Particles Optics Theory, in: *E-nano-newsletter*, Issue 6 (PHANTOMS Foundation, Madrid). <[www.phantomsnet.net/Foundation/newsletters.php](http://www.phantomsnet.net/Foundation/newsletters.php)>.
191. T.T. Tsong, 1978. Field ion image formation. *Surf. Sci.* **70**, 211–233.
192. A.M. James and M.P. Lord (eds), 1992. *Macmillan's Chemical and Physical Data* (Macmillan, London).
193. D.R. Lide (ed.), 2000. *CRC Handbook of Chemistry and Physics*, 81st edition (CRC Press, Boca Raton).
194. T. Durakiewicz, S. Halas, A. Arko, J.J. Joyce and D.P. Moore, 2001. Electronic work-function calculations of polycrystalline metal surfaces revisited. *Phys. Rev. B* **164**, 045101.
195. W.F. Gale and T.T. Totmeier (eds), 2004. *Smithells Metals Reference Book*, 8th edition (Elsevier, Amsterdam).

---

# 3 Gas Field Ionization Sources

*Richard G. Forbes*

## CONTENTS

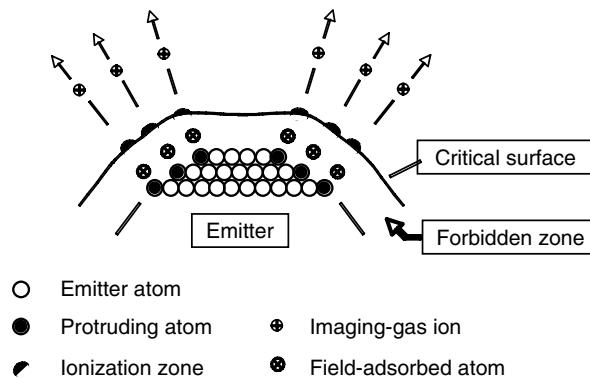
3.1	Introduction.....	88
3.2	Advantages, a Challenge and a Trade-Off.....	91
3.2.1	Advantages of a Gas Field Ionization Source.....	91
3.2.2	The Gas-Pressure Trade-Off .....	91
3.2.3	Secondary Electrons.....	92
3.3	Technological Development.....	92
3.4	Gas Field Ionization Fundamentals .....	93
3.4.1	Preliminaries .....	93
3.4.1.1	Emitter Formation.....	93
3.4.1.2	Field Definitions.....	93
3.4.1.3	The Ionoptical Surface.....	93
3.4.1.4	The Real-Source Current-Density Distribution.....	94
3.4.2	Gas-Atom Behavior .....	94
3.4.2.1	The Critical Surface and the Firmly Field-Adsorbed Layer.....	94
3.4.2.2	The Long-Range Polarization Potential-Energy Well .....	95
3.4.2.3	Typical Gas-Atom History .....	95
3.4.2.4	Best Image Field and Best Source Field .....	96
3.4.3	Ion Generation .....	96
3.4.3.1	Ion Generation Theory.....	96
3.4.3.2	Energy Spreads .....	97
3.5	Theory of Emission Current .....	97
3.5.1	Introduction .....	97
3.5.2	Supply Current and Effective Capture Area .....	99
3.5.3	Ionization Regimes.....	99
3.5.4	Gas Temperature at Ionization .....	100
3.5.5	Supply-and-Capture Regime .....	100
3.6	Basic Real-Source Data .....	101
3.6.1	Emission-Site Radius.....	101
3.6.2	Field Ion Microscope Resolution Criterion .....	103
3.6.3	Illustrative Values .....	103
3.7	The Spherical Charged Particle Emitter .....	103
3.7.1	Basic Ideas .....	103
3.7.1.1	The Optical Model .....	103
3.7.1.2	Ions on Radial Trajectories .....	103
3.7.1.3	The Effect of Transverse Velocity.....	103
3.7.1.4	The Blurred Beam.....	105
3.7.2	Objects and Machines .....	105
3.7.2.1	Optical Objects Generated by the Spherical Charged Particle Emitter.....	105

3.7.2.2	Machines Based on the Müller Emitter .....	106
3.7.2.3	Field Ion Microscope Resolving Power .....	107
3.7.3	Source Sizes and Related Topics .....	107
3.7.3.1	Virtual Source for a GFIS-Based Machine.....	107
3.7.3.2	Transverse Zero-Point Energy .....	108
3.7.3.3	Minimum Value for $\rho_2'$ .....	108
3.7.3.4	Effect on Total Energy Distribution.....	108
3.7.3.5	Other Effects of Ion Energy Spread.....	108
3.8	The Role of the Weak Lens.....	108
3.8.1	Introduction .....	108
3.8.2	Angular Magnification .....	109
3.8.3	Transverse Magnification .....	110
3.8.4	Müller Emitter Source Sizes .....	111
3.8.5	Projection Magnification .....	111
3.8.6	Field Ion Microscope Image-Spot Size .....	111
3.8.7	Numerical Trajectory Analyses.....	112
3.9	Aberrations.....	112
3.9.1	Space Charge.....	112
3.9.2	Local Angular Distortion .....	113
3.9.3	Diffraction at the Beam Acceptance Aperture .....	113
3.9.4	Spherical and Chromatic Aberration.....	113
3.9.5	Gas Field Ionization Source Radius .....	113
3.9.6	Column Aberrations .....	114
3.10	Source Properties .....	114
3.10.1	Alternative Figure of Merit .....	115
3.11	Summary and Discussion.....	116
3.11.1	A Speculation about New Machines of Nanotechnology.....	117
	Acknowledgment .....	117
A.1	Appendix: Corrected Sounthon Gas-Supply Theory .....	117
A.1.1	Preliminary Definitions.....	118
A.1.2	Supply to a Hemisphere.....	118
A.1.3	Supply to a Cylinder .....	118
A.1.4	Supply to a Cone.....	119
A.1.5	Total Captured Flux for Müller Emitter .....	120
A.1.6	Numerical Illustrations.....	120
A.2	Appendix: Glossary of Special Terms .....	122
	List of Abbreviations .....	124
	References .....	125

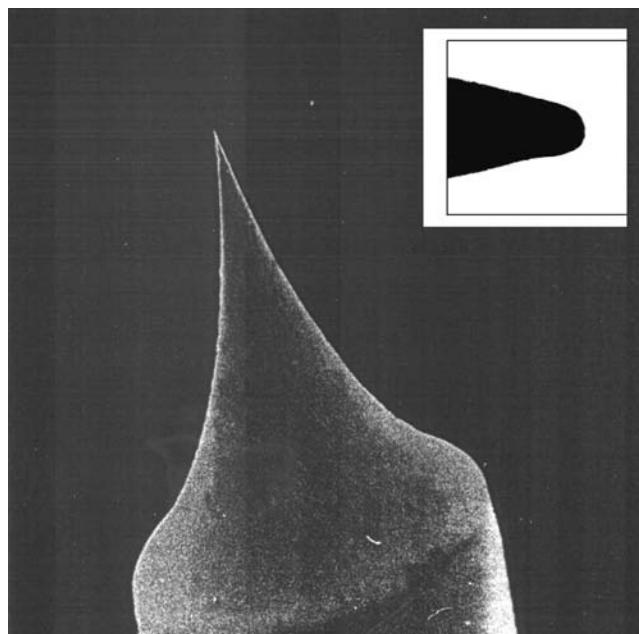
### 3.1 INTRODUCTION

This chapter aims to describe the physics and optics of gas field ionization (GFI) emitters. These differ in two main ways from the liquid metal ion source (LMIS) described in Chapter 2: the ionizant atoms are initially free, rather than bound, and the ions are formed by surface field ionization (FI) slightly above the emitter surface, rather than by field evaporation at the surface.

The ionizant is supplied as neutral gas atoms or molecules. These are attracted to a needle-like field emitter by polarization forces, captured by thermal accommodation, and funneled toward the needle tip by the polarization potential-energy (PPE) well. Surface FI takes place in a thin *ionization layer* just outside a *critical surface* somewhat above the needle's tip ( $\sim 500$  pm above the edge of the metal atom for the helium-on-tungsten [He-on-W] system). In this layer there are strong variations across the surface in the generated ion current density. These relate to the existence of



**FIGURE 3.1** Schematic diagram illustrating the local physical situation at the apex of a helium (He) field ion emitter. A weakly bound rarified layer of He imaging-gas atoms *bounces about* on top of this structure but is not shown. The firmly field-adsorbed atoms are present only at temperatures below  $\sim 100$  K.



**FIGURE 3.2** A conventional tungsten Müller emitter (inset: electron micrograph of apex).

zones of relatively intense ionization above the protruding atoms. From each such *ionization zone* a narrow ion beam emerges. Figure 3.1 illustrates this.

This GFI emitter is used in two related but distinct ways: as an emitter for a projection microscopy (field ion microscopy [FIM]) and as a source for a focused ion beam (FIB) system, called a gas field ionization source (GFIS). The source physics and optics are common, but the operating and focusing conditions are different. Much of the source physics was established long ago when FIM was being developed.<sup>1-7\*</sup>

The needle-like object (see Figure 3.2) to which a high electric field is applied has a fundamental role in the charged particle (CP) optics of many ion and electron emission devices. Here,

---

\* Note that aspects of the physics are seriously incomplete in some of the earlier discussions of field ion emission and that some issues are still not fully settled.

it is called a *Müller emitter*, after the inventor<sup>8</sup> of field electron microscopy, FIM, and atom-probe tomography. Depending on the context, a Müller emitter generates ions or electrons, and the emission can be thermally induced, field induced, photon induced, or mixed. The basic CP optics is common because (in the absence of space charge) particle trajectories are independent of the charge-to-mass ratio and do not depend on whether charge is positive or negative.

The essential feature of a Müller emitter is that the emitted particle moves in an electrostatic field that can be treated as central field of force; consequently (unlike photons), the particle path is curved. Some other sources (e.g., pointed hairpins) can also be treated like this. A valuable property is that the emitted beam can be focused to a fine spot. For this reason, many machines of modern nanotechnology use Müller emitters.

The commonalities of CP optical behavior and the common ion generation process for FIM and the GFIS allow existing knowledge about electron emitters and FIM to be brought together with GFIS specifics to make a unified theory of their ionoptical properties. This chapter presents a first attempt at a synthesis of this kind. The term *field emission* as used here covers both field electron and field ion emission.

Central optical questions for FIM were: how do we resolve atoms (solved qualitatively over 50 years ago though the numerics are still poor), and how do we analyze field ion micrographs (solved long ago). Central questions for a GFIS are: how much current can we get into the beam, and what is the final specimen spot size; what is the energy spread of the total-ion-energy distribution; what are the values of related optical parameters, particularly virtual source size and reduced brightness; and what aberrations does the source have.

A precautionary remark is necessary. The FIM and the GFIS are devices with very complex physics. It is operationally impossible to carry out first-principle prediction of many ionoptical quantities for three reasons. First, the exact potentials in which electrons and atoms move at the emitter surface are not clearly known, and some relevant physical processes are very messy and not fully understood. Second, the quantum-mechanical models for the field ionization of gas atoms near a surface are incomplete and badly in need of updating. Third, the gas kinetics is so complex that accurate calculations are operationally impossible in their own right. Notwithstanding this, useful estimates can be made. But the primary role of theory is to provide physical understanding, leaving quantitative details to be filled in by technological experimentation and development.

The main origins of this chapter are as follows. The theory of the emission process is drawn mainly from the work of the present authors,<sup>3,7,9–11</sup> Southon,<sup>12,13</sup> Müller and Tsong,<sup>2</sup> Van Eekelen,<sup>14</sup> Castilho and Kingham,<sup>4</sup> and Witt and Müller.<sup>15</sup> The general properties of field ion micrographs are described in textbooks on field ion microscopy and atom-probe tomography,<sup>2,5,6</sup> but the articles of Smith,<sup>16</sup> Southworth,<sup>17</sup> and Walls are of specific interest. The approach to the optics of Müller emitters derives from Hawkes and Kasper's standard textbook on electron optics<sup>18</sup> and indirectly from the work of Ruska,<sup>19</sup> Gomer,<sup>1,20</sup> and Wiesner and Everhart.<sup>21–23</sup> Tondare<sup>24</sup> has presented a useful review of GFIS research and development up to early 2005 with many references. Liu and Orloff<sup>25</sup> have begun the process of formalizing GFIS optical properties. A stimulus for this chapter has been the development<sup>26</sup> by the ALIS Corporation (now part of Carl Zeiss SMT) of a commercial helium scanning ion microscope (He SIM) using a pyramidal built-up GFIS.

There are clear gains in this synthesis. In FIM, a micrograph is a map of the positions of emitter surface atoms and of identifiable crystal facets. Ball models (or computerized equivalents) give data about the positions of surface atoms and crystal facets on the emitter surface. Comparisons of map and model give quantitative information about CP projection optics and (potentially) emitter optical aberrations. Hitherto, there has been little use of such information in mainstream CP optics. In the other direction, long-established experimental facts about the optics of field ion micrographs can be given a ready explanation in terms of mainstream ideas.

More generally, in the same way that development of electron microscopy by Ruska and others in the 1930s stimulated the quantitative development of electron source optics, the development of the He SIM may stimulate the integration and development of GFI optics.

The general plan of the chapter is as follows. Sections 3.1 through 3.3 provide an introduction, Sections 3.4 and 3.5 outline the underlying physics, Section 3.6 deals with the size of the real source, Sections 3.7 through 3.10 deal with ionoptical issues, and Section 3.11 provides a summary and discussion. Appendix A.1 provides a corrected theory of gas supply and Appendix B.1 a glossary of special terms used and a list of acronyms.

The conventions of the international system of measurement are used, including the rationalized meter-kilogram-second (rmks) equation system; equations in older literature have been converted to rmks form. In addition, for dimensional consistency in equations, the symbol  $n_1$  is used to formally denote the atomic level unit of the SI quantity *amount of substance*; it is best read here as *one atom*.

## 3.2 ADVANTAGES, A CHALLENGE AND A TRADE-OFF

### 3.2.1 ADVANTAGES OF A GAS FIELD IONIZATION SOURCE

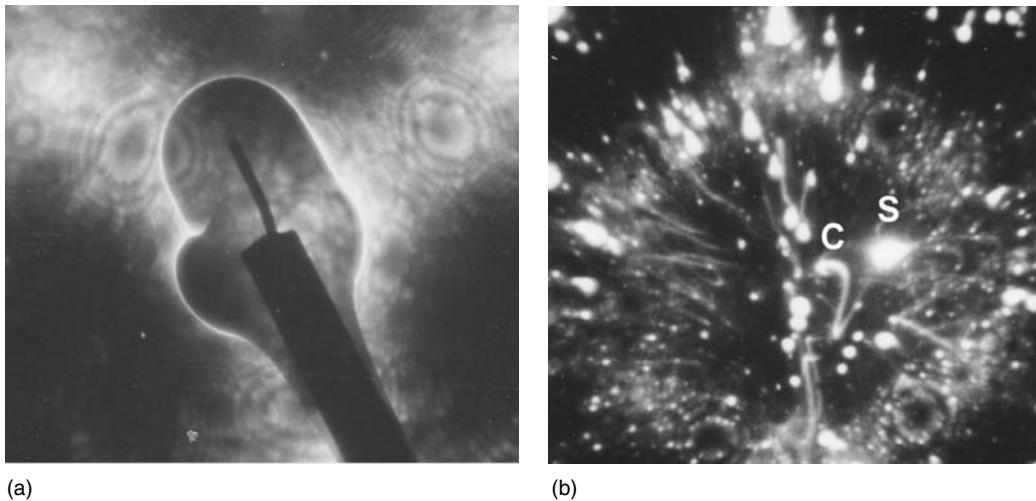
Up till now, the most successful ion source has been the LMIS, widely used in FIB machines. By comparison, a GFIS advantage is that source physics imposes no minimum current, so there is no problem of large minimum virtual source size. In fact, the carefully designed and built-up GFIS may be closer to the perfect point ion source than any competing technology. The GFIS can exceed the LMIS in brightness.<sup>26,27</sup> Another advantage is that  $\text{He}^+$  and  $\text{H}_2^+$  ions are less destructive on impact than metal ions and better suited to scanning ion imaging.

### 3.2.2 THE GAS-PRESSURE TRADE-OFF

Potential technical problems are that emitted current density and angular current density are locally nonuniform and that primary ions might undergo charge exchange and scattering whilst in transit. Getting adequate beam current has been a historical problem for the GFIS. Emission current is proportional to the background gas pressure in the emitter enclosure, but secondary scattering processes go as the square of pressure.<sup>9</sup> Increasing pressure beyond a certain limit may bring the unwelcome design problems of removing fast neutrals and slow ions from the transmitted beam and coping with small primary ion deflections.

Secondary processes are more likely close to the emitter (where the gas concentration is higher) but can occur all along the path in those parts of an ionoptical column not separately pumped. In low-temperature FIM, image spots can be very bright and gas concentrations near the emitter particularly high. Neutral-generated image spots and tails and secondary-ion-generated *comets* are illustrated in Figure 3.3; these effects and *nebulosity* around image spots and general image mistiness have all been observed and explained<sup>7,9</sup> as secondary phenomena related to primary ion transit. It is also known<sup>12,28</sup> that in an old-style FIM if the screen is made positive relative to the electrode surrounding the emitter, then a small electron current flows to the screen if the pressure is sufficiently great. At 78 K, Feldman and Gomer<sup>28</sup> saw such effects at pressures above  $\sim 5 \times 10^{-4}$  Torr ( $7 \times 10^{-2}$  Pa). This current must come from impact ionization events caused by the primary ion beam. An implication is that at high gas pressures, currents measured at the detector may have a secondary ion contribution.

In principle, all types of GFIS are susceptible to these effects. At the time of writing it is not clear how big a problem (if at all) they will pose for GFIS-based FIB-type machines. However, these effects can be reduced by emitter shape design. A GFIS needle should *catch* as much gas as practicable and then ionize it only at a small number of apex sites. This reduces the background pressure needed to provide a given beam current. Suitable shapes should have large values of the parameter  $A_c$  defined in Section 3.5.2 (subject to any trade-offs needed elsewhere).



**FIGURE 3.3** Charge-exchange effects occurring at high imaging-gas pressure (3 mTorr He, near 80 K) (in both cases, ions are deflected sideways by a subsidiary wire electrode several centimeters below the emitter). (a) Formation of image spots by fast neutrals (seen within the bright boundary surrounding the wire electrode) (positive voltage on wire). (b) A comet C associated with a bright image-spot S (wire earthed). The conical tail on S is due to fast neutrals formed very close to the emitter. The comet is due to slow ions: its shape can be explained only if some ions are formed on the detector side of the wire electrode (several centimeters from the emitter). (From R.G. Forbes, 1971. PhD thesis, University of Cambridge.)

### 3.2.3 SECONDARY ELECTRONS

As with the LMIS, positive ion impact onto the emitter enclosure or other surface generates secondary electrons that travel back to the anode, if system design allows, and may confuse emission current measurements. Worse, because extraction voltages are high (often  $\sim$ 20 kV) the electrons may generate X-rays. The emitter support structure can be a significant X-ray source (see Ref. 2, p. 177), and in some circumstances precautions might need thinking about.

## 3.3 TECHNOLOGICAL DEVELOPMENT

Much GFIS development (and before the arrival of channel-plate intensifiers, much FIM development) were explorations of the aforementioned trade-offs. Thus, an aim of the author's PhD work in the mid1960s (with Alan Cottrell and Mike Southon as advisers) was to explore whether operating a He-on-W FIM at 4.2 K would improve gas supply without undesirable secondary effects.\* With the GFIS, initial exploratory work<sup>29–32</sup> supported the viability of microprobe applications. Then attempts to capture enough gas and generate enough beam current began. Attempts were initially made to modify substrate shape to develop brighter  $H_2^+$  sources.<sup>33–35</sup> Attempts were also made<sup>36,37</sup> to extract sufficient current from the tungsten (W) 111 planes of a conventional emitter (these planes emit strongly at some imaging fields). The possibility of using emission from an ultrafine metallurgical precipitate has been explored.<sup>38</sup> Gas has been diffused to an emitter via a porous needle.<sup>39</sup> And an emitter has been attached to a capillary tube down which gas is passed.<sup>40</sup> But the most promising developments involved a needle with a built-up apex that emitted only from one or a few emission sites.

\* The secondary effects were undesirable—in particular we lost most of the image spots, which was not helpful for a microscopy—but much was learnt about how the He-on-W FIM worked.

Three built-up needle shapes have been developed as potential GFIS emitters: the atomically sharp emitter ending in a single atom;<sup>41,42</sup> the supertip,<sup>27,43–45</sup> which is a small field-enhancing protrusion on an underlying Müller emitter; and the pyramidal W(111)-oriented apex ending in three atoms,<sup>26</sup> commercialized by the ALIS Corporation (later, part of Carl Zeiss SMT) as the *Atomic level ion source* (ALIST<sup>TM</sup>). The single-atom-apex emitter, although it works, has made no headway as an ion emitter. The supertip GFIS has been operated with several imaging gases; the ALIS uses He<sup>+</sup> ion emission.

The pyramidal apex shape seems to be superior. This is probably partly because the number of atomic emission sites at the apex is better controlled, partly because the shape is a dynamic equilibrium one for a particular applied voltage and temperature.<sup>26</sup> This means that, like the Binh design for a scanning microscopy probe,<sup>46</sup> if it degrades in operation then it can be renewed *in situ*. Because the most competitive emitter design seems to be this pyramidal apex, illustrative numerical information given here relates to the He-on-W ionization system it currently uses. The next section examines the complex basic physics of GFI devices.

## 3.4 GAS FIELD IONIZATION FUNDAMENTALS

### 3.4.1 PRELIMINARIES

#### 3.4.1.1 Emitter Formation

A sharp needle can be prepared by electropolishing and then by cleaning and smoothing by heating and *field evaporation* in vacuum. The needle tip can be *electroformed* by heating with a high electric field applied: this causes field-induced migration of atoms. The basic driving force is the same as for an LMIS: the change in shape decreases the electrical Gibbs function (Forbes–Ljepojevic thermodynamic potential<sup>47</sup>) by increasing capacitance between the needle and its surroundings. However, details of atomic migration kinetics are important, and fabrication procedures can be complex and proprietary. In the supertip case, the protrusion is formed by field-emitting electrons in the presence of He; He<sup>+</sup> ions formed by electron-impact ionization bombard the emitter and help metal atoms surmount activation energy barriers for migration.<sup>44</sup> Electroforming has a long history<sup>46,48–51</sup> and is also used for preparing field electron emitters<sup>48,50</sup> and scanning probes.<sup>46</sup>

#### 3.4.1.2 Field Definitions

FIM uses the concept of an *applied field*. This refers to the field slightly above the emitter apex, outside the range of the very local field variations close to the surface. FIM image appearance alters strongly with applied voltage.<sup>7,10</sup> The FIM image is sharpest at a *best image voltage* (BIV) when the applied field is equal to the *best image field* (BIF).

For any particular imaging gas, the experimental condition of BIF for a liquid-nitrogen cooled emitter is easily recognized visually and is apparently much the same for all metals. For He-on-W imaging at BIF, experiments by Sakurai and Müller<sup>52</sup> allocated the value 45 V/nm to the field above the W(110) plane. For theoretical consistency with existing calculations, this value is used here, although Sakurai and Müller found the field above the W(111) plane to be slightly higher ( $\sim 50$  V/nm).

For a pyramidal built-up emitter, the applied field can be identified with the maximum field in the critical surface (see Section 3.4.1.3) above the apex site or one of the apex sites. It is provisionally assumed that the GFIS operates under gas conditions similar to FIM BIF, and the value 45 V/nm is used as an estimate of the related applied field. (This assumption will be reexamined later.)

#### 3.4.1.3 The Ionoptical Surface

With GFI, the ion generation formally resembles a chemical reaction. For the He-on-W system, the reaction is the trivial one:  $M_W + He \rightarrow M_W^- + He^+$ , where the symbol M<sub>W</sub> means *tungsten metal*.

The critical surface used in GFI theory is the *crossing surface* for the relevant reaction for the situation where the topmost gas-atom electron makes a radiationless tunneling transition to the metal Fermi level without generation or absorption of heat. This critical surface becomes the GFI *real ionoptical surface* (the surface where ions are assumed to be created).

Physically, the critical surface is not an exact electrostatic equipotential, and there is no physical requirement for electric field lines to cross it normally. However, basic ionoptical models (and associated detailed trajectory analyses) would usually model the critical surface as an electrostatic equipotential. Very small discrepancies, of order 10 meV or less for the He-on-W system, are thereby introduced into the assessment of ion energies. The possibility also exists of localized ionoptical aberrations, but these are currently thought to be small for the He-on-W system and are disregarded here.

### 3.4.1.4 The Real-Source Current-Density Distribution

With GFI, one has to deal with a six-dimensional distribution involving the position and velocity of the gas-atom nucleus. The probability per unit volume of finding the nucleus at position  $X$  (integrated over all velocities and expressed as *atoms per unit volume*) is termed the *gas concentration*,  $C_g(X)$ . Because atom wavelengths are much shorter than electron wavelengths, there can be significant spatial variations (on the scale of tens of picometers) both in  $C_g$  and in the rate-constant for ionization  $P_e(X)$  (called here the *electron tunneling rate-constant*). The product  $j_3(X) = (e/n_1)C_g(X)P_e(X)$  is called the *ionization density*. Here,  $e$  is the elementary charge and  $n_1$  a formal way of writing *one atom*.

Because ionization zones are narrow (of order 20 pm), the three-dimensional (3-D) ionization distribution can be converted to a 2-D distribution of *real-source current density* ( $J_S$ ) in the critical surface by integrating  $j_3$  along a tube of space (of infinitesimal cross section) that starts from the critical surface and passes out through the ionization zone, following the trajectory of a hypothetical ion created at rest at critical surface position  $X^{cr}$ .  $j_3$  falls off with distance along the tube, so the result can be written as  $J_S(X^{cr}) = (e/n_1) C_g(X^{cr}) P_e(X^{cr}) \delta(X^{cr})$ . Here,  $\delta(X^{cr})$  is a decay constant for  $j_3$  for fall-off along the tube and is nearly constant. So the variation in  $J_S(X^{cr})$  is well approximated by the variation in  $C_g(X^{cr})P_e(X^{cr})$ .

## 3.4.2 GAS-ATOM BEHAVIOR

### 3.4.2.1 The Critical Surface and the Firmly Field-Adsorbed Layer

By definition, when a gas-atom nucleus is at the *critical surface*, its topmost electron level aligns with the metal Fermi level. Ionization zones lie just outside the critical surface. For nuclear positions inside this surface, the topmost gas-atom electron level drops below the Fermi level, so there are no vacant states for the electron to tunnel into. So there is a *forbidden zone* between the metal atoms and the critical surface where normal FI cannot happen.

Hence, below  $\sim 100$  K, each protruding metal atom is normally covered by a *firmly field-adsorbed* gas atom.<sup>9,53-57</sup> Bonding is localized and due mainly to strong local polarization forces, with a maximum over the metal atom, but also involves chemical (exchange) effects. These occur because the applied field lifts the topmost gas-atom electron level into the metal conduction band.

In low-temperature GFI, the gas atoms subject to ionization move on top of this field-adsorbed layer, which has three main roles. It reduces the activation energy for (the nucleus of) a fully accommodated imaging-gas atom to get into an ionization zone; reduces the activation energy for an atom to escape from the emitter un-ionized; and provides an intermediate collision partner that helps an incoming *hot* gas atom lose energy more quickly than at a bare surface. This last effect helps to reduce beam energy spread at temperatures below  $\sim 100$  K (see Section 3.4.3.2). The firmly field-adsorbed layer (which is quasi-solid) and the mobile layer (which is gaseous) should be seen as different thermodynamic phases of helium and may have different effective temperatures.

### 3.4.2.2 The Long-Range Polarization Potential-Energy Well

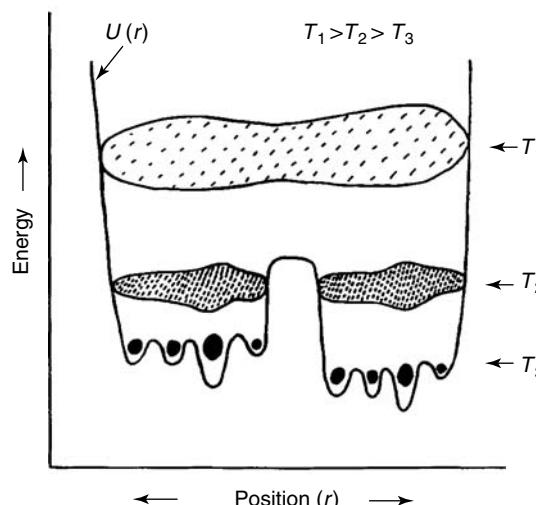
In an electric field  $F$ , away from the surface, a neutral gas atom has a PPE  $U = -(1/2)\alpha_{\text{gas}}F^2$ , where  $\alpha_{\text{gas}}$  is the gas-atom polarizability. For He  $\alpha_{\text{gas}} = 0.143 \text{ mev V}^{-2} \text{ nm}^2$ ; so at BIF (45 V/nm) a He atom coming directly from space to the emitter cap hits the emitter surface with kinetic energy (KE)  $\sim 0.14 \text{ eV}$ . For an emitter surface at temperature  $T_{\text{em}}$ , this is much greater than the equipartition energy  $(1/2)k_B T_{\text{em}}$  (0.013 eV at 300 K and 0.0034 eV at 78 K). So there is a high chance that after its first bounce on the rough emitter surface, the incoming atom will be trapped in the long-range PPE well that surrounds the needle tip and will then lose energy via further bounces. Similar arguments apply to the part of the emitter shank where the local surface field  $F_s$  is sufficiently high.

The PPE well extends into space by a distance equal to several or many cap radii (depending on gas temperature and emitter shape). It can have complex local structure close to the surface. The well helps catch atoms and draw them to the needle. It extends down the side of the needle with the field  $F_b$  at the well edge given by some appropriate capture-oriented criterion of the form  $(1/2)\alpha_{\text{gas}}F_b^2 = \gamma_{\text{gas}}k_B T_{\text{bk}}$ , where  $T_{\text{bk}}$  is the *background* gas temperature (i.e., the gas temperature in the *gas space* enclosing the emitter but well away from it).  $\gamma_{\text{gas}}$  is taken below as 1.

In the PPE well, the neutral-atom concentration is higher than the background value and increases as the needle apex is approached. Emitted ions have to travel through this gas pocket. This is also true for an LMIS though the effect is less pronounced.

### 3.4.2.3 Typical Gas-Atom History

The general gas conditions under which a GFIS operates seem similar (but not identical) to those for an FIM. The typical gas-atom history in these conditions has three main stages: capture, accommodation, and diffusion. (1) The atom is captured on the needle shank and then moves to its tip, heating up as it does so, because it gains KE from the PPE well. Certainly at low needle temperatures, much more of the gas supply is captured by the shank than by the cap (see Appendix A.1). (2) As illustrated in Figure 3.4, this hot captured gas then cools by transferring KE to the substrate when the atoms bounce and accumulates into the higher-field regions above the needle tip. As it cools further, the gas partitions into *confines*, which are extended well-like substructures in the



**FIGURE 3.4** Schematic diagram illustrating how energy accommodation takes place for captured imaging gas funneled to the emitter cap. The gas is initially hot and relatively evenly distributed across the surface. As the gas cools, it *partitions* with more cool gas ending up in the high-field (deep polarization potential-energy well) regions of the emitter.

PPE well, and gas atoms become temporarily trapped in these confines. (3) As the gas becomes fully accommodated, across-surface diffusion takes place relatively close to the surface, initially within confines and then between them, as gas concentrations build toward those characteristic of a thermodynamic equilibrium at  $T_{\text{em}}$ , across the needle tip as a whole. This description is derived from detailed analysis of voltage and temperature dependences in FIM images.<sup>3,7,9,10</sup>

### 3.4.2.4 Best Image Field and Best Source Field

The specific gas condition for FIM BIF occurs toward the end of the accommodation stage. BIF is a trade-off between blurring of the image spots in the (111) regions and the onset of image spots in other image regions.<sup>7</sup> The blurring of the (111) image spots was originally attributed by the present author to *hot ionization* (i.e., FI close to the critical surface but before the gas had fully accommodated), but it is now thought that softening of tunneling rate-constant variations across the surface<sup>4</sup> as field increases may be part of the explanation.

The specific gas condition will be different for operation of a (111)-oriented GFIS. Starting from low field, one needs to increase the field (first to increase the tunneling rate-constant and then to maximize the gas supply) until one of the following happens: either (a) blurring of the current-density variations above the emitter apex increases the size (or reduces the brightness) of the beam spot unacceptably; (b) ionization sets in at non-apex emission sites (robbing the apex sites of part of the gas supply); (c) increase in the energy spread leads to an unacceptable increase in column chromatic aberrations; or (d) the incidence of undesirable secondary processes becomes unacceptable. Restraint (d) can be removed by lowering the gas pressure (but this reduces brightness). The author's guess is that (a) and (b) will set in more or less together before (c), probably with (a) happening first at a point that corresponds to the end of the accommodation stage.

The field at which these *best source conditions* occurs can be termed the *best source field* (BSF) and the corresponding voltage for a particular source the *best source voltage* (BSV). The calculations of Van Eekelen (see Ref. 14, Figure 4) suggest that if BIF is  $\sim 45$  V/nm, then BSF would be  $\sim 40$  V/nm. However, the ALIS is (111)-oriented, rather than the usual orientation of (110) for a tungsten emitter and (certainly for field-evaporated emitter shapes) has a higher field over it. If we use the Sakurai and Müller calibration result<sup>52</sup> that the field over the (111) plane is 50 V/nm when the field over the (110) plane is 45 V/nm, and reduce this 50 V/nm by multiplying by the factor 40/45, then the outcome is 44 V/nm. For present purposes we round this to 45 V/nm and use this as the BSF for an ALIS. (This allocation may need to change if field calibration experiments are carried out.)

### 3.4.3 ION GENERATION

#### 3.4.3.1 Ion Generation Theory

A full theory of gas-atom FI requires wave mechanics for both electron and nuclear motion. However, a quasi-classical approach is customary, which treats electron tunneling by wave mechanics but gas-atom motion classically. This approach, based on the formula  $j_3(X) = (e/n_1)C_g(X)P_e(X)$ , was presented earlier.  $P_e(X)$  depends both on the *average field in the tunneling barrier*  $F_{\text{av}}$ , and on the distance  $x$  of the nucleus from the emitter's *electrical surface*;<sup>58,59</sup>  $C_g(X)$  depends on the local field at  $X$  but also on the gas condition.

A formula attributed to Gomer<sup>1</sup> is often used to calculate the tunneling rate-constant  $P_e$ , but is an approximation of limited accuracy. Other approximations are discussed in Refs 3 and 11, but currently the best treatment is that of Lam and Needs.<sup>60</sup> In principle, all formulas need to be written<sup>3</sup> in terms of  $x$  and  $F_{\text{av}}$  or equivalent parameters, but they may be written formally as

$$P_e = v_e \exp(-\mu b I^{3/2} / F) \quad (3.1)$$

where  $b$  is the second Fowler–Nordheim constant ( $\approx 6.830890 \text{ eV}^{-3/2} \text{ V}^{-1} \text{ nm}$ ),  $I$  the gas-atom ionization energy ( $\approx 24.587 \text{ eV}$  for He),  $v_e$  an *attempt frequency* usually taken as  $10^{15}\text{--}10^{16} \text{ s}^{-1}$ ,  $F$  the local field at the position of the gas-atom nucleus, and  $\mu$  the tunneling exponent correction factor.  $\mu$  is  $\sim 0.6$  but is a function of  $F$  and is a sensitive function of position both across and normal to the surface. These  $\mu$ -variations are responsible both for FIM resolving ability and for the fall-off in  $P_e$  that gives rise to the high-temperature ion energy full-width-at-half-maximum (FWHM). In this simplified form, the necessary distinction between  $F$  and  $F_{av}$  is absorbed into the expression for  $\mu$ . (If  $\mu$  is treated as constant, then Equation 3.1 should be used only for illustrative calculations of how the average tunneling rate-constant varies with applied field.)

Exact calculation of  $C_g(X)$  is operationally impossible due to the complexity of the gas kinetics. But, in circumstances where the gas can be treated very locally as in effective thermodynamic equilibrium at temperature  $T_g$ , the ratio of gas concentrations at two points A and B close to each other (where the local fields are  $F_A$  and  $F_B$ ) is given from Maxwell–Boltzmann statistics as

$$C_{gA}/C_{gB} = \exp\left[\frac{1}{2}\alpha_{\text{gas}}(F_A^2 - F_B^2)/k_B T_g\right] \quad (3.2)$$

A built-up GFIS works because  $F$  and  $P_e$  are always relatively high above the emission sites, as compared with other sites on the needle tip. The GFIS comes into normal operation when the applied field at these sites reaches the BSF, and  $P_e$  at these sites reaches its relevant operating value.

### 3.4.3.2 Energy Spreads

The FWHM of the ion energy distribution for He FI depends on several factors, including local applied field. The FWHM can be as low as 0.29 eV (measured<sup>61</sup> for a W[100] zone-line decoration atom at 45 K). The FWHM at BIF depends on the local environment and is affected by  $T_{\text{em}}$ . Thus, for  $\text{He}^+$  emission from sites on a W(111) plane, the FWHM was reported<sup>61</sup> as 417 meV at  $T_{\text{em}} = 80 \text{ K}$  but 747 meV at 200 K.

The physical FWHM depends on how  $C_g$  and  $P_e$  vary with distance from the critical surface: the factor with more rapid fall-off has greater influence. Fall-off in  $P_e$  was long believed more important, but existing theory<sup>11,62</sup> for  $P_e$  cannot convincingly deliver an FWHM below 1 eV at BIF. Also,  $P_e$  is not temperature dependent. Thus, the aforementioned results suggest that at low temperature fall-off in  $C_g$  determines the physical FWHM. This and a later work<sup>63</sup> link the larger FWHM at higher temperatures to the absence of any firmly field-adsorbed layer at these temperatures. When the layer is present, the imaging gas cools more rapidly during accommodation, so it is colder at ionization and concentrated more near the surface. This leads to a lower physical FWHM.

In the high-temperature limit, the FWHM is determined by fall-off in  $P_e$  and has a measured value of  $\sim 1 \text{ eV}$  (which is compatible with the lowest plausible theoretical predictions<sup>11</sup> for He FI). This should be low enough to avoid severe chromatic aberration effects in FIB-type applications even if operation were at room temperature.

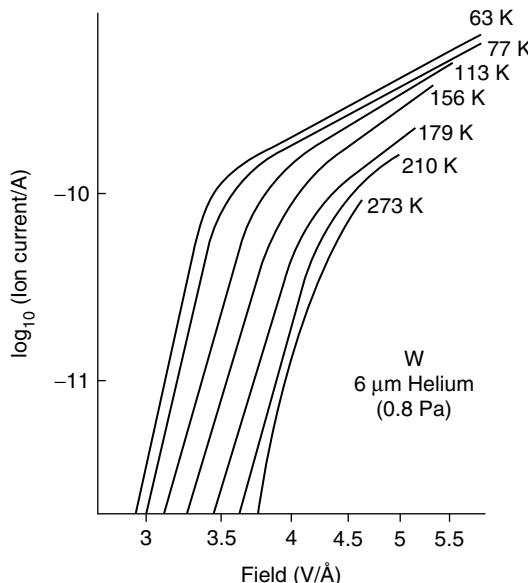
In this discussion, a subtle distinction between the FWHM of the normal energy distribution and that of the total energy distribution has been neglected. This is justified in another section.

An FWHM of 1 eV at a field of 45 V/nm implies that ionization zones have thickness of order 20 pm, much less than a He atom diameter. Ionization occurs when the atomic nucleus is in the zone.

## 3.5 THEORY OF EMISSION CURRENT

### 3.5.1 INTRODUCTION

The total emission current  $i$  is directly proportional to the kinetic-theory equilibrium gas flux density  $Z_{\text{bk}}$  distant from the emitter. Hence,  $i$  is proportional to the background gas pressure  $p_{\text{bk}}$



**FIGURE 3.5** Current/voltage characteristics for helium-on-tungsten (He-on-W) system at various emitter temperatures (measurements by Southon<sup>12,13</sup>). (For these measurements, the background gas temperature  $T_{\text{bk}}$  is now known to have been greater than the emitter temperature  $T_{\text{em}}$ )

and also varies as  $T_{\text{bk}}^{-1/2}$ , where  $T_{\text{bk}}$  is the background gas temperature. For a GFIS, emission current control is normally via adjustment of  $p_{\text{bk}}$ , as the source should be operated at a fixed applied field, which implies operation at a fixed applied voltage.

The current  $i$  varies with applied voltage  $V$ . This can also be regarded as a variation with the applied field  $F$ . Figure 3.5 shows Southon's well-known  $i$ - $V$  characteristics for He,<sup>12,13</sup> taken for different emitter temperatures (but the background gas temperature was, almost certainly, somewhat higher). At low emitter temperature, but not at room temperature, the characteristic has a knee.

Exact calculation of emission current is operationally impossible due to gas-kinetic complexities, but the main features of  $i$ - $V$  characteristics can be reproduced by modeling. It is convenient to write

$$i \approx i_{\text{SAC}}(F, T_{\text{em}}, T_{\text{bk}} \mid \text{capture properties}) \cdot \Pi_i(F, T_{\text{em}}) \quad (3.3)$$

where  $i_{\text{SAC}}$  measures the flux of neutral atoms delivered to the emitting cap at the start of the accommodation stage, expressed as a supply-and-capture (SAC) current (by multiplying the flux by  $e/n_i$ ), and  $\Pi_i$  is the probability that a neutral atom, trapped at the cap at the start of the accommodation stage, will be ionized. Both  $i_{\text{SAC}}$  and  $\Pi_i$  are functions of  $F$  and  $T_{\text{em}}$ ; in addition,  $i_{\text{SAC}}$  depends on emitter shape and surface condition and on the whole thermal environment (if this is not isothermal). Equation 3.3 ignores a very small contribution from FI of gas atoms on individual paths that (if no FI occurred) would involve one impact with the cap surface and then escape directly back into space.

For simplicity, nearly all theories assume that the emitter and background gas temperatures are equal, although this is not normally true in FIM. However, formulas are stated in another section in terms of what physically seems the more appropriate temperature. Note that the effective gas temperature at ionization ( $T_g$ ) may be different from both  $T_{\text{em}}$  and  $T_{\text{bk}}$  (see Section 3.5.4).

### 3.5.2 SUPPLY CURRENT AND EFFECTIVE CAPTURE AREA

The supply current  $i_{\text{SAC}}$  can be written as

$$i_{\text{SAC}} = A_c J_{\text{bk}} \quad (3.4a)$$

where  $A_c$  is the *effective capture area* for the needle.  $J_{\text{bk}}$  is the *equivalent background supply-current density*: this is the current-density equivalent of the background gas flux density  $Z_{\text{bk}}$ , assuming one elementary positive charge per atom, and is given by

$$J_{\text{bk}} \equiv (e / n_1) Z_{\text{bk}} = e (2\pi m k_B T_{\text{bk}})^{-1/2} p_{\text{bk}} \equiv a_g p_{\text{bk}} \quad (3.4b)$$

The parameter  $a_g$  is defined by Equation 3.4b and is constant for given source arrangements.

The most common arrangement for a GFI emitter is to cool it to liquid nitrogen temperature. Strictly, the boiling point of liquid nitrogen is 77.36 K. But there are often temperature differences present between the refrigerant and the emitter and between the refrigerant and the thermal shields surrounding the emitter, and it is common practice to round the temperature up to 78 or 80 K, even when thermal contacts are thought to be good. The reference temperature used here is 78 K. For He at 78 K,  $a_g = a_{78} \approx 2.39 \times 10^4 \text{ A m}^{-2} \text{ Pa}^{-1} = 3.19 \times 10^6 \text{ A m}^{-2} \text{ Torr}^{-1} = 3.19 \text{ pA } \mu\text{m}^{-2} \mu\text{Torr}^{-1}$ . For He at  $T_{\text{bk}}$ ,  $a_g = a_{78} (78 \text{ K})^{1/2} T_{\text{bk}}^{-1/2}$ .

The parameter  $A_c$  has the dimensions of area but is greater than the physical gas-collecting area due to the gas-capturing properties of the long-range PPE well.  $A_c$  is a function of apex field and the thermal environment. Feldman and Gomer's experimental work<sup>28</sup> suggests that for an 80 nm radius emitter at 45 V/nm at  $T_{\text{bk}} = 78 \text{ K}$ ,  $A_c$  might have a value of order of magnitude 1  $\mu\text{m}^2$  (see Appendix A.1).

### 3.5.3 IONIZATION REGIMES

Following Gomer,<sup>1</sup>  $\Pi_i$  can be formally written as

$$\Pi_i = k_i / (k_i + k_e) \quad (3.5)$$

where  $k_i$  and  $k_e$  are hypothetical rate-constants for ionization of the gas atom when at the emitter cap and its un-ionized escape, respectively.

Clearly, expression 3.5 changes (at  $k_i = k_e$ ) from  $\Pi_i \approx k_i/k_e$  for  $k_i \ll k_e$  to  $\Pi_i \approx 1$  for  $k_i \gg k_e$ . These limits define ionization regimes that can be called the *dynamic low-field equilibrium* (DLFE) regime ( $k_i \ll k_e$ ) and the SAC regime ( $k_i \gg k_e$ ).

At very high fields, there is, in principle, a *free-space field ionization* (FSFI) regime in which almost all approaching gas atoms are field ionized before reaching the cap, either on or above the needle shank or in space well above the cap. This FSFI regime cannot be reached with He FI on refractory metal emitters, because substrate field evaporation happens first but can be observed<sup>7,9</sup> with hydrogen FI. For most potential emitter contaminants, such as water and nitrogen molecules, FSFI occurs at fields well below the He BIF/BSF and ensures that contaminants do not reach the cap of an operating GFIS.

The limiting expressions for  $i$  are as follows:

$$i \sim (k_i / k_e) i_{\text{SAC}} = (e / n_1) n_{\text{cap}} k_i \quad (k_i \ll k_e) \quad (\text{DLFE limit}) \quad (3.6)$$

$$i = i_{\text{SAC}} \sim A_c a_g p_{\text{bk}} \quad (k_i \gg k_e) \quad (\text{SAC limit}) \quad (3.7)$$

where  $n_{\text{cap}}$  is the total count of gas atoms present at the cap in the accommodation and diffusion stages at a particular instant of time.

Note that, although they are sometimes used, the detailed expressions for  $k_e$  in the older FIM textbooks<sup>1,2</sup> are fundamentally incorrect, because the underlying model violates the principle of conservation of energy (by assuming that un-ionized escape by hopping does not require activation energy). The limiting formulas 3.6 and 3.7 are both formally correct, but for Southon's results there is a separate difficulty concerning  $n_{\text{cap}}$ .

This difficulty is that the slopes of the low-field steep parts of Southon's  $i$ - $V$  characteristics do not exhibit the temperature dependence that simple theory predicts. Jousten et al.,<sup>64</sup> who carefully investigated GFIS  $i$ - $V$  characteristics, also noted this. In measurements on He emission from a W super-tip, Böhringer et al.<sup>44</sup> obtained results similar to Southon's when the gas temperature  $T_{\text{bk}}$  was held constant but the emitter temperature  $T_{\text{em}}$  was altered. But, when  $T_{\text{bk}}$  and  $T_{\text{em}}$  were altered together, the expected temperature dependence in the slope of the  $i$ - $V$  characteristic was present.<sup>44,64</sup>

The present author used Southon's microscope in the 1960s. Almost certainly, Southon's results were strongly influenced by poor thermal contact between the refrigerated glass finger and the loosely fitting slide-on copper thermal shield surrounding the emitter. This poor contact caused  $T_{\text{bk}}$  to be significantly greater than  $T_{\text{em}}$ . Also, Feldman and Gomer<sup>28</sup> suggested this as the origin of discrepancies between their results and Southon's. (See also the discussion in Ref. 2, pp. 36–40.)

The implication of these results seems to be that if  $T_{\text{em}} \neq T_{\text{bk}}$ , then in the  $k_i \ll k_e$  regime the mean gas concentration at the emitter surface ( $n_{\text{cap}}$ ) is in dynamic rather than thermodynamic equilibrium with the background gas concentration. Thus, it is not valid to determine  $n_{\text{cap}}$  by applying the usual Maxwell–Boltzmann equilibrium formula with  $T$  set equal to  $T_{\text{em}}$ . At the time of writing, it is not clearly known how to calculate  $n_{\text{cap}}$  in a nonisothermal environment. But this should not affect GFIS theory because the GFIS operates in the SAC regime.

### 3.5.4 GAS TEMPERATURE AT IONIZATION

Van Eekelen's<sup>14</sup> detailed (but effectively one-dimensional [1-D]) numerical modeling of the gas dynamics found that the knee of the  $i$ - $V$  characteristics corresponds to the end of the accommodation stage, that BIF occurs in the SAC regime, and that at BIF the gas is hot at the point of ionization (i.e., at the point in neutral gas history at which ionization occurs). This last conclusion was also reached independently both by Chen and Seidman<sup>65</sup> (from quantitative analysis of image-spot size dependence on emitter temperature) and by Forbes<sup>7</sup> (from qualitative analysis of FIM image contrast).

Castilho and Kingham<sup>4</sup> subsequently reinterpreted the experiments of Chen and Seidman to suggest that the gas temperature ( $T_g$ ) at BIF was much lower than hitherto assumed (they estimated between 100 and 120 K for an emitter at liquid nitrogen temperature 77 K). But the discrepancy with the Van Eekelen analysis (which found that at BIF the accommodation is 77% complete, i.e.,  $T_g \approx 400$  K) has never been fully resolved. Part of the explanation may be the 1-D nature of Van Eekelen's analysis; Southon's analysis<sup>12</sup> showed that 3-D trapping effects were important. For a liquid-nitrogen-cooled emitter at FIM BIF, the present author now favors a  $T_g$ -estimate closer to 100 K than 400 K.

To allow for continuing uncertainty, data are given in another section for both 78 and 300 K. These figures represent the behavior at liquid nitrogen temperature and at room temperature of a GFIS at BSF. The behavior of an FIM at BIF is presumed to lie between the limits shown but much closer to the 78 K values than to the 300 K values.

### 3.5.5 SUPPLY-AND-CAPTURE REGIME

The effective capture area  $A_c$  can also be written as

$$A_c = 2\pi r_a^2 \sigma_{\text{Mc}} \quad (3.8)$$

where  $r_a$  is the radius of the assumed hemispherical emitter apex and  $\sigma_{Mc}$  is the (total) captured-flux enhancement factor for the Müller emitter. Southon, in an extensive unpublished analysis in his PhD thesis,<sup>12</sup> derived complex approximate formulas for  $\sigma_{Mc}$ ; this derivation is outlined (and corrected) in Appendix A.1. In another approach, Van Eekelen numerically analyzed a detailed gas-kinetic model that used a field-distance relationship derived from hyperboloidal geometry, as already noted. Recently, Liu and Orloff<sup>66</sup> offered a third approach, based on the idea that the gas population on the shank is in approximate thermodynamic equilibrium with the background gas population. All models can account for the experimental field dependence of emission current in the SAC-limited regime though in slightly different ways.

Unfortunately, applying the older models, especially to pyramidal emitters, is problematic at the time of writing. The models suggest trends but have poorly tested predictive power. Currently, it seems best to treat  $A_c$  as an empirical figure of merit, collate empirical values, and look for empirical relationships between  $A_c$  and emitter shape parameters (e.g., the apex radius and taper angle before buildup). Note that valid values of  $A_c$  can be measured only when the apparatus is measuring the whole of the emission current from the emitter and includes no secondary contributions.

### 3.6 BASIC REAL-SOURCE DATA

This section provides data relating to emission-site size and FIM resolution criteria. As long as the crystal facet is not too large, a conventional FIM can just resolve the atoms in the W(111) net plane. The smallest interatom spacing in this plane is  $s_{111} = 447.6$  pm and the corresponding *half-separation*  $\rho_{sep} = 223.8$  pm. Consider a hypothetical spherical field emitter of radius  $r_a = 80$  nm: a small feature of length  $\rho$  in its surface subtends an angle  $\alpha$  at its center given by

$$\alpha = \rho / r_a \quad (3.9)$$

The *angular half-separation* corresponding to  $\rho_{sep}$  is  $\alpha_{sep} = 2.80$  mrad.

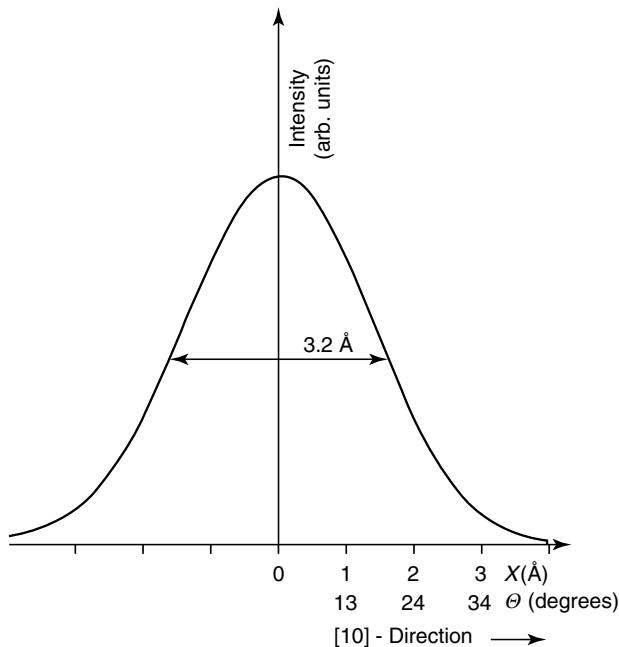
#### 3.6.1 EMISSION-SITE RADIUS

Witt and Müller<sup>15</sup> have measured intensity profiles from field ion images. Where individual image spots can be resolved, they find that the profiles have the shape of a *rounded triangle with small tails* (see Figure 3.6). This suggests modeling a single-atom emission site on the W(111) facet as a disc of radius  $s_{111}/4$  in the critical surface, centered over the atom. This disc contains ~50% of the site current. For an 80 nm radius emitter, the radius and angular radius of the disc would be 112 pm and 1.40 mrad, respectively. Similarly, a three-atom cluster such as that shown in Figure 3.7 can be modeled as a disc of radius  $(1/2 + 1/\sqrt{3})s_{111}/2 \cong 0.539s_{111} \cong 241$  pm, and angular radius 3.01 mrad.

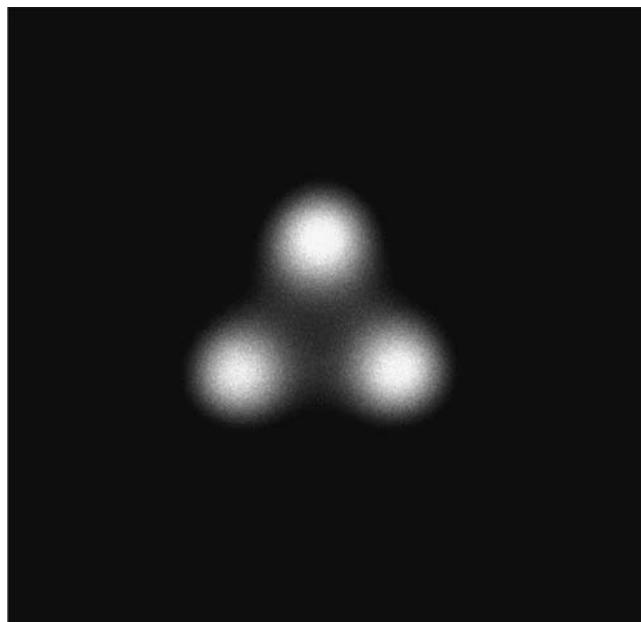
More generally, one can define a *fall-off radius* (FORD)  $\rho_0$  and a *fall-off half-angle* (FOHA)  $\alpha'_s$  about the central axis of an emission site, such that the related radius or cone contains some defined proportion (e.g., 50 or 90%) of the site current. The radii and half-angles just derived provide estimates appropriate to an emitter of radius 80 nm and a criterion of 50%.

These can be compared with the estimated site FORD for protruding single atoms. Thus, for a W atom on top of a (110) plane, Witt and Müller deduced a FORD of 140 pm by deconvoluting experimental measurements of the intensity variation and a FORD of  $135 \pm 10$  pm from the temperature dependence of spot size.

Theoretical estimates by Castilho and Kingham,<sup>4</sup> based on a hemisphere-on-a-plane electrostatic model, give FORD (at 45 V/nm) ranging from 117 to 167 pm (for hemisphere radii ranging from 137 to 250 pm). Taken as a whole, given the approximate nature of the estimates, the picture is fairly consistent.



**FIGURE 3.6** Measured helium (He) ion intensity profile (at the detector) across the field ion microscopy (FIM) image spot corresponding to an isolated tungsten atom on top of a W(110) facet at an emitter temperature of 63 K. (Courtesy J. Witt and K. Müller, preprint, later published as Ref. 15.)



**FIGURE 3.7** Emission from the three atoms in the W(111) facet on the top of the electroformed pyramidal apex of an atomic level ion source<sup>26</sup> (ALIS). (Reproduced from Carl Zeiss SMT, from their Web site. With permission.)

### 3.6.2 FIELD ION MICROSCOPE RESOLUTION CRITERION

For the W(111) facet, let  $A$  and  $M$  label points on the critical surface directly above an atom and directly above a point midway between two atoms that are a distance  $s_{111}$  apart. From GFI theory,<sup>3</sup> the rate-constant ratio  $P_{eA}/P_{eM}$  has been estimated theoretically as  $\sim 2.7$ . Both at 78 and 300 K, the gas-concentration ratio is close to unity, so the current-density ratio  $J_{SA}/J_{SM}$  is  $\sim 2.7$  in both cases. This is much more than sufficient for adjacent atoms in the (111) net plane to be resolved in the current-density variations in the critical surface. However, this signal is blurred by transverse-velocity effects (see Section 3.7.1.3). To assess whether blurring will prevent FIM from resolving atoms, we provisionally use the criterion that the FOHA for the blurred emission site ( $\alpha'_T$ ) has to become greater than the angular half-separation of adjacent emission sites. Given the assumptions mentioned earlier, this implies the test  $\alpha'_T > 2\alpha'_s$ .

### 3.6.3 ILLUSTRATIVE VALUES

To illustrate the application of formulas, numerical data will be provided in another section for a single-atom emission site on the W(111) facet, assuming  $r_a = 80$  nm. To allow consistency of calculation, illustrative values of relevant parameters will be given to three significant figures. The accuracy to which these parameter values are actually known is much less and never seriously better than 10%, if that.

## 3.7 THE SPHERICAL CHARGED PARTICLE EMITTER

### 3.7.1 BASIC IDEAS

#### 3.7.1.1 The Optical Model

In their standard textbook on electron optics, Hawkes and Kasper<sup>18</sup> discuss electron emission from Müller emitters. The approach here builds on theirs but uses slightly different notation. Figure 3.8 shows the two-stage optical model used. This comprises (1) a hypothetical spherical charged particle emitter (SCPE) of radius  $r_a$  fitted to the real emitter's apex and (2) a weak lens. The SCPE forms images of the emission site (and of itself); the weak lens then transforms the SCPE behavior to that of a Müller emitter. We discuss basic issues using the SCPE and then adjust numerical values via the lens. Some aberrations have to be dealt with separately.

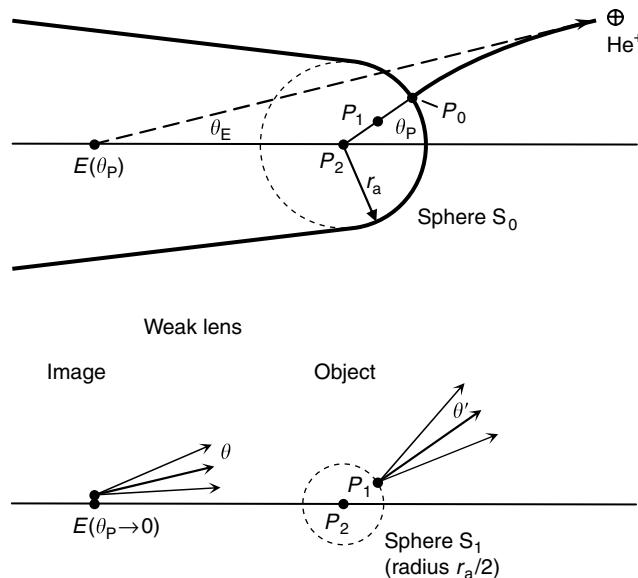
The hypothetical SCPE is an extraordinary optical element, key to the behavior of Müller emitters and related machines, but with no analogy in photon optics. It has been analyzed by Ruska,<sup>19</sup> Gomer,<sup>1,20</sup> Everhart,<sup>21</sup> and Hawkes and Kasper<sup>18</sup> in various contexts, but no treatment is complete for ion emitters. In what follows, the word *ion* is used, but electrons behave the same unless indicated otherwise. Angles and parameters that relate to the SCPE are shown primed; those relating to the weak lens are shown double-primed; and the corresponding angles and parameters that relate to the Müller emitter as a whole are shown unprimed.

#### 3.7.1.2 Ions on Radial Trajectories

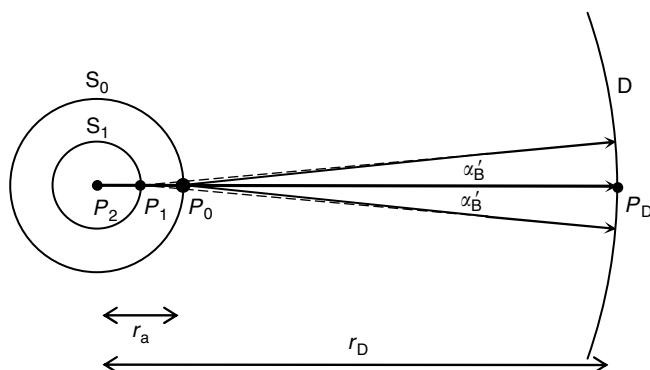
Figure 3.9 shows an SCPE  $S_0$  surrounded by a detector sphere  $D$ . An ion emitted radially from point  $P_0$  on  $S_0$  travels radially to point  $P_D$  on  $D$ . If all ions traveled radially, the emission would appear to come from a virtual point source at the center  $P_2$  of  $S_0$  with an appropriate angular intensity distribution, and the current-density distribution arriving at  $D$  would be proportional to that ( $J_S$ ) leaving SCPE. These are the *unblurred distributions*.

#### 3.7.1.3 The Effect of Transverse Velocity

In reality, emitted ions have transverse KE  $\kappa$ . Motion in a central force field conserves angular momentum; after a radial distance of about  $10r_a$  the transverse KE is largely converted to radial KE,



**FIGURE 3.8** Schematic diagram (not to scale) illustrating the optical behavior of an ideal Müller emitter. The emitter has a quasi-cone-like shape with its apex modeled as a section of an exact spherical charged particle emitter (SCPE). A real source is shown at point  $P_0$  on the emitter apex. The SCPE forms an image of this at point  $P_1$ . The effect of the cone-like shank is represented as the behavior of a weak lens: this is considered to form an image of  $P_1$  in the Gaussian image plane, that is, in the plane of point  $E(\theta_p \rightarrow 0)$ . The weak lens compresses angles and reduces the lateral size of extended objects (thus, the image of  $P_1$  is closer to the optical axis than  $P_1$  is). In reality, the distance  $P_2$  to  $E(\theta_p \rightarrow 0)$  is much greater than this diagram suggests.



**FIGURE 3.9** Schematic diagram (not to scale) illustrating the optical behavior of a spherical charged particle emitter (SCPE); for simplicity, an on-axis beam of ions is used. The ions emerging from a point  $P_0$  on the real emitting sphere  $S_0$  (of radius  $r_a$ ) are blurred out into a narrow cone-like pencil of ions that (when it reaches the detector  $D$ ) appears to diverge from point  $P_1$  on sphere  $S_1$  (of radius  $r_a/2$ ). The blurring is due to the transverse ion velocity distribution at emission. The cone of fall-off half-angle (FOHA)  $\alpha'_B$  contains a specified percentage of the ions from  $P_0$ ; a notional value of 50% is used here. For clarity, angles are exaggerated. See Table 3.3 for typical values of  $\alpha'_B$ .

and effectively the ions then travel in straight lines. (The trajectory actually approximates to a hyperbola.) Solution of the equation of motion shows that ions emitted from  $P_0$  appear to diverge from a point  $P_1$  on a sphere  $S_1$  of radius  $r_a/2$ . Ions with mean transverse KE  $\kappa_{av}$  have trajectories with a limiting cone of half-angle  $\alpha'_B$ , as shown in Figure 3.9, called here the *blurring cone*. For the GFIS at BSF, the gas is taken to be in thermal equilibrium with the emitter, with  $\kappa_{av} = k_B T_g = k_B T_{em}$ .

For the FIM at BIF, where the gas is not in thermal equilibrium with the emitter,  $T_g$  is defined by  $k_B T_g \equiv \kappa_{av}$ .

In the equation of motion, an important factor is the field in the vicinity of the emitter apex (the higher the field, the greater the acceleration, and the less time there is for the ion to move sideways). As the SCPE is an embedded part of a model for a Müller emitter, the field  $F_a$  at the SCPE apex must be written as<sup>18</sup>

$$F_a = V / k_a r_a \quad (3.10)$$

where  $V$  is the applied voltage and  $k_a$  a *shape factor*. For an isolated sphere  $k_a = 1$ . For real Müller emitters usually  $5 < k_a < 7$ . In the present case, where we take  $V = 20$  kV,  $r_a = 80$  nm,  $F_a = 45$  V/nm, self-consistency requires that we take  $k_a \approx 5.56$ . When discussing an SCPE fitted to a Müller emitter apex, the shape factor for the Müller emitter must be put into relevant formulas.

### 3.7.1.4 The Blurred Beam

The blurring-cone FOHA  $\alpha'_B$  is given, to good approximation, by

$$\alpha'_B(T_g) = 2(k_a \kappa_{av} / eV)^{1/2} = 2(k_B T_g / eFr_a)^{1/2} \quad (3.11)$$

Illustrative values are  $\alpha'_B(78$  K) = 2.73 mrad and  $\alpha'_B(300$  K) = 5.36 mrad. In the broadened beam, the defined proportion of the site current (e.g., 50%) now lies within a cone of FOHA  $\alpha'_T$  with its apex at a point  $P_3$  that lies between  $P_1$  and  $P_2$ .  $\alpha'_T$  can be estimated by the usual quadrature formula

$$\alpha'_T = \{(\alpha'_S)^2 + (\alpha'_B)^2\}^{1/2} \quad (3.12)$$

Illustrative values, using this formula, are  $\alpha'_T(78$  K) = 3.07 mrad and  $\alpha'_T(300$  K) = 5.54 mrad. In fact, this formula is not a good approximation because the unblurred beam profile is not Gaussian, and (in FIM) the gas is not in thermal equilibrium. In principle, convolution should be used.

## 3.7.2 OBJECTS AND MACHINES

### 3.7.2.1 Optical Objects Generated by the Spherical Charged Particle Emitter

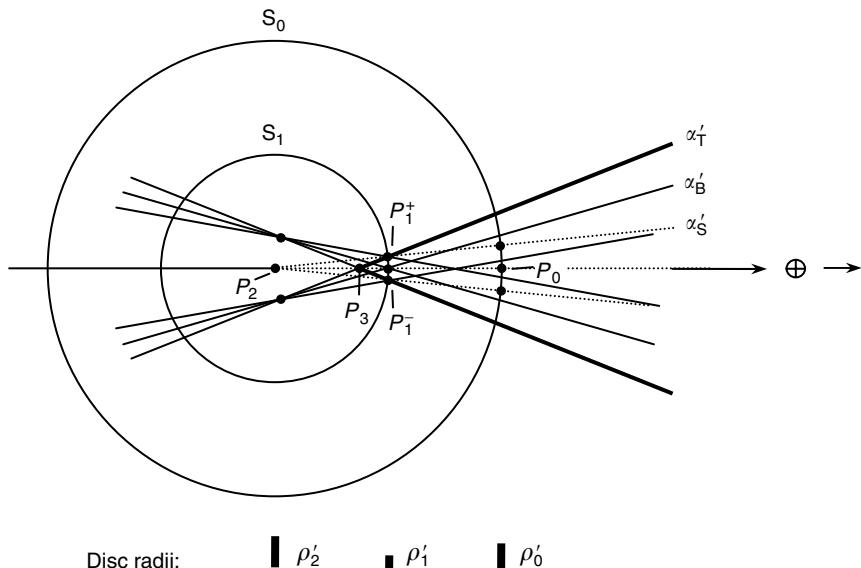
In Figure 3.10, the blurring-cone generators are inserted for point  $P_1$  and also for the points  $P_1^+$  and  $P_1^-$  that mark the boundaries of the original real object, and all resulting rays are projected backward beyond the sphere center. Two discs of minimum confusion result, one of radius  $\rho'_1 = \alpha'_S r_a / 2$  centered at  $P_1$  (usually called the SCPE *Gaussian image* but called here the *first conceptual object*), and another of radius  $\rho'_2 = \alpha'_B r_a / 2$  centered at the SCPE center  $P_2$  (usually called the SCPE *crossover* but called here the *second conceptual object*). The first of these is the physical image of the real object of radius  $\rho_0 = \alpha'_S r_a$  on  $S_0$ , so the SCPE has exerted a *transverse magnification*  $M'$  defined as and given by

$$M' \equiv \rho'_1 / \rho_0 = 0.5 \quad (3.13)$$

The *angular magnification* of the SCPE (defined for radially projected ions) is  $m' = 1$ .

A parameter of interest is the *blurring ratio*  $m'_B$  given by

$$m'_B = \alpha'_B / \alpha'_S = \rho'_2 / \rho'_1 = (2 / \rho_0)(k_B T_g r_a / eF_a)^{1/2} \quad (3.14)$$



**FIGURE 3.10** Schematic diagram illustrating the formation of *conceptual optical objects* by a spherical charged particle emitter (SCPE). The real object lies on sphere  $S_0$  with notionally 50% of the current emitted from a disc of fall-off radius (FORD)  $\rho_0'$  that subtends a fall-off half-angle (FOHA)  $\alpha'_S$  at  $P_2$ . Owing to transverse-velocity effects (with blurring-cone FOHA  $\alpha'_B$ ), there is an apparent object (the *first conceptual object*, or *Gaussian image*) on sphere  $S_1$  of FORD  $\rho_1'$  ( $=\rho_0'/2$ ) with notionally 50% of the emission current lying within a cone of FOHA  $\alpha'_T$  apparently projected from point  $P_3$ . In addition, there is disc of minimum confusion (of FORD  $\rho_2'$ ) in the vicinity of point  $P_2$  (the *second conceptual object* or *crossover*). This diagram shows the case where  $\alpha'_B > \alpha'_S$  and  $\rho_1' < \rho_2'$ . To allow detail to be shown, angles are exaggerated.

This parameter,  $m'_B$  is relevant both to the resolution of atoms in the FIM and to the focusing of a GFIS-based FIB machine. Illustrative values are  $m'_B(78\text{ K}) = 1.95$  and  $m'_B(300\text{ K}) = 3.83$ .

A *blurring magnification*  $m'_T$  can be defined by

$$m'_T = \alpha'_T / \alpha'_S \quad (3.15)$$

This parameter relates to the increase in spot size in the FIM. Using the quadrature approximation, illustrative values are  $m'_T(78\text{ K}) = 2.19$  and  $m'_T(300\text{ K}) = 3.96$ .

### 3.7.2.2 Machines Based on the Müller Emitter

In principle, the SCPE and Müller emitter give rise to microscopes/machines of three general kinds. Machines of the first kind are the field electron microscope and FIM. These are projection machines that look at the current density in a plane some distance from the emitter. These current densities are proportional to those existing in the plane of the aperture in machines of the second and third kinds and so provide information of interest to their machine optics. The field ion image spots are blurred, magnified versions of the real objects on  $S_0$ , obtained by projecting the corresponding first conceptual objects (Gaussian images) on  $S_1$  via the weak lens.

Machines of the second kind include electron microscopes and SIMs (including FIB machines). Here, the normal procedure is to use the second conceptual object as a *nearly point bright source* of FORD  $\rho_2'$ . The blurred FOHA ( $\alpha'_T$ ) of the projected first conceptual object (as seen in FIM) defines a *natural angular aperture*. This is more significant for ion emitters, as  $\alpha'_S$  and  $\alpha'_T$

are normally much larger for electron emitters. But beam acceptance half-angles will normally be less than  $\alpha'_T$ , for both electron and ion emitters.

In machines of the third kind, the first conceptual object would be focused to give (in principle) a magnified and unblurred image of the real object, thereby displaying information about how the rate of the chemical reaction noted earlier varies with position in the critical surface. As far as the author is aware, ion machines of this kind have never been specifically considered (and they might have severe problems with column aberrations and be difficult to focus correctly). But they are of theoretical interest because if operated at sufficiently low temperatures they might be able to act as *vibrational wavefunction magnifiers*. (That is, they would magnify the probability-current variations determined by the transverse component of the departing ion's wavefunction.) The focusing of emitted ions, irrespective of transverse velocity at emission (except as constrained by the beam acceptance aperture and aberrations), means that we can get (nearly) precise information without violating the Heisenberg uncertainty principle.

### 3.7.2.3 Field Ion Microscope Resolving Power

If each of two adjacent emission sites has unblurred FOHA  $\alpha'_S$  and total FOHA  $\alpha'_T$  and has angular half-separation  $\alpha'_{sep} = 2\alpha'_S$ , then prediction is that the beams will seriously overlap if  $m'_B > \sqrt{3} \cong 1.73$ . Illustrative values derived previously were  $m'_B(78 \text{ K}) = 1.95$  and  $m'_B(300 \text{ K}) = 3.83$ . This suggests that there should be resolution difficulties. In practice, atoms on W(111) net planes can just be resolved by conventional FIM near 80 K if the plane is small enough; so probably a more sophisticated theory, based on convolution and a more sophisticated resolution criterion, is needed.

Nevertheless, Equation 3.14 shows correctly that (a) for given apex field and interatom half-separation  $\rho_0$ , blurring is reduced by using lower operating temperatures and smaller emitters (factors key to the invention of FIM), and (b) for given values of  $T_g$  and  $r_a$  there is a minimum interatom separation that can be resolved.

For Equations 3.13 and 3.14, the various values discussed previously have been compiled into Table 3.3, in Section 3.11.

## 3.7.3 SOURCE SIZES AND RELATED TOPICS

### 3.7.3.1 Virtual Source for a GFIS-Based Machine

From Equations 3.13 and 3.14, the FORD of the second conceptual object (crossover) is

$$\rho'_2 = m'_B M' \rho_0 = (k_B T_g r_a / e F_a)^{1/2} \quad (3.16)$$

Note that  $\rho'_2$  depends on the emitter radius but not on the FORD  $\rho_0$  of the real object. Illustrative values are:  $\rho'_2(78 \text{ K}) = 91 \text{ pm}$  and  $\rho'_2(300 \text{ K}) = 214 \text{ pm}$ . These should be compared with the FORD of the first conceptual object (Gaussian image),  $\rho'_1 = M' \rho_0 = \rho_0/2 = 56 \text{ pm}$ : the second object is larger. Thus, the GFIS is different from the field electron emission case, where usually  $\rho'_2 \ll \rho'_1$ . If the rest of the optics were good enough, the operator could choose either (a) to focus on the Gaussian image and work with a small spot with (possibly) significant intensity variations across it or (b) to focus on the crossover and work with a larger, more uniform spot. However, in the present stage of machine design, it seems that column aberrations will have the greater influence on the final beam spot size, and it is not easy to predict what the best working arrangements will turn out to be (or they might be different for different applications).

As  $\rho_0$  and  $F_a$  are fixed by earlier choices, Equation 3.16 suggests that  $\rho'_2$  can be reduced by using smaller emitters and by working at lower temperatures; further,  $\rho'_2$  becomes smaller than  $\rho'_1$  if  $m'_B < 1$ . As column aberrations are technical challenges (and their influence might, e.g., be reduced by software-based methods), it is of interest to explore what factors place a lower limit on  $\rho'_2$ .

### 3.7.3.2 Transverse Zero-Point Energy

$\alpha'_B$  cannot be decreased indefinitely by working at a lower temperature, because a lower limit is set by the wave mechanical behavior of the nucleus. (Strictly, it is the behavior of the ion core as represented by the position of the nucleus.) Because the ionization discs in the critical surface are of limited size, the transverse component of the vibrational wavefunction must have a peaked transverse shape during the ionization process, and this has associated KE. This is equivalent to the existence of transverse zero-point energy and imposes a lower limit  $k_B T_0$  on the transverse ion KE at the point of ionization and thus on  $\alpha'_B$ . There is a slightly higher temperature  $T_1$  at which transverse wave mechanical effects need to be considered to estimate  $\alpha'_B$  well. This is an old problem, closely related to the issue of minimum spot size in an FIM. For the He-on-W system,  $T_1$  has been estimated<sup>9</sup> as not greater than  $\sim 40$  K (by assuming the nucleus moves in a square box of side 200 pm). So this effect should have little influence on a GFIS operating near 80 K or higher. The corresponding temperature  $T_0$  is 16 K. (Obviously, these are rough estimates.)

In the literature, this effect has been described as *associated with the uncertainty principle*. (Also, some treatments add thermal blurring and uncertainty principle terms in quadrature, but this seems a poor approximation, certainly when  $T_g$  is many times  $T_1$ .)

### 3.7.3.3 Minimum Value for $\rho'_2$

The other factor that limits the crossover FORD  $\rho'_2$  is the emitter radius  $r_a$ . For  $r_a = 80$  nm,  $T_0 \approx 16$  K yields  $\rho'_2 = 50$  pm.  $\rho'_2$  becomes lower if  $r_a$  is lower, but it is unclear whether the optical theory used here is valid for extremely low radii. A provisional approach is to allocate a hypothetical minimum effective value to  $r_a$ , say 10 nm. Taking  $T_0 \approx 16$  K then yields  $\alpha'_B \approx 3.50$  mrad, and  $\rho'_2 \approx 18.0$  pm. In practice, due to column aberrations, there seems little possibility of reaching values as low as this. But this figure demonstrates that (unlike LMIS-based machines) GFIS-based FIB-type machines will not be limited by the source size. It suggests that a well-built GFIS could be closer to an ideal point source than any obvious competing technology and underlines the importance of reducing source and column aberrations and any other technical problems, such as mechanical vibration.

### 3.7.3.4 Effect on Total Energy Distribution

With an SCPE, the conversion of transverse to radial KE means that a measured energy distribution will always be a total energy distribution. With a field electron emitter, this has a dramatic effect, which was first noted by Young.<sup>67</sup> With a field ion emitter, where the dominant origin of the energy spread is the distribution of ionization at various distances from the surface, the energy spread due to this effect (typically 0.25–1 eV) will not be markedly affected by the addition of a small transverse thermal KE (of order 0.007 eV at 78 K and 0.025 eV at 300 K).

### 3.7.3.5 Other Effects of Ion Energy Spread

Energy spreads due to a spread of ionization normal to the surface are equivalent to having a range of values for  $r_a$ . For the GFIS, this range is  $\sim 20$  pm. Any resulting blurring effects should be small for the SCPE, though there will be chromatic aberration effects in the column.

## 3.8 THE ROLE OF THE WEAK LENS

### 3.8.1 INTRODUCTION

As already noted, the emitter shank has an optical influence that can be represented as a weak lens. This lens forms virtual images (e.g., near E in Figure 3.8) of the conceptual objects produced by the SCPE (e.g., at  $P_1$ ), and these images then serve as virtual objects for the FIB or SIM column. Points near E also serve as projection centers in FIM.

As with the SCPE, this lens has special properties, because CPs and photons obey different physical laws. It also has aberrations. *Ideal Müller emitters* should exist that have optimum properties for the formation of paraxial beams. Ideally, this section should define all these properties and assess residual aberrations. This enterprise has only just begun; this section reports some useful basic progress.

### 3.8.2 ANGULAR MAGNIFICATION

Suppose that, in its central region, the Müller emitter apex resembles a sphere of radius  $r_a$ . Consider the trajectories of CPs (*ions*) that leave the surface normally, at positions defined by angle  $\theta_p$  as shown in Figure 3.8. These trajectories are bent *inward* (toward the optical axis) by the field configuration but more so near the emitter where the field is strongest and the ions are slowest. Away from the emitter (beyond a distance often thought of as  $10r_a$ ), the ions effectively travel in straight lines, at angle  $\theta_E$  to the optical axis. These trajectories may be projected back to intersect the optical axis at a point  $E(\theta_p)$ . The position of  $E(\theta_p)$  depends slightly on  $\theta_p$ ; in the lower diagram in Figure 3.8, the limiting point  $E(\theta_p \rightarrow 0)$  defines the Gaussian image plane.

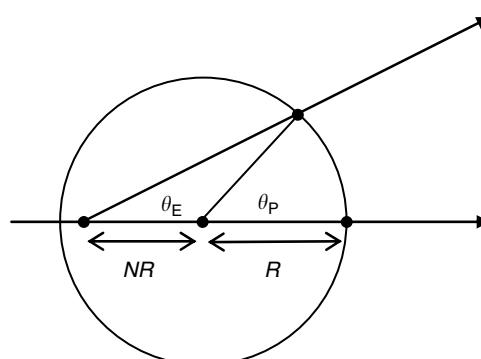
For the Müller emitter, a *Müller-type angular magnification*  $m(\theta_p)$  is defined by the *linear projection condition*

$$\tan \theta_E = m(\theta_p) \cdot \theta_p \quad (3.17)$$

This is equivalent to a well-known formula<sup>68</sup> used in FIM. However, there are at least two other ways of defining angular magnification, namely, the photon-optics method based on  $\tan \theta_E = m_H(\theta_p) \cdot \tan \theta_p$  and the *angular ratio method* based on  $\theta_E = m_W(\theta_p) \cdot \theta_p$  ( $m_W$  is the Wiesner and Everhart<sup>22</sup> parameter  $K$ ). The issue here is one of principle: in CP optics which of these parameters ( $m$ ,  $m_H$ ,  $m_W$ ) is more nearly constant as a function of angle  $\theta_p$ ? All three methods are, of course, equivalent for very small angles of  $\theta_p$ , but they differ for large angles. Very large projection angles are used in FIM, so the issue of which form of definition is appropriate is significant.

A central experimental fact of the crystallographic analysis of field ion micrographs is that (on average) the Müller-type angular magnification  $m$  is very nearly constant for values of  $\theta_p$  up to  $\sim 50^\circ$ .<sup>6,17,69</sup> In practice in FIM, the quantity cited is not  $m$  but a parameter  $N$  that relates to the limiting position (as  $\theta_p \rightarrow 0$ ) of the apparent crystallographic projection center for small facets that appear at known values of  $\theta_p$  (see Figure 3.11). This is related to  $m$  by

$$m = 1/(1 + N) \quad (3.18)$$



**FIGURE 3.11** Schematic diagram illustrating the definition of the parameter  $N$  used in the crystallographic analysis of field ion micrographs.

This near-constancy of  $m(\theta_p)$  is also found in the numerical trajectory analyses of Smith and Walls<sup>16</sup> (for ions, using a hyperboloidal emitter).

Hawkes and Kasper<sup>18</sup> report similar results for electrons, using Equation 3.17, and note that if the beam from a Müller emitter is to have good optical quality, then the Abbe sine condition must also hold. So it is also required that

$$\sin \theta_E = m(\theta_p) \cdot \sin \theta_p \quad (3.19)$$

Hawkes and Kasper then expand the trigonometric functions in Equations 3.17 and 3.19 to third order to yield a condition for optimum behavior:

$$(\theta_E - \theta_E^3/6) / (\theta_p - \theta_p^3/6) \approx (\theta_E + \theta_E^3/3) / \theta_p \quad (3.20)$$

This gives  $m \approx \theta_E/\theta_p \approx 1/\sqrt{3} \approx 0.578$ . So an ideal Müller emitter should have  $m \approx 1/\sqrt{3}$ .

In FIM, a typical experimental value<sup>17</sup> is  $N = 0.83$ , which yields  $m \approx 0.55$ , so these FIM emitters are apparently close to ideal. At present, the technological reasons for this are not fully understood.

This parameter,  $m$ , is the overall angular magnification of the Müller emitter for ions that leave a spherical apex normally. As the SCPE had angular magnification  $m' = 1$ , we deduce that the ideal weak lens has angular magnification  $m'' = m/m' \approx 0.578$ .

### 3.8.3 TRANSVERSE MAGNIFICATION

The weak lens will make images of radii  $\rho_1$  and  $\rho_2$ , respectively, of the first and second conceptual objects produced by the SCPE. For small objects, a transverse magnification  $M''$  can be defined for the weak lens by

$$M'' \equiv \rho_1 / \rho'_1 = \rho_2 / \rho'_2 \quad (3.21)$$

As an ion moves away from the emitter, it moves from a region where it effectively sees the potential distribution of a sphere to one where it sees the potential distribution of the Müller emitter as a whole. Hawkes and Kasper note that, optically, this is equivalent to a change in refractive index where the ratio of the refractive indexes is given by

$$\text{Ratio} = (V / F_a r_a)^{1/2} = k_a^{1/2} \quad (3.22)$$

where  $k_a$  is the shape factor used earlier. So, for the weak lens here, the Smith–Helmholtz formula of photon optics (e.g., Ref. 70, Section 4.4.5) should be replaced, certainly for small angles, by the formula

$$m'' M'' k_a^{1/2} = 1 \quad (3.23)$$

For field ion emitters, normally  $5 < k_a < 7$ ; we assume  $k_a = 6$  as an *ideal* value here. Hence

$$M'' = 1/(k_a^{1/2} m'') \approx 1/\sqrt{2} \approx 0.707 \quad (3.24)$$

So ideal Müller emitters might be expected to have overall transverse magnification ( $M = \rho_1/\rho_0$ ) given approximately by  $M = M' M'' \approx \sqrt{2}/4 \approx 0.354$ .

### 3.8.4 MÜLLER Emitter Source Sizes

For consistency in our illustrative values, however, we need to use the value  $k_a = V/F_a r_a = 5.56$ , which (with  $m'' = 1/\sqrt{3}$ ) yields  $M'' \approx 0.735$ ,  $M \approx 0.367$ . This gives the following source radius estimates. The first virtual object (the Gaussian image of the atomic emission site) has estimated FORD  $\rho_1 \approx 41$  pm. The second virtual object (the crossover) has for  $r_a = 80$  nm, the FORD values  $\rho_2(78\text{ K}) \approx 80$  pm,  $\rho_2(300\text{ K}) \approx 152$  pm, and an estimated lower limit (for  $r_a = 10$  nm and  $T_g = 16\text{ K}$ ) of 13 pm. The following approximate direct formula for  $\rho_2$  (also derived by Hawkes and Kasper) can be obtained:

$$\rho_2 \approx r_a (3k_B T_g / eV)^{1/2} \quad (3.25)$$

These estimates should be treated with caution because emitters vary, and because the effects of the pyramidal buildup have not been taken into account.

### 3.8.5 PROJECTION MAGNIFICATION

If a small or smallish feature of length  $l_0$  in the surface of the Müller emitter converts into a feature of length  $l$  on a *projection plane* at a distance  $r_E$  from (point E in) the emitter, then the feature's *projection magnification*  $M_{\text{FIM}}$  is defined by

$$M_{\text{FIM}} = l / l_0 \quad (3.26)$$

The symbol  $M_{\text{FIM}}$  is used because this is the transverse magnification that applies to FIM, although it would also apply to the plane of the aperture in a Wehnelt-type electrode arrangement<sup>18</sup> for a GFIS. For a spherical emitter apex of radius  $r_a$ , the feature of length  $l_0$  subtends at P<sub>2</sub> an angle  $\delta\theta_p = l_0/r_a$ , and the linear projection condition, Equation 3.17, provides that this angle converts into distance  $l = r_E m \delta\theta_p$ , so we obtain

$$M_{\text{FIM}} = m r_E / r_a \quad (3.27)$$

In FIM,  $M_{\text{FIM}}$  typically has values between  $10^6$  and  $10^7$ . In practice,  $r_E$  is very nearly equal to the distance  $r_D$  shown in Figure 3.9.

In FIM literature, the angular magnification  $m$  in Equation 3.27 is written in the form  $1/\beta_{\text{FIM}}$ , and  $\beta_{\text{FIM}}$  is called the *compression factor*. It is reported<sup>6</sup> that  $\beta_{\text{FIM}}$  is typically  $\sim 1.5$ . The theoretical and the experimental (FIM) estimates of  $m$  made previously were in the range 0.55–0.58, so one would predict  $\beta_{\text{FIM}}$  to be approximately in the range 1.7–1.8. Obviously, 1.5 is close to this, but there seems to be a small discrepancy that remains to be explained.

In practice, it is found in FIM that regional variations in emitter curvature can lead to regional variations in projection magnification  $M_{\text{FIM}}$  across the image by up to a factor of two. (For discussion, see Ref. 6.) It is also possible for small protrusions to cause local magnification effects, as discussed in Section 3.9.2.

Also, in some types of FIM geometry, it is possible for additional weak-lens effects associated with the overall microscope geometry to influence the final projection magnification as observed at the image detector. Unplanned effects of this kind should not occur with a GFIS.

### 3.8.6 FIELD ION MICROSCOPE IMAGE-SPOT SIZE

Consider an emission site of radius  $\rho_0$  on S<sub>0</sub>. From Equations 3.15 and 3.17, with  $\delta\theta_p = \alpha_T' = m_T' \rho_0 / r_a$ , the radius  $l_T$  of a blurred image spot on the projection plane at distance  $r_E$  from E is

$$l_T = m'_T m (r_E / r_a) \rho_0 = m'_T M_{\text{FIM}} \rho_0 \quad (3.28)$$

$m'_T$  is the additional factor that occurs as a result of transverse-velocity blurring and specifies the *local* increase in size of small individual features that the human eye judges by the FORD. (Hence the name blurring magnification given earlier.)

### 3.8.7 NUMERICAL TRAJECTORY ANALYSES

Numerical trajectory analyses lead to more detailed results (but shape parameters should be selected so that calculated outcomes, especially  $m$  and  $M$ , are compatible with the experimental FIM data). Parabolic and hyperbolic models have been used, notably in FIM projection theory, but the best simple model is the Smythe–Dyke *sphere-on-orthogonal-cone* (SOC) model,<sup>71,72</sup> which can be fitted to real emitter shapes by choice of two parameters. Wiesner and Everhart<sup>22,23</sup> have used the SOC model for trajectory analysis, and Liu and Orloff<sup>25</sup> have used it for aberration analysis. Wiesner and Everhart find that  $0.35 < M < 0.53$  and that their parameter  $K (= \theta_E/\theta_p)$  which becomes equal to  $m$  as  $\theta_p \rightarrow 0$ , lies in the range  $0.39 < K < 0.64$ . The values obtained earlier lie within these ranges. Thus, analytical, numerical, and experimental estimates of the magnifications associated with the basic behavior of traditional field emitters are all compatible.

Wiesner and Everhart also find<sup>23</sup> that (for the range of parameters they investigated) the positions of the virtual objects generated by the Müller emitter are between 0.6 and 0.2 mm behind the emitter apex and that the first and second virtual objects are separated by a distance of about  $Mr_a/m$  (i.e.,  $\sim 100$  nm), with the second object further from the apex than the first object.

In general, field-evaporated emitter endforms are not spherical but have a slightly flattened emitter apex. Emitters with a pyramidal built-up apex have a different shape. An outstanding problem is the development of better detailed understanding of the regional ionoptical effects that occur when the emitter apex is not quasi-spherical. Additional magnification factors will arise that will be part of the detailed theory of GFIS/FIB performance, but the experience of FIM is that compensation effects may also occur.<sup>6</sup> These problems will have to be solved by detailed numerical trajectory analyses that are well designed and implemented.

## 3.9 ABERATIONS

The aberration data here relate to the virtual images formed by the weak lens, which serve as virtual objects for the FIB/SIM column. Except for angular distortion, they are presented as the disc radii (FORD) for the effect operating alone. Strictly, the aberrations of the emitter and the column should be treated in a fully unified fashion, but this is beyond the scope of this chapter. A discussion of FIB column aberrations with related references can be found in the book by Orloff, Utlaut, and Swanson.<sup>73</sup>

Detailed technical design data for He SIMs are not available at the time of writing. For illustrative purposes, data are given for the value  $\alpha_o = 1$  mrad, where  $\alpha_o$  is the half-angle subtended by the beam acceptance aperture at the emitter (strictly, at point E( $\theta_p \rightarrow 0$ )).

### 3.9.1 SPACE CHARGE

With the LMIS the specimen spot size is limited by trajectory-displacement effects that occur near the emitter due to high space charge above it. Such effects (sometimes called *lateral broadening*) can also limit the performance of field electron emission sources.<sup>74</sup> Unpublished calculations by the present author suggest that the He GFIS is most unlikely to suffer from any space-charge problem, even at a site emission current of  $\sim 100$  pA. No problems of this kind have ever been reported in He GFI (though it is also true that space-charge effects have not specifically been looked for).

### 3.9.2 LOCAL ANGULAR DISTORTION

Small surface protrusions cause additional local magnification, in the sense that the emitted beam diverges more than it otherwise would. The effect in the FIM image (or in the plane of the beam aperture) is geometrically similar to transverse-velocity blurring except that it applies to ions emitted radially from nearby positions on the emitter surface. The result is that an extra factor  $m_X$  should be included in magnification formulas due to this local angular distortion. For a small hemisphere on a parabola, Rose<sup>75</sup> suggested the ansatz

$$m_X \approx 1.1 (r_a / r_{loc})^{1/2} \quad (3.29)$$

where  $r_{loc}$  is the local radius of curvature of the protrusion, but quantitative use of this formula should be cautious.

### 3.9.3 DIFFRACTION AT THE BEAM ACCEPTANCE APERTURE

The diffraction aberration disc has FORD  $\rho_d$  given by

$$\rho_d = 0.61\lambda / \alpha_0 \quad (3.30)$$

where  $\lambda$  is the He-ion wavelength (0.10 pm for an extraction voltage of 20 kV). This yields  $\rho_d = 6.1$  pm.

### 3.9.4 SPHERICAL AND CHROMATIC ABERRATION

Liu and Orloff<sup>25</sup> tabulate spherical ( $C_{si}$ ) and chromatic ( $C_c$ ) aberration coefficients for various shape-parameter ( $n, \gamma$ ) combinations. The parameters  $n = 0.16$ ,  $\gamma = 3$  are used here to describe a typical emitter shape. For  $F_a = 40$  V/nm and  $r_a = 80$  nm (interpolated), coefficient values for Equations 3.31 and 3.32 are:  $C_s \approx 1.08$  mm and  $C_c \approx 90$   $\mu\text{m}$ . (Note that  $C_s = 2C_{si}$ .) For  $\Delta E = 1$  eV and  $V = 20$  kV, the aberration discs have FORD  $\rho_s$  and  $\rho_c$  given by

$$\rho_s = \frac{1}{2} C_s \alpha_0^3 = 0.5 \text{ pm} \quad (3.31)$$

$$\rho_c = C_c \alpha_0 \Delta E / eV = 4.5 \text{ pm} \quad (3.32)$$

Liu and Orloff also report that results depend significantly on a term  $E_a x_c$  in their equations. This is the voltage equivalent of the work needed to form the ion. It is not clear to the writer that this term is needed, as the ionoptical surface in which the ions start with only thermal energy is the critical surface. An alternative approach would identify an equipotential of the SOC model with the critical surface. It remains to be seen if significantly different values for  $C_s$  and  $C_c$  would result. However, the preceding estimates are already very low; so conclusions would not be qualitatively changed.

### 3.9.5 GAS FIELD IONIZATION SOURCE RADIUS

At 78 K, the estimated object FORD before aberration was between 41 and 80 pm, dependent on the focusing conditions. The figures here suggest that beam-aperture diffraction is not significant, and that GFIS radius would be affected by conventional spherical and chromatic *emitter* aberrations only if  $\alpha_0$  were more than  $\sim 4$  mrad. As the natural angular aperture of a one-atom emission site

has been estimated as 1.77 mrad (for an 80 nm radius emitter), it makes no design sense to use an acceptance half-angle as large as 4 mrad. So, in the case under discussion, the effective source size (as far as the emitter is concerned) seems to be determined by the actual source size chosen as a result of focusing procedures.

Obviously, this argument does not take into account any additional effects associated with local magnification at the pyramidal built-up apex; the author's guess is that these are unlikely to be big enough to change the conclusion.

### 3.9.6 COLUMN ABERRATIONS

Notte (private communication) reports that, on the image side, the Carl Zeiss SMT He SIM uses a beam convergence angle  $\alpha_i \sim 0.3$  mrad. Orloff (private communication) suggests that a He SIM column might have aberration coefficients (referred to the image side) of  $C_s = 300$  mm and  $C_c = 30$  mm. Assuming a 40 keV beam at the specimen, and  $\Delta E = 1$  eV as before, yields the results  $\rho_s \sim 4$  pm and  $\rho_c \sim 230$  pm. These figures are very speculative but suggest that, at present, the spot size in a He SIM would be limited by chromatic aberration in the column. This conclusion is consistent with the comment of Joy et al.<sup>76</sup> that the Carl Zeiss machine uses a low convergence angle in order to limit chromatic aberration.

Joy et al.<sup>76</sup> consider that the ionoptical performance is dominated by the demagnified source size. But this conclusion seems to be on the basis of an assumed emitter source size of 500 pm, which is much greater than the values derived earlier. It remains the author's view that at this stage of development, machine performance is most likely determined by chromatic aberration in the column and nonoptical mechanical effects such as vibration. A more careful and informed treatment of He SIM column aberrations is clearly needed but is beyond the scope of this chapter.

## 3.10 SOURCE PROPERTIES

To compare GFIS performance with other types of ion source, the parameters of most interest are the beam current, the brightness  $B$  and the reduced brightness  $B_R$ , and the best-focused spot size. Brightness is defined by  $B = i/A\Omega$ , where a current  $i$  is emitted into a solid angle  $\Omega$  from an area  $A$ . In normal situations, including the Müller emitter, reduced brightness is defined as  $B_R = B/V$ , but for an SCPE, which is an embedded part of a Müller emitter, it needs to be defined as  $B_R = B k_a/V$ .

Background gas pressure and temperature affect emission current and should be stated when comparisons with other sources are made. Obviously, brightness cannot be increased indefinitely by increasing gas pressure, because beyond a certain limit (not clearly known at the time of writing), the occurrence of secondary gas-phase processes will degrade performance.

A design issue is the proportion  $\beta$  of the total emission current that is accepted by the beam-defining aperture. This depends on (a) how much of the total emission current the site in use draws and (b) the proportion of the site current that is accepted by the aperture. With the ALIS source, if only one of the three apex sites is used,  $\beta$  is roughly estimated by  $(\alpha_o/\alpha_T)^2/6$  ( $\approx 5\%$  for  $\alpha_o = 1$  mrad at 78 K). Because the angular intensity in the beam may vary somewhat with angle, this formula is not precise.

For the ALIS GFIS, Ward et al. report<sup>26</sup> the optical characteristics of their system. The available beam currents lie between 1 fA and  $\sim 800$  pA. Other available optical data are compared with those for an LMIS in Table 3.1. Clearly, the reported brightness values are significantly higher than the LMIS and are slightly higher or comparable with values of field electron emitter brightness recently reported.<sup>74</sup> The basic reason for the superior GFIS performance, compared to the LMIS, is the smaller virtual source size.

**TABLE 3.1**  
**Comparison of Properties of Optical Systems Using a Helium Gas Field Ionization Source at 78 K and a Gallium Liquid Metal Ion Source**

Property	He GFIS + SIM <sup>26</sup>	Ga LMIS + FIB <sup>73</sup>
Gas pressure	a	n/a
Virtual source radius (pm)	150	25000
Emission current (pA)	a	2000
Energy spread (eV)	1	5
Extraction voltage (kV)	20	10
Angular current density ( $\mu\text{A}/\text{sr}$ )	2.5	10
Brightness ( $\text{A m}^{-2} \text{sr}^{-1}$ )	$4 \times 10^{13}$	$10^{10}$
Reduced brightness ( $\text{A m}^{-2} \text{sr}^{-1} \text{V}^{-1}$ )	$2 \times 10^9$	$10^6$
Predicted best spot size (nm)	0.25	6 [Ref. 77]

<sup>a</sup> Data not available.

*Source:* Unless otherwise indicated, illustrative data are from Ward B.W. et al., *J. Vac. Sci. Technol.*, 24, 2871–2874, 2006; Orloff J. et al. in *High Resolution Focused Ion Beams: FIB and Its Applications*, Kluwer Academic, New York, 2003.

### 3.10.1 ALTERNATIVE FIGURE OF MERIT

For comparing different GFI sources, it may be better to use an alternative parameter  $G_{\text{FI}}$  defined (in the context of the SCPE) by

$$G_{\text{FI}} = B_{\text{R}} / a_{\text{g}} p_{\text{bk}} = i k_{\text{a}} / A \Omega V a_{\text{g}} p_{\text{bk}} \quad (3.33)$$

where  $i$ ,  $A$ ,  $\Omega$ ,  $V$  are a consistent set of quantities, and  $a_{\text{g}}$  and  $p_{\text{bk}}$  appear in Equation 3.4. This definition takes out the dependence on background gas pressure and temperature. Assume that the machine is focused on the first conceptual object with FORD  $\rho'_1$ , corresponding to a single-atom site that emits a proportion  $\beta_s$  of the captured gas flux. In the context of the SCPE, if the half-angle  $\alpha'_T$  ( $= m'_T \alpha'_S$ ) contains half the current from this site, we obtain

$$G_{\text{FI}} = (\beta_s A_c / 2) (k_{\text{a}} / \pi^2 \rho'_1 r_a^2 m'_T^2 \alpha'_S^2 V) \quad (3.34)$$

Using  $\alpha'_S = \rho_0 / r_a$ ,  $\rho'_1 = \rho_0 / 2$ , and  $F_{\text{a}} = V / k_{\text{a}} r_a$  gives

$$G_{\text{FI}} = 2 \beta_s A_c r_a / \pi^2 m'_T^2 F_{\text{a}} \rho_0^4 \quad (3.35)$$

for which convenient units are  $\text{V}^{-1} \text{ sr}^{-1}$ .  $G_{\text{FI}}$  has values of order  $10^8 \text{ V}^{-1} \text{ sr}^{-1}$  with  $A_c$  and  $m'_T$  the only parameters that are difficult to estimate well (see Appendix A.1 for  $A_c$ ), though  $A_c$  can be measured for any well-defined thermal environment.

In practice, if a single-atom emission site is being used, it looks as though  $A_c$  might be the only parameter that is easy to improve on. Increase in  $r_a$  involves a corresponding increase in extraction voltage and presumably in applied voltages in general, which is likely to be technically undesirable. Obviously, if a value is available for  $G_{\text{FI}}$ , a corresponding estimate of reduced brightness can be made by multiplying  $G_{\text{FI}}$  by  $a_g p_{\text{bk}}$ .

### 3.11 SUMMARY AND DISCUSSION

This chapter has presented an integrated account of the physics and optics of the GFIS. We have concentrated exclusively on sources that produce finely focused beams, but other forms of GFIS exist. For example, the Spindt volcano-type GFIS<sup>78</sup> can find application in specialized forms of mass spectrometry. Also, there is a deuterium GFIS as part of the handheld, pyroelectric cold-nuclear-fusion device reported by Naranjo et al.<sup>79</sup>

In the CP optics of the Müller emitter, the linear projection condition replaces the definition of angular magnification used in photon optics, and the resulting predictions about emitter lens-type behavior (compiled in Table 3.2) are well supported by the experimental facts of FIM. These predictions have been used to derive GFIS ionoptical characteristics as compiled in Table 3.3.

A GFIS differs from other field emission sources in that source brightness varies sharply with position across the surface, and—within a practical limit set by gas-phase scattering effects—is alterable by an external variable (the background gas pressure). To assess the merits of an individual GFIS, a new characterization parameter  $G_{\text{FI}}$  has been defined by dividing the reduced brightness by the background gas flux density (expressed as a supply-current density, typically in A/m<sup>2</sup>).

**TABLE 3.2**  
**Magnification Properties of the Spherical Charged Particle Emitter (SCPE), the Weak Lens, and the Müller Emitter (for the Ideal Case and the Present Case)**

Property	SCPE	Weak Lens	Müller Emitter
Angular magnification	$m' = 1$	$m'' \approx 1/\sqrt{3} \cong 0.577$	$m = m'm'' \approx 1/\sqrt{3} \cong 0.577$
Transverse magnification ( $k_a = 6$ )	$M' = 0.5$	$M'' \approx 1/\sqrt{2} \cong 0.707$	$M = M'M'' \approx \sqrt{2}/4 \cong 0.354$
Transverse magnification ( $k_a = V/r_a F_a = 5.56$ )	$M' = 0.5$	$M'' \cong 0.735$	$M = M'M'' \cong 0.367$

**TABLE 3.3**  
**Definitions and Illustrative Values for Parameters Related to a Müller Emitter and Its Embedded Spherical Charged Particle Emitter (SCPE)**

Parameter	Symbol	Parameters for Embedded SCPE		Parameters for Müller Emitter			
		Definition	$T_g = 78 \text{ K}$	$T_g = 300 \text{ K}$	Definition	$T_g = 78 \text{ K}$	$T_g = 300 \text{ K}$
<i>Angles (in mrad)</i>							
Source fall-off half-angle (FOHA)	$\alpha'_S$	$= \rho_0/r_a$	1.40	1.40	$\alpha_S = m''\alpha'_S$	0.81	0.81
Blurring FOHA	$\alpha'_B$	$= 2(k_B T_g / eF_a r_a)^{1/2}$	2.73	5.36	$\alpha_B = m''\alpha'_B$	1.58	3.09
Total FOHA	$\alpha'_T$	$= \alpha'_B \times \alpha'_S$	3.07*	5.54*	$\alpha_T = m''\alpha'_T$	1.77*	3.20*
Blurring ratio	$m'_B$	$= \alpha'_B/\alpha'_S$	1.95	3.83	$m_B = \alpha_B/\alpha_S = m'_B$	1.95	3.83
Blurring magnification	$m'_T$	$= \alpha'_T/\alpha'_S$	2.19*	3.96*	$m_T = \alpha_T/\alpha_S = m'_T$	2.19*	3.96*
<i>Radii (in pm)</i>							
Fall-off radius (FORD) of first object (Gaussian image)	$\rho'_1$	$= M'\rho_0 = \rho_0/2$	56	56	$\rho_1 = m''\rho'_1$	41	41
FORD of second object (cross-over)	$\rho'_2$	$= m'_B M' \rho_0$ $= (k_B T_g r_a / eF_a)^{1/2}$	109	214	$\rho_2 = m''\rho'_2$	80	158

The values are based on a sphere of radius  $r_a = 80 \text{ nm}$ , a disc-like real object of FORD  $\rho_0 = 112 \text{ pm}$ , and weak-lens magnifications of  $m'' = 0.578$  and  $M'' = 0.735$ . The angles/radii given contain roughly 50% of the emission current. Values with an asterisk are derived from quadrature formulas.  $T_g$  is the mean gas-atom temperature at ionization.

$G_{FI}$  includes a parameter  $A_c$ , the effective capture area, that can be estimated from measurement of the GFIS apex radius before buildup, using the theory in the Appendix A.1, but is better measured in separate well-defined experiments. It is also clear that brightness is maximized by designing emitter shape to ensure that the whole captured current is emitted via a single atomic-sized emission site at the emitter apex and by selecting the beam from the central portion of this site.

For a column without aberrations, the specimen spot size would apparently be minimized by focusing on the first virtual object (the Gaussian image) at room temperature (though there may be stability issues involved) but at low temperatures on the second virtual object (the crossover). A lower limit to the crossover FORD is set by the transverse zero-point energy associated with the He-ion vibrational wavefunction during the emission process. For the site under discussion, this limit has been roughly estimated as equivalent to a temperature of 16 K. Thus, the minimum cross-over FORD is roughly estimated as  $\sim 40$  pm for an 80 nm radius emitter,  $\sim 15$  pm for a 10 nm radius emitter. In practice, for the foreseeable future, it seems likely that the specimen spot size will be limited by column aberrations (probably mainly chromatic aberration) or nonoptical considerations such as mechanical vibration.

At the time of writing, the only GFIS-based FIB-type machine on the market is the Carl Zeiss SMT He SIM.<sup>26</sup> The optical properties reported for this machine (and corresponding properties for an LMIS) are listed in Table 3.1. The claim that the He GFIS is brighter than a gallium LMIS seems to be well justified. Early experimental use of the He SIM suggests that it will be a useful addition to the observational tools of science, especially in the materials area.<sup>76,80</sup>

### 3.11.1 A SPECULATION ABOUT NEW MACHINES OF NANOTECHNOLOGY

Since, even at present, the claimed spot size is less than a nanometer, a popular name for machines like the He SIM could be *picoprober*. Scientifically, it is not impossible that picoprobers could be designed for other noble gases or for small molecules<sup>81</sup> such as N<sub>2</sub> or CH<sub>4</sub>, provided that there is no substantial chemical interaction between the molecule and the substrate or adsorbates on it and provided that one could overcome any problems associated with the failure of free-space FI to keep the emitter surface free from contaminants at the lower BSFs.

If such things turn out to be technologically possible (which is far from guaranteed), then the possibility seems to exist of machines containing multiple picoprober units, for different chemical species, focused on the same small specimen area. With or without the additional presence of reactive gases at the specimen surface (as in some FIB applications<sup>73</sup>), there may be the possibility of new modes of inducing well-localized chemical effects by using high-impact-energy ion bombardment. Technologically, such machines could be very difficult to build and commission, but they seem less of a science-fiction concept than Drexler's molecular assemblers.<sup>82-84</sup>

In short, the ultimate limitations of this third and final generation of optical machines—the ionoptical machines—are yet to be explored.

## ACKNOWLEDGMENT

I thank Jon Orloff and Peter Hawkes for advice and for helpful comments on a draft version of this chapter, but the responsibility for any oversights remains mine.

## A.1 APPENDIX: CORRECTED SOUTHON GAS-SUPPLY THEORY

This appendix presents a corrected, and more user-friendly, version of the main results of Southon's unpublished theory of gas supply to a field ion emitter.<sup>12</sup> Despite the lapse of time and the appearance of later analyses,<sup>14,64,66,85,86</sup> this remains the most wide-reaching theoretical investigation of the issues involved. But, in its present state, it still has limited predictive power.

### A.1.1 PRELIMINARY DEFINITIONS

Here  $Z$  denotes flux density,  $R$  denotes flux,  $\sigma$  denotes flux enhancement factor, and  $\Phi$  is defined by

$$\Phi = \frac{1}{2} \alpha_{\text{gas}} F^2 / k_B T_{\text{bk}} \quad (\text{A.1.1})$$

where  $\alpha_{\text{gas}}$  is the gas-atom polarizability,  $F$  the emitter surface field,  $k_B$  Boltzmann's constant, and the approaching gas atoms are drawn from a population taken to be in thermodynamic equilibrium at temperature  $T_{\text{bk}}$ . Subscripts have the following specific meanings: H = hemisphere, S = sphere, C = cylinder, I = integrated over truncated cone, M = Müller emitter total value, 0 = zero-field value, i = incident value, c = captured value, and a = apex value.

### A.1.2 SUPPLY TO A HEMISPHERE

If no polarization effects were operating, the flux  $R_{\text{H}0}$  incident onto half a spherical emitter of radius  $r_a$  would be given by

$$R_{\text{H}0} = 2\pi r_a^2 Z_{\text{bk}} \quad (\text{A.1.2})$$

where  $Z_{\text{bk}}$  is background gas flux density. When polarization effects operate, the *incident flux*  $R_{\text{Hi}}$  is

$$R_{\text{Hi}} = \sigma_{\text{Si}} R_{\text{H}0} \quad (\text{A.1.3})$$

where  $\sigma_{\text{Si}}$  is the *incident flux enhancement factor* for a sphere. Southon showed that

$$\sigma_{\text{Si}} = \exp(-\Phi_a) + (\pi\Phi_a)^{1/2} \operatorname{erf}(\Phi_a^{1/2}) \approx (\pi\Phi_a)^{1/2} \quad (\text{A.1.4})$$

where  $\Phi_a$  is the value applicable to the sphere surface (and hence to the emitter apex when this is modeled as a hemisphere). This formula replaced earlier formulas<sup>1,87</sup> and is accepted as correct.<sup>28</sup> The mathematical approximation is good when  $\Phi_a > \sim 2$ , which is nearly always the case in practical situations.

In principle, the quantity of interest is the *captured flux*  $R_{\text{Hc}}$ , given by

$$R_{\text{Hc}} = \Pi_S R_{\text{Hi}} = \Pi_S \sigma_{\text{Si}} R_{\text{H}0} \equiv \sigma_{\text{Sc}} R_{\text{H}0} \quad (\text{A.1.5})$$

where  $\Pi_S$  is the *mean probability of capture* for a spherical emitter, and  $\sigma_{\text{Sc}}$  ( $\equiv \Pi_S \sigma_{\text{Si}}$ ) is the *captured-flux enhancement factor* for a sphere. In practical situations the approximations  $\Pi_S \approx 1$ ,  $\sigma_{\text{Sc}} \approx \sigma_{\text{Si}}$  are normally used for the emitter apex.

### A.1.3 SUPPLY TO A CYLINDER

For a cylinder of radius  $r$  and length  $l$ , the incident flux  $R_{\text{C}0}$  in the absence of polarization effects is  $R_{\text{C}0} = 2\pi rl Z_{\text{bk}}$ . Southon showed that, when polarization effects operate, the incident flux  $R_{\text{Ci}}$  is given in terms of an enhancement factor  $\sigma_{\text{Ci}}$  by

$$R_{\text{Ci}} = \sigma_{\text{Ci}} R_{\text{C}0} = [2(\Phi/\pi)^{1/2} + \{1 - \operatorname{erf}(\Phi^{1/2})\} \exp(\Phi^{1/2})] R_{\text{C}0} \quad (\text{A.1.6a})$$

$$\approx 2(\Phi/\pi)^{1/2} R_{\text{C}0} \quad (\text{A.1.6b})$$

The approximation represents the asymptotic form as  $\Phi$  becomes large, is adequate in practical situations, and is used because it simplifies the mathematics.

Derivation of an expression for the probability of capture  $\Pi_C$  for an atom hitting a cylinder is not straightforward; after analysis, Southon suggested the semiheuristic form

$$\Pi_C \sim c_0 \Phi^{1/2} (\Phi \leq \Phi_1) \quad \Pi_C = 1 \quad (\Phi > \Phi_1) \quad (\text{A.1.7})$$

where  $c_0$  is related to the (thermal) accommodation coefficient  $a_0$  for the atom-surface collision by

$$c_0 = 0.46 a_0^{1/4} \quad (\text{A.1.8})$$

$\Phi_1$  is the value of  $\Phi$  for which  $\Pi_C = 1$  and is given by

$$\Phi_1 = 1/c_0^2 = 4.73 a_0^{-1/2} \quad (\text{A.1.9})$$

#### A.1.4 SUPPLY TO A CONE

Consider a truncated cone, of radius  $r_a$  at the smaller end, and let  $l$  be length measured from the smaller end. For the captured flux ( $dR_{lc}$ ) to an element of the cone of radius  $r$  and length  $dl$ , Southon proposed the ansatz

$$dR_{lc} = 2\pi r dl \times \sigma_{Ci} \Pi_C Z_{bk} \quad (\text{A.1.10})$$

where  $\sigma_{Ci}$  and  $\Pi_C$  are the parameters for a cylinder of radius  $r$ . For a cone of half-angle  $\eta$ , we have  $r = r_a + l \tan \eta$  and  $dr = \tan \eta dl$ , so

$$dR_{lc} = (2\pi r dr / \tan \eta) \times \sigma_{Ci} \Pi_C Z_{bk} \quad (\text{A.1.11})$$

Southon next assumed that surface field  $F$  falls off with distance along the cone in such a fashion that

$$\Phi / \Phi_a = (F / F_a)^2 = (r_a / r)^2 \quad (\text{A.1.12})$$

This is not geometrically correct but seems an adequate zeroth-order approximation. This leads to the relation  $r dr = -(r_a^2 \Phi_a^2 / 2) \Phi^{-2} d\Phi$  and hence to

$$dR_{lc} = -(\pi r_a^2 Z_{bk} \Phi_a / \tan \eta) \Phi^{-2} d\Phi \times \sigma_{Ci} \Pi_C \quad (\text{A.1.13})$$

This expression has to be integrated over the relevant part of the cone surface from the point where  $r = r_a$  and  $\Phi = \Phi_a$  to a point on the cone (at which  $\Phi = \Phi_b$ ) where the captured atoms cannot (on average) get to the emitter apex without being desorbed back into the gas phase. Reversing the limits of integration to get rid of the minus sign and using Equations A.1.2 and A.1.7, we obtain

$$R_{lc} = R_{H0} (\Phi_a / \pi^{1/2} \tan \eta) \int_{\Phi_b}^{\Phi_a} \Phi^{-3/2} \Pi_C d\Phi \quad (\text{A.1.14})$$

This equation replaces equations 50 and 51 in Ref. 12.

Two cases now arise. If  $\Phi_a \leq \Phi_1$ , so that  $\Pi_C < 1$  at all points, then Equation A.1.7 yields

$$R_{lc} = R_{H0}(\Phi_a / \pi^{1/2} \tan \eta) c_0 \int_{\Phi_b}^{\Phi_a} \Phi^{-1} d\Phi \quad (A.1.15a)$$

$$= R_{H0}[(\Phi_a / \pi^{1/2} \tan \eta) c_0 \ln(\Phi_a / \Phi_b)] \equiv \sigma_{lc1} R_{H0} \quad (A.1.15b)$$

where  $\sigma_{lc1}$  is an enhancement factor defined by this equation.

If  $\Phi_a > \Phi_1$ , the integral in Equation A.1.14 is split into two to become

$$R_{lc} = (R_{H0} \Phi_a / \pi^{1/2} \tan \eta) \left\{ c_0 \int_{\Phi_b}^{\Phi_1} \Phi^{-1} d\Phi + \int_{\Phi_1}^{\Phi_a} \Phi^{-3/2} d\Phi \right\} \quad (A.1.16a)$$

$$= R_{H0}[(\Phi_a / \pi^{1/2} \tan \eta) \{c_0 \ln(\Phi_1 / \Phi_b) + 2(\Phi_1^{-1/2} - \Phi_a^{-1/2})\}] \equiv \sigma_{lc2} R_{H0} \quad (A.1.16b)$$

where  $\sigma_{lc2}$  is an enhancement factor defined by Equation A.1.16b. This equation replaces equation 52 in Ref. 12.

In the absence of detailed relevant information, the choice of the value of  $\Phi_b$  is somewhat arbitrary. We follow Southon in putting  $\Phi_b = 1$  and also in taking  $\eta = 13^\circ$ , so  $\tan \eta = 0.231$ .

### A.1.5 TOTAL CAPTURED FLUX FOR MÜLLER Emitter

The total captured flux for the emitter is estimated by adding the contributions from the hemispherical cap and the conical shank (ignoring fine details of geometry) and is given formally by

$$R_{Mc} = (\sigma_{Sc} + \sigma_{lc}) R_{H0} \equiv \sigma_{Mc} R_{H0} = \sigma_{Mc} (2\pi r_a^2) Z_{bk} \quad (A.1.17)$$

$\sigma_{Mc}$  is the captured-flux enhancement factor for the Müller emitter as a whole, and  $\sigma_{lc}$  is either  $\sigma_{lc1}$  or  $\sigma_{lc2}$ , depending on the relative sizes of  $\Phi_1$  and  $\Phi_a$ . From the earlier definition of the effective area of capture  $A_c$ , it can be seen that

$$A_c = \sigma_{Mc} (2\pi r_a^2) \quad (A.1.18)$$

Although this treatment is not exact, it does exhibit the physics of estimating the gas supply to a field ion emitter and brings out many of the difficulties involved. Obviously, it is a significant approximation to derive enhancement factors in the way that Southon did. In reality, the polarized gas atoms move in the PPE distribution due to the whole Müller emitter, and the true field fall-off down the shank is different from that assumed. But, for the trajectories, the only alternative that is obviously superior is a massively complicated Monte-Carlo-type simulation of atomic trajectories and behavior. This was unthinkable in the 1960s and is probably still beyond what can be accomplished satisfactorily.

### A.1.6 NUMERICAL ILLUSTRATIONS

Some features of the real situation can be illustrated using Southon's model. A particular difficulty relates to accommodation coefficients. A simplistic classical formula is

$$a_0 = 4mM/(m + M)^2 \quad (A.1.19)$$

where  $m$  and  $M$  here are the masses (or relative atomic masses) of the He atom and the object it collides with. On the emitter shank, this is quite likely to be an adsorbed oxygen atom, so we would have  $m \approx 4$ ,  $M \approx 16$ ,  $a_0 \sim 0.64$ ,  $c_0 \sim 0.41$ , and  $\Phi_1 \sim 5.9$ ; but, if we assumed collision with a tungsten atom, then we would have  $M \approx 184$ ,  $a_0 \sim 0.083$ ,  $c_0 \sim 0.25$ , and  $\Phi_1 \sim 16$ . Theoretical results are sensitive to the value of  $\Phi_1$  and so depend significantly on the assumptions about the accommodation process. It is far from obvious what the surface state of the shank of a real, dirty emitter is and what parameters should be used to make the model applicable. What this model (or any other present model) generates, therefore, are illustrative values.

To obtain numerical values, we continue with the oxygen-based parameters, which seem more realistic. Values derived are collated in Table A.1.1 for  $T_{bk} = 300, 78, 20$ , and  $5\text{ K}$ . As before, values are shown to three significant figures for consistency of calculation, but the uncertainties of the situation are so great that the only accuracy claimed here is that the final result is “probably within an order of magnitude, or possibly better.”

It can be seen that the total captured-flux enhancement factor  $\sigma_{Mc}$  increases significantly as the background gas temperature is decreased due to the gas captured on the conical emitter shank. Because  $\sigma_{Ic}/\sigma_H > 1$ , the theory also indicates that more gas is captured on the shank than at the apex; this is predicted even at room temperature, but the effect is stronger at lower temperatures. This interpretation of experimental  $i$ - $V$  characteristics was first suggested for hydrogen by Becker.<sup>88</sup> Hydrogen field ion micrographs taken near  $80\text{ K}$  confirm<sup>7,9</sup> that most of the gas supply comes from the shank (see Figure 9 in Ref. 7). This is also consistent with a time-delay phenomenon noted by Tsong and Müller.<sup>89</sup>

For comparison, values derived from Southon’s uncorrected theory are shown in the last two lines of Table A.1.1. These show the same qualitative trends as the corrected theory but make them much stronger at lower temperatures. Jousten et al. noted<sup>64</sup> that the uncorrected Southon theory predicts enhancement factors larger than those observed experimentally.

**TABLE A.1.1**

**Working Table for the Calculation of the (Total) Captured-Flux Enhancement Factor  $\sigma_{Mc}$  for a Müller Emitter, for Various Values of the Background Gas Temperature  $T_{bk}$**

Origin	Parameter	Value at $T_{bk} =$			
		300 K	78 K	20 K	5 K
See text	$\Phi_b$	1.00	1.00	1.00	1.00
Equation A.1.9	$\Phi_1$	5.91	5.91	5.91	5.91
Equation A.1.3	$\Phi_a$	5.60	21.5	84.0	336
Equation A.1.14	$\Phi_a/\pi^{1/2}\tan\eta$	13.7	52.6	205	821
See text	Is $\Phi_a > \Phi_1$ ?	No	Yes	Yes	Yes
Equation A.1.15b	$c_o \ln(\Phi_a/\Phi_b)$	0.709	n/a	n/a	n/a
Equation A.1.16b	$c_o \ln(\Phi_1/\Phi_b)$	n/a	0.731	0.731	0.731
Equation A.1.16b	$2(\Phi_1^{-1/2} - \Phi_a^{-1/2})$	n/a	0.392	0.604	0.713
Rows above	Sum of integrals	0.709	1.12	1.34	1.44
$\times \Phi_a/\pi^{1/2}\tan\eta$	$\sigma_{Ic}$	9.70	59.1	274	1190
Equation A.1.4	$\sigma_{Hc}$	4.19	8.23	16.2	32.5
Equation A.1.17	$\sigma_{Mc}$	13.9	67.3	290	1220
Rows above	$\sigma_{Ic}/\sigma_{Hc}$	2.31	7.18	16.9	36.5
Ref. 12	$\sigma_{Mc}$ (Southon)	29.9	697	7240	65550
Ref. 12	$\sigma_{Ic}/\sigma_{Hc}$ (Southon)	2.01	17.3	47.6	109

$c_o$  is based on collision with an oxygen atom and is taken as 0.41, and  $\eta$  is taken as  $13^\circ$ .

Ideally, one would like to compare these theoretical predictions with experimental results from a range of emitters, but the number of suitable (well-defined) sets of results is actually very small. In terms of background gas temperature, the best-defined experiments are those of Feldman and Gomer,<sup>28</sup> who immersed a complete microscope in refrigerant. From their Figure 12, for  $r_a = 28.7$  nm,  $T_{bk} = 78$  K,  $p_{gas} = 2.38 \mu\text{m Hg}$  (mTorr),  $F_a = 45 \text{ V/nm}$ , we obtain  $A_c = 0.058 \mu\text{m}^2$  and  $\sigma_{Mc}(\text{exp}) = 11$ ; the predicted value, using corrected theory is  $\sigma_{Mc}(\text{theor}) = 67$ . From their Figure 13, for  $r_a = 44.5$  nm,  $T_{bk} = 78$  K,  $p_{gas} = 2.18 \mu\text{m Hg}$  (mTorr),  $F_a = 43 \text{ V/nm}$ , we obtain  $A_c = 0.27 \mu\text{m}^2$  and  $\sigma_{Mc}(\text{exp}) = 22$ ; the predicted value is  $\sigma_{Mc}(\text{theor}) = 61$ . On the face of it, these figures suggest that corrected Southon theory still predicts results that are somewhat too high numerically. However, the predicted variation in  $\sigma_{Mc}$  with temperature is in broad agreement with the high-field end of the measurements shown in Figure 5 of Ref. 44.

If the trend in the Feldman and Gomer results were continued, one might expect an 80 nm radius emitter to have at 45 V/nm, at  $T_{bk} = 78$  K, an  $A_c$  value of order of magnitude  $1 \mu\text{m}^2$ .

Some general points emerge from this analysis. (1) It is the effective temperature of the *incoming* gas atoms that is important in determining the captured flux (the emitter temperature, if different, will have a much smaller effect). (2) In a normal FIM configuration, the apex of the emitter nearly always points toward room-temperature objects, although the sides of the emitter may be exposed to refrigerated objects. For such instruments, there is a theoretical case for calculating  $\sigma_{Hc}$  on the basis of  $T_{bk} = 300$  K but  $\sigma_{Ic}$  on the basis of refrigerant temperature, although this is not done here. (3) The  $i$ - $V$  measurements reported in GFIS literature may be sufficient for technical purposes, but few are sufficiently well characterized for good theory/experiment comparisons. It is important to measure the *total* emission current (disregarding secondary contributions) in a well-characterized thermal situation as well as the current from prominent emitting features. (4) For GFIS brightness, the gain in enhancement factor in going from 300 to 78 K is possibly not worth having (due to the complications of cryogenics) if the beam current can be increased by other means, but the reduction in energy spread may be important. (5) As the captured flux depends on the physical state of the emitter shank, there may be issues of GFIS current stability if the state of the shank changes with time during operation.

As already emphasized, this gas-supply treatment is intended not as an exact model but more as a template into which different model assumptions (e.g., about accommodation or about field fall-off down the shank) can be fitted when a wider range of well-characterized experimental results becomes available to test and adjust it.

## A.2 APPENDIX: GLOSSARY OF SPECIAL TERMS

**Applied field,  $F$**  Electric field just above the emitter apex, outside the range of very local fields due to the atomic structure ( $F$  is often not well defined but can be defined more carefully in specific cases).

**Background gas temperature,  $T_{bk}$**  Gas atoms approaching the emitter (or some part of it) are assumed to be drawn from a population in thermodynamic equilibrium at temperature  $T_{bk}$ .

**Best image field/voltage** The field/voltage at which a field ion image is sharpest.

**Best source field/voltage** The field/voltage at which a gas field ionization source gives optimum performance for a given background gas pressure.

**Comet** Bright streak on field ion image formed by secondary ions generated by charge exchange along the path of the primary ion beam from a particular emission site.

**Compression factor,  $\beta_{FIM}$**  Parameter used in field ion microscope literature, equal by definition to  $m^{-1}$ , where  $m$  is the angular magnification of the Müller emitter.

**Critical surface**, When the gas-atom nucleus lies in the critical surface, the topmost gas-atom electron can make a radiationless energy-conserving transition to the emitter Fermi level.

**Crossover** Here, the disc of minimum confusion determined by ion trajectory back-projection, to exist at the center of a spherical charged particle emitter (or an optically formed image of this disc).

**Dynamic low-field equilibrium (DLFE) regime** Observable gas-supply regime that occurs at low emitter fields (below the gas field ionization source operating field) when the emitter temperature and background gas temperature are different.

**Effective capture area,  $A_c$**  A parameter that appears in the theory of gas supply to a field ion emitter and hence in the theory of GFIS brightness: the captured gas flux is  $A_c$  times the background gas flux density.

**Electrical surface (strictly applicable only to atomically flat surfaces)** The surface in space at which “the electric field would seem to start, if it were assumed to be constant”: this concept is used in one-dimensional models to make sure that the potential energy term  $eFx$  has the correct limiting value as  $x$  becomes large: physically, the electrical surface is close to the outer edge of the emitter’s electron-charge clouds: see Refs 58 and 59 for more careful discussions.

**Electroformation** The process of formation of a pointed shape by heating an object in the presence of a high electric field, sometimes in the presence of bombarding particles.

**Electron tunneling rate-constant,  $P_e$**  Probability per unit time that field ionization of a gas atom will take place by electron tunneling.

**Equivalent background supply-current density,  $J_{bk}$**  The current-density equivalent of the background gas flux density  $Z_{bk}$ , assuming one elementary positive charge per atom.

**Fall-off half-angle (FOHA)** The beam half-angle that contains a defined proportion of the current emitted from a given site (a notional 50% is used here).

**Fall-off radius (FORD)** A radius, associated with a real or virtual emitting object that contains a defined proportion of the current emitted from the object (a notional 50% is used here).

**Field evaporation** The process in which an emitter atom is removed from the emitter surface by the action of a very high electric field alone: for tungsten, this occurs at a field of  $\sim 57$  V/nm.

**Firmly field-adsorbed layer** Set of gas atoms each bound to the surface, immediately above a protruding emitter atom, by a mixture of polarization and chemical (exchange) forces.

**First conceptual object** The Gaussian image formed by the (hypothetical) spherical charged particle emitter used as an ionoptical model for the apex region of a Müller emitter.

**Forbes–Ljepojevic thermodynamic potential (or electrical Gibbs function)** Thermodynamic potential (or *free energy*) analogous to the usual mechanical Gibbs function, but with mechanical (*pdv*) external-work term replaced by an electrical external-work term of form “–voltage  $\times$  d(charge)”: this free energy applies when a change in electrical capacitance occurs in a thermodynamic system to which an external battery or high-voltage generator is connected.

**Forbidden zone** The region between the emitter surface and the critical surface where normal field ionization is forbidden, because there are no vacant states for electrons to tunnel into.

**Free-space field ionization (FSFI)** Field ionization that takes place in space away from the ionization zones close to the critical surface.

**Gas-atom history** The statistical (probabilistic) concept that represents what the *typical time-dependent behaviors* of gas atoms trapped by the polarization potential-energy well would be if they were not field ionized (assumes gas pressures are so low that gas-atom collisions in space are rare).

**Gas concentration,  $C_g(X)$**  Probability per unit volume of finding a gas-atom nucleus in an infinitesimally small volume of space at  $X$ , expressed as the quantity *atoms per unit volume*.

**Gaussian image** Here, the *physical* image determined by ion trajectory back-projection.

**Hot ionization** Field ionization in ionization zones in circumstances where the mean gas-atom kinetic energy parallel to the emitter surface is greater than  $k_B T_{\text{em}}$ , where  $T_{\text{em}}$  is the emitter temperature.

**Ionization density,  $j(X)$**  Charge generated per unit volume at point  $X$  in space above the emitter.

**Ionization zone** Small disc-shaped region of space somewhat above a substrate atom, just outside the critical surface: statistically, in field ion microscope/gas field ionization source operation, the majority of field ionization events take place when a gas-atom nucleus is in an ionization zone.

**Ionoptical surface** The surface in which the ions are deemed to start motion with thermal energy only: for gas field ionization the ionoptical surface is identified with the critical surface.

**Linear projection condition, Equation 3.17** Mathematical condition that specifies that in the action of a charged particle lens for large angles,  $\tan\theta_E$  is directly proportional to  $\theta_p$ , where  $\theta_p$  and  $\theta_E$  are the object and image angles, respectively (see Figure 3.8).

**Müller emitter** The pointed needle-like metal or liquid object that generates electron or ion emission at or above its pointed tip (see Figure 3.2).

**Müller-type angular magnification** Angular magnification as defined by Equation 3.17: this differs from the (Helmholtz-type) definition of angular magnification used in photon optics, because it reflects the different physical behaviors of photons and charged particles.

**Natural angular aperture** The concept that recognizes that most of the emission current from a particular emission site lies within a small emission half-angle: for convenience,  $\alpha_T$  is used as a measure.

**$n_1$**  Formally, the amount of substance of a system containing one entity: informally, *one atom or one ion or one molecule*, depending on context, like the SI unit the mole  $n_1$  has the formal dimension *amount of substance*.

**Picoprober** Suggested popular name for scanning ion microscope with subnanometer resolution.

**Polarization potential-energy well** Region of space around the apex of a Müller emitter where the gas-atom polarization potential energy is greater in magnitude than its thermal energy: this well can extend several cap radii into space and trap gas atoms.

**Projection magnification,  $M_{\text{FIM}}$**  Ratio of feature length observed in the projection plane to length of feature on the emitter surface.

**Projection plane** A distant plane where the (unfocused) ion current distribution from a Müller emitter is observed.

**Quasiclassical theory** A theoretical approximation in which wave mechanics is formally applied to the motion of the topmost gas-atom electron but not to the motion of the gas-atom nucleus.

**Second conceptual object** The crossover formed by the (hypothetical) spherical charged particle emitter used as an ionoptical model for the apex region of a Müller emitter (see Figure 3.10).

**Supertip** Small field-enhancing protrusion formed on the apex region of a Müller emitter.

**Supply-and-capture (SAC) regime** A gas-supply regime where it is assumed that (nearly) all captured gas atoms are field ionized close above the surface of the emitter apex.

**Transverse zero-point energy** Energy associated with the lowest-energy mode of the transverse part of the wavefunction for ion-core behavior during emission.

## LIST OF ABBREVIATIONS

ALIS™	Atomic level ion source
BIF	Best image field
BIV	Best image voltage

BSF	Best source field
BSV	Best source voltage
CP	Charged particle
DLFE	Dynamic low-field equilibrium
FI	Field ionization
FIB	Focused ion beam
FIM	Field ion microscope (or field ion microscopy)
FOHA	Fall-off half-angle
FORD	Fall-off radius
FSFI	Free-space field ionization
FWHM	Full width at half maximum
GFI	Gas field ionization
GFIS	Gas field ionization source
He	Helium
KE	Kinetic energy
LMIS	Liquid metal ion source
PPE	Polarization potential energy
SAC	Supply-and-capture
SCPE	Spherical charged particle emitter
SIM	Scanning ion microscope
W	Tungsten

## REFERENCES

1. R. Gomer, 1963. *Field Emission and Field Ionization* (Harvard University Press, Cambridge, MA).
2. E.W. Müller and T.T. Tsong, 1969. *Field Ion Microscopy: Principles and Applications* (Elsevier, New York).
3. R.G. Forbes, 1985. Seeing atoms: The origins of local contrast in field-ion images. *J. Phys. D: Appl. Phys.* **18**, 973–1018.
4. C.M.C. de Castilho and D.R. Kingham, 1987. Resolution of the field-ion microscope. *J. Phys. D: Appl. Phys.* **20**, 116–124.
5. T.T. Tsong, 1990. *Atom-Probe Field Ion Microscopy* (Cambridge University Press, Cambridge, UK).
6. M.K. Miller, A. Cerezo, M.G. Heatherington and G.D.W. Smith, 1996. *Atom Probe Field Ion Microscopy* (Clarendon Press, Oxford).
7. R.G. Forbes, 1996. Field-ion imaging old and new. *Appl. Surf. Sci.* **94/95**, 1–16.
8. A.J. Melmed, 2003. Erwin Müller. *US National Academy of Sciences – Biographical memoirs* (National Academies Press, Washington, D.C.), Vol. 82, 198–219.
9. R.G. Forbes, 1971. Field ion microscopy at very low temperatures. PhD thesis, University of Cambridge.
10. R.G. Forbes, 1971. A theory of field-ion imaging: I – A quasi-classical site-current formula; II – On the origin of site-current variations. *J. Microsc.* **96**, 57–61 and 63–76.
11. R.G. Forbes, 1986. Towards an unified theory of helium field ionization phenomena. *J. de Physique* **47**, Colloque C2, 3–10.
12. M.J. Southon, 1963. Image formation in the field ion microscope. PhD thesis, University of Cambridge.
13. M.J. Southon and D.G. Brandon, 1963. Current–voltage characteristics of the helium field-ion microscope. *Philos. Mag.* **8**, 579–591.
14. H.A.M. Van Eekelen, 1970. The behaviour of the field-ion microscope: A gas dynamical calculation. *Surf. Sci.* **21**, 21–44.
15. J. Witt and K. Müller, 1986. Computer-aided measurement of FIM intensities. *J. de Physique* **47**, Colloque C2, 465–470.
16. R. Smith and J.M. Walls, 1978. Ion trajectories in the field-ion microscope. *J. Phys. D: Appl. Phys.* **11**, 409–419.
17. H.N. Southworth and J.M. Walls, 1978. The projection geometry of the field-ion image. *Surf. Sci.* **76**, 129–140.

18. P.W. Hawkes and E. Kasper, 1996. *Principles of Electron Optics Vol. 2: Applied Geometrical Optics* (Academic, London), Chapter 45.
19. E. Ruska, 1933. Zur fokussierbarkeit von kathodenstrahlbündeln grosser augsgangquerschnitte. *Z. Physik.* **83**, 684–697.
20. R. Gomer, 1952. Velocity distribution of electrons in field emission. Resolution in the projection microscope. *J. Chem. Phys.* **20**, 1772–1776.
21. T.E. Everhart, 1967. Simplified analysis of point-cathode electron sources. *J. Appl. Phys.* **38**, 4944–4957.
22. J.C. Wiesner and T.E. Everhart, 1973. Point-cathode electron sources – electron optics of the initial diode region. *J. Appl. Phys.* **44**, 2140–2148.
23. J.C. Wiesner and T.E. Everhart, 1974. Point-cathode electron sources – electron optics of the initial diode region: Errata and addendum. *J. Appl. Phys.* **45**, 2797–2798.
24. V.N. Tondare, 2005. Quest for high brightness, monochromatic noble gas ion sources. *J. Vac. Sci. Technol.* **23**, 1496–1508.
25. X. Liu and J. Orloff, 2005. A study of optical properties of gas phase field ionization sources. *Adv. Imaging Electron Phys.* **138**, 147–176.
26. B.W. Ward, J.A. Notte and N.P. Economou, 2006. Helium ion microscope: A new tool for nanoscale microscopy and metrology. *J. Vac. Sci. Technol.* **24**, 2871–2874.
27. S. Kalbitzer, 1999. Bright ion beams for the nuclear microprobe. *Nucl. Instrum. Methods Phys. Res. B* **158**, 53–60.
28. U. Feldman and R. Gomer, 1966. Some observations on low-temperature field ion microscopy. *J. Appl. Phys.* **6**, 2380–2390.
29. R. Levi-Setti, 1974. Proton scanning microscopy: Feasibility and promise. In: *Scanning Electron Microscopy 1974* (O. Johari and I. Corvin, eds) (IIT Research Institute, Chicago, IL), pp. 125–134.
30. W.H. Escivitz, T.R. Fox and R. Levi-Setti, 1976. Scanning transmission ion microscope with a field ion source. *Proc. Natl. Acad. Sci.* **72**, 1826–1828.
31. J.H. Orloff and L.W. Swanson, 1976. Study of a field ionization source for microprobe applications. *J. Vac. Sci. Technol.* **12**, 1209–1213.
32. J. Orloff and L.W. Swanson, 1977. A scanning ion microscope with a field ionization source. *Scan. Electron Microsc.* **1**, 57–62.
33. G.R. Hanson and B.M. Siegel, 1980. H<sub>2</sub> and rare gas field ion source with high angular current. *J. Vac. Sci. Technol.* **16**, 1876–1878.
34. P.R. Schwobel and G.R. Hanson, 1984. Localized field ion emission using adsorbed hydrogen films on <110>-oriented tungsten field emitters. *J. Appl. Phys.* **56**, 2101–2105.
35. J.A. Kubby and B.M. Siegel, 1986. High resolution structuring of emitter tips for the gaseous field ionization source. *J. Vac. Sci. Technol. B* **4**, 120–125.
36. T. K. Horiuchi, T. Itakura and H. Ishikawa, 1988. Emission characteristics and stability of a helium field-ion source. *J. Vac. Sci. Technol. B* **6**, 937–940.
37. T. Itakura, K. Horiuchi and N. Nakayama, 1991. Microprobe of helium ions. *J. Vac. Sci. Technol. B* **9**, 2596–2901.
38. M.K. Miller and S.J. Sijbrandij, 1999. A novel approach to gaseous field ion sources for focused ion beam applications. *Ultramicroscopy* **79**, 225–230.
39. T. Teraoka, H. Nakane and H. Adachi, 1994. Hydrogen permeability of palladium needles and possibility of new field ion source. *Jpn. J. Appl. Phys.* **33**, L1110–L1112.
40. E. Salançon, Z. Hammadi, and R. Morin, 2003. A new approach to gas field ion sources. *Ultramicroscopy* **95**, 183–188.
41. H.-W. Fink, 1986. Mono-atomic tips for scanning tunneling microscopy. *IBM J. Res. Develop.* **30**, 460–465.
42. H.-W. Fink, 1988. Point source for ions and electrons. *Phys. Scripta* **38**, 260–263.
43. K. Jousten, K. Böhringer, R. Börret and S. Kalbitzer, 1988. Growth and current characteristics of stable protrusions on tungsten field ion emitters. *Ultramicroscopy* **26**, 301–312.
44. K. Böhringer, K. Jousten and S. Kalbitzer, 1988. Development of a high brightness gas field ionization source. *Nucl. Instrum. Methods Phys. Res. B* **30**, 289–292.
45. R. Börret, K. Böhringer and S. Kalbitzer, 1990 Current–voltage characteristic of a gas field ion source with a supertip. *J. Phys. D: Appl. Phys.* **23**, 1271–1277.
46. V.T. Binh, N. Garcia and S.T. Purcell, 1996. Electron field emission from atomic sources: Fabrication, properties, and applications of nanotips. *Adv. Imaging Electron Phys.* **95**, 63–153.

47. N.N. Ljepojevic and R.G. Forbes, 1995. Variational thermodynamic derivation of the formula for pressure difference across a charged conducting liquid surface and its relation to the thermodynamics of electrical capacitance. *Proc. R. Soc. Lond. A* **450**, 177–192.
48. W.P. Dyke, F.M. Charbonnier, R.W. Strayer, R.L. Floyd, J.P. Barbour and J.K. Trolan, 1960. Electrical stability and life of the heated field emission cathode. *J. Appl. Phys.* **31**, 790–805.
49. W. Polanschütz and E. Krautz, 1974. Feldionenmikroskopie bei höheren Temperaturen. *Z. Metallkde* **65**, 623–636.
50. L.W. Swanson, 1976. Comparative study of the zirconiated and built-up W thermal-field cathode. *J. Vac. Sci. Technol.* **12**, 1228–1233.
51. Y.A. Vlasov, O.L. Golubev and V.N. Shrednik, 1989. Equilibrium and stationary shapes of heated metallic emitters in a strong electric field. *Growth of Crystals* Vol. 19, 3–14. (Translated in 1993 by Consultants Bureau, New York, from Russian original printed by “Nauka”, Moscow.)
52. T. Sakurai and E.W. Müller, 1977. Field calibration using the energy distribution of a free-space field ionization. *J. Appl. Phys.* **48**, 2618–2625.
53. E.W. Müller, S.B. McLane and J.A. Panitz, 1969. Field adsorption and desorption of helium and neon, *Surf. Sci.* **17**, 430–438.
54. T.T. Tsong and E.W. Müller, 1970. Field adsorption of inert gas atoms on field ion emitter surfaces. *Phys. Rev. Lett.* **25**, 911–913.
55. K. Nath, H.J. Kreuzer and A.B. Anderson, 1986. Field adsorption of rare gases. *Surf. Sci.* **176**, 261–283.
56. R.G. Forbes, H.J. Kreuzer and R.L.C. Wang, 1996. On the theory of helium field adsorption. *Appl. Surf. Sci.* **94/95**, 60–67.
57. R.L.C. Wang, H.J. Kreuzer and R.G. Forbes, 1996. Field adsorption of helium and neon on metals: An integrated theory. *Surf. Sci.* **350**, 183–205.
58. N.D. Lang and W. Kohn, 1973. Theory of metal surfaces: Induced surface charge and image potential. *Phys. Rev. B* **7**, 3541–3550.
59. R.G. Forbes, 1999. The electrical surface as the centroid of the surface-induced charge. *Ultramicroscopy* **79**, 25–34.
60. S.C. Lam and R.J. Needs, 1992. Calculation of ionization rate-constants for the field-ion microscope. *Surf. Sci.* **277**, 359–369.
61. N. Ernst, G. Bozdech, H. Schmidt, W.A. Schmidt and G.L. Larkins, 1993. On the full-width at half-maximum of field-ion energy distributions. *Appl. Surf. Sci.* **67**, 111–117.
62. S.C. Lam and R.J. Needs, 1994. Theory of field ionization. *Appl. Surf. Sci.* **76/77**, 61–69 (and references therein).
63. W.A. Schmidt, Yu. Suchorski and J.H. Block, 1994. New aspects of field adsorption and accommodation in field ion imaging. *Surf. Sci.* **301**, 52–60.
64. K. Jousten, K. Böhringer and S. Kalbitzer, 1988. Current–voltage characteristics of a gas field ion source. *Appl. Phys. B* **46**, 313–323.
65. Y.C. Chen and D. N Seidman, 1971. On the atomic resolution of a field ion microscope, *Surf. Sci.* **26**, 61–84.
66. X. Liu and J. Orloff, 2005. Analytical model of a gas phase field ionization source. *J. Vac. Sci. Technol. B* **23**, 2816–2820.
67. R.D. Young. Theoretical total-energy distribution of field-emitted electrons. *Phys. Rev.* **113**, 110–114.
68. R.W. Newman, R.C. Sanwald and J.J. Hren, 1967. A method for indexing field ion micrographs. *J. Sci. Instrum.* **44**, 828–830.
69. T.J. Wilkes, G.D.W. Smith and D.A. Smith, 1974. On the quantitative analysis of field ion micrographs. *Metallography* **7**, 403–430.
70. M. Born and E. Wolf, 1975. *Principles of Optics* (Pergamon, Oxford).
71. W.R. Smythe, 1949. *Static and Dynamic Electricity* (McGraw Hill, New York).
72. W.P. Dyke, J.K. Trolan, W.W. Dolan and G. Barnes, 1953. The field emitter: Fabrication, electron microscopy, and electric field calculations. *J. Appl. Phys.* **24**, 570–576.
73. J. Orloff, M. Utlaut and L.W. Swanson, 2003. *High Resolution Focused Ion Beams: FIB and Its Applications* (Kluwer Academic, New York).
74. G.A. Schwind, G. Magera and L.W. Swanson, 2006. Comparison of parameters for Schottky and cold field emission sources. *J. Vac. Sci. Technol. B* **24**, 2897–2901.
75. D.J Rose, 1956. On the magnification and resolution of the field emission electron microscope. *J. Appl. Phys.* **27**, 215–220.

76. D.C. Joy, B.J. Griffin, J. Notte, L. Stern, S. McVey, B. Ward and C. Fenner, 2007. Device metrology with high-performance scanning ion beams. *Proc. SPIE* **6518**, 65181I.
77. P.W. Hawkes and B. Lençova, 2006. Charged particles optics theory. *E-nano-newsletter*, Issue 6 (PHANTOMS Foundation, Madrid). Available on [www.phantomsnet.net/Foundation/newsletters.php](http://www.phantomsnet.net/Foundation/newsletters.php).
78. W.A. Berth and C.A. Spindt, 1977. Characteristics of a volcano field ion quadrupole mass spectrometer. *Int. J. Mass Spectrom. Ion Phys.* **25**, 183–198.
79. B. Naranjo, J.K. Gimzewski and S. Putterman, 2005. Observation of nuclear fusion driven by a pyroelectric crystal. *Nature* **434**, 1115–1117.
80. B.J. Griffin and D.C. Joy, 2007. Imaging with the He scanning ion microscope and with low voltage field emission STEM – a comparison using carbon nanotube, platinum thin film, cleaved molybdenum disulphide samples and metal standards. *Acta Microscopica* **16** (Suppl. 2), 3–4.
81. H.D. Beckey, 1971. *Field Ionization Mass Spectrometry* (Pergamon, Oxford).
82. K.E. Drexler, 1996. *Engines of Creation: The Coming Era of Nanotechnology* (Anchor books, New York).
83. R.E. Smalley, 2001. Of chemistry, love and nanobots. *Sci. Am.* September, 56–57.
84. K.E. Drexler and R.E. Smalley, 2003. Nanotechnology: Drexler and Smalley make the case for and against ‘molecular assemblers.’ *Chem. Eng. News* (American Chemical Society, Washington D.C.) **81**(48), 37–42.
85. V.A. Nazarenko, 1970. Calculation of the current–voltage characteristics of a field-ion microscope. *Int. J. Mass Spectrom. Ion Phys.* **5**, 63–70.
86. H. Iwasaki and S. Nakamura, 1975. Ion current generation in the field ion microscope. I. Dynamic approach. II. Quasi-static approach. *Surf. Sci.* **52**, 588–596 and 597–614.
87. E.W. Müller, 1960. Field ionization and field ion microscopy. *Adv. Electron. Electron Phys.* **13**, 83–179.
88. J.A. Becker, 1958. Study of surfaces by using new tools. In: *Solid State Physics: Advances in Research and Applications* (F. Seitz and D. Turnbull, eds) (Academic, New York), Vol. 7, 379–424.
89. T.T. Tsong and E.W. Müller, 1966. Current–voltage characteristics by image photometry in a field-ion microscope. *J. Appl. Phys.* **37**, 3065–3070.

---

# 4 Magnetic Lenses for Electron Microscopy

*Katsushige Tsuno*

## CONTENTS

4.1	Introduction.....	129
4.2	Design Procedure of Magnetic Lenses .....	132
4.2.1	Design Procedure of Pole Pieces.....	132
4.2.2	Design Procedure of the Magnetic Circuit.....	136
4.2.2.1	Design of the Coil .....	136
4.2.2.2	Design of a Pole and a Yoke.....	140
4.2.3	Magnetic Materials.....	141
4.2.3.1	Saturation Magnetic Flux Densities.....	141
4.2.3.2	Homogeneity of Magnetic Properties of Lens Materials.....	143
4.3	Examples of Magnetic Lens Design .....	145
4.3.1	Aberration Correctors.....	145
4.3.1.1	Generation of Multipole Field Components with a Dodecapole.....	145
4.3.1.2	Hexapole Spherical Aberration Corrector with Transfer Doublet .....	145
4.3.1.3	Combined Electrostatic and Magnetic Quadrupole Lenses as a Chromatic Aberration Corrector.....	148
4.3.2	Objective Lens Design of Low-Energy Electron Microscope/Photoelectron Emission Microscope .....	150
4.3.3	Lotus Root Lens as a Multibeam Electron Lithography System .....	152
4.3.4	Various Objective Lenses for Low-Voltage Scanning Electron Microscope .....	153
4.3.5	Combined Electrostatic and Magnetic Lenses for Low-Voltage Scanning Electron Microscope .....	156
	For Further Information.....	157
	References .....	157

## 4.1 INTRODUCTION

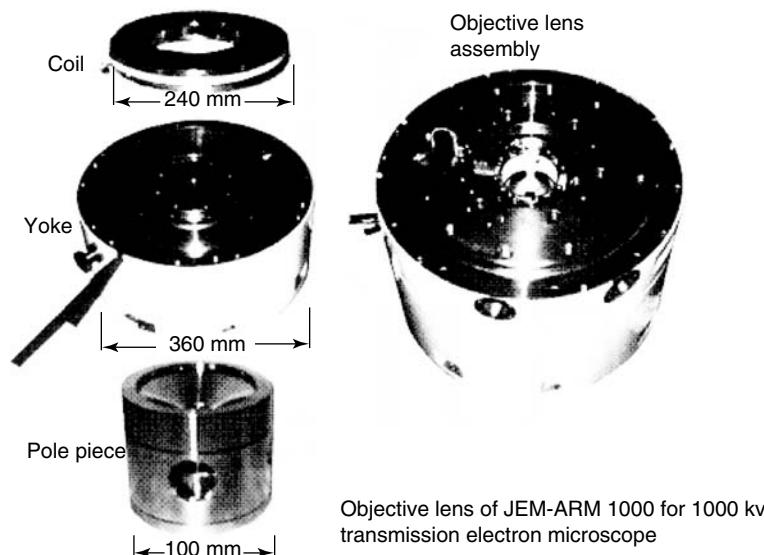
At the end of nineteenth century, it was found that an axially symmetric magnetic field has a focusing effect on an electron beam in a cathode ray oscilloscope: it acts as a lens. The effect is similar to that of a glass lens on light. This effect was first investigated by Busch in 1926, both theoretically and experimentally (Cosslett, 1991). Two important results were described in his theory. The first result is that the focal length of the lens decreases with an increase of magnetic flux density. If a lens is made not just of a coil of wire, but formed from a coil wrapped around a ferromagnetic (iron) yoke, the concentration of the magnetic flux by the iron yoke makes a very effective lens. The second result is that the electron beam rotates in the magnetic field. This rotation is simply proportional to the integral of the field strength and does not depend on the field distribution. Ruska and Knoll, who are well known as the inventors of the transmission electron microscope (TEM) in 1932, were the first to recognize the importance of the concentration of the field distribution by an iron

yoke and applied it to the production of lenses (Ruska, 1980). The concentration of the magnetic field made it possible to produce an electron microscope with magnification higher than the optical microscope. Their idea of the magnetic lens is still the basis of lenses at present (Mulvey, 1984).

Glaser showed, in his famous *bell-shaped* model calculation, that the condition for obtaining the minimum focal length and the minimum aberrations of a magnetic lens used in a TEM is to put the specimen at the middle of the gap of a pair of symmetrical pole pieces (Cosslett, 1991). Until Glaser's proposal, the specimen was placed in front of the objective lens. Half the field before the specimen is not used to form the image in the Glaser lens, so the excitation of objective lens has to be greatly increased. Moreover, the prefield above the specimen acts as a strong condenser lens. Such a strong objective lens, called the *condenser-objective* lens, was made by Riecke and Ruska in 1961 (Ruska, 1980). Unfortunately, the full power of the condenser-objective lens was not fully exploited at that time, because high-resolution electron microscopy still faced many technical problems, and the objective-lens aberrations were not the limiting factor for the resolution.

The resolution of an electron microscope is limited by (1) the stability of the instrument, in particular, the high voltage and the lens currents and (2) external disturbances such as vibration, contamination, charging, fluctuation of stray magnetic fields, and astigmatism produced by the lack of machining accuracy or the nonuniformity of the magnetic properties of the pole-piece material used. Between the 1940s and the 1960s, astigmatism introduced by the lack of machining accuracy was the main factor limiting the resolution. A stigmator was first installed in a commercial microscope in 1961–1962 to compensate the astigmatism (Kanaya, 1985).

A new era for high-resolution microscopy began in the 1970s after the technical problems had been solved. At first, high resolution was attained by using high-voltage microscopes ( $>500$  kV), because the resolution is proportional to three-fourth power of the electron wavelength and one-fourth power of the spherical aberration coefficient  $C_s$ . Figure 4.1 shows the objective lens of the 1000 kV TEM installed at Berkeley, California, in 1983 (Tsuno and Honda, 1983), which is a typical example of a magnetic lens. The lens consists of coils, poles, a yoke, and pole pieces. In this figure the upper half of the yoke can be seen in the assembled lens. The upper and lower pole pieces are fixed by a Be–Cu nonmagnetic spacer to maintain the mechanical accuracy of the lens bore center. Such a lens is called a *pole-piece lens* by Ruska. A resolution of 0.1 nm was attained using a 1250 kV microscope (Phillipp et al., 1994).



**FIGURE 4.1** An example of the objective lens and its components. (Photographs courtesy of Mr. H. Watanabe of JEOL Engineering Ltd.)

Up to about 1980, electron-optical calculations were primarily made analytically using a model axial field distribution such as the *bell* shape. The design of the lens shape was made by experience, which was only obtained after many experimental trials. Design technology of the objective lens using the finite element method (FEM) and a numerical calculation method for the electron-optical properties of magnetic lenses have been developed during the 1970s (Munro, 1973), and these methods made it possible to produce high-resolution microscopes with resolution between 0.17 and 0.21 nm using modest voltages (200 kV) or medium voltages (300–400 kV).

Owing to the rapid development of superconducting magnets, most of the high-field iron yoke electromagnets disappeared during the 1970s. Electron microscopes are an exception as they still use high-field iron yoke electromagnets. The reason for superconducting magnets not being commonly used in magnetic lenses is that the magnetic lens requires not only a high magnetic field but also a narrow (highly confined) field distribution, which cannot be generated by a superconducting coil. A superconducting shield lens using the Meissner effect has been used to make the narrow magnetic field (Lefranc et al., 1982). However, the technology for the production of such a lens was very difficult and inappropriate for commercial production. Recent development of superconducting material production technology and high-temperature superconductors offers a chance for the revival of the superconducting shielding electron lenses, however.

During the 1990s, successful application of aberration correctors resulted in a significant revolution in the field of electron lenses (see also Chapter 12). The first successful correctors were made for spherical and chromatic aberrations for a scanning electron microscope (SEM) (Zach and Haider, 1995), and this was followed by a spherical aberration corrector for TEM or scanning TEM (STEM) slightly afterward (Haider et al., 1995). The spherical aberration corrector is widely used in high-resolution TEM or STEM at present. Two methods are in use for correcting spherical aberration at high-accelerating voltage TEM or STEM. One uses magnetic hexapoles with transfer round lenses (Haider et al., 1995). Another system combines four magnetic quadrupoles and at least three octupoles (Krivanek et al., 1999). The hexapole system is used in both TEM and STEM, but the quadrupole-octupole system is exclusively used in STEM, due to its off-axis aberrations.

There are four methods mainly used in correcting both spherical and chromatic aberrations at low accelerating voltages for SEM, low-energy electron microscope (LEEM)/photoelectron emission microscope (PEEM) and focused ion beam systems (FIB): (1) combined magnetic and electrostatic quadrupoles can create negative chromatic aberration and octupoles generate aperture aberrations (Zach and Haider, 1995); (2) electrostatic quadrupoles and octupoles combined with retarding potential generate chromatic aberration, and octupoles are used to generate negative spherical aberration (Weiβbäcker and Rose, 2001); (3) a dispersion-free double focus Wien filter creates negative spherical and chromatic aberrations (Tsuno et al., 2005); and (4) an electrostatic mirror combined with a magnetic beam separator (Preikszas and Rose, 1997) can correct both spherical and chromatic aberrations.

Correction of chromatic aberration has not been successful in TEM or STEM at high voltages. It is now being tested by means of a quadrupole-octupole corrector system similar to the SEM corrector used at low voltages. At present, a monochromator is used to reduce the chromatic effect in TEM. Once both spherical and chromatic aberrations have been corrected, it will be possible to use objective lenses with a large bore and a large gap. Another recent approach is to use a low-voltage TEM, because the wavelength limit comes at a very low voltage, and the ability of accelerated electrons to transit the specimen is sufficient even at low voltages for the thin specimens needed for high-resolution observation. The low-voltage TEM may solve the difficulty of high voltages in generating a  $C_c$  corrector.

In the first edition of this chapter, objective lenses of high-resolution TEMs and low-voltage SEMs were mainly described. However, since the late 1990s, the main topic in the field of electron lenses has been the aberration correction. In high-resolution microscopy, interest is now not on the design of the shape of the lens, but on the design of aberration correctors. Electron lens design technology has spread to various electron-optical instruments such as positron microscopes, PEEMs,

and spin-polarized electron emission microscopes. Magnetic electron lenses are used to concentrate those electrons and illuminate them on the specimen. In such magnetic lenses, various special properties are required, which were not common in the usual TEM lenses. In this second edition, lenses for aberration correctors and various kinds of lenses for various purposes will be described (see also Chapters 6 and 12). Various applications of lens designs prior to the 1980s, which are not described here, are to be found in the book of Hawkes and Kasper (1989).

## 4.2 DESIGN PROCEDURE OF MAGNETIC LENSES

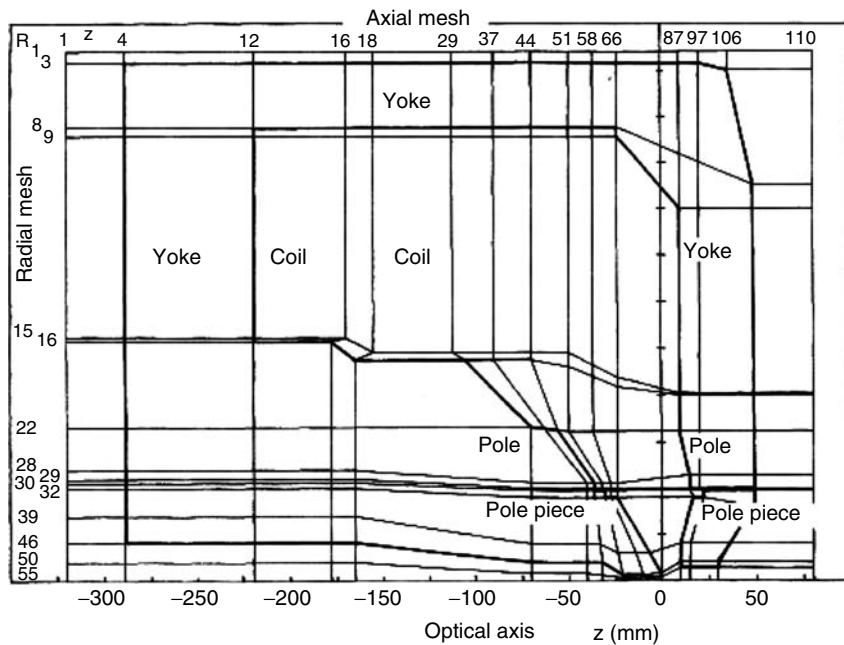
Design of a magnetic lens consists of two parts: (1) design of electron optics and (2) design of magnetic circuits including coils, yokes, and poles. Designers of electron optics easily forget about the second part. If experimentally obtained electron-optical properties are far from what is expected, the second part must be checked. The first step of the lens design is the optical design (Plies, 1994) to determine the optimum pole-piece shape and ampere-turns (AT) under given external conditions. The second step is the coil design. The main task here is the estimation of heat conduction. The third is the design of yokes and poles. The tool most often used is the FEM software. Because the technology of designing iron-cored electromagnets has largely disappeared due to the replacement of the electromagnets by the superconducting magnets, there are no new guide books for magnet design. Therefore, the technique necessary for the magnet design of electron lenses will be described here in detail.

### 4.2.1 DESIGN PROCEDURE OF POLE PIECES

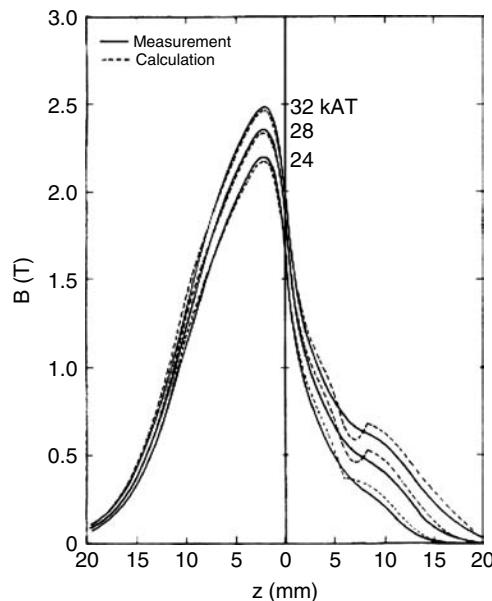
The numerical design procedure of electron lenses is as follows:

1. FEM calculation of magnetic flux density distributions in a model lens
2. Calculation of axial magnetic field distributions of the lens changing the excitation of the coil
3. Calculation of paraxial rays using the Runge-Kutta method with the axial magnetic field distributions
4. Calculation of aberration integrals using the Simpson method with the axial magnetic field distributions and the paraxial rays
5. Repetition of calculations 1–4 by changing the pole-piece shape to meet the optical requirements.

The FEM developed by Munro (see Appendix A) is based on a differential form in which the whole area (including empty space around the pole pieces) is divided into finite elements. In another approach, the integral form of the FEM was developed by Trowbridge (Biddlecombe and Trowbridge, 1984). In the Munro approach, a coarse mesh is generated using radial and axial lines; an example of coarse mesh lines for the lens of Figure 4.1 is shown in Figure 4.2. The coarse meshes are then divided into several fine meshes; this is done by the program automatically. The total number of fine meshes is limited by the size of the memory available and by the allowable computation time. The interval of the fine mesh is small at the gap of the lens, and it increases gradually toward the yoke and the pole, to save on the total number of meshes and hence computation time. It is important to avoid a sudden change of the mesh spacings (more than about a factor of two in the lines per millimeter), otherwise the accuracy of the field distribution will be compromised. The broken lines of Figure 4.3 show the calculated axial field distributions of the lens shown in Figure 4.1 (Tsuno and Honda, 1983). The solid lines are obtained by experiment. The hump in the calculated axial field distribution at the end point of the hole is an artifact caused by the sudden change of the mesh size. The exponentially increasing fine mesh layout developed by Hill and Smith (1982b) decreases such artifacts. An exponentially increasing automatic fine mesh generator was developed



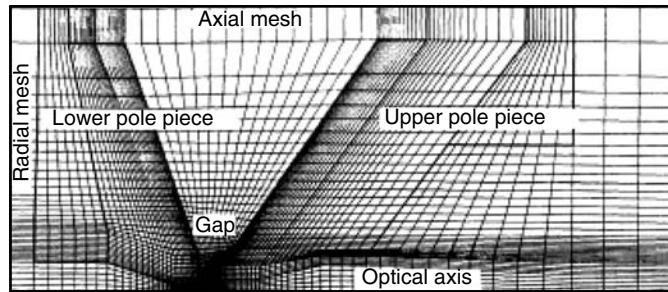
**FIGURE 4.2** Coarse mesh lines of the lens shown in Figure 4.1. (From Tsuno, K. and Honda, T., *Optik*, 64(4), 373, 1983. With permission.)



**FIGURE 4.3** Measured and calculated axial field distributions of the lens of Figure 4.1. (From Tsuno, K. and Honda, T., *Optik*, 64(4), 374, 1983. With permission.)

by the present author (see Lencová and Lenc, 1986). An example is shown in Figure 4.4. In this case, the gap was selected as the finest mesh region for the  $z$ -direction, and the optical axis (lowest line of the figure) was selected for the finest mesh region for axial mesh.

The incomplete Cholesky conjugate gradient (ICCG) method was introduced for solving the matrix of the FEM by Lencová and Lenc (1986). Their software reduces the memory and the



**FIGURE 4.4** An example of logarithmically increased fine mesh.

computation time about 1/100 compared with the previous software. The ICCG method is very effective for a regular matrix including many zeros. The  $(2I - 1)(2J - 1)$  matrix reduces to  $10I$ , where  $I$  and  $J$  are the radial and the axial fine mesh numbers, respectively.

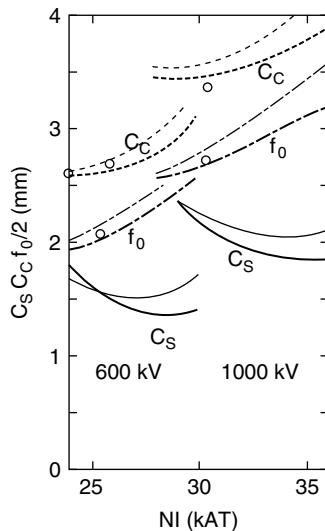
Magnetic saturation is treated using the Newton–Raphson method. The calculation is done in three steps. The first step is made without taking into account the magnetic saturation. A very high magnetic flux density is attained at the tip of the pole piece. The second calculation has to be made under the assumption of very low permeability, because of the high saturation in the first step, and results in a low magnetic flux density. The third calculation is made using very high permeabilities corresponding to the low flux densities. Thus, the computation never converges when the lens is severely saturated. Lencová and Lenc (1986) introduced a gradual method to solve this problem. The calculation starts at a level of excitation below saturation or partial saturation. The potential values of all the mesh points are stored in a file, which is used in the calculation of higher (increasing) AT. After the second calculation of the AT, the effect of the magnetic saturation is taken into account gradually and there is no sudden change of permeability. Only two or three iterations are sufficient to reach a convergence.

The difference in the peak field values between experiment and calculation was due to insufficient mesh points up to the 1980s, due to lack of computational power. However, this difficulty disappeared with the rapid increase of the memory size of personal computers. At present, the difference depends on the difference of the magnetization curves used in the calculation from the experiment. Unfortunately, the magnetic property of permendur, a widely used alloy for lens fabrication, is different from ingot to ingot and depends on the history of the heat and the mechanical treatments. Therefore, even if we measure the magnetization curve using a test piece made of the same ingot, it is not possible to know the actual magnetization curve of the lens itself, because of the differences induced by machining, heat treatment, and the size of the material.

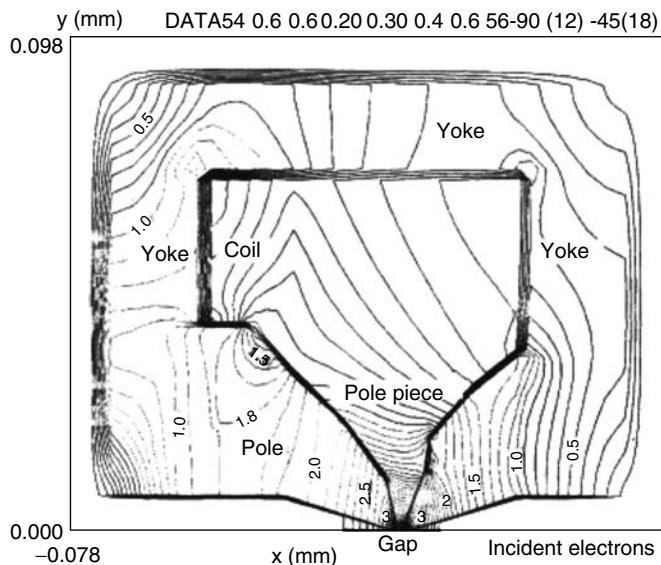
Figure 4.5 shows optical properties obtained from the calculated axial magnetic field distribution (thin lines) and the calculated values using the measured axial field distributions (bold lines) at 600 and 1000 kV (Tsuno and Honda, 1983). Circles indicate the optical properties obtained experimentally.

A model yoke, which has a small coil with high current density, is used during the optimization of the pole-piece shape. The purpose of using the model yoke is that we have not yet designed the coil and therefore have no knowledge about the size of the actual yoke. The length and width of the yoke can be determined only after the design of the coil. An example of the flux density distribution in the model yoke is shown in Figure 4.6. Basic points of a good flux density distribution are as follows:

1. Flux density contour lines are perpendicular to the optical axis and decrease with increasing distance from the gap.
2. In the gap, the flux density contour lines are nearly parallel to the optical axis.



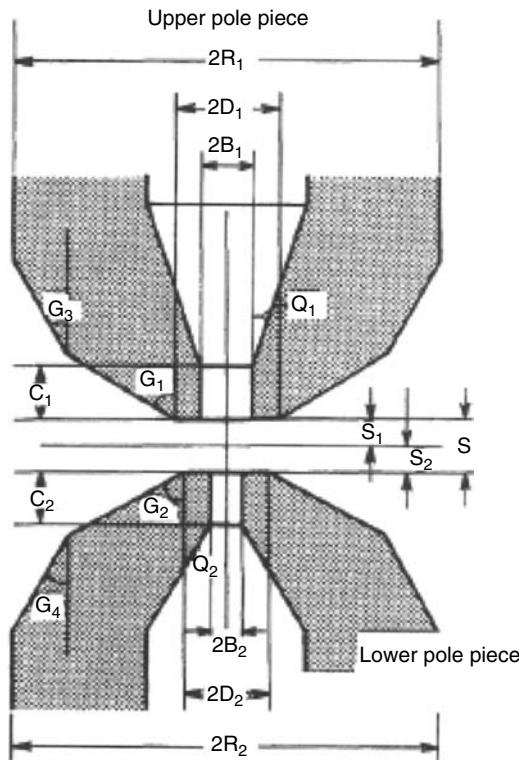
**FIGURE 4.5** Measured and calculated optical properties of the lens of Figure 4.1. (From Tsuno, K. and Honda, T., *Optik*, 64(4), 375, 1986. With permission.)



**FIGURE 4.6** Flux density contour map of the model lens.

The excitation NI (N, number of turns of the coil; and I, the excitation current) is chosen to give the optimum optical values at the specimen position.

In the case of Figure 4.6, two-stage tapers are used in both pole pieces. The first-stage tapers  $G_1$  and  $G_2$  shown in Figure 4.7 are important for obtaining a small spherical aberration coefficient ( $C_s$ ). The small value for  $G_1$  and the large value for  $G_2$  provide a low  $C_s$  value (in this case  $C_s = 0.35$  mm at 200 kV). If a single taper is used in both pole pieces, the flux density of the upper pole piece will be very high, and severe magnetic saturation does not allow sufficient field strength in the gap. To avoid saturation at the knee of the upper pole piece, a large second-taper  $G_3$  (see Figure 4.7) is important. On the other hand, the flux density in the lower pole piece is reasonable because of the



**FIGURE 4.7** Symbols of transmission electron microscope pole pieces.

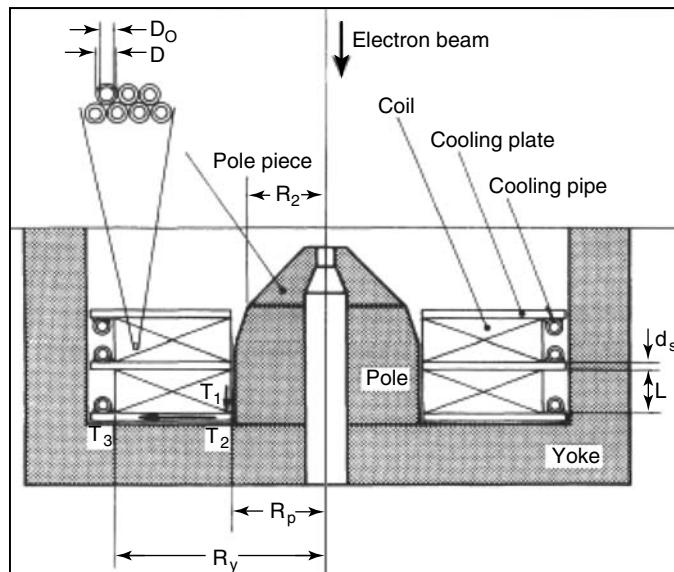
large taper  $G_2$ . The taper decreases continuously and there are no problems from the point of view of magnetic saturation. However, if the taper angles of both pole pieces are large, the magnetic resistance in the space between the tapered parts of the pole piece is small, and most of the flux leaks through the tapered part of the pole piece. This results in a reduction of the flux in the gap region. Therefore, it is desirable to keep a large distance between the upper and the lower pole pieces at the tapered portion of the lens. The decreased second-taper angle  $G_4$  is selected to keep a large distance between both pole pieces.

#### 4.2.2 DESIGN PROCEDURE OF THE MAGNETIC CIRCUIT

Figure 4.8 shows a schematic drawing of a half lens in which yoke, pole, and coils are indicated. To design a whole lens, the size of the coil has to be determined. The length of the pole and the radius of the yoke are determined so as to enclose the coil. If the excitation of the lens is small, the radius of the pole at the bottom side  $R_p$  can be made equal to  $R_2$  (or  $R_1$ ). However, if the excitation is large,  $R_p$  has to be made larger than  $R_2$  (or  $R_1$ ) to avoid magnetic saturation at the bottom of the pole: the radius  $R_p$  must be changed and the coil must be redesigned. The total shape is determined after some iterations of the design of the coil and the pole.

##### 4.2.2.1 Design of the Coil

The coil is made of a copper wire covered with an insulating coating. The total diameter  $D$  of the copper wire (e.g.,  $D = 1.102 \text{ mm}$ ; see Figure 4.8) is the sum of the diameter of copper  $D_0$  ( $D_0 = 1.00 \text{ mm}$ ) and twice the thickness of the insulator. There are two kinds of coils. One has a circular cross section (cylindrical wires) and the other a flat cross section (rectangular wires). The former is mainly used in low-current electromagnets whose maximum current is  $<20 \text{ A}$  whereas the latter is



**FIGURE 4.8** Symbols and names of the magnetic circuit of the magnetic lens. Only the lower half of the lens is shown.

used in magnets with higher current. High-voltage electron microscopes (1000 kV or more) use the flat rectangular wires, and ordinary voltage microscopes (100–200 kV or less) use cylindrical wires. For the middle-voltage microscopes (300–400 kV), cooling efficiency is not sufficient for the cylindrical wires, and the electrical resistance of the wire is too low for the flat rectangular coils. For those microscope lenses, a thin copper sheet with a narrow thickness and a large width is sometimes used to both optimize the cooling efficiency and obtain a reasonable resistance of the wire. Such copper sheet coils (tape-winding coil) were used by Mulvey (1982) for his miniature lenses. Although the excitation NI is determined during the electron-optical design, there are a variety of choices in N and I.

Recently it has been necessary that some lenses need to be constructed without water cooling. There are a variety of applications that do not allow the use of water cooling, for example, where a high voltage is present, in a high-vacuum chamber, use of the lens in a movable environment, and so on. In such cases, the temperature increase and the total wattage must be kept low, and it is necessary to think about the heat conduction away from the lens.

As shown as  $T_1$ ,  $T_2$ , and  $T_3$  in Figure 4.8, the process of the heat flow in the coil and the cooling plate is considered as follows:

1. In the coil, joule heat generated by the electrical resistance (temperature  $T_1$  at the middle of the coil) propagates to the cooling plate along the axis (temperature  $T_2$ ); ( $T_1 \rightarrow T_2$ ).
2. In the cooling plate, the heat propagates radially to the cooling pipe (temperature  $T_3$  at the pipe); ( $T_2 \rightarrow T_3$ ).

The circular line where the copper cooling pipe is soldered is at the lowest temperature  $T_3$ . If two cooling pipes are used at both sides of the cooling plates, the middle line of the plate has the highest temperature  $T_2$ . In the latter case, the temperature difference in the cooling plate will be 1/2–1/3 of the former case. All joule heat  $q_0$  generated in the coil has to be removed by the cooling water. The coil must not have thermal contact with the pole or the yoke but must have contact with the cooling plates. In the following calculation, the inside radius of the coil  $R_p$  is assumed to be equal to the outer radius of the pole for simplicity, although there is a gap between the pole and the coil to separate them thermally in the actual lens.

Because the heat conduction of copper is 1000 times larger than that of the insulation material, heat conductivity of the coil is practically determined by the insulator. When the coil has a small width  $L$  and a large thickness ( $R_y - R_p$ ), condition 1 is fulfilled. The heat conduction is determined by the total thickness  $N_c(D - D_0)$  of the insulator in the radial direction, where  $N_c$  is the number of turns in the axial direction. If the thickness  $d_s$  of the cooling plate is small enough compared with its outer radius, condition 2 is fulfilled.

Now, we shall take the temperature of the cooling water ( $T_3$ ) to be  $T_3 = 20^\circ\text{C}$  and the highest allowable value of the temperature in the coil to be  $T_1 = 100^\circ\text{C}$ . It is possible to choose  $T_1 = 200\text{--}400^\circ\text{C}$ , if we use a special coil. However, to prevent the drift of the image, it is desirable to choose a lower value  $T_1$  in the electron microscope design. The temperature difference in this case is  $T_1 - T_3 = 80^\circ\text{C}$ .

#### 4.2.2.1.1 Heat Transfer in the Coil and the Cooling Plates

The following analysis follows that of Kroon (1968). The equilibrium temperature distribution of a medium can be expressed through the diffusion equation

$$\nabla^2 T = -q_0/M \quad (4.1)$$

where  $M$  is the thermal conductivity and  $q_0$  the amount of heat generated per second per cubic millimeter. In a coil,  $M$  is a function of the position because the coil consists of a repetitive structure of copper, insulator, and empty space. The diffusion equation is simplified if the average thermal conductivity  $M_c$  of the coil is introduced. For cylindrical wires,  $M_c$  is ~10 to 15 times that of the thermal conductivity of the insulating coating. Polyester has  $M = 1.8 \times 10^{-4} \text{ W/K} \cdot \text{mm}$ , and  $M_c$  can be assumed to be  $M_c = 2 \times 10^{-3} \text{ W/K} \cdot \text{mm}$ . On the other hand,  $M(\text{Cu}) = 0.4 \text{ W/K} \cdot \text{mm}$ . The maximum temperature  $T_1$  is given by

$$T_1 = T_2 + (q_0/2M_c)(L/2)^2 \quad (4.2)$$

The total heat  $q_0$  is found using the power  $P$  spent in the coil.

$$P = \pi(R_y^2 - R_p^2)Lq_0 \quad (4.3)$$

where  $L$  is the length of the coil (see Figure 4.8). Thus

$$T_1 - T_2 = LP / \left[ 8\pi M_c (R_y^2 - R_p^2) \right] \quad (4.4)$$

By using the well-known equation for the power

$$P = I^2 R_c \quad (4.5)$$

the resistance of the coil  $R_c$  is written as

$$R_c = s_0 L_{\text{total}} / S \quad (4.6)$$

where  $s_0$  is the electrical conductivity of copper, which depends on the temperature  $T$  (in degrees centigrade) according to

$$s_0(T) = 1.7241 \times 10^{-5} (1 + 0.00393T) \Omega \text{mm} \quad (4.7)$$

At an average copper temperature of  $50^\circ\text{C}$ ,  $s_0(50^\circ\text{C}) = 2.063 \times 10^{-5} \Omega\text{mm}$ .  $S$  is the cross section of the copper wire.

$$S = f(R_y - R_p)L/N \quad (4.8)$$

and  $L_{\text{total}}$  is the total length of the coil

$$L_{\text{total}} = f\pi L \left( R_y^2 - R_p^2 \right) / S \quad (4.9)$$

where the space factor  $f$  depends on the shape and the size of the coil

$$f = 3^{1/2} \pi (D_0/2)^2 / [6(D/2)^2] \quad (4.10)$$

We then obtain the final temperature difference

$$T_1 - T_2 = s_0 N^2 I^2 / [8M_c f (R_y - R_p)^2] = 0.167 \times 10^{-2} (NI)^2 / (R_y - R_p)^2 \quad (4.11)$$

The diffusion equation in the cooling plate is simplified using condition 2:

$$d^2T/dr^2 + (1/r)dT/dr = -Q_0/M_s d_s \quad (4.12)$$

where  $M_s$  is the thermal conductivity of the copper plate  $M_s = 0.4 \text{ W/K mm}$  and  $Q_0 (\text{W/mm}^2)$  the amount of heat flowing into the cooling plate. If  $Q_0 = q_0$  is satisfied, all the heat generated in the coil is removed through the cooling plate, and, therefore, there is no temperature increase in the lens. The solution of the diffusion equation for one cooling pipe is

$$T_3 - T_2 = (P/4\pi M_s d_s) \left[ 1 - \left[ 2R_0^2 / (R_y^2 - R_p^2) \right] \ln(R_y/R_p) \right] \quad (4.13)$$

If we put  $P = 1000 \text{ VA}$ ,  $R_p = 45 \text{ mm}$ ,  $R_y = 105 \text{ mm}$ , and  $d_s = 8 \text{ mm}$ , then the obtained value of  $T_3 - T_2$  is  $<20^\circ\text{C}$ . If we use a two pipe cooling plate, the maximum temperature difference will be  $<10^\circ\text{C}$ .

#### 4.2.2.1.2 Power Supply and Selection of the Wire for the Coil

If we assume  $T_1 = 100^\circ\text{C}$ ,  $T_3 = 20^\circ\text{C}$ ,  $T_3 - T_2 = 10^\circ\text{C}$ , and  $NI = 15,000 \text{ AT}$ , Equations 4.11 and 4.13 imply  $R_y - R_p = 68.5 \text{ mm}$ . For the value of  $R_p = 45 \text{ mm}$ , the value of  $R_y$  is obtained from Equation 4.4 as  $R_y = 113.5 \text{ mm}$ . If  $R_y$  is larger than the expected value, it can be reduced by dividing the coil into two pieces as shown in Figure 4.8. These calculations tell us an important result, namely, the parameter obtained from the consideration of the heat flow, the height of the coil  $R_y$ . The diameter of the wire  $D_0$  and the power VI (V, voltage and I, current of the power supply) are still undetermined.

Now, power  $P$  is expressed as

$$P = VI = s\pi(R_y + R_p)N^2 I^2 / [f(R_y - R_p)L] \quad (4.14)$$

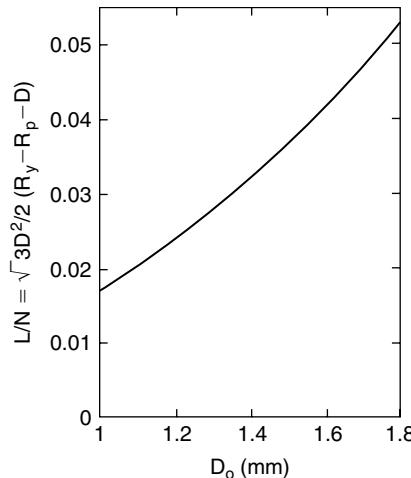
In the present case,  $VI = 34,224.7/\text{L}$ .

The next parameter we should decide on is the power supply. There are two typical cases for the power supply: (1) high current and low voltage (low impedance) and (2) low current and high voltage (high impedance). High-impedance electromagnets using a few kilovolts and a few amperes were common when vacuum tubes were used as power supplies. However, after the development of power transistors, low-impedance magnets became popular because of their compact size and low-cost performance. In the case of 100–200 kV electron microscopes, middle-impedance power supplies with 80–190 V and 1–20 A are used. The selection of the impedance is mainly driven by economics.

For example, if we have a power supply of  $15 \text{ A} \times 85 \text{ V}$ ,  $L$  is derived from Equation 4.14 as  $L = 26.85 \text{ mm}$ . The number of turns  $N$  can be determined using the following equation:

$$L = 3^{1/2} ND^2 / [2(R_y - R_p - D)] \quad (4.15)$$

For the case of  $NI = 13,000 \text{ AT}$ ,  $N$  is found to be 867 turns.



**FIGURE 4.9** Length of the coil versus the diameter of the wires used as a coil.

**TABLE 4.1**

**The Relation between the Outer Diameter of Coil D and the Diameter of Copper  $D_0$  (mm)**

$D_0$	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6	1.8
$D$	0.246	0.456	0.672	0.882	1.102	1.304	1.508	1.712	1.914

*Note:* These values are obtained from the Japanese industrial standards (JIS).

Figure 4.9 shows the relation of Equation 4.15. As  $L/N = 26.85/867 = 0.031$ , in our case, the diameter of the copper part of the wire  $D_0$  is found to be 1.4 mm. ( $D_0$  is selected to be a discrete value.) The relation between  $D$  and  $D_0$  is shown in Table 4.1.  $L$  is recalculated using  $D = 1.508$  mm, which corresponds to  $D_0 = 1.4$  mm. The obtained value is  $L = 27.6$  mm. The number of turns for the axial direction  $N_c$  is obtained from

$$N_c = \text{Integer}(L/D) = 18 \quad (4.16)$$

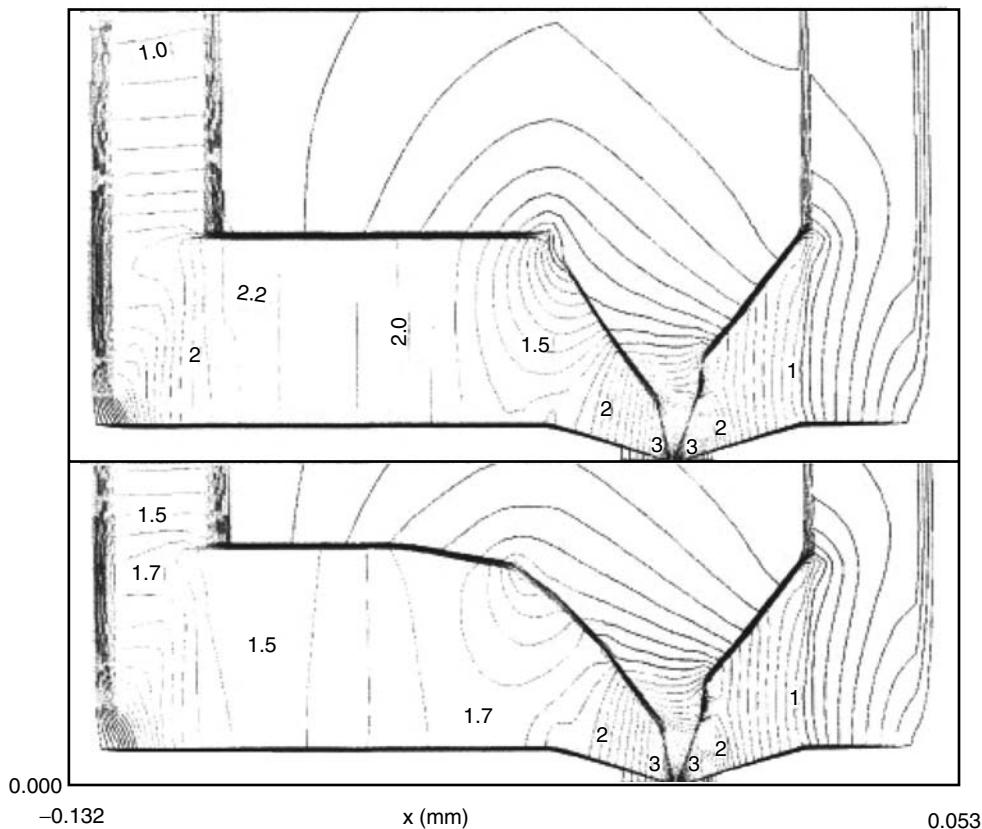
The number of turns for the radial direction  $N_h$  is

$$N_h = \text{Integer}(N/N_c) + 1 = 49 \quad (4.17)$$

Thus, all the values necessary for the coil design have been derived.

#### 4.2.2.2 Design of a Pole and a Yoke

Using the length  $L$  and the outer radius  $R_y$  of the coil, the dimensions of the pole and the yoke are now determined. The upper part of Figure 4.10 shows the calculated result of the flux density distributions in the central half of the lens using the actual dimensions of the coil ( $R_y$  and  $L$ ). In this case,  $R_p$  is set equal to the radius of the root of the pole piece  $R_2$  to show what happens in the pole when the excitation of the lens is large. The highest flux density reaches 2.2 T around the root of the pole, which exceeds the saturation flux density of pure iron (2.14 T). The saturation occurs as a result of the root radius of the pole being too small. However, at the region near the pole pieces the flux density is relatively low. It is possible to decrease  $R_p$  instead of increasing  $R_y$ . The lower part of Figure 4.10 shows the revised flux density contour map in which  $R_p$  is increased and  $R_2$  is decreased. As seen from the lower part of the figure, the magnetic saturation at the root of the pole has disappeared.



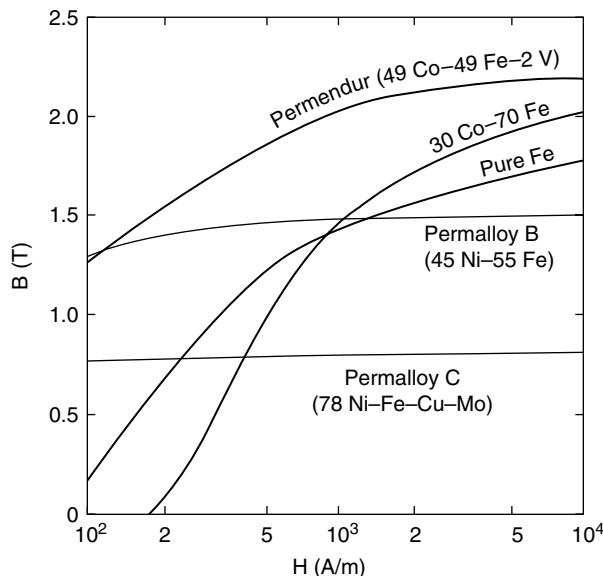
**FIGURE 4.10** Flux density contour lines of a lens. Upper, the first design; lower, improved design to reduce the saturation at the root of the pole.

The desired maximum flux density of pure iron, which is used for the yoke and the pole, is 1.7 T. If the flux density of pure iron exceeds this value, a leakage flux appears along the bore and may cause astigmatism in the image. The desirable flux density in the yoke is between 0.7 and 1.2 T at its highest permeability. The magnetic flux density distribution in the yoke or in the pole is greatly changed by a small change of its radius. It is quite important to check the flux density distribution in the whole lens.

#### 4.2.3 MAGNETIC MATERIALS

##### 4.2.3.1 Saturation Magnetic Flux Densities

Various magnetization curves for materials that are used as lens materials are shown in Figure 4.11. Pure iron is a popular magnetic lens material, the reason being its high saturation magnetic flux density  $B_s$  and its reasonable cost. Although the saturation magnetic flux density  $B_s$  of pure iron is  $B_s = 2.14$  T, it has been thought that the technical saturation is only 1.7 T. When the flux density in the pole exceeds this value, it is better to use higher  $B_s$  material. The reason is high crystal magnetic anisotropy of iron—the first anisotropy constant  $K_1$  is listed in Table 4.2 (Hall, 1960). It is quite difficult to magnetize the crystal in the  $\langle 111 \rangle$  direction. Therefore, the permeability of polycrystalline pure iron is very low. The permeability of pure iron increases when the crystalline size of the material is increased. When the purity of iron is increased, the size of the crystal is increased and so the permeability is increased. Therefore, it has been believed that higher purity and larger crystal size



**FIGURE 4.11** Magnetization curves for lens materials.

**TABLE 4.2**  
**The First Anisotropy Constant  $K_1$  of Iron and Iron–Cobalt Alloys**

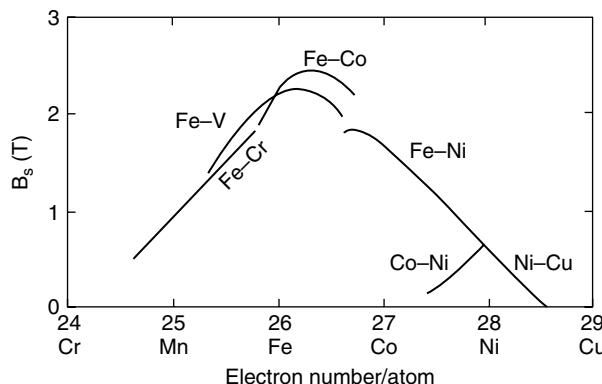
Pure Iron ( $J/m^3$ )	30% Co–Fe ( $J/m^3$ )	50% Co–Fe (permendur) ( $J/m^3$ )
$4.8 \times 10^4$	$3.5 \times 10^4$	$0 \times 10^4$ (slow cooled) $1.3 \times 10^4$ (quenched)

*Source:* From Hall, R.C., *J. Appl. Phys.*, 31(5), 157S, 1960. With permission.

are necessary as a magnetic material. However, as will be shown later, this is not true for magnetic lens material.

Permendur (49% Co–49% Fe–2% V alloy) is usually used as the pole-piece material when high field strength ( $>2$  T) is necessary in the gap. The difference of saturation flux density  $B_s$  between pure iron and permendur ( $B_s = 2.3$  T) is only 10%. The significance of permendur is its high permeability and high saturation flux density. Owing to the low technical saturation of pure iron, permendur is widely used as a magnetic lens material. The anisotropy constant decreases with the increase of the cobalt content as shown in Table 4.2. At the concentration near Fe:Co = 1:1, it forms an ordered alloy when it is cooled slowly. At the weight ratio of 1:1 of Fe:Co, which is the concentration of permendur, the anisotropy constant is nearly zero. Vanadium is added to improve the machinability of permendur. A nonmagnetic  $\gamma$ -phase appears in Fe–Co–V ternary alloy (Chen, 1961). The content of the  $\gamma$ -phase is different according to the difference in the vanadium concentration and the mechanical and heat treatments. Therefore,  $B_s$  depends on the percentage of the nonmagnetic phase. The magnetization curve of permendur is quite different for different production runs, as was indicated by Mulvey and Tahir (1990).

Figure 4.12 shows the famous Slater–Pauling curve (Slater, 1937; Pauling, 1938). The saturation magnetic flux density depends on the electron number per atom. The maximum of the magnetic flux density appears at 30–35% Co–Fe ( $B_s = 2.43$  T) (Pshechenkova and Skokov, 1966). This material is sometimes used as a pole-piece material because it has the highest  $B_s$ . However, in low external field  $H$ , the magnetic flux density  $B$  is always lower than permendur as can be seen in Figure 4.11.



**FIGURE 4.12** Slater–Pauling curve.

The material is very brittle, has high hardness, and is very difficult to machine. Therefore, the material is used for the production of extremely high magnetic field. The rare earth metal holmium has high saturation flux density ( $B_s = 2.75$  T) at liquid-helium temperature and was once used in an actual TEM (Bonjour, 1975).

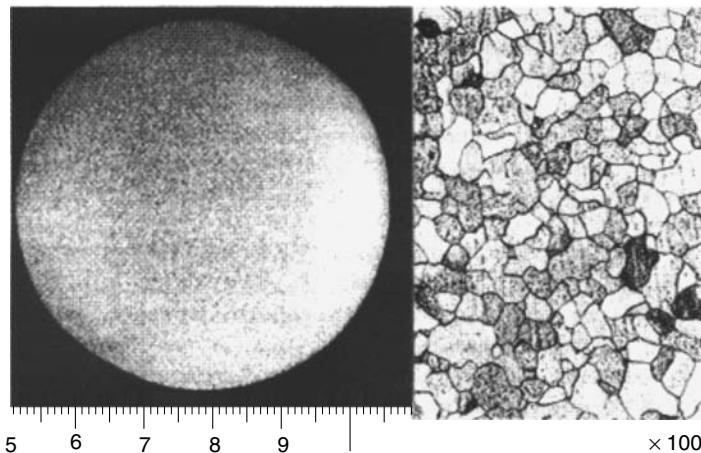
Permalloy is a good material to use to avoid magnetic hysteresis and to obtain high magnetic-shielding properties. There are two permalloys. One is permalloy C (78% Ni–Fe–Mo–Cu alloy) and the other permalloy B (45% Ni–Fe). The former has the highest permeability in metallic magnetic materials when it is produced as a thin sheet. The  $B_s$  of permalloy C is only 0.8 T and the material is quite sensitive to mechanical stress and heat treatment. According to the author's experience, permalloy B is better than permalloy C as a magnetic lens material. This is because the permeability of permalloy B is high enough that it can be used in a lens, and it is not so sensitive to stress. Moreover, the  $B_s$  of permalloy B is about twice that of permalloy C (1.5 T, see Figure 4.11).

The use of permalloy B has increased in recent years to avoid residual magnetic fields and for the use of lenses in a high-vacuum environment. It is also used for multipoles of aberration correctors and monochromators. When pure iron is used as a pole piece or a multipole, a residual magnetic field strength of a few militeslas cannot be avoided. However, such a residual magnetic field can cause a deflection of the beam, especially in the case of aberration correctors. It is advantageous to use permalloy B in such a case. Moreover, permalloy is believed to be a more suitable material than the pure iron in high vacuum, because of its similarity in the atomic concentration to stainless steel.

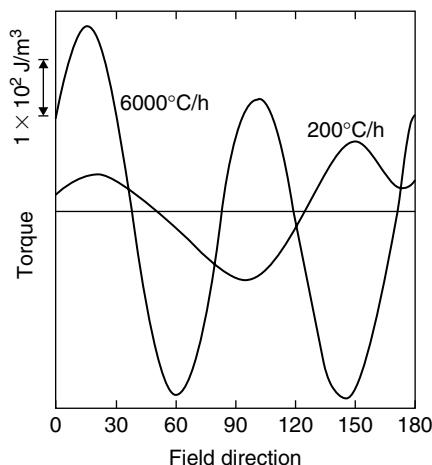
#### 4.2.3.2 Homogeneity of Magnetic Properties of Lens Materials

It has long been known that inhomogeneity of magnetic properties of pole-piece material has a severe influence on the onset of astigmatism. One cause of astigmatism is a deviation of the magnetic field distribution in a lens from position to position. It is well known from experience that the crystallite size is quite important in pure iron pole pieces. Larger crystallite sizes (such as 1 mm) cause an inhomogeneous distribution of the permeability, because the crystal magnetic anisotropy of iron is quite large (see Table 4.2). The allowable crystal size of pure iron is <0.1 mm (30–50  $\mu\text{m}$  is desirable). The positional dependence of the permeability causes an anisotropic distribution of the leakage field distribution in the bore of the lens. This anisotropic leakage field distribution is an origin of astigmatism.

The amount of impurities such as carbon and manganese is usually kept very small in pure iron to improve its magnetic properties. However, the absence of these impurities results in a large crystal size. A special pure iron rod, whose crystal size is small and whose distribution is uniform over the cross section of the bar, was successfully developed in the 1960s. The iron rod, whose macro- and microstructures are shown in Figure 4.13, was developed by T. Kato of Daido Special



**FIGURE 4.13** Macro- and microstructures of the pure iron specially developed as a pole-piece material.



**FIGURE 4.14** Torque curves of a permendur used as a pole piece, which shows directional magnetic properties of the material.

Metals Co. Ltd. as an electron lens material. Small amounts of C or Mn are added to minimize the crystal size.

It is better to use permendur in the region where the flux density exceeds 1.7 T to avoid the appearance of leakage fields. In the case of permendur, the crystal size is not so critical as in pure iron because the crystal magnetic anisotropy of permendur is nearly zero (see Table 4.2). A size of 0.2–0.5 mm is allowable. However, homogeneous forging or rolling is critically important to avoid the appearance of astigmatism when using this material. The induced magnetic anisotropy appears due to directional forging. It is well known that the cold rolling of the magnetic alloy induces a directional ordering and produces an induced uniaxial magnetic anisotropy. In the case of permendur,  $1 \times 10^4 \text{ J/m}^3$  order of uniaxial magnetic anisotropy is observed (English et al., 1967). This is about one-fifth of the crystalline anisotropy of iron. Therefore, the production method should be axially symmetric to avoid induced magnetic anisotropy. Figure 4.14 shows torque curves obtained from a forged permendur rod. The amount of the uniaxial anisotropy component is estimated to be of the order of  $10^2 \text{ J/m}^3$ . One of the curves is obtained by quenching after an 800°C heat treatment. When the specimen is quenched, a large fourfold symmetry component caused by the crystalline anisotropy appears even in a polycrystal.

The crystal size of permalloy sometimes reaches a few millimeters. However, the negligible crystal anisotropy reduces the influence of the crystal size on astigmatism, and the maximum allowable size is larger than that in the case of pure iron. A problem with the large crystal size of permalloy is the possibility of fracturing of a crystallite from the top of the pole piece or the multipole.

## 4.3 EXAMPLES OF MAGNETIC LENS DESIGN

### 4.3.1 ABERRATION CORRECTORS

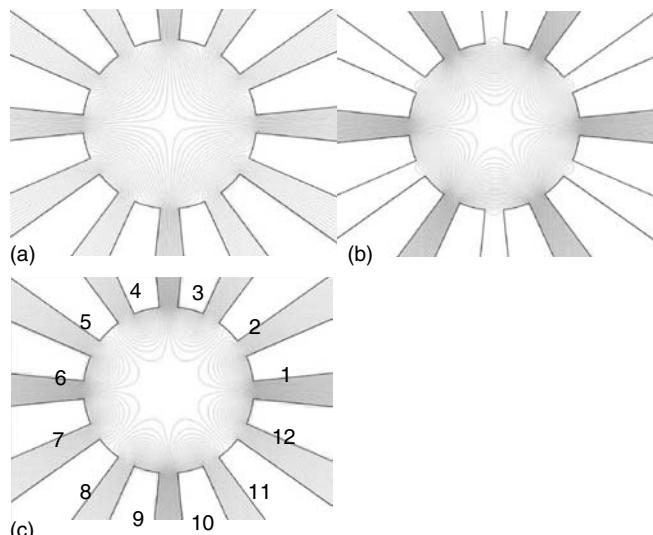
An aberration corrector made of two hexapoles combined with transfer round lenses can be used as a spherical aberration corrector, and combined magnetic and electrostatic quadrupole lenses can be used as a chromatic aberration corrector. In the former case, the role of the round lenses (the first transfer doublet) is to transfer the diffraction pattern to the corrector. The second transfer doublet is used to reverse the electron trajectory to cancel out the second-order aberrations induced by the hexapole. In the latter case, quadrupole lenses make the beam shape astigmatic in the x- and the y-directions and generate negative chromatic aberration in each direction.

#### 4.3.1.1 Generation of Multipole Field Components with a Dodecapole

As is well known, a round lens cannot have negative or zero spherical or chromatic aberration. A multipole lens is therefore used to create negative aberrations. Figures 4.15a through 4.15c show cross sections of the pole regions of a dodecapole and its flux distributions excited for generating quadrupole (Figure 4.15a), hexapole (Figure 4.15b), and octupole (Figure 4.15c) field components. Excitations of each pole for generating various components are listed in Table 4.3. A dodecapole corrector is used to compensate machining errors and assembly errors of multipoles (although it is necessary to use a lot of power supplies).

#### 4.3.1.2 Hexapole Spherical Aberration Corrector with Transfer Doublet

A hexapole generates a deflection field as shown in Figure 4.16. The upper electron beam (A in the figure) is deflected toward the upper pole and the lower beam (B in the figure) toward the center. The magnetic field becomes stronger when the beam approaches the pole surface and weaker when



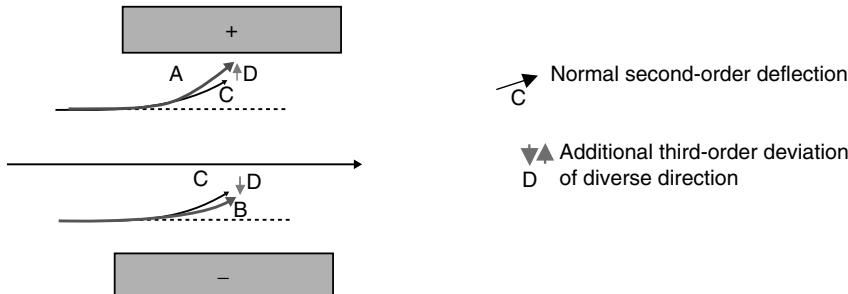
**FIGURE 4.15** (a) Quadrupole, (b) hexapole, and (c) octupole magnetic fields created by dodecapole.

**TABLE 4.3**  
**Generation of Magnetic Multipole Components**  
**by Using a Dodecapole Shown in Figure 4.15**

Poles	Angles	$B_2$	$B_3$	$B_4$
1	15	$NI_2/2$	$NI_3$	$NI_4$
2	45	$NI_2$	$NI_3$	0
3	75	$NI_2/2$	$-NI_3$	$-NI_4$
4	105	$-NI_2/2$	$-NI_3$	$NI_4$
5	135	$-NI_2$	$NI_3$	0
6	165	$-NI_2/2$	$NI_3$	$-NI_4$
7	195	$NI_2/2$	$-NI_3$	$NI_4$
8	225	$NI_2$	$-NI_3$	0
9	255	$NI_2/2$	$NI_3$	$-NI_4$
10	285	$-NI_2/2$	$NI_3$	$NI_4$
11	315	$-NI_2$	$-NI_3$	0
12	345	$-NI_2/2$	$-NI_3$	$-NI_4$

$B_2$ , quadrupole;  $B_3$ , hexapole; and  $B_4$ , octupole components.

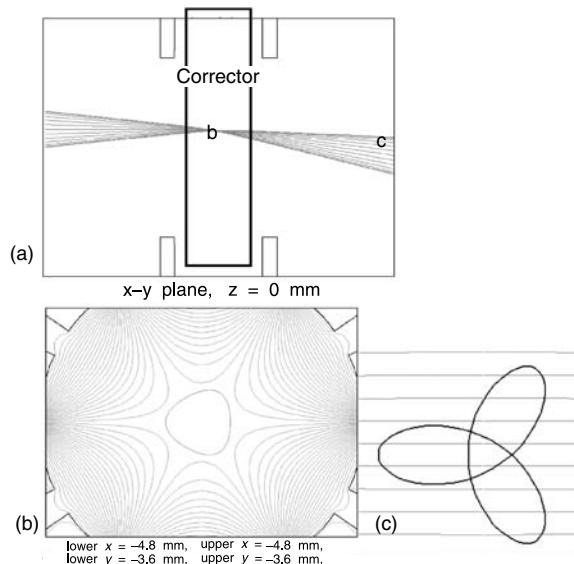
Source: From Tsuno, K., Ioanoviciu, D., Martínez, G., *J. Microscopy*, 217, 205–215, 2005. With permission.



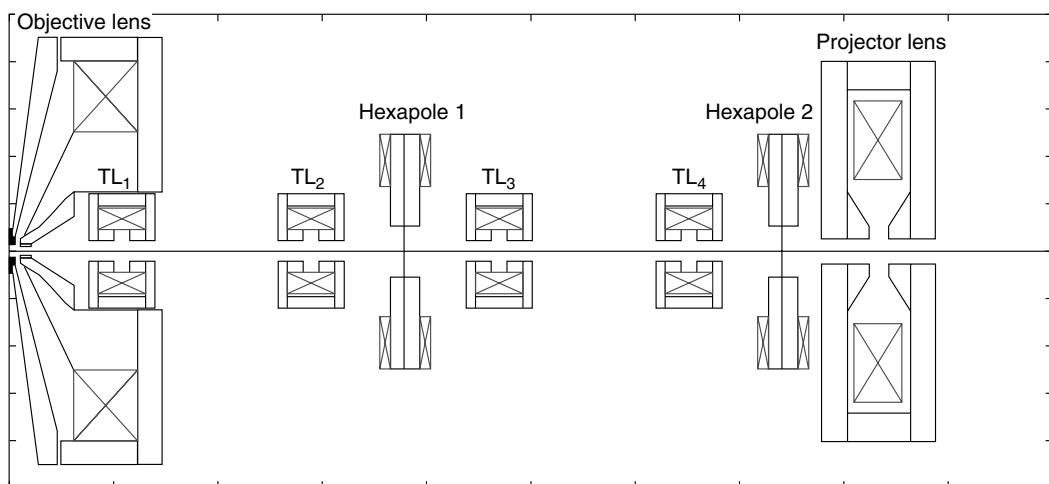
**FIGURE 4.16** A simple model of deflecting electron beams by a hexapole field and its generation of axially symmetric spherical aberration.

it approaches the optical axis. The beam A is deflected more when it approaches the pole. On the other hand, in the beam B, the bending angle becomes smaller when it approaches the center. The additional deflection is always in the radial direction. This additional axially symmetric deflection is the negative spherical aberration. In the case of nonaxially symmetric system, there are four aperture aberrations, which correspond to the spherical aberration of the round lens system. However, when the hexapole corrector is used, all of the four aperture aberrations are of the same value and do not have to be adjusted. The round input electron beam is deformed to a triangular shape due to the deflection of the beam. This shape is due to the second-order geometrical aberrations. A double hexapole and a transfer doublet are used to cancel these second-order geometrical aberrations. The reason for the necessity of the transfer doublet is that the hexapoles have no lens action. Cancellation of the second-order geometrical aberrations in the double focus lens is well known.

Figure 4.17a shows electron trajectories of a beam focusing 1 mm upstream from the center of the hexapole. As is clearly seen, the beam is deflected downward. Figures 4.17b and 4.17c show beam shapes (aberration figures) at the middle of and after exiting from the hexapole, respectively, together with the contour lines of the magnetic potentials. A round beam becomes a triangle at the



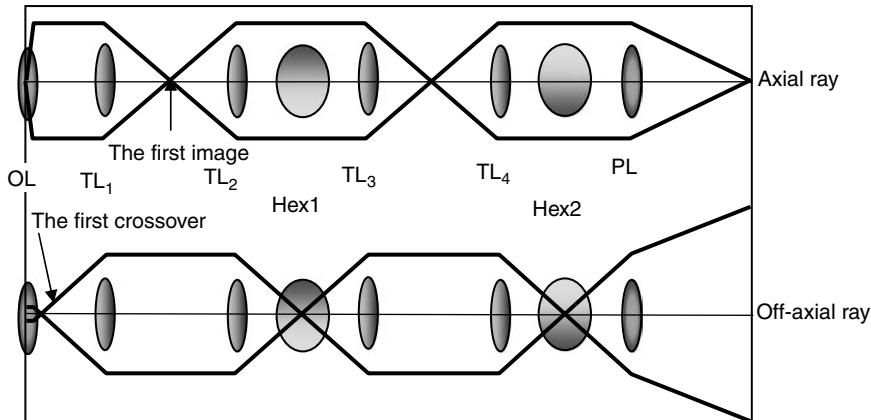
**FIGURE 4.17** (a) Electron beam deflection inside the hexapole corrector but 1 mm up from the center; (b) beam shape at the center of the corrector, indicated b in (a); and (c) the beam shape after exiting from the corrector as indicated c in (a).



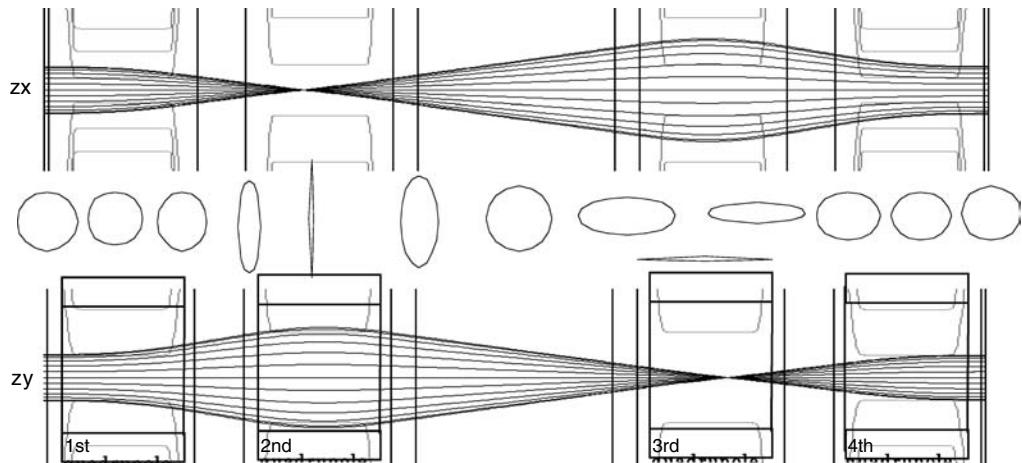
**FIGURE 4.18** Configuration of the objective lens of low-energy electron microscope, the first (TL<sub>1</sub> and TL<sub>2</sub>) and the second (TL<sub>3</sub> and TL<sub>4</sub>) transfer doublets, two hexapole correctors (hexapoles 1 and 2), and projector lens.

center and is deformed to a leaf shape at the exit of the first hexapole. The second hexapole is used to cancel out the distorted beam shape.

As an example of spherical aberration correction, the corrector is connected to a LEEM accelerated at 20 kV. Cross section of the system, including the objective lens of the LEEM, hexapole corrector, and two transfer doublets, is shown in Figure 4.18. Two principal rays are shown in Figure 4.19. The reflected beam at the specimen is accelerated between the specimen and the first anode. It focuses at the bottom of the objective lens when it is used without the corrector. The off-axial ray that left the specimen parallel to the axis makes a crossover at the focal point of the objective lens. In the system with the corrector, two round lenses (the first transfer doublet TL<sub>1</sub>



**FIGURE 4.19** Schematic diagram of electron trajectories of the axial ray (electrons exited from the center of the specimen) and the field ray (electron exited parallel to the optical axis from off-axial region of the specimen) from the specimen to the detection plane through the corrector.

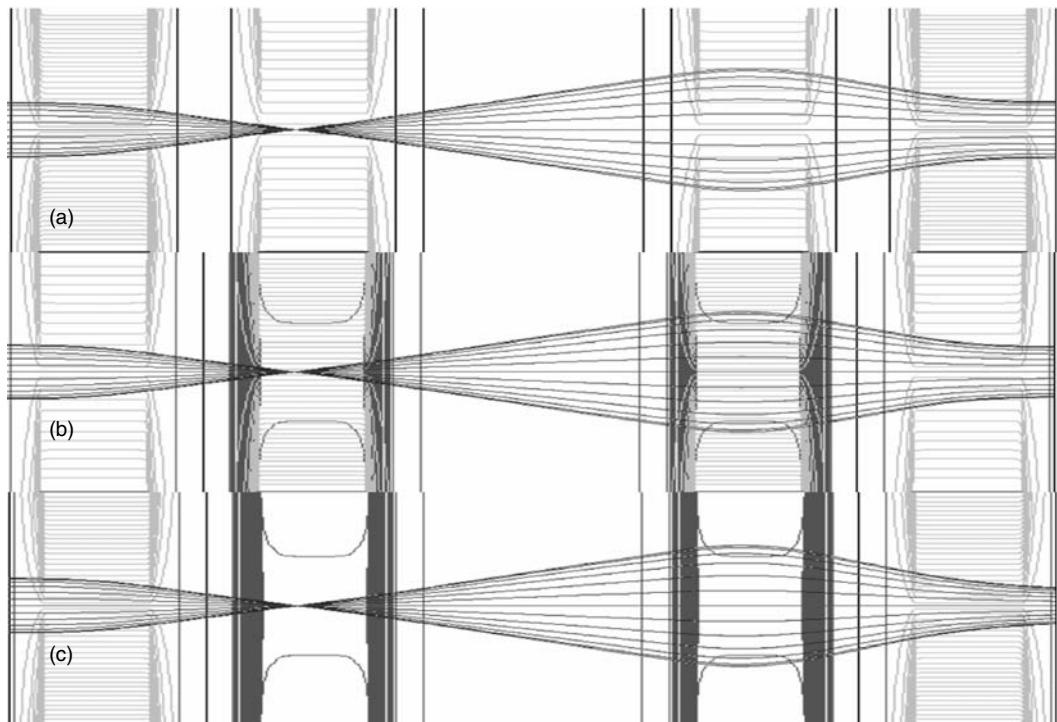


**FIGURE 4.20** Electron trajectories on the zx and zy planes for the four magnetic quadrupole lens system. Beam shapes (aberration figures) at various positions are also shown.

and TL<sub>2</sub>) are located inside the objective lens. The first doublet is used to transfer the first crossover of the beam to the center of the first hexapole. The reason why the hexapole corrector should be located at the crossover point is that the off-axial beam should pass through the optical axis at the center of the corrector to create no off-axis aberrations. After passing through the first hexapole, there are again two round lenses (the second transfer doublet TL<sub>3</sub> and TL<sub>4</sub>). The second transfer doublet reverses the beam and the beam inputs parallel to the second hexapole. It makes the second focus of the beam to cancel the second-order geometrical aberrations.

#### 4.3.1.3 Combined Electrostatic and Magnetic Quadrupole Lenses as a Chromatic Aberration Corrector

A combined electrostatic and magnetic quadrupole lens system is described as an example of a chromatic aberration corrector. Four quadrupole lenses are used to form a basic first order focusing ray trajectory. Electron ray trajectories and the beam shape in each step are shown in Figure 4.20. At the first quadrupole lens, the y-directional beam diverges and the x-directional beam converges.

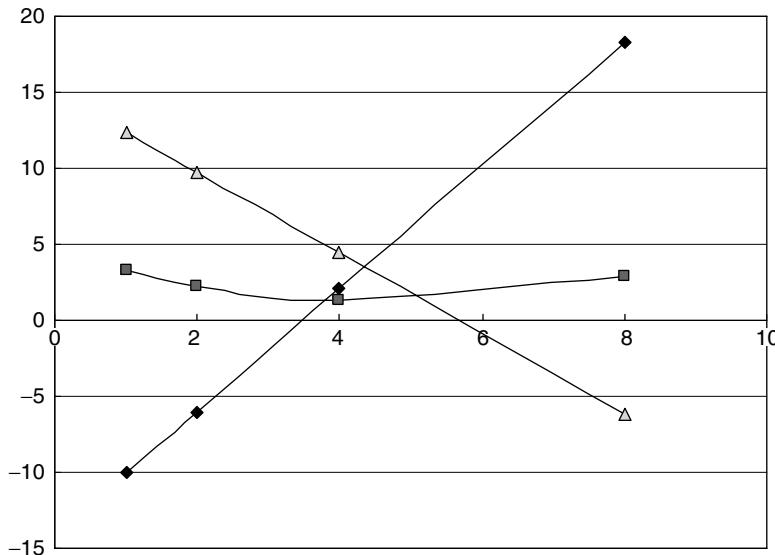


**FIGURE 4.21** (a)–(c) Electron trajectories of combined electrostatic and magnetic quadrupoles. The first and the fourth quadrupoles are magnetic and are at constant excitation. The second and the third quadrupoles are the combined quadrupoles but are excited only by the magnetic field (a) and only by the electrostatic field (c). In the case of (b), half magnetic and half electrostatic excitations are adjusted to keep the same first-order trajectory.

The x-directional beam has a focus at the second quadrupole. The beam has an elongated shape in the y-direction at this plane. The second quadrupole lens has a focusing action in the x-direction and forms a focus at the third quadrupole. The beam has an elongated shape in the x-direction in the third quadrupole. After passing through the fourth quadrupole, the beam forms a stigmatic focus. Although Figure 4.20 is shown with magnetic quadrupoles, a similar principal ray can be generated using electrostatic quadrupoles.

It is not possible to produce a negative chromatic aberration in such a focusing system, because the second and the third quadrupoles act as convex lenses, and the aberrations generated by the first quadrupole are canceled out at the fourth quadrupole. To make a negative chromatic aberration corrector, it is necessary to use a concave lens. As described in Chapter 6, an electrostatic quadrupole produces twice the chromatic aberration of a magnetic quadrupole. If the electrostatic quadrupole lens is used as a concave lens and the magnetic quadrupole lens acts as a convex lens at the same position, it is possible to keep the same first-order optical properties as shown in Figure 4.21, after adjusting the amount of excitation of the quadrupoles. If the concave lens action is increased by increasing the voltage of the electrostatic quadrupole, the negative chromatic aberration is increased. The additional lens action of the electrostatic lens can be compensated by increasing the excitation of the magnetic quadrupole. Thus, the chromatic aberration can be changed while keeping the electron trajectory the same as in Figure 4.20.

Figures 4.21a through 4.21c show zx-electron trajectories of four quadrupole system with changing excitation of the second and third quadrupoles. In the case of Figure 4.21a, only the magnetic quadrupoles are excited similarly as shown in Figure 4.20 and in Figure 4.21c, only the electrostatic quadrupoles are excited at the second and the third quadrupoles. In the case of Figure 4.21b, half of



**FIGURE 4.22** Magnetic excitation NI2 dependence of  $C_s$ ,  $C_c$ , and the voltage of the electrostatic V2 quadrupoles. Triangles indicate  $C_c$ , light rectangles indicate  $C_s$  and dark rectangles indicate NI2.

the magnetic quadrupole excitation and half of the electrostatic excitation are combined. As can be seen from Figures 4.21a through 4.21c, all three trajectories are the same.

Figure 4.22 shows the magnitudes of  $C_c$  and  $C_s$  and the excitation of the electrostatic quadrupole component ( $V_2$ ) plotted against the excitation of the magnetic quadrupole ( $NI_2$ ) for the second and the third quadrupoles. As the excitation of the magnetic quadrupole, is increased the excitation of the electrostatic quadrupoles linearly increases and  $C_c$  decreases linearly until it becomes negative. On the other hand,  $C_s$  always has positive values and does not change much with the increase of the quadrupole excitation.

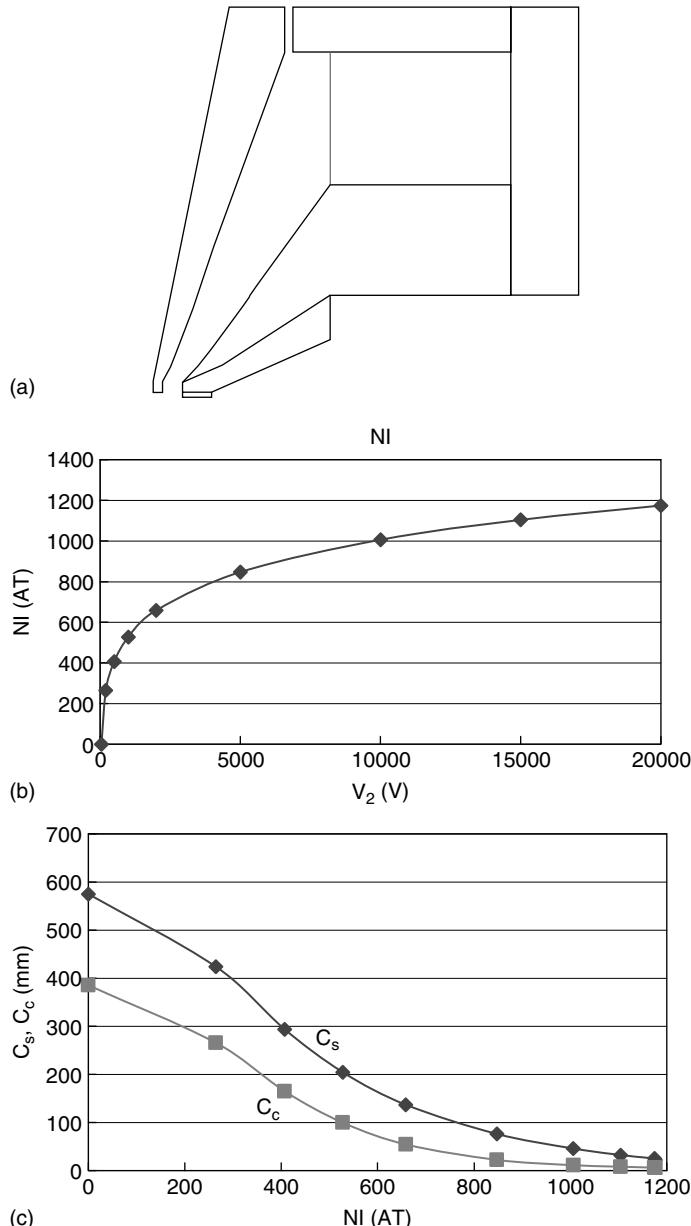
#### 4.3.2 OBJECTIVE LENS DESIGN OF LOW-ENERGY ELECTRON MICROSCOPE/PHOTOELECTRON EMISSION MICROSCOPE

Electron microscopes can be classified into four categories; two are the direct imaging systems: TEMs and reflection or emitted electron microscopes (LEEM/PEEM). The other two categories of microscopes use scanning beams: STEMs and SEMs. Among these four electron microscopes, LEEM/PEEM has not been so widely used. From the point of view of magnetic electron lens design, the main interest for TEM/SEM/STEM microscope lenses is aberration correction, at present. However, magnetic electron lens design itself is still under development in LEEM/PEEM.

The objective lens of LEEM/PEEM is similar to the one for low-voltage SEM (Frank and Müllerová, 1999). This is because LEEM/PEEM uses large bulk specimens similar to SEM. Electrons are emitted from the specimen with very low acceleration voltages (from a few volts to a hundred volts). These low-voltage electrons are accelerated up to a few or a few tens of kilovolts inside the objective lens. The specimen of a LEEM/PEEM floats at a high voltage, similar to the cathode of the electron gun; electrons emitted from the specimen at low voltages are accelerated to the final accelerating voltage inside the objective lens. Another important point in the LEEM/PEEM design is the objective aperture. The focal plane of the objective lens is inside the acceleration field. It is not possible to insert a moving objective aperture in such an acceleration field, so the aperture is usually inserted after the intermediate lens crossover plane. A drawback of this aperture position is the demagnification of the diffraction pattern. A modification would be to separate the objective

lens into two parts and locate the aperture plane after the first part of the objective lens, where the aperture can be put at earth potential. Another possibility is to insert a transfer doublet lens inside the objective lens, to transfer the aperture position outside the objective lens as was shown in the example of  $C_s$  corrector (Figures 4.18 and 4.19).

Figure 4.23a shows an objective lens of LEEM/PEEM (Bauer, 2001). There are at least two electrodes and therefore three potentials in the system (specimen voltage  $V_1$ , the first electrode voltage  $V_2$ , and the second electrode voltage  $V_3$ ). In the present case, the beam is accelerated up to  $V_3 = 20$  kV. Figure 4.23b shows the relation between  $V_2$  and the magnetic excitation of the lens NI.



**FIGURE 4.23** (a) An objective lens same as Figure 4.18, (b) first electrode voltage ( $V_2$ ) dependence of the magnetic excitation (NI) of the objective lens, and (c) magnetic excitation (NI) dependence of  $C_s$  and  $C_c$  of the low-energy electron microscope lens.

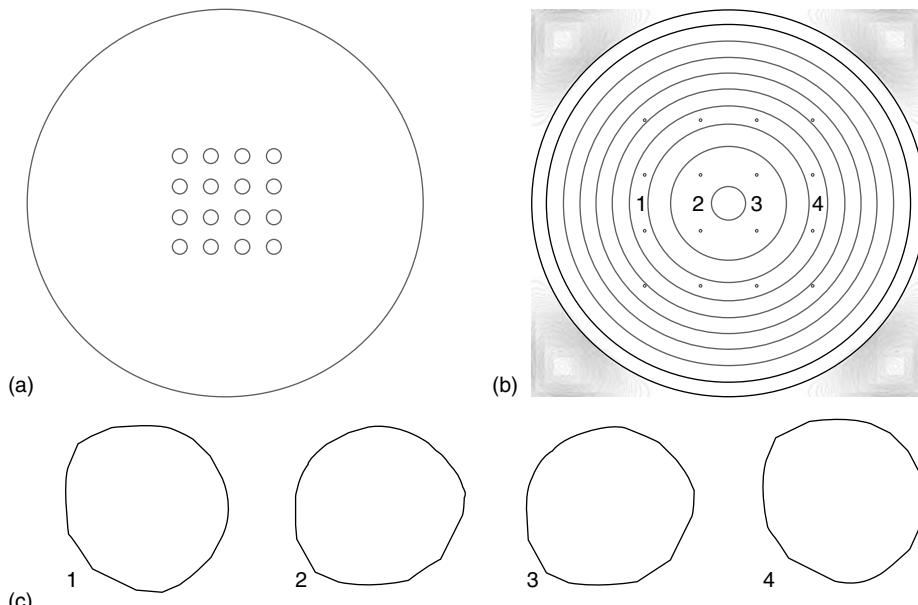
When  $NI = 0$ , it means the objective lens is all electrostatic and  $V_2 = 41.9$  V. When  $V_2$  is increased  $NI$  also increases. When  $V_2$  equals  $V_3$ , the excitation of the magnetic lens becomes  $NI = 1174$  AT. An interesting thing is that, this value is nearly the same as the excitation of a TEM/SEM lens at 20 kV. When the same objective lens shape is used in a TEM/SEM at 20 kV, the excitation is  $NI = 1176$  AT.

The gap of the magnetic lens is located between the first and the second electrodes. The magnetic field is applied to electrons that are accelerated from  $V_2$  to 20 kV. At low accelerating voltage  $V_2$ , the necessary magnetic excitation is small and as  $V_2$  increases the necessary magnetic excitation  $NI$  becomes large.

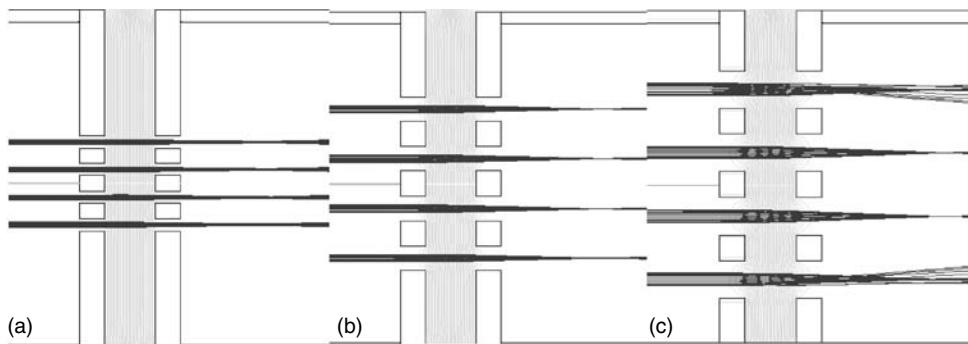
Figure 4.23c shows the spherical and the chromatic aberration coefficients  $C_s$  and  $C_c$  plotted against the magnetic excitation  $NI$ . Both  $C_s$  and  $C_c$  decrease with increasing  $NI$ . When the electrostatic lens is used ( $NI = 0$ ),  $C_s = 575$  mm and  $C_c = 385$  mm are obtained. These values decrease to  $C_s = 25$  mm and  $C_c = 5.9$  mm at  $V_2 = 20$  kV. These values are very similar to the values of the TEM/SEM lens where  $C_s = 22$  mm and  $C_c = 5.9$  mm. From the calculation shown in Figure 4.23, one concludes that the acceleration of the beam from the specimen to the objective lens does not destroy the resolution. Lens properties improve when the electrostatic lens is supplemented by the magnetic lens. The results suggest that the resolution of LEEM/PEEM can be improved to the same level of resolution as TEM/SEM. In this study, we did not use a magnetic immersion lens, which is usually the case in a TEM objective lens. If we use the magnetic immersion lens in LEEM/PEEM,  $C_s$  and  $C_c$  must be further reduced.

#### 4.3.3 LOTUS ROOT LENS AS A MULTIBEAM ELECTRON LITHOGRAPHY SYSTEM

Coulomb interaction limits the beam current in single column direct-writing electron beam lithography systems. To increase the throughput of an electron beam exposure system, several multibeam and multicolumn systems have been proposed. As an example of a multibeam system, a lens developed by Haraguchi et al. (2002) is analyzed. They developed an objective lens with numerous holes in the pole-piece bore with a single coil system as shown in Figure 4.24a. They call it as *lotus root*.



**FIGURE 4.24** (a) Pole-piece bore of the lotus root lens, (b) 16 beam patterns written by the lotus root lens, and (c) beam profiles at the focus of 1, 2, 3, and 4 lenses shown in (a) and (b).



**FIGURE 4.25** Beam trajectories of three lotus root lenses with different distances between each hole and with the different diameters of hole. The outer diameters of the main pole-piece bore are the same.

*lens* (LRL). The LRL has more than 10 holes. The name is given by the similarity of the pole shape to a vegetable lotus root.

Figures 4.25a through 4.25c compare three examples of electron rays of LRL by changing the distance and diameter of holes. As can be seen from the comparison, we can find a good condition of small aberrations, where off-axial beams form a fine focus. In the case of Figure 4.25a, large aberrations prevent the formation of a small beam size. On the other hand, in the case of Figure 4.25c, the central beam has a small focused beam but the outside beams have large aberrations. The middle case Figure 4.25b shows good focal properties with small beams from all of the holes.

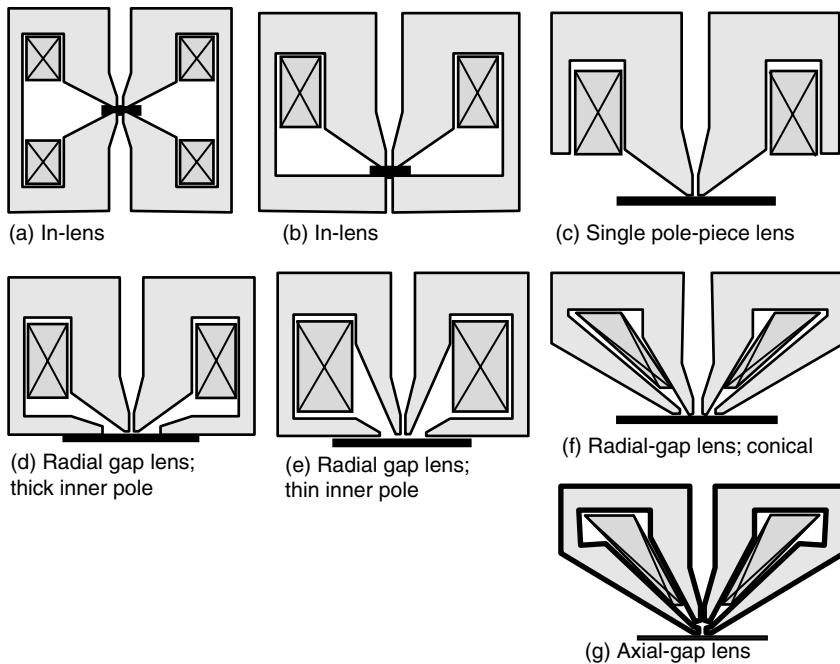
Figures 4.24b and 4.24c show spot diagrams of 16 beams for the case Figure 4.25b. Nearly the same round spots are arranged regularly as shown in Figure 4.24b. There is no distortion and no blur in the image. Circular lines in Figure 4.24b indicate the same potential lines of magnetic field. Nearly circular lines 1, 2, 3, and 4 in Figure 4.24c are the beam profiles corresponding 1, 2, 3, and 4 of Figure 4.24b. It is possible to make a multibeam lithography system using LRL. The beam shapes are nearly circular but distorted a little. It is not clear whether such distortion can be tolerated in the case of a lithography system.

#### 4.3.4 VARIOUS OBJECTIVE LENSES FOR LOW-VOLTAGE SCANNING ELECTRON MICROSCOPE

Although the objective lens is the most important part of a TEM for determining the resolution, the electron gun is more important in determining the resolution of SEM. Tungsten hairpin filaments were used in the first SEMs, and the resolution was about 20 nm in the 1960s. Resolution improved in the 1970s as a result of the use of the field emission gun (FEG) (Komoda and Saito, 1972; see also Chapters 2 and 9). The resolution has improved gradually to 2.5 nm in 1973, 2 nm in 1981, 0.7 nm in 1985, and at last 0.4 nm in recent years.

From Figures 4.26a through 4.26g, various SEM objective lenses are classified systematically (Tsuno, 1999). A typical objective lens, which is called a conical lens, is shown in Figure 4.26f. The specimen is located beneath the objective lens and is outside the objective lens magnetic field (out of lens) similar to the LEEM/PEEM lens shown in Figure 4.23a. The secondary electron detector is located at the side of the specimen.

During the 1970s, a scanning coil was added to a TEM and the instrument was used as both a STEM and a SEM (Koike et al., 1971). The resolution of the SEM mode was typically 2 nm. Two points are significant in comparing the SEM mode of a TEM with the ordinary SEM. The first is the immersion objective lens (Figures 4.26a and 4.26b). Microscopes with such an objective lens are called as *in-lens* SEM. The second point is the position of the detector, which is set above the objective lens. Secondary electrons spiral up through the objective lens along the flux lines of the magnetic field (Tsuno, 1999) and are caught by the electric field applied to the detector. The resolution of the in-lens field emission SEM reached 0.7 nm in 1985 (Nagatani et al., 1987). The lens in Figure 4.26b



**FIGURE 4.26** Various types of scanning electron microscope lenses.

was built by Nakagawa et al. (1991). In their case, the distance between the lower pole piece and the specimen is fairly large, because it is necessary to insert the specimen stage in this gap.

During the 1970s, it was found that low voltages are very useful for nonconductive specimens. However, the resolution of the SEM at low voltages ( $\sim 1$  kV) was quite poor. Since the beginning of the 1980s, the necessity of high-resolution, low-voltage SEMs with a large specimen stage for observing semiconductor wafers has increased rapidly, necessitating the use of FEGs. Design of the objective lens is especially important in the low-voltage SEM to take advantage of the FEG. A single pole-piece lens proposed by Mulvey in 1973 (Mulvey, 1982) has advantages in low-voltage SEM, although he proposed it for TEM. Characteristics of the single pole-piece lens are (1) large specimen area due to the single pole piece located near the specimen and (2) low aberrations due to the immersion of the specimen in-lens field. A drawback of the single pole-piece lens is the large excitation required. However, for the low-voltage SEM, this is not a problem. This type of lens was applied to a low-voltage SEM by Pawley (1984) and Yonezawa et al. (1994).

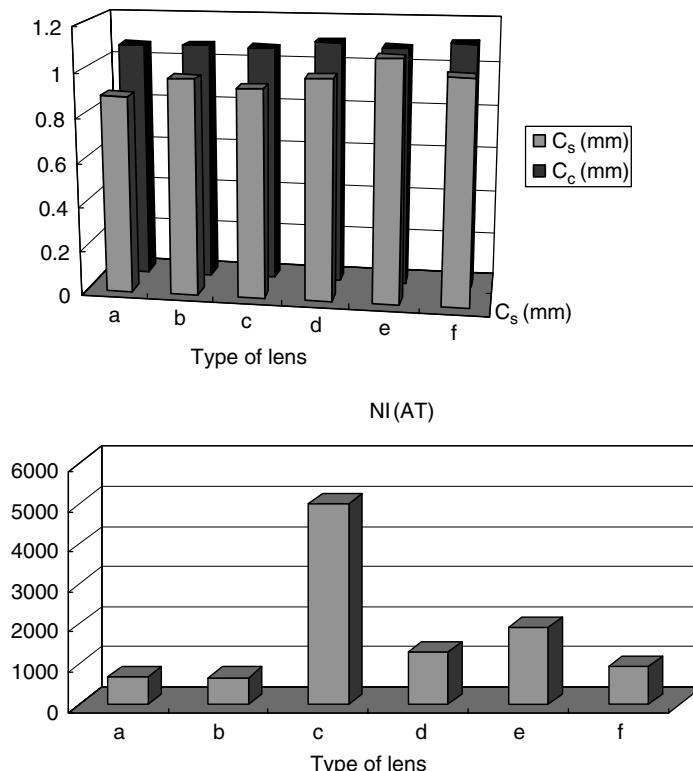
Hill and Smith (1982a) proposed a type of single pole-piece lens (Figure 4.26c) different from the original single pole-piece lens proposed by Mulvey. The lens is located in front of the specimen and the incident beam illuminates the specimen through the objective lens. Mulvey's lens has advantages in terms of optical properties. However, from the point of view of application to a large specimen observation, lens in Figure 4.26c has an advantage. Mulvey's single pole-piece lens prevents the large specimen placed near the field maximum. Shao and Lin (1989) applied the lens in Figure 4.26c to a low-voltage SEM. The lens properties were  $C_s = 1.3$  mm and  $C_c = 1.0$  mm, and resolution of 1.7 nm was predicted at 1 kV at 2 mm working distance (WD, the distance between the pole piece and the specimen) at 2070 AT.

Tang and Song (1990) proposed the lenses in Figures 4.26d and 4.26e. These lenses were obtained by adding an outer yoke to the lens in Figure 4.26c. However, in this lens, the outer pole piece is flat. Tang called such a lens the *side pole-gap lens*, because the gap faces not to the optical axis but to the specimen plane. This was a new idea of an *immersion lens* under relatively low excitation compared with the single pole-piece lenses. Tang expected very low spherical and chromatic aberration

coefficients based on a model calculation. The significant point of the lens is that the axial magnetic field is highly asymmetrical and has a small half width. The specimen is located in the tail of the field distribution. The difference between the lens in Figure 4.26c and the lenses in Figures 4.26d and 4.26e is simply the existence of the outer yoke. The field is applied to the specimen because the gap faces not to the optical axis but to the specimen. Sato et al. (1993) followed Tang and applied the immersion lens in Figure 4.26f to his SEM design. He obtained  $C_s = 3.2$  mm and  $C_c = 2.7$  mm with WD = 3 mm. The excitation of lens in Figure 4.26f is decreased greatly compared with the lens in Figure 4.26c. Sato's lens can be used in an accelerating voltage of up to 12 kV.

If we compare the lens in Figure 4.26f with the conventional one in Figure 4.26g, the shape of the yoke is nearly the same. Only the pole piece is different. In case of the lens in Figure 4.26f, both pole pieces have the same axial length, and the magnetic field is distributed in the radial direction and applied to the specimen. On the other hand, in the case of the lens in Figure 4.26g, the inner pole piece is shorter than the outer one, but the outer pole-piece bore is nearly the same as the inner one. The field is applied to the optical axis. The final objective lens shape for the low-voltage SEM ends up similar to the shape of the conventional lens. We shall call the lens in Figure 4.26f a *radial-gap lens* and the ordinary lens in Figure 4.26g an *axial-gap lens*. Recently, a lens combining radial and axial lenses with two exciting coils was reported. The lens in Figure 4.26f is used at high magnification, and the lens in Figure 4.26g is used at low magnification and for magnetic specimens.

Comparison of lens properties  $C_s$  and  $C_c$  is made for all the radial-gap lenses as shown in Figures 4.26a through 4.26f, with the same gap length ( $2s = 4$  mm), bore diameter ( $2b = 4$  mm), and accelerating voltage ( $U_o = 1$  kV). The optical properties and the excitation of these lenses are shown in Figures 4.27a and 4.27b, respectively. Although the differences in its shape are very large,  $C_s$  and  $C_c$  do not change significantly among the radial-gap lenses. All of the radial-gap lenses show



**FIGURE 4.27** Spherical ( $C_s$ ) and chromatic ( $C_c$ ) aberration coefficients for various radial-gap lenses and their excitations.

nearly  $C_s = 1$  mm and  $C_c = 1$  mm. The difference is only around 0.2 mm. However, as shown in Figure 4.27b, the excitation is changed significantly. The excitation of the lens in Figure 4.26c is about 10 times larger than that of the lens in Figure 4.26a. The lens in Figure 4.26c was introduced for a large specimen stage. Other lenses (Figures 4.26d through 4.26f) were introduced to reduce the excitation while keeping a large specimen area without sacrificing optical properties. The final result was the lens in Figure 4.26f, where a large specimen with a specimen tilt can be imaged under relatively low excitation. The outer shape of the yoke is the same as the conventional lens of axial-gap lens in Figure 4.26g.

#### 4.3.5 COMBINED ELECTROSTATIC AND MAGNETIC LENSES FOR LOW-VOLTAGE SCANNING ELECTRON MICROSCOPE

In Section 4.3.3, the combined acceleration field and magnetic objective lens for LEEM/PEEM was discussed. There is another group of lenses specialized for low-voltage SEM in which a retarding electric field is combined with a magnetic objective lens. In the case of LEEM/PEEM, the beam is accelerated in the objective lens because of the direct imaging: electrons are emitted from the specimen and form an image. On the other hand, in the case of SEM, an electron beam forms a probe on the specimen after deceleration. The direction of the beam is different when taking the aberrations into account. There are a variety of combination of fields as reviewed by Müllerová and Lenc (1992) and Tsuno et al. (1996). The most common method is the immersion electrostatic lens (C in Table 4.4) in which the specimen is floating on a high voltage. Type C is advantageous, because a deceleration field located between the specimen and the first electrode is very effective in reducing the aberration coefficients. However, from the operational point of view, there are some difficulties in immersion electrostatic lenses. They require a flat specimen and it is not possible to tilt the specimen.

In the case of B in Table 4.2, the specimen is located at earth potential and in field free space. Frosien et al. (1989) have succeeded in obtaining good optical properties in an actual SEM using such a lens. The electric field follows the magnetic field sequentially in their lens. The specimen and the first electrode are set at earth potential to avoid leakage of the electric field to the specimen. They obtained  $C_s = 3.7$  mm and  $C_c = 1.8$  mm at WD = 4 mm, at 500 V beam potential. Resolution of 3 nm has been attained at 1 kV beam potential (Frosien et al., 1994).  $C_s$  and  $C_c$  are reduced to one-tenth of those of the pure magnetic lens by adding the electric retarding field.

---

**TABLE 4.4**  
**Classification of Scanning Electron Microscope Objective Lenses**

Combined magnetic and electrostatic	Immersion magnetic (A)	Sequential	Yau et al. (1981); $C_s = 0.1$ and $C_c = 0.08$
	Mixed		Beck et al. (1995); $C_s = 0.77$ , and $C_c = 0.62$
	Both out of lens (B)	Sequential	1. Frosien et al. (1989); 1 kV, $C_s = 3.7$ , $C_c = 1.8$ , d = 11, NI = 402, and WD = 4 2. Frosien et al. (1994); d = 3 nm, 50° tilt, and WD = 6 mm
	Mixed		Pries and Schweizer (1987); 0.7–2 kV, $C_s = 8$ , $C_c = 3$ , d = 100, and NI = 467
Immersion electrostatic (C)	Sequential		1. Paden and Nixon (1968); 30 kV–5 eV 2. Yau et al. (1981); 10 kV–100 V, $C_s = 0.4$ , $C_c = 0.1$ , and NI = 665 AT
	Mixed		Yau et al. (1981); 10–9.9 kV, $C_s = 0.1$ , and $C_c = 0.08$

Source: Partly after Müllerová, I. and Lenc, M., *Ultramicroscopy*, 41(4), 399–410, 1992.

---

There is to date no lens in which a magnetic immersion lens is combined with a retarding electric field. In the case of a pure magnetic lens, the magnetic immersion lens is very effective for reducing  $C_s$  and  $C_c$ . However, according to Frosien, it is unnecessary to use a magnetic immersion lens in the combined magnetic and electrostatic lenses. The values of  $C_s$  and  $C_c$  are already low enough in the combined field lenses even when neither is an immersion lens.

## FOR FURTHER INFORMATION

The book *Magnetic Electron Lenses*, edited by Hawkes (1982), is useful for the understanding of the technology described here. In this book, aberration theory and CAD methods, which are necessary to design magnetic lenses, are covered (as in this handbook). There are three chapters written by Lenz, Riecke, and Mulvey about the practical design examples of magnetic lenses and their optical properties that are especially valuable.

The book *Laboratory Magnets* (Kroon, 1968) is a useful source for details about the design methods of magnetic circuits. Section 4.2 in this chapter follows Kroon's book. However, this book was written before CAD methods became familiar, so it is useful only for the purpose of learning about electromagnets.

As for materials from which magnetic lenses are fabricated, there are as yet no books. Because magnetic lens material represents only a small fraction of the market for magnetic materials, no one has yet bothered to write one. However, lens material is actually very special, as noted in this chapter. Specification sheets from various manufacturers of magnetic materials are useful to learn the magnetization curves of materials. However, you must be very careful to note that most of the magnetization curves are obtained from cold-rolled sheets, which have superior properties to the magnetic properties of bulk materials made by hot forging.

If you want to study examples of lenses that were designed or produced in the past, the *Proceedings of the International and European Congresses on Electron Microscopy* are useful. There are a lot of examples of magnetic lenses in those proceedings, most of which are not published in the usual journals and books. Details of those proceedings are found in Hawkes's book *Magnetic Electron Lenses* and in *Principles of Electron Optics* by Hawkes and Kasper (1989). In this book, you can find many examples of the special magnetic lenses that are not described here.

There are as yet no books about aberration correction, so the aberration and corrector sections of the revised edition of this book are useful for understanding the principles of aberration correction.

## REFERENCES

- Bauer, E. 2001. Photoelectron spectroscopy: present and future, *J. Electron Spectroscopy Related Phenomena*, 114–116, 975–987.
- Beck, S., Plies, E., and Schiebel, B. 1995. Low-voltage probe forming columns for electrons, *Nucl. Instrum. Methods Phys. Res.*, A363(1,2), 31–42.
- Biddlecombe, C. S. and Trowbridge, C. W. 1984. CAE of electromagnetic devices, *Computer-Aided Eng. J.*, 4, 84–90.
- Bonjour, P. 1975. High field superconducting electron lenses with holmium polepieces, *J. Phys. E: Sci. Instrum.*, 8(10), 764–768.
- Chen, C. W. 1961. Metallurgy and magnetic properties of an Fe–Co–V alloy, *J. Appl. Phys.*, 32(3), 348S–355S.
- Cosslett, V. E. 1991. Fifty years of instrumental development of the electron microscope, in *Advances in Optical and Electron Microscopy*, Barer, R. and Cosslett, V. E. Eds., Academic Press, London, 215–267.
- English, A. T., Chin, G. Y., and Von Neida, A. R. 1967. Effect of ordering on rolling-induced magnetic anisotropy in FeCo–2V, *J. Appl. Phys.*, 38(3), 997–998.
- Frank, L. and Müllerová, I. 1999. Strategies for low- and very-low-energy SEM, *J. Electron Microscopy*, 48, 205–219.

- Frosien, J., Lanio, S., and Feuerbaum, H. P. 1994. High precision electron optical system for absolute and CD-measurements on large substrates, *Nucl. Instrum. Methods*, A393(1,2), 25–30.
- Frosien, J., Plies, E., and Anger, K. 1989. Compound magnetic and electrostatic lenses for low-voltage supplies, *J. Vac. Sci. Technol.*, B7(6), 1874–1877.
- Haider, M., Braunshausen, G. and Schwan, E. 1995. Correction of the spherical aberration of a 200 kV TEM by means of a hexapole-corrector, *Optik*, 99, 167–179.
- Hall, R. C. 1960. Magnetic anisotropy and magnetostriction of ordered and disordered cobalt–iron alloys, *J. Appl. Phys.*, 31(5), 157S–158S.
- Haraguchi, T., Sakazaki, T., Hamaguchi, S. and Yasuda H. 2002. Development of electromagnetic lenses for multielectron beam lithography system, *J. Vac. Sci. Technol.*, B20, 2726–2729.
- Hawkes, P. W. and Kasper, E. 1989. *Principles of Electron Optics*, Vol. 2, Academic Press, London.
- Hill, R. and Smith, K. C. A. 1982. Application of the Cambridge interactive electron lens analysis system, *CIELAS. Inst. Phys. Conf. Ser.*, No. 61, Institute of Physics, chap. 2, 71–74.
- Hill, R. and Smith, K. C. A. 1982. The single-pole lens as a scanning electron microscope objective, in *Scanning Electron Microscopy II*, SEM Inc., AMF O'Hare, Chicago, 465–471.
- Kanaya, K. 1985. Reminiscences of the development of electron optics and electron microscope instrumentation in Japan, in *The Beginnings of Electron Microscopy*, Hawkes, P. W. Ed., Academic Press, Orlando, FL, 317–386.
- Koike, H., Ueno, K., Suzuki, M., Matsuo, T., Aita, S., and Shibatomi, K. 1971. High resolution scanning device for the JEM-100B electron microscope, *JEOL News*, 9e(3), 20–21.
- Komoda, T. and Saito, S. 1972. Experimental resolution limit in the secondary electron mode for a field emission source scanning electron microscope, *Scanning Electron Microscopy*, Part I, *Proceedings of the Fifth Annual Scanning Electron Microscope Symposium*, IIT Research Institute, Chicago, IL, 129–136.
- Krivanek, O. L., Dellby, N., and Lupin, A. R. 1999. Towards sub-Å electron beams, *Ultramicroscopy*, 78, 1–11.
- Kroon, D. J. 1968. *Laboratory Magnets*, Philips Technical Library, Eindhoven, The Netherlands.
- Lefranc, G., Knapek, E., and Dietrich, I. 1982. Superconducting lens design, *Ultramicroscopy*, 10(1/2), 111–123.
- Lencová, B. and Lenc, M. 1986. A finite element method for the computation of magnetic electron lenses, *Scanning Electron Microscopy III*, SEM Inc., AMF O'Hare, Chicago, IL, 897–915.
- Müllerová, I. and Lenc, M. 1992. Some approaches to low-voltage scanning electron microscopy, *Ultramicroscopy*, 41(4), 399–410.
- Mulvey, T. 1982. Unconventional lens design, in *Magnetic Electron Lenses*, Hawkes, P. W. Ed., Springer-Verlag, Berlin, 359–420.
- Mulvey, T. 1984. *Magnetic Electron Lens II, Electron Optical Systems*, SEM Inc., Chicago, IL, 15–27.
- Mulvey, T. and Tahir, K. 1990. Recent development in electron-optical instrumentation and computer-aided design, *Ultramicroscopy*, 33(1), 1–13.
- Munro, E. 1973. Computer-aided design of electron lenses by the finite element method, in *Image Processing and Computer-Aided Design in Electron Optics*, Hawkes, P. W. Ed., Academic Press, London, 284–323.
- Nagatani, T., Saito, S., Sato, M., and Yamada, M. 1987. Development of an ultra high resolution scanning electron microscope by means of a field emission source and in-lens system, *Scanning Microsc.*, 1(3), 901–909.
- Nakagawa, H., Nomura, N., Koizumi, T., Anazawa, N., and Harafuji, K. 1991. A novel high-resolution scanning electron microscope for the surface analysis of high-aspect-ratio three dimensional structures, *Jpn. J. Appl. Phys.*, 30(9A), 2112–2127.
- Paden, R. S. and Nixon, W. C. 1968. Retarding field scanning electron microscopy, *J. Sci. Instrum. (J. Phys. E)*, 1(11), 1073–1080.
- Pauling, L. 1938. The nature of the interatomic forces in metals, *Phys. Rev.*, 54(12), 899–904.
- Pawley, J. 1984. Low voltage scanning electron microscopy, *J. Microsc.*, 136(10), 45–68.
- Philipp, F., Höschen, R., Osaki, M., Möbus, G., and Rühle, M. 1994. New high-voltage atomic resolution microscope approach 1 Å point resolution installed in Stuttgart, *Ultramicroscopy*, 56(1–3), 1–10.
- Plies, E. 1994. Electron optics of low-voltage electron beam testing and inspection. Part I: simulation tools, in *Advances in Optical and Electron Microscopy*, Vol. 13, Mulvey, T. Ed., Academic Press, London.
- Plies, E. and Schweizer, M. 1987. Calculation of fields and secondary electron trajectories in a new in-lens spectrometer for electron beam testing, *Siemens Forsch.-u. Entwickl.-Bet. Bd.*, 16(1), 30–36, 1987.

- Preikszas, D. and Rose, H. 1997. Correction properties of electron mirrors, *J. Electron Microscopy*, 46, 1–9.
- Pshechenkova, G. V. and Skokov, A. D. 1966. Magnetic properties and hardness of iron–cobalt alloys, *Fiz. Met. Metalloved.*, 21(3), 339–345.
- Ruska, E. 1980. *The Early Development of Electron Lenses and Electron Microscopy*, Translated by Mulvey, T., S. Hirzel Verlag, Stuttgart.
- Sato, M., Todokoro, H., and Kageyama, K. 1993. A snorkel type conical objective lens with E cross B field for detecting secondary electrons, *SPIE Charged-Particle Optics*, 2014, 17–23.
- Shao, Z. and Lin, P. S. D. 1989. High-resolution low-voltage electron optical system for very large specimens, *Rev. Sci. Instrum.*, 60(11), 3434–3441.
- Slater, J. C. 1937. Electronic structure of alloys, *J. Appl. Phys.*, 8(6), 385–390.
- Tang, T.-T. and Song, J.-P. 1990. Side pole-gap magnetic electron lenses, *Optik*, 84(3), 108–112.
- Tsuno, K. 1999. Optical design of electron microscope lenses and energy filters, *J. Electron Microscopy*, 48, 801–820.
- Tsuno, K., Handa, N., and Matsumoto, S., 1996. Immersion lenses for low voltage SEM and LEEM, *SPIE Charged-Particle Optics*, 2522, 243–252.
- Tsuno, K. and Honda, T. 1983. Magnetic field distribution and optical properties of asymmetrical objective lenses for an atomic resolution high-voltage electron microscope, *Optik*, 64(4), 367–378.
- Tsuno, K., Ioanoviciu, D., and Martínez, G., 2005. Third-order aberration theory of Wien filters for monochromators and aberration correctors, *J. Microscopy*, 217, 205–215.
- Weißbäcker, C. and Rose, H., 2001, 2002, Electrostatic correction of the chromatic and of the spherical aberration of charged-particle lenses (part I, part II), *J. Electron Microscopy*, 50, 383–390, 51, 45–51.
- Yau, Y. W., Pease, R. F., Iranmanesh, A. A., and Polasko, K. J. 1981. Generation and application of finely focused beams of low-energy electrons, *J. Vac. Sci. Technol.*, 19(4), 1048–1052.
- Yonezawa, A., Ooi, M., Morita, S., Takahashi, S., Adachi, T., and Kato, J. 1994. Development of semiconductor process evaluation FE-SEM using the single pole-piece objective lens with conical polepiece, in *132nd Committee Electron and Ion Beam Science and Technology*, Vol. 128, Japan Society for the Promotion of Science, 153–158.
- Zach, J. and Haider, M., 1995. Aberration correction in a low voltage SEM by a multipole corrector, *Nucl. Instrum. Methods Phys. Res.*, A363, 316–325.



---

# 5 Electrostatic Lenses

*Bohumila Lencová*

## CONTENTS

5.1	Introduction .....	162
5.2	Optical Properties of Electrostatic Lenses.....	166
5.2.1	Evaluation of Lens Properties .....	166
5.2.2	The Equation of Motion and the Trajectory Equation .....	168
5.2.3	Electrostatic Lens as a Thick Lens .....	170
5.2.4	Types of Electrostatic Lenses.....	171
5.2.5	Matrix Description of the Trajectory and Models of Some Electrostatic Lenses....	172
5.2.5.1	Matrix Description.....	172
5.2.5.2	Lens Action of an Aperture Lens .....	172
5.2.5.3	Lens Action of the Homogeneous Field .....	173
5.2.5.4	Model of Lens Action (Cathode Lens, Immersion Lens, and Unipotential Lens).....	174
5.2.6	Aberrations of Electrostatic Lenses .....	176
5.2.7	Typical Properties of Round Electrostatic Lenses .....	177
5.2.7.1	Two-Electrode Immersion Lenses.....	177
5.2.7.2	Three-Electrode Unipotential Lenses .....	178
5.2.7.3	Multielement and Zoom Lenses (Movable Lens). ....	179
5.2.7.4	Immersion Objective Lenses .....	179
5.2.7.5	Electron Mirror .....	181
5.2.7.6	Grid and Foil Lenses .....	181
5.2.8	Cylindrical and Astigmatic Lenses.....	183
5.2.9	Optimization .....	183
5.3	Practical Design of Electrostatic Lenses .....	185
5.3.1	Design Problems.....	185
5.3.2	Electrical Insulation and Breakdown in the Vacuum Gap.....	187
5.3.3	Insulator Materials and Design .....	188
5.3.4	Manufacturing and Alignment Accuracy .....	190
5.3.5	Environmental and System Considerations .....	191
5.4	Examples.....	192
5.4.1	Electrostatic Transmission Electron Microscopes .....	192
5.4.2	Low-Voltage Transmission Electron Microscopy .....	194
5.4.3	Low-Energy and Photoemission Electron Microscopes .....	194
5.4.4	Scanning Electron Microscopes (Compound and Retarding Lenses).....	196
5.4.5	Other Applications of Electrostatic Lenses.....	199
5.4.6	Ion Microscopy and Lithography .....	199
5.5	Conclusions .....	200
5.6	Further Information .....	200
	Acknowledgments.....	201
	References .....	201

## 5.1 INTRODUCTION

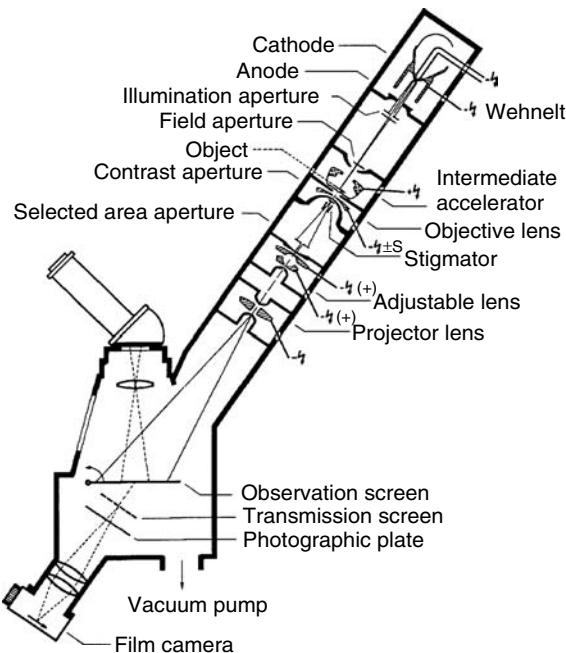
In light optics, lenses are made of transparent materials, usually with axial symmetry, and the beam of light is focused by refraction at the surfaces making up the lens. In particle optics, the paths of the charged particles are changed on passing through an electromagnetic field. Electromagnetic fields of rotational symmetry can produce a stigmatic image of objects emitting or transmitting electrons or ions placed on or close to the axis. Electrostatic lenses usually consist of charged electrodes with cylindrical bores precisely arranged on a common axis. Other types of electron lenses that do not possess rotational symmetry will be mentioned only briefly.

The foundations of electron optics, including electrostatic lenses, were laid in the article of Busch (1926), which stated that magnetic as well as electrostatic rotationally symmetric fields focus cathode rays (electrons) and presented the equation of motion in the paraxial region. The focal length of magnetic lenses was first derived in his following article (Busch, 1927). The existence of an electrostatic lens as an imaging element was not clearly stated; this possibility was first put forward by Brüche (1930, 1932). The focal length in the thin lens approximation was derived by Johannson and Scherzer (1933). Independently, Davisson and Calbick (1931, 1932) in two short communications under the title “Electron lenses” estimated the focal length of an aperture lens, in the first case, in fact, with an error making the lenses appear weaker by a factor of two.

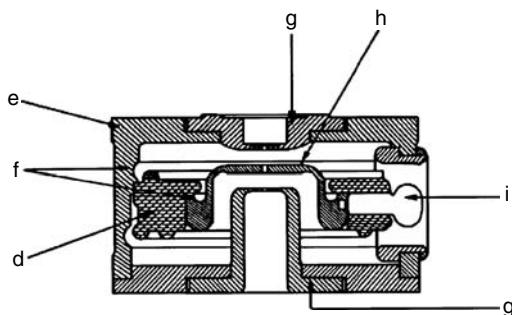
Most progress in electron optics was largely through the experimental work by Knoll, Ruska, and coworkers at the High Voltage Laboratory of the Technische Hochschule (now Technical University) and by Brüche and coworkers at the AEG Research Laboratory, Berlin from 1931 onward. Both groups showed the first images of various objects with electron lenses: the former group used magnetic lenses; the latter group employed electrostatic lenses. Both groups built two-lens electron microscopes and used them to study the structure of emitting cathode surfaces, a technique refined in particular by Brüche and coworkers to high perfection. At the same time, the theory of electron optics was formulated and perfected, especially by the work of Glaser and of Scherzer, and, as early as 1934, the whole area was presented in book form (Brüche and Scherzer, 1934). Most of the electron optical theory and the principles of lens design and use were established within a decade.

At the AEG laboratories in Berlin, efforts were concentrated on electrostatic optics. The group headed by E. Brüche included O. Scherzer, A. Recknagel (involved in mirror and cathode lens theories), and H. Boersch and H. Mahl, designers of electrostatic microscopes. Brüche (1932) and Brüche and Johannson (1932) published early micrographs from an emission microscope with electrostatic imaging. Electrostatic lenses were applied in other instruments such as Braun (cathode ray) tubes, image pick-up tubes, and image converters. An electrostatic transmission electron microscope (TEM) was developed in the late 1930s. The last electrostatic microscope designed by the AEG group in Mosbach, the EM8 launched in 1956 (then produced by Zeiss, who took microscope production over from AEG), is shown in cross section in Figure 5.1 (Mahl et al., 1956). Electrostatic TEMs designed in the 1940s in the United States at G.E. Laboratories, by Farrand Optical Co., and in France by CSF have been mostly forgotten (see Cosslett, 1951) and the latest electrostatic microscopes designed in Jena (Hahn, 1955) and by Elektros in the United States (Rempfer et al., 1972) often go unnoticed. Unfortunately, the development of the electrostatic microscope was terminated before it could profit from much of the technological progress started in the late 1960s, such as ultrahigh vacuum (UHV) technology, new electrode and insulating materials, field-emission guns, and significant progress in electronics. TEMs with magnetic lenses offering shorter focal length, smaller aberrations (thus better resolution), more flexibility in their excitation, and use of higher beam voltage completely supplanted them.

Electrostatic TEMs worked at a maximum energy of 50–70 keV because of high-voltage breakdown problems; the best resolution obtained was 2.0 nm. The resolution limit can be evaluated from a simple estimate based on the mode of operation of an electrostatic objective lens, where the thin object has to be placed outside the strong electrostatic field to decrease the electrostatic force on it and to allow sample tilt without disturbing the lens field. The electrical breakdown limit allows



**FIGURE 5.1** Cross section of EM8, the last electrostatic transmission electron microscope (TEM) of the firm Zeiss in Germany. It contained many innovations: an inclined column for operational convenience, an objective lens with an intermediate accelerator according to Möllenstedt, an adjustable lens allowing a magnification change in the ratio 1:3:10, and an advanced electronics and vacuum system. Despite all this, the EM8 microscope was never sold commercially. (From Mahl, H. et al., *Electron Microsc.*, 35, 1956. With permission.)



**FIGURE 5.2** Objective lens of the KM5 electrostatic transmission electron microscope (TEM) of TTC, Switzerland that could be operated at 70 kV. The insulator (d), made of epoxy resin mixed powdered quartz, was designed so that all metal–insulator–vacuum junctions occur at places of low field strength with the curved surface (f). Lens housing (e) and outer electrodes (g) are held at ground potential; central electrode (h) is held at cathode potential. High voltage is fed from the side through (i). (From Gribi, M. et al., *Optik*, 16(2), 74, 1959. With permission.)

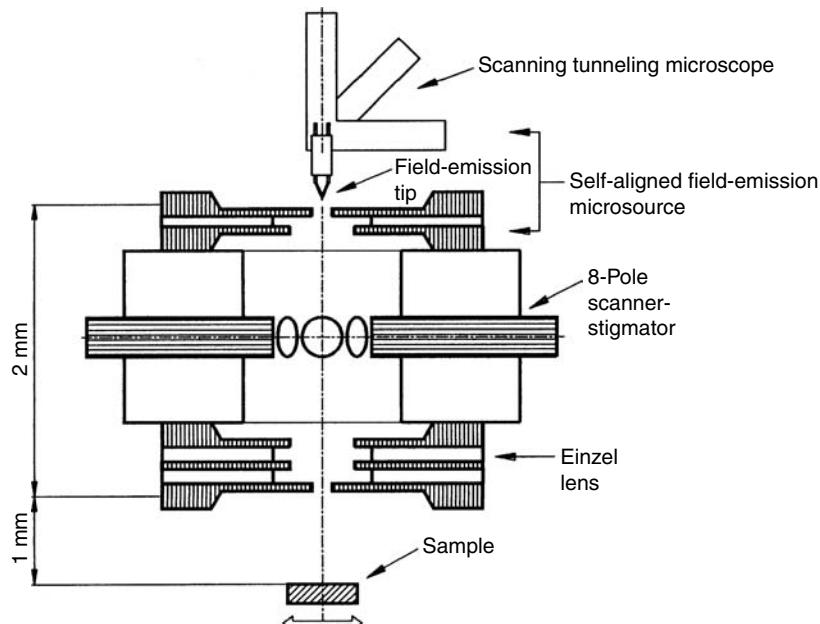
a minimum electrode distance of 2–4 mm for voltages of 40–70 kV (the highest voltage, 70 keV, was used by Gribi et al. (1959) in a commercial microscope of Trüb, Täuber & Cie. AG; the cross section through their objective lens is shown in Figure 5.2). After adding the interelectrode distance, electrode thicknesses, and the small distance to the object, the resulting minimum focal distance cannot be much less than 6–7 mm. The spherical aberration coefficient  $C_s$  was measured to be 50 mm by Hahn (1955), and thus the value of  $(C_s \lambda^3)^{1/4}$  is  $\sim 1.7$  nm for 50 keV; one can thus expect  $\sim 1$  nm resolution. With the intermediate acceleration proposed by Möllenstedt (1955), it is possible

to double the beam energy as it passes through the sample as shown in Figure 5.1, but this does not allow one to improve the resolution (the EM8 model with intermediate accelerator was never sold commercially). If the electrons are decelerated to  $\sim 5$  keV before they reach the sample (Heinemann and Möllenstedt, 1967), the spherical aberration coefficient of the objective lens decreases and thus the image resolution does not actually suffer at the lower voltage; at the same time the contrast of thin biological samples improves significantly.

In low-energy electron microscopy (LEEM) and photoemission electron microscopy (PEEM)—reviewed by Rempfer et al. (1991), Veneklasen (1992), Bauer (1994, 2007), and Feng and Scholl (2007)—the electron optical systems use a cathode lens as a critical element in which the primary electrons are decelerated and secondary electrons or photoelectrons are accelerated. These microscopes are intended for surface studies and are thus of UHV design. Other lenses in the electron optical column can then also be electrostatic because they can be more easily employed in a UHV environment whereas magnetic lenses usually have to be put outside the vacuum. Thus electrostatic lenses once again are being used for imaging, with typical beam voltages of 20 kV.

For low voltages it is possible to produce electrostatic lenses with miniature dimensions and so, once again, electrostatic lenses are being applied where space is a premium, such as in micro-columns of scanning electron microscopes (SEMs) where the tip-to-sample distance is only 3 mm (see Figure 5.3) (Murray et al., 1991). A two-dimensional array of such devices can be used for high throughput electron beam lithography. As a source of electrons *self-aligned* field-emission (SAFE) sources based on scanning tunneling microscope style tips are used. Extreme miniaturization has been proposed by Koops et al. (1995) with the *disposable* low-voltage electron optical system of miniature dimensions that can be grown on a substrate using additive lithography.

Any change of particle energy requires the use of some electrostatic optics, such as the retarding optics in scanning devices where the energy of the particles is decreased just in front of the specimen; this was reviewed by Müllerová and Lenc (1992) and Müllerová and Frank (2003). Low energy has to be reached either just in front of the sample, which forms part of a cathode lens, or in a lens immediately preceding the sample. Designs of this type often combine a decelerating lens with

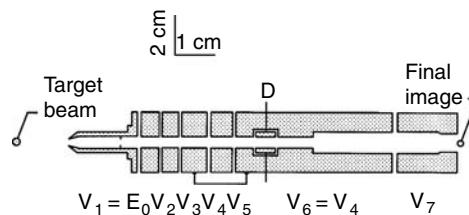


**FIGURE 5.3** Schematic cross section of a micro-scanning electron microscope column that can be operated at 1 keV. For the alignment of the field-emission gun a scanning tunneling microscope is used. (From Murray, L. P. et al., *J. Vac. Sci. Technol. B*, 9(6), 2955, 1991. With permission.)

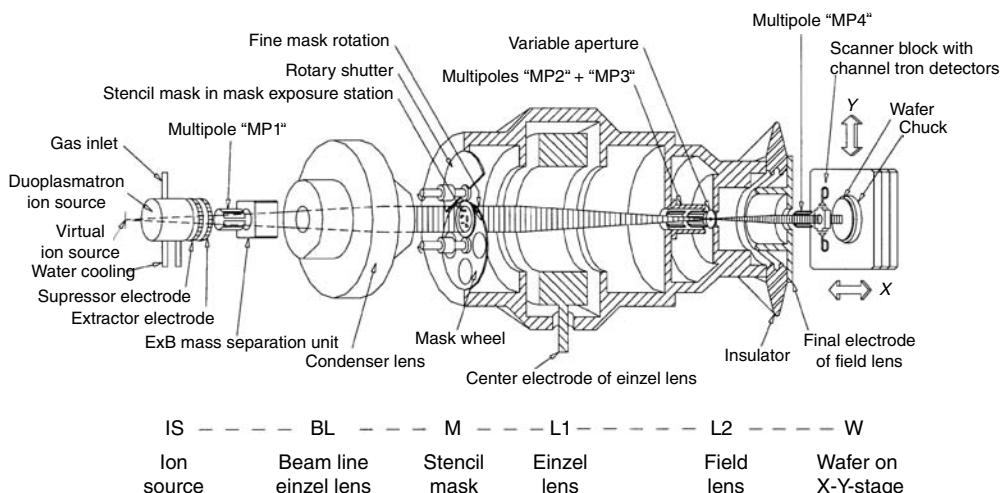
magnetic focusing, and they are used for scanning electron microscopy and electron beam testing (Frosien et al., 1989, 1995; Meisburger et al., 1992).

Interest in electrostatic optics was revived around 1970 owing to increasing demands on electrostatic lenses in electron (Roy, 1990; Erskine, 1995) and mass spectrometers (Liebl, 1989), where they are used to transport the beam from the sample to the input of the spectrometer (King, 1995). An example of a low-voltage seven-electrode lens system (Chutjian, 1979) is shown in Figure 5.4. Electrostatic optics is traditionally used in cathode ray tubes (Martin, 1986) and image pick-up tubes (Oku et al., 1988).

So far, we have discussed electrostatic lenses for electrons. However, they can also be used to focus massive particles such as ions and, unlike round magnetic lenses which have rather weak focusing properties for ions, the behavior of electrostatic lenses is independent of the particle mass. Relativistic effects have to be taken into account only for electrons at energies above, say, 50 keV (for protons at energies above 100 MeV). Thus, with the availability of high-brightness field-ionization sources (see Chapters 2 and 3), scanning ion microscopy with mass analysis became feasible. Nanopores as small as 4 nm can be drilled with a focused Ga ion beam in 20 nm thick membranes (Gierak et al., 2007). Electrostatic focused ion beam instruments are in much demand by the electronics industry that needs ion beam failure analysis, circuit editing capability, and lithographic mask repair systems (see Chapter 11). Ion beam lithography at sub-0.1  $\mu\text{m}$  resolution has been demonstrated with quite a large ion optical system, shown in Figure 5.5, which is used for projection ion beam lithography (Stengl et al., 1992), an improved version was also discussed by Loeschner et al.



**FIGURE 5.4** Typical input lens assembly of an electron spectrometer. A seven-electrode cylindrical lens (D denotes an auxiliary deflector) is used to transport the beam from the target to the final image, where the electrons arrive at an energy given by the voltage  $V_7$  of the last electrode. (From Chutjian, A., *Rev. Sci. Instrum.*, 50(3), 348, 1979. With permission.)



**FIGURE 5.5** Schematics of the Ion Microfabrication Systems GmbH projection ion beam lithography system Alpha5x. The total length of the system from stencil mask to wafer is 2 m. (From Stengl, G. et al., *J. Vac. Sci. Technol. B*, 10(6), 2825, 1992. With permission.)

(2001). It works at 55–75 keV beam energy, imaging a stencil mask onto a wafer 2 m away at five times demagnification; resolution and distortion are below 0.1  $\mu\text{m}$  on a  $12 \times 12$  mm exposure field. Electrostatic optics has also found applications in ion beam accelerators (Hyder, 1990); these instruments working at 10–34 MeV are beyond the scope of the present chapter, although we can learn some of the newest technologies for electrode production and high-voltage insulation from them.

The imaging properties of electrostatic lenses and their off-axis aberrations are less well known than for magnetic lenses, and the properties in the context of the thin lens approximation have been much less systematically studied, partly because of the premature loss of interest in electrostatic optics. Also, unlike magnetic lenses where the geometry is basically given by only a few geometric parameters (if the lenses do not saturate), so that *universal* curves showing dependence of the lens focal properties and aberrations on the lens excitation can be drawn, electrostatic lenses are difficult to characterize. Even for a simple three-electrode unipotential lens (where the beam energy changes only inside the lens) there are many geometric parameters involved (thickness of electrodes, inter-electrode distances and diameters), and the electrode shapes (e.g., rounding of the electrodes) also affect the potential distribution. The electron optical properties of electrostatic lenses will be dealt with in Section 5.2 and illustrated in terms of several selected lens geometries. Section 5.3 will be devoted to a number of practical problems in lens design such as electrical breakdown, mechanical limits, and adjustments, that is, factors that have to be observed when the lenses are applied. Finally, examples of the application of electrostatic lenses will be given in Section 5.4.

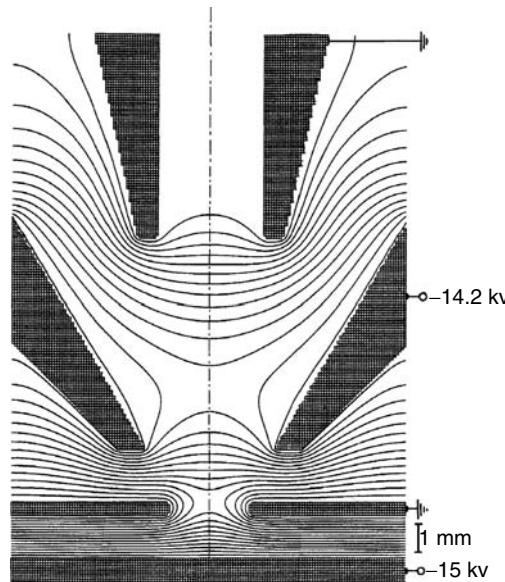
## 5.2 OPTICAL PROPERTIES OF ELECTROSTATIC LENSES

### 5.2.1 EVALUATION OF LENS PROPERTIES

Evaluation of the properties of electrostatic lenses can usually be separated into two parts, the derivation of the potential distribution followed by the calculation of the electron optical properties, mostly by using only the potential along the optic axis. We will consider the evaluation of lenses in this order.

The major problem is the solution of the Laplace equation with high accuracy. Field computation methods used in the past were not very advanced and of limited use (Weber, 1967), although the analog methods based on electrolytic tank or resistor networks were quite advanced and could provide potential values with reasonable accuracy (Francken, 1967). Potential distributions of simple analytical form have been found for some geometries which approximated the true field distribution quite well. For lenses made of cylinders of equal diameter the axial potential could be obtained by Bessel function expansions (Bertram, 1942). The Glaser–Schiske bell-shaped field (Glaser and Schiske, 1955; Kanaya et al., 1966) was often used for the computation of properties of electrostatic lenses. However, the range of geometries that could be studied in this way was rather limited (Mulvey and Wallington, 1973; Hawkes and Kasper, 1989, Chapter 35).

Suitable computer-aided design methods became available for potential computations only after 1970. Now it is easy to calculate the field of electrostatic lenses numerically on personal computers (see Appendix A). For rotationally symmetric lenses, if the correct definition of boundary conditions is used, smooth fields and better than 0.1% accuracy in both potential and its first derivative can be obtained by any of the available numerical methods. The finite difference method (FDM) is traditionally the easiest to use; SIMION is quite a popular program for this and can use square meshes with several million points (Dahl and Delmore, 1994; Dahl, 2000), a significant improvement over the previous *standard* 16,000 (Dahl and Delmore, 1988; Heddle, 1991, 2000), (Figure 5.6). More freedom in the selection of mesh is provided by the finite element method (FEM). For example, the first-order FEM using meshes with graded step size and with up to 1,000,000 points runs on a fast PC within ~1 min (Lencová, 1995a; Lencová and Zlámal, 2007). Accuracy of the potential or field can be simply determined by comparing the potential values with the results computed in a mesh using twice the number of points in each direction (Lencová, 2002).



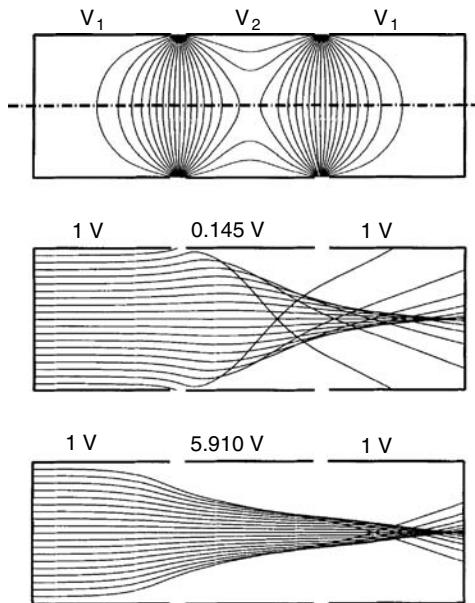
**FIGURE 5.6** Arrangement of the electrodes of a tetrode electrostatic lens of a low-energy electron microscope. The figure shows voltage applied to lens electrodes and equipotentials as computed by the program SIMION (Dahl and Delmore, 1988). (From Liebl, H. and Senftinger, B., *Ultramicroscopy*, 36(1–3), 91–98, 1991. With permission.)

The second-order FEM allows one to use curved as well as straight lines (Zhu and Munro, 1995). The highest accuracy in potential and field values is provided by the charge density method (Renau et al., 1982; Kasper, 1987; Martinez and Sancho, 1991; Read, 1996). Cubric et al. (1998) compared the accuracy of all three methods on suitably selected examples.

The computation of particle trajectories and optical properties of given potential distribution is simpler. The electron optical properties can be obtained by either accurate *direct ray-tracing* as shown in Figure 5.7 (although this procedure is more time-consuming) or the much faster solution of the paraxial trajectory equation and the evaluation of aberration integrals, which involves only the knowledge of the potential distribution on the axis of rotational symmetry (see Section 5.2.2).

The early studies of the properties of electrostatic lenses largely relied on experimental methods. For this purpose, electron optical benches were used (see the review by Skocylas et al. (1994)), in which the position of the lens could be changed. The measurements were based on the *shadow method* developed by Spangenberg and Field (1942)—see Rempfer (1985), Baranova and Yavor (1989), and Hanszen and Lauer (1967). In this method the shadow image of two meshes, one in front and another behind the lens, is recorded, and from the image size and distortion the properties of the lens can be evaluated. The advantage of this approach was that the lens being measured was actually used later; the disadvantage was the time-consuming nature of such a procedure, especially for higher accuracy. Exceptionally high measurement precision was reached by Rempfer (1985) who published the focal properties and the aberration of focal position and focal distance for a number of lens geometries, from which the value of the spherical aberration and distortion coefficients can be derived. The design of a PEEM was based on these measurements (Dunham et al., 1994).

In fact, so many parameters are involved in the description of electrostatic lens geometry that there is no complete catalog of lens data available. A number of practical geometries used in electron spectrometry were studied in detail in Harting and Read (1976) and Heddle (1991). Older data and measurements (Hanszen and Lauer, 1967; El Karch and El Karch, 1970) often do not provide sufficient accuracy, and usually they also do not differentiate between objective (real) and projective



**FIGURE 5.7** Equipotentials in a simple electrostatic lens made up of three cylinders with a small separation. Ray-traces are shown focusing a parallel beam in either a decelerating or accelerating voltage arrangement on the central electrode so that the Gaussian image of an object at infinity lies on the right-hand electrode.

(asymptotic) lens properties. Most measured data do not include lens aberrations or these measurements are not accurate enough.

### 5.2.2 THE EQUATION OF MOTION AND THE TRAJECTORY EQUATION

The electron optical properties characterizing the motion of a charged particle with charge  $e$  and rest mass  $m$  in an electrostatic field can be derived either from the relativistic equation of motion of charged particles

$$\frac{d}{dt} \left( \frac{m\vec{v}}{\sqrt{1 - v^2/c^2}} \right) = e(\vec{E} + \vec{v} \times \vec{B}) \quad (5.1)$$

or from the trajectory equation (see Chapter 6 for more details). The electrostatic field  $\mathbf{E} = -\nabla\Phi$  is in the direction of the normal to the equipotential lines shown in Figure 5.6 or at the top of Figure 5.7.

Lens action in an axially symmetric field is due to the radial component of the field intensity  $\mathbf{E}_r$ . In the vicinity of the axis it can be expressed in terms of the distribution of electrostatic potential  $\Phi(z)$  on the axis of rotational symmetry as  $\mathbf{E}_r = -r\Phi''(z)/2$ ; that is, for small  $r$  it is proportional to the distance from the symmetry axis, implying lens action (focal length independent of  $r$ ). At the same time the speed of the charged particle in the  $z$ -direction changes due to the accelerating or decelerating action of the  $z$ -component of the electrostatic field, so that the description of the behavior of electrostatic lenses is not simple. As a rule, the slope of the trajectory  $r'$  is small. Recall that the trajectory equation can be cast into the paraxial approximation (supposing small  $r$  and  $r'$ ) where it becomes

$$r'' + \frac{\gamma}{2} \frac{\Phi'}{\Phi} r' + \frac{\gamma}{4} \frac{\Phi''}{\Phi} r = 0 \quad (5.2)$$

$r$  could equally well be  $x$  or  $y$ ,  $\gamma = \gamma(z) = 1 + 2 \varepsilon \Phi(z)$ ,  $\varepsilon = |e|/(2mc^2)$ , and  $\hat{\Phi}$  denotes the *relativistically corrected potential*  $\hat{\Phi} = \hat{\Phi}(z) = \Phi(z) \cdot (1 + \varepsilon \Phi(z))$ . The sign and the zero values of the potential  $\Phi(z)$  are chosen so that  $|e|\hat{\Phi}$  equals the kinetic energy of the particle. From this expression it follows that there is a very weak dependence of the trajectory in the given electrostatic field on the mass of the particle, appearing only in the parameter  $\varepsilon$ . At an energy of 20 MeV for protons the correction is only 1%, so that for ions we can always work in the nonrelativistic limit and put  $\varepsilon = 0$  and  $\gamma = 1$ . For electrons  $\varepsilon = 10^{-6} \text{ V}^{-1}$  and a 1% relativistic correction already applies at 10 keV. In practice, the relativistic effects are mostly worth considering at electron energy 50 keV or more.

The coordinate (Picht) transformation  $R(z) = r(z)[\hat{\Phi}(z)/\hat{\Phi}(z_o)]^{1/4}$  transforms the paraxial equation into the form

$$R'' + \frac{2 + \gamma^2}{16} \left( \frac{\Phi'}{\hat{\Phi}} \right)^2 R = 0 \quad (5.3)$$

This equation is mathematically equivalent to Equation 5.2, and it is suitable for numerical integration as well as for the derivation of some properties of electrostatic lenses.

Many useful properties important for applications in charged particle optics are implied in Equation 5.2:

1. Any solution of Equation 5.2, which is a linear, homogeneous, second-order ordinary differential equation, can be expressed as a linear combination of two basic solutions. In optics the standard choice of the basic solutions is  $r_a(z)$  and  $r_b(z)$ , defined in the object plane  $z_o$ , which can be real or asymptotic. In the real case,

$$r_a(z_o) = 0 \quad r'_a(z_o) = 1 \quad r_b(z_o) = 1 \quad r'_b(z_o) = 0 \quad (5.4)$$

whereas for an asymptotic object position the initial conditions are specified in the starting point of the field  $z_s$  as

$$r_a(z_s) = z_o - z_s \quad r'_a(z_s) = 1 \quad r_b(z_s) = 1 \quad r'_b(z_s) = 0 \quad (5.5)$$

The image plane  $z = z_i$  is either real at the position where the imaging trajectory  $r_a$  crosses the axis (i.e.,  $r_a(z_i) = 0$ ) or asymptotic, obtained at the end of the field  $z_e$  as  $z_i = z_e - r_a(z_e)/r'_a(z_e)$ . The two basic solutions in the image (or end) plane determine the basic optical parameters: the image focal length  $f_i$ , the magnification  $M$ , and the angular magnification  $M_A$  as follows:

$$f_i = \frac{1}{r'_b(z_i)} \quad M = r_b(z_i) \quad M_A = r'_a(z_i) \quad (5.6)$$

2. For any two solutions of Equation 5.2 there exists an invariant (Wronski determinant), namely, for  $r_a(z)$  and  $r_b(z)$  it holds that

$$\hat{\Phi}^{1/2}(z) \{r'_a(z)r_b(z) - r_a(z)r'_b(z)\} = \hat{\Phi}^{1/2}(z_o) \quad (5.7)$$

Let us define the index of refraction in the object and the image space as  $n_o = \hat{\Phi}^{1/2}(z_o)$  and  $n_i = \hat{\Phi}^{1/2}(z_i)$ , respectively. Then if the previous equation is evaluated at  $z = z_i$ ,  $n_i M M_A = n_o$ , which is a perfect analogy of the Smith–Helmholtz (or Lagrange–Helmholtz) formula in light optics.

3. For a nonrelativistic beam, the change of all voltages by the same factor does not change the optical properties. Further, if all dimensions are scaled by the same factor, all optical properties also change by the same factor, and thus it is possible to present the optical properties divided by the same characteristic dimension (a length of the lens or a diameter of some critical electrode).

### 5.2.3 ELECTROSTATIC LENS AS A THICK LENS

As in light optics, we can describe the electrostatic lens as a thick lens characterized by positions of the principal planes  $H_o$  and  $H_i$  and the focal points  $F_o$  and  $F_i$ . The ray parallel to the axis in the object space is refracted at the image principal plane  $H_i$  into the image focal point  $F_i$ , the ray coming through the object focal point  $F_o$  is refracted at the object principal plane  $H_o$  parallel to the axis in the image space—see Figure 5.8. The object and image positions satisfy the lens equation

$$\frac{f_o}{a} + \frac{f_i}{b} = 1 \quad (5.8)$$

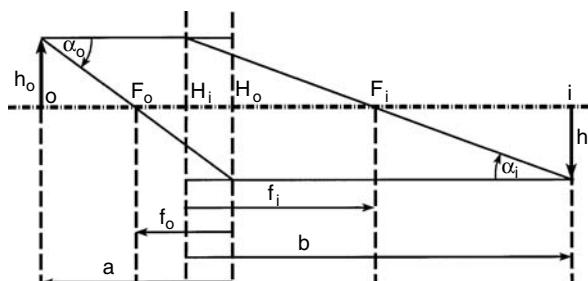
where the object focal length is  $f_o = f_i n_o / n_i$ , and the distances in the object space are measured from the object principal plane, that is,  $f_o = z_{H_o} - z_{F_o}$  and  $a = z_{H_o} - z_o$ , and distances in the image space are measured from the image principal plane, that is,  $f_i = z_{F_i} - z_{H_i}$  and  $b = z_i - z_{H_i}$ . The angular and the direct magnifications are given by

$$M_A = \frac{\alpha_i}{\alpha_o} = -\frac{a}{b} \quad M = \frac{h_i}{h_o} = -\frac{f_o}{f_i} \frac{b}{a} \quad (5.9)$$

From the known values of  $f_i$ ,  $M$ , and  $M_A$  the remaining characteristic quantities can be derived. The position of the image focal point  $z = z_{F_i}$  is given as the intersection of the line coming from the image point  $r_b(z_i)$  with slope  $-1/f_i$  with the optic axis. The intersection of this line with the line traveling parallel to the axis at unit distance gives the position of the image principal plane  $z_{H_i}$ . The relation  $z_{H_o} = z_o - M_A(z_i - z_{H_i})$ , where  $M_A$  is the angular magnification, gives the position of the object principal plane. The position of the object focal plane can be found from the relation for the object focal length  $z_{F_o} = z_{H_o} - (n_o/n_i)f_i$ .

The expression for the optical strength  $1/f_i$  of the lens can be obtained by integrating Equation 5.3 for the trajectory  $R_b(z)$ :

$$\frac{1}{f_i} = \frac{\hat{\Phi}_o^{1/4}}{16\hat{\Phi}_i^{1/4}} \int_{z_o}^{z_i} \frac{(2 + \gamma^2)\Phi'^2}{\hat{\Phi}^2} R_b(z) dz + \frac{M\gamma_i \Phi'_i}{4\hat{\Phi}_i} - \left( \frac{\hat{\Phi}_o}{\hat{\Phi}_i} \right)^{1/4} \frac{\gamma_o \Phi'_o}{4\hat{\Phi}_o} \quad (5.10)$$



**FIGURE 5.8** Notation of local length and principal and focal planes to define the optical properties of a thick electrostatic electron lens with crossed principal planes.

In the nonrelativistic approximation, the optical power of the electrostatic lens is independent of the charge and mass of the particle. In the relativistic case, the optical power becomes smaller as the kinetic energy of the particle becomes higher. In the thin lens approximation it is supposed that the trajectory  $R_b(z)$  in the lens region (region of the nonzero field intensity) does not substantially differ from unity; if the object and the image planes lie in the region of zero field intensity, the electrostatic lens is always convergent. Weak lenses, where the trajectory does not cross the axis inside the lens field, are also convergent because of the form of Equation 5.3, and the principal planes are crossed. If the object lies in the field (an important case being the cathode lens), the lens can be divergent. Similarly, if the lens becomes too strong, that is, the particle kinetic energy becomes much higher on passing through the lens or the particles are brought close to a standstill, the lens may again become divergent. The particles cannot travel in a region with negative  $\Phi$  (their kinetic energy cannot be negative), and so they can be reflected in the lens; this is an important case of an *electron mirror*.

Gaussian characteristics and also aberration coefficients of the lenses can be real or asymptotic. Basically two cases are distinguished in charged particle optics.

1. *Objective lens characteristics (real object, asymptotic image).* The ray is formed by a real trajectory  $r_a(z)$  starting at the object plane  $z = z_o$ ; the image plane  $z = z_i$  is obtained from the tangent line at the end of the nonzero field region  $z = z_c$  at intersection with the optic axis.
2. *Projective lens characteristics (asymptotic object, asymptotic image).* The imaging ray  $r_a(z)$  is a straight line starting from the plane  $z = z_s$  where the nonzero field region begins and points to the object plane  $z = z_o$ , and from the end plane of the lens region  $z = z_e$  we get the image plane  $z = z_i$  as before.

#### 5.2.4 TYPES OF ELECTROSTATIC LENSES

*Immersion lenses* in light optics have a different index of refraction on each side of the lens; in electron optics they have an overall accelerating or decelerating action; lenses using only two electrodes are a typical case. We suppose that the field in front of the lens and behind it is zero to consider the lens as a separate electron optical element. Contrary to optical lenses, where the change of index of refraction is mostly abrupt at the surface of refraction, in particle optics the index of refraction changes smoothly in the lens. The index of refraction can also reach a much larger range of values. Another important application of lenses of this type is an immersion objective lens, the first imaging lens immediately following the sample. If the sample is placed in a strong accelerating field, we call these lenses *cathode lenses*.

Three-electrode lenses are mostly operated as *unipotential lenses* (sometimes also called as *einzel lenses*), where the energy is the same in object and image space, and the focusing action is performed by changing the beam energy on passing through the lens. They can also work as immersion lenses, and the overall acceleration or deceleration is given by the voltages on the outside electrodes. In a three-electrode immersion lens, the central electrode is used to focus the image, but the image magnification is also given by the immersion ratio and by the lens geometry. To select both the magnification for a given object and image position with given immersion ratio, we need at least one additional free parameter, and we must introduce at least one additional electrode. In this way it is possible to realize *zoom lenses*. In practice, zoom lenses have five or more electrodes (see e.g., Figure 5.4) so that we can also control the beam aperture by limiting the diameter of the beam. In electron spectroscopy, multielectrode lenses often realize *energy scanning*, in which the beam energy on the lens exit changes but the beam magnification is preserved.

### 5.2.5 MATRIX DESCRIPTION OF THE TRAJECTORY AND MODELS OF SOME ELECTROSTATIC LENSES

#### 5.2.5.1 Matrix Description

Despite some complications related to the description of the behavior of electrostatic lenses, it is possible to derive simplified models based on the behavior of particles in a field at an aperture separating two regions of different electric fields and on the solution of the trajectory equation in a homogeneous field. An aperture lens as such does not exist as a self-contained element, as we cannot separate its action from the electric field around it, but it is quite a useful concept, early studied by Davisson and Calbick (1931, 1932). Homogeneous fields can be found in some parts of the electrostatic lens or in the vicinity of a flat electrode. With a simple application of matrix formalism we shall present a model of a cathode lens, formed by the homogeneous field between a sample and an aperture, and models of an immersion lens made by two apertures and of a thick unipotential three-electrode lens.

The Gaussian description of a general optical system is given in the most elegant way by using a *matrix representation* (the complete representation that also includes the aberrations is given in Chapter 6). The transfer matrix  $\hat{T}$  gives the position  $r_2$  and the slope  $r'_2$  of the ray in the plane  $z = z_2$  in terms of the position  $r_1$  and the slope  $r'_1$  of the ray in the plane  $z = z_1$ :

$$\begin{pmatrix} r_2 \\ r'_2 \end{pmatrix} = \hat{T} \begin{pmatrix} r_1 \\ r'_1 \end{pmatrix} \quad \hat{T} = \begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{pmatrix} \quad (5.11)$$

The general optical system is formed by a sequence of drift (field-free) and lens (field) regions, which is represented by sequential multiplication of the respective matrices, namely, by a matrix  $\hat{D}$  for a drift space from  $z = z_1$  to  $z_2$  and by a matrix  $\hat{L}$  for the lens with the image focal length  $f_i$  transferring a ray from the object principal plane  $z = z_{H_o}$  with the object space refraction index  $n_o$  onto the image principal plane  $z = z_{H_i}$  with the image space refraction index  $n_i$  (i.e.,  $\Phi_i^{1/2}$ ):

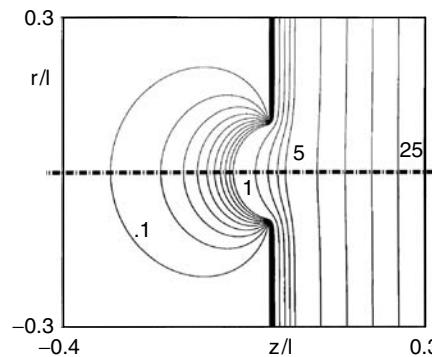
$$\hat{D} = \begin{pmatrix} 1 & z_2 - z_1 \\ 0 & 1 \end{pmatrix} \quad \hat{L} = \begin{pmatrix} 1 & 0 \\ -\frac{1}{f_i} & \frac{n_o}{n_i} \end{pmatrix} \quad (5.12)$$

More details of the matrix description are given in Dahl (1973) or Wollnik (1987); it is used also for other elements (prisms, quadrupoles, etc.).

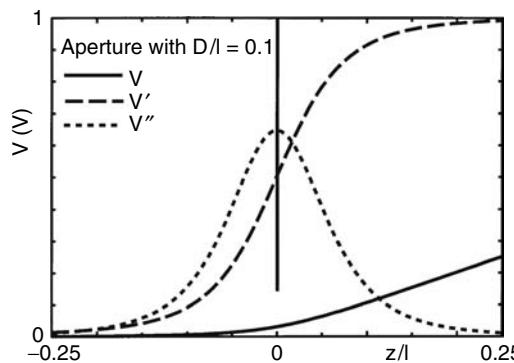
#### 5.2.5.2 Lens Action of an Aperture Lens

The aperture lens is the simplest type of electrostatic lens; equipotentials are shown in Figure 5.9. An electrode with a small circular aperture at  $z = z_a$  separates two regions with homogeneous fields in the  $z$ -direction. There is a very short region from  $z = z_1 = z_{a-0}$  to  $z = z_2 = z_{a+0}$  in which  $\Phi'' \neq 0$ , as shown in Figure 5.10. Supposing  $\Phi_2 = \Phi_1 = \Phi_a$  and  $r_2 = r_1 = r_a$ , the integration of Equation 5.2 over this region gives

$$r'_2 = r'_1 - \frac{(1 + 2\varepsilon\Phi_a)r_a}{4\Phi_a(1 + \varepsilon\Phi_a)}(\Phi'_2 - \Phi'_1) \quad (5.13)$$



**FIGURE 5.9** Equipotentials in the vicinity of an aperture, relevant to the discussion of the properties of an aperture lens (see also Figure 5.10).



**FIGURE 5.10** Distribution of potential  $V(z)$  on the axis together with its first and second derivatives in the vicinity of a round aperture (computed for two apertures at 0 and 100 V placed at a distance 1 apart and having radii  $l/10$ , calculated with the first-order finite element method programs). (From Lencová, B., *Nucl. Instrum. Methods Phys. Res.*, A363(1–2), 190–197, 1995. With permission.)

In the nonrelativistic limit ( $\epsilon = 0$ ) the transfer matrix has the form of a thin lens with the optical strength  $1/f = -\Delta r'/r_a = (\Phi'_2 - \Phi'_1)/4\Phi_a$ . The transfer matrix is then

$$\hat{T} = \hat{L} = \begin{pmatrix} 1 & 0 \\ -\frac{\Phi'_2 - \Phi'_1}{4\Phi_a} & 1 \end{pmatrix} \quad z_{H_0} = z_{H_i} = z_a \quad (5.14)$$

So a change of the field at the aperture acts as a lens that can be positive or negative.

### 5.2.5.3 Lens Action of the Homogeneous Field

There is an analytical solution of the trajectory equation for the motion in a homogeneous field. Integrating Equation 5.2 in the region from  $z = z_1$  to  $z_2$  in which  $\Phi'' = 0$ , we get

$$\begin{aligned} r'_2 &= \left( \frac{\Phi_1(1 + \epsilon\Phi_1)}{\Phi_2(1 + \epsilon\Phi_2)} \right)^{1/2} r'_1 \\ r_2 &= r_1 + 2r'_1(z_2 - z_1) \frac{\Phi_1(1 + \epsilon\Phi_1)}{(\Phi_2 - \Phi_1)(\epsilon\Phi_1)^{1/2}} \ln \frac{(1 + \epsilon\Phi_2)^{1/2} + (\epsilon\Phi_2)^{1/2}}{(1 + \epsilon\Phi_1)^{1/2} + (\epsilon\Phi_1)^{1/2}} \end{aligned} \quad (5.15)$$

In the nonrelativistic limit the transfer matrix has the form

$$\hat{T} = \begin{pmatrix} 1 & \frac{2(z_2 - z_1)\Phi_1^{1/2}}{\Phi_2^{1/2} + \Phi_1^{1/2}} \\ 0 & \frac{\Phi_1^{1/2}}{\Phi_2^{1/2}} \end{pmatrix} \quad (5.16)$$

This matrix can be written as a product of three matrices, describing the motion in free space from  $z = z_1$  to a thin immersion lens with zero optical strength at  $z = z_{H_o}$  and again the free space motion from  $z = z_{H_i}$  to  $z_2$  as follows:

$$\hat{T} = \hat{D}_2 \hat{L} \hat{D}_1 = \begin{pmatrix} 1 & z_2 - z_{H_i} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \frac{\Phi_1^{1/2}}{\Phi_2^{1/2}} \end{pmatrix} \begin{pmatrix} 1 & z_{H_o} - z_1 \\ 0 & 1 \end{pmatrix} \quad (5.17)$$

$$z_{H_i} = z_{H_o} = \frac{z_2 \Phi_1^{1/2} + z_1 \Phi_2^{1/2}}{\Phi_1^{1/2} + \Phi_2^{1/2}}$$

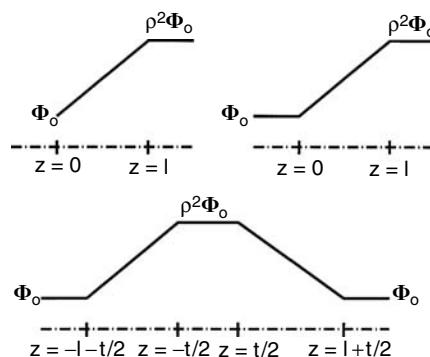
#### 5.2.5.4 Model of Lens Action (Cathode Lens, Immersion Lens, and Unipotential Lens)

The combination of aperture lenses and the lens made by a homogeneous field can be used for modeling basic types of electrostatic lenses with electrodes made of thin apertures (see Figure 5.11).

The model of a *cathode lens* consists of a homogeneous field from  $z = 0$  to  $l$ , where the potential changes from  $\Phi_o$  to  $\Phi_i = \rho^2 \Phi_o$ , followed by an aperture at  $z = l$  with zero field behind it. Multiplying the appropriate matrices we get the transfer matrix, from which the optical characteristics of the model cathode lens as a thick diverging (usual case of  $\rho \geq 1$ ) or converging ( $0 < \rho \leq 1$ ) lens are

$$z_{H_o} = \frac{2l}{\rho + 1} \quad z_{H_i} = l \quad f_i = -4l \frac{\rho^2}{\rho^2 - 1} \quad f_o = -4l \frac{\rho}{\rho^2 - 1} \quad (5.18)$$

For  $\rho > 1$  the cathode lens is always divergent; the position of the image focal point  $z_{F_i} = z_{H_i} + f_i = -3l - 4l/(\rho^2 - 1)$  does not move for usual large values of acceleration ratios. The position of the image plane and image and angular magnifications are



**FIGURE 5.11** Notation of coordinates and potentials for electrostatic lens models of a cathode lens, immersion lens, and thick unipotential lens used for the derivation of formulas for lens properties.

$$z_i = l - \frac{4l\rho^2}{(3\rho-1)(\rho+1)} = \frac{l(1-2M)\rho}{\rho+1} \quad (5.19)$$

$$M = \frac{2\rho}{3\rho-1} \quad M_A = \frac{3\rho-1}{2\rho^2} \quad MM_A\rho = 1$$

For  $\rho \gg 1$  we see that  $M \rightarrow 2/3$ ,  $M_A \rightarrow 0$ , and  $z_i \rightarrow -l/3$  (Lenc and Müllerová, 1992).

The model of an *immersion lens*, that is, a lens in which the energy of the particle changes, consists of two aperture lenses positioned at  $z = 0$  and  $z = l$  with a zero field in front of the first aperture and behind the second aperture, and a homogeneous field in between. This model was used by Zworykin et al (1945). With the same notation as for the cathode lens we get

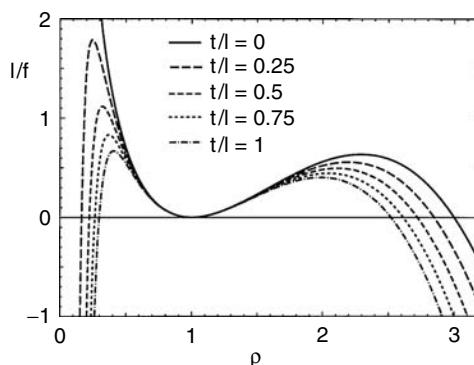
$$\begin{aligned} z_{H_0} &= -\frac{4l}{3(\rho^2-1)} & f_o &= \frac{8l\rho(\rho+1)}{3(\rho^2-1)^2} \\ z_{H_i} &= -\frac{(\rho+3)l}{3(\rho^2-1)} & f_i &= \frac{8l\rho^2(\rho+1)}{3(\rho^2-1)^2} \end{aligned} \quad (5.20)$$

so that the immersion lens always behaves like a thick converging lens.

The model of a symmetrical *unipotential lens* can be obtained as a combination of two immersion lenses, and it is also possible to model in this way a lens with a thick central electrode of thickness  $t$ . The model consists of four aperture lenses positioned at  $z = -l - t/2$ ,  $z = -t/2$ ,  $z = t/2$ , and  $z = l + t/2$  (see Figure 5.11) with a zero field in front of the first aperture and behind the last aperture and homogeneous fields of opposite signs between the first two and the last two apertures. There is no field between the two apertures in the center and for  $t = 0$  the two intermediate apertures coincide. The potential on the outside apertures is  $\Phi_o$ , the potential of the inner aperture is  $\Phi_i = \rho^2 \Phi_o$ . It holds that

$$\begin{aligned} z_{H_0} = -z_{H_i} &= \frac{(\rho-1)^2[16l^2\rho^2 + 2l(\rho+1)(7\rho-1)t + (\rho+1)^2t^2]}{2(\rho+1)[8\rho^2l(3-\rho) - 3(\rho+1)(\rho-1)^2t]} \\ f_o = f_i &= \frac{64\rho^3(\rho+1)l^2}{3(\rho^2-1)^2[8\rho^2l(3-\rho) - 3(\rho+1)(\rho-1)^2t]} \end{aligned} \quad (5.21)$$

This equation simplifies if  $t = 0$ , where for  $0 < \rho < 3$  the unipotential lens behaves like a thick converging lens, and for  $\rho > 3$  it behaves like a thick diverging lens. The dependence of  $1/f_o$  on  $\rho$  for several values of lens thickness is shown in Figure 5.12. The validity and applicability of the aforementioned models will be checked in Figure 5.16.



**FIGURE 5.12** Optical strength  $1/f$  for a unipotential lens model for various thicknesses of the central electrode.

### 5.2.6 ABERRATIONS OF ELECTROSTATIC LENSES

The complete treatment of aberrations in electron optics that also covers electrostatic lenses is given in Chapter 6. In most cases studied in practice, only the axial aberrations are considered: the third-order spherical aberration coefficient  $C_S$  and the first-order chromatic aberration  $C_C$ , for which the error in the image plane is defined as

$$\delta r_i = MC_S \alpha_o^3 + MC_C \alpha_o \Delta V/V \quad (5.22)$$

The coefficients  $C_S$  and  $C_C$  depend, for a lens with given electrode voltages, on the reciprocal magnification  $m = 1/M$  as

$$C_S = C_{S0} + C_{S1}m + C_{S2}m^2 + C_{S3}m^3 + C_{S4}m^4 \quad C_C = C_{C0} + C_{C1}m + C_{C2}m^2 \quad (5.23)$$

For weak lenses it is possible to perform the integration of aberration coefficients under certain simplifying assumptions, namely, that the Picht trajectory in Equation 5.3 does not change appreciably in the lens field. Then it is possible to show that the axial aberration coefficients for both immersion lenses (Lenc and Lencová, 1997) and unipotential lenses (see Hawkes et al., 1995) can be expressed as

$$C_{C0} = 2f_o \quad C_{S0} = \frac{f_o^3}{LD} \quad f_o = \frac{L}{(\rho^2 - 1)^2} \quad (5.24)$$

Here  $L$  and  $D$  are characteristic lengths of the lens and aberration, respectively, given by the normalized axial potential  $\psi(z)$  as

$$L = \frac{16}{3 \int \psi'^2 dz} \quad D = \frac{64}{5L^2 \int \psi''^2 dz} \quad \psi(z) = \frac{\Phi(z) - \Phi_o}{(\rho^2 - 1)\Phi_o} \quad (5.25)$$

Crewe (1991a) in a study of unipotential lenses proposed that  $f = L/(2\ln \rho)$ , which behaves similarly to formula 5.25, and showed that for selected shapes of normalized axial potential  $C_{C0} \approx 2f$ ,  $C_{S0} \approx 20f^3/L^2$ , whereas for magnetic lenses  $C_C \approx f$ ,  $C_S \approx 5f^3/L^2$ ; thus, magnetic lenses have smaller axial aberrations for the same focal length: chromatic aberration by a factor of two, spherical aberration by a factor of four. A similar study has also been performed for weak three-electrode lenses (Crewe, 1991b).

For immersion lenses it is sufficient to study the lens behavior and give the lens data in only one regime. From the coefficients of axial aberrations in the accelerating regime for  $\rho^2 = U_2/U_1 > 1$  we get the coefficients in decelerating mode, denoted by an asterisk, as

$$C_{Ci}^* = C_{C(2-i)} \rho^3 \quad C_{Sj}^* = C_{S(4-j)} \rho^3 \quad i = 0, 1, 2, \quad j = 0, \dots, 4 \quad (5.26)$$

For symmetrical lenses there are thus three independent coefficients that give the spherical aberration (Hawkes and Lencová, 2002). The dependence of polynomial coefficients on magnification for weak lenses can be simplified to (Heddle, 1991, 2000; Ura, 1994; Lenc and Lencová, 1997).

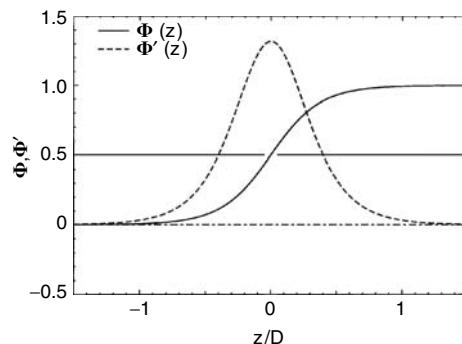
$$C_S = C_{S0} \left(1 - \frac{1}{\sigma M}\right)^4 \quad C_C = C_{C0} \left(1 - \frac{1}{\sigma M}\right)^2 \quad (5.27)$$

where for unipotential lenses  $\sigma = 1$  and for immersion lenses  $\sigma = \rho^{1/2} = (\Phi_i/\Phi_o)^{1/4}$ .

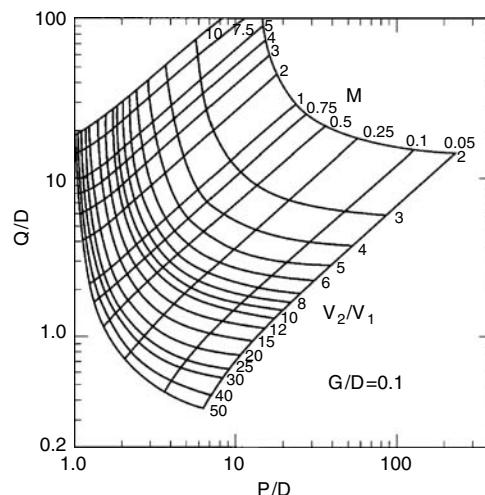
### 5.2.7 TYPICAL PROPERTIES OF ROUND ELECTROSTATIC LENSES

#### 5.2.7.1 Two-Electrode Immersion Lenses

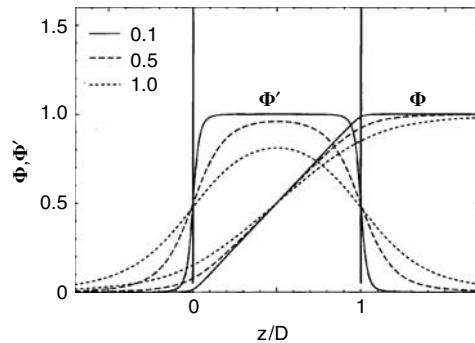
Two-electrode lenses of the simplest geometry are usually made of either two cylinders or two apertures of equal diameter. These lenses are immersion lenses, because the energy of the particle changes on passing through the field of the lens. In the case of cylinder (tube) lenses, the most frequently studied geometry is that a small distance  $G$  given by  $G/D = 0.1$  separates cylinders of equal diameter  $D$ . The axial potential and the field for this ratio are shown in Figure 5.13. This distribution does not change much for a smaller cylinder separation or if the two cylinders are made of thin or thick electrodes. For higher values of  $G/D$  the actual geometry of the gap region and that of the surrounding electrodes has to be taken into account. The electron optical properties of such a lens are easy to calculate. Cylinder lenses are the most frequently used lenses in electron spectrometry. Their properties are often presented in the form of so-called P/Q curves, where  $P$  and  $Q$  are the object side and image side distances from the center of the lens, instead of using the curves of focal distance  $f$  and position of the focal plane  $F$ . These distances are usually scaled by the lens diameter (Figure 5.14).



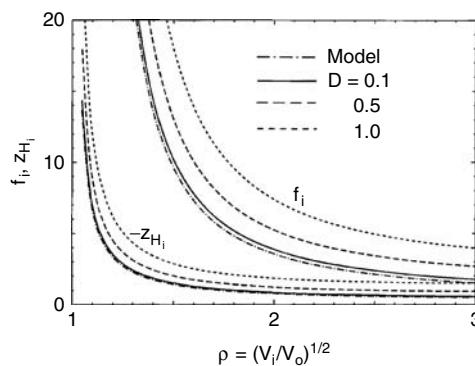
**FIGURE 5.13** Axial potential and its first derivative for a two-tube lens of unit diameter with  $G/D = 0.1$ .



**FIGURE 5.14** P/Q curves characterize the imaging properties of a two-tube lens with  $G/D = 0.1$ .  $P$  is the object distance and  $Q$  the image distance from the lens center. The voltage required for imaging conditions as well as image magnification can be obtained from this curve. (From Read, F. H. et al., *J. Phys. E: Sci. Instrum.*, 4(9), 629, 1971. With permission.)



**FIGURE 5.15** Axial potential and its first derivative for a two-aperture lens with unit distance between the two apertures, for several values of the diameter D.



**FIGURE 5.16** The comparison of the image focal distance and image principal plane position with an analytic model given by Equation 5.20 for a two-aperture lens with a unit distance between the apertures.

Lenses made of two thin apertures of equal diameter D and separated by a distance L are also quite frequently studied. In this case during computation of the electrostatic field one has to take into account that it is necessary to close the lens by attaching to it a cylinder of suitably chosen diameter (it has to be more than 3D so as not to affect the field significantly) and whose outer ends are closed at a distance of at least 5D. Either D or L can be used for scaling the displayed properties. The potential and field for several values of D/L are shown in Figure 5.15. For small values of D/L, the focal distance is expressed by Equation 5.20. For the same values of D/L as in Figure 5.15, Figure 5.16 gives a set of focal distances for the accelerating regime of the immersion lens. As discussed in Section 5.2.5, the immersion lens is always a thick convergent lens.

The understanding of two-electrode lenses provides a basis for the computation of multielectrode lenses. In practice, all geometric parameters (electrode thickness, rounding of the electrodes) of the lenses and all electrodes in the vicinity of the lens itself have to be taken into account. The geometric parameters also influence lens aberrations. Generally, as follows from Equation 5.25, the lenses with smooth and slowly decreasing fields, that is, lenses with larger diameter, have lower coefficients of spherical aberration, whereas the chromatic aberration coefficient is rather fixed to the lens focal length. For detailed studies of lens properties, see DiChio et al., (1974), Harting and Read (1976), and Heddle (1991).

### 5.2.7.2 Three-Electrode Unipotential Lenses

Three-electrode lenses can either have the same voltage on the outer electrodes, when they are called unipotential lenses, or all three voltages applied to them can be different. Again, the most

frequently studied geometries are those of three cylinders, used mostly in electron spectroscopy, or three apertures of equal diameter. A number of three-electrode lenses studied experimentally by Rempfer (1985) are still used in imaging systems. The properties of commonly used two- and three-electrode lenses were published by Harting and Read (1976); see also Hawkes and Kasper (1989) for a complete list of references, or Heddle (1991).

The range of optical properties of these lenses can be much broader than it is in the case of immersion lenses. For example, a model for the potential for a symmetrical unipotential lens presented in Figure 5.11 shows that the unipotential lenses can be either convergent or divergent (Figure 5.12). Unipotential lenses operated close to a mirror mode allow only a selected energy region of the beam to pass through the lens, so that the lens is used as a band-pass filter. The study of filters requires quite a high computational accuracy for both the potential evaluation and for ray-tracing; the most recent studies are those of Stenzel (1971), Lenz (1973), and Niemitz (1980).

Immersion lenses with three electrodes use the potential on the central electrode to focus the beam, whereas the potentials on the outer electrodes define the overall ratio of beam energy in object and image space. A large number of asymmetric lens geometries were studied in connection with the design of the lens in front of a field-emission electron or ion gun and for probe-forming lenses in ion beam systems. In these lenses, generally, the central electrode of the lens should be placed closer to the source in the gun lens or closer to the sample in the probe-forming lens; see, for example, Riddle (1978), Orloff and Swanson (1979), Mayer and Gaukler (1987), and Burghard et al. (1989).

### 5.2.7.3 Multielement and Zoom Lenses (Movable Lens)

In three-electrode immersion lenses the central electrode is used for focusing, and thus over a certain range it can fulfill the condition for zoom lenses where the object and image positions are fixed for changing immersion ratio. Also to keep the magnification of the image at about the same level, it is necessary to add at least one additional electrode. The system of electrostatic lenses usually has to incorporate an aperture limiting the angular beam extent, and often another aperture limiting the beam size (see, e.g., Figure 5.4). Such an aperture (or both of them) is placed inside a central longer electrode at such a position that it does not significantly influence the potential distribution. A larger number of articles are devoted to the design of such systems, and we have selected just the most representative ones. Martinez and Sancho (1983) and Martinez et al. (1983) studied four-cylinder electrostatic lenses and their use for energy scanning at constant image position and magnification. Five-element lenses were studied by Heddle and Papadovassilakis (1984) and Trager-Cowan et al. (1990a). An interesting idea of *moving* the lens by cutting the cylinder electrode into a number of short, closely spaced elements to which appropriate voltages are supplied, thus allowing the focusing field to be arbitrarily shifted, was proposed by Read (1983), and studied by Trager-Cowan et al. (1990b). The discussion of the design of a seven-electrode lens is saved for the section on optimization (Section 5.2.9).

### 5.2.7.4 Immersion Objective Lenses

One of the possibilities for minimizing the axial aberration coefficients involves the use of immersion objective lenses. Electrostatic immersion objective lenses are characterized by the fact that the object is positioned in the low potential region of an accelerating immersion lens (in a direct imaging system) or of a decelerating immersion lens (in a probe-forming system). When the object or emissive cathode is *immersed* in an electrostatic field, which is usually strong, we often speak of *cathode lenses* (Hawkes and Kasper, 1989, section 35.1). The first theoretical studies on immersion electrostatic objective lenses were published by Recknagel (1941). In fact, the first SEM built in the United States used an immersion electrostatic lens (Zworykin et al., 1942, 1945).

At first, cathode lenses were studied mostly for their use as objective lenses of emission electron microscopes (Möllenstedt and Lenz, 1963). Nowadays, immersion objective lenses form a basic part of many electron and ion optical systems. The use of immersion objective lenses in low-voltage SEM

was discussed by Reimer (1993) and in a review article by Müllerová and Lenc (1992) and Müllerová and Frank (2003). Their use in LEEM and PEEM was described by Chmelík et al. (1989), Griffith and Engel (1991), Rempfer et al. (1991), Veneklasen (1992), and Bauer (1994). The applications in low-energy electron beam inspection and lithographic systems were reviewed by Plies (1994) and in ion probe microscopes by Liebl (1989).

Since Recknagel's first article, the calculations have been repeated by many authors (a detailed study of the cathode lens is given by Lenc and Müllerová, 1992) without changing the essential result that the minimum values of aberration coefficients are given by the value of the electrostatic field intensity  $\mathbf{E}$  and the potential  $\Phi_o$  giving the kinetic energy of the electron at the object plane,

$$C \approx S \approx \frac{\Phi_o}{E} \quad (5.28)$$

If only a weak electric field at the object is tolerable, an aperture held at the same voltage as that of the sample is placed above the sample at a distance  $w$ . The values of the axial aberration coefficients are of the order of the working distance  $w$ , that is, the distance from the object to the first electrode at object potential.

An immersion objective lens can be regarded as an accelerating lens followed by a focusing lens (unipotential electrostatic lens, magnetic lens, or a compound lens). The values of the axial aberration coefficients are given essentially by the values of the accelerating lens, because the angular aperture of the focusing lens is much smaller. The overlap of the magnetic field with the electrostatic field of the accelerating lens does not change the aberration coefficients significantly, as can easily be seen from the analytically solvable model with quadratic electrostatic potential superimposed on a homogeneous magnetic field (Lenc, 1995). The axial aberration coefficients of the accelerating lens  $S_{(a)}$  (spherical aberration) and  $C_{(a)}$  (chromatic aberration) are given to a remarkably good approximation by the simple formula

$$S_{(a)} \approx C_{(a)} \approx \frac{1}{2} \int_{z_0}^{z_a} \left( \frac{\Phi_o}{\Phi(z)} \right)^{1/2} \left[ \frac{\Phi_o}{\Phi(z)} - \frac{\Phi_o}{\Phi_a} \right] dz \quad (5.29)$$

For the complete immersion objective lens, we have to add the aberrations of the focusing lenses  $S_{(f)}$  and  $C_{(f)}$ ; the addition rules for the total value of the coefficient give

$$S = S_{(a)} + \left( \frac{\Phi_o}{\Phi_a} \right)^{3/2} \frac{1}{M_{(a)}^4} S_{(f)} \quad C = C_{(a)} + \left( \frac{\Phi_o}{\Phi_a} \right)^{3/2} \frac{1}{M_{(a)}^2} C_{(f)} \quad (5.30)$$

Here  $M_{(a)}$  is the magnification of the accelerating lens for which in the weak lens approximation we have

$$\frac{1}{M_{(a)}} = 1 - \frac{1}{4\Phi_o} \int_{z_0}^{z_a} \left( \frac{\Phi_o}{\Phi(z)} \right)^{1/2} (\Phi''(z) + \eta^2 B(z)^2) \int_{z_0}^z \left( \frac{\Phi_o}{\Phi(\zeta)} \right)^{1/2} d\zeta dz \quad (5.31)$$

From the expression for the axial aberration coefficients it is clear that the immersion objective lenses used at very low voltage are very different from those used in field-free low-voltage modes (classification by Rose and Preikszas, 1992). In the first case, the cathode lens can provide values of aberration coefficients small enough for high-resolution work

$$S_{(a)} \approx C_{(a)} \approx l \frac{\Phi_o \Phi_a^{1/2} - \Phi_o^{1/2}}{\Phi_a^{1/2} + \Phi_o^{1/2}} \quad (5.32)$$

Here  $l$  is the distance over which the potential changes from  $\Phi_o$  to  $\Phi_a$ . With typical electron energy of electrons in the object plane of 10 eV, an immersion ratio of 1000 can easily be obtained and the

values of aberration coefficients are approximated by Recknagel's formula. Typical design examples are collected in Chmelík et al. (1989) and Müllerová and Lenc (1992); other examples are given by Rose and Preikszas (1992), Meisburger et al. (1992), and Hordon et al. (1993).

In the second case, the aberration coefficients are approximated by

$$S_{(a)} \approx C_{(a)} \approx \frac{1}{2} \left( w + l \left( \frac{\Phi_o}{\Phi_a} \right)^{1/2} \right) \left( 1 - \frac{\Phi_o}{\Phi_a} \right) \quad (5.33)$$

With a typical energy of 100 eV to 1 keV on the sample, the usual immersion ratio is 100 to 10, and the contribution of the focusing lens to the aberration coefficients cannot be neglected. For electrons, the use of compound magnetic and electrostatic lenses could be beneficial in this case. A typical design is described by Frosien et al. (1989), later used in the Zeiss Gemini SEM (Weimer and Martin, 1994). A more detailed study of the optimized compound lens is provided by Preikszas and Rose (1995).

### 5.2.7.5 Electron Mirror

In an electron mirror, the particles are brought to a standstill and reflected. Mirrors are theoretically very attractive as the sign of the axial aberrations and can be opposite to that of the corresponding aberrations of rotationally symmetric lenses (see also Chapter 6). Recently, they were reconsidered in connection with LEEMs (Rempfer, 1990) where they can be used for the correction of axial aberrations; the theory of mirrors was completely rewritten by Rose and Preikszas (1995) and Preikszas and Rose (1997). The derivation of optical properties of mirrors represents some problems in computations, as at the return point of electron trajectory the condition of small  $r'$  breaks down. In Figure 5.17 the behavior of a mirror is shown, operated at 1:1 magnification; the aberration coefficients derived from ray-tracing give similar values to those of aberration theory. By adding two electrodes to the mirror of Figure 5.17, a so-called tetrode mirror is obtained, where the two major axial aberrations can be corrected for a large range of their ratio and thus for different imaging conditions (Preikszas, 1995).

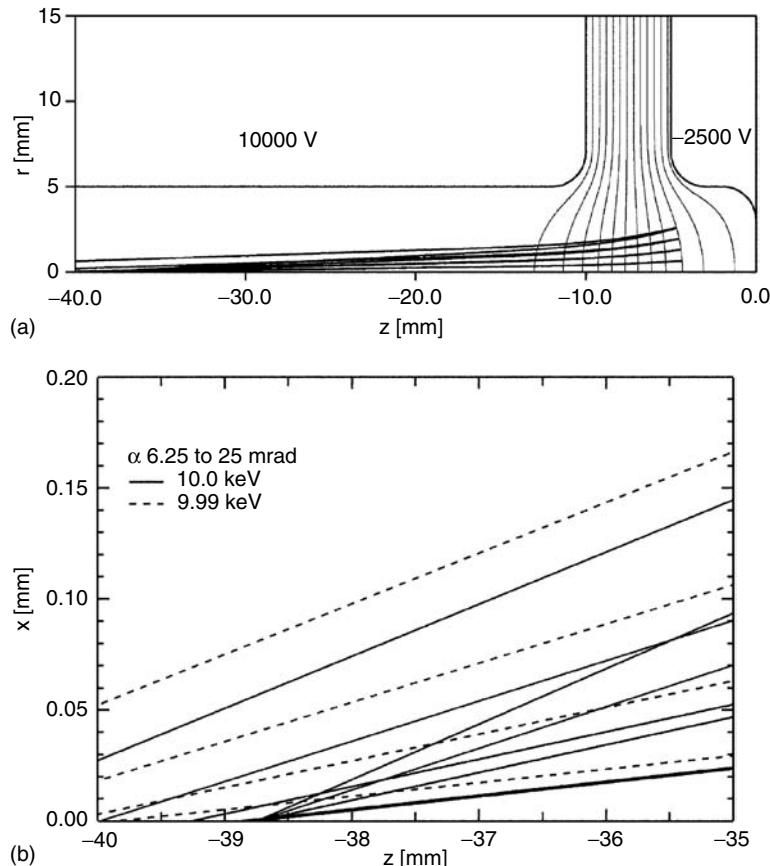
Another use of mirrors is in ion spectrometers, and an extensive study of three-cylinder lenses for use as mirrors was done by Berger and Baril (1982), Berger (1983), and Boulanger and Baril (1990).

### 5.2.7.6 Grid and Foil Lenses

Foil lenses are made of thin films transparent to electrons, and if they are a part of an electrode to which a suitable voltage is applied, they allow us to overcome one of the obstacles given by Scherzer's theorem, namely, the continuity of the electric field on the axis. In this way it is possible in principle (though difficult in practice) to realize lenses of very low spherical aberration—see Meisburger and Jacobsen (1982). An overview of attempts to correct spherical aberration in TEM and STEM instruments by adding a foil lens onto the magnetic objective lens is given by Hanai et al. (1984). The most recent study of foil corrector in an SEM has been investigated by van Aken et al. (2004) (see Chapter 6 for more information).

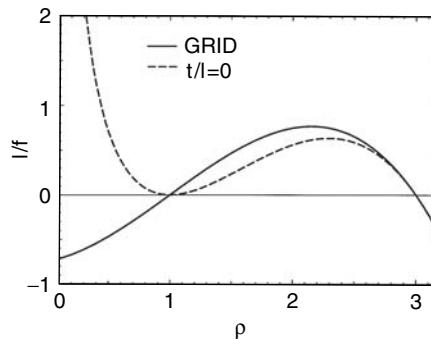
Grid lenses (sometimes also called gauze lenses) realize an auxiliary equipotential with discontinuity of the field with the help of a grid or gauze with small openings. It is possible to approximate the behavior of the unipotential gauze lens by a model similar to what was discussed in Section 5.2.5 but with the aperture lens at the position of the gauze missing; then the focal length and the position of the focal planes are given by

$$z_{H_o} = -z_{H_i} = \frac{l(\rho - 1)^2}{(\rho + 1)(3 - \rho)} \quad f_o = f_i = \frac{4l}{(\rho + 1)(\rho - 1)(3 - \rho)} \quad (5.34)$$



**FIGURE 5.17** The upper part of the figure shows electron trajectories starting at  $z = -38.7$  mm in a diode electrostatic mirror (Preikszas and Rose, 1997). On the zero equipotential, the electrons are mirrored, and the outer-most rays starting with 50 mrad angle leave the mirror almost parallel to the axis. From the ray trace positions in the image we can derive the correct values of axial aberration coefficients. In the lower part of the figure the ray traces up to maximum 25 mrad are shown for two energies in more detail. Clearly the sign of the third order spherical aberration and the first order chromatic aberration are opposite to that of the electrostatic lens shown in Figure 5.7.

and, as shown in Figure 5.18, such lenses are stronger than lenses without the grid, in particular, for small acceleration ratios, as discussed by Hanszen and Lauer (1967), and divergent for  $\rho < 1$ . This property was thus utilized for focusing on linear accelerators of ions, and due to their small geometric aberrations they were applied in older image pick-up tubes (Verster, 1963). In cathode ray tubes, grid lenses are often used for postdeflection amplification (Martin, 1986). Recently, grid lenses were applied in the newest version of the ion beam projection system (Chalupka et al., 1994) (the previous version is shown in Figure 5.5). A very attractive idea is that of correcting spherical aberration with a grid of special shape, outlined more than 50 years ago by Zworykin et al. (1945). This was actually used for some time in electron spectrometry (Read, 1982), and a similar arrangement was studied by Kato and Sekine (1995). The openings in the grid inevitably slightly disturb the lens performance; the effect of the grid openings (so-called *facet lenses*) can be modeled as aperture lenses with inverse focal distance  $1/f = -\Delta\Phi/(4\Phi)$  (Equation 5.13). Williams et al. (1995) used a three-dimensional charge density program and ray-tracing to model the effect of the facet lenses for several types of meshes. Kato and Sekine (1996) investigated the effect of mesh openings and concluded that the presence of holes does not alter the effectiveness of the aberration correction by meshes.



**FIGURE 5.18** Behavior of a gauze unipotential lens compared with that of a thin aperture lens (comparison of the two models).

### 5.2.8 CYLINDRICAL AND ASTIGMATIC LENSES

In some applications it is not necessary to provide stigmatic focusing, and cylindrical or astigmatic lenses can be used. For a review of these lenses see Chapter 39 in Hawkes and Kasper (1989) or the review of Baranova and Yaror (1989).

The simplest equivalents of cylindrical lenses in optics are electrostatic lenses of planar symmetry where the potential is independent of one of the coordinates perpendicular to the beam axis. In a first approximation their focal distance is half that of rotation symmetric lenses of the same cross section, as stated very early by Davisson and Calbick (1932). Electron optical properties of some typical geometries are given by Harting and Read (1976).

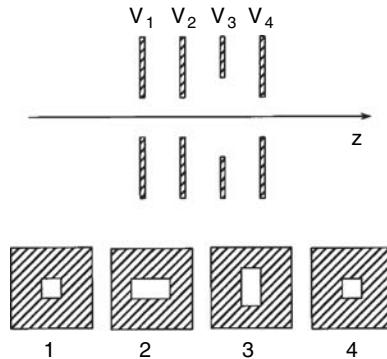
Quadrupole lenses are stronger focusing than rotationally symmetric lenses, and so they are mostly used in high-energy beams. Because they focus differently in two perpendicular directions, they can be used to modify the input beam of mass or energy spectrometers. To get a stigmatic image, three or more lenses have to be used. Ribbon beams of some ion sources are also best manipulated with these lenses. Transaxial lenses that have rotationally symmetric electrodes are often studied but the beam axis is perpendicular to the axis of rotational symmetry.

One of the reasons for interest in multipole lenses is the possibility of compensating some of the axial aberrations (see Chapters 6 and 12). Okayama (1990) showed an application of a self-aligned electrostatic quadrupole correction-lens system, with three quadrupoles and two aperture electrodes, that allows one to correct spherical aberration of a probe-forming lens.

Another important class of lenses is crossed-lenses, using rectangular holes instead of circular openings. Modern wide-band oscilloscope tubes use crossed-lenses to provide increased deflection and to provide a smaller beam due to partial correction of spherical aberration. They are characterized by their simple construction, short length, ease of tuning, and they allow a flexible control of astigmatism. Recently, Baranova and Read (1994) studied a four-electrode crossed-lens shown in Figure 5.19; for the potential computation they used a program based on the charge density method in three dimensions. Five-electrode crossed-lenses and their aperture aberrations were studied by Baranova et al. (1996). Carillon and Gauthé (1964) realized a velocity analyzer with a three-electrode crossed-lens in a TEM. A cylindrical electrostatic lens was utilized as a high-resolution velocity analyzer by Möllenstedt (1949).

### 5.2.9 OPTIMIZATION

With so many geometric parameters of individual lenses or systems of lenses involved, it is natural that the need for optimization is high. Most of the optimizations of electrostatic lenses and systems have been achieved by systematic investigation by establishing the dependence of the optical properties on geometric or electrical variables. A trial-and-error approach can often provide an



**FIGURE 5.19** The geometry of the crossed four-electrode lens studied by Baranova and Read. (From Baranova, L. H. and Read, F. H., *Rev. Sci. Instrum.*, 65(6), 1994, 1994. With permission.)

adequate design, in particular, if design constraints such as available space and voltages do not leave much freedom for experimenting. An example of a study of unipotential lenses is given by Shimizu (1983). The effects of the various lens geometric parameters on the performance of lenses was systematically investigated for lenses in front of a field-emission gun by Orloff and Swanson (1979) and for a probe-forming electrostatic lens by Kurihara (1985) and Tsumagari et al. (1988). It is shown that a suitable lens design can be found without having to use optimization methods.

At the moment we do not have a fully general optimization method, but rather a number of approaches each applied to a specific problem. Frequently, standard optimization methods are applied to find an optimum electrostatic lens design. Glatzel and Lenz (1988) prefer to use parameters related directly to the potentials on electrodes and to the shape and position of the field-forming electrodes as optimization variables. They succeeded in improving published results of Szilagyi (1983), which was claimed to be an optimum design of an immersion lens for ion beam lithography. Benez et al. (1995) also varied the geometric parameters, using the optimization based on a genetic algorithm, and documented the success of the method by improving the resolution of the well-known asymmetric lens of Orloff and Swanson (1979).

The calculation of the potential distribution was too time-consuming to be included inside an optimization loop. In particular, attempts have been concentrated on finding an approximation to the optimum axial potential distribution, from which it is eventually possible to derive the shape of electrodes. An overview of these efforts is given by Szilagyi (1988), who mainly tried to find an optimum axial potential expressed by a cubic spline. A different approach, the so-called second-order electrode method, was developed using a simple model to get the axial potential from the potential of the electrodes and radii of selected points on the electrodes with variable length. The radii and lengths are also used as parameters of the optimization, so that the result is an optimum electrode shape—see Adriaanse et al. (1989) and van der Steen et al. (1990). The use of this method for the design of a multimode transport lens is given by van der Stam et al. (1993). The axial field distribution suitable for finding a minimum of axial aberrations for compound lenses is discussed by Preikszas and Rose (1995). In any approach based on the search of axial potential or field distribution it is necessary to check the results with an accurate computation of properly reconstructed electrodes; otherwise the optimization methods remain controversial and the results of these optimizations sometimes look suspicious.

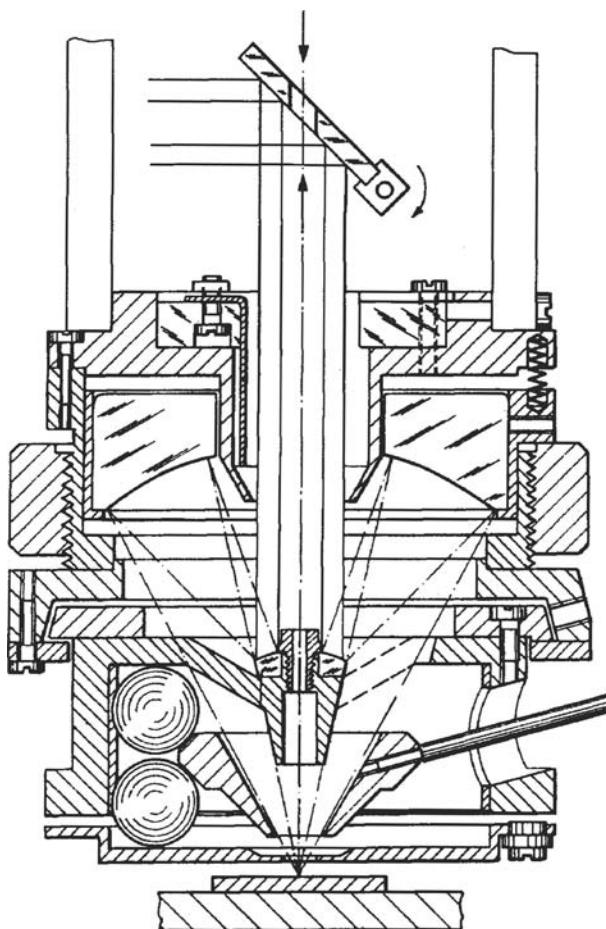
More recently Degenhardt (1997), who optimized the design of a four-lens transfer optics for a PEEM, showed that axial potential of the lenses he used can be successfully approximated with analytical functions that represent the potential of an aperture lens (see Hawkes and Kasper, 1989, for details). With known analytical expression for the axial potential it is possible to calculate quickly optical properties and easily find the optimum electrode position, geometry, and voltage for given imaging conditions. Such an approach cannot be used for lens systems made of

cylinders. The matrix approach used by Chutjian (1979) can only be used if the cylinder length is much larger than the diameter, which is obviously not the case of the system shown in Figure 5.4. Boesten (1988) systematically modified the voltage settings of the individual lenses for Chutjian's geometry to obtain better performance of the system. He used an approach previously pioneered by Fink and Kisker (1980) who developed a fast ray-tracing procedure that could, for known potential distribution of unit voltage on each of the lens electrodes in turn, obtain an optimum setting of voltages on individual electrodes for given requirements.

### 5.3 PRACTICAL DESIGN OF ELECTROSTATIC LENSES

#### 5.3.1 DESIGN PROBLEMS

The region of the electrostatic lens analyzed in the preceding section is the part that is used for the focusing of electron or ion beams; it lies close to the axis and it usually forms only a small part of the total volume of the actual electron lens (Figure 5.20). In general, the lens is a complex



**FIGURE 5.20** Cross section of the low-energy electron microscope (LEEM) objective lens design of Liebl and Senftinger (1991). The part of the lens shown in Figure 5.6, which is important for electron optical performance, forms only a small part of the actual lens. The center electrode is held between six sapphire balls (only two of them appear in this section). The adjustable spherical mirrors of the Schwarzschild objective are used for optical observation and UV irradiation. The mirror at the top is swung out of the way for incoming UV irradiation from above. (From Liebl, H. and Senftinger, B. 1991, *Ultramicroscopy*, 36(1–3), 91–98. With permission.)

three-dimensional structure: high voltage is fed in from the side and the insulators themselves may not possess rotational symmetry; the electrodes may be perforated for better evacuation. All electrodes have to be machined with high accuracy and aligned on a common axis, and the electrostatic lens as a whole must be mechanically rigid so that individual electrodes do not move with respect to each other when the electron optical system is assembled, transferred, or during bake-out. One of the chief advantages of electrostatics is that the volume of the electrodes does not contribute to the field distribution. In principle, any metal can fulfill the requirements for producing an equipotential, although in practice the selection is much more restricted. This is much simpler than in the case of magnetic lenses, where the field shape depends on lens excitation because of saturation, where there are hysteresis effects, and where the inhomogeneity of the material produces parasitic leakage fields. For electron lenses and high-accuracy ion beam systems we have to shield the system from external magnetic fields by mu-metal or active shielding.

The encyclopedic two-volume tables of von Ardenne (1956, 1964) summarized much of the first 30 years of knowledge of particle optics instrumentation and provided a complete set of data on materials and techniques used to produce lenses for electron microscopy at that time; they are still very useful, although since then many new materials and technologies have been developed and new data were obtained for a wider range of lens geometries.

Lenses used in spectroscopy (electron impact studies) are usually of simpler design because they work at low voltages, so that sharp edges do not lead to electrical breakdown. They have to work over quite a wide range of voltages, and thus shaped electrodes do not lead to better overall performance. Ease of fabrication and assembly is therefore preferred and electrode geometry is based on cylinders or thin apertures mostly of equal diameter (Read, 1979). To achieve uniform surface potentials and reduced electron reflections, it is common to coat electrodes seen by electrons with colloidal graphite in isopropanol, preferably applied as an aerosol with a dry air propellant and followed by a mild baking. For electrodes, soft materials such as Al and Cu are often used, but as they are more demanding on assembly and susceptible to deformation and surface damage, it is often better to use other materials (see Section 5.3.2). Electrostatic shields around the critical regions (scattering or target regions) must be used to guarantee not only potential uniformity but also as a barrier against stray electrons or other particles. Insulation at low voltages is not critical, and it is normally achieved by means of precision glass or ruby spheres and alumina rods; machinable glass-ceramic is a solution when special shapes are needed, but it is necessary to screen all insulators near the beam. Considerable care must be taken in the choice of nuts, bolts, and screws so that they can be annealed and are nonmagnetic (materials like Cu–Be are preferred). The design of electron spectrometers was reviewed by Roy (1990) and Erskine (1995); the review also includes a discussion of the relevant electron optics (see also King, 1995).

The main obstacle in the design of high performance electrostatic lenses is caused by problems connected with electrical insulation and voltage breakdown. We have to distinguish between the problems related to discharge in the vacuum gap, initiated by field emission from electrodes, and to surface breakdown in insulators; these will be discussed in the following two sections, including the proper selection of suitable materials and conditioning. Often, *rule of thumb* knowledge and experience are more valuable than theory.

For ion optical systems, care must be taken that lens electrodes are not exposed to ions. Damage to the electrode surface caused by deposition or sputtering of ions may create a rough surface and hence possible new field-emission sites. Exchangeable apertures must be incorporated in the system to limit the beam size. For electrons, a different argument applies: under the impact of high-energy electrons contaminants may decompose; in particular, oil contaminants (from pumps) produce an irregular insulating film. In bad vacuum, positive ions incident on an insulating contamination film can polarize it and enhance field emission. It is again necessary to shield the insulators used in the lens from a possible direct impact of the beam.

### 5.3.2 ELECTRICAL INSULATION AND BREAKDOWN IN THE VACUUM GAP

The main contribution to the performance of the lens comes from the parts of the electrode close to the optical axis. Maximum field strength needs to be produced only in the vicinity of the vacuum gap between the electrodes in the axial region. In the peripheral region, where the electrodes are supported and insulated, a larger spacing can be allowed. In the vacuum gap electron field emission plays an important role. Although we need fields  $\sim 3 \times 10^9$  V/m to obtain field emission from metal tips, it is in practice impossible to reach such a high value for a macroscopic electrode. The field values that can be used in practice are between 10 and 20 kV/mm ( $10\text{--}20 \times 10^6$  V/m) and depend on the choice of electrode material and the surface preparation of the electrodes. Higher values are seldom used; for example, Aihara et al. (1988) investigated field strength well above 20 kV/mm, and Tsumagari et al. (1988) allowed 37 kV/mm. Möllenstedt and Laauser (1980) studied the high-voltage stability of steel electrodes in a UHV of  $2 \times 10^{-8}$  Pa, and they could use a separation of  $\sim 50\text{--}100$   $\mu\text{m}$  for electrodes at 10–20 kV. In their study, it was possible to keep the field strength up to 140 kV/mm; no surface damage was observed for electrodes made of hardened roller-bearing steel with  $\sim 1\%$  Cr, carefully polished and cleaned—an improvement by a factor of 10 over an electrostatic microscope operating at  $10^{-3}$  Pa. The discharge current had to be limited to a few nanoamperes by a resistor of 150 M $\Omega$  in series with the high-voltage supply.

As electrode materials, nonmagnetic (304 or 310) stainless steel, Ti, or a titanium alloy such as IMI Ti-318, Mo, or a Cu-Be alloy are preferred. Hard materials are required to minimize damage due to field-emission events, which can cause local heating. The material should not only be physically hard but also resistant to corrosion effect such as oxidation. For example, steel and Ti are covered by a strong insulating stable oxide film 5 nm thick that consequently limits the switch-on of the field emission. The materials must be acceptable for UHV technology (nongassy) and have good machinability; for modern UHV microscopes titanium is often preferred—see Adamec et al. (1998) and Hartel et al. (2002). Electrode edges must be rounded to avoid local enhancement of field. An important step in the production of electrodes is the polishing of electrodes; mechanical polishing should be replaced by chemical or electrochemical polishing to get reliable surface operation, but great care must be taken during electrochemical polishing to avoid creating asperities that can lead to field emission. Careful ultrasonic cleaning and several rinsing steps should be followed by the assembly of the lens in a clean, dust-free environment.

Before reaching actual high-voltage breakdown, the discharge between electrodes is often started by so-called microdischarges. Microdischarges are bursts of current flowing between electrodes in vacuum; they last from 0.010 to 10 ms and collapse by sputtering or deabsorption of the contributing atoms during the discharge; therefore, microdischarges only seldom end in a flashover. Stable prebreakdown currents from *cold* emission sites are possible at fields as low as 1 kV/mm: we have to consider that the field is enhanced by a factor 10–1000 at local irregularities. Field-emission sites on the electrodes are created during manufacturing; the main causes are metallic *whiskers* and particulate microstructure or contamination embedded in and attached to electrode surfaces, for example, due to the polishing process and dust accumulated during assembly. For ion lenses, the electrode surface can be damaged by the deposition or sputtering of ions that may create roughness. Localized contaminants enhance the field intensity on electrodes. Microparticles can also be generated by sparks. The emission process is accompanied by the subsequent generation of x-rays.

Conditioning of electrodes is necessary to improve the lens performance and to obtain the most stable insulating strength. It is a process used to quench as many as possible of the sources of pre-breakdown current and primary microparticle events so that the number of potential hazards to the stability of the gap is reduced to a level where the operational performance is acceptable. Several types of electrode conditioning are used. *Current conditioning* means increasing the voltage in small steps and allowing the prebreakdown currents to stabilize. Normally, we try to reach  $\sim 25\%$  more than the intended operating voltage. During the current conditioning, micropromtrusions lose particles,

adsorbed gases are removed, or the protrusions are thermally blunted and their characteristics are modified so that they are no longer harmful. *Spark conditioning* with fast sparks (nanosecond high-voltage pulses) can effectively remove microprotrusions, for example, after a complete collapse of voltage by the formation of an arc. The energy dissipated must be low to prevent irreversible damage. A familiar preparative technique used to generate clean surfaces and remove surface contaminants is *glow discharge conditioning*. At a pressure of 0.1–1 Pa of a suitable inert gas, such as Ar, a low-voltage alternating current glow discharge is developed. If this is applied too frequently or too long, the surface can be adversely affected. Similarly, *gas conditioning*, for example, Ar at 0.1 Pa with high voltage on, allows low currents of a few tens of microamperes to run for some time to remove the field-emission sites by bombarding them away with positive ions. These processes do not remove the microparticles, which can be eliminated with a combination of oxidization (heating in air) followed by magnetron sputtering with Ar and vacuum baking. Damage and deconditioning can occur as a consequence of vacuum problems or outgassing during minor discharges. In the worst case, erosion and anode melting is caused by the field-emitted electrons, which may produce sputtered material on electrodes and insulators. Of course, not all these methods can easily be used in a given device.

More information, details, and limits of each of the conditioning techniques are given by Latham (1995). Previous experience with the design of electrostatic lenses was summarized by Hanszen and Lauer (1967). Stability of a focused ion beam system, the effect of electrode materials, and conditioning procedures are discussed by Aihara et al. (1988). Much experience has recently been obtained in electrostatic accelerators (Joy, 1990) and with superconducting radio-frequency cavities used in particle-accelerating and storage systems (Latham and Xu, 1990; Latham, 1995). Field strengths up to 50–60 kV/mm can be used with bulk accelerator electrodes if copper is replaced by high-purity Nb, mechanical polishing is replaced by chemical or electrochemical polishing techniques; also ultrasonic cleaning and several rinsing steps must be applied. The last steps consist of *in situ* procedures—high-temperature annealing to 1200°C or melting the electrode surface by high-current arc conditioning, followed by gas conditioning in an atmosphere of *conditioning gas* such as He at  $10^{-3}$  to  $10^{-2}$  Pa under field for 10–20 min.

### 5.3.3 INSULATOR MATERIALS AND DESIGN

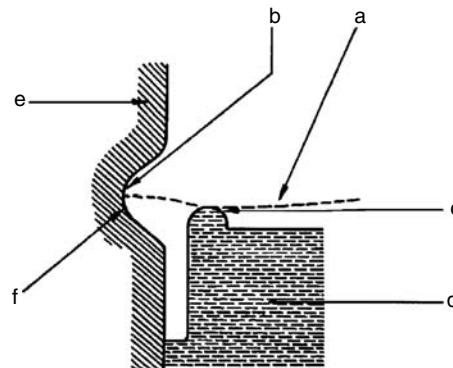
The electrodes are supported and insulated in the peripheral regions of the lens, where a spacing larger than that used in the vacuum gap can be allowed. In larger electrostatic accelerators, insulator gradients 1.5–3 kV/mm can be reached (Hyder, 1990; Joy, 1990), but the terminal voltage (15–35 MV) is much higher. This is about three times less than the value considered *safe* in electron lenses for a lens voltage of tens of kilovolts. The volume strength of the insulator materials is in the range of 40–100 kV/mm. The insulation strength along the surface of solid insulators in vacuum is invariably lower than either their volume strength or the strength of the vacuum gap. The surface gradient is lowered by increased insulator length and by corrugations. For high (MV) acceleration tubes and guns of MV microscopes (Ohye et al., 1993), the voltage insulation length has to be subdivided into multiple short ring-like sections separated by metallic electrodes between which the voltage is distributed.

Secondary emission processes dominate the surface breakdown of insulators. Often the initiating mechanism lies at the triple junction of solid dielectric, vacuum, and electrode, mostly at the negative electrode (cathode). A major effort must be directed to reducing the electric field intensity in the region of cathode–insulator junction. Any vacuum gap in this region will experience a field intensification approximately  $\epsilon_r$  times ( $\epsilon_r$  being the relative permittivity of the dielectric). The remedy is to seal the dielectric to the metal electrode with a medium whose dielectric constant and conductivity are higher than those of the insulator itself: this can be accomplished by a soft metal film such as aluminum, lead, or indium; even a thermoplastic or thermosetting dielectric may be applicable. From experimental investigation (Shannon et al., 1965), it follows that it is possible to sustain voltage gradients of 8 kV/mm without flashover after suitable conditioning under electric stress with (1) a void-free contact between insulator and electrode (e.g., with polyvinyl acetate cement) and

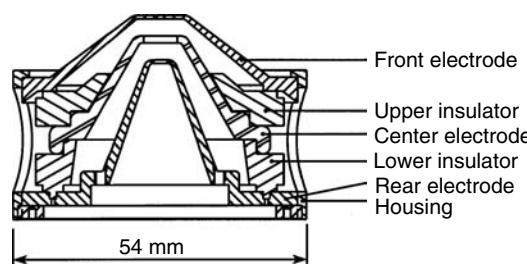
(2) an insulator and electrode contact which reduces the field at the insulator-negative electrode junction and provides barriers along the insulator surface.

The field strength at the junction can be reduced by a suitable geometric design (Latham, 1995) and by using an insulator material of low permittivity (Aihara et al., 1988). Beyond this, a positive charge on the insulator surface caused by secondary emission of electrons must be avoided. By choosing a proper angle between insulator surface and field direction, it is possible to obtain an uncharged or only negatively charged insulator surface near the junction. If such an angle cannot be realized, the distance between the negative electrode and the point of impact of electrons on the insulator (see Figure 5.21) should be so large that a sufficiently high potential difference occurs. Then, the energy of the electrons striking the insulator is high enough for the secondary emission factor to fall below unity (Boersch et al., 1963). As mentioned before, a corrugated surface increases the surface length of the insulator and helps to suppress the charge transport along the surface. A negative charge on the insulator raises the potential difference near the positive electrode and may lead to the stripping of electrons from the insulator surface. This effect can be suppressed by suitable geometrical design of the insulator and the positive electrode. An additional undercut at the positive electrode helps to stop any surface currents (Rempfer, 1985)—see Figure 5.22.

Microdischarges are induced by cascades of positive and negative ions and electrons that release each other by impact from contamination layers on the electrodes (Boersch et al., 1961; 1963): they are extinguished if the voltage is lowered by using a resistor in the feed line: a  $50\text{-M}\Omega$  resistance



**FIGURE 5.21** The arrangement of the electrode and insulator in the vicinity of the triple junction between negative electrode, insulator, and vacuum is important for the stable performance of the high voltage. The rounded shape (f) in the negative electrode (e) reduces the field strength and the bulge (c) on the insulator (d) prevents electrons from being stripped off from the negatively charged insulator. Field-emitted electrons (a) arriving at the insulator are sufficiently fast for the secondary emission coefficients to fall below 1. (From Gribi, M. et al., 1959. *Optik*, 16(2), 72.)



**FIGURE 5.22** Cross section of an objective lens used in a photoemission microscope designed by Rempfer; the swept-back front electrode allows direct illumination with UV light. (From Rempfer, G. F., 1985. *J. Appl. Phys.*, 57(7), 2385. With permission.)

shortens the voltage dip to  $10^{-4}$  s. When the discharge current increases, electrode material is evaporated and spark-overs with complete voltage breakdown can occur. The evaporation of the electrode material must be reduced with the aid of a protective resistance, and the charge transported by one spark is limited by the capacitance of the lens system (thus it should be held as small as possible). In a number of experimental studies, Boersch et al. (1963) described a model for discharge along an insulating surface by electrons hopping along the surface. The possibility of electron multiplication must therefore be taken into account.

Good materials for the insulator between the high-voltage supply line and lens chamber are nonporous ceramics, glass and glass-ceramics, and synthetics such as *acryl glass* and *epoxy resin*, which can be worked easily on a lathe. For bulk insulators in accelerators, borosilicate glass or high-alumina ceramics are used. The advantage of glass is its transparency and lower dielectric constant. Insulators made of rexolite (crosslinked polystyrene) work well at 40 kV (Rempfer, 1985). Modern machinable glass-ceramic materials such as MACOR<sup>TM</sup> are often employed.

The most common means of insulating in electrostatic optics is to use ceramic rods, ceramic rings, and ceramic, sapphire, or ruby balls. The insulation based on rods, used in particular for lower voltages, provides quite an economic design: three to six rods are used as a rule, arranged at a regular angle; individual electrodes must have precise holes, and the rods are fixed with screws to outer electrodes. Intermediate electrodes are either fixed to the rods or, for higher precision, auxiliary balls or insulating spaces are used. Disadvantages are the loss of accuracy for longer systems, difficulties in setting the distances between the electrodes precisely, and the fragility of the rods.

The design based on insulating balls allows high precision to be reached and ensures a good mechanical stiffness of the system. Precisely machined ceramic, ruby, or sapphire balls are available in sizes from 1 to 12 mm with accuracy better than 1  $\mu\text{m}$ , and they are mechanically and chemically stable. Two electrodes are separated with three balls placed at the vertices of an equilateral triangle. If the electrode needs insulation on both sides, the other set of balls is rotated by 30°. The shape of the sphere is also quite advantageous as long as the angle between the conductor and the ball is less than 30°; that is, the hole in which the insulator rests has a diameter smaller than that of the ball radius or the size of the gap is about equal to the sphere radius. The groups of electrodes are usually fixed with four screws. By using precise numerically controlled machines for defining the ball positions, it is possible to reach high accuracy, better than 0.01 mm. An example of a design based on balls is given in Figure 5.20. In practice, the design is usually overdetermined and thus it allows small tolerances of manufacturing in angle and radius (see, e.g., Hartel et al., 2002).

Insulation with the help of ceramic rings or lathe-machined structures leads to more precision in interelectrode spacing, better setup, and higher stiffness. The disadvantage is that the number of commercially available shapes and ceramic models is relatively small. MACOR can be machined into complicated shapes inexpensively with readily available tools, is nonporous, chemically stable, and UHV compatible. Attention must be given to the bonding process between the electrodes and the insulator. For more demanding applications, the designer should try to reduce the field at the triple junction and introduce on the insulator surface corrugations to stop charge transport along the surface.

When considering insulation in cables, any damage usually propagates along the surface, and cleaning may remove this. To avoid the problems, proper design dimensions of the high-voltage feedthrough are required.

### 5.3.4 MANUFACTURING AND ALIGNMENT ACCURACY

In real lenses, ellipticity of the bores in electrodes is often observed, and misalignments and tilts may occur between electrodes. This introduces asymmetric fields and generates additional parasitic aberrations (see Chapter 6, which gives a survey of the literature, and Hawkes and Kasper (1989) for a more complete treatment). The review of Yavor (1993) also covers more general systems, such as prisms and spectrometers.

The fields can be computed by a perturbation method based on Sturrock's principle (Sturrock, 1951). Small misalignment and tilt produce additional weak deflection fields, whereas the elliptical defect produces a weak quadrupole field. They can be calculated with FDM (Janse, 1971; Munro, 1988) or FEM (Edcombe, 1991). The interest in the computation of parasitic effects grew recently in particular in connection with ion beam lithography, where electrostatic lenses are used in combination with deflectors. The treatment of parasitic fields and aberration formulas is given by Liu and Zhu (1990), Kurihara (1990), and Tsumagari et al. (1988, 1991).

Ellipticity is the best-known parasitic effect—and, fortunately, it can easily be removed by a stigmator. This effect is mostly caused by unround bores. The validity of Sturrock's principle for the evaluation of unround holes has been verified by an independent check with a three-dimensional computed potential (Rouse, 1994); for ellipticity of 5%, the agreement with theory was better than 1% (notice that typically ellipticity below 0.1% is required). No three-dimensional modeling of other effects such as tilt and electrode displacement has so far been published.

A number of parasitic effects are time dependent—for example, charging of dust particles and contamination layers, temperature drifts; they are thus difficult to distinguish and characterize (see Latham, 1995). Surface charging may develop in the column because of impact of primary or secondary particles. There are hardly any estimates of these effects, because they depend on too many parameters such as the composition of the residual vacuum, the quality, thickness, and composition of surfaces involved, and particle energy and angle. A model permitting the calculation of position errors of electron beams in an electron beam lithography system due to surface charging in the beam vicinity has been developed by Langner (1990), who also estimated the space charge effects due to ions generated by the beam and secondary electrons.

### 5.3.5 ENVIRONMENTAL AND SYSTEM CONSIDERATIONS

Formerly, electrostatic lenses were operated in a rather poor oil vacuum, typically  $10^{-3}$  to  $10^{-2}$  Pa. Oil vapors accumulate and under the impact of charged particles form contamination layers. In the initial stage of their formation they actually help to suppress the field-emission sites on electrodes. When these layers grow thicker and become polarized by the ion impact, field emission is enhanced by ions released from these layers as they crack. A number of studies done by Boersch et al. (1961, 1963, 1966) involve the investigation of electron hopping on insulators, research on the timescale of microdischarges and separation of the role of ions. Often the high-voltage breakdown reliability could be improved by allowing a slight flow of gas around the insulators; this was observed by Gribi et al. (1959) who applied a cold discharge gun where some input of gas is actually needed to keep up the discharge.

For reliable long-term operation, the total vacuum pressure and, in particular, the partial pressure of hydrocarbons must be kept low. This reduces the possibility of growth of contaminating layers where surface charges can develop. Heating of inner surfaces in the close vicinity of the beam (e.g., apertures) is essential to delay the growth of these layers. Extended area surface charging by secondary electrons, coupled with the growth of insulating layers, causes long-term position drifts; other types of surface and space charges cause transient position errors.

With the use of UHV technology, the formation of contamination is decreased. For the removal of contamination the established procedure is bake-out to 150 or 250°C. The systems become more stable, and electrical breakdown problems become less critical (Latham, 1995).

The focal length of electrostatic lenses is insensitive to voltage variation if all the voltages are derived from a common high-voltage supply. For this reason, unipotential design was applied to the whole electrostatic microscope, where all voltages were derived from a single high-voltage supply. Focusing is often done by the change of specimen position (Rempfer et al., 1972; Delong, 1992); although this puts some extra demands on specimen manipulation, voltage stabilities required are much lower. For fine focusing, the voltage on the central electrode can be changed with the help of a battery or a resistor chain containing a potentiometer. Good performance of an electrostatic TEM

was obtained with a voltage stability of only 0.1% (Hahn, 1955). However, microdischarges affect the function of an electrostatic lens system because of the local voltage drop on one of the lenses. The high-voltage system of the EM8 was the most advanced (Schluge, 1954). Nowadays, voltage fluctuations  $<10^{-5}$  are easy to achieve with much smaller power supplies and at a lower cost.

Ideally, electrostatic lenses do not consume or dissipate any power and there is no need for water-cooling. The lower weight of the column and lower resolution requirements make vibration problems less severe. Operation of electrostatic lenses and removal of astigmatism is more straightforward as the beam does not rotate in an electrostatic lens. In connection with a data acquisition system, slow drifts can also be removed (Rempfer et al., 1972).

Most parts of an electrostatic system are made of nonmagnetic materials, and thus the only problem is to shield the beam from external magnetic fields. This is mostly achieved by placing single or double mu-metal shield around the column.

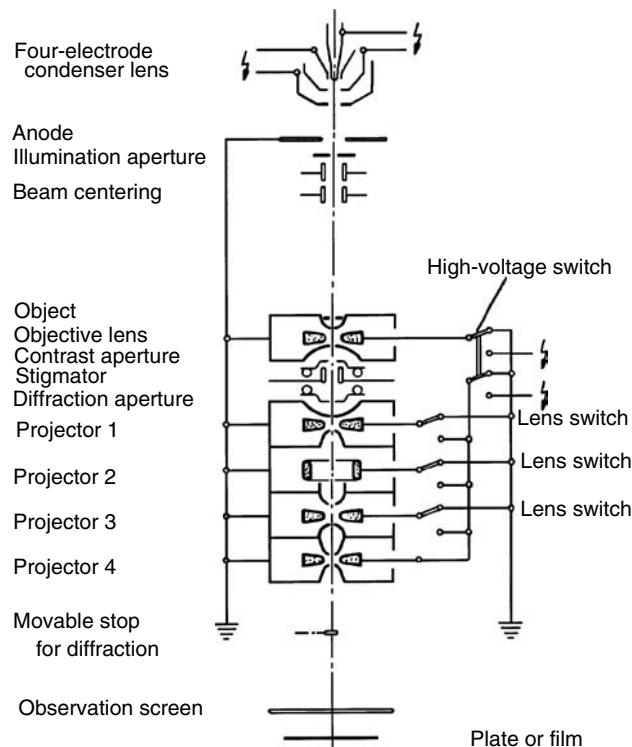
## 5.4 EXAMPLES

### 5.4.1 ELECTROSTATIC TRANSMISSION ELECTRON MICROSCOPES

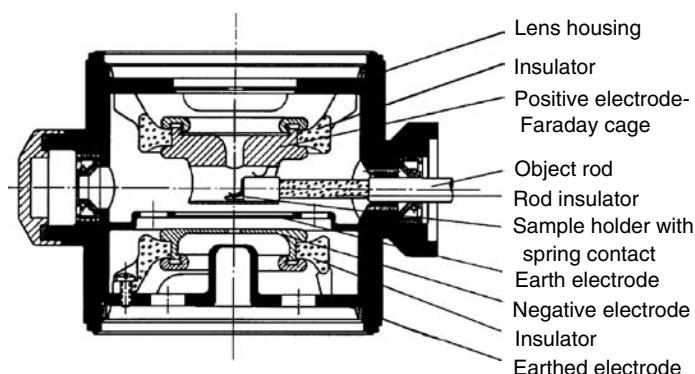
In the early days, certain advantages of electrostatic lenses were emphasized. It was easier to produce high-quality electrostatic lenses than magnetic (especially objective) lenses. Moreover, the use of electrostatic lenses did not conflict with the patent rights of Siemens and of Knoll and Ruska. All voltages could be derived from a single high-voltage supply. Most lenses in the electron optical system were used as einzel lenses—unipotential lenses with the central electrode held at the cathode potential. Thus, their focal distance was fixed, and only a limited number of fixed magnifications were available. On the other hand, as the focal length does not change if all voltages change by the same ratio, the high-voltage supplies need not be stabilized below 0.1% for satisfactory performance. The power consumption of electrostatic lenses is negligible, and there is no need for water-cooling such as in the case of magnetic lenses. The beam in electrostatic lens does not rotate, and so specimen movement, astigmatism correction, and taking of stereo pair images are straightforward.

An overview of early electrostatic microscopes is given by Cosslett (1951). He also stressed disadvantages such as greater susceptibility to contamination, slow accretion of insulating deposits under the electron beam, inherently higher aberrations of electrostatic lenses, the use of einzel lenses making focusing and magnification changes difficult, and more complicated mechanical solution of the specimen stage because of focusing by moving the specimen along the axis. Therefore, microscopes were designed for routine performance, and the maximum resolution achieved was only  $\sim 2$  nm.

Later, several firms produced new models, working mostly at 50 keV beam energy; for this energy the film material used for recording the images provided the optimum performance (Möllenstedt, 1955). Zeiss Jena (Hahn, 1955) designed a microscope Elmi-D, and sold over 100 units. The microscope used four projector lenses. The three intermediate einzel lenses could be switched on and off, and thus seven magnifications between 2,000 and 30,000 times and a diffraction regime were available; with variable voltage on the central electrode of the lens preceding the last projector lens an *optical* magnification range of 800–1200 times was also available (see Figure 5.23). The EM8 model of Zeiss Oberkochen was also sold to about 60 laboratories (Schluge, 1954). The improved version shown in Figure 5.1, inclined for operational convenience, was unfortunately never commercialized (Mahl et al., 1956). It also tried to overcome the inflexibility of the einzel lens concept by employing an adjustable four-electrode projector lens (*Regellinse*) (Rang and Weitsch, 1956). Two fixed intermediate voltages could be connected to the central electrodes, and the microscope magnification could be changed in the ratio 10:3:1. Of particular interest was the intermediate accelerator (*Zwischenbeschleuniger*) of Möllenstedt (1955) shown in Figure 5.24. The object is placed in the center of a Faraday cage that can be connected to a high-voltage supply of the same magnitude and opposite polarity to that of the rest of the lenses; thus, the beam passing through the sample has effectively twice the beam energy used in the rest of the column; this solution allowed the maximum energy in an electrostatic microscope. The highest voltage



**FIGURE 5.23** Schematic arrangement of Elmi-D2 electrostatic transmission electron microscope (TEM) produced by VEB Carl Zeiss Jena, German Democratic Republic. A tetrode condenser lens is used for illumination. The voltages on each of the three intermediate projector lenses could be switched on and off, so that magnification in six stages between 2,000 and 30,000 plus diffraction at a fixed camera length were available. Magnification in the light-optical range 800–1200 was possible with projector 3 excited with only a part of the voltage on the cathode.  $C_s$  of the objective lens was measured as 50 mm. (From Hahn, E., 1955. *Jenaer Jahrb.*, I, 42. With permission.)



**FIGURE 5.24** An ingenious idea about how to reach 100 keV beam energy in an electrostatic objective lens where insulators withstand only 50 kV was proposed by Möllenstedt. The lens housing and outer electrodes were earthed, and the sample was placed in the middle of a Faraday cage held at a positive potential of 50 kV, whereas the focusing electrode was held at the negative voltage of the same value. The electrons pass through the sample with twice the maximum electron energy used in the rest of the column. (From Möllenstedt, G., 1955. *Optik*, 12(10), 451. With permission.)

in a single step, 70 kV, was used in the microscope produced by Trüb, Täuber & Cie AG (Gribi et al., 1959). The electrostatic objective lens is shown in Figure 5.2. The microscope used a magnetic condenser lens and a magnetic intermediate projective lens. Two technical drawbacks of the microscope reported by Mühlethaler (1990) resulted in the termination of its production: the cold cathode did not provide a sufficiently bright and homogeneous spot at high magnification and the objective lens was not equipped with a stigmator. The last attempt to produce an electrostatic transmission microscope was by Elektros Corporation in the United States (Rempfer et al., 1972). Despite many advanced technological solutions and elegant design, compactness, and mobility of the microscope, only a limited number of units were produced.

Magnetic objective lenses of high quality can be designed as a result of significant technological progress in producing, processing, and machining magnetic materials, summarized in Chapter 4. Even if the microscopes are often used in applications that could easily be covered by electrostatic microscopes at lower cost, the difficulties in reaching higher magnifications and better resolution killed electrostatic microscopes off in the late 1960s, when most of the new materials, UHV technology, and computer program had only begun to have an impact on microscope design and performance. Only the designs of LEEM and PEEM discussed in Section 5.4.3 profited indirectly from the electrostatic TEM development.

#### 5.4.2 LOW-VOLTAGE TRANSMISSION ELECTRON MICROSCOPY

The contrast in thin biological samples increases in a TEM as the beam voltage decreases. Wilska (1964) used a magnetic TEM at voltages between 6 and 10 kV; his research set limits on low-voltage operation and showed that the microscope needs an electrostatic filter lens to eliminate most inelastically scattered electrons (Wilska et al., 1970). Low-voltage beams produced by a thermionic cathode do not provide enough beam intensity; the reduced beam brightness is further decreased by space charge effects. Recently, Delong (1992) (see also Delong et al., 2000) designed another TEM working at 5 keV: sufficient intensity of the illumination beam is secured with a Schottky-emission electron gun. This microscope uses a miniature magnetic lens, excited by cobalt-samarium permanent magnets, that fulfills the role of condenser and condenser-objective lens with axial aberration coefficients below 1 mm (see Adamec et al., 1995). Electron optical magnification of 300 times is provided by an electrostatic lens system, easy to manufacture for 5 kV voltage, and further magnification is provided by an optical microscope; thus, the electron optical part is unusually small.

Heinemann and Möllenstedt (1967) used the idea employed in the intermediate accelerator of Figure 5.24 to decelerate the electrons so that they pass a thin object at 1–10 keV energy. The axial aberration coefficients related to the specimen plane are quite low, ~0.5 mm for 1 keV, and  $C_s = 5 \text{ mm}$  and  $C_C = 1.8 \text{ mm}$  at 5 keV, thus allowing, in principle, 1 nm resolution. The beam in the rest of the column is held at the operating voltage of 50 kV, and thus it is not susceptible to external disturbances. More details on the advantages of beam deceleration are discussed in Section 5.4.4 for the optics of low-voltage SEMs.

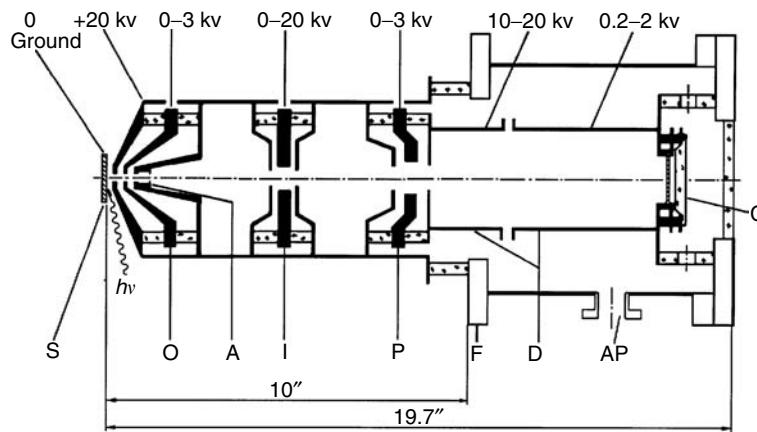
#### 5.4.3 LOW-ENERGY AND PHOTOEMISSION ELECTRON MICROSCOPES

Primary electrons slowed down in front of the specimen, scattered at the specimen at low energy, or even reflected before they reach the specimen, and electrons emitted from the sample by thermoemission, photoemission, or various other secondary emission processes allow imaging of the surface structure that is sensitive to composition of the uppermost layers of the sample. As shown by Delong and Drahoš (1971), in the back focal plane of the cathode lens, a (low-energy) diffraction pattern of constant size is produced, allowing the use of various diffraction techniques, including selected area diffraction. A historical overview up to 1991 is provided by Griffith and Engel (1991); reviews of the relevant instrumental problems are given by Veneklasen (1992) and Bauer (1994), most recently by Bauer (2007) and Feng and Scholl (2007).

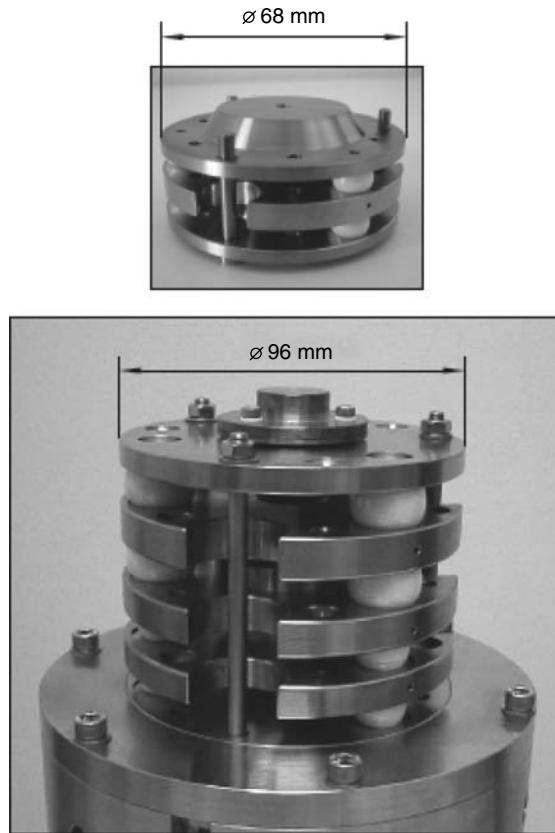
The critical electron optical element is the cathode lens, discussed in Section 5.2.5. The acceleration field of the cathode lens with length  $l$  is a divergent lens that places the image plane at a distance  $l/3$  and the focal plane at  $3l$  behind the specimen. The cathode lens is followed by an electrostatic or a magnetic electron lens; this objective lens must have quite a large focal length. For this purpose electrostatic lenses provide electron optical properties comparable with those of a magnetic lens, see Chmelík et al. (1989). An example of the electrostatic lens designed by Rempfer is given in Figure 5.22: the swept-back entrance electrode allows direct specimen illumination by ultraviolet light. Another design example has already been shown in Figure 5.20. Recent analysis of the objective lens with a SIMION program is given by Feng et al. (2002) and Bernheim (2006).

In a LEEM we have to illuminate the specimen with an auxiliary electron beam; therefore, a beam splitter is an important element, immediately following the objective lens. It is either a magnetic prism (Rose, 1978) or an imaging Wien filter (Rose, 1987; Tsuno, 1993). To reduce the problems with high-voltage insulation and also to reduce the cost of high-voltage supplies, the microscopes are designed for a convenient compromise value of 10–20 keV. For imaging magnetic lenses can be used (Veneklasen, 1992): because of UHV design needed for surface studies they have to be placed outside vacuum, and so their inner diameter is limited by the outer diameter of the liner tube; thus they do not operate at low focal length. Therefore, an electrostatic system can be conveniently used as it can be placed inside a UHV vessel and operated at optimum lens strength. A small size LEEM designed by Adamec et al. (1998) can be mounted on a standard 6 in. flange and fitted as an attachment to a UHV vessel. Small dimensions for this system are achieved by the use of 5 or 10 keV beam electrostatic optics, and a magnetic prism with deflection of only 20° between the primary and the imaging beams.

In a photoemission microscope the primary electron beam is not needed. An all-electrostatic PEEM is described by Rempfer et al. (1991), Skoczylas et al. (1991), Tromp and Reuter (1991), Marx et al. (1994), and Dunham et al. (1994). An example of a photoemission microscope that can be placed on a UHV flange and thus also combined with other surface analytical techniques is shown in Figure 5.25 (Engel et al., 1991); the image is recorded on a multichannel plate, where the beam arrives at an optimum low energy, an interesting design feature. In addition, the sample is held at the earth potential and the voltages on the outer electrodes of the lenses are *floating*. The PEEM described by Latham (1995) was also used in the field-emission mode for the study of high-voltage prebreakdown effects.



**FIGURE 5.25** Schematic arrangement of an older electrostatic photoemission microscope that can be added to an ultrahigh vacuum system. Object S is held at ground potential, objective (cathode) lens O is equipped with an aperture A. Intermediate lens I and projector lens P together with a decelerating tube lens D image the sample onto a microchannel plate image intensifier C. The whole design fits onto an 8 in. flange F; AP is a port for an auxiliary pump. (From Engel, W. et al., *Ultramicroscopy*, 36(1-3), 148–153, 1991. With permission.)



**FIGURE 5.26** An example of a modern electrostatic lens design, field lens, and electron mirror of SMART. The insulation and precise position of electrodes is based on precision ceramic spheres. (From Hartel et al., *Adv. Imag. Electron Phys.*, 120, 41–133, 2002. With permission.)

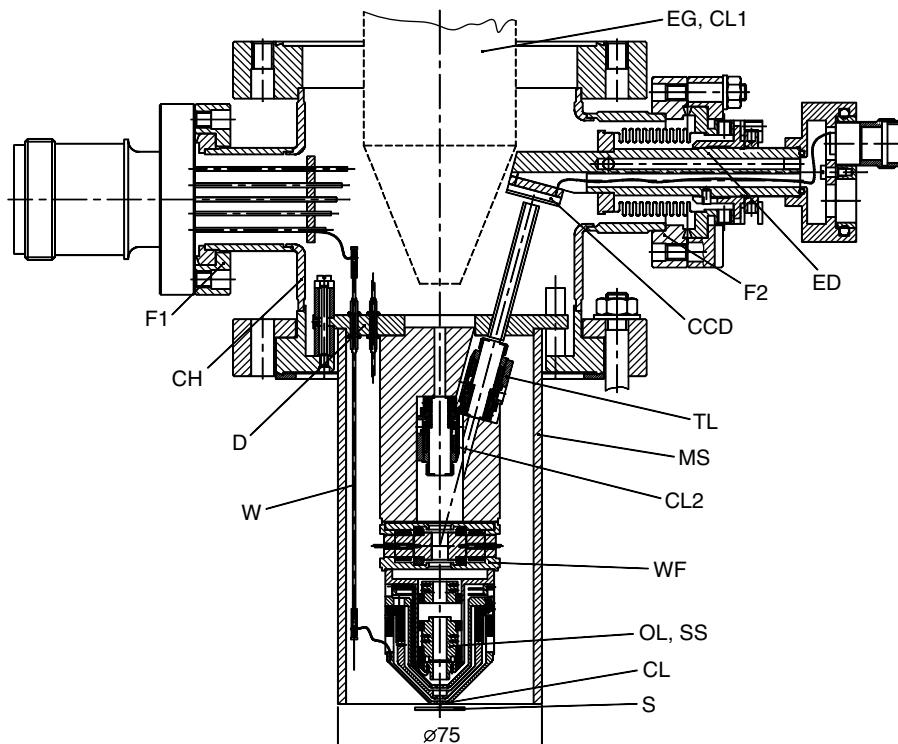
A number of projects are currently concerned with the design of a PEEM or LEEM that should reach optimum performance by the correction of axial aberrations. This correction, which can be done with the help of an electrostatic mirror, was discussed in Section 5.2.7.5. (Skoczyłas et al., 1991; Rose and Preikszas, 1994). The corrected optics is the basis of an ambitious SMART project (Fink et al., 1997) (schematically shown in Figure 6.14 of Chapter 6). A thorough analysis of electron optical problems of the microscope, its design, testing, and a discussion including details of manufacturing and tolerancing of the critical elements of the system, are summarized in a review by Hartel et al. (2002). Figure 5.26 shows two electrostatic elements used in the SMART system, an intermediate lens preceding the mirror corrector, and the mirror corrector itself.

#### 5.4.4 SCANNING ELECTRON MICROSCOPES (COMPOUND AND RETARDING LENSES)

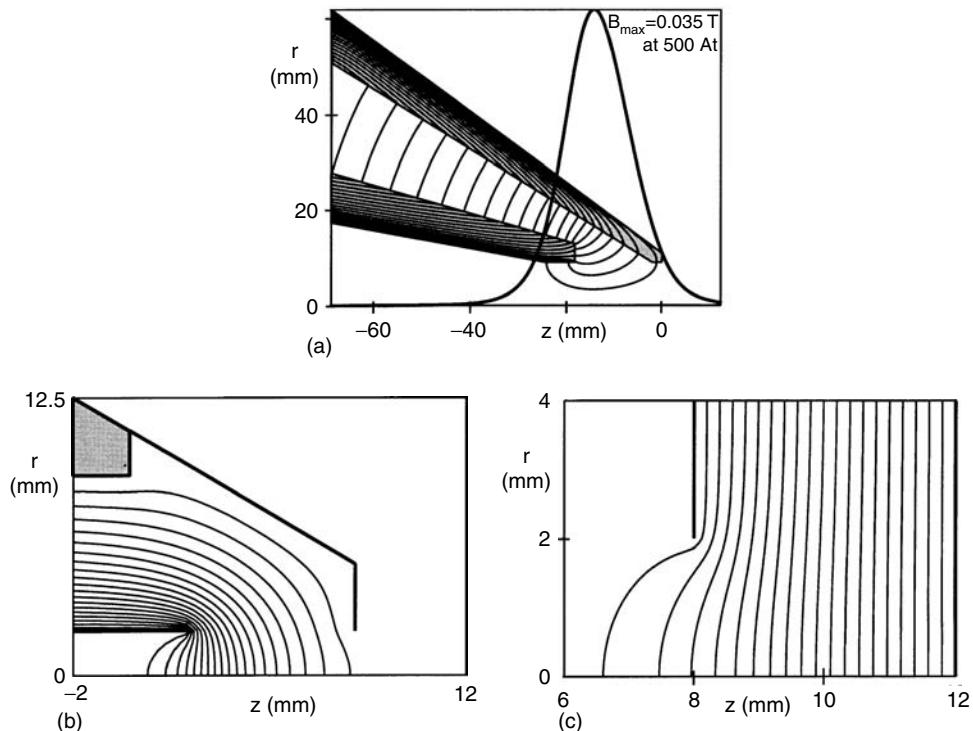
The first SEMs made outside Germany used electrostatic lenses. McMullan (1953) applied both an electrostatic condenser and an objective lens, whereas Zworykin et al. (1942) even employed an immersion objective lens. All recent commercial SEMs use magnetic lenses. It is now even feasible to produce positron microprobes that can be combined with an electron beam of the same energy to adjust the optical system (Seijbel et al., 1992); in magnetic lenses the two beams rotate in different directions, but this effect can be compensated for by the computer control. With the electrostatic system, the polarity of lenses must be changed. If decay time measurements have to be performed, it is necessary to use electrostatic optics (Roach et al., 1995).

Recently, scanning electron devices with low-energy electrons impinging on the specimen became interesting for the electronics industry because of lower damage and reduced charging of electronic components (Reimer, 1993). If the beam is held at low energy in the whole column, space charge effects decrease the beam brightness significantly (Barth et al., 1990). Therefore the beam must be held at higher energy (above 10 keV), and slowed in front of the sample. The electron optics of the two basic solutions was discussed in Section 5.2.7.4: deceleration with an immersion lens just before the specimen or using the specimen as a *cathode*. The first approach with deceleration was applied by Frosien et al. (1989, 1995); its main advantage is in the fact that the field on the sample is much lower than with the cathode lens. The second approach based on a cathode lens was discussed by Munro et al. (1988), Meisburger et al. (1992), Lenc and Müllerová (1992), Hordon et al. (1993), and reviewed by Müllerová and Frank (2003), see also Paden and Nixon (1968). The disadvantage of the cathode lens is that the electric field on the specimen is quite high, which is not acceptable in some applications, in particular in the study of electronic components. For very low energies on the sample only the cathode lens is applicable, because the immersion electrostatic lens above the sample becomes too strong for high deceleration ratios.

Collected secondary particles follow a path similar to that of the primary electron beam and they have to be separated from this beam; the design of the detection system must thus be included in the design of the lens system. Meisburger et al. (1992) used for this purpose a Wien filter that deflects most signal electrons toward a detector. A more ambitious system design is shown in Figure 5.27



**FIGURE 5.27** Critical parts of the scanning low-energy electron microscope (LEEM) system. The primary beam electrons (PE) are focused by an auxiliary condenser lens (CL) into the center of the Wien filter (WF); the objective lens (OL) with electrostatic deflectors can be combined with a cathode lens if the primary electrons are decelerated by an electrostatic field between OL and the specimen (S). The signal electrons (SE) are deflected from the primary beam and with an electrostatic transfer lens (TL) they are transferred to the detector; its ground electrode is split into four sections that allow centering of the beam. (From Vlček, 2005. unpublished. With permission.)

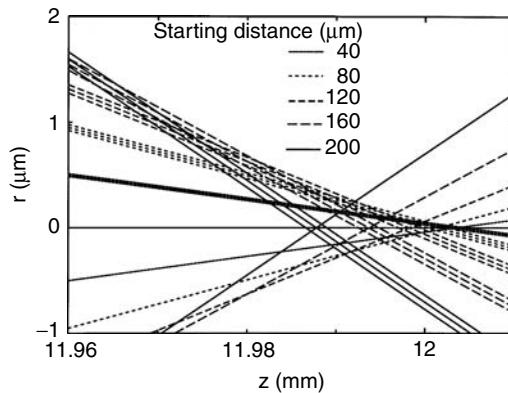


**FIGURE 5.28** Two possibilities of combining a magnetic lens (a) with an electrostatic lens. In the first case an electrostatic immersion lens (b) is used where electrons are decelerated well above the specimen, so that there is no strong electrostatic field at the sample itself; notice also an overlap of the electrostatic and magnetic fields. In the second case a cathode lens (c) arrangement is used; the electrons are decelerated in a strong field just above the sample. (From Lencová, B., *Proc. Multinational Congress Electron. Microsc.*, Stará Lesná, Slovakia, 19–24, 1995. With permission.)

(Vlček et al., 2005). The design uses the existing commercial electrostatic two-lens column with a Schottky cathode as an illumination part of a low-energy SEM. The crucial element in the new electron optical setup is the weak Wien filter, which separates the signal electrons by deflecting them by 15 degrees from the primary beam. The objective lens resembles the LEEM design of Adamec et al. (1998). An auxiliary condenser lens transfers the beam crossover to the Wien filter and the transfer lens is used to focus the back focal plane of the objective lens on a 2D detector, for example, a fast back-side illuminated CCD element with  $80 \times 80$  pixels (Horáček, 2003); the evaluation of intensities in reflected beam and the angular or energy distribution should enable observation of contrast in a scanning microscope similar to that observed in a LEEM.

The magnetic objective lens can be combined with any of the two types of electrostatic lenses, as illustrated in Figure 5.28 (Lencová, 1995b). The electron optical properties can be visualized by displaying the electron ray-traces as shown in Figure 5.29. At 1 keV energy on the sample, the two approaches result in similar optical properties.

In the low-voltage SEM, the resolution can be improved by employing the correction of spherical and chromatic aberrations with multipole correctors in an electrostatic immersion lens, an idea proposed by Zach (1990). The electrostatic detector objective lens (EDOL) combines the deceleration of the electron beam, a multipole corrector for chromatic and spherical aberration correction, and a detection system for secondary electrons. Zach and Haider (1994) successfully demonstrated that the correction of spherical and chromatic aberrations is possible in a low-voltage SEM. The commercial production of the corrected scanning microscope started only recently (Uno et al., 2005). See also Chapter 12.



**FIGURE 5.29** Axial aberration coefficients can be derived from electron trajectories starting at different heights parallel to the axis before reaching the magnetic lens. The trajectories demonstrate the effect of spherical aberration, and at the top of the figure also that of chromatic aberration (the electrons are traced with energies differing by  $\pm 0.5$  eV from the nominal energy). The analysis of the beam in the vicinity of a sample shows that for the geometry of Figure 5.26 and deceleration from 20 to 1 keV both concepts provide similar results: the image side  $C_S$  and  $C_C$  coefficients are 5.0 and 2.3 mm for the compound lens (b), 6.3 and 0.7 mm for the cathode lens (c), respectively. The magnetic lens alone would have  $C_S = 141$  mm and  $C_C = 28$  mm for the same sample position. The analysis confirmed the computations of Chmelfk et al. (1989). (From Lencová, B., Proc. Multinational Congress Electron. Microsc., Stará Lesná, Slovakia, 19–24, 1995. With permission.)

Miniature SEM columns like the one shown in Figure 5.3, operating at a voltage below 1 keV (Murray et al., 1991), also represent an advanced concept in electron beam lithography, where structures below 100 nm can be made with 10 nm accuracy very efficiently if many such columns are operated in parallel. Electron optical properties, tolerances of lens constructions in such small dimensions (bore roundness, surface flatness, edge smoothness, etc.), and alignment (concentricity, parallelism, etc.) were investigated by Kern and Chang (1994). Electron sources at these dimensions have a spherical aberration coefficient of only 10  $\mu\text{m}$ , and so they can provide three orders of magnitude higher brightness than conventional field-emission sources.

#### 5.4.5 OTHER APPLICATIONS OF ELECTROSTATIC LENSES

Electrostatic optics is an important part of modern electron spectrometers for surface studies. A system of several lenses, consisting mostly of simple cylinders, is used to transport a beam from the sample to the input of an electron spectrometer, and at the same time it provides the energy scanning of the spectrum if the analyzer voltages are fixed. A review by King (1995) provides an introduction to electron optics specific to this type of application. In the same series Erskine (1995) summarizes design aspects for input optics for electron spectrometers. Technical aspects of electrostatic lenses used at the input of electron spectrometers were given in Section 5.3.1, and further references can also be found in Roy (1990) and Heddle (1991, 2000).

#### 5.4.6 ION MICROSCOPY AND LITHOGRAPHY

Scanning ion optical devices for ion lithography are covered in Chapter 11. Focused ion beam systems using field-emission ion sources have become important for the semiconductor industry, for failure analysis, mask repair, and circuit modification. As with electron probes, decelerating ion optical systems, investigated by Narum and Pease (1988) and Aihara et al. (1989), should be able to provide a focused ion beam at relatively low energies for localized low-energy ion beam deposition or shallow doping ion implantation. The smallest holes (nanopores) with 4 nm diameter have been fabricated with Ga ions in 20 nm SiC membranes (Gierak et al., 2007). In scanning systems, the sequential nature of

the signal collection limits the speed with which we can obtain an image. For example, the image in a LEEM should be obtained in a shorter time than it is possible with a scanning system (Veneklasen, 1991). Similar arguments apply also for imaging ion systems reviewed by Slodzian (1980). The spatial distribution of ion species can be obtained faster with a scanning ion microscope providing a simultaneous image of many image points in parallel (Slodzian et al., 2003).

Projection ion lithography can compete with the speed and cost of other lithographic techniques (Stengl et al., 1992; Chalupka et al., 1994; Loeschner et al., 2001) also because the ions are more efficient in the modification of the properties of a resist. For an illustration, the first of such ion optical devices is shown in Figure 5.5.

## 5.5 CONCLUSIONS

Electrostatic lenses are widely used as transport lens systems in the spectrometry of electron or ion beams as well as in ion beam optical systems. The renewal of production of electrostatic TEMs is not very likely, although they may be feasible for some routine microscopy where ultimate resolution is not a goal. Low-voltage TEMs still have to find their application niche, and intermediate acceleration/deceleration pioneered by Heinemann and Möllenstedt (1967) can be one of the approaches. LEEMs and PEEMs are very useful for surface studies, where a UHV environment is needed, and this goes well with electrostatic optics. If such instruments are used as attachments, the electron optical demands are usually not very high. To increase the resolution, aberration correction was proposed, and several projects are now aimed at this goal. The correction of axial aberrations in a low-voltage SEM using an electrostatic objective lens has already been demonstrated. Low-voltage SEMs and inspection systems for electronics are required, and further development in this direction can be expected. A number of specimens easily withstand the electrostatic field as used in a cathode lens arrangement, and this solution allows achievement of lower landing energies and also low aberrations. Certainly at miniature dimensions, in spite of some problems with signal detection, electrostatic optics fulfills all demands. The optics and theory of electrostatic lenses still have to be further improved. Disadvantages in the design and use of electrostatic optics include the problems connected with electrical breakdown and insulation and the disturbing influence of contamination. Perhaps the largest revival of electrostatic optics is in high-resolution focused ion beams.

## 5.6 FURTHER INFORMATION

The history of electrostatic optics can be followed in the recollections of Ruska (1980) and Cosslett (1987); Hawkes and Kasper (1989) give a thorough overview of older references. The explosive development in the first decade of electron optics may be appreciated by comparing the first book on electron optics by Brüche and Scherzer (1934) with the following book by Brüche and Recknagel (1941). The book edited by Busch and Brüche (1937) that contains contributions, partly reprinted from journals, of leading scientists in the area is interesting. The book gives an overview not only of experimental and theoretical articles on electron optics and microscopy of that time but also of other electron optical elements (cathode ray tubes, pick-up tubes, image converters, multipliers, first television). The historical roots of electrostatic optics itself are somewhat obscure; see the comments in Busch's article in Busch and Brüche (1937), and some rivalry between the two Berlin groups can be felt in the appendix to Brüche and Recknagel (1941). Fortunately, priority claims in this area do not have the bitter consequences of Rüdenberg's patent (Ruska, 1980). The literature of that time is rather sparse, and detailed chronological comparison of the developments is not possible. Important early articles, such as Busch (1927), Davisson and Calbick (1931, 1932), Zworykin et al. (1942), Möllenstedt (1949), and McMullan (1953), are reprinted in SPIE Milestone Series (Hawkes, 1994). Outstanding older textbooks are by Zworykin et al. (1945), Cosslett (1951), Glaser (1952), and Grivet (1972).

As for sources of further references, almost 300 articles are cited in chapter 35 of Hawkes and Kasper (1989), on rotationally symmetric electrostatic lenses; very valuable, in particular, is an overview and critical assessment of used field models. Further chapters of the book cover quadrupoles, cylindrical lenses, mirrors and prisms, mentioned only briefly in this chapter. Over 200 articles are referenced by Baranova and Yavor (1989), where references to more exotic lenses and Russian literature can be found, also covered partly by Szilagyi (1988). Measurements of electron optical properties are reviewed by Hanszen and Lauer (1967), Rempfer (1985), as well as Baranova and Yavor (1989). For other older results a good starting point is the review of Hanszen and Lauer (1967) as well as Mulvey and Wallington (1973).

## ACKNOWLEDGMENTS

Support of the grant agency of the Czech Republic (projects 202/94/1100, 202/94/0565, and 102/05/0886) is acknowledged. The author is indebted to Tom Mulvey, Aston University, U.K., Frank H. Read, University of Manchester, U.K., and P. W. Hawkes, CNRS Toulouse, for their language corrections and valuable remarks. The author thanks G. Rempfer, Portland University, U.S., G. Stengl, Ion Microfabrication Systems GmbH, Vienna, Austria and I. Vlček of Institute of Scientific Instruments, Brno for providing figures used in this chapter.

## REFERENCES

- Adamec, P., Bauer, E., and Lencová, B., 1998. Compact low-energy electron microscope for surface imaging, *Rev. Sci. Instrum.*, 69(10), 3583–3587.
- Adamec, P., Delong A., and Lencová, B., 1995. Miniature magnetic electron lenses with permanent magnets, *J. Microsc.*, 179(2), 129–132.
- Adriaanse, J. P., van der Steen, H. W. G., and Barth, J. E., 1989. Practical optimization of electrostatic lenses, *J. Vac. Sci. Technol.*, B7(4), 651–666.
- Aihara, R., Kasahara, H., Sawaragi, H., Shearer, M. H., and Thompson, W. B., 1989. Optical system for a low-energy focused ion beam, *J. Vac. Sci. Technol.*, B7(1), 79–82.
- Aihara, R., Obata, M., Sawaragi, H., and Shearer, M. H., 1988. Stabilization of an electrostatic lens for a focused ion beam system, *J. Vac. Sci. Technol.*, B6(3), 958–961.
- Aken, R. H. van, Lenc, M. and Barth, J. E., 2004. Aberration integrals for the low-voltage foil corrector. *Nucl. Instrum. Meth. Phys. Res. A*, 519(1–2), 205–215.
- Ardenne, M. von, 1956. *Tabellen für Elektronenphysik, Ionenphysik und Übermikroskopie*, VEB Deutscher Verlag der Wissenschaften, Berlin.
- Ardenne, M. von, 1964. *Tabellen zur angewandten Physik*, VEB Deutscher Verlag der Wissenschaften, Berlin.
- Baranova, L. A. and Read, F. H., 1994. Electrostatic electron-optical crossed lens with controlled astigmatism, *Rev. Sci. Instrum.*, 65(6), 1993–1997.
- Baranova, L. A. and Yavor, S. Ya., 1989. The optics of round and multipole electrostatic lenses, *Adv. Electron. Electron Phys.*, 76, 1–207.
- Baranova, L. A., Yavor, S. Ya., and Read, F. H., 1996. Crossed aperture lenses for the correction of chromatic and aperture aberrations, *Rev. Sci. Instrum.* 67(3), 756–760.
- Barth, J. E., Jansen, G. H., and Kruit, P., 1990. Low-voltage diffraction-limited probe current: Limits due to e–e interactions in intermediate crossover, *Electron Microsc.*, 1, 394–395.
- Bauer, E., 1994. Low energy electron microscopy, *Rep. Prog. Phys.*, 57(9), 895–938.
- Bauer, E., 2007. LEEM and SPLEEM, in *Science of Microscopy* (P. W. Hawkes and J. C. H. Spence, Eds), Springer, New York, pp. 605–656.
- Benez, A., Plies E., and Ziegler, W., 1995. Optimierung einer elektrostatischen Transferlinse mit genetischen Algorithmen, *Optik*, 100(Suppl. 6), 3.
- Berger, C., 1983. Design of rotationally symmetrical electrostatic mirror for time-of-flight mass spectrometry, *J. Appl. Phys.*, 54(7), 3699–3703.
- Berger, C. and Baril, M., 1982. Studies of three-cylinder electrostatic mirrors and lenses, *J. Appl. Phys.*, 53(6), 3950–3956.

- Bernheim, M., 2006. Image acquisition with immersion objective lenses using electrons emitted with several tenths of an electron volt energies: Towards high spatial resolution ESCA analysis, *Ultramicroscopy*, 106(4–5), 398–412.
- Bertram, S., 1942. Calculation of axially symmetric fields, *J. Appl. Phys.*, 13, 496–502.
- Boersch, H., Hamisch, H., and Ehrlich, W., 1963. Oberflächenentladungen über Isolatoren im Vakuum, *Z. Angew. Phys.*, 15(6), 518–525.
- Boersch, H., Hamisch, H., and Schirrmüller, H., 1966. Negative Sekundärionen in Mikroüberschlägen, *Z. Angew. Phys.*, 22(1), 5–7.
- Boersch, H., Hamisch, H., and Wiesner, S., 1961. Über elektrische Mikroüberschläge im Vakuum, *Z. Angew. Phys.*, 13(10), 450–456.
- Boesten, L., 1988. Modified operation of the seven-element electrostatic Chutjian-Kuyatt lens system, *Rev. Sci. Instrum.*, 59(2), 233–237.
- Boulanger, P. and Baril, M., 1990. Design of a new focused multipassage mass spectrometer, *Nucl. Instrum. Meth. Phys. Res.*, A298(1–3), 161–170.
- Brüche, E., 1930. Strahlen langsamer Elektronen und ihre technische Anwendung, in *Forschung und Technik* (W. Petersen, Ed.), Verlag von Julius Springer, Berlin.
- Brüche, E., 1932. Elektronenmikroskop, *Naturwissenschaften*, 20(1), 49.
- Brüche, E. and Johannson, H., 1932. Elektronenoptik und Elektronenmikroskop, *Naturwissenschaften*, 20(21), 353–358.
- Brüche, E. and Recknagel, A., 1941. *Elektronengeräte*, Verlag von Julius Springer, Berlin.
- Brüche, E. and Scherzer, O., 1934. *Geometrische Elektronenoptik*, Verlag von Julius Springer, Berlin.
- Burghard, R. A., Swanson, L. W., and Orloff, J., 1987. Performance comparison of electrostatic lenses for field emission ion and electron sources, *J. Vac. Sci. Technol.*, B5(1), 364–371.
- Busch, H., 1926. Berechnung der Bahn von Kathodenstrahlen im axialsymmetrischen elektromagnetischen Felde, *Ann. Phys.*, 81(25), 974–993.
- Busch, H., 1927. Über die Wirkung der Konzentrierungsspule bei der Braunschen Röhre, *Arch. Elektrotech-nik.*, 18(6), 583–594.
- Busch, H. and Brüche, E., 1937. *Beiträge zur Elektronenoptik*, Verlag J. A. Barth, Leipzig.
- Carillon, A. and Gauthe, B., 1964. Réalisation d'un analyseur de vitesse électrostatique comportant une lentille à électrodes "croisées," *Electron Microsc.*, A, 51–52.
- Chalupka, A., Stengl, G., Buschbeck, H., Lammer, G., Vonach, J., Fischer, R., Hammel, E., Löschner, H., Nowak, R., Wolf, P., Finkelstein, W., Hill, R. W., Berry, I. L., Harriot, L. R., Melngailis, J., Randall, J. N., Wolfe, J. C., Stroh, H., Wollnik, H., Mondelli, A. A., Petillo, J. J., and Leung, K., 1994. Novel electrostatic column for ion projection lithography, *J. Vac. Sci. Technol.*, B12(6), 3513–3517.
- Chmelík, J., Veneklasen, L., and Marx, G., 1989. Comparing cathode lens configurations for low energy electron microscopy, *Optik*, 83(5), 155–160.
- Chutjian, A., 1979. Geometries and focal properties of two electron-lens systems useful in low-energy scattering, *Rev. Sci. Instrum.*, 50(3), 347–355.
- Cosslett, V. E., 1951. *Practical Electron Microscopy*, Butterworths Scientific Publications, London.
- Cosslett, V. E., 1987. Fifty years of instrumental development of the electron microscope, *Adv. Opt. Electron Microsc.*, 10, 215–267.
- Crewe, A. V., 1991a. An algebraic approach to the symmetrical einzel lenses, *Optik*, 88(3), 118–125.
- Crewe, A. V., 1991b. A new characterization of the magnetic lens, *Optik*, 89(2), 70–74.
- Cubric, D., Lencová, B., Read, F. H., and Zlámal J., 1998. Comparison of FDM, FEM and BEM for electrostatic charged particle optics, *Nucl. Instrum. Meth. Phys. Res.*, A427(1–2), 357–362.
- Dahl D. A., 2000. SIMION for the personal computer in reflection, *Int. J. Mass Spectrom.*, 200(1), 3–25.
- Dahl, D. A. and Delmore, J. E., 1988. *The SIMION PC/PS2 User's Manual*, Version 4.0. Idaho Nat. Eng. Lab. Rep. ECG-CS-7233.
- Dahl, D. A. and Delmore, J. E., 1994. *SIMION 5.0 Notes*, Appendix to Idaho Nat. Eng. Lab. Rep. EGG-CS-7233.
- Dahl, P., 1973. *Introduction to Electron and Ion Optics*. Academic Press, New York.
- Davisson, C. J. and Callick, C. J., 1931. Electron lenses, *Phys. Rev.*, 38(3), 585.
- Davisson, C. J. and Callick, C. J., 1932. Electron lenses, *Phys. Rev.*, 42(4), 580.
- Degenhardt, R., 1997. Optimization of electrostatic lens systems: Fast approximative calculation of symmetric and asymmetric einzel lenses, *Proc. SPIE*, 3155, 70–77.
- Delong, A., 1992. Low voltage TEM, *Electron Microsc.*, I, 79–81.
- Delong, A. and Drahoš, V., 1971. Low energy electron diffraction in an emission electron microscope, *Nat. Phys. Sci.*, 230, 196–197.

- Delong, A., Hladil, K., Kolařík, V. and Pavelka P., 2000. Low voltage electron microscope I. – design, in *Proc. EUREM 2000* (P. Tománek and R. Kolařík, Eds), II97–II98.
- DiChio, D., Natali, S. V., and Kuyatt, C. E., 1974. Focal properties of two-tube electrostatic lens for large and near-unity voltage ratios, *Rev. Sci. Instrum.*, 45(4), 559–565.
- Dunham, D., Deslodge, D. M., Rempfer, G. F. Skocylas, W. P., and Tonner, B. P., 1994. Evaluation of aberration coefficients of practical electrostatic lenses for X-ray absorption microspectroscopy and imaging, *Nucl. Instrum. Meth. Phys. Res.*, A347(1–3), 441–445.
- Edgcombe, C. J., 1991. On the perturbation of trajectories by small displacement, *Optik*, 86(4), 120–128.
- El Kareh, A. B. and El Kareh, J. C. J., 1970. *Electron Beams, Lenses and Optics*, Academic Press, New York, 2 vols.
- Engel, W., Kordesch, M. E., Rotermund, H. H., Kubala, S., and Vonoertzen, A., 1991. A UHV-compatible high-resolution photoelectron emission microscope for applications in surface science, *Ultramicroscopy*, 36(1–3), 148–153.
- Erskine, J. L., 1995. Electron energy analyzers, in *Experimental Methods in the Physical Sciences*, Vol. 29A (R. Celotta and T. Lucatorto, Eds), Academic Press, San Diego, pp. 209–230.
- Feng, J., Padmore, H., Wei, D. H., Anders, S., Wu, Y., Scholl, A., and Robin, D., 2002. Modeling the acceleration field and objective lens for an aberration corrected photoemission electron microscope, *Rev. Sci. Instrum.*, 73(3), 1514–1517.
- Feng, J. and Scholl, A., 2007. Photoemission electron microscopy (PEEM), in *Science of Microscopy* (P. W. Hawkes and J. C. H. Spence, Eds), Springer, New York, pp. 657–695.
- Fink, J. and Kisker, E., 1980. A method for rapid calculation of electron trajectories in multielement electrostatic cylinder lenses, *Rev. Sci. Instrum.*, 51(7), 918–920.
- Fink, R., Weiss, M. R., Umbach, E., Preikszas, D., Rose, H., Spehr, R., Hartel, P., Engel, W., Degenhardt, R., Wichtendahl, R., Kuhlenbeck, H., Erlebach, W., Ihmann, K., Schlögl, R., Freund, H.-J., Bradshaw A. M., Lilienkamp, G., Schmidt, T., Bauer, E., Benner, G., 1997. SMART: A planned ultrahigh-resolution spectromicroscope for BESSY II. *J. Electron Spectrosc. Relat. Phenom.*, 84, 231–250.
- Francken, J. C., 1967. Analogical methods for resolving Laplace's and Poisson's equations, in *Focusing of Charged Particles*, Vol. I (A. Septier, Ed.), Academic Press, New York, pp. 101–162.
- Frosien, J., Lanio S., and Feuerbaum, H. P., 1995. High precision electron optical system for absolute and CD-measurements on large substrates, *Nucl. Instrum. Meth. Phys. Res.*, A363(1–2), 25–30.
- Frosien, J., Plies, E., and Anger, K., 1989. Compound magnetic and electrostatic lenses for low-voltage applications, *J. Vac. Sci. Technol.*, B7(6), 1874–1877.
- Gierak, J., Madouri, A., Biance, A. L., Bourhis, E., Patriarche, G., Ulysse, C., Lucot, D., Lafosse, X., Auvray, L., Bruchhaus, L., and Jede, R., 2007. Sub-5 nm FIB direct patterning of nanodevices. *Microelectr. Eng.*, 84, 779–783.
- Glaser, W., 1952. *Grundlagen der Elektronenoptik*, Springer Verlag, Wien.
- Glaser, W. and Schiske, P., 1955. Die Bildfehler dritter Ordnung einer typischen elektrostatischen Einzellinse, *Optik*, 12(5), 233–245.
- Glatzel, U. and Lenz, F., 1988. Optimization of electrostatic immersion lenses, *Optik*, 79(1), 15–18.
- Gribi, M., Thürkauf, M., Villiger, W., and Wegmann, L., 1959. Ein 70 kV-Elektronenmikroskop mit kalter Kathode und elektrostatischer Linse, *Optik*, 16(2), 65–86.
- Griffith, O. H. and Engel, W., 1991. Historical perspective and current trends in emission microscopy, mirror electron microscopy and low-energy electron microscopy, *Ultramicroscopy*, 36(1–3), 1–28.
- Grivet P., 1972. *Electron Optics*, Second Edition, Pergamon, London.
- Hahn, E., 1955. Die elektronenoptischen Elemente des Elmi D2, *Jenaer Jahrbuch.*, I, 35–53.
- Hanai, T., Hibino, M., and Maruse, S., 1984. Reduction of the probe diameter through the correction of spherical aberration by the foil lens, *J. Electron Microsc.*, 33(4), 329–334.
- Hanszen, K.-J. and Lauer, R., 1967. Electrostatic lenses, in *Focusing of Charged Particles*, Vol. I (A. Septier, Ed.), Academic Press, New York, pp. 251–307.
- Hartel, P., Preikszas D., Spehr R., Müller, H. and Rose H., 2002. Mirror corrector for low-voltage electron microscopes. *Adv. Electron. Electron Phys.*, 120, 41–133.
- Harting, E. and Read, F. H., 1976. *Electrostatic Lenses*, Elsevier, Amsterdam.
- Hawkes, P. W. (Ed.), 1994. *Selected Papers on Electron Optics* (SPIE Milestone Series, Vol. MS 94), SPIE Optical Engineering Press, Bellingham, WA.
- Hawkes, P. W. and Kasper, E., 1989. *Principles of Electron Optics*, Vols. 1 and 2, Academic Press, London.
- Hawkes, P. W. and Lencová, B., 2002. Studies on the asymptotic cardinal elements and aberration coefficients of symmetric electrostatic lenses, *Optik* 113(2) 78–82.

- Hawkes, P. W., Lencová, B., and Lenc, M., 1995. Approximate expressions for electron lens properties, *J. Microsc.*, 179(2), 145–157.
- Heddle, D. W. O., 1991. *Electrostatic Lens Systems*, Adam Hilger, Bristol.
- Heddle, D. W. O., 2000. *Electrostatic Lens Systems*, Second Edition, Institute of Physics Publishing, Bristol and Philadelphia.
- Heddle, D. W. O. and Papadovassilakis, N., 1984. The magnification behaviour of a five-element electrostatic lens, *J. Phys. E: Sci. Instrum.*, 17(7), 599–605.
- Heinemann, K. and Möllenstedt, G., 1967. Elektronenmikroskopie mit langsamem Elektronen mittels Zwischenverzögerer, *Optik*, 26(1), 11–21.
- Horáček, M., 2003. Charge-coupled device area detector for low energy electrons, *Rev. Sci. Instrum.* 74 (10), 3379–3384.
- Hordon, L. S., Huang, Z., Maluf, N., Browning, R., and Pease, R. F. W., 1993. Limits of low-energy electron optics, *J. Vac. Sci. Technol.*, B11(6), 2299–2303.
- Hyder, H. R. McK., 1990. Large electrostatic accelerators, *Nucl. Instrum. Meth. Phys. Res.*, A287(1,2), 1–6.
- Janse, J., 1971. Numerical treatment of electron lenses with perturbed axial symmetry, *Optik*, 33(3), 270–281.
- Johannson, H. and Scherzer, O., 1933. Über die elektrische Elektronensammellinse, *Z. Phys.*, 80(3,4), 183–192.
- Joy, T., 1990. Electrostatic accelerator tubes—recent progress and future directions, *Nucl. Instrum. Meth. Phys. Res.*, A287(1,2), 48–63.
- Kanaya, K., Kawakatsu, H., Yamazaki, H., and Shimizu, K., 1966. Electron optical properties of three-electrode electron lenses, *J. Sci. Instrum.*, 43(7), 416–429.
- Kasper, E., 1987. An advanced method of field calculation in electron optical systems with axisymmetric boundaries, *Optik*, 77(1), 3–12.
- Kato, M. and Sekine, T., 1995. Spherical aberration correction of electrostatic lenses using spherical meshes, *J. Vac. Sci. Technol.*, A13(4), 2255–2260.
- Kato, M. and Sekine, T., 1996. Evaluation of the dispersive nature of meshes used for the spherical aberration correction of electrostatic lenses, *J. Vac. Sci. Technol.*, A14(2), 423–461.
- Kern, D. P. and Chang, T. H. P., 1994. Miniaturized electron optics: Basics and applications, *Electron Microsc.*, I, 149–152.
- King G. C., 1995. Electron and ion optics, in *Experimental Methods in the Physical Sciences*, Vol. 29A (R. Celotta and T. Lucatorto, Eds), Academic Press, San Diego, pp. 189–207.
- Koops, H. W. P., Munro, E., Rouse, J., Kretz, J., Rudolph, M., Weber, M., and Dahm, G., 1995. Miniature low voltage beam systems producible by combined lithographies, *Nucl. Instrum. Meth. Phys. Res.*, A363(1–2), 1–9.
- Kurihara, K., 1985. A focused ion beam system for submicron lithography, *J. Vac. Sci. Technol.*, B3(1), 41–44.
- Kurihara, K., 1990. Aberrations due to misalignment in electrostatic lens and deflection systems, *J. Vac. Sci. Technol.*, B8(1), 452–445.
- Langner, G., 1990. A modeling approach to charging in e-beam columns, *J. Vac. Sci. Technol.*, B8(6), 1711–1715.
- Latham, R. V. (Ed.), 1995. *High Voltage Vacuum Insulation (Basic Concept and Technological Practice)*, Academic Press, New York.
- Latham, R. V. and Xu, N. S., 1990. “Electron pin-holes”: The limiting defect for insulating high voltages by vacuum, a basis for new cold cathode electron sources, *Vacuum*, 20(18), 1173–1181.
- Lenc, M., 1995. Immersion objective lenses for very low energy electron microscopy, in *Proc. Multinational Congress Electron. Microsc.*, Stará Lesná, Slovakia, 103–104.
- Lenc, M. and Lencová, B., 1997. Third order geometrical and first order chromatic aberrations of electrostatic lenses, a comparison of formulas, *Optik*, 105(3), 121–128.
- Lenc, M. and Müllerová, I., 1992. Electron optical properties of a cathode lens, *Ultramicroscopy*, 41(4), 411–417.
- Lencová, B., 1995a. Computation of electrostatic lenses and multipoles by the first order finite element method, *Nucl. Instrum. Meth. Phys. Res.*, A363(1–2), 190–197.
- Lencová, B., 1995b. CAD in electron optics, in *Proc. Multinational Congress Electron. Microsc.*, Stará Lesná, Slovakia, 19–24.
- Lencová, B., 2002. Accuracy of electrostatic lens computations with FOFEM, *Ultramicroscopy* 93(3–4), 263–270.

- Lencová, B. and Zlámal, J., 2007. The development of EOD program for the design of electron optical devices. *Microsc. Microanal.*, 13 (Suppl. 3), 2–3.
- Lenz, F., 1973. Computer-aided design of electron optical systems, in *Image Processing and Computer-Aided Design in Electron Optics* (P. W. Hawkes, Ed.), Academic Press, London, pp. 274–282.
- Liebl, H., 1989. Ion probe microscopy, *Adv. Opt. Electron Microsc.*, 11, 101–151.
- Liebl, H. and Senftinger, B., 1991. Low-energy electron microscope of novel design, *Ultramicroscopy*, 36(1–3), 91–98.
- Liu, H. and Zhu, X., 1990. Numerical computation of the error effect in electron beam focusing, *Optik*, 84(4), 123–136.
- Loeschner, H., Stengl, G., Kaesmaier, R., and Wolter, A., 2001. Characterization of a process development tool for ion projection lithography. *J. Vac. Sci. Technol.* B19(6), 2520–2524.
- Mahl, H., Volkmann, H., and Weitsch, W., 1956. Über ein neues elektrostatisches Gebrauchs-Elektronenmikroskop, *Electron Microsc.*, 34–37.
- Martin, A., 1986. Cathode ray tubes for industrial and military applications, *Adv. Electron. Electron Phys.*, 67, 183–328.
- Martinez, G., and Sancho, M., 1983. Four-cylinder electrostatic lens I: Potential distribution and focal properties, *J. Phys. E: Sci. Instrum.*, 16(7), 625–630.
- Martinez, G. and Sancho, M., 1991. Application of the integral equation method to the analysis of electrostatic potentials and electron trajectories, *Adv. Electron Phys.*, 81, 1–41.
- Martinez, G., Sancho, M., and Read, F. H., 1983. Four-cylinder electrostatic lens II: Energy scanning at constant image position and magnification, *J. Phys. E: Sci. Instrum.*, 16(7), 631–635.
- Marx, G., Von Przychowski, M. D., Krömker, B., Ziethen, Ch., and Schönhense, G., 1994. Construction of a UHV-emission electron microscope with preparation chamber, *Electron Microsc.*, I, 239–240.
- Mayer, H. P. and Gaukler, K. H., 1987. Berechnungen elektrostatischer Abbildungssysteme für Flüssigmetall-Feldionenquellen, *Optik*, 77(3), 129–134.
- McMullan, D., 1953. An improved scanning electron microscope for opaque specimens, *Proc. IEE*, 100(75 part II), 245–249.
- Meisburger, W. D., Brodie, A. D., and Desai, A. A., 1992. Low-voltage electron-optical system for the high-speed inspection of integrated circuits, *J. Vac. Sci. Technol.* B10(6), 2804–2808.
- Meisburger, W. D. and Jacobsen, E. H., 1982. Electrostatic lens of very low spherical aberration, *Optik*, 62(4), 359–366.
- Möllenstedt, G., 1949. Die elektrostatische Lines als hochauflösender Geschwindigkeitsanalysator, *Optik*, 5(8/9), 499–517.
- Möllenstedt, G., 1955. 100keV-Elektronen im elektrostatischen Elektronenmikroskop (Zwischenbeschleuniger), *Optik*, 12(10), 441–466.
- Möllenstedt, G. and Laauser, H., 1980. High voltage stability of steel electrodes in an ultra-high vacuum at extremely small electrode separations, *Vacuum*, 30(7), 283–289.
- Möllenstedt, G. and Lenz, F., 1963. Electron emission microscopy, *Adv. Electron. Electron Phys.*, 18, 251–329.
- Mühlethaler, K., 1990. The beginning of electron microscopy in Zürich, in *History of Electron Microscopy in Switzerland* (J. R. Günter, Ed.), Birkhäuser Verlag, Basel, 59–82.
- Müllerová, I. and Frank, L., 2003. Scanning low-energy electron microscopy, *Adv. Electron. Electron Phys.*, 128, 309–443.
- Müllerová, I. and Lenc, M., 1992. Some approaches to low-voltage scanning electron microscopy, *Ultramicroscopy*, 41(4), 399–410.
- Mulvey, T. and Wallington M. J., 1973. Electron lenses, *Rep. Prog. Phys.*, 36(4), 347–421.
- Munro, E., 1988. Finite difference programs for computing tolerances for electrostatic lenses, *J. Vac. Sci. Technol.*, B6(3), 941–948.
- Munro, E., Orloff, J., Rutherford, R., and Wallmark, J., 1988. High-resolution, low-energy beams by means of mirror optics, *J. Vac. Sci. Technol.*, B6(6), 1971–1976.
- Murray, L. P., Stauffer, U., Bassous, E., Kern, D. P., and Chang, T. H. P., 1991. Experimental evaluation of a scanning tunneling microscope-microlens system, *J. Vac. Sci. Technol.*, B9(6), 2955–2961.
- Narum, D. H. and Pease, R. F. W., 1988. A variable energy focused ion beam system for *in situ* microfabrication, *J. Vac. Sci. Technol.*, B6(3), 966–973.
- Niemitz, H. P., 1980. Theoretische Untersuchung von elektrostatischen Drei-Elektroden-Filterlinsen, Dissertation, U. Tübingen.
- Ohye, T., Morita, C., and Shimoyama, H., 1993. Computer simulation of electron optical characteristics of accelerating tube for high voltage electron microscope, *Proc. SPIE*, 2014, 36–44.

- Okayama, S., 1990. A new type of quadrupole correction lens for electron-beam lithography, *Nucl. Instrum. Meth. Phys. Res. A*, 298(1–3), 488–495.
- Oku, K., Maruyama, M., and Fukushima, M., 1988. Electrostatic focusing and electrostatic deflection system with reduced aperture aberrations, *J. Vac. Sci. Technol.*, B6(6), 1999–2002.
- Orloff, J. and Swanson, L. W., 1979. An asymmetric electrostatic lens for field emission microprobe applications, *J. Appl. Phys.*, 50(4), 2494–2501.
- Paden, R. S. and Nixon, W. C., 1968. Retarding field scanning electron microscopy, *J. Sci. Instrum. (J. Phys. E)*, ser. 2, 1(11), 1073–1080.
- Plies, E., 1994. Electron optics of low-voltage electron beam testing and inspection. Part I: Simulation tools, *Adv. Opt. Electron Microsc.*, 13: 123–242.
- Preikszas, D., 1995. Korrektur des Farb- und Öffnungsfehlers eines Niederspannungs-Elektronenmikroskops mit Hilfe eines Elektronenspiegels. Dissertation, TH Darmstadt.
- Preikszas, D. and Rose, H., 1995. Procedures for minimizing the aberrations of electromagnetic compound lenses, *Optik*, 100(4), 179–187.
- Preikszas, D. and Rose, H., 1997. Correction properties of electron mirrors, *J. Electron Microsc.*, 46(1), 1–9.
- Rang, O. and Weitsch, W., 1956. Ein regelbares elektrostatisches Vierelektroden-Projektiv, *Optik*, 13(5), 201–208.
- Read, F. H., 1979. Electrostatic lenses, *Radiat. Eff.*, 44(1), 129–136.
- Read, F. H., 1982. Charged particle spectrometers, *US Patent* 4,358,680.
- Read, F. H., 1983. A “movable” electrostatic lens, *J. Phys. E: Sci. Instrum.*, 16(7), 636–642.
- Read, F. H., 1996. CPO-2D and CPO-3D programs, from RB Consultants Ltd. available from Integrated Sensors Ltd., PO Box 88, Manchester M60 1QD, UK.
- Read, F. H., Adams, A., and Soto-Montiel, J. R., 1971. Electrostatic cylinder lenses, I: Two-element lenses, *J. Phys. E: Sci. Instrum.*, 4(9), 625–632.
- Recknagel, A., 1941. Theorie des elektrischen Elektronenmikroskops für Selbststrahler, *Z. Physik*, 117(11,12), 689–708.
- Reimer, L., 1993. *Image Formation in Low-Voltage Scanning Electron Microscopy*, SPIE Optical Engineering Press, Bellingham, WA.
- Rempfer, G. F., 1985. Unipotential electrostatic lenses: Paraxial properties and aberrations of focal length and focal point, *J. Appl. Phys.*, 57(7), 2385–2401.
- Rempfer, G. F., 1990. A theoretical study of the hyperbolic electron mirror as a correcting element for spherical and chromatic aberration in electron optics, *J. Appl. Phys.*, 67(10), 6027–6040.
- Rempfer, G. F., Connell, R., Mercer, L., and Louisette, I., 1972. An electrostatic transmission electron microscope, *Am. Lab.*, 4, 39–45.
- Rempfer, G. F., Skoczyłas, W. P., and Griffith, O. H., 1991. Design and performance of a high-resolution photo-electron microscope, *Ultramicroscopy*, 36(1–3), 196–221.
- Renau, A., Read, F. H., and Brunt, J. N. H., 1982. The charge-density method of solving electrostatic problems with and without the inclusion of space-charge, *J. Phys. E: Sci. Instrum.*, 15(3), 347–354.
- Riddle, G. H. N., 1978. Electrostatic einzel lenses with reduced spherical aberration for use in field-emission guns, *J. Vac. Sci. Technol.*, 15(3), 857–860.
- Roach, T., Bakshi, A., and Kanter, K. F., 1995. Electrostatic microfocusing limits for positron beams, *Meas. Sci. Technol.*, 6(5), 496–501.
- Rose, H., 1978. Aberration correction of homogeneous magnetic deflection systems, *Optik*, 51(1), 15–38.
- Rose, H., 1987. The retarding Wien filter as a high-performance imaging filter. *Optik*, 77(1), 26–34.
- Rose, H. and Preikszas, D., 1992. Outline of a versatile corrected LEEM, *Optik*, 92(1), 31–44.
- Rose, H. and Preikszas, D., 1994. Corrected low-energy electron microscope for multi-mode operation, *Electron Microsc.*, I, 197–198.
- Rose, H. and Preikszas, D., 1995. Time-dependent perturbation formalism for calculating the aberrations of systems with large ray gradients, *Nucl. Instrum. Meth. Phys. Res.*, A363(1–2), 301–316.
- Rouse, J., 1994. Three-dimensional computer modeling of electron optical systems, *Adv. Opt. Electron. Microsc.*, 13, 1–121.
- Roy, D., 1990. Design of electron spectrometers, *Rep. Prog. Phys.*, 53(12), 1621–674.
- Ruska, E., 1980. The early development of electron lenses and electron microscopy (translated by T. Mulvey), *Microsc. Acta Suppl.*, 5, 1–140.
- Schluge, H., 1954. Das AEG-ZEISS-Elektronenmikroskop, *Zeiss Werkzeitsch.*, 104–112.

- Seijbel, L., Kruit, P., van Veen, A., and Schut, H., 1992. Scanning positron microanalysis, *Electron Microsc.*, 1, 231–232.
- Shannon, J. P., Philip, S. F., and Trump, J. G., 1965. Insulation of high voltage across solid insulators in vacuum, *J. Vac. Sci. Technol.*, 2(3), 234–239.
- Shimizu, K., 1983. Determination of the optimum electrode shape of electrostatic einzel lenses, *Jpn. J. Appl. Phys.*, 22(10), 1623–1626.
- Skoczyłas, W. P., Rempfer, G. F., and Griffith, O. H., 1991. A proposed modular imaging system for photo-electron and electron probe microscopy with aberration correction and for mirror microscopy and low-energy electron microscopy, *Ultramicroscopy*, 36(1–3), 252–261.
- Skoczyłas, W. P., Rempfer, G. F., and Griffith, O. H., 1994. Electron optical benches for in-line and branched systems. A new bench designed for mirror-based aberration correction and low energy electron microscopy, *Rev. Sci. Instrum.*, 65(10), 3183–3193.
- Slodzian, G., 1980. Microanalyzers using secondary ion emission, *Adv. Electron. Electron Phys.*, Suppl. 13B, 1–44.
- Slodzian, G., Hillion, F., Stadermann, F. J., and Horreard, F., 2003. Oxygen isotopic measurements on the Cameca Nanosims 50. *Appl. Surf. Sci.* 203–204, 798–801.
- Spangenberg, K. and Field, L. M., 1942. Some simplified methods of determining the optical characteristics of electron lenses, *Proc. IRE*, 30, 138–144.
- Stam, M. A. J. van der, Barth, J. E., and Kruit, P., 1993. Design of a multi-mode transport lens with optimization program SOEM, *Proc. SPIE*, 2014, 45–56.
- Steen, H. W. G. van der, Barth, J. E., and Adriaanse, J. P., 1990. Engineering constraints and computer-aided optimization of electrostatic lens systems, *Nucl. Instrum. Meth. Phys. Res.*, A298, 377–382.
- Stengl, G., Bösch, G., Chalupka, A., Fegerl, J., Fischer, R., Lammer, G., Löschner, H., Malek, L., Nowak, R., Traher, C., Wolf, P., Mauger, P., Shimkunas, A., Sen, S., and Wolfe, J. C., 1992. Ion projector wafer exposure results at 5× ion-optical reduction obtained with nickel and silicon stencil masks, *J. Vac. Sci. Technol.* B10(6), 2824–2828.
- Stenzel, W., 1971. Die elektrostatische Linse als Energiefilter, *Optik*, 32(4), 352–374.
- Sturrock, P. A., 1951. Perturbation characteristic functions and their application to electron optics, *Philos. Trans. R. Soc. London*, A210, 269–289.
- Szilagyi, M., 1983. Improvement of electrostatic lenses for ion beam lithography, *J. Vac. Sci. Technol.*, B1(4), 1137–1140.
- Szilagyi, M., 1988. *Electron and Ion Optics*, Plenum, New York.
- Trager-Cowan, C., Kay, S. M., and Heddle, D. W. O., 1990a. Multiple-element electrostatic lenses: 1. The five-element lens, *Meas. Sci. Technol.*, 1(8), 738–746.
- Trager-Cowan, C., Miller, R. B., Kay, S. M., and Heddle, D. W. O., 1990b. Multiple-element electrostatic lenses: I. A multiple disc lens, *Meas. Sci. Technol.*, 1(8), 747–750.
- Tromp, R. M. and Reuter, M. C., 1991. Design of a new photo-emission/low-energy electron microscope for surface studies, *Ultramicroscopy*, 36(1–3), 99–106.
- Tsumagari, T., Ohiwa, H., and Noda, T., 1988. Design of a low-aberration lens for focused ion beams, *J. Vac. Sci. Technol.*, B6(3), 949–952.
- Tsumagari, T., Ohiwa, H., and Noda, T., 1991. Unified mechanical aberration theory of electrostatic focusing-deflection systems, *Electron. Commun. Jpn.*, 74(1), 90–98.
- Tsuno, K., 1993. Electron optical analysis of a retarding Wien filter for electron spectroscopic imaging, *Rev. Sci. Instrum.*, 64(3), 659–666.
- Uno, S., Honda, K., Nakamura, N., Matsuya, M., and Zach, J., 2005. Aberration correction and its automatic control in scanning electron microscopes, *Optik*, 116(9), 438–448.
- Ura, K., 1994. Magnification dependence of aberration coefficients in electrostatic bipotential lens, *J. Electron Microsc.*, 43(6), 406–407.
- Veneklasen, L. H., 1991. Scanning versus direct imaging emission microscopy, *Ultramicroscopy*, 36(1–3), 63–75.
- Veneklasen, L. H., 1992. The continuing development of low-energy electron microscopy for characterizing surfaces, *Rev. Sci. Instrum.*, 63(12), 5513–5532.
- Verster, J. L., 1963. On the use of gauzes in electron optics, *Philips Res. Rep.*, 18(6), 465–605.
- Vlček, I., Lencová, B., and Horáček, M., 2005. Design of detector optics for a low energy SEM. *Proc. Microscopy Conf. 2005 – Dreiländertagung Davos (CH)*, Paul Scherrer Institute, Villigen, 59.

- 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*, Vol. I (A. Septier, Ed.), Academic Press, New York, 45–99.
- Weimer, E. and Martin, J.-P., 1994. Development of a new ultra-high performance scanning electron microscope, *Electron Microsc.*, I, 67–68.
- Williams, D. L., Read, F. H., and Bowring, N. J., 1995. Defocusing of charged particle beams transmitted through meshes, *Nucl. Instrum. Meth. Phys. Res.*, A363(1–2), 120–123.
- Wilska, A. P., 1964. Some basic features of a new low-voltage electron microscope, *Electron Microsc.*, 29–30.
- Wilska, A. P., Hanszen, K.-J., and Lauer, R., 1970. A high efficiency 5-electrode filter lens for low voltage electron microscopy, *Electron Microsc.*, II, 29–30.
- Wollnik, H., 1987. *Optics of Charged Particles*, Academic Press, Orlando.
- Yavor, M. I., 1993. Methods for calculation of parasitic aberrations and machining tolerances in electron optical systems, *Adv. Electron. Electron Phys.*, 86, 225–281.
- Zach, J., 1990. Resolution limits in low voltage scanning electron microscopes using retarding objective lenses, *Nucl. Instrum. Meth. Phys. Res.*, A298(1–3), 255–259.
- Zach, J. and Haider, M., 1994. Correction of spherical and chromatic aberration in a LVSEM, *Electron Microsc.*, I, 199–201.
- Zhu, X. and Munro, E., 1995. Second-order finite element method and its practical application in charged particle optics, *J. Microsc.*, 179(2), 170–180.
- Zworykin, V. K., Hillier, J., and Snyder, R. L., 1942. A scanning electron microscope, *ASTM Bull.*, 117(8), 15–23.
- Zworykin, V. K., Morton, G. A., Ramberg, E. G., Hillier, J., and Vance, A. W., 1945. *Electron Optics and the Electron Microscope*, Wiley, New York.

---

# 6 Aberrations

*Peter W. Hawkes*

## CONTENTS

6.1	Introduction.....	210
6.2	Methods of Calculating Aberrations.....	211
6.3	Geometric and Chromatic Aberration Coefficients.....	221
6.3.1	Introduction .....	221
6.3.2	Round Lenses .....	221
6.3.3	Quadrupole Lenses.....	249
6.3.4	Sextupoles.....	252
6.3.5	Superimposed Deflection Fields and Round Lenses .....	254
6.3.6	Mirrors and Cathode Lenses .....	259
6.3.7	Systems with Curved Optic Axes.....	265
6.3.8	Wien Filters .....	271
6.4	Aberration Representation and Symmetry .....	286
6.4.1	Representation .....	286
6.4.2	Symmetry .....	290
6.5	Parasitic Aberrations.....	292
6.6	Aberration Correction.....	294
6.6.1	Introduction .....	294
6.6.2	Minimization.....	296
6.6.3	Correction.....	297
6.6.3.1	Quadrupoles and Octopoles .....	297
6.6.3.2	Sextupoles .....	298
6.6.3.3	Correctors of Chromatic Aberration.....	300
6.6.3.4	General Multipole Correctors .....	301
6.6.3.5	Mirrors .....	304
6.6.3.6	Other Techniques .....	305
	For Further Information.....	310
	Acknowledgments.....	310
	References.....	311
A.1	Appendix: A Brief Introduction to Differential Algebra .....	339

## 6.1 INTRODUCTION\*

The very notion of optics presupposes that the elements being used for focusing, deflecting, or otherwise manipulating beams of light or charged particles have some dominant effect, which can be characterized by means of a small number of quantities, perturbed by minor effects likewise capable of simple characterization. For most optical elements—round and quadrupole lenses and magnetic and electrostatic deflecting elements—the dominant effect is *linear* and the quantities that characterize it are the familiar cardinal elements of Gaussian optics: focal lengths, positions of foci, principal planes, and nodal planes. The perturbations are then the *geometric aberrations*, which measure any departure from linear behavior and are characterized by the (geometric) aberration coefficients. These coefficients form groups of terms of increasingly higher order and it is fortunately rarely necessary in electron optics to go beyond the terms of the two lowest orders.

The dominant linear behavior may be perturbed by other effects than these geometric aberrations, which are merely an indication of the inadequacy of the linear approximation. Electron lenses are very sensitive to any variation in the energy of the incident particles; in the language of glass lenses, which has been adopted in particle optics, electron lenses suffer from *chromatic aberration*. This, too, is characterized by coefficients, and, again, only the lowest degree or at worst the two lowest degrees usually need be considered.

Finally, the focusing effect expected of a lens or other optical element may well be modified, usually detrimentally, by mechanical or other imperfections. No *round lens* has perfect circular symmetry. The magnetic material of which pole pieces are made may not be perfectly homogeneous. Alignment of a sequence of lenses is never exact. The consequences of such constructional flaws are known as *parasitic aberrations*. They can be characterized by coefficients and indeed the corresponding classification is important for understanding how they can be compensated or avoided. Nevertheless, numerical values of the parasitic aberration coefficients are not often encountered, since they will be different for each lens and may not be constant in time, especially if they are a consequence of misalignment or poor adjustment of some part of a complicated unit.

Unlike glass lenses, which of course occupy a well-defined volume of space, electron lenses have no clear limits, and, in the case of magnetic objective lenses, it is common to situate the specimen near the center of the magnetic field. Only the part of the field downstream from the object thus participates in the image-forming process; the part upstream plays the role of a final condenser lens. Beyond the objective of an electron microscope (or any other multistage instrument) will be a sequence of lenses, each magnifying the intermediate image furnished by its predecessor. In a microscope, these lenses are the intermediate and projector lenses. Unlike the case of the objective lens, the entire field participates in the image formation for these lenses. We are therefore obliged to distinguish between the *real* cardinal elements and aberration coefficients of objective lenses (in which an actual specimen is present) and the *asymptotic* cardinal elements and aberration coefficients of all other lenses (condensers, intermediate, and projector lenses) in which the relation between the asymptotes to rays incident on the lens and the asymptotes to rays emerging from the lens needs to be characterized.

In this account of aberrations, we first consider methods of calculation. Formulas are available for most of the coefficients in the form of aberration integrals, and it is reasonably straightforward to evaluate these formulas numerically. We then provide explicit information about the geometric and chromatic aberrations of the main types of optical element. This is complemented by an

\* The incorporation of aberration correctors and monochromators into commercial electron microscopes is at the heart of many of the changes to this chapter. The fifth-order aberrations of round lenses and the aberrations of sextupoles are now covered and long sections are devoted to the Wien filter, the family of  $\Omega$ -filters and to electron mirrors. The section on chromatic aberrations has been improved. The section on parasitic aberrations has been extended. The sections on aberration corrections have been largely rewritten. New references have been added throughout. In several cases, authors have been kind enough to send me files containing their published aberration coefficients, sometimes with errors or typos removed, so that any differences that may be noted are deliberate.

account of parasitic aberrations: with the rapid growth of computer-controlled units for the automatic alignment and adjustment of electron microscopes, there has been a resurgence of interest in these defects. A section is then devoted to the dependence of aberration coefficients on system parameters and in particular on symmetry.

Various methods have been proposed and tested, admittedly with unequal thoroughness, to correct the principal aberrations of round lenses. These culminated in the successful correction of spherical aberration in the 1990s. We describe these correctors and their descendants, which are now available commercially, in reasonable detail. Other possible ways of achieving correction are summarized only briefly. Once spherical aberration had been overcome, the chromatic aberration became a more serious obstacle to subångström imaging. Two ways of combating this problem have been pursued: (a) incorporation of monochromators into microscopes; such devices reduce the energy spread of the electron beam by removing electrons whose energy falls outside the passband of the monochromator and (b) design of correctors of the chromatic aberration. The former has the obvious disadvantage that the current is reduced while the latter must be compatible with the  $C_s$  corrector. Correctors are also discussed in Chapters 4 and 12.

It is not possible to give full derivations of all the formulas discussed here. These are often provided in Hawkes and Kasper (1989) and references to equations in those books are given in the form (Hawkes and Kasper, equation number).

Last, a word about terminology. Terms such as *second order*, *third degree* are used somewhat carelessly in aberration studies, and in most situations this is of little importance and the meaning is obvious. In dispersive systems, however, the coupling between chromatic and geometric effects becomes important, and it is prudent to adopt a terminology that prevents confusion. We follow that employed by Rose and Krah (1995), which coincides largely with earlier usage for the primary aberrations. An *nth-order aberration* contains geometric ray parameters (typically object coordinates and gradients), the sum of whose exponents is  $n$ . The exponent of any chromatic parameter is called the *degree* of the aberration—order + degree = *rank*.

## 6.2 METHODS OF CALCULATING ABERRATIONS

The properties of optical elements are established by first deriving paraxial equations, which determine the behavior of particles that remain close to and, in all except electron mirrors, travel at small angles to the optic axis. The solutions of these paraxial equations enable us to calculate the associated cardinal elements.

To go beyond the paraxial approximation, several methods have been developed. The simplest method is to set out from the usual equations of motion for charged particles in electrostatic and magnetic fields and, instead of limiting these equations to the lowest order terms, which yields the paraxial approximation, retain the next higher-order group of terms. For elements with a straight optic axis, for example, the exact equations of motion can be put into the form (Hawkes and Kasper, 3.22)

$$\begin{aligned} x'' &\equiv \frac{d^2x}{dz^2} = \frac{\rho^2}{2\hat{\Phi}} \left( \frac{\partial \hat{\Phi}}{\partial x} - x' \frac{\partial \hat{\Phi}}{\partial z} \right) + \frac{\eta\rho^2}{\hat{\Phi}^{1/2}} (\rho B_y - y' B_t) \\ y'' &\equiv \frac{d^2y}{dz^2} = \frac{\rho^2}{2\hat{\Phi}} \left( \frac{\partial \hat{\Phi}}{\partial y} - y' \frac{\partial \hat{\Phi}}{\partial z} \right) + \frac{\eta\rho^2}{\hat{\Phi}^{1/2}} (-\rho B_x - x' B_t) \end{aligned} \quad (6.1)$$

in which  $x(z)$  and  $y(z)$  are the Cartesian coordinates of an electron at the point  $z$  on the optic axis. The dimensionless quantity  $\rho$  denotes  $(1 + x'^2 + y'^2)^{1/2}$  and  $\eta$  is a constant given by

$$\eta = \left( \frac{e}{2m_0} \right)^{1/2} \cong 3 \times 10^5 \text{ C}^{1/2} \text{ kg}^{-1/2} \quad (6.2)$$

where  $-e$  is the charge on the electron and  $m_0$  its rest mass. The electrostatic field is described by the potential  $\Phi(x, y, z)$  and

$$\hat{\Phi} = \Phi(1 + \varepsilon\Phi) \quad (6.3)$$

where

$$\varepsilon = \frac{e}{2m_0c^2} \cong 1 \text{ MV}^{-1} \quad (6.4)$$

The magnetic field is described by  $\mathbf{B}$ , which has components  $(B_x, B_y, B_z)$ ; the tangential component of  $\mathbf{B}$ , denoted by  $B_t$ , is defined by

$$B_t = \frac{(x'B_x + y'B_y + B_z)}{\rho} \quad (6.5)$$

It is easy to expand the electrostatic potential  $\Phi(x, y, z)$  and the components of the field  $\mathbf{B}$  about the optic axis. For  $\Phi$ , we find

$$\begin{aligned} \Phi(x, y, z) &= \phi(z) - \frac{1}{4}(x^2 + y^2)\phi'' + \frac{1}{64}(x^2 + y^2)^2\phi^{(4)} - \frac{1}{2304}(x^2 + y^2)^3\phi^{(6)} \\ &\quad - xF_1(z) - yF_2(z) + \frac{1}{8}(x^2 + y^2)(xF_1'' + yF_2'') - \frac{1}{192}(x^2 + y^2)^2(xF_1^{(4)} + yF_2^{(4)}) \\ &\quad + \frac{1}{2}(x^2 - y^2)p_2(z) + xyq_2(z) - \frac{1}{24}(x^4 - y^4)p_2'' - \frac{1}{12}xy(x^2 + y^2)q_2'' \\ &\quad + \frac{1}{768}(x^2 + y^2)^2\{(x^2 - y^2)p_2^{(4)} + 2xyq_2^{(4)}\} \\ &\quad - \frac{1}{6}p_3(z)x(x^2 - 3y^2) - \frac{1}{6}q_3(z)y(3x^2 - y^2) + \frac{1}{96}(x^2 + y^2)\{x(x^2 - 3y^2)p_3'' + y(3x^2 - y^2)q_3''\} \\ &\quad + \frac{1}{24}p_4(z)(x^4 - 6x^2y^2 + y^4) + \frac{1}{6}q_4(z)xy(x^2 - y^2) \\ &\quad - \frac{1}{480}(x^2 + y^2)\{(x^4 - 6x^2y^2 + y^4)p_4'' + 4xy(x^2 - y^2)q_4''\} \\ &\quad - \frac{1}{120}\{x(x^4 - 10x^2y^2 + 5y^4)p_5 + y(5x^4 - 10x^2y^2 + y^4)q_5\} \\ &\quad + \frac{1}{720}\{(x^2 - y^2)(x^4 - 14x^2y^2 + y^4)p_6 + 2xy(3x^2 - y^2)(x^2 - 3y^2)q_6\} \\ &= \phi - \frac{r^2}{4}\phi'' + \frac{r^4}{64}\phi^{(4)} - \frac{r^6}{2304}\phi^{(6)} \\ &\quad - r(F_1 \cos \varphi + F_2 \sin \varphi) + \frac{r^3}{8}(F_1'' \cos \varphi + F_2'' \sin \varphi) - \frac{r^5}{192}(F_1^{(4)} \cos \varphi + F_2^{(4)} \sin \varphi) \\ &\quad + \frac{r^2}{2}(p_2 \cos 2\varphi + q_2 \sin 2\varphi) - \frac{r^4}{24}(p_2'' \cos 2\varphi + q_2'' \sin 2\varphi) + \frac{r^6}{768}(p_2^{(4)} \cos 2\varphi + q_2^{(4)} \sin 2\varphi) \\ &\quad - \frac{r^3}{6}(p_3 \cos 3\varphi + q_3 \sin 3\varphi) + \frac{r^5}{96}(p_3'' \cos 3\varphi + q_3'' \sin 3\varphi) \end{aligned}$$

$$\begin{aligned}
& + \frac{r^4}{24} (p_4 \cos 4\varphi + q_4 \sin 4\varphi) - \frac{r^6}{480} (p_4'' \cos 4\varphi + q_4'' \sin 4\varphi) \\
& - \frac{r^5}{120} (p_5 \cos 5\varphi + q_5 \sin 5\varphi) \\
& + \frac{r^6}{720} (p_6 \cos 6\varphi + q_6 \sin 6\varphi)
\end{aligned} \tag{6.6}$$

The second representation, in cylindrical polar coordinates  $(r, \varphi, z)$ , shows that elements having rotational symmetry about the optic axis (the  $z$ -axis) are wholly characterized by the function  $\phi(z)$ . We shall see that  $p_2(z)$  and  $q_2(z)$  are associated with quadrupoles and  $p_4(z)$  and  $q_4(z)$  with octopoles. The functions  $F_1(z)$  and  $F_2(z)$  arise in deflection systems while  $p_3(z)$  and  $q_3(z)$  occur when a threefold symmetry is present (sextupoles); fivefold symmetry corresponds to  $p_5(z)$  and  $q_5(z)$  and sixfold symmetry (dodecapoles) to  $p_6(z)$  and  $q_6(z)$ .

For magnetic fields, we have

$$\begin{aligned}
B_x = & -\frac{x}{2} B'(z) + \frac{x}{16} (x^2 + y^2) B''' - \frac{x}{384} (x^2 + y^2)^2 B^{(5)} + B_1(z) - \frac{1}{8} (3x^2 + y^2) B_1'' \\
& + \frac{1}{192} (x^2 + y^2)(5x^2 + y^2) B_1^{(4)} - \frac{1}{4} xy B_2'' + \frac{1}{48} (x^2 + y^2) xy B_2^{(4)} \\
& - xP_2 - yQ_2 + \frac{1}{6} x^3 P_2'' + \frac{y}{12} (3x^2 + y^2) Q_2'' - \frac{x}{384} (x^2 + y^2)(3x^2 - y^2) P_2^{(4)} \\
& - \frac{y}{384} (x^2 + y^2)(5x^2 + y^2) Q_2^{(4)} + \frac{1}{2} (x^2 - y^2) P_3 + xy Q_3 \\
& - \frac{1}{96} (5x^4 - 6x^2y^2 - 3y^4) P_3'' - \frac{xy}{24} (3x^2 + y^2) Q_3'' - \frac{x}{6} (x^2 - 3y^2) P_4 \\
& + \frac{y}{6} (y^2 - 3x^2) Q_4 + \frac{x}{240} (3x^4 - 10x^2y^2 - 5y^4) P_4'' + \frac{y}{120} (5x^4 - y^4) Q_4'' \\
& + \frac{1}{24} (x^4 - 6x^2y^2 + y^4) P_5 + \frac{xy}{6} (x^2 - y^2) Q_5 \\
& - \frac{x}{120} (x^4 - 10x^2y^2 + 5y^4) P_6 - \frac{y}{120} (5x^4 - 10x^2y^2 + y^4) Q_6
\end{aligned} \tag{6.7a}$$

$$\begin{aligned}
B_y = & -\frac{y}{2} B'(z) + \frac{y}{16} (x^2 + y^2) B''' - \frac{y}{384} (x^2 + y^2)^2 B^{(5)} + B_2(z) - \frac{1}{8} (3y^2 + x^2) B_2'' \\
& - \frac{1}{4} xy B_1'' + \frac{1}{192} (x^2 + y^2)(x^2 + 5y^2) B_2^{(4)} + \frac{xy}{48} (x^2 + y^2) B_1^{(4)} - xQ_2 + yP_2 \\
& - \frac{1}{6} y^3 P_2'' + \frac{x}{12} (x^2 + 3y^2) Q_2'' - \frac{y}{384} (x^2 + y^2)(x^2 - 3y^2) P_2^{(4)} \\
& - \frac{x}{384} (x^2 + y^2)(5y^2 + x^2) Q_2^{(4)} + \frac{1}{2} (x^2 - y^2) Q_3 - xy P_3 \\
& - \frac{1}{96} (3x^4 + 6x^2y^2 - 5y^4) Q_3'' + \frac{xy}{24} (x^2 + 3y^2) P_3'' - \frac{x}{6} (x^2 - 3y^2) Q_4 \\
& - \frac{y}{6} (y^2 - 3x^2) P_4 + \frac{x}{120} (x^4 - 5y^4) Q_4'' - \frac{y}{240} (5x^4 + 10x^2y^2 - 3y^4) P_4''
\end{aligned}$$

$$\begin{aligned}
& + \frac{1}{24}(x^4 - 6x^2y^2 + y^4)Q_5 - \frac{xy}{6}(x^2 - y^2)P_5 \\
& - \frac{x}{120}(x^4 - 10x^2y^2 + 5y^4)Q_6 + \frac{y}{120}(5x^4 - 10x^2y^2 + y^4)P_6
\end{aligned} \tag{6.7b}$$

$$\begin{aligned}
B_z = & B - \frac{1}{4}(x^2 + y^2)B'' + \frac{1}{64}(x^2 + y^2)^2 B^{(4)} - \frac{1}{2304}(x^2 + y^2)^3 B^{(6)} \\
& + xB'_1 + yB'_2 - \frac{1}{8}(x^2 + y^2)(xB'''_1 + yB'''_2) + \frac{1}{192}(x^2 + y^2)^2 (xB^{(5)}_1 + yB^{(5)}_2) \\
& - \frac{1}{2}(x^2 - y^2)P'_2 - xyQ'_2 + \frac{1}{24}(x^2 + y^2)\{(x^2 - y^2)P'''_2 + 2xyQ'''_2\} \\
& - \frac{1}{768}(x^2 + y^2)^2 \{(x^2 - y^2)P^{(5)}_2 + 2xyQ^{(5)}_2\} + \frac{1}{6}x(x^2 - 3y^2)P'_3 \\
& - \frac{1}{6}y(y^2 - 3x^2)Q'_3 - \frac{1}{96}(x^2 + y^2)\{x(x^2 - 3y^2)P'''_3 - y(y^2 - 3x^2)Q'''_3\} \\
& - \frac{1}{24}(x^4 - 6x^2y^2 + y^4)P'_4 - \frac{1}{6}xy(x^2 - y^2)Q'_4 + \frac{1}{480}(x^2 + y^2)(x^4 - 6x^2y^2 + y^4)P'''_4 \\
& + \frac{1}{120}xy(x^4 - y^4)Q''_4 \\
& + \frac{1}{120}\{x(x^4 - 10x^2y^2 + 5y^4)P'_5 + y(5x^4 - 10x^2y^2 + y^4)Q'_5\} \\
& - \frac{1}{720}\{(x^2 - y^2)(x^4 - 14x^2y^2 + y^4)P'_6 + 2xy(3x^2 - y^2)(x^2 - 3y^2)Q'_6\}
\end{aligned} \tag{6.7c}$$

The corresponding vector potential has components

$$\begin{aligned}
A_x = & -\frac{y}{2} \left[ B - \frac{1}{8}(x^2 + y^2)B'' + \frac{1}{192}(x^2 + y^2)^2 B^{(4)} \right] + \frac{1}{4}(x^2 - y^2)B'_2 \\
& - \frac{1}{48}(x^4 - y^4)B''_2 + \frac{1}{1536}(x^2 + y^2)(x^4 - y^4)B^{(5)}_2 - \frac{xy}{2}B'_1 \\
& + \frac{1}{24}xy(x^2 + y^2)B'''_1 - \frac{1}{768}xy(x^2 + y^2)^2 B^{(5)}_1 - \frac{1}{12}y(y^2 - 3x^2)P'_2 \\
& - \frac{1}{12}x(x^2 - 3y^2)Q'_2 + \frac{1}{192}(x^2 + y^2)\{y(y^2 - 3x^2)P'''_2 \\
& + x(x^2 - 3y^2)Q'''_2\} + \frac{(x^2 + y^2)^2}{7680}\{(3x^2 - y^2)yP^{(5)}_2 - (x^2 - 3y^2)xQ^{(5)}_2\} - \frac{1}{12}xy(x^2 - y^2)P'_3 \\
& + \frac{1}{48}(x^4 - 6x^2y^2 + y^4)Q'_3 + \frac{1}{960}(x^2 + y^2)\{4xy(x^2 - y^2)P'''_3 \\
& - (x^4 - 6x^2y^2 + y^4)Q'''_3\} + \frac{1}{240}\{y(5x^4 - 10x^2y^2 + y^4)P'_4 \\
& - x(x^4 - 10x^2y^2 + 5y^4)Q'_4\} - \frac{(x^2 + y^2)}{5760}\{(5x^4 - 10x^2y^2 + y^4)yP'''_4 \\
& - (x^4 - 10x^2y^2 + 5y^4)xQ'''_4\} - \frac{1}{1440}\{2xy(3x^4 - 10x^2y^2 + 3y^4)P'_5 \\
& - (x^2 - y^2)(x^4 - 14x^2y^2 + y^4)Q'_5\}
\end{aligned} \tag{6.8a}$$

$$\begin{aligned}
A_y = & \frac{x}{2} \left[ B - \frac{1}{8}(x^2 + y^2)B'' + \frac{1}{192}(x^2 + y^2)^2 B^{(4)} \right] \\
& + \frac{1}{4}(x^2 - y^2)B'_1 - \frac{1}{48}(x^4 - y^4)B'''_1 + \frac{1}{1536}(x^2 + y^2)(x^4 - y^4)B^{(5)}_1 \\
& + \frac{xy}{2}B'_2 - \frac{xy}{24}(x^2 + y^2)B'''_2 + \frac{1}{768}xy(x^2 + y^2)^2 B^{(5)}_2 \\
& - \frac{x}{12}(x^2 - 3y^2)P'_2 - \frac{y}{12}(3x^2 - y^2)Q'_2 + \frac{1}{192}(x^2 + y^2)\{x(x^2 - 3y^2)P'''_2 + y(3x^2 - y^2)Q'''_2\} \\
& - \frac{(x^2 + y^2)^2}{7680}\{(x^2 - 3y^2)xP^{(5)}_2 + (3x^2 - y^2)yQ^{(5)}_2\} + \frac{1}{48}(x^4 - 6x^2y^2 + y^4)P'_3 \\
& + \frac{xy}{12}(x^2 - y^2)Q'_3 - \frac{1}{960}(x^2 + y^2)\{(x^4 - 6x^2y^2 + y^4)P'''_3 + 4xy(x^2 - y^2)Q'''_3\} \\
& - \frac{x}{240}(x^4 - 10x^2y^2 + 5y^4)P'_4 - \frac{y}{240}(5x^4 - 10x^2y^2 + y^4)Q'_4 \\
& + \frac{(x^2 + y^2)}{5760}\{(x^4 - 10x^2y^2 + 5y^4)xP'''_4 + (5x^4 - 10x^2y^2 + y^4)yQ'''_4\} \\
& + \frac{1}{1440}(x^2 - y^2)(x^4 - 14x^2y^2 + y^4)P'_5 + \frac{xy}{720}(3x^4 - 10x^2y^2 + 3y^4)Q'_5
\end{aligned} \tag{6.8b}$$

$$\begin{aligned}
A_z = & -xB_2(z) + yB_1(z) + \frac{1}{8}(x^2 + y^2)(xB''_2 - yB''_1) - \frac{1}{192}(x^2 + y^2)^2(xB^{(4)}_2 - yB^{(4)}_1) \\
& + \frac{1}{2}(x^2 - y^2)Q_2 - xyP_2 - \frac{1}{24}(x^4 - y^4)Q''_2 + \frac{xy}{12}(x^2 + y^2)P''_2 \\
& + \frac{1}{768}(x^2 + y^2)^2\{(x^2 - y^2)Q^{(4)}_2 - 2xyP^{(4)}_2\} - \frac{x}{6}(x^2 - 3y^2)Q_3 \\
& - \frac{y}{6}(y^2 - 3x^2)P_3 + \frac{1}{96}(x^2 + y^2)\{x(x^2 - 3y^2)Q''_3 + y(y^2 - 3x^2)P''_3\} \\
& + \frac{1}{24}(x^4 - 6x^2y^2 + y^4)Q_4 - \frac{xy}{6}(x^2 - y^2)P_4 - \frac{1}{480}(x^2 + y^2) \\
& \times \{(x^4 - 6x^2y^2 + y^4)Q''_4 - 4xy(x^2 - y^2)P''_4\} - \frac{1}{120}\{x(x^4 - 10x^2y^2 + 5y^4)Q_5 \\
& - y(5x^4 - 10x^2y^2 + y^4)P_5\} + \frac{1}{720}\{(x^2 - y^2)(x^4 - 14x^2y^2 + y^4)Q_6 \\
& - 2xy(3x^2 - y^2)(x^2 - 3y^2)P_6\}
\end{aligned} \tag{6.8c}$$

In the work of Harald Rose (in particular, Rose, 2003a), a slightly different notation is adopted; the correspondence (already listed in Hawkes, 2007) is as follows. Rose (2003a) uses the expansion

$$\varphi(w, z) = \operatorname{Re} \sum_{m=0}^{\infty} \sum_{l=0}^{\infty} (-)^l \frac{m!}{l!(m+l)!} \left( \frac{w\bar{w}}{4} \right)^l \bar{w}^m \Phi_m^{[2l]}(z) \quad (6.9)$$

in which  $w = x + iy$ , and  $[2l]$  signifies  $2l$ -fold differentiation with respect to  $z$ . An overbar indicates the complex conjugate.

This leads to the following correspondences:

$$\begin{aligned} \phi(z) &= \Phi_0(z) \\ F_1(z) &= -\Phi_1^{(r)} & F_2(z) &= -\Phi_1^{(i)} & \Phi_1 &= -(F_1 + iF_2) \\ p_2(z) &= 2\Phi_2^{(r)} & q_2(z) &= 2\Phi_2^{(i)} & \Phi_2 &= \frac{1}{2}(p_2 + iq_2) \\ p_3(z) &= -6\Phi_3^{(r)} & q_3(z) &= -6\Phi_3^{(i)} & \Phi_3 &= -\frac{1}{6}(p_3 + iq_3) \\ p_4(z) &= 24\Phi_4^{(r)} & q_4(z) &= 24\Phi_4^{(i)} & \Phi_4 &= \frac{1}{24}(p_4 + iq_4) \\ p_5(z) &= -120\Phi_5^{(r)} & q_5(z) &= -120\Phi_5^{(i)} & \Phi_5 &= -\frac{1}{120}(p_5 + iq_5) \\ p_6(z) &= 720\Phi_6^{(r)} & q_6(z) &= 720\Phi_6^{(i)} & \Phi_6 &= \frac{1}{720}(p_6 + iq_6) \end{aligned} \quad (6.10)$$

and we have written

$$\Phi_m = \Phi_m^{(r)} + i\Phi_m^{(i)}$$

For the vector potential, the gauge adopted by Rose is the same as that used here, and the components of the vector potential are given by

$$\begin{aligned} A &= A_x + iA_y = \sum_{m=0}^{\infty} \sum_{l=0}^{\infty} \frac{(-)^l}{2i} \frac{m!}{l!(m+1+l)!} \left( \frac{w\bar{w}}{4} \right)^l w^{m+1} \bar{\Psi}_m^{[2l+1]}(z) \\ A_z &= \operatorname{Im} \sum_{m=0}^{\infty} \sum_{l=0}^{\infty} (-)^l \frac{m!}{l!(m+l)!} \left( \frac{w\bar{w}}{4} \right)^l \bar{w}^m \Psi_m^{[2l]}(z) \end{aligned} \quad (6.11)$$

The correspondence is now

$$\begin{aligned} \Psi_1 &= -(B_1 + iB_2) \\ \Psi_2 &= \frac{1}{2}(P_2 + iQ_2) \\ \Psi_3 &= -\frac{1}{6}(P_3 + iQ_3) \\ \Psi_4 &= \frac{1}{24}(P_4 + iQ_4) \\ \Psi_5 &= -\frac{1}{120}(P_5 + iQ_5) \\ \Psi_6 &= \frac{1}{720}(P_6 + iQ_6) \end{aligned} \quad (6.12)$$

For systems with *rotational symmetry*, Equations 6.1 collapse to the paraxial form

$$\begin{aligned} X'' + \frac{\gamma\phi'}{2\hat{\phi}}X' + \frac{\gamma\phi''}{4\hat{\phi}}X + \frac{\eta B}{\hat{\phi}^{1/2}}Y' + \frac{\eta B'}{2\hat{\phi}^{1/2}}Y &= 0 \\ Y'' + \frac{\gamma\phi'}{2\hat{\phi}}Y' + \frac{\gamma\phi''}{4\hat{\phi}}Y - \frac{\eta B}{\hat{\phi}^{1/2}}X' - \frac{\eta B'}{2\hat{\phi}^{1/2}}X &= 0 \end{aligned} \quad (6.13)$$

in which we have temporarily used  $X$  and  $Y$  instead of  $(x, y)$  for the transverse coordinates;  $\gamma$  denotes  $1 + 2\varepsilon\phi$  and  $\hat{\phi} = \phi(1 + \varepsilon\phi)$ . These coupled differential equations can be separated by introducing *rotating* coordinates:

$$X + iY = (x + iy)\exp\{i\theta(z)\} \quad (6.14)$$

with

$$\theta' = \frac{\eta B}{2\hat{\phi}^{1/2}} \quad (6.15)$$

whereupon Equations 6.13 become

$$\begin{aligned} x'' + \frac{\gamma\phi'}{2\hat{\phi}}x' + \frac{\gamma\phi'' + \eta^2 B^2}{4\hat{\phi}}x &= 0 \\ y'' + \frac{\gamma\phi'}{2\hat{\phi}}y' + \frac{\gamma\phi'' + \eta^2 B^2}{4\hat{\phi}}y &= 0 \end{aligned} \quad (6.16)$$

If we retain the next higher-order terms in the numerous series expansions that lead to Equations 6.16, we obtain equations of the form

$$\begin{aligned} \frac{d}{dz}(\hat{\phi}^{1/2}x') + \frac{\gamma\phi'' + \eta^2 B^2}{4\hat{\phi}^{1/2}}x &= \Lambda_1 \\ \frac{d}{dz}(\hat{\phi}^{1/2}y') + \frac{\gamma\phi'' + \eta^2 B^2}{4\hat{\phi}^{1/2}}y &= \Lambda_2 \end{aligned} \quad (6.17)$$

in which the complicated expressions  $\Lambda_1$  and  $\Lambda_2$  consist of numerous terms of third order in  $x, y$ , and their derivatives (Hawkes and Kasper, 24.7). The method of variation of parameters gives the solutions immediately. If  $g(z)$  and  $h(z)$  are the solutions of Equations 6.16 that satisfy the following initial conditions in some plane,  $z = z_o$ , which we refer to as the object plane,

$$\begin{aligned} g(z_o) &= h'(z_o) = 1 \\ g'(z_o) &= h(z_o) = 0 \end{aligned} \quad (6.18)$$

then

$$x(z) = x_og(z) + x'_oh(z) + \frac{h(z)}{W} \int_{z_o}^z \Lambda_1 g(\zeta) d\zeta - \frac{g(z)}{W} \int_{z_o}^z \Lambda_1 h(\zeta) d\zeta \quad (6.19)$$

with a similar expression for  $y(z)$ . In the image plane conjugate to  $z_o$ , which we denote by  $z = z_i$ , we have  $h(z_i) = 0$  and  $g(z_i) = M$ , the transverse magnification. The (constant) Wronskian  $W$  is given by

$$\begin{aligned} W &= \hat{\phi}^{1/2}(gh' - g'h) \\ &= \hat{\phi}_o^{1/2} = \hat{\phi}_i^{1/2} M h'_i \end{aligned} \quad (6.20)$$

and finally

$$\begin{aligned} \frac{x(z_i) - Mx_o}{M} &= -\frac{1}{W} \int_{z_o}^{z_i} \Lambda_1 h(z) dz \\ \frac{y(z_i) - My_o}{M} &= -\frac{1}{W} \int_{z_o}^{z_i} \Lambda_2 h(z) dz \end{aligned} \quad (6.21)$$

The quantity on the left-hand side of each of these expressions is the distance of the point of arrival of an electron in the image plane calculated with the correction term  $\Lambda_1$  or  $\Lambda_2$  from the point of arrival of the same electron calculated in the paraxial approximation. This quantity is the total geometric aberration. The presence of the magnification  $M$  in the denominator is a reminder that it is usual to refer the aberrations back to the object plane.

The strength of the foregoing approach, commonly known as the *trajectory method*, is that it is simple and straightforward, but it also has a weakness. We shall see that, for many optical elements, certain of the aberration coefficients are related very simply, and these relations are not revealed by this method. This drawback is remedied in the eikonal method, which, although marginally more complicated, is distinctly more powerful. In this method we derive the various equations not directly from a form of Newton's law but by writing down the Euler equations of a variational expression of the form

$$\int_{P_1}^{P_2} M dz \rightarrow \text{extremum} \quad (6.22)$$

which is to all intents and purposes identical with Fermat's principle. By analogy, the function  $M$ , defined by

$$M = \hat{\Phi}^{1/2}(1 + x'^2 + y'^2) - \eta(A_x x' + A_y y' + A_z) \quad (6.23)$$

is commonly known as the refractive index (although it is obviously not dimensionless). By expanding  $M$  as a power series in  $x$ ,  $y$ , and their derivatives and collecting up terms of the same order, we may write

$$M = M^{(2)} + M^{(4)} + M^{(6)} + \dots \quad (6.24)$$

On retaining only the second-order terms  $M^{(2)}$ , Equation 6.22 will give the paraxial Equations 6.16. A perturbation calculus (Glaser, 1952, 1956; Sturrock, 1951a, 1952, 1955; Hawkes, 1965b; Rose 1968, 2003a, 2008a; Rose and Petri, 1971; Plies and Typke, 1978; Plies, 2002; Hawkes and Kasper, chapter 22) enables us to derive the aberrations by differentiation. There are no redundant terms here, so that the equality or proportionality of any aberration coefficients will be immediately obvious.

The chromatic aberration may be treated in a similar fashion. For the primary chromatic aberration, the perturbation term  $M^{(c)}$  (corresponding to  $M^{(4)}$  in the case of the geometrical aberrations) is given by  $(\partial M^{(2)}/\partial\phi)\Delta\phi + (\partial M^{(2)}/\partial B)\Delta B$ ; in the rotating coordinate system,

$$M^{(c)} = C_1(x^2 + y^2) + C_2(x'^2 + y'^2) + C_3(xy' - x'y) \quad (6.25)$$

where

$$\begin{aligned} C_1 &= \frac{\phi'' + \gamma\eta^2 B^2}{16\hat{\phi}^{3/2}} \Delta\phi - \frac{\eta^2 B}{4\hat{\phi}^{1/2}} \Delta B \\ C_2 &= \frac{\gamma}{4\hat{\phi}^{1/2}} \Delta\phi \\ C_3 &= \frac{\eta B}{4} \left( \frac{\gamma}{\phi} \Delta\phi - 2 \frac{\Delta B}{B} \right) \end{aligned} \quad (6.26)$$

These two methods provide expressions for the aberration coefficients, as we shall see later, in the form of integrals, the integrands containing the field functions ( $\phi, B, \dots$ ) and their derivatives, as well as the paraxial solutions ( $g, h$ ) and their derivatives. A completely different approach has appeared with the widespread availability of powerful computing facilities. In its simplest form, this consists in solving the *exact* trajectory equations with very high accuracy for a set of rays that leave the object planes at various distances from the axis and with various gradients, including very small values in both cases. If the accuracy is good enough, the total aberration is simply the distance between the point of arrival of a paraxial electron and of one far from the axis and with a nonnegligible gradient. If enough rays are calculated, the aberration coefficients can be extracted from the resulting distances (Lencová and Zlámal, 2005, 2006, 2007; Lencová and Oral, 2006; Munro et al., 2006a). Crude though this technique may seem, it can be useful for systems for which the aberration integrals are numerous or very complicated or even unknown; alternatively, it can be used to provide an independent check of coefficients already calculated using the integrals. An obvious drawback to this approach is that the last stage of the calculation involves forming the difference between two small quantities. An ingenious way of avoiding this has been found by Kasper (1987a,b) who established a differential equation for this difference, which thus emerges directly when the differential equation is solved and does not have to be evaluated from two other calculations. For further details, we refer to Kasper's papers or to Hawkes and Kasper (section 33.4). Another disadvantage is that it provides no insight into the character of each aberration coefficient, notably any symmetries and any limitations on the permitted values of the coefficient. Yet another way of calculating the aberration coefficients numerically is based on *differential algebra*, a form of analysis introduced by Robinson (1961, 1966) and pursued by many others, see Mahwin (1994). This approach, first used in charged particle optics by Berz (1987) and presented methodically in Berz (1989, 1999), is coming into widespread use in electron optics (Liu, 2004b, 2006; Wang et al., 2002a,b, 2004, 2008a,b; Munro et al., 2006b,c; Cheng et al., 2006). In this method, the derivatives of field or potential needed for progressively more accurate ray tracing and hence for obtaining aberration coefficients of increasing rank are calculated to high accuracy in a novel way. In the original form of the method, it was necessary to know the axial field or potential extremely accurately so that the derivatives could be calculated correctly; model distributions are thus ideal whereas measured or computed fields must be fitted to suitable analytical functions before the method can be implemented. Typical choices of such basis functions are the Hermite polynomials and spline functions. The published examples show that this is often satisfactory, but it can be objected that the comparisons so far made are by no means exhaustive. More recently, however, Kang et al. (2007) have extended differential algebra

in such a way that fields calculated numerically can be used directly without recourse to models; in view of the elaborate interpolation that is required, it is not, however, obvious why differential algebra should be thought superior to exact ray tracing and fitting. A short introduction to the algebra is given in Appendix A.1, where other references will be found.

Yet another numerical tool is the matrizant, introduced by Dymnikov, and useful for very complicated structures (Dymnikov and Hellborg, 1993; Dymnikov, 1995, 1999; Dymnikov and Martínez, 1997; Dymnikov et al., 1998).

A procedure known as the  $\tau$ -variation method was originally introduced by Monastyrski and Shchelev (see Kulikov et al., 1985) and is presented in Greenfield et al. (2006) and Greenfield and Monastyrskiy (2009). It was first used to study temporal aberrations in cathode lenses and has been further developed for this purpose by Zhou et al. (2005a–c, 2006), where another approach to these aberrations is also described, the direct integral method. It is however more versatile than this application might suggest and is used by Greenfield et al. to study parasitic aberrations.

A very different way of investigating the effect of aberrations has been introduced by Pozzi, inspired by the multislice method of studying electron propagation in specimens; for details see Pozzi (1995) and di Sebastiano and Pozzi (1996).

The trajectory and eikonal methods allow us to establish formulas for the primary aberrations, which are of third order or first degree in systems with straight optic axes. The next higher approximation leads to fifth-order geometrical aberrations, third-degree chromatic aberrations, and, in principle, to fifth-rank aberrations, since mixed chromatic and geometric effects can then occur. With the development of correctors of third-order geometrical aberrations and ways of combating chromatic aberration, the fifth-rank aberrations have become important. Once again, the corresponding aberration coefficients can be calculated by deriving aberration integrals or by exact ray tracing. For details of the method of calculating the aberration integrals, we refer to Hawkes and Kasper (part IV). Expressions for all the fifth-order geometrical aberration coefficients of round lenses have been published by Hawkes (1965b), Li and Ni (1988) and Liu (2004a). See also Ximen et al. (1987). Here, we reproduce those of Liu, which have been checked by comparison with values obtained by a different approach for a test field; they are considerably simpler than those published by Hawkes (1965b) and moreover, some printing errors have been detected in the latter. Expressions for the fifth-order spherical aberration coefficient were derived by Wu (1957) and Archard (1960), who went on to evaluate it numerically for Glaser's bell-shaped field (see Equation 6.41): "As analytical integration would be tedious and not very useful, the integration has been done step by step using a digital computer, the Ferranti Mercury, taking only five minutes to work out  $C_5$  for ten values of  $k$ ." His values are not too different from those tabulated by Ximen and Liu (1997). Curves showing  $C_5$  as a function of excitation for this model are to be found in Rose (1968), where a slip in Archard's calculation is identified.

The fifth-order spherical aberration coefficient has been evaluated by some authors, neglecting the contribution arising from the combination of third-order aberrations and the paraxial equations. It is therefore of interest to know how large this contribution is. There is hardly any information on this point with the exception of the article by Ximen and Liu (1997), who tabulate the total aberration and the part generated by substituting the paraxial solution in the fifth-order terms in the refractive index (or sixth-order terms in the eikonal) separately for Glaser's bell-shaped field. For this case at least, the contribution from the combination term is very much smaller than the 'fifth-order' contribution, and a calculation based only on the latter will hence not be seriously misleading. Whether this is true more generally is not known. With this caveat in mind, we draw attention to the publication of Li et al. (2002) and the earlier articles of Li et al. (1995) and Mu et al. (1999), in which values of  $C_5$  are given for a range of excitations and gap/bore ( $S/D$ ) ratios using a model field tabulated by El-Kareh and El-Kareh (1970):

$$B(z) = \frac{2\mu_0 NI}{\pi S} \int_0^\infty \frac{\sin(Sx/D)x \cos(2xz/D)}{xI_0(x)} dx$$

Only the ‘fifth-order’ part is calculated. The behavior of  $C_5$  as a function of the excitation parameter is similar to that found by Ximen and Liu for all values of  $S/D$ : a rapid fall from a large positive value to a very small negative value, with little further change.

The expressions given by Liu and Hawkes use the rotating coordinate system. Another set of aberration integrals has been derived by Zhu (to be published), in this case in a fixed coordinate system; here the aberration coefficients of deflectors as well as round lenses are calculated, and the rotating system is not appropriate when deflection fields are present.

Even higher-order aberrations, in particular, the counterparts of spherical aberration, have been investigated in some detail by Matsuya et al. (1995, 1997). They use the Lie algebraic method (see Section 6.3.7) and differential algebra to calculate the coefficients for two magnetic field distributions (Glaser’s bell-shaped field and an exponential model) up to 11th order and to search for relations between the coefficients. In a later article (Matsuya and Nakagawa, 2004), the higher-order aberrations of a scanning electron microscope (SEM) equipped with an aberration corrector inspired by that of Zach are analyzed. A full treatment of the electrostatic retarding lens that may be placed before the specimen is given. High-order aberrations are also considered by Gao and Mu (1999) and Sang and Mu (2005).

This is a suitable point at which to mention the use of canonical coordinates. Aberration calculations are a little less laborious if, instead of characterizing rays in terms of position and gradient, we use the canonically conjugate quantities position and momentum. The results are not very different from those found using position and gradient for the primary aberrations, notably the third-order aberrations of round lenses. The relation between the two sets of results is, however, more complicated when we proceed to aberrations of higher order or rank, such as the fifth-order aberrations of round lenses; it is given explicitly by Rose (2003a, section 5.3.5). The canonical coordinates, originally used extensively by Sturrock, have been employed in numerous publications by Ximen but are not considered further here; for much detail, see the works of Ximen and Liu (Ximen, 1990a,b, 1991a,b, 1995, 1996; Liu and Ximen, 1992; Liu, 2002; Shao and Ximen, 1992; Ximen and Liu, 1996, 1997, 1998). Ximen and Liu (2000a,b) consider third- and fifth-order aberrations for Glaser’s bell-shaped field and for a new field model; the latter is of interest only for test purposes as the electrostatic field does not have the required form  $A + \phi(z)$ .

When calculating the secondary (or higher-order) aberrations numerically, it is important to recall that the ray gradient at the object is given by the tangent of the angle  $\alpha$  between the ray and the optic axis. If the results of exact ray tracing are fitted to the angle instead of its tangent, the value will not agree with that given by the aberration integrals (Lencová et al., 2008). The relation between the values of the coefficient when  $\alpha$ ,  $\tan \alpha$ , and  $\sin \alpha$  are used is given in Rose (1968).

## 6.3 GEOMETRIC AND CHROMATIC ABERRATION COEFFICIENTS

### 6.3.1 INTRODUCTION

In this section, formulas will be given for the aberration coefficients of the principal electron optical elements: round lenses, quadrupoles, octopoles, sextupoles, combined focusing and deflecting elements, energy analyzers, and mirrors. It is not possible to deal with all of these elements in full detail here, and we therefore concentrate on those that are most commonly needed.

### 6.3.2 ROUND LENSES

We recall that, at least in the very important case of magnetic lenses, it is necessary to distinguish between real and asymptotic aberrations. For the former, it is usual to employ not the coordinates of position and gradient at the object plane, as we did in Section 6.2, but the position coordinates in the object plane and the position coordinates in the aperture plane, which will in practice be situated in or close to the real image focal plane of the lens. We denote the position of this plane by  $z = z_a$ ,

with transverse coordinates  $(x_a, y_a)$ . It is then more convenient to work with the paraxial solutions  $s(z)$  and  $t(z)$  that satisfy the conditions

$$\begin{aligned} s(z_o) &= t(z_a) = 1 \\ s(z_a) &= t(z_o) = 0 \end{aligned} \quad (6.27)$$

These are of course related linearly to  $g(z)$  and  $h(z)$ . We note that, like  $h(z)$ ,  $t(z_i) = 0$  and like  $g(z)$ ,  $s(z_i) = M$ . In terms of these rays, the aberration in the image plane has the following form:

$$\begin{aligned} \Delta x_i &= \frac{x(z_i) - Mx_o}{M} = x_a [Cr_a^2 + 2KV + 2kv + (F - A)r_o^2] \\ &\quad + x_o (Kr_a^2 + 2AV + av + Dr_o^2) - y_o (kr_a^2 + aV + dr_o^2) \\ \Delta y_i &= \frac{y(z_i) - My_o}{M} = y_a [Cr_a^2 + 2KV + 2kv + (F - A)r_o^2] \\ &\quad + y_o (Kr_a^2 + 2AV + av + Dr_o^2) + x_o (kr_a^2 + aV + dr_o^2) \end{aligned} \quad (6.28)$$

so that

$$\Delta u_i \equiv \Delta x_i + \Delta y_i = \begin{pmatrix} u_o^* \\ 2u_a^* \end{pmatrix} \begin{pmatrix} D + id & F & K - ik \\ \frac{A + ia}{2} & K + ik & \frac{C}{2} \end{pmatrix} \begin{pmatrix} u_o^2 \\ u_o u_a \\ u_a^2 \end{pmatrix} \quad (6.29)$$

Here, we have written

$$\begin{aligned} r_a^2 &= x_a^2 + y_a^2 \\ r_o^2 &= x_o^2 + y_o^2 \\ V &= x_o x_a + y_o y_a = \frac{1}{2}(u_o u_a^* + u_o^* u_a) \\ v &= x_o y_a - x_a y_o = \frac{i}{2}(u_o u_a^* - u_o^* u_a) \end{aligned} \quad (6.30)$$

The quantities  $A, a, C, D, d, F, K$ , and  $k$  are the (real) geometric aberration coefficients.

$$\begin{aligned} \Delta u_i &= Cr_a^2 u_a \quad (\text{spherical aberration}) \\ &\quad + 2(K + ik)r_a^2 u_o + (K - ik)u_a^2 u_o^* \quad (\text{coma}) \\ &\quad + (A + ia)u_o^2 u_a^* \quad (\text{astigmatism}) \\ &\quad + Fr_o^2 u_a \quad (\text{field curvature}) \\ &\quad + (D + id)r_o^2 u_o \quad (\text{distortion}) \end{aligned} \quad (6.31)$$

Of these aberrations, the spherical aberration is the most important for objective lenses because it does not vanish for object points on the axis. In a magnifying system, the zone in the immediate vicinity of the optic axis is the most important close to the object, for it is here that the rays may be inclined at relatively steep angles; after several stages of magnification, the zone of interest is larger but the angles involved will be proportionately much smaller with the result that for intermediate and projector lenses, the asymptotic distortion is the most important aberration. For high-resolution work, the objective lens coma may become serious, but this aberration can fortunately be eliminated.

It is usual to express the spherical aberration (referred back to the object) in terms of gradient at the object plane:

$$\begin{aligned}\Delta x_i &= C_s x'_o (x'^2 + y'^2) \\ \Delta y_i &= C_s y'_o (x'^2 + y'^2)\end{aligned}\tag{6.32}$$

in which

$$C_s = \frac{C}{t'^3} = \frac{1}{\hat{\phi}_o^{1/2}} \int_{z_o}^{z_i} (L_1 h^4 + 2L_2 h^2 h'^2 + L_3 h'^4) dz\tag{6.33}$$

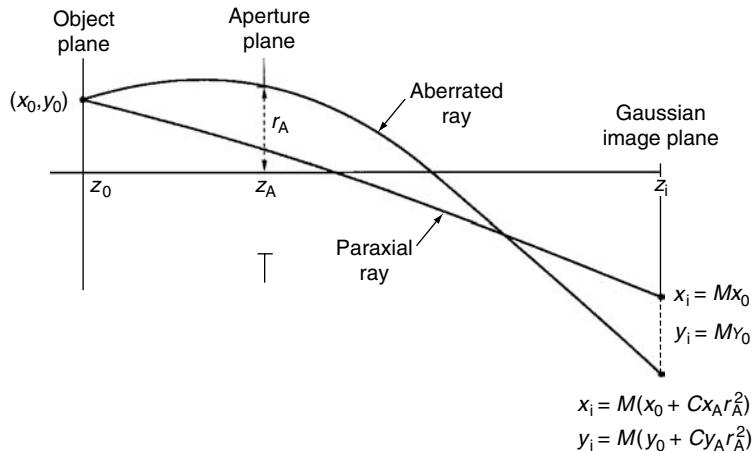
and

$$\begin{aligned}L_1 &= \frac{1}{32\hat{\phi}^{1/2}} \left( \frac{\phi''^2}{\hat{\phi}} - \gamma \phi^{(4)} + \frac{2\gamma\phi''\eta^2 B^2}{\hat{\phi}} + \frac{\eta^4 B^4}{\hat{\phi}} - 4\eta^2 B B'' \right) \\ L_2 &= \frac{1}{8\hat{\phi}^{1/2}} (\gamma\phi'' + \eta^2 B^2) \\ L_3 &= \frac{1}{2} \hat{\phi}^{1/2}\end{aligned}\tag{6.34}$$

The following quantities appear in the integrals for the anisotropic aberration coefficients  $a$ ,  $k$ , and  $d$ :

$$\begin{aligned}P &= \frac{\eta}{16\hat{\phi}^{1/2}} \left( \frac{\gamma\phi''B}{\hat{\phi}} - B'' + \frac{\eta^2 B^3}{\hat{\phi}} \right) \\ Q &= \frac{\eta B}{4\hat{\phi}^{1/2}} \\ R &= \frac{\eta^2 B^2}{8\hat{\phi}^{3/2}}\end{aligned}\tag{6.35}$$

Returning to Equation 6.28, we see that, for a circular aperture of radius  $r_A$ , the image of an object point  $(x_o, y_o)$  will not be a point situated at  $(x_i = Mx_o, y_i = My_o)$  but a circular disk of radius  $MCr_A^3$  irrespective of the object coordinates. We shall see that the sign of  $C$  is such that the outermost rays are focused more strongly than the paraxial rays (Figure 6.1). It is usual to define the longitudinal



**FIGURE 6.1** The effect of spherical aberration.

spherical aberration as the distance  $\zeta_i$  from the paraxial image plane to the plane in which the ray that touches the rim of the aperture intersects the axis. This is easily shown to be given by

$$\zeta_i = -\frac{MCr_A^2}{t'_i} = -\frac{M^2C}{t'_o} \left( \frac{\hat{\phi}_i}{\hat{\phi}_o} \right)^{1/2} r_A^2 \quad (6.36)$$

and corresponds to a distance  $\zeta_o$  in object space

$$\zeta_o = \frac{\zeta_i}{M^2} = -\frac{C}{t'_o} \left( \frac{\hat{\phi}_i}{\hat{\phi}_o} \right)^{1/2} r_A^2 \quad (6.37)$$

The radius of the spherical disk is in fact smaller than  $MCr_A^3$  in a plane between the paraxial image plane and the plane  $z = z_i - \zeta_i$ . This *disk of least confusion* is found to lie in the plane  $z = z_i - \zeta$ , where

$$\zeta = \frac{3}{4} |\zeta_i| \quad (6.38)$$

and its radius is  $MCr_A^3/4$ .

It is important to note that the spherical aberration coefficient of a given lens will be different when it is reversed. The relation between the *forward* and *backward* aberration coefficients is mostly needed for magnetic lenses, because these lenses are used as magnifying lenses in transmission electron microscopes (TEMs) and as demagnifying lenses in probe-forming instruments such as the scanning transmission electron microscope (STEM) and the SEM. A full discussion is given in Hawkes and Kasper (equations 24.52 through 24.55). The essential result is that for a magnetic lens with high-magnification spherical aberration coefficient  $C_s$ , the probe radius  $r_p$  when the same lens is used for demagnification is given by

$$r_p = C_s \theta_p^3 \quad (6.39)$$

in which  $\theta_p$  denotes the greatest angle at the probe.

The formula for  $C_s$  can be written in a large number of ways by partial integration of the various terms. An entirely general expression and many particular cases are listed in Hawkes and Kasper (equations 24.59 through 24.70). For magnetic lenses, two useful representations are

$$C_s = \frac{1}{48} \int_{z_o}^{z_i} \left( 5 \frac{\eta^2 B'^2}{\hat{\phi}} - \frac{\eta^2 BB''}{\hat{\phi}} + 4 \frac{\eta^4 B^4}{\hat{\phi}^2} \right) h^4 dz \quad (6.40a)$$

$$C_s = \frac{1}{16} \int_{z_o}^{z_i} \left( \frac{\eta^4 B^4}{\hat{\phi}^2} h^4 + 2(hB' + h'B)^2 \frac{\eta^2 h^2}{\hat{\phi}} + 2 \frac{\eta^2 B^2}{\hat{\phi}} h^2 h'^2 \right) dz \quad (6.40b)$$

The first of these has the merit that the paraxial solution  $h(z)$  appears undifferentiated. The second reveals immediately that  $C_s$  is *always positive*. This result, obtained by Otto Scherzer (1936), is so important that it is known as *Scherzer's theorem*. It shows that there is no hope of finding a lens geometry for which  $C_s$  vanishes. At best, we can seek the polepiece configuration for which  $C_s$  is smallest. Walter Glaser, another great theoretician of the first decades of electron optics, found Scherzer's result so unpalatable that he devoted considerable effort to disproving it. In particular, he inquired (Glaser, 1940) whether there was a field  $B(z)$  for which the part of the integrand of Equation 6.40a in parentheses vanished by regarding this part as a differential equation for  $B(z)$ . Unfortunately, but perhaps not surprisingly, the resulting field distribution is not suitable for focusing purposes. The electrostatic case was subsequently analyzed by Recknagel (1941a), with similar and equally unsurprising results.

Extensive sets of curves and tables showing the dependence of  $C_s$  on the lens geometry and excitation have been published (see Chapters 4 and 5). The integrals can be evaluated in closed form for a particular field model, also introduced by Glaser (1941), the so-called *bell-shaped field*:

$$B(z) = \frac{B_o}{1 + (z/a)^2} \quad (6.41)$$

Although this does not reproduce the field distribution in real lenses very faithfully, the results are useful for understanding the variation of  $C_s$  with lens geometry. For this field

$$\begin{aligned} \frac{C_s}{a} &= \left[ \frac{\pi k^2}{4\omega^3} - \frac{4k^2 - 3}{4(4k^2 + 3)} \cos\left(2\psi_o - \frac{\pi}{\omega}\right) \sin\left(\frac{\pi}{\omega}\right) \right] \operatorname{cosec}^4 \psi_o \\ &= C_4 m^4 + C_3 m^3 + C_2 m^2 + C_1 m + C_0 \end{aligned} \quad (6.42)$$

and when  $M \rightarrow \infty$

$$\frac{C_s}{a} = C_0 = \left[ \frac{\pi k^2}{4\omega^3} - \frac{4k^2 - 3}{8(4k^2 + 3)} \sin\left(\frac{2\pi}{\omega}\right) \right] \operatorname{cosec}^4\left(\frac{\pi}{\omega}\right)$$

in which  $k^2 = \eta^2 B_0^2 a^2 / 4\phi$  and  $\omega^2 = 1 + k^2$ ;  $z = a \cot \psi$  and  $m = 1/M$ . The coefficients  $C_j$  are obtained by setting

$$\cot \psi_o = m \operatorname{cosec}\left(\frac{\pi}{\omega}\right) + \cot\left(\frac{\pi}{\omega}\right) \quad (6.43)$$

giving

$$\begin{aligned} C_4 &= C_0 = \left[ \frac{\pi k^2}{4\omega^3} - \frac{4k^2 - 3}{8(4k^2 + 3)} \sin\left(\frac{2\pi}{\omega}\right) \right] \operatorname{cosec}^4\left(\frac{\pi}{\omega}\right) \\ C_3 &= C_1 = \left\{ \frac{\pi k^2}{\omega^3} \cot\left(\frac{\pi}{\omega}\right) - \frac{4k^2 - 3}{4(4k^2 + 3)} \left[ 3 + \cos\left(\frac{2\pi}{\omega}\right) \right] \right\} \operatorname{cosec}^3\left(\frac{\pi}{\omega}\right) \\ C_2 &= \left[ \frac{\pi k^2}{2\omega^3} (1 + 3 \cot^2\left(\frac{\pi}{\omega}\right)) - \frac{3(4k^2 - 3)}{2(4k^2 + 3)} \cot\left(\frac{\pi}{\omega}\right) \right] \operatorname{cosec}^2\left(\frac{\pi}{\omega}\right) \end{aligned} \quad (6.44)$$

Lists of numerical values are given by El-Kareh and El-Kareh (1970).

For electrostatic lenses, convenient forms of  $C_s$  are

$$\begin{aligned} C_s &= \frac{1}{32\phi_o^{1/2}} \int_{z_o}^{z_i} \phi^{1/2} \left[ -\frac{\phi''' \phi'}{2\phi^2} + 2 \frac{\phi''^2}{\phi^2} + 5 \left( \frac{\phi'}{\phi} \right)^4 - 5 \frac{\phi'' \phi'^2}{\phi^3} \right] h^4 dz \\ &= \frac{1}{64\phi_o^{1/2}} \int_{z_o}^{z_i} \phi^{1/2} (4\psi'^2 + 3\psi^4 - 5\psi^2\psi' - \psi\psi'') h^4 dz \\ &= \frac{1}{32\phi_o^{1/2}} \int_{z_o}^{z_i} \phi^{1/2} \tilde{C} h^4 dz \end{aligned} \quad (6.45)$$

in which

$$\psi = \phi'/\phi \quad (6.46)$$

and

$$\tilde{C} = \left( \frac{\phi''}{\phi} + \frac{\phi'}{\phi} \frac{h'}{h} - \frac{5}{4} \frac{\phi'^2}{\phi^2} \right)^2 + \frac{3}{2} \left( \frac{\phi''}{\phi} + \frac{\phi'}{\phi} \frac{h'}{h} - \frac{\phi'^2}{\phi^2} \right)^2 + \frac{2\phi'^2}{\phi^2} \left( \frac{h'}{h} + \frac{5\phi'}{6\phi} \right)^2 + \frac{\phi'^4}{36\phi^4} \quad (6.47)$$

Here, we have listed only nonrelativistic expressions; the relativistic forms are to be found in Hawkes and Kasper (chapter 24). As for the magnetic case, we have given, first, formulas in which derivatives of  $h(z)$  are absent from the integrand and, second, a version that shows that for electrostatic lenses too,  $C_s$  is positive definite. The same can be shown to be true for lenses in which both magnetic and electrostatic fields are present; Rose (1967, 1968) has proved that the relativistically correct formula can likewise be written as a positive definite quantity. An error in his expression has been detected and the correct form is to be found in Preikszas and Rose (1995).

If an approximation is sufficient, for preliminary design purposes, for example, then some observations of Crewe (1991a–c) can be useful. For symmetric electrostatic lenses with no overall accelerating or retarding effect (einzel lenses), the following quantities vary remarkably little over a very wide range of excitations (1991a):

$$f \ln \phi_m \quad \frac{C_s}{f^3} \quad \frac{C_c \phi_m^{1/2}}{f} \quad (6.48)$$

in which  $\phi_m$  is the axial potential at the center of the lens (the length of which is scaled to unity). For three-element electrostatic lenses that may have different potentials in object and image spaces, Crewe (1991c) finds that

$$\begin{aligned} f &= \frac{1.09L\phi_i^{1/4}}{(\ln\phi_m)^2 \left(1 - \rho + \frac{1}{2}\rho^2\right)} l(\phi_i, \phi_m) \quad \text{with } \rho = \frac{\ln\phi}{\ln\phi_m} \\ C_s &= \frac{17.5f^3\phi_m^{1/2}}{L^2\phi_i^{1/4}} m(\phi_i, \phi_m) \\ C_c &= \frac{2f\phi_i^{3/4}}{\phi_m^{1/2}} n(\phi_i, \phi_m) \end{aligned} \quad (6.49)$$

in which  $\phi_i$  is the potential in image space and  $L$  is a characteristic length. The functions  $l$ ,  $m$ , and  $n$  remain close to unity for a wide range of excitations. The axial potential is set equal to unity in object space. For magnetic lenses, Crewe (1991b) gives equally simple approximate formulas of the form

$$f = C_c \propto \frac{\phi_o}{(NI)^2} \quad C_s \propto f^3 \quad (6.50)$$

All these approximate expressions are based on the thin-lens approximation, which is not discussed here. For further examination of this approximation, see the work recapitulated in Glaser (1952, 1956) and that of Der-Shvarts and Makarova (1972, 1973), Hawkes (1980b), Renau and Heddle (1986) with comment by Hawkes (1987) and by Renau and Heddle (1987) and the critical examination of Crewe's proposals by Hawkes et al. (1995); see also Ura (1994).

In connection with these approximate formulas, we draw attention to articles by Rempfer (1985, 1999a,b, 2000) who claimed that the number of independent aberrations of round lenses was fewer than had been believed. The analysis on which this claim was based was, however, shown to be unsound (Hawkes, 1999, 2000) and numerical analysis confirmed that the coefficients are related as claimed by Rempfer only in the thin-lens approximation (Hawkes and Lencová, 2002), a result that was already well known (e.g., Hawkes, 1987).

If the spherical aberration cannot be eliminated by skillful design, how small can it be made? A careful study of this important question was made by Tretner (1959), who established the limits within which the coefficients of spherical ( $C_s$ ) and chromatic aberrations ( $C_c$ , discussed later) will lie for practical lenses. His results are recapitulated in full in Hawkes and Kasper (section 36.4), where his curves are reproduced, and reconsidered in Section 6.6.2.

Although the spherical aberration is the most serious aberration of objective and probe-forming lenses, the other aberrations cannot be neglected. We now list one form of the aberration integrals, together with expressions for Glaser's bell-shaped field. For many other forms of these expressions, see Hawkes and Kasper (chapter 24). A particularly useful article by Lencová and Lenc (1997) provides comparisons of the numerous published formulas for the primary geometric and chromatic aberration coefficients of magnetic lenses.

### Coma

$$\begin{aligned} K &= \frac{1}{\phi_o^{1/2}} \int_{z_o}^{z_i} [L_1 st^3 + L_2 (st)'tt' + L_3 s't'^3] dz \\ k &= \int_{z_o}^{z_i} (Pt^2 + Qt'^2) dz \end{aligned}$$

### Astigmatism and field curvature

$$F = \frac{1}{\phi_o^{1/2}} \int_{z_o}^{z_i} [2L_1 s^2 t^2 + L_2 (st)'^2 + 2L_3 s'^2 t'^2 + R] dz$$

$$A = \frac{1}{\phi_o^{1/2}} \int_{z_o}^{z_i} (L_1 s^2 t^2 + 2L_2 s' t' + L_3 s'^2 t'^2 - R) dz$$

$$a = 2 \int_{z_o}^{z_i} (Pst + Qs't') dz$$

$$R = \frac{\eta^2 B^2}{8\hat{\phi}^{1/2}}$$

### Distortion

$$D = \frac{1}{\phi_o^{1/2}} \int_{z_o}^{z_i} [L_1 s^3 t + L_2 (st)' ss' + L_3 s'^3 t'] dz$$

$$d = \int_{z_o}^{z_i} (Ps^2 + Qs'^2) dz$$

Since the position of the aperture plane  $z = z_a$  appears in the definitions of the rays  $s(z)$  and  $t(z)$ , the values of the aberration coefficients will change if another aperture plane is selected. For a different aperture plane  $z = \bar{z}_a$ , the rays  $s$  and  $t$  become  $\bar{s}(z)$  and  $\bar{t}(z)$ , satisfying the conditions

$$\bar{s}(z_o) = \bar{t}(\bar{z}_a) = 1 \quad \bar{s}(\bar{z}_a) = \bar{t}(z_o) = 0 \quad (6.51)$$

$\bar{s}(z)$  and  $\bar{t}(z)$  must be expressible as linear combinations of  $s(z)$  and  $t(z)$ , and it is easily shown that

$$\begin{aligned} \bar{s}(z) &= s(z) - \frac{s(\bar{z}_a)}{t(\bar{z}_a)} t(z) \\ \bar{t}(z) &= \frac{t(z)}{t(\bar{z}_a)} \end{aligned} \quad (6.52)$$

For this new position of the aperture, the coma coefficient  $K(\bar{z}_a)$  is proportional to  $K(z_a) - s(\bar{z}_a)C_s(z_a)/t(\bar{z}_a)$ . If therefore we place the aperture at the point for which

$$\frac{s(\bar{z}_a)}{t(\bar{z}_a)} = \frac{K(z_a)}{C_s(z_a)} \quad (6.53)$$

the coma will vanish. This is known as the coma-free point,  $z = z_k$ .

The expressions for the coefficients of coma, astigmatism, field curvature, and distortion for Glaser's bell-shaped field as given in Glaser (1952) with a few minor corrections are listed in the following equations; the notation  $k = \eta B_0 a / 2\phi^{1/2}$  is so standard that we have retained it and denote the anisotropic coma coefficient in this list by  $\tilde{k}$ .

$$\begin{aligned}
K &= \frac{1}{4a^2} \frac{\sin^2 \psi_a}{\sin^3 \omega(\psi_a - \psi_o)} \left[ \pi k^2 \cos \omega(\psi_a - \psi_o) - \frac{\omega^2}{4\omega^2 - 1} \sin \pi \omega \{(4\omega^2 - 7)\omega \cos(\psi_o + \psi_i) \right. \\
&\quad \left. \cos \omega(\psi_a - \psi_o) + 3 \sin(\psi_o + \psi_i) \sin \omega(\psi_a - \psi_o)\} \right] \\
&\xrightarrow{M \rightarrow \infty} \frac{1}{4a^2} \frac{\sin^2 \psi_a}{\sin^3 \omega \psi_a} \left[ \pi k^2 \cos \omega \psi_a - \frac{\omega^2}{4\omega^2 - 1} \sin \left( \frac{\pi}{\omega} \right) \right. \\
&\quad \left. \times \left\{ (4\omega^2 - 7)\omega \cos \left( \frac{\pi}{\omega} \right) \cos \omega \psi_a + 3 \sin \left( \frac{\pi}{\omega} \right) \sin \omega \psi_a \right\} \right] \\
\tilde{k} &= \frac{k}{4a^2} \frac{\sin^2 \psi_a}{\sin^2 \omega(\psi_a - \psi_o)} \left\{ \pi \frac{2\omega^2 - 1}{2\omega} - \frac{\omega^2(\omega^2 - 2)}{k^2} \sin \frac{\pi}{\omega} \cos(\psi_o + \psi_i) \right\} \\
&\xrightarrow{M \rightarrow \infty} \frac{k}{8a^2} \frac{\sin^2 \psi_a}{\sin^2 \omega \psi_a} \left\{ \pi \frac{2\omega^2 - 1}{\omega} - \frac{\omega^2(\omega^2 - 2)}{k^2} \sin \left( \frac{2\pi}{\omega} \right) \right\} \tag{6.54}
\end{aligned}$$

$$\begin{aligned}
A &= \frac{1}{4a^2} \frac{\sin \psi_o \sin \psi_a}{\sin^3 \omega(\psi_o - \psi_a)} \left\{ \pi k^2 \cos 2\omega(\psi_o - \psi_a) - \frac{\omega^2}{4\omega^2 - 1} \sin \frac{\pi}{\omega} \{ \omega(4\omega^2 - 7) \right. \\
&\quad \left. \cos(\psi_o + \psi_i) \cos 2\omega \cos(\psi_a - \psi_o) \} - 3 \sin(\psi_o + \psi_i) \sin 2\omega(\psi_a - \psi_o) \right\} \\
&\xrightarrow{M \rightarrow \infty} \frac{1}{4a^2} \frac{\sin(\pi/\omega) \sin \psi_a}{\sin^3 \omega \psi_a} \left[ \pi k^2 \cos 2\omega \psi_a - \frac{\omega^2}{4\omega^2 - 1} \sin \frac{\pi}{\omega} \right. \\
&\quad \left. \left\{ \omega(4\omega^2 - 7) \cos \left( \frac{\pi}{\omega} \right) \cos 2\omega \psi_a + 3 \sin \left( \frac{\pi}{\omega} \right) \sin 2\omega \psi_a \right\} \right] \\
a &= \frac{k}{2a^2} \frac{\sin \psi_o \sin \psi_a}{\sin^2 \omega(\psi_o - \psi_a)} \left[ -\frac{\pi}{2\omega} (2\omega^2 - 1) \cos \omega(\psi_a - \psi_o) + \frac{\omega}{k^2} \sin \frac{\pi}{\omega} \right. \\
&\quad \left. \{ (\omega^2 - 2)\omega \cos(\psi_o + \psi_i) \cos \omega(\psi_a - \psi_o) + \sin(\psi_o + \psi_i) \sin \omega \psi_a - \psi_o \} \right] \\
&\xrightarrow{M \rightarrow \infty} \frac{k}{2a^2} \frac{\sin(\pi/\omega) \sin \psi_a}{\sin^2 \omega \psi_a} \left[ \frac{\pi}{2\omega} (2\omega^2 - 1) \cos \omega \psi_a - \frac{\omega}{k^2} \sin \frac{\pi}{\omega} \right. \\
&\quad \left. \left\{ (\omega^2 - 2)\omega \cos \left( \frac{\pi}{\omega} \right) \cos \omega \psi_a + \sin \left( \frac{\pi}{\omega} \right) \sin \omega \psi_a \right\} \right] \tag{6.55}
\end{aligned}$$

$$\begin{aligned}
2A - F &= \frac{\omega k^2}{a^2} \frac{\sin \psi_o \sin \psi_a}{\sin \omega(\psi_a - \psi_o)} \left\{ \frac{\pi}{\omega} - \sin \frac{\pi}{\omega} \cos(\psi_o + \psi_i) \right\} \\
&\xrightarrow{M \rightarrow \infty} \frac{\omega k^2}{2a^2} \frac{\sin(\pi/\omega) \sin \psi_a}{\sin \omega \psi_a} \left( \sin \frac{2\pi}{\omega} - \frac{2\pi}{\omega} \right)
\end{aligned}$$

$$\begin{aligned}
D &= \frac{k}{4a^2} \frac{\sin^2 \psi_o}{\sin^3 \omega(\psi_a - \psi_o)} \left[ \left\{ \pi - \omega \sin \left( \frac{\pi}{\omega} \right) \cos(\psi_o + \psi_i) \right\} \cos \omega(\psi_a - \psi_o) \right. \\
&\quad + \frac{\omega}{k^2} \frac{2\omega^2 + 1}{4\omega^2 - 1} \sin \left( \frac{\pi}{\omega} \right) \cos(\psi_o + \psi_i) \cos 3\omega(\psi_a - \psi_o) \\
&\quad \left. - \frac{3}{k^2} \frac{\omega^2}{4\omega^2 - 1} \sin \left( \frac{\pi}{\omega} \right) \sin(\psi_o + \psi_i) \sin 3\omega(\psi_a - \psi_o) \right] \\
&\xrightarrow{M \rightarrow \infty} \frac{k^2}{8a^2} \frac{\sin^2(\pi/\omega)}{\sin^3 \omega \psi_a} \left\{ \left( \omega \sin \frac{2\pi}{\omega} - 2\pi \right) \cos \left( \psi_a - \frac{\pi}{\omega} \right) \right. \\
&\quad \left. + \frac{\omega}{k^2} \frac{2\omega^2 + 1}{4\omega^2 - 1} \sin \frac{2\pi}{\omega} \cos 3\omega \psi_a - \frac{6}{k^2} \frac{\omega^2}{4\omega^2 - 1} \sin^2 \frac{\pi}{\omega} \sin 3\omega \psi \right\} \quad (6.56) \\
d &= \frac{k}{4a^2} \frac{\sin^2 \psi_o}{\sin^2 \omega(\psi_a - \psi_o)} \left[ \pi \frac{2\omega^2 - 1}{2\omega} - \omega^2 \sin \frac{\pi}{\omega} \cos(\psi_o + \psi_i) \right. \\
&\quad \left. + \frac{\omega}{k^2} \sin \frac{\pi}{\omega} \{ \omega \cos(\psi_o + \psi_i) \cos 2\omega(\psi_o - \psi_a) + \sin(\psi_o + \psi_i) \sin 2\omega(\psi_o - \psi_a) \} \right] \\
&\xrightarrow{M \rightarrow \infty} \frac{k}{8a^2} \frac{\sin^2(\pi/\omega)}{\sin^2 \omega \psi_a} \left\{ \pi \frac{2\omega^2 - 1}{\omega} - \omega^2 \sin \frac{2\pi}{\omega} \right. \\
&\quad \left. + \frac{\omega^2}{k^2} \sin \frac{2\pi}{\omega} \cos 2\omega \psi_a - \frac{2\omega}{k^2} \sin^2 \frac{\pi}{\omega} \sin 2\omega \psi_a \right\}
\end{aligned}$$

For the fifth-order aberrations of round lenses, we follow Liu (2004a) closely. A few minor errors in the published formulas have been corrected (Liu, private communication 2007)

$$\begin{aligned}
\Delta \mathbf{r}_5 &= (A_5 \mathbf{r}'_o + a_5 \mathbf{r}'^{**}_o) (\mathbf{r}'_o \cdot \mathbf{r}'_o)^2 \\
&\quad + (B_{51} \mathbf{r}_o + b_{51} \mathbf{r}^{*}_o) (\mathbf{r}'_o \cdot \mathbf{r}'_o)^2 + (B_{52} \mathbf{r}'_o + b_{52} \mathbf{r}'^{**}_o) (\mathbf{r}'_o \cdot \mathbf{r}'_o) (\mathbf{r}_o \cdot \mathbf{r}'_o) \\
&\quad + (C_{51} \mathbf{r}_o + c_{51} \mathbf{r}^{*}_o) (\mathbf{r}'_o \cdot \mathbf{r}'_o) (\mathbf{r}_o \cdot \mathbf{r}'_o) + (C_{52} \mathbf{r}'_o + c_{52} \mathbf{r}'^{**}_o) (\mathbf{r}'_o \cdot \mathbf{r}'_o) (\mathbf{r}_o \cdot \mathbf{r}_o) \\
&\quad + (C_{53} \mathbf{r}'_o + c_{53} \mathbf{r}'^{**}_o) (\mathbf{r}_o \cdot \mathbf{r}'_o)^2 \\
&\quad + (D_{51} \mathbf{r}_o + d_{51} \mathbf{r}^{*}_o) (\mathbf{r}'_o \cdot \mathbf{r}'_o) (\mathbf{r}_o \cdot \mathbf{r}_o) + (D_{52} \mathbf{r}'_o + d_{52} \mathbf{r}^{*}_o) (\mathbf{r}_o \cdot \mathbf{r}'_o)^2 \\
&\quad + (D_{53} \mathbf{r}'_o + d_{53} \mathbf{r}'^{**}_o) (\mathbf{r}_o \cdot \mathbf{r}'_o) (\mathbf{r}_o \cdot \mathbf{r}_o) \\
&\quad + (E_{51} \mathbf{r}_o + e_{51} \mathbf{r}^{*}_o) (\mathbf{r}_o \cdot \mathbf{r}'_o) (\mathbf{r}_o \cdot \mathbf{r}_o) + (E_{52} \mathbf{r}'_o + e_{52} \mathbf{r}'^{**}_o) (\mathbf{r}_o \cdot \mathbf{r}_o)^2 \\
&\quad + (F_5 \mathbf{r}_o + f_5 \mathbf{r}^{*}_o) (\mathbf{r}_o \cdot \mathbf{r}_o)^2
\end{aligned} \quad (6.57)$$

The coefficients denoted by capital letters are isotropic aberrations, those denoted by lower-case letters are anisotropic.  $A_5$  and  $a_5$  characterize the fifth-order spherical aberration,  $B_5$  and  $b_5$  are coma-like,  $E_5$  and  $e_5$  resemble astigmatism and field curvature, and  $F_5$  and  $f_5$  are distortions. The others, which depend on  $(\text{angle})^3 \cdot (\text{position})^2$  and  $(\text{angle})^2 \cdot (\text{position})^3$ , have no simple analogs in third order.

Each of the analytical expressions for the aberration coefficients can be written in the unified form

$$X = -\frac{1}{\phi_o^{1/2}} \int_{z_o}^{z_i} \left\{ X_i - \frac{1}{\phi_o^{1/2}} (X_{c0} + X_{c1}r_z^2 + X_{c2}r_zr_\beta + X_{c3}r_\beta^2) \right\} dz \quad (6.58)$$

where  $X$  stands for each individual aberration coefficient and is a function of  $F_{lmn}$  and  $F_{lmn}^*$ .

### *Isotropic aberration coefficients*

$$\begin{aligned} A_{5i} &= 6F_{030} \\ A_{5c0} &= -12(F_{011}\varepsilon_{020} - F_{020}\varepsilon_{011} - 4F_{020,o}F_{020}) \\ A_{5c1} &= -3 \left( \frac{\phi}{\phi_o} \right)^{1/2} (F_{011}^2 + F_{010}^{*2}) \\ A_{5c2} &= 24 \left( \frac{\phi}{\phi_o} \right)^{1/2} F_{020}F_{011} \\ A_{5c3} &= -48 \left( \frac{\phi}{\phi_o} \right)^{1/2} F_{020}^2 \end{aligned} \quad (6.59)$$

$$\begin{aligned} B_{51i} &= F_{021} \\ B_{51c0} &= -8F_{020}\varepsilon_{110} - 8F_{000}\varepsilon_{020} + F_{011}\varepsilon_{011} + 5F_{010}^*\varepsilon_{010}^* + 4F_{011,o}F_{020} + 8F_{020,o}F_{011} \\ B_{51c1} &= - \left( \frac{\phi}{\phi_o} \right)^{1/2} [2F_{011}(F_{110} + F_{002}) + F_{010}^*F_{001}^*] \\ B_{51c2} &= \left( \frac{\phi}{\phi_o} \right)^{1/2} [8F_{020}(F_{110} + F_{002}) + 3F_{011}^2 - F_{001}^{*2}] \\ B_{51c3} &= -12 \left( \frac{\phi}{\phi_o} \right)^{1/2} F_{020}F_{011} \end{aligned} \quad (6.60)$$

$$\begin{aligned} B_{52i} &= 4B_{51i} \\ B_{52c0} &= 4[4F_{020}\varepsilon_{110} - 2(2F_{110} + F_{002})\varepsilon_{020} - F_{011}\varepsilon_{011} \\ &\quad + 6F_{020}\varepsilon_{002} - F_{010}^*\varepsilon_{010}^* + 8F_{011,o}F_{020} + 4F_{020,o}F_{011}] \\ B_{52c0} &= 4B_{52c0} \\ B_{52c0} &= 4B_{52c0} \\ B_{52c0} &= 4B_{52c0} \end{aligned} \quad (6.61)$$

$$C_{51i} = 2F_{012}$$

$$C_{51c0} = 2(4F_{011}\varepsilon_{110} + 4F_{020}\varepsilon_{101} - 4F_{101}\varepsilon_{020} - 3F_{002}\varepsilon_{011}$$

$$+ F_{011}\varepsilon_{002} + 3F_{001}^*\varepsilon_{010} + F_{010}^*\varepsilon_{001} + 4F_{002,o}F_{020}$$

$$+ 4F_{011,o}F_{011} + 4F_{020,o}F_{002} - 4F_{010,o}^*F_{010}^*)$$

$$C_{51c1} = -4\left(\frac{\phi}{\phi_o}\right)^{1/2} [F_{101}F_{011} + (2F_{110} + F_{002})F_{002} - F_{100}^*F_{010}^*] \quad (6.62)$$

$$C_{51c2} = 4\left(\frac{\phi}{\phi_o}\right)^{1/2} [4F_{101}F_{020} + 2(F_{110} + 2F_{002})F_{011} - 3F_{010}^*F_{001}^*]$$

$$C_{51c3} = -8\left(\frac{\phi}{\phi_o}\right)^{1/2} (F_{011}^2 + 2F_{020}F_{002} - F_{010}^{*2})$$

$$C_{52i} = 4F_{120}$$

$$C_{52c0} = -2F_{011}\varepsilon_{110} + 12F_{020}\varepsilon_{101} - 4F_{101}\varepsilon_{020} - 2F_{110}\varepsilon_{011}$$

$$- 5F_{001}^*\varepsilon_{010}^* + 3F_{010}^*\varepsilon_{001}^* + 8F_{020,o}F_{110} + 24F_{110,o}F_{020}$$

$$+ 2F_{011,o}F_{011} + 18F_{010,o}^*F_{010}^*$$

$$C_{52c1} = -2\left(\frac{\phi}{\phi_o}\right)^{1/2} (4F_{110}^2 + 2F_{101}F_{011} + 6F_{100}^*F_{010}^* + F_{001}^{*2}) \quad (6.63)$$

$$C_{52c2} = 16\left(\frac{\phi}{\phi_o}\right)^{1/2} (F_{101}F_{020} + F_{110}F_{011} + F_{010}^*F_{001}^*)$$

$$C_{52c3} = -2\left(\frac{\phi}{\phi_o}\right)^{1/2} (16F_{110}F_{020} + F_{110}^2 + 9F_{010}^{*2})$$

$$C_{53i} = C_{51i}$$

$$C_{53c0} = 4[4F_{020}\varepsilon_{101} - (2F_{110} + F_{002})\varepsilon_{011} + F_{011}\varepsilon_{002} - F_{010}^*\varepsilon_{001}^*]$$

$$+ 4F_{002,o}F_{020} + 2F_{011,o}F_{011} - 2F_{010,o}^*F_{010}^*] \quad (6.64)$$

$$C_{53c1} = C_{51c1}$$

$$C_{53c2} = C_{51c2}$$

$$C_{53c3} = C_{51c3}$$

$$\begin{aligned}
D_{51i} &= F_{111} \\
D_{51c0} &= 16F_{020}\varepsilon_{200} + 4F_{110}\varepsilon_{110} + F_{011}\varepsilon_{101} - 3F_{101}\varepsilon_{011} \\
&\quad + 7F_{010}^*\varepsilon_{100}^* + 3F_{100}^*\varepsilon_{010}^* - F_{001}^*\varepsilon_{001}^* + 2F_{011,o}F_{110} \\
&\quad + 4F_{101,o}F_{020} + 4F_{110,o}F_{011} + 2F_{011,o}F_{002} \\
&\quad + 2F_{001,o}F_{010}^* + 3F_{010,o}F_{001}^* \\
D_{51c1} &= -\left(\frac{\phi}{\phi_o}\right)^{1/2} [4F_{011}F_{200} + 2(F_{002} + 3F_{110})F_{101} + 5F_{100}^*F_{001}^*] \tag{6.65} \\
D_{51c2} &= \left(\frac{\phi}{\phi_o}\right)^{1/2} [4(F_{110} + 2F_{002})F_{110} + 16F_{200}F_{020} + 6F_{101}F_{011} + 2F_{100}^*F_{010}^* + 3F_{001}^{*2}] \\
D_{51c3} &= -\left(\frac{\phi}{\phi_o}\right)^{1/2} [4F_{101}F_{020} + 2(3F_{110} + F_{002})F_{011} + 5F_{010}^*F_{001}^*]
\end{aligned}$$

$$\begin{aligned}
D_{52i} &= 3F_{003} \\
D_{52c0} &= 4(2F_{011}\varepsilon_{101} - F_{101}\varepsilon_{011} - F_{002}\varepsilon_{002} - 2F_{010}^*\varepsilon_{100}^* \\
&\quad + F_{100}^*\varepsilon_{010}^* + F_{001}^*\varepsilon_{001}^* + 2F_{002,o}F_{011} + 2F_{011,o}F_{002} \\
&\quad - 2F_{001,o}F_{010}^* - F_{010,o}F_{001}^*) \\
D_{52c1} &= -12\left(\frac{\phi}{\phi_o}\right)^{1/2} (F_{002}F_{101} - F_{001}^*F_{100}^*) \tag{6.66} \\
D_{52c2} &= 12\left(\frac{\phi}{\phi_o}\right)^{1/2} (F_{101}F_{011} + F_{002}^2 - F_{100}^*F_{010}^* - F_{001}^{*2}) \\
D_{52c3} &= -12\left(\frac{\phi}{\phi_o}\right)^{1/2} (F_{011}F_{002} - F_{010}^*F_{001}^*)
\end{aligned}$$

$$\begin{aligned}
D_{53i} &= 2D_{51i} \\
D_{53c0} &= 2[16F_{020}\varepsilon_{002} - 2(2F_{110} + F_{002})\varepsilon_{110} + 3F_{011}\varepsilon_{101} \\
&\quad - F_{101}\varepsilon_{011} - 2F_{110}\varepsilon_{002} + F_{010}^*\varepsilon_{100}^* - 3F_{100}^*\varepsilon_{010}^* \\
&\quad - F_{001}^*\varepsilon_{001}^* + 2F_{001,o}F_{110} + 4F_{101,o}F_{020} \\
&\quad + 2(2F_{110,o} + F_{002,o})F_{011} + 4F_{001,o}F_{010}^* + F_{010,o}F_{001}^*] \tag{6.67}
\end{aligned}$$

$$D_{53c1} = 2D_{51c1}$$

$$D_{53c2} = 2D_{51c2}$$

$$D_{53c3} = 2D_{51c3}$$

$$E_{51i} = 2F_{102}$$

$$\begin{aligned} E_{51c0} &= 2[8F_{011}\varepsilon_{200} - 2F_{101}\varepsilon_{110} + (2F_{110} + F_{002})\varepsilon_{101} \\ &\quad - 3F_{101}\varepsilon_{002} + 2F_{100}^*\varepsilon_{001}^* + 2F_{002,o}F_{110} + 2F_{101,o}F_{011} \\ &\quad + 2(F_{110,o} + F_{002,o})F_{002} - 2F_{100,o}^*F_{010}^*] \\ E_{51c1} &= -8\left(\frac{\phi}{\phi_o}\right)^{1/2} (F_{101}^2 + 2F_{200}F_{002} - F_{100}^{*2}) \end{aligned} \quad (6.68)$$

$$\begin{aligned} E_{51c2} &= 4\left(\frac{\phi}{\phi_o}\right)^{1/2} [2(2F_{002} + F_{110})F_{101} + 4F_{200}F_{011} - 3F_{100}^*F_{001}^*] \\ E_{51c3} &= -4\left(\frac{\phi}{\phi_o}\right)^{1/2} [F_{101}F_{011} + (2F_{110} + F_{002})F_{002} - F_{100}^*F_{010}^*] \end{aligned}$$

$$E_{52i} = 2F_{210}$$

$$\begin{aligned} E_{52c0} &= 8F_{011}\varepsilon_{200} - 2F_{101}\varepsilon_{110} - 2F_{110}\varepsilon_{101} + F_{001}^*\varepsilon_{100}^* \\ &\quad - 3F_{100}^*\varepsilon_{001}^* + F_{110,o}F_{110} + 2F_{101,o}F_{011} \\ &\quad + 6F_{100,o}^*F_{010}^* + F_{001,o}^*F_{001}^* \\ E_{52c1} &= -\left(\frac{\phi}{\phi_o}\right)^{1/2} (16F_{200}F_{110} + F_{101}^2 + 9F_{100}^{*2}) \end{aligned} \quad (6.69)$$

$$\begin{aligned} E_{52c2} &= 8\left(\frac{\phi}{\phi_o}\right)^{1/2} (F_{110}F_{101} + F_{200}F_{011} + F_{100}^*F_{001}^*) \\ E_{52c3} &= -\left(\frac{\phi}{\phi_o}\right)^{1/2} (4F_{110}^2 + 2F_{101}F_{011} + 6F_{100}^*F_{010}^* + F_{001}^{*2}) \end{aligned}$$

$$F_{5i} = F_{201}$$

$$\begin{aligned} F_{5c0} &= 8(F_{110} + F_{002})\varepsilon_{200} - 3F_{101}\varepsilon_{101} + F_{100}^*\varepsilon_{100}^* \\ &\quad + 2F_{101,o}F_{110} + 2F_{101,o}F_{002} + F_{100,o}^*F_{001}^* \\ F_{5c1} &= -12\left(\frac{\phi}{\phi_o}\right)^{1/2} F_{200}F_{101} \end{aligned} \quad (6.70)$$

$$\begin{aligned} F_{5c2} &= \left(\frac{\phi}{\phi_o}\right)^{1/2} [8(F_{110} + F_{002})F_{200} + 3F_{101}^2 - F_{100}^{*2}] \\ F_{5c3} &= -\left(\frac{\phi}{\phi_o}\right)^{1/2} [2(F_{110} + F_{002})F_{101} + F_{100}^*F_{001}^*] \end{aligned}$$

*Anisotropic aberration coefficients*

$$\begin{aligned} a_{5i} &= 0 \\ a_{5c0} &= -4(F_{010}^* \varepsilon_{020} + F_{020}^* \varepsilon_{010}) \end{aligned} \quad (6.71)$$

$$a_{5c1} = 0$$

$$a_{5c2} = 0$$

$$a_{5c3} = 0$$

$$b_{51i} = 5F_{020}^*$$

$$\begin{aligned} b_{51c0} &= -8F_{001}^* \varepsilon_{020} + 3F_{010}^* \varepsilon_{011} - 7F_{011}^* \varepsilon_{010} + 12F_{020}^* \varepsilon_{001} \\ &\quad + 36F_{010,o}^* F_{020} + 24F_{020,o}^* F_{010} \end{aligned}$$

$$b_{51c1} = -5\left(\frac{\phi}{\phi_o}\right)^{1/2} (2F_{110}^* F_{010} + F_{011}^* F_{001}) \quad (6.72)$$

$$b_{51c2} = 20\left(\frac{\phi}{\phi_o}\right)^{1/2} (F_{011}^* F_{010} + F_{020}^* F_{001})$$

$$b_{51c3} = -60\left(\frac{\phi}{\phi_o}\right)^{1/2} F_{020}^* F_{010}$$

$$b_{52i} = -\frac{4}{5} b_{51i}$$

$$b_{52c0} = 4(-F_{010}^* \varepsilon_{011} + F_{011}^* \varepsilon_{010} - 4F_{020}^* \varepsilon_{001} - 8F_{010,o}^* F_{020} - 4F_{020,o}^* F_{010})$$

$$b_{52c1} = -\frac{4}{5} b_{51c1} \quad (6.73)$$

$$b_{52c2} = -\frac{4}{5} b_{51c2}$$

$$b_{52c3} = -\frac{4}{5} b_{51c3}$$

$$c_{51i} = 4F_{011}^*$$

$$\begin{aligned} c_{51c0} &= 2[4F_{010}^* \varepsilon_{110} - 4F_{100}^* \varepsilon_{020} - 3F_{001}^* \varepsilon_{011} + 3F_{010}^* \varepsilon_{002} \\ &\quad + 12F_{020}^* \varepsilon_{100} - (4F_{110} + 5F_{002}) \varepsilon_{010} + F_{011}^* \varepsilon_{001} \\ &\quad + 12F_{001,o}^* F_{020} + 8F_{010,o}^* F_{011} + 8F_{011,o}^* F_{010} + 4F_{020,o}^* F_{001}] \end{aligned} \quad (6.74)$$

$$c_{51c1} = -8\left(\frac{\phi}{\phi_o}\right)^{1/2} [F_{011}^* F_{100} + F_{101}^* F_{010} + (2F_{110} + F_{002}) F_{001}]$$

$$c_{51c2} = 8\left(\frac{\phi}{\phi_o}\right)^{1/2} [4F_{020}^* F_{100} + 2(F_{110} + 2F_{002}) F_{010}^* + 3F_{011}^* F_{001}]$$

$$c_{51c3} = -32\left(\frac{\phi}{\phi_o}\right)^{1/2} (F_{011}^* F_{010} + F_{020}^* F_{001})$$

$$\begin{aligned}
c_{52i} &= -\frac{1}{4} c_{51i} \\
c_{52c0} &= -6F_{010}^* \epsilon_{110} - 4F_{100}^* \epsilon_{020} + F_{001}^* \epsilon_{011} - 12F_{020} \epsilon_{100}^* \\
&\quad - 2(F_{110} - 2F_{002}) \epsilon_{010}^* - F_{011} \epsilon_{001}^* - 4F_{001,o}^* F_{020} \\
&\quad - 6F_{010,o}^* F_{011} - 2F_{011,o}^* F_{010} - 4F_{020,o}^* F_{001} \tag{6.75}
\end{aligned}$$

$$c_{52c1} = -\frac{1}{4} c_{51c1}$$

$$c_{52c2} = -\frac{1}{4} c_{51c2}$$

$$c_{52c3} = -\frac{1}{4} c_{51c3}$$

$$\begin{aligned}
c_{53i} &= -\frac{1}{2} c_{51i} \\
c_{53c0} &= 4[-F_{010}^* \epsilon_{002} - 4F_{020} \epsilon_{100}^* + (2F_{110} + F_{002}) \epsilon_{010}^* \\
&\quad - F_{011} \epsilon_{001}^* - 4F_{001,o}^* F_{020} - 2F_{010,o}^* F_{011} - 2F_{011,o}^* F_{010}] \tag{6.76}
\end{aligned}$$

$$c_{53c1} = -\frac{1}{2} c_{51c1}$$

$$c_{53c2} = -\frac{1}{2} c_{51c2}$$

$$c_{53c3} = -\frac{1}{2} c_{51c3}$$

$$d_{51i} = 3F_{110}^*$$

$$\begin{aligned}
d_{51c0} &= -2F_{001}^* \epsilon_{110} + 3F_{010}^* \epsilon_{011} - 3F_{100}^* \epsilon_{011} + 3F_{011} \epsilon_{100}^* \\
&\quad - 3F_{101} \epsilon_{010}^* - 2F_{110} \epsilon_{001}^* + 6F_{010,o}^* F_{110} + 12F_{100,o}^* F_{020} \\
&\quad + 2F_{001,o}^* F_{011} + 12F_{110,o}^* F_{010} + F_{011,o}^* F_{001} \tag{6.77}
\end{aligned}$$

$$d_{51c1} = -3 \left( \frac{\phi}{\phi_o} \right)^{1/2} (6F_{110} F_{100}^* + 4F_{200} F_{010}^* + F_{101} F_{001}^*)$$

$$d_{51c2} = 12 \left( \frac{\phi}{\phi_o} \right)^{1/2} (F_{011} F_{100}^* + F_{101} F_{010}^* + F_{110} F_{001}^*)$$

$$d_{51c3} = -3 \left( \frac{\phi}{\phi_o} \right)^{1/2} (4F_{020} F_{100}^* + 6F_{110} F_{010}^* + F_{011} F_{001}^*)$$

$$d_{52i} = 3F_{002}^*$$

$$\begin{aligned}
d_{52c0} &= 4(2F_{010}^* \epsilon_{101} - F_{100}^* \epsilon_{011} - F_{001}^* \epsilon_{002} + 2F_{011} \epsilon_{100}^* \\
&\quad - F_{101} \epsilon_{010}^* - F_{002} \epsilon_{001}^* + 2F_{001,o}^* F_{011} + F_{010,o}^* F_{002} \\
&\quad + 2F_{002,o}^* F_{010} + F_{011,o}^* F_{001}) \tag{6.77}
\end{aligned}$$

$$d_{52c1} = -12 \left( \frac{\phi}{\phi_o} \right)^{1/2} (F_{002}F_{100}^* + F_{101}F_{001}^*)$$

$$d_{52c2} = 12 \left( \frac{\phi}{\phi_o} \right)^{1/2} (F_{011}F_{100}^* + F_{101}F_{010}^* + 2F_{002}F_{001}^*) \quad (6.78)$$

$$d_{52c3} = -12 \left( \frac{\phi}{\phi_o} \right)^{1/2} (F_{002}F_{010}^* + F_{011}F_{001}^*)$$

$$d_{53i} = -\frac{2}{3}(d_{51i} + d_{52i})$$

$$\begin{aligned} d_{53c0} = & 2[-3F_{010}^*e_{101} - F_{100}^*e_{011} + F_{001}^*e_{002} - 7F_{011}e_{100}^* \\ & + 3F_{101}e_{010}^* + F_{002}e_{001}^* - 2F_{010,o}^*F_{110} \\ & - 4F_{100,o}^*F_{020} - 4F_{001,o}^*F_{011} - 2F_{010,o}^*F_{002} \end{aligned} \quad (6.79)$$

$$- 2(2F_{110,o} + F_{002,o})F_{010}^* - F_{011,o}F_{001}^*]$$

$$d_{53c1} = -\frac{2}{3}(d_{51c1} + d_{52c1})$$

$$d_{53c2} = -\frac{2}{3}(d_{51c2} + d_{52c2})$$

$$d_{53c3} = -\frac{2}{3}(d_{51c1} + d_{51c2})$$

$$e_{51i} = 2F_{101}^*$$

$$\begin{aligned} e_{51c0} = & 2[8F_{010}^*e_{200} - 2F_{100}^*e_{110} - 3F_{100}^*e_{002} + (2F_{110} \\ & + F_{002})e_{100}^* - 2F_{101}e_{001}^* + 2F_{001,o}^*F_{110} + 2F_{100,o}^*F_{011} \\ & + F_{001,o}^*F_{002} + 2F_{101,o}^*F_{010}^* + (F_{110,o} + F_{002,o})F_{001}^*] \end{aligned}$$

$$e_{51c1} = -16 \left( \frac{\phi}{\phi_o} \right)^{1/2} (F_{101}F_{100}^* + F_{200}F_{001}^*) \quad (6.80)$$

$$e_{51c2} = 4 \left( \frac{\phi}{\phi_o} \right)^{1/2} [2(F_{110} + 2F_{002})F_{100}^* + 4F_{200}F_{010}^* + 3F_{101}F_{001}^*]$$

$$e_{51c3} = -4 \left( \frac{\phi}{\phi_o} \right)^{1/2} (4F_{011}F_{100}^* + F_{101}F_{010}^* + (2F_{110} + F_{002})F_{001}^*)$$

$$e_{52i} = -\frac{1}{2}e_{51i}$$

$$\begin{aligned} e_{52c0} = & -8F_{010}^*e_{200} - 2F_{100}^*e_{110} + F_{001}^*e_{101} \\ & - 2(3F_{110} + 2F_{002})e_{100}^* + 3F_{101}e_{001}^* - 2F_{001,o}^*F_{110} \\ & - 2F_{100,o}^*F_{011} - F_{001,o}^*F_{002} - 2F_{101,o}^*F_{010}^* - 2F_{110,o}^*F_{001}^* \end{aligned}$$

$$\begin{aligned}
e_{52c1} &= -\frac{1}{2} e_{51c1} \\
e_{52c2} &= -\frac{1}{2} e_{51c2} \\
e_{52c3} &= -\frac{1}{2} e_{51c3}
\end{aligned} \tag{6.81}$$

$$\begin{aligned}
f_{5i} &= F_{200}^* \\
f_{5c0} &= 4F_{001}^* e_{200} - 3F_{100}^* e_{101} - F_{101} e_{100}^* + 2F_{100,o}^* F_{110} + F_{101,o} F_{001}^* \\
f_{5c1} &= -12 \left( \frac{\phi}{\phi_o} \right)^{1/2} F_{200} F_{100}^* \\
f_{5c2} &= 4 \left( \frac{\phi}{\phi_o} \right)^{1/2} (F_{101} F_{100}^* + F_{200} F_{001}^*) \\
f_{5c3} &= - \left( \frac{\phi}{\phi_o} \right)^{1/2} (2F_{110} F_{100}^* + F_{101} F_{001}^*)
\end{aligned} \tag{6.82}$$

In these equations we have written

$$\begin{aligned}
e_{lmn} &= \int_{z_o}^z F_{lmn} dz \quad \begin{cases} l,m,n = 0,1,2 & l+m+n = 2 \\ l,m,n = 0,1,2,3 & l+m+n = 3 \end{cases} \\
e_{lmn}^* &= \int_{z_o}^z F_{lmn}^* dz \quad \begin{cases} l,m,n = 0,1 & l+m+n = 1 \\ l,m,n = 0,1,2 & l+m+n = 2 \end{cases}
\end{aligned} \tag{6.83}$$

The quantities  $F_{lmn}$  are as follows:

$$\begin{aligned}
F_{200} &= F_{LMN}(g^4, g^2 g'^2, g'^4) \\
F_{110} &= 2F_{LMN} \left( h^2 g^2, \frac{h^2 g'^2 + h'^2 g^2}{2}, h'^2 g'^2 \right) - K\phi_o \\
F_{101} &= 4F_{LMN} \left( hg^3, \frac{hgg'^2 + h'g'g^2}{2}, h'g'^3 \right)
\end{aligned} \tag{6.84}$$

$$\begin{aligned}
F_{020} &= F_{LMN}(h^4, h^2 h'^2, h'^4) \\
F_{011} &= 4F_{LMN} \left( h^3 g, \frac{hgh'^2 + h'g'h^2}{2}, h'^3 g' \right) \\
F_{002} &= 4F_{LMN}(h^2 g^2, hg h' g', h'^2 g'^2) + K\phi_o \\
F_{100}^* &= F_{PQ}(g^2, g'^2) \\
F_{010}^* &= F_{PQ}(h^2, h'^2) \\
F_{001}^* &= 2F_{PQ}(gh, g'h')
\end{aligned} \tag{6.85}$$

$$\begin{aligned}
F_{300} &= F_L(g^6, g^4g'^2, g^2g'^4, g'^6) \\
F_{210} &= 3F_L \left[ g^4h^2, \frac{g^2(g^2h'^2 + 2g'^2h^2)}{3}, \frac{g'^2(g'^2h^2 + 2g^2h'^2)}{3}, g'^4h'^2 \right] + F_N(g^2, g'^2) \\
F_{201} &= 6F_L \left[ g^5h, \frac{g^3g'(gh' + 2gh')}{3}, \frac{gg'^3(g'h + 2gh')}{3}, g'^5h' \right] \\
F_{120} &= 3F_L \left[ g^2h^4, \frac{h^2(h^2g'^2 + 2g^2h'^2)}{3}, \frac{h'^2(g^2h'^2 + 2g'^2h^2)}{3} g'^2h'^4 \right] + F_N(h^2, h'^2) \\
F_{111} &= 12F_L \left[ g^3h^3, \frac{gh(h^2g'^2 + ghg'h' + g^2h'^2)}{3}, \frac{g'h'(g'^2h^2 + ghh'g' + g^2h'^2)}{3}, h'^3g'^3 \right] \\
&\quad + 2F_N(gh, g'h') \tag{6.86}
\end{aligned}$$

$$\begin{aligned}
F_{102} &= 12F_L \left[ g^4h^2, \frac{g^2hg'(g'h + 2gh')}{3}, \frac{gh'g'^2(gh' + 2g'h)}{3}, g'^4h'^2 \right] - F_N(g^2, g'^2) \\
F_{030} &= F_L(h^6, h^4h'^2, h^2h'^4, h'^6) \\
F_{021} &= 6F_L \left[ gh^5, \frac{h^3h'(g'h + 2gh')}{3}, \frac{hh'^3(gh' + 2g'h)}{3}, g'h'^5 \right] \\
F_{012} &= 12F_L \left[ g^2h^4, \frac{gh^2h'(gh' + 2g'h)}{3}, \frac{g'hh'^2(g'h + 2gh')}{3}, h'^4g'^2 \right] - F_N(h^2, h'^2) \\
F_{003} &= 8F_L(g^3h^3, g^2h^2g'h', ghg'^2h'^2, g'^3h'^3) - 2F_N(gh, g'h') \\
F_{200}^* &= F_M(g^4, g^2g'^2, g'^4) \\
F_{110}^* &= 2F_M \left( h^2g^2, \frac{h^2g'^2 + h'^2g^2}{2}, h'^2g'^2 \right) - K_1\phi_o^{3/2} \\
F_{101}^* &= 4F_M \left( hg^3, \frac{ghg'^2 + g'h'g^2}{2}, h'g'^3 \right) \tag{6.87} \\
F_{020}^* &= F_M(h^4, h^2h'^2, h'^4) \\
F_{011}^* &= 4F_M \left( h^3g, \frac{ghh'^2 + g'h'h^2}{2}, h'^3g' \right) \\
F_{002}^* &= 4F_M(g^2h^2, ghg'h', g'^2h'^2) + K_1\phi_o^{3/2}
\end{aligned}$$

Here

$$\begin{aligned}
F_{LMN} &= -\frac{L_1}{4}x_1 - \frac{L_2}{4}x_2 - \frac{L_3}{4}x_3 \\
F_{PQ} &= \phi_o^{1/2}(-Px_1 - Qx_2) \\
F_L &= -L_4x_1 - L_5x_2 - L_6x_3 - L_7x_4 \\
F_M &= \phi_o^{1/2}(-M_1x_1 - M_2x_2 - M_3x_3) \\
F_N &= \phi_o(-N_1x_1 - N_2x_2) \tag{6.88}
\end{aligned}$$

in which

$$\begin{aligned}
L_4 &= -\frac{1}{9216\phi^{5/2}}[9\eta^6B^6 + 9\eta^4\phi''B^4 - 9\eta^2(\phi''^2 - \phi\phi^{(4)})B^2 \\
&\quad - 12\eta^2\phi^2B^{(4)}B - 9\phi''^3 + 9\phi\phi''\phi^{(4)} - 2\phi^2\phi^{(6)}] \\
L_5 &= -\frac{3\eta^4B^4 + 2\eta^2\phi''B^2 - \phi''^2 + \phi\phi^{(4)}}{256\phi^{3/2}} \\
L_6 &= -\frac{3\eta^2B^2 + \phi''}{64\phi^{1/2}} \\
L_7 &= -\frac{\phi^{1/2}}{16} \\
M_1 &= -\frac{\eta[9\eta^4B^5 + 6\eta^2\phi''B^3 - 3(\phi''^2 - \phi\phi^{(4)})B - 2\phi^2B^4]}{768\phi^{5/2}} \tag{6.89} \\
M_2 &= -\frac{\eta B(3\eta^2B^2 + \phi\phi'')}{32\phi^{3/2}} \\
M_3 &= -\frac{3\eta B}{16\phi^{1/2}} \\
N_1 &= -\frac{\eta^2B^2(3\eta^2B^2 + \phi\phi'')}{64\phi^{5/2}} \\
N_2 &= -\frac{3\eta^2B^2}{16\phi^{3/2}} \\
K_1 &= -\frac{\eta^3B^3}{16\phi^{5/2}}
\end{aligned}$$

The definitions of  $L_1$ ,  $L_2$ ,  $L_3$ ,  $P$ , and  $Q$  are given in Equations 6.34 and 6.35. Finally, we note that  $F_{lmn,o}$  denotes the value of  $F_{lmn}$  at the object plane; for  $F_{200}$ , for example, we have

$$F_{200,o} = F_{200} \Big|_{z=z_o} = -\frac{L}{4} \Big|_{z=z_o} = -\frac{1}{128\phi_o^{1/2}} \left\{ \frac{(\phi_o'' + \eta^2B_o^2)^2}{\phi_o} - \phi_o^{(4)} - 4\eta^2B_oB_o'' \right\} \tag{6.90}$$

We now turn to the *asymptotic* coefficients. Expressed in terms of position and gradient in the (asymptotic) object plane, these coefficients may be written as

$$\begin{pmatrix} C \\ K \\ A \\ F \\ D \end{pmatrix} = \begin{pmatrix} C_4 & C_3 & C_2 & C_1 & C_0 \\ 0 & K_3 & K_2 & K_1 & K_0 \\ 0 & 0 & A_2 & A_1 & A_0 \\ 0 & 0 & F_2 & F_1 & F_0 \\ 0 & 0 & 0 & D_1 & D_0 \end{pmatrix} \begin{pmatrix} m^4 \\ m^3 \\ m^2 \\ m \\ 1 \end{pmatrix} \tag{6.91}$$

$$\begin{pmatrix} k \\ a \\ d \end{pmatrix} = \begin{pmatrix} k_2 & k_1 & k_0 \\ 0 & a_1 & a_0 \\ 0 & 0 & d_0 \end{pmatrix} \begin{pmatrix} m^2 \\ m \\ 1 \end{pmatrix} \quad (6.92)$$

in which  $m$  again denotes reciprocal magnification ( $m = 1/M$ ). The coefficients  $C_i$ ,  $K_i$ ,  $A_i$ ,  $F_i$ ,  $D_i$ ,  $k_i$ ,  $a_i$ , and  $d_0$  are independent of object position and hence of magnification; they are characteristics of the lens alone, like its focal length. They, in turn, may be expressed in terms of a small number of integrals as follows:

### Spherical aberration

$$C = C_4m^4 + C_3m^3 + C_2m^2 + C_1m + C_0 \quad (6.93)$$

where

$$\begin{aligned} C_4 &= i_1 f_o^4 \\ C_3 &= -4i_2 f_o^4 - r^2 f_o / 2 \\ C_2 &= 2(i_3 + i_4) f_o^4 \\ C_1 &= -4i_5 f_o^4 - f_o / 2 \\ C_0 &= i_o f_o^4 \end{aligned} \quad (6.94)$$

### Coma

$$K = K_3m^3 + K_2m^2 + K_1m + K_0 \quad (6.95)$$

where

$$\begin{aligned} K_3 &= -i_1 f_o^3 \\ K_2 &= 3i_2 f_o^3 + r^2 / 2 \\ K_1 &= -(i_3 + i_4) f_o^3 \\ K_0 &= i_5 f_o^3 \end{aligned} \quad (6.96)$$

For the anisotropic components

$$k = k_2m^2 + k_1m + k_0 \quad (6.97)$$

where

$$\begin{aligned} k_2 &= i_7 f_o^2 \\ k_1 &= -i_8 f_o^2 \\ k_0 &= i_9 f_o^2 \end{aligned} \quad (6.98)$$

*Astigmatism and field curvature*

$$A = A_2 m^2 + A_1 m + A_0 \quad (6.99)$$

$$F = F_2 m^2 + F_1 m + F_0 \quad (6.100)$$

where

$$\begin{aligned} A_2 &= F_2/2 = i_1 f_o^2 \\ A_1 &= F_1/2 = -2i_2 f_o^2 - r^2/2f_o \\ A_0 &= i_3 f_o^2 \quad F_0 = i_4 f_o^2 \end{aligned} \quad (6.101)$$

For the anisotropic components

$$a = a_1 m + a_0 \quad (6.102)$$

where

$$\begin{aligned} a_1 &= -2i_7 f_o \\ a_0 &= i_8 f_o \end{aligned} \quad (6.103)$$

*Distortion*

$$D = D_1 m + D_0 \quad (6.104)$$

where

$$\begin{aligned} D_1 &= -i_1 f_o \\ D_0 &= i_2 f_o + \frac{1}{2f_i^2} \end{aligned} \quad (6.105)$$

For the anisotropic component

$$d = d_o = i_7 \quad (6.106)$$

where

$$r = \left( \frac{\hat{\phi}_o}{\hat{\phi}_i} \right)^{1/2} = \frac{f_o}{f_i} \quad (6.107)$$

More compactly, we may write

$$\begin{pmatrix} C \\ K \\ A \\ F \\ D \end{pmatrix} = Q \begin{pmatrix} m^4 \\ m^3 \\ m^2 \\ m \\ 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} k \\ a \\ d \end{pmatrix} = q \begin{pmatrix} m^2 \\ m \\ 1 \end{pmatrix} \quad (6.108)$$

in which

$$Q = \begin{pmatrix} i_1 f_o^4 & -4i_2 f_o^4 - \frac{1}{2} r^2 f_o & 2(i_3 + i_4) f_o^4 & -4i_5 f_o^4 - \frac{1}{2} f_o & i_6 f_o^4 \\ 0 & -i_1 f_o^3 & 3i_2 f_o^3 + \frac{1}{2} r^2 & -(i_3 + i_4) f_o^3 & i_5 f_o^3 \\ 0 & 0 & i_1 f_o^2 & -2i_2 f_o^2 - \frac{r}{2f_o} & i_3 f_o^2 \\ 0 & 0 & 2i_1 f_o^2 & -4i_2 f_o^2 - \frac{r}{f_o} & i_4 f_o^2 \\ 0 & 0 & 0 & -i_1 f_o & i_2 f_o + \frac{1}{2f_i^2} \end{pmatrix} \quad (6.109)$$

and

$$q = \begin{pmatrix} i_7 f_o^2 & -i_8 f_o^2 & i_9 f_o^2 \\ 0 & -2i_7 f_o^2 & i_8 f_o \\ 0 & 0 & i_7 \end{pmatrix} \quad (6.110)$$

The functions  $i_1$  through  $i_9$  denote the following integrals:

*Electrostatic lenses (nonrelativistic expressions)*

$$\begin{aligned} i_1 &= \int \Lambda_e G^4 dz \\ i_2 &= \int \Lambda_e G^3 \bar{G} dz - \frac{1}{8f_o f_i^2} \\ i_3 &= \int \Lambda_e G^2 \bar{G}^2 dz - \frac{\phi_o^{1/2}}{24f_o^2} \int \frac{\phi''}{\phi^{3/2}} dz \\ i_4 &= 2 \int \Lambda_e G^2 \bar{G}^2 dz - \frac{\phi_o^{1/2}}{24f_o^2} \int \frac{\phi''}{\phi^{3/2}} dz \\ i_5 &= \int \Lambda_e G \bar{G}^3 dz - \frac{1}{8f_o^3} \\ i_6 &= \int \Lambda_e \bar{G}^4 dz \\ i_7 = i_8 = i_9 &= 0 \end{aligned} \quad (6.111)$$

*Magnetic lenses*

$$\begin{aligned}
i_1 &= \int \Lambda_m G^4 dz \\
i_2 &= \int \Lambda_m G^3 \bar{G} dz - \frac{1}{8f^3} \\
i_3 &= \int \Lambda_m G^2 \bar{G}^2 dz - \frac{1}{6f^2} \int \frac{\eta^2 B^2}{\hat{\phi}} dz \\
i_4 &= 2 \int \Lambda_m G^2 \bar{G}^2 dz + \frac{1}{6f^2} \int \frac{\eta^2 B^2}{\hat{\phi}} dz \\
i_5 &= \int \Lambda_m G \bar{G}^3 dz - \frac{1}{8f^3} \\
i_6 &= \int \Lambda_m \bar{G}^4 dz \\
i_7 &= \int \Lambda_a G^2 dz \\
i_8 &= 2 \int \Lambda_a G \bar{G} dz \\
i_9 &= \int \Lambda_a \bar{G}^2 dz
\end{aligned} \tag{6.112}$$

The functions  $\Lambda_e$ ,  $\Lambda_m$ , and  $\Lambda_a$  are as follows:

$$\Lambda_e = \frac{1}{64} \left( \frac{\phi}{\phi_o} \right)^{1/2} \left[ 4 \left( \frac{\phi''}{\phi} \right)^2 - \frac{\phi' \phi'''}{\phi^2} - 10 \left( \frac{\phi'}{\phi} \right)^2 \frac{\phi''}{\phi} + 10 \left( \frac{\phi'}{\phi} \right)^4 \right] \tag{6.113}$$

$$\Lambda_m = \frac{1}{48} \left( \frac{4\eta^4 B^4}{\hat{\phi}^2} + 5 \frac{\eta^2 B'^2}{\hat{\phi}} - \frac{\eta^2 BB''}{\hat{\phi}} \right) \tag{6.114}$$

$$\Lambda_a = \frac{1}{16} \left( \frac{\eta B''}{\hat{\phi}^{1/2}} + \frac{2\eta^3 B^3}{\hat{\phi}^{3/2}} \right) \tag{6.115}$$

$G(z)$  and  $\bar{G}(z)$  are the solutions of the paraxial equation routinely used in the study of asymptotic properties. They are the rays incident on the lens parallel to the axis on the object and image sides at unit distance from the axis:

$$\lim_{z \rightarrow -\infty} G(z) = 1 \quad \lim_{z \rightarrow \infty} \bar{G}(z) = 1 \tag{6.116}$$

For symmetric lenses, with a suitable choice of origin,  $G(z) = \bar{G}(-z)$ , and three of the integrals are no longer independent.

$$i_1 = i_6 \quad i_2 = i_5 \quad \text{and} \quad i_7 = i_9 \tag{6.117}$$

Ohye et al. (2006) give relativistic forms of  $C_4$ ,  $C_3$ ,  $C_2$ ,  $C_1$ , and  $C_0$  for electrostatic lenses together with a few numerical values.

Expressions of this type have the advantage that the effect of system symmetry is easily seen, essentially thanks to the choice of the paraxial solutions  $G(z)$  and  $\bar{G}(z)$ . We shall see that the design

of more complicated systems, notably imaging energy analyzers with an  $\Omega$ -shaped optic axis, is facilitated by a similar choice.

The distortion has a particularly simple structure, and in the case of high magnification, we simply have  $D = D_0$ . We have seen that it is necessary to introduce a rotating coordinate system to uncouple the paraxial equations. A further example of this rotational effect is that magnetic lenses suffer from three anisotropic aberrations in addition to the five isotropic aberrations of electrostatic lenses, which are already familiar from the theory of glass lenses. The *anisotropic* distortion is, as shown by Equation 6.106, independent of the magnification. For comment on the elimination of distortion, see Lambrakis et al. (1977), Marai and Mulvey (1977), and Tsuno and Harada (1981a,b).

The aberration coefficients can be written in many different ways, and we draw particular attention to the expressions published by Lencová and Lenc (1990, 1994), which lend themselves well to programming. The authors give formulas for all the geometric and chromatic aberrations of magnetic (1990) and electrostatic (1994) round lenses in an arbitrary plane and in the Gaussian image plane in terms of a small number of generating functions. Aberrations of position and gradient are both considered, and systems with and without a limiting aperture are likewise both envisaged. All these formulas are very space-consuming and cannot be included here; they are, however, easily found. To convince the reader of their convenience, we reproduce the generating functions, in terms of which the aberrations are expressed, and the geometrical aberrations expressed in terms of them. For magnetic lenses, two integrals,  $I_1$  and  $I_2$ , are required for the geometric aberrations and, anticipating the next section, one integral  $C(p, q)$  for the chromatic aberrations.

$$\begin{aligned} I_1(p, q, r, s) &= \frac{1}{2} \int_{z_o}^z \left\{ (p'q' - b^2 pq)(r's' - b^2 rs) + (bpq)'(brs)' \right\} d\zeta \\ I_2(p, q) &= \frac{1}{2} \int_{z_o}^z \left\{ b(p'q' + b^2 pq) + \frac{1}{4} b'(pq)' - \frac{1}{4} (b'pq)' \right\} d\zeta \\ C(p, q) &= - \int_{z_o}^z b^2 pq d\zeta \end{aligned} \quad (6.118)$$

The symbols  $p$ ,  $q$ ,  $r$ , and  $s$  may be  $g(z)$  or  $h(z)$  and  $b(z)$  denotes  $\eta B(z)/2\phi^{1/2}$ . For electrostatic lenses, the generating functions are

$$\begin{aligned} I_1(p, q, r, s) &= \frac{1}{2\hat{\phi}_o^{1/2}} \int_{z_o}^z \left[ \left( p'q' - \frac{\gamma\phi''}{8\hat{\phi}} pq \right) \left( r's' - \frac{\gamma\phi''}{8\hat{\phi}} rs \right) \right. \\ &\quad \left. + \frac{2+\gamma^2}{192} \left[ 3\frac{\phi''}{\phi} pq + 2\frac{\phi'}{\hat{\phi}}(pq)' \right] \left[ 3\frac{\phi''}{\hat{\phi}} rs + 2\frac{\phi'}{\hat{\phi}}(rs)' \right] \right. \\ &\quad \left. - \frac{2+\gamma^2}{48} \left( \frac{\phi'}{\hat{\phi}} \right)^2 (pq)'(rs)' - \frac{3\gamma\phi'^2\phi''}{64\hat{\phi}^3} pqrs \right] d\zeta \end{aligned} \quad (6.119)$$

$$\begin{aligned} I_2(p, q) &= \int_{z_o}^z \left( \frac{\gamma\phi''}{32\hat{\phi}} pq \right)' d\zeta \\ C(p, q) &= -\frac{1}{2} \hat{\phi}_o^{1/2} \int_{z_o}^z \frac{1}{\hat{\phi}^{1/2}} \left( \gamma p'q' + \frac{\phi''}{\hat{\phi}} pq \right) d\zeta \end{aligned}$$

In terms of these generating functions, the aberration coefficients are as follows:

$$\begin{aligned}
 C_s &= I_1(h, h, h, h) + C_{se} \\
 K &= I_1(h, g, h, h) + K_e \\
 F &= I_1(h, g, h, g) + I_1(h, h, g, g) + \{F_e \text{ or } F_m\} \\
 A &= I_1(h, h, g, g) + A_e \\
 D &= I_1(h, g, g, g) + D_e \\
 E &= I_1(g, g, g, g) + E_e \\
 k &= I_2(h, h) \\
 a &= 2I_2(g, h) \\
 d &= I_2(g, g)
 \end{aligned} \tag{6.120}$$

in which the following terms are to be retained only for electrostatic lenses:

$$\begin{aligned}
 C_{se} &= h_i'^4 \frac{f_i}{2f_o} (z_i - z_2) + \frac{1}{2} (z_l - z_o) \\
 K_e &= -h_i'^3 \frac{1}{2f_o} (z_i - z_2) \\
 F_e &= h_i'^2 \frac{1}{f_i f_o} (z_i - z_2) + \frac{\gamma \phi_o^{1/2}}{16} \int_{z_1}^{z_2} \frac{\phi''}{\hat{\phi}^{3/2}} dz \\
 A_e &= h_i'^2 \frac{1}{2f_i f_o} (z_i - z_2) \\
 D_e &= -h_i' \frac{1}{2f_i^2 f_o} (z_i - z_2) \\
 E_e &= \frac{1}{2f_i^3 f_o} (z_i - z_2)
 \end{aligned} \tag{6.121a}$$

and

$$F_m = \frac{\eta^2}{4\hat{\phi}} \int_{z_o}^{\hat{z}} B^2 d\zeta \tag{6.121b}$$

These are the real coefficients for magnetic lenses and the asymptotic coefficients for electrostatic lenses. For the latter, the field is assumed to be negligible outside some entry plane ( $z = z_1$ ) and some exit plane ( $z = z_2$ ). Similar relations are given in the articles of Lencová and Lenc for the aberrations of gradient, the chromatic aberrations, and numerous other cases.

### Chromatic aberration

The aberration  $\Delta u_i$  is now given by

$$\Delta u_i = \{C_c u'_o + (C_D + iC_\theta)u_o\} \left( 2 \frac{\Delta B_0}{B_0} - \frac{\Delta \hat{\phi}_0}{\hat{\phi}_0} \right) \quad (6.122)$$

for magnetic lenses, and

$$\Delta u_i = -(C_c u'_o + C_D u_o) \frac{\Delta \hat{\phi}_0}{\hat{\phi}_0} \quad (6.123)$$

for electrostatic lenses. In the magnetic case

$$C_c = \frac{\eta^2}{4\hat{\phi}_o} \int_{z_o}^{z_i} B^2 h^2 dz \quad C_D = \frac{\eta^2}{4\hat{\phi}_o} \int_{z_o}^{z_i} B^2 g h dz \quad C_\theta = \frac{\eta}{4\hat{\phi}_o^{1/2}} \int_{z_o}^{z_i} B dz \quad (6.124a)$$

For electrostatic lenses, we have (nonrelativistically)

$$\begin{aligned} C_c &= \int_{z_o}^{z_i} \frac{3}{8} \left( \frac{\phi_o}{\phi} \right)^{1/2} \left( \frac{\phi'}{\phi} \right)^2 h^2 dz \\ C_D &= \int_{z_o}^{z_i} \frac{3}{8} \left( \frac{\phi_o}{\phi} \right)^{1/2} \left( \frac{\phi'}{\phi} \right)^2 g h dz + \frac{1}{4} \left( \frac{\phi_o}{\phi_i} - 1 \right) \end{aligned} \quad (6.124b)$$

The variations  $\Delta\phi_0$  and  $\Delta B_0$  are assumed to be small changes in the accelerating voltage  $\phi_0$  and in the maximum magnetic field  $B_0$ , respectively. Provided that the lens yoke does not become saturated,  $\Delta B_0/B_0 = \Delta(NI)/NI$ , where  $NI$  is the number of ampère-turns in the lens coil.

The chromatic aberration coefficient  $C_c$  is similar to the spherical aberration coefficient  $C_s$  in that it creates a circular blur for each object point, wherever the latter is situated, even on the optic axis. It is therefore important for objective lenses, and the preceding formulas show immediately that it is positive definite. The coefficient  $C_D$  is a distortion, important primarily for projector lenses.

The asymptotic form of the chromatic aberration coefficients may be expressed as follows:

$$\begin{aligned} \begin{pmatrix} u_i + \Delta u_i \\ u'_i + \Delta u'_i \end{pmatrix} &= T \begin{pmatrix} u_o \\ u'_o \end{pmatrix} \\ T &= \begin{pmatrix} M - M(C_D + iC_\theta)\Delta & -MC_c\Delta \\ c + \tilde{C}_D\Delta & rm + \tilde{C}_c\Delta \end{pmatrix} \end{aligned}$$

and as usual,  $r = f_o/f_i = (\hat{\phi}_o/\hat{\phi}_i)^{1/2}$ ;  $c$  denotes the convergence,  $c = -1/f_i$ . We have written

$$\Delta = 2 \frac{\Delta B_0}{B_0} - \frac{\Delta \hat{\phi}_0}{\hat{\phi}_0} \quad (6.125)$$

In the top row (aberrations of position), the coefficients are given by

Magnetic	Electrostatic
$C_c = \frac{\eta^2}{4\hat{\phi}} \int B^2 H^2 dz$	$C_c = \hat{\phi}_0^{1/2} \int EH^2 dz$
$C_D = \frac{\eta^2}{4\hat{\phi}} \int B^2 GH dz$	$C_D = \bar{C}_D + C_e$
$C_\theta = \frac{\eta}{4\hat{\phi}^{1/2}} \int Bdz$	$\bar{C}_D = \hat{\phi}_0^{1/2} \int EGH dz$
	$C_e = \frac{\hat{\phi}_0}{4} \left( \frac{\gamma_i}{\hat{\phi}_i} - \frac{\gamma_o}{\hat{\phi}_o} \right) + \frac{1}{4}(r^2\gamma_i - \gamma_o)$

in which

$$E = \frac{\gamma(3 + 2\varepsilon\hat{\phi})\phi'^2}{8\hat{\phi}^{5/2}} \quad (6.126)$$

$C_e$  of course vanishes for lenses with no overall accelerating effect, for which  $\phi_i = \phi_o$ .

In the second row of the matrix  $\mathbf{T}$ , the aberrations of gradient appear

Magnetic	Electrostatic
$\tilde{C}_c = mC_D + C_c/f$	$\tilde{C}_c = rm(\bar{C}_D - C_e) + C_c/f_i$
$\tilde{C}_D = C_D/f + mC_z$	$\tilde{C}_D = (\bar{C}_D + C_e)/f_i + rmC_z$
$C_z = \frac{\eta^2}{4\hat{\phi}_o} \int B^2 G^2 dz$	$C_z = \hat{\phi}_o^{1/2} \int EG^2 dz$

The coefficients can again be written as polynomials in object position or reciprocal magnification (Shimoyama, 1982; Hawkes, 1982):

$$\begin{pmatrix} C_c \\ C_D \\ C_\theta \end{pmatrix} = \begin{pmatrix} f_o^2 C_2 & -2f_o^2 C_1 & f_o^2 C_0 \\ 0 & -f_o C_2 & f_o C_1 + C_e \\ 0 & 0 & C_\theta \end{pmatrix} \begin{pmatrix} m^2 \\ m \\ 1 \end{pmatrix} \quad (6.127)$$

in which, for magnetic lenses,

$$C_2 = m_{20} \quad C_1 = m_{11} \quad \text{and} \quad C_0 = m_{02} \quad (6.128)$$

where

$$m_{ij} = \frac{\eta^2}{4\hat{\phi}_o} \int B^2 G^i \bar{G}^j dz \quad (6.129)$$

For electrostatic lenses

$$C_2 = e_{20} \quad C_1 = e_{11} \quad \text{and} \quad C_0 = e_{02} \quad (6.130)$$

where

$$e_{ij} = \hat{\phi}_o^{1/2} \int \frac{\gamma(3 + 2\varepsilon\hat{\phi})}{8\hat{\phi}^{5/2}} \phi'^2 G^i \bar{G}^j dz \quad (6.131)$$

In the nonrelativistic approximation, these collapse to

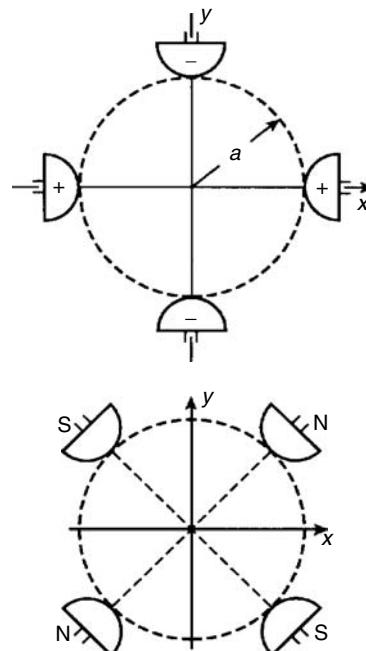
$$e_{ij} = \frac{3}{8} \phi_o^{1/2} \int \frac{\phi'^2}{\phi^{5/2}} G^i \bar{G}^j dz \quad C_e = \frac{\phi_o}{4} \left( \frac{1}{\phi_i} - \frac{1}{\phi_o} \right) = \frac{1}{4}(r^2 - 1) \quad (6.132)$$

The aberrations of gradient are given explicitly for electrostatic lenses by Lencová and Lenc (1994) and are used to establish the aberration coefficients of doublets in Hawkes (2004); some related numerical work is reported in Lencová and Hawkes (2004).

### 6.3.3 QUADRUPOLE LENSES

These lenses have planes of symmetry and antisymmetry (Figure 6.2), and, when oriented as shown in the figure, electrostatic quadrupoles are characterized by  $p_2(z)$  and magnetic ones by  $Q_2(z)$ . The paraxial equations for  $x(z)$  and  $y(z)$  are then uncoupled but, unlike those for round lenses, they are not identical. Explicitly, we have (nonrelativistically)

$$\begin{aligned} x'' + \frac{\phi'}{2\phi} x' + \frac{\phi'' - 2p_2 + 4\eta Q_2 \phi^{1/2}}{4\phi} x &= 0 \\ y'' + \frac{\phi'}{2\phi} y' + \frac{\phi'' + 2p_2 - 4\eta Q_2 \phi^{1/2}}{4\phi} y &= 0 \end{aligned} \quad (6.133)$$



**FIGURE 6.2** The appropriate orientation of electrostatic (*above*) and magnetic (*below*) quadrupole lenses.

The familiar cardinal elements can thus be used to characterize such lenses, but two sets of elements are now necessary, one in the  $x-z$  plane, the other in the  $y-z$  plane. In theory, real and asymptotic elements must be distinguished, but, in practice, it is not very likely that a real object will be inserted into a quadrupole field. We shall therefore consider only the asymptotic situation here; full discussion of the real elements is to be found in the literature (Melkich, 1947; Glaser, 1956; Dušek, 1959; Hawkes, 1966, 1970e; Yavor, 1968; Baranova and Yavor, 1986, 1989).

Since the cardinal elements are usually different in the  $x-z$  and  $y-z$  planes, a point object will in general be imaged astigmatically. The object plane will be conjugate to some plane  $z = z_{xi}$  if we use the cardinal elements corresponding to the  $x-z$  plane and to some other plane  $z = z_{yi}$  if we use those for the  $y-z$  plane. Quadrupoles are commonly employed in sets of two, four, or more elements, and we therefore need the notion of astigmatic object as well as astigmatic image (see chapter 19 of Hawkes and Kasper for more extensive explanation of this concept).

The lower symmetry of quadrupole lenses means that more coefficients are needed to characterize their aberrations in full. Here, we consider only the (asymptotic) *aperture aberrations*. In place of Equation 6.21, we now write

$$\begin{aligned}\Delta x_i &= \frac{x(z_{xi}) - M_x x_o}{M_x} \\ \Delta y_i &= \frac{y(z_{yi}) - M_y y_o}{M_y}\end{aligned}\quad (6.134)$$

where  $M_x$  denotes the magnification between  $z_{xi}$  and  $z_{xo}$  and  $M_y$  that between  $z_{yi}$  and  $z_{yo}$ . The coordinate  $x_o$  and gradient  $x'_o$  are measured in the plane  $z = z_{xo}$  while  $y_o$  and  $y'_o$  are measured in  $z = z_{yo}$ . Then,  $\Delta x_i = (0030)x_o'^3 + (0012)x_o'y_o'^2$  and  $\Delta y_i = (0003)y_o'^3 + (0021)x_o'^2y_o'$  in which, for constant  $\phi(z)$

$$\begin{aligned}(0030) &= \int_{-\infty}^{\infty} (F - \bar{F}) H_x^4 dz \\ (0003) &= \int_{-\infty}^{\infty} (F + \bar{F}) H_y^4 dz \\ (0012) &= \int_{-\infty}^{\infty} [KH_x^2 H_y^2 + LH_x H_y (H_x H_y' - H_y H_x')] dz + \frac{1}{4} (d_i m_x^2 m_y^2 - d_o) \\ (0021) &= \int_{-\infty}^{\infty} [KH_x^2 H_y^2 + LH_x H_y (H_x H_y' - H_y H_x')] dz - \frac{1}{4} (d_i m_x^2 m_y^2 - d_o)\end{aligned}\quad (6.135)$$

and (nonrelativistically)

$$\begin{aligned}F &= \frac{1}{6\phi^2} (2p_2^2 - 4\eta p_2 Q_2 \phi^{1/2} + 3\eta^2 Q_2^2 \phi) - 2\Xi \\ &= \frac{\beta^4 q^2}{6} (3n^2 - 2n + 3) - 2\tau\omega \\ \bar{F} &= \frac{1}{48\phi} (p_2'' - 2\eta Q_2'' \phi^{1/2}) \\ &= -\frac{1}{24} \beta^2 q''\end{aligned}$$

$$\begin{aligned}
K &= -\frac{1}{2\phi^2}(p_2 - \eta Q_2 \phi^{1/2})^2 + 6\Xi \\
&= -\frac{1}{2}\beta^4 q^2(n^2 - 2n + 1) + 6\tau\omega \\
L &= \frac{3}{8} \frac{p'_2 - 2\eta Q'_2 \phi^{1/2}}{\phi} \\
&= -\frac{3}{4}\beta^2 q'_2
\end{aligned} \tag{6.136}$$

$$\begin{aligned}
d_i &= z_{ix} - z_{iy} \\
d_o &= z_{ox} - z_{oy}
\end{aligned} \tag{6.137}$$

In the foregoing formulas, we have included a representation that is convenient when electrostatic and magnetic quadrupole fields never overlap or when they fit exactly (Ovsyannikova and Yavor, 1965). We then write

$$\begin{aligned}
p_2(z) &= p_{20}q(z) & Q_2(z) &= Q_{20}q(z) \\
p_4(z) &= p_{40}\omega(z) & Q_4(z) &= Q_{40}\omega(z)
\end{aligned} \tag{6.138}$$

in which  $q(z)$  and  $\omega(z)$  are normalized functions that have a maximum value of unity. We then define excitation parameters

$$\beta_E^2 = \frac{p_{20}}{2\phi} \quad \beta_M^2 = \frac{\eta Q_{20}}{\phi^{1/2}} \quad \beta^2 = \beta_M^2 - \beta_E^2 \quad n = \frac{\beta_E^2}{\beta_M^2 - \beta_E^2} \tag{6.139}$$

For purely electrostatic quadrupoles, therefore,  $n = -1$ ; for magnetic quadrupoles  $n = 0$ . For the octopoles, we write

$$\tau_E = -\frac{p_{40}}{48\phi} \quad \tau_M = \frac{\eta Q_{40}}{24\phi^{1/2}} \quad \tau = \tau_m - \tau_E \tag{6.140}$$

so that

$$\Xi = \frac{1}{48} \left( \frac{\phi}{\phi_i} \right)^{1/2} \left( \frac{p_4}{\phi} - \frac{2\eta Q_4}{\phi^{1/2}} \right) = \tau\omega(z) \tag{6.141}$$

The structure of the chromatic aberrations has an interesting feature. We have

$$\begin{aligned}
\Delta x_i &= (C_{cx}x'_o + C_{Mx}x_o) \frac{\Delta\phi}{\phi_0} \\
\Delta y_i &= (C_{cy}y'_o + C_{My}y_o) \frac{\Delta\phi}{\phi_0}
\end{aligned} \tag{6.142}$$

for variations of the accelerating voltage  $\Delta\phi_0$ . We shall not assume that  $\phi(z)$  is constant here. The coefficients are given by

$$\begin{aligned} C_{cx} &= \int (k_\phi + k_q) H_x^2 dz \\ C_{cy} &= \int (k_\phi - k_q) H_y^2 dz \\ C_{Mx} &= \int (k_\phi + k_q) G_x H_x dz \\ C_{My} &= \int (k_\phi - k_q) G_y H_y dz \end{aligned} \quad (6.143)$$

where

$$\begin{aligned} k_\phi &= -\frac{3}{8} \left( \frac{\phi_o}{\phi} \right)^{1/2} \frac{\phi'^2}{\phi^2} \\ k_q &= \left( \frac{\phi_o}{\phi} \right)^{1/2} \frac{p_2 - \eta Q_2 \phi^{1/2}}{2\phi} \end{aligned} \quad (6.144)$$

For pure quadrupoles, in the notation introduced earlier, we find

$$\begin{aligned} C_{cx} &= \frac{1}{2}(n-1)\beta^2 \int q(z) H_x^2 dz \\ C_{cy} &= -\frac{1}{2}(n-1)\beta^2 \int q(z) H_y^2 dz \\ C_{Mx} &= \frac{1}{2}(n-1)\beta^2 \int q(z) G_x H_x dz \\ C_{My} &= -\frac{1}{2}(n-1)\beta^2 \int q(z) G_y H_y dz \end{aligned} \quad (6.145)$$

If, therefore,  $n = 1$ , which implies  $\beta_M^2 = 2\beta_E^2$ , a combined electrostatic and magnetic quadrupole designed so that the two field functions have the same shape will be *achromatic*. This condition was first established by Kel'man and Yavor (1961), rediscovered by Septier (1963) and generalized by Hawkes (1964, 1965a); see also Yavor et al. (1964).

### 6.3.4 SEXTUPOLES

Like octopoles, sextupoles are used not in isolation but in conjunction with one or more round lenses. The sextupoles do not affect the paraxial equations, which will be the same as Equations 6.16; it is convenient to adopt a slightly different pair of solutions from those of Equation 6.18, namely,  $h(z)$ , which, as before respectively,

$$h(z_o) = 0 \quad h'(z_o) = 1 \quad (6.146a)$$

and  $k(z)$

$$k(z_o) = 1 \quad k(z_k) = 0 \quad (6.146b)$$

where  $z = z_k$  is the coma-free point discussed earlier (Equation 6.53). The expression for a general ray becomes

$$u(z) = \Omega_h h(z) + \Omega_k k(z) \quad (6.147)$$

in which  $\Omega_h$  and  $\Omega_k$  are simply related to the usual object coordinates (for ample details, see Rose, 2003a, especially equation 5.244, which is followed here closely; for another account, see Liu, 1999):

$$\begin{aligned} \Omega_k &= u_o \\ \Omega_h &= \frac{p_o}{p_o(u_o = 0)} - u_o k'(z_o) \end{aligned} \quad (6.148)$$

and  $p_o$  is the momentum.

When this general ray passes through a sextupole lens, it will be deviated through a distance  $\Delta u$ , which has the form

$$\Delta u = \bar{\Omega}_h^2 u_{11} + \bar{\Omega}_h \bar{\Omega}_k u_{12} + \bar{\Omega}_k^2 u_{22} \quad (6.149)$$

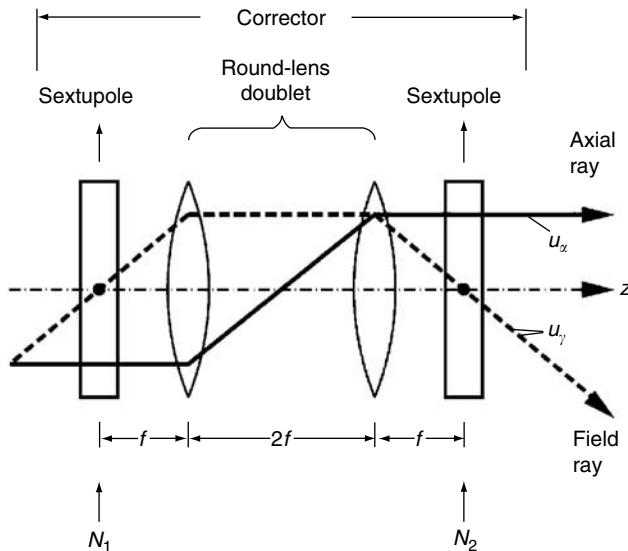
in which

$$\begin{aligned} u_{11}(z) &= \frac{1}{2} \left( h(z) \int_{z_o}^z H h^2 k \, dz - k(z) \int_{z_o}^z H h^3 \, dz \right) \\ u_{12}(z) &= \left( h(z) \int_{z_o}^z H h k^2 \, dz - k(z) \int_{z_o}^z H h^2 k \, dz \right) \\ u_{22}(z) &= \frac{1}{2} \left( h(z) \int_{z_o}^z H k^3 \, dz - k(z) \int_{z_o}^z H h k^2 \, dz \right) \end{aligned} \quad (6.150)$$

The function  $H(z)$  characterizes the field distribution in the sextupoles; in the most general case (in which both electrostatic and magnetic sextupoles may be present),  $H$  is given by

$$H(z) = -\frac{\exp(-3i\chi)}{\hat{\phi}_o^{1/2}} \left( \frac{(1 + \varepsilon\phi)(p_3 + iq_3)}{\hat{\phi}^{1/2}} + i\eta(P_3 + iQ_3) \right) \quad (6.151)$$

where  $\chi(z)$  characterizes the usual rotation in magnetic lenses (retained here because the sextupole corrector will be associated with a magnetic lens, relative to which the paraxial solutions have been defined).



**FIGURE 6.3** The simplest sextupole arrangement. No second-order aberrations are introduced outside the corrector. (After Rose, 2003a. Courtesy of the author and Springer Verlag.)

A single sextupole will introduce unwanted second-order effects. These effects can be avoided by using two or more sextupoles in series. All the three contributions will be canceled if the four integrals

$$\int_{z_o}^{z_e} H(z) h^{3-n} k^n dz \quad n = 0, 1, 2, \text{ and } 3 \quad (6.152)$$

vanish simultaneously. The form of these four conditions, in two of which  $h(z)$  is an even function while  $k(z)$  may be odd, and in the other two the situation is reversed, indicates that symmetry can be used to eliminate all four quantities. The simplest arrangement is shown in Figure 6.3. Here, one ray is symmetrical within each sextupole but antisymmetric about the center plane of the combination; the other ray is antisymmetric about the midplane of each sextupole but symmetric about the center plane of the whole combination. The members of the round-lens doublet have equal and opposite excitations, chosen in such a way that the centers of the sextupoles are conjugates, with magnification  $-1$ . The sextupole strength is a free parameter and can hence be used to cancel the spherical aberration of an adjoining round lens.

### 6.3.5 SUPERIMPOSED DEFLECTION FIELDS AND ROUND LENSES

A separate section on these mixed systems is provided because of their immense practical importance in the field of nanolithography (Munro, 1980; Plies, 1982, 1994). Pure deflection systems are not accorded a separate section here—they are investigated at length by Kaashoek (1968) and Haantjes and Lubben (1957, 1959) and described in Hawkes and Kasper (chapter 32). In the most general case, in which electrostatic and magnetic round lens and deflecting fields are present, the paraxial equations take the form

$$w'' = A_1 w^* + A_2 w + A_3 w' + A_4 \quad (6.153)$$

in which (nonrelativistically)

$$w = X + iY$$

$$\begin{aligned} A_1 &= -\frac{1}{2} \frac{Q_E}{\phi} + i\eta \frac{Q_M}{\phi^{1/2}} \\ A_2 &= -\frac{1}{4} \frac{\phi''}{\phi} + i\eta \frac{B'}{2\phi^{1/2}} \\ A_3 &= -\frac{1}{2} \frac{\phi'}{\phi} + i\eta \frac{B}{\phi^{1/2}} \\ A_4 &= -\frac{1}{2} \frac{F_T}{\phi} - i\eta \frac{B_T}{\phi^{1/2}} \end{aligned} \quad (6.154)$$

where

$$F_T = F_1 + iF_2 \quad B_T = B_1 + iB_2 \quad Q_E = p_2 + iq_2 \quad \text{and} \quad Q_M = P_2 + iQ_2 \quad (6.155)$$

Note that we have used  $w$  instead of  $u$  because the rotating coordinates are not employed here. The various special cases of Equation 6.153 are each of interest in particular situations. Thus, the round electrostatic lens field is unlikely to be allowed to overlap any electrostatic deflecting fields, although this situation is examined by Cleaver (1987) and Ohiwa et al. (1981). The term in  $A_1 w^*$  can usually be neglected since its effect can be eliminated by means of a stigmator. The general solution of Equation 6.153 has the form

$$w(z) = w_o \sigma(z) + w'_o \tau(z) + I_c m(z) + U_c e(z) \quad (6.156)$$

in which we have written

$$F_j(z) = U_j a_j(z) \quad U_c = U_1 + iU_2 \quad B_j(z) = \mu_o I_j b_j(z) \quad I_c = I_1 + iI_2 \quad (6.157)$$

The functions  $m(z)$  and  $e(z)$  are given by

$$\begin{aligned} m(z) &= -\frac{i\eta\mu_o}{\phi_o^{1/2}} \int b(\zeta) \{ \tau(z)\sigma^*(\zeta) - \sigma(z)\tau^*(\zeta) \} d\zeta \\ e(z) &= -\frac{1}{2\phi_o^{1/2}} \int \frac{a(\zeta)}{\phi^{1/2}(\zeta)} \{ \tau(z)\sigma^*(\zeta) - \sigma(z)\tau^*(\zeta) \} d\zeta \end{aligned} \quad (6.158)$$

The aberrations are then obtained by one of the methods presented earlier. Variation of parameters is employed in the authoritative articles of Munro (1974) and Chu and Munro (Munro and Chu, 1982; Chu and Munro, 1982) and of Plies (1982, 1994) and Plies and Elstner (1989). In the worst case, 56 complex geometric aberration coefficients are obtained, but this falls to nine if the Gaussian spot size can be ignored relative to the magnetic deflection and there is no electrostatic component (it is important to consult the article of Plies and Elstner on this point). In practice, the

Gaussian spot size is indeed likely to be negligible, but both magnetic and electrostatic deflection fields may well be present, magnetic to explore a large expanse and electrostatic because of the fast response. There are then 27 geometric coefficients, as listed in Table 6.1, the formulas for which can be written in the compact form.

$$\begin{aligned}
 C_0 &= F(\tau, \tau, \tau^*) \\
 \tilde{K}_0^{*(m)} &= F(\tau, \tau, m^*) \quad \tilde{K}_0^{*(e)} = F(\tau, \tau, e^*) \\
 2\tilde{K}_0^{(m)} &= F(\tau, m, \tau^*) + F(m, \tau, \tau^*) \\
 2\tilde{K}_0^{(e)} &= F(\tau, e, \tau^*) + F(e, \tau, \tau^*) \\
 \tilde{A}_0^{(m)} &= F(m, m, \tau^*) \quad \tilde{A}_0^{(e)} = F(e, e, \tau^*) \\
 \tilde{A}_0^{(me)} &= F(m, e, \tau^*) + F(e, m, \tau^*) \\
 F_0^{(m)} &= F(\tau, m, m^*) + F(m, \tau, m^*) \\
 F_0^{(e)} &= F(\tau, e, e^*) + F(e, \tau, e^*) \\
 F_0^{(me)} &= F(\tau, m, e^*) + F(m, \tau, e^*) \\
 F_0^{(em)} &= F(\tau, e, m^*) + F(e, \tau, m^*) \\
 \tilde{D}_0^{(m)} &= F(m, m, m^*) \quad \tilde{D}_0^{(e)} = F(e, e, e^*) \tag{6.159} \\
 \tilde{D}_0^{(mme^*)} &= F(m, m, e^*) \quad \tilde{D}_0^{(m^*ee)} = F(e, e, m^*) \\
 \tilde{D}_0^{(mm^*e)} &= F(m, e, m^*) + F(e, m, m^*) \\
 \tilde{D}_0^{(mee^*)} &= F(e, m, e^*) + F(m, e, e^*) \\
 \tilde{K}_4^{(m)} &= G(\tau^*, \tau^*, m^*) \quad \tilde{K}_4^{(e)} = G(\tau^*, \tau^*, e^*) \\
 \tilde{A}_4^{(m)} &= 2G(\tau^*, m^*, m^*) \quad \tilde{A}_4^{(e)} = 2G(\tau^*, e^*, e^*) \\
 \tilde{A}_4^{(me)} &= 2G(\tau^*, m^*, e^*) + 2G(\tau^*, e^*, m^*) \\
 \tilde{D}_4^{(m)} &= G(m^*, m^*, m^*) \quad \tilde{D}_4^{(e)} = G(e^*, e^*, e^*) \\
 \tilde{D}_4^{(me)} &= G(m^*, m^*, e^*) + 2G(m^*, e^*, m^*) \\
 \tilde{D}_4^{(em)} &= G(e^*, e^*, m^*) + 2G(e^*, m^*, e^*)
 \end{aligned}$$

in which

$$F(x_1, x_2, x_3) = F_1 + \frac{1}{\phi_i^{1/2} \tau_i'^* v_{1i} v_{2i} v_{3i}^*} \int_{z_o}^{z_i} \phi^{1/2} (F_2 + F_3) dz \tag{6.160}$$

**TABLE 6.1****Third-Order Geometric Aberrations of a Combined Focusing and Dual-Channel Deflection System with Magnetic Main-Field Deflection and Electrostatic Secondary Deflection**

	Axial Aberration	Magnetic Deflection Aberration	Electrostatic Deflection Aberration	Mixed Deflection Aberration
Spherical aberration	$C_0 s^2 s^*$	—	—	—
Coma	—	$\tilde{K}_0^{(m)*} s^2 w^{(m)*}$	$\tilde{K}_0^{(e)*} s^2 w^{(e)*}$	—
	—	$2\tilde{K}_0^{(m)*} ss^* w^{(m)}$	$2\tilde{K}_0^{(e)*} ss^* w^{(e)}$	—
Astigmatism and field curvature	—	$\tilde{F}_0^{(m)} sw^{(m)} w^{(m)*}$	$\tilde{F}_0^{(e)} sw^{(e)} w^{(e)*}$	$\tilde{F}_0^{(em)} sw^{(m)*} w^{(e)}$
	—	—	—	$\tilde{F}_0^{(me)} sw^{(m)} w^{(e)*}$
	—	$\tilde{A}_0^{(m)} s^* w^{(m)2}$	$\tilde{A}_0^{(e)} s^* w^{(e)2}$	$\tilde{A}_0^{(me)} s^* w^{(m)} w^{(e)}$
Distortion	—	$\tilde{D}_0^{(m)} w^{(m)2} w^{(m)*}$	$\tilde{D}_0^{(e)} w^{(e)2} w^{(e)*}$	$\tilde{D}_0^{(mme*)} w^{(m)2} w^{(e)*}$
	—	—	—	$\tilde{D}_0^{(mm*e)} w^{(m)} w^{(m)*} w^{(e)}$
	—	—	—	$\tilde{D}_0^{(mee*)} w^{(m)} w^{(e)} w^{(e)*}$
	—	—	—	$\tilde{D}_0^{(m*ee)} w^{(m)*} w^{(e)2}$
Fourfold coma	—	$\tilde{K}_4^{(m)} s^2 w^{(m)*}$	$\tilde{K}_4^{(e)} s^2 w^{(e)*}$	—
Fourfold astigmatism	—	$\tilde{A}_4^{(m)} s^* w^{(m)*2}$	$\tilde{A}_4^{(e)} s^* w^{(e)*2}$	$\tilde{A}_4^{(me)} s^* w^{(m)*} w^{(e)*}$
Fourfold distortion	—	$\tilde{D}_4^{(m)} w^{(m)*3}$	$\tilde{D}_4^{(e)} w^{(e)*3}$	$\tilde{D}_4^{(me)} w^{(m)*2} w^{(e)*}$
	—	—	—	$\tilde{D}_4^{(em)} w^{(m)*} w^{(e)*2}$

Gaussian spot size negligible.

where

$$F_1 = \frac{n_1 n_2 n_3}{\phi_i^{1/2} \tau_i'^* v_{1i} v_{2i} v_{3i}} \left[ \phi^{1/2} \tau'^* v_2 \left\{ v_1 v_3^* \left( \frac{\phi''}{32\phi} - \frac{i\eta B'}{16\phi^{1/2}} \right) + v_1 \left( \frac{s_3 a}{16\phi} + \frac{i\eta m_3 b}{8\phi^{1/2}} \right) + v_3^* \left( \frac{s_1 a}{8\phi} + \frac{i\eta m_1 b}{4\phi^{1/2}} \right) \right\} \right]_{z_o}^{z_i} \quad (6.161)$$

$$\begin{aligned} F_2 = & \frac{1}{2} \tau'^* x'_1 x'_2 x'_3^* + \frac{3}{64} \left( \frac{\phi''}{\phi} \right)^2 \tau^* x_1 x_2 x''_3^* \\ & - \frac{\phi''}{32\phi} (\tau''^* x_1 x_2 x_3^* + 2\tau^* x'_1 x_2 x_3^* + \tau^* x_1 x_2 x'''_3^* + 2\tau'^* x_1 x_2 x'_3^* + 2\tau^* x'_1 x'_2 x''_3^*) \\ & + \frac{\phi'' \phi'}{32\phi^2} (\tau'^* x_1 x_2 x_3^* + 2\tau x'_1 x_2 x_3^* + \tau x_1 x_2 x'_3^*) - \frac{3\phi'' \phi'^2}{128\phi^3} \tau^* x_1 x_2 x_3^* \\ & + \frac{i\eta B'}{16\phi^{1/2}} (\tau''^* x_1 x_2 x_3^* - 2\tau^* x''_1 x_2 x_3^* + \tau^* x_1 x_2 x'''_3^* + 2\tau'^* x_1 x_2 x'_3^* - 2\tau^* x'_1 x'_2 x''_3^*) \end{aligned} \quad (6.162)$$

$$\begin{aligned}
F_3 = & \frac{s_1 a}{8\phi} \left\{ \tau^* \left( \frac{\phi''}{\phi} x_2 x_3^* + \frac{s_2 a}{\phi} x_3^* + \frac{s_3 a}{\phi} x_2 \right) \right. \\
& - (\tau''^* x_2 x_3^* + \tau^* x_2'' x_3^* + \tau^* x_2 x_3''' + 2\tau'^* (x_2 x_3^*)') \\
& + \frac{\phi'}{\phi} (\tau^* x_2 x_3^*)' - \frac{3\phi'^2}{4\phi^2} \tau^* x_2 x_3^* \left. \right\} + \frac{s_2 a}{16\phi} x_3^* \left( 4\tau'^* x_1 + \frac{\phi''}{\phi} \tau^* x_1 \right) \\
& + \frac{s_3 a}{16\phi} \left\{ \frac{3\phi''}{2\phi} \tau^* x_1 x_2 - \tau^* x_1 x_2 - \tau''^* x_1 x_2 - 2\tau x_1'' x_2 - 2\tau^* x_1' x_2' \right. \\
& + \frac{\phi'}{\phi} x_2 (\tau'^* x_1 + 2\tau^* x_1') - \frac{3\phi'^2}{4\phi^2} \tau^* x_1 x_2 \left. \right\} \\
& - \frac{i\eta b m_1}{4\phi^{1/2}} (\tau''^* x_2 x_3^* - \tau x_2'' x_3^* + \tau x_2 x_3''' + 2\tau'^* x_2 x_3^*) \\
& - \frac{i\eta b m_3}{8\phi^{1/2}} (\tau''^* x_1 x_2 - 2\tau^* x_1'' x_2 - 2\tau^* x_1' x_2') \tag{6.163}
\end{aligned}$$

and

$$G(x_1^*, x_2^*, x_3^*) = \frac{1}{2\phi_i^{1/2} \tau'^* x_{1i}^* x_{2i}^* x_{3i}^*} \int_{z_o}^{z_i} \left( \frac{s_3 p_e}{2\phi} + \frac{\eta m_3 P_m}{\phi^{1/2}} \right) \tau^* x_1^* x_2^* dz \tag{6.164}$$

For each of the possible meanings of  $w_j$ , the quantities  $v_j$ ,  $n_j$ ,  $m_j$ , and  $s_j$  that occur in these definitions take the following values:

$$\begin{array}{cccccc}
w_j & v_j & n_j & m_j & s_j \\
\tau & \tau' & 0 & 0 & 0 \\
\sigma & \sigma & 1 & 0 & 0 \\
m & m & 1 & 1 & 0 \\
e & e & 1 & 0 & 1
\end{array} \tag{6.165}$$

The field functions  $p_e$  and  $P_m$  that appear in  $G(x_1^*, x_2^*, x_3^*)$  denote the contributions from the field components with threefold symmetry:

$$p_e = \frac{p_3 - iq_3}{U_c} \quad P_m = -i \frac{P_3 - iQ_3}{I_c} \tag{6.166}$$

For the more general cases in which the electrostatic and magnetic deflection fields are not aligned and Gaussian spot size cannot be neglected, see the work of Chu and Munro cited earlier and the very thorough studies of Plies (1982, 1994) and Plies and Elstner (1989). The aberrations in planes other than the Gaussian image plane are considered in detail by Zhang and Cleaver (1996).

The asymptotic aberrations of position and gradient of superimposed round lens and deflection fields have been fully analyzed by Hawkes (1989), who has also established the addition rules for combining such elements (1991). Tang (1986a) too has studied such components, and his work includes the chromatic aberrations as well. Li (1994) has reconsidered these formulas and has, in particular, obtained expressions from which high derivatives of field functions have been

eliminated. Note that printing errors survived in Li's earlier article on the same subject (Li, 1993) and the article Li (1994) is to be preferred; see also Li and Ximen (1982) and Ximen (1986, 1988). For other studies on various aspects of deflection elements, see Glaser (1949), Lencová (1981, 1988), and Lenc and Lencová (1988). Combinations of lenses and deflectors are also discussed by Uno et al. (1995), Dowker et al. (1995), Zhu et al. (1999), and Ren et al. (2007). Higher order aberrations are considered by Li (1986).

### 6.3.6 MIRRORS AND CATHODE LENSES

The term *electron mirror* can lead to confusion for in electron optics a distinction needs to be made between two situations. In the first, the electrostatic retarding field is so high that electrons incident on it are halted and retrace their path essentially close to the same optic axis. This is the case, for example, in an electrostatic einzel lens when the potential of the central electrode is such that the electrons cannot pass the barrier. In the other situation, the electrons are incident on the repeller field at an angle to the normal that is not at all negligible and return along paths that are far from the incident ones. This is simply a special case of a system with a curved optic axis and is extremely common in certain types of analyzer, for example, the cylindrical mirror analyzer. In the light-optical situation, it is natural to refer to both arrangements as mirrors, since there is no physical reason to distinguish between them (in geometric optics at least). But for electrons, they are very different, since in the first case the electrons come to a halt at the mirror and are accelerated away from rest in the opposite direction; close to the turning point, their gradient is not small and one of the conditions that we regard as characteristic of the paraxial approximation is not satisfied. In the second case, the electrons are not brought to a standstill; the velocity component normal to the mirror does, of course, fall to zero and change sign, but this is of no great consequence since the behavior of the electron beam will be studied relative to a smooth curved optic axis.

We are concerned in this section with mirrors in the first sense. The breakdown of the paraxial conditions just mentioned means that they require special treatment, and this can also be employed to study cathode lenses, devices in which a surface emitting electrons is situated inside an electrostatic field, which serves to accelerate and focus the particles. The problem here is that, close to the cathode, the electron velocity will be low and the gradients of the trajectories may be steep. It is thus obvious that this region has much in common with the zone in a mirror in which the electron direction is reversed, and, indeed, the optics of cathode lenses is essentially that of the second half of a mirror.

Several ways of avoiding the problems created by the large gradients close to the turning points of mirrors have been proposed. The first approach is to establish paraxial equations and hence aberration formulas with time (or some simple function of time) as independent variable. This has been investigated very thoroughly (Recknagel, 1936, 1937 and for cathode lenses, Recknagel, 1941b; see also section 42 of Glaser, 1952; or chapter 18 of Hawkes and Kasper). Among the other solutions to the difficulty is a very ingenious transformation whereby the trajectory of an arbitrary electron is measured relative to that of a reference electron, traveling on the axis with a given initial energy. The position coordinates of other electrons will then characterize their off-axis distance and the axial distance from the plane of the reference particle. The associated theory has been very extensively studied by Kel'man et al. (Kel'man et al., 1971a–c, 1972a,b, 1972–4, 1973; Sekunova and Yakushev, 1975; Sekunova, 1977; Daumenov et al., 1978, 1981; Bimurzaev et al., 1999, 2003, 2004; Bimurzaev and Yakushev, 2004), who have listed the aberration coefficients for rotationally symmetrical systems of both mirrors and cathode lenses. A connected account is to be found in Yakushev and Sekunova (1986). Mirrors have been reconsidered and entirely recalculated by Rose and Preikszas (1995). For the study of aberrations especially, this approach seems definitely preferable to the use of a timelike variable. This theory has been employed in connection with cathode lenses by Takaoka and Ura (1984, 1989, 1990; Takaoka, 1995). See also Kolesnikov and Monastyrskii (1988) and Monastyrskii (1989).

In this theory, the axial coordinate  $z$  and the time are eliminated by introducing two new variables,  $\zeta$  and  $\xi$  (in the Russian work  $\eta$  is used rather than  $\xi$  but here we reserve  $\eta$  for the constant  $(e/2m_0)^{1/2}$ ). The time derivative of  $\zeta$  is equal to the velocity of the reference particle (traveling on the optic axis):

$$\frac{d\zeta}{dt} = 2\sigma\eta\phi^{1/2}(\zeta) \quad (6.167)$$

and  $\sigma = 1$  when the reference particle is traveling in the positive direction (toward the mirror) and  $\sigma = -1$  in the opposite situation. Formally,  $\sigma = (dz/dt)/|dz/dt|$ .

The other variable  $\xi$  is defined by

$$z = \zeta + \xi \quad (6.168)$$

After some calculation, it is found that the (rotating) Cartesian coordinates  $x$  and  $y$  satisfy the familiar paraxial equation, perturbed by aberration terms:

$$\phi u'' + \frac{1}{2}\phi'u' + \frac{\phi'' + \eta^2 B^2}{4}u = S^{(3)} \quad 6.169$$

in which primes now indicate differentiation with respect to  $\zeta$ . The expression for  $S^{(3)}$  contains the usual terms in  $x$  and  $y$  as well as a term involving  $\xi$ :

$$S^{(3)} = \frac{1}{32}R_4(x^2 + y^2)u - \frac{1}{4}R'_2\xi u + \frac{1}{8}\eta B''C_2u$$

in which

$$\begin{aligned} R_2 &= \phi'' + \eta^2 B^2 \\ R_4 &= \phi^{(4)} + 4\eta^2 BB'' \\ C_2 &= \frac{1}{2}i\sigma\phi^{1/2}(u\bar{u}' - \bar{u}u') \end{aligned}$$

The variable  $\xi$  satisfies an equation of the form

$$2\phi\xi' - \phi'\xi = s^{(3)} \quad (6.170)$$

where

$$s^{(3)} = -\phi u'\bar{u}' - \frac{1}{4}R_2(x^2 + y^2) - \eta BC_2 + \varepsilon \quad (6.171)$$

The quantity  $\varepsilon$  is the small energy of the particle at the point at which the energy of the reference particle is zero. The difficulties that arise in the traditional theories in the vicinity of the zone of reflection of the electrons are avoided here, and  $S^{(3)}$  and  $s^{(3)}$  behave in such a way that the familiar perturbation theory can be employed. This is essentially due to the smallness of  $\phi$  in the reflection zone, where the gradients become large, whereas the gradients are small where  $\phi$  is not small.

The paraxial equations (6.169 and 6.170 with  $S^{(3)} = s^{(3)} = 0$ ) have solutions  $p(\zeta)$  and  $g(\zeta)$ , in which

$$g(\zeta) = q(\zeta)\phi^{1/2}(\zeta) \quad (6.172)$$

so that

$$\phi q'' + \frac{1}{2}\phi'q' + \frac{1}{4}(3\phi'' + \eta^2B^2)q = 0 \quad (6.173)$$

We denote the point at which  $\phi$  vanishes by  $z_k$  and set

$$p(\zeta_k) = q(\zeta_k) = 1 \quad (6.174)$$

and so

$$\begin{aligned} p'(\zeta_k) &= -\frac{\phi_k''}{2\phi'_k} - \frac{\eta^2B_k^2}{4\phi'_k} \\ q'(\zeta_k) &= -\frac{\phi_k''}{2\phi'_k} - \frac{\eta^2B_k^2}{12\phi'_k} \end{aligned} \quad (6.175)$$

The Wronskian tells us that

$$\phi^{1/2}(pg' - p'g) = \frac{1}{2}\phi'_k \quad (6.176)$$

The general solution of the paraxial equation is hence of the form

$$w(\zeta) = \alpha p(\zeta) + \beta g(\zeta) \quad (6.177)$$

and it will be convenient to distinguish the trajectories after reflection at the mirror thus

$$\tilde{w} = w(z, \tilde{\sigma}, \tilde{\alpha}, \tilde{\beta}) \quad (6.178)$$

and

$$\tilde{\sigma} = -\sigma \quad \tilde{\alpha} = \alpha \quad \tilde{\beta} = -\beta \quad (6.179)$$

A straight ray leaving the object plane ( $z = z_o$ ,  $\zeta = \zeta_o$ ) at a point  $u_o$  with gradient  $u'_o$  may be written as

$$u = \frac{u_o - u'_o g_o / g'_o}{R p'_o} p(z) + \frac{u_o - u'_o p_o / p'_o}{R g'_o} g(z) \quad (6.180)$$

where  $R = p_o/p'_o - g_o/g'_o$ .

From this, we see that the incident ray (minus sign) will return along the same path if the final term vanishes and hence if

$$u_o = u'_o p_o / p'_o \quad (6.181)$$

in which case, all the incident rays intersect the optic axis at the point  $\zeta = \zeta_c$  at which  $p(\zeta_c) = 0$ . The corresponding point  $z = z_c$  is thus the *center of curvature* of the mirror.

If, on the other hand, the first term on the right-hand side of Equation 6.180 vanishes, which implies that

$$u_o = u'_o g_o / g'_o \quad (6.182)$$

then the incident and reflected rays will be symmetric about the optic axis and their asymptotes will intersect at the *vertex* of the mirror,  $z = z_v$ , where  $g(z_v) = 0$ . We see that the quantity denoted by  $R$  in Equation 6.180 is the *radius of curvature* of the mirror,  $R = z_v - z_c$ ; when  $\sigma R > 0$ , the mirror is convergent and when  $\sigma R < 0$ , it is divergent. By similar arguments, we can show that the focal length is given by  $R/2$  and the focus is situated midway between the vertex and the center of curvature. If as usual we denote the image plane conjugate to the object plane by  $z_i$ , we have

$$p_o g_i + p_i g_o = 0 \quad (6.183a)$$

which can be rewritten in the familiar form (the principal plane coincides with the vertex).

$$\frac{1}{z_o - z_c} + \frac{1}{z_i - z_c} = \frac{2}{R} = \frac{1}{f} \quad (6.183b)$$

After a considerable amount of algebra, given in full by Yakushev and Sekunova (1986), integrals for the aberration coefficients are found. Here, we merely list the results including the improved form of the spherical aberration coefficient given by Bimurzaev et al. (2004).

In the image plane, we have

$$\begin{aligned} \Delta u_i = & C(x'_o{}^2 + y'_o{}^2)u'_o \quad (\text{spherical aberration}) \\ & + 2(K + ik)(x'_o{}^2 + y'_o{}^2)u_o + (K - ik)(x'_o{}^2 - y'_o{}^2 + 2ix'_o y'_o)u_o^* \quad (\text{coma}) \\ & + (A + ia)r_o^2(x'_o - iy'_o) \quad (\text{astigmatism}) \\ & + Fr_o^2u'_o \quad (\text{field curvature}) \\ & + (D + id)r_o^2u_o \quad (\text{distortion}) \\ & + \frac{\epsilon}{\phi_o} \{u'_o K_{1r} + u_o (K_{2r} + ik_r)\} \end{aligned} \quad (6.184)$$

in which

$$\begin{aligned}
 C &= \frac{1}{R^3} \left\{ J_1 Z_{ov}^4 + 2(2J_2 + J_3) Z_{ov}^2 Z_{oc}^2 + J_4 Z_{oc}^4 + 3Z_{ov}^2 + Z_{oc}^2 \right\} \\
 K &= -\frac{1}{R^3} \left\{ J_1 Z_{ov}^3 + 2(2J_2 + J_3) Z_{ov} Z_{oc} Z_{of} + J_4 Z_{oc}^3 + Z_{ov} Z_{oc} (2Z_{ov} + Z_{oc}) \right\} \\
 A &= \frac{1}{R^3} \left\{ J_1 Z_{ov}^2 + 4J_2 Z_{of}^2 + 2J_3 Z_{ov} Z_{oc} + J_4 Z_{oc}^2 + J_7 R^2 + Z_{ov} (Z_{ov} + 2Z_{oc}) \right\} \\
 F &= \frac{1}{R^3} \left\{ 2J_1 Z_{ov}^2 + 4J_2 (Z_{of}^2 + Z_{ov} Z_{oc}) + J_3 (Z_{ov} + Z_{oc})^2 + 2J_4 Z_{oc}^2 - J_7 R^2 + 2Z_{ov} (Z_{ov} + 2Z_{oc}) \right\} \\
 D &= -\frac{1}{R^3} \left\{ J_1 Z_{ov} + 2(2J_2 + J_3) Z_{of} + J_4 Z_{oc} + 3Z_{ov} \right\} \\
 K_{1r} &= \frac{\phi_o}{R} (J_8 Z_{ov}^2 + J_9 Z_{oc}^2) \\
 K_{2r} &= -\frac{\phi_o}{R} (J_8 Z_{ov} + J_9 Z_{oc}) \\
 k &= \frac{1}{R^2} (J_5 Z_{ov}^2 + J_6 Z_{oc}^2) \\
 a &= -\frac{2}{R^2} (J_5 Z_{ov} + J_6 Z_{oc}) \\
 d &= \frac{1}{R^2} (J_5 + J_6) \\
 k_r &= \phi_o J_{10}
 \end{aligned} \tag{6.185}$$

and  $J_1 - J_{10}$  denote

$$\begin{aligned}
 J_1 &= \frac{g'_o}{8\phi'_k p'^3} \int_{z_k}^{z_o} \frac{p}{\phi^{1/2}} \left\{ R_4 p^3 + 4R'_2 p^2 p' + 32p'' (\phi p'^2 + R_2 p^2/4) \right\} dz \\
 J_2 &= \frac{1}{8\phi'_k p'_o g'_o} \int_{z_k}^{z_o} \frac{g}{\phi^{1/2}} \left\{ R_4 p^2 g + 4R'_2 p p' g + 32p'' (\phi p' g' + R_2 p g/4) \right\} dz \\
 J_3 &= \frac{1}{8\phi'_k p'_o g'_o} \int_{z_k}^{z_o} \frac{q}{\phi^{1/2}} \left\{ \phi R_4 p^2 q + 2R'_2 p^2 (2\phi q' - \phi' q) + 32\phi q'' (\phi p'^2 + R_2 p^2/4) \right\} dz \\
 J_4 &= \frac{P'_o}{8\phi'_k g'^3} \int_{z_k}^{z_o} \phi^{1/2} q \left\{ \phi R_4 q^3 + 2R'_2 q^2 (2\phi q' - \phi' q) + 32q'' (\phi g'^2 + R_2 g^2/4) \right\} dz \\
 J_5 &= \frac{\sigma\eta}{8p'^2} \int_{z_k}^{z_o} \frac{p}{\phi^{1/2}} \{B'' p + 4B' p' + 8B p''\} dz
 \end{aligned}$$

$$\begin{aligned}
J_6 &= -\frac{\sigma\eta}{16p_o'^2} \int_{z_k}^{z_o} \frac{q}{\phi^{1/2}} \{5\phi'B'q - 4\phi(B'q' + 4Bq'')\} dz \\
J_7 &= \frac{\eta^2}{8p_o'g_o'} \int_{z_k}^{z_o} \frac{1}{\phi^{3/2}} \{2B_k\phi B' + B(\phi'_k B - B_k\phi')\} dz \\
J_8 &= -\frac{4g_o'}{\phi'_k p_o'} \int_{z_k}^{z_o} \frac{pp''}{\phi^{1/2}} dz \\
J_9 &= -\frac{4p_o'}{\phi'_k g_o'} \int_{z_k}^{z_o} \phi^{1/2} qq'' dz \\
J_{10} &= \frac{\sigma\eta}{2\phi'_k} \int_{z_k}^{z_o} \frac{1}{\phi^{3/2}} \{2\phi B' + B(\phi'_k - \phi')\} dz
\end{aligned} \tag{6.186}$$

The distances  $Z$  are as follows:

$$\begin{aligned}
Z_{ov} &= z_o - z_v & Z_{oc} &= z_o - z_c & Z_{oF} &= z_o - z_F \\
Z_v &= z - z_v & Z_c &= z - z_c & Z_F &= z - z_F & Z_o &= z - z_o
\end{aligned} \tag{6.187}$$

In the special case of an electrostatic mirror operating at high magnification, for which the object plane and the focus coincide, the expression for  $C_s$  collapses to

$$C_s = C = \tilde{J}_1 + \tilde{J}_2 + \tilde{J}_3 + \delta \tag{6.188}$$

in which convenient forms of the four contributions are (Bimurzaev et al., 2004):

$$\begin{aligned}
\tilde{J}_1 &= \frac{1}{256p_o'^4\phi_o'^{1/2}} \int_{z_k}^{z_o} \left[ \phi^{-1/2} \left\{ \frac{\phi'''\phi'}{2\phi} (p^4 - 1) + 32pp''f_1 \right\} - \frac{1}{2} \{\phi'' - \phi'' - \phi'''(z - z_k)\} \left( \frac{p^4\phi'}{\phi^{3/2}} \right)' \right] dz \\
\tilde{J}_2 &= \frac{1}{(16p_o'g_o')^2\phi_o'^{1/2}} \int_{z_k}^{z_o} \left[ (\phi'' - \phi) \left\{ \frac{2pq}{\phi^{1/2}} \left( \frac{5}{2}\phi'pq + \phi'_k \right) \right\}' + 64q(p''f_2 + \phi^{1/2}q''f_1) \right] dz \\
\tilde{J}_3 &= \frac{1}{256g_o'^4\phi_o'^{1/2}} \int_{z_k}^{z_o} \left\{ \frac{7}{2}(\phi'' - \phi_k'')(\phi^{1/2}q^4\phi')' + 32\phi^{1/2}qq''f_3 \right\} dz \\
\delta &= \frac{\phi'_k}{16\phi_o'^{1/2}p_o'g_o'} \left( \frac{3}{2} + \frac{\phi''\phi'^2}{(16p_o'g_o')^2\phi_o'^2} \right) - \frac{\phi'''_k}{256p_o'^4\phi_o'^{1/2}}
\end{aligned} \tag{6.189}$$

If a magnetic field is also present, Equation 6.188 must be replaced by

$$C_s = \tilde{J}_1 + \tilde{J}_2 + \tilde{J}_3 + \frac{\eta^2}{64p_o'^4\phi_o'^{1/2}} (\hat{J}_1 + \hat{J}_2 + \hat{J}_3) + \delta \tag{6.190}$$

in which (Bimurzaev, 2006, personal communication)

$$\begin{aligned}\hat{J}_1 &= \int_{z_k}^{z_o} \frac{p^2 B}{\phi^{1/2}} (p^2 B'' + 2pp'B') dz = \int_{z_k}^{z_o} \left( \frac{p^2 B}{\phi^{1/2}} \right)' (B'_k - p^2 B') dz \\ \hat{J}_2 &= 2 \int_{z_k}^{z_o} \phi^{1/2} pqB (3pqB'' + 4p'qB' + 2pq'B') dz = - \int_{z_k}^{z_o} pqB' \{6\phi^{1/2}(pqB)' + B\phi^{-1/2}(2pq\phi' + \phi'_k)\} dz \\ \hat{J}_3 &= \int_{z_k}^{z_o} \phi^{1/2} q^2 B \{\phi(q^2 B'' + 2qq'B') - \phi'q^2 B'\} dz = - \int_{z_k}^{z_o} \phi^{1/2} q^2 B' \{\phi(q^2 B)' + 5\phi'q^2 B/2\} dz\end{aligned}\quad (6.191)$$

For more details, we refer to the articles cited earlier, and in particular to those of Yakushev and Sekunova (1986), Rose and Preikszas (1995), Takaoka (1995), Bimurzaev and Yakushev (2004), and Zhukov and Zavyalova (2006), who examined a combined electrostatic–magnetic mirror. For earlier work on these topics, or other approaches, see Artsimovich (1944), Ximen (1957), Ximen et al. (1983), Ioanovicu et al. (1989), and Lenc and Müllerová (1992a,b) and for a fresh investigation, Wang et al. (2007a,b).

### 6.3.7 SYSTEMS WITH CURVED OPTIC AXES

Round lenses, quadrupoles, and mirrors are the elements routinely found in focusing systems and image-forming instruments, notably microscopes and probe-forming devices for microscopy, lithography, or other purposes. They are frequently accompanied by devices in which the axis is curved, typically approximately circular or a sequence of circular segments, and which are hence dispersive. These are the magnetic and electrostatic prisms and the  $\Omega$  and associated filters to be found in mass spectrometers and energy analyzers. The axis is curved in high-energy particle accelerators, in which the particles circulate many times before reaching the desired energy. The optics of such devices was at first expressed in much the same language as that of low-energy instruments with straight axes, and we refer to texts such as Brown (1968), Brown and Servranckx (1985, 1987), and Carey (1987) for accounts of this approach. More recently, these have been complemented by newer theories, better suited to the study of the high-order perturbations that can build up in circular accelerators; these new theories are, like their predecessors, based on Hamiltonian mechanics but take explicit advantage of the body of theory associated with Lie algebra. We refer to the work of Dragt (Dragt, 1987, 1990; Dragt and Forest, 1986; Dragt et al., 1986; Healy and Dragt, 1989; Forest, 1998) for a clear account of these ideas. Yet another approach has been developed by Berz (Berz, 1987, 1989, 1990, 1995, 1999; Makino and Berz, 1997, 1999; Berz and Makino, 2004) and the associated program suites are widely used. For a variant, see Berdnikov and van der Stam (1995).

We can say no more about the optics of accelerators here and give only a superficial introduction to the aberrations of prisms; for full details, see the work of Wollnik (1967, 1987) with particular reference to mass spectrometers, where references to individual articles are to be found, notably the papers of Matsuo (1975) and Nakabushi et al. (1983), and Rose and Krah (1995), and Rose (2003a), where  $\Omega$ -filters, W-filters, and mandoline filters are explored in detail. The contribution of fringing fields to the aberrations is explored in Hartmann and Wollnik (1995). A very complete bibliography is included in part X of Hawkes and Kasper.

The principal difference between the theory presented in earlier sections and that of systems with curved axes is the form of the field expansions. In the worst possible case, in which the axis is a skew curve, tensor analysis is indispensable and no one has ventured far beyond the lowest orders. In the systems considered here, the optic axis lies in a plane, and there exist trajectories lying in the same plane close to the axis. This is a considerable simplification and the various expansions required are easy to generate once the fact that the element of length is now of the form  $ds^2 = dx^2 + dy^2 + (1 - x/R)^2 dz^2$  has been appreciated. We shall not give these expansions

here; they can be found in the work of Rose (2003a) and in Hawkes and Kasper (chapter 51). The local radius of curvature  $R$  is given by  $-\phi^{1/2}/\eta B_2$ . We draw attention to the remarkably clear account of such systems, going up to the third-rank aberrations, by Plies (2002). In this, the earlier fundamental work of Plies and Typke (1978) is recapitulated and considerably extended.

The nature of the aberrations of such systems may be understood by expressing the position and momentum of a particle in an arbitrary plane as a power series in the initial values of these quantities and a chromatic parameter,  $E/E_0$ , where  $E$  denotes the particle energy and  $E_0$  the nominal energy of the beam, usually the mean or median value; mass variation may also be included when appropriate. The primary aberrations are now quadratic in the initial values, and, at worst, nine coefficients are needed to describe position aberrations in one direction, six in the other. The aberration coefficients of each rank are interrelated more or less straightforwardly. These interrelations emerge naturally if the eikonal method is employed but need to be established separately when the trajectory method or simple power-series expansion is used (e.g., Wollnik and Berz, 1985). This is the role of what is called the *symplectic condition* in the Lie algebra approach. The relation between these various ways of studying aberrations has been examined in depth by Rose (1987a) and a helpful analysis of the role of symmetry in all these situations is to be found in Zeitler (1990).

The formulas published by Matsuo et al. have been reexamined by Toyoda et al. (1995), who have used a computer algebra package (MATHEMATICA) to establish all the aberration coefficients up to third rank for electrostatic and magnetic sector fields (we understand *rank* to include not only chromatic effects but also those resulting from mass variations). Specifically, they solve the path equations

$$\begin{aligned} x'' + k_x x &= (X|\gamma)\gamma + (X|k)k + (X|xx)x^2 + (X|x\gamma)x\gamma + (X|xk)xk + (X|yy)y^2 \\ &\quad + (X|x'x')x'^2 + (X|y'y')y'^2 + (X|\gamma\gamma)\gamma^2 + (X|\gamma k)\gamma k + (X|k k)k^2 \\ &\quad + (X|xxx)x^3 + (X|xyy)xy^2 + (X|xx'x')xx'^2 + (X|xy'y')xy'^2 + (X|x'y'y')x'y'y' \\ &\quad + \{(X|xx)\gamma + (X|xxk)k\}x^2 + \{(X|yy)\gamma + (X|yyk)k\}y^2 \\ &\quad + \{(X|x'x')\gamma + (X|x'x'k)k\}x'^2 + \{(X|y'y')\gamma + (X|y'y'k)k\}y'^2 \\ &\quad + \{(X|x\gamma\gamma)\gamma^2 + (X|xyk)\gamma k + (X|k k)k^2\}x \\ &\quad + (X|\gamma\gamma\gamma)\gamma^3 + (X|\gamma\gamma k)\gamma^2 k + (X|\gamma k k)\gamma k^2 + (X|k k k)k^3 \end{aligned} \quad (6.192a)$$

$$\begin{aligned} y'' + k_y y &= (Y|xy)xy + (Y|x'y')x'y' + \{(Y|y)\gamma + (Y|yk)k\}y \\ &\quad + (Y|xx)y^2 + (Y|yyy)y^3 + (Y|xx'y')xx'y' + (Y|x'x'y)x'^2y + (Y|yy'y')yy'^2 \\ &\quad + \{(Y|xy)\gamma + (Y|xyk)k\}xy + \{(Y|x'y')\gamma + (Y|x'y'k)k\}x'y' \\ &\quad + \{(Y|y\gamma\gamma)\gamma^2 + (Y|y\gamma k)\gamma k + (Y|y k k)\gamma k^2\}y \end{aligned} \quad (6.192b)$$

in which  $\gamma$  and  $k$  characterize variations in mass and energy. (Here and later, we have modified their notation slightly. In Equation 6.192a,  $D_k$  becomes  $(X|k)$  and in 6.192b,  $D_k$  becomes  $(Y|k)$ ; in Equations 6.193 through 6.195, subscripted coefficients are again replaced by coefficients of the form  $(xl\dots)$  and  $(yl\dots)$ .) To first order, Toyoda et al. write

$$\begin{aligned} x_I &= (xlx)x_0 + (xlx')x'_0 + (xl\gamma)\gamma + (xlk)k \\ x'_I &= (x'lx)x_0 + (x'lx')x'_0 + (x'l\gamma)\gamma + (x'lk)k \\ y_I &= (yl)y_0 + (yl'y')y'_0 \\ y'_I &= (y'ly)y_0 + (y'ly')y'_0 \end{aligned} \quad (6.193)$$

in which the functions  $(x|xx) \dots (y|yy')$  are sinusoidal functions in the sharp-cutoff (SCOFF) approximation. Toyoda et al. set the object plane at  $z = 0$ , and the subscripts are hence written as zero instead of  $o$ . To second order,

$$\begin{aligned} x_{II} &= (x|xx)x_0^2 + (x|xx')x_0x'_0 + (x|xy\gamma)x_0\gamma + (x|xx\kappa)x_0\kappa \\ &\quad + (x|yy)y_0^2 + (x|yy')y_0y'_0 + (x|x'x')x_0'^2 + (x|y'y')y_0'^2 \\ &\quad + (x|x'\gamma)x_0'\gamma + (x|x'\kappa)x_0'\kappa + (x|yy\gamma)\gamma^2 + (x|y\kappa\gamma)\gamma\kappa + (x|y\kappa\kappa)\kappa^2 \\ y_{II} &= (y|xy)x_0y_0 + (y|xy')x_0y'_0 + (y|xy'y)x_0'y_0 + (y|xy'y')x_0'y'_0 \\ &\quad + (y|yy\gamma)y_0\gamma + (y|y\kappa\gamma)y_0\kappa + (y|y\gamma\kappa)y_0'\gamma + (y|y\gamma\kappa)y_0'\kappa \end{aligned} \quad (6.194)$$

with similar expressions for  $x'_{II}$  and  $y'_{II}$ , and to third order,

$$\begin{aligned} x_{III} &= (x|xxx)x_0^3 + (x|xyy)x_0y_0^2 \\ &\quad + (x|xxx')x_0^2x'_0 + (x|xyy')x_0y_0y'_0 + (x|x'yy)x_0'y_0^2 \\ &\quad + (x|xx'x')x_0x_0'^2 + (x|xy'y')x_0y_0'^2 + (x|x'yy')x_0'y_0y'_0 \\ &\quad + (x|x'x'x')x_0'^3 + (x|x'y'y')x_0'y_0'^2 \\ &\quad + \{(x|xx\gamma)\gamma + (x|xx\kappa)\kappa\}x_0^2 + \{(x|yy\gamma)\gamma + (x|yy\kappa)\kappa\}y_0^2 \\ &\quad + \{(x|xx\gamma)\gamma + (x|xx\kappa)\kappa\}x_0x'_0 + \{(x|yy\gamma)\gamma + (x|yy\kappa)\kappa\}y_0y'_0 \\ &\quad + \{(x|x'x'\gamma)\gamma + (x|x'x'\kappa)\kappa\}x_0'^2 + \{(x|y'y'\gamma)\gamma + (x|y'y'\kappa)\kappa\}y_0'^2 \\ &\quad + \{(x|x'x'\gamma)\gamma + (x|x'x'\kappa)\kappa\}x_0'^2 + \{(x|y'y'\gamma)\gamma + (x|y'y'\kappa)\kappa\}y_0'^2 \\ &\quad + \{(x|xy\gamma)\gamma^2 + (x|xy\kappa)\gamma\kappa + (x|xx\kappa)\kappa^2\}x_0 + \{(x|x'\gamma\gamma)\gamma^2 + (x|x'\gamma\kappa)\gamma\kappa + (x|x'\kappa\kappa)\kappa^2\}x'_0 \\ &\quad + (x|y\gamma\gamma)\gamma^3 + (x|y\gamma\kappa)\gamma^2\kappa + (x|y\kappa\kappa)\gamma\kappa^2 + (x|y\kappa\kappa\kappa)\kappa^3 \end{aligned} \quad (6.195)$$

$$\begin{aligned} y_{III} &= (y|xyy)x_0^2y_0 + (y|yyy)y_0^3 \\ &\quad + (y|xyy')x_0^2y'_0 + (y|xx'y)x_0x'_0y_0 + (y|yyy')y_0^2y'_0 \\ &\quad + (y|xx'y')x_0x'_0y'_0 + (y|xy'x'y)x_0'^2y_0 + (y|yy'y')y_0y'_0^2 \\ &\quad + (y|xy'x'y')x_0'^2y'_0 + (y|y'y'y')y_0'^3 \\ &\quad + \{(y|xy\gamma)\gamma + (y|xy\kappa)\kappa\}x_0y_0 + \{(y|xy\gamma)\gamma + (y|xy\kappa)\kappa\}x'_0y_0 \\ &\quad + \{(y|xy\gamma)\gamma + (y|xy\kappa)\kappa\}x_0y'_0 + \{(y|xy\gamma)\gamma + (y|xy\kappa)\kappa\}x'_0y'_0 \\ &\quad + \{(y|yy\gamma)\gamma^2 + (y|y\gamma\kappa)\gamma\kappa + (y|y\kappa\kappa)\kappa^2\}y_0 \\ &\quad + \{(y|y\gamma\gamma)\gamma^2 + (y|y\gamma\kappa)\gamma\kappa + (y|y\kappa\kappa)\kappa^2\}y'_0 \end{aligned} \quad (6.196)$$

again with similar expressions for  $x'_{III}$  and  $y'_{III}$ . The resulting coefficients are too voluminous to be included in this chapter, and, indeed, only a selection is included in the article by Toyoda et al., from

whom the remainder may be obtained on request. These expressions are reproduced here to show what information is available and to indicate the nature of the aberrations of each order and degree. They also illustrate vividly the disadvantage of the trajectory method: the numerous interrelations between the primary aberration coefficients, which can be established by invoking the symplectic condition, are not apparent here as they are in the expressions obtained by the eikonal method.

A very important special case of this basic theory is the family of energy filters for the production of filtered images (images formed with particles in a given energy range) and spectra (energy-loss spectra from a given zone of the specimen and hence from a given region of the image). The early design of Castaing and Henry has been largely superseded by several mirror-free designs, known as  $\Omega$ -filters,  $\alpha$ -filters, S-filters, W-filters, or mandoline filters according to the shape of the optic axis (Figure 6.4). Both magnetic and electrostatic designs are known—here we concentrate on magnetic filters (surveyed by Tsuno, 2001; see also Tsuno et al., 1995 [chicane filter], 1999). The numerous possibilities are shown in Figure 6.4. The full theory of these devices is to be found in the work of Rose (Rose, 1978; Rose and Pejas, 1979; Rose and Krahl, 1995; Plies, 2002; Rose, 2003a, 2008a) and a useful *précis* is given by Rouse et al. (1997). The latter lists the aberration coefficients in terms of generating functions, and it is these expressions that we reproduce here. Note that there are small differences between the definitions and meanings of symbols in the key publications just cited. For recent work on the mandoline filter, see Essers et al. (2008).

The choice of paraxial solutions is dictated by the fact that two sets of planes are of physical significance: first, the *object plane*,  $z = z_{\bar{o}}$ , which is an intermediate image plane conjugate to the object plane of the microscope in which the filter has been placed and second, the *diffraction plane*  $z = z_D$ , conjugate to the diffraction plane  $z = z_d$  of the objective lens of the microscope, which coincides with the back focal plane. The plane  $z = z_{\bar{o}}$  will in turn be conjugate to the final image plane of the instrument whereas the plane  $z = z_D$  will be conjugate to a further diffraction plane, which is referred to as the *slit plane* ( $z = z_s$ ) since it is the plane in which an energy-selecting slit will be placed. One set of paraxial solutions,  $h_x$  and  $h_y$ , satisfies only a slightly modified form of the usual boundary conditions:

$$\begin{aligned} h_x(z_o) &= h_y(z_o) = 0 \\ h'_x(z_{\bar{o}}) &= h'_y(z_{\bar{o}}) = 1 \end{aligned} \quad (6.197a)$$

and hence

$$h_x(z_{\bar{o}}) = h_y(z_{\bar{o}}) = h_x(z_i) = h_y(z_i) = 0 \quad (6.197b)$$

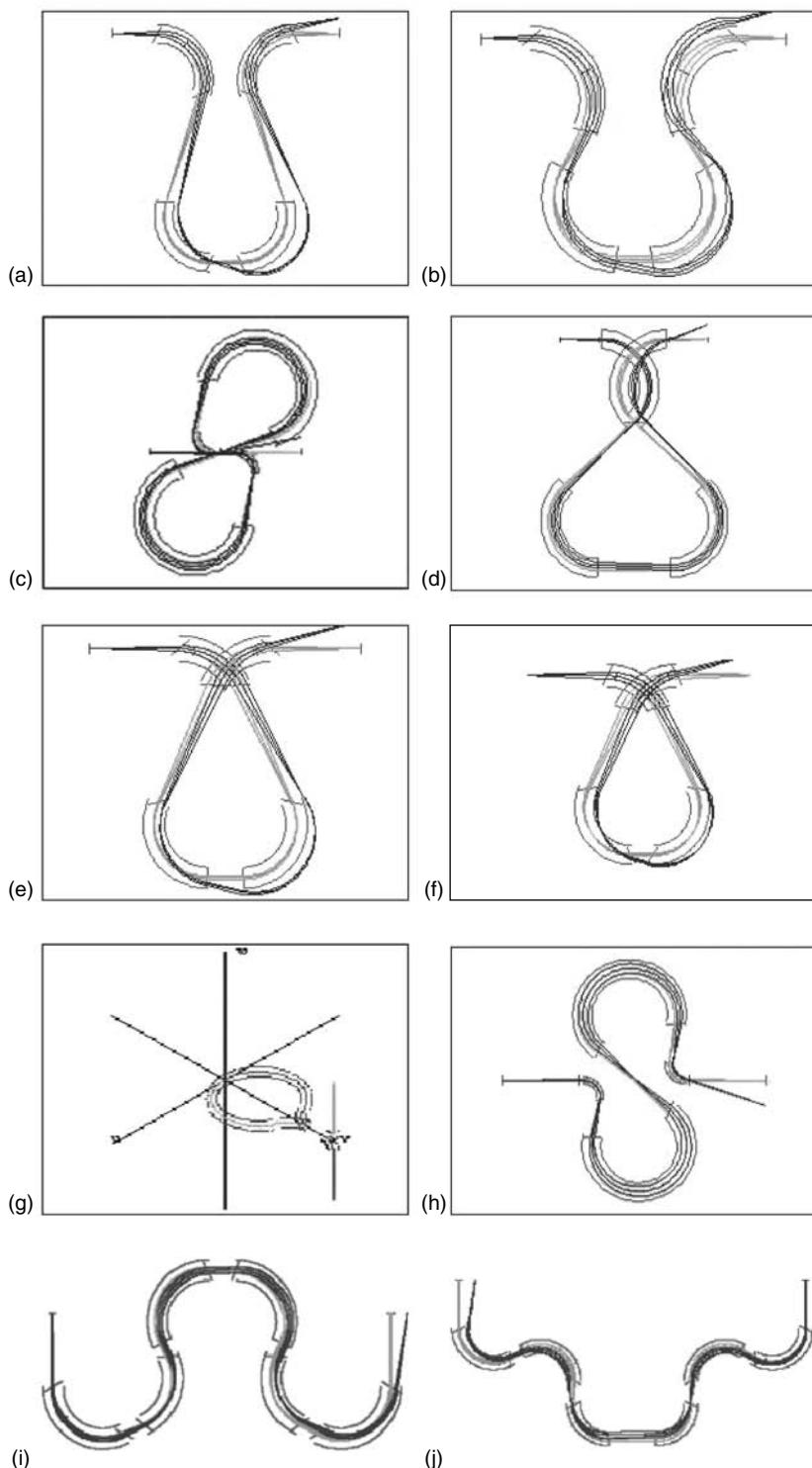
The other set of paraxial solutions,  $d_x(z)$  and  $d_y(z)$ , satisfies the conditions

$$\begin{aligned} d_x(z_D) &= d_y(z_D) = 0 \\ d'_x(z_D) &= d'_y(z_D) = 1 \end{aligned} \quad (6.197c)$$

We follow Rouse et al. in expressing the aberrations in terms of the angles  $\alpha$  and  $\beta$  at the image plane and  $\gamma$  and  $\delta$  at the slit plane (note that Equations 6.192 through 6.196, “ $\gamma$ ” denotes mass variation).

The primary aberrations take the form

$$\begin{aligned} \Delta x(z = z_i) &= \int_{z_o}^{z_i} \{ (\alpha\alpha\alpha)\alpha^2 + 2(\alpha\alpha\gamma)\alpha\gamma + (\alpha\gamma\gamma)\gamma^2 \\ &\quad + \frac{1}{2}(\alpha\beta\beta)\beta^2 + (\alpha\beta\delta)\beta\delta + \frac{1}{2}(\alpha\delta\delta)\delta^2 \\ &\quad + (\alpha\alpha\kappa)\alpha\kappa + (\alpha\gamma\kappa)\gamma\kappa + (\alpha\kappa\kappa)\kappa^2 \} dz \end{aligned}$$



**FIGURE 6.4** (See color insert following page 340.) Tableau showing the optic axes and typical rays in the various imaging filters: (a)  $\Omega$ -filter, A-type; (b)  $\Omega$ -filter, B-type; (c) infinity filter; (d) mandoline filter; (e)  $\alpha$ -filter, A-type; (f)  $\alpha$ -filter, B-type; (g)  $\phi$ -filter; (h) S-filter; (i) twin-column W-filter; and (j) variant twin-column geometry. (Courtesy of Dr. K. Tsuno.)

$$\begin{aligned}\Delta y(z = z_i) = & \int_{z_o}^{z_i} \{(\alpha\beta\beta)\alpha\beta + (\alpha\beta\delta)\alpha\delta + (\gamma\beta\beta)\gamma\beta + (\gamma\beta\delta)\gamma\delta \\ & + (\beta\beta\kappa)\beta\kappa + (\beta\delta\kappa)\delta\kappa\} dz\end{aligned}$$

$$\begin{aligned}\Delta x(z = z_s) = & \int_{z_D}^{z_s} \{(\alpha\alpha\gamma)\alpha^2 + 2(\alpha\gamma\gamma)\alpha\gamma + (\gamma\gamma\gamma)\gamma^2 \\ & + \frac{1}{2}(\gamma\beta\beta)\beta^2 + (\gamma\beta\delta)\beta\delta + \frac{1}{2}(\gamma\delta\delta)\delta^2 \\ & + (\alpha\gamma\kappa)\alpha\kappa + (\gamma\gamma\kappa)\gamma\kappa + (\gamma\kappa\kappa)\kappa^2\} dz\end{aligned}$$

$$\Delta y(z = z_s) = \int_{z_D}^{z_s} \{(\alpha\beta\delta)\alpha\beta + (\alpha\delta\delta)\alpha\delta + (\gamma\beta\delta)\gamma\beta + (\gamma\delta\delta)\gamma\delta + (\beta\delta\kappa)\beta\kappa + (\delta\delta\kappa)\delta\kappa\} dz \quad (6.198)$$

The coefficients appearing here are generated by the functions  $F(x_1, x_2, x_3)$ ,  $G(x_1, y_2, y_3)$ ,  $C_1(x_1, x_2)$ ,  $C_2(y_1, y_2)$ , and  $C_3(x_1)$ , defined by

$$\begin{aligned}F(x_1, x_2, x_3) &= -\frac{1}{24}(4Q_3 + 7B_2Q_2)x_1x_2x_3 - \frac{1}{8}B'_2x_1x_2x'_3 - \frac{1}{2}B_2x_1x'_2x'_3 \\ G(x_1, y_2, y_3) &= \frac{1}{8}(4Q_3 + 3B_2Q_2)x_1y_2y_3 - \frac{3}{8}B'_2x'_1y_2y_3 + \frac{1}{4}B'_2x_1y_2y'_3 - \frac{1}{2}B_2x_1y'_2y'_3 \\ C_1(x_1, x_2) &= -\frac{1}{4}(B_2^2 + Q_2)x_1x_2 \\ C_2(y_1, y_2) &= \frac{1}{4}Q_2y_1y_2 \\ C_3(x_1) &= -\frac{1}{4}B_2x_1\end{aligned} \quad (6.199)$$

and are as follows:

$$\begin{aligned}(\alpha\alpha\alpha) &= 3F(h_x, h_x, h_x) \\ (\alpha\alpha\gamma) &= F(h_x, h_x, d_x) + F(h_x, d_x, h_x) + F(d_x, h_x, h_x) \\ (\alpha\gamma\gamma) &= F(h_x, d_x, d_x) + F(d_x, h_x, d_x) + F(d_x, d_x, h_x) \\ (\gamma\gamma\gamma) &= 3F(d_x, d_x, d_x) \\ (\alpha\beta\beta) &= 2G(h_x, h_y, h_y) \\ (\alpha\beta\delta) &= G(h_x, h_y, d_y) + G(h_x, d_y, h_y) \\ (\alpha\delta\delta) &= 2G(h_x, d_y, d_y) \\ (\gamma\beta\beta) &= 2G(d_x, h_y, h_y) \\ (\gamma\beta\delta) &= G(d_x, h_y, d_y) + G(d_x, d_y, h_y)\end{aligned}$$

$$\begin{aligned}
(\gamma\delta\delta) &= 2G(d_x, d_y, d_y) \\
(\alpha\alpha\kappa) &= 2\{F(h_x, h_x, x_\kappa) + F(h_x, x_\kappa, h_x) + F(x_\kappa, h_x, h_x) + C_1(h_x, h_x)\} \\
(\alpha\gamma\kappa) &= F(h_x, d_x, x_\kappa) + F(h_x, x_\kappa, d_x) + F(d_x, h_x, x_\kappa) + F(d_x, x_\kappa, h_x) + F(x_\kappa, h_x, d_x) + F(x_\kappa, d_x, h_x) \\
&\quad + C_1(h_x, d_x) + C_1(d_x, h_x) \\
(\alpha\kappa\kappa) &= F(h_x, x_\kappa, x_\kappa) + F(x_\kappa, h_x, x_\kappa) + F(x_\kappa, x_\kappa, h_x) + C_1(h_x, x_\kappa) + C_1(x_\kappa, h_x) + C_3(h_x) \\
(\gamma\gamma\kappa) &= 2\{F(d_x, d_x, x_\kappa) + F(d_x, x_\kappa, d_x) + F(x_\kappa, d_x, d_x) + C_1(d_x, d_x)\} \\
(\gamma\kappa\kappa) &= F(d_x, x_\kappa, x_\kappa) + F(x_\kappa, d_x, x_\kappa) + F(x_\kappa, x_\kappa, d_x) + C_1(d_x, x_\kappa) + C_1(x_\kappa, d_x) + C_3(d_x) \\
(\beta\beta\kappa) &= 2G(x_\kappa, h_y, h_y) + 2C_2(h_y, h_y) \\
(\beta\delta\kappa) &= G(x_\kappa, h_y, d_y) + G(x_\kappa, d_y, h_y) + C_2(h_y, d_y) + C_2(d_y, h_y) \\
(\delta\delta\kappa) &= 2G(x_\kappa, d_y, d_y) + 2C_2(d_y, d_y)
\end{aligned} \tag{6.200}$$

Not all analyzers have curved optic axes. The most important type with a straight axis is the Wien filter, in which an electrostatic deflection field is counterbalanced by a magnetic field. The properties of these filters are examined in Section 6.3.8.

### 6.3.8 WIEN FILTERS

The Wien filter is a device in which transverse electric and magnetic fields are present. In its simplest and idealized form, the field distributions are such that electrons of a given energy pass through undeflected; for all other energies, the electrons are deflected and can hence be intercepted. Such a device can be used as a monochromator or as an energy analyzer or even as an aberration corrector. This picture is oversimplified for several reasons, the most obvious of which is that it is not easy to create the necessary matching field distributions. We first present the theory of these devices and then indicate some of the configurations that are being used in practice.

We consider a system described by a rotationally symmetric electrostatic field, electrostatic and magnetic quadrupole, sextupole and octopole components, and transverse fields characterized by  $F_1$  for the electrostatic field and  $B_2$  for the magnetic field (see the expansions 6.6 through 6.8).

The paraxial equations of motion take the form

$$\begin{aligned}
x'' + \frac{\phi'}{2\phi} x' + \left( \frac{\phi''}{4\phi} + \frac{F_1^2}{4\phi^2} - \frac{p_2}{2\phi} + \frac{\eta Q_2}{\phi^{1/2}} \right) x &= \frac{\eta B_2}{\phi^{1/2}} - \frac{F_1}{2\phi} \\
y'' + \frac{\phi'}{2\phi} y' + \left( \frac{\phi''}{4\phi} + \frac{p_2}{2\phi} - \frac{\eta Q_2}{\phi^{1/2}} \right) y &= 0
\end{aligned} \tag{6.201}$$

The deflection terms on the right-hand side of Equations 6.201 vanish if

$$F_1 = 2\phi^{1/2}\eta B_2 \tag{6.202}$$

which may also be written as

$$F_1 = vB_2 \tag{6.203}$$

where  $v$  denotes the velocity given by

$$v = 2\phi^{1/2}\eta \quad (6.204)$$

These are the forms of the *Wien condition* (Wien, 1897/8, 1898). When this condition is satisfied, Equations 6.201 simplify to

$$\begin{aligned} x'' + \frac{\phi'}{2\phi}x' + \left( \frac{\phi''}{4\phi} + \frac{F_1^2}{4\phi^2} - \frac{p_2}{2\phi} + \frac{\eta Q_2}{\phi^{1/2}} \right)x &= 0 \\ y'' + \frac{\phi'}{2\phi}y' + \left( \frac{\phi''}{4\phi} + \frac{p_2}{2\phi} - \frac{\eta Q_2}{\phi^{1/2}} \right)y &= 0 \end{aligned} \quad (6.205)$$

This equation closely resembles the paraxial equations of quadrupoles and, like them, the Wien filter will not have the same strength in the  $x-z$  and  $y-z$  planes. For the system to provide *stigmatic focusing*, another condition must be satisfied:

$$\frac{F_1^2}{4\phi} - p_2 + 2\phi^{1/2}\eta Q_2 = 0 \quad (6.206)$$

All the principal studies of Wien filters derive these two conditions but their subsequent development varies. Some impose the Wien condition but make no immediate further simplifications; others assume that both conditions are satisfied; some proceed on the assumption that the SCOFF approximation is adequate. Here we begin by imposing only the Wien condition and then make further assumptions as special cases. The treatment is nonrelativistic, in view of the fact that electrostatic fields are present. It will soon become apparent that Wien filters are frequently used in pairs or larger groups and we devote a little space to such situations.

Inclusion of electrons with slightly different energies from that of the reference particle adds a further term to the paraxial equations satisfying the Wien condition:

$$\begin{aligned} x'' + \frac{\phi'}{2\phi}x' + \left( \frac{\phi''}{4\phi} + \frac{F_1^2}{4\phi^2} - \frac{p_2}{2\phi} + \frac{\eta Q_2}{\phi^{1/2}} \right)x &= \frac{F_1\kappa}{4\phi_0} \\ y'' + \frac{\phi'}{2\phi}y' + \left( \frac{\phi''}{4\phi} + \frac{p_2}{2\phi} - \frac{\eta Q_2}{\phi^{1/2}} \right)y &= 0 \end{aligned} \quad (6.207)$$

in which

$$\kappa = \Delta\phi/\phi_0 \quad (6.208)$$

It is usual to assume that the electrostatic potential  $\phi$  is constant within the Wien filter so that finally, we have

$$\begin{aligned} x'' + \left( \frac{F_1^2}{4\phi^2} - \frac{p_2}{2\phi} + \frac{\eta Q_2}{\phi^{1/2}} \right)x &= \frac{F_1\kappa}{4\phi_0} + W_{2x} \\ y'' + \left( \frac{p_2}{2\phi} - \frac{\eta Q_2}{\phi^{1/2}} \right)y &= W_{2y} \end{aligned} \quad (6.209)$$

in which we have retained terms of the next higher order in  $x$ ,  $y$ , and their derivatives, from which the primary geometrical aberrations can be derived. First, however, we write down the paraxial

solutions of Equation 6.209 ( $W_{2x} = W_{2y} = 0$ ),

$$\begin{aligned} x^{(1)}(z) &= x_o g_x(z) + x'_o h_x(z) + \kappa x_\kappa(z) \\ y^{(1)}(z) &= y_o g_y(z) + y'_o h_y(z) \end{aligned} \quad (6.210)$$

in which

$$x_\kappa(z) = h_x(z) \int_{z_o}^z \frac{F_1 g_x}{4\phi_0} d\zeta - g_x(z) \int_{z_o}^z \frac{F_1 h_x}{4\phi_0} d\zeta \quad (6.211)$$

We now assume that the system forms a *stigmatic* image in some plane  $z = z_i$  even though the stigmatic focusing condition is not satisfied. On retaining the perturbation terms  $W_{2x}$  and  $W_{2y}$ , we have

$$\begin{aligned} x_i(z) &= x^{(1)}(z_i) + \Delta x(z_i) \\ y_i(z) &= y^{(1)}(z_i) + \Delta y(z_i) \end{aligned} \quad (6.212)$$

in which

$$\begin{aligned} \Delta x &= A_{11}x'^2_o + A_{22}y'^2_o + A_{33}x_o^2 + A_{44}y_o^2 + A_{13}x_o x'_o + A_{24}y_o y'_o \\ &\quad + C_{cx}x'_o \kappa + C_{tx}x_o \kappa + C_{c3}\kappa^2 \\ \Delta y &= B_{11}x_o y_o + B_{22}x'_o y'_o + B_{12}x_o y'_o + B_{21}x'_o y_o + C_{cy}y'_o \kappa + C_{ty}y_o \kappa \end{aligned} \quad (6.213)$$

The aberration coefficients as given by Liu and Tang (1995) are as follows (referred back to the object plane):

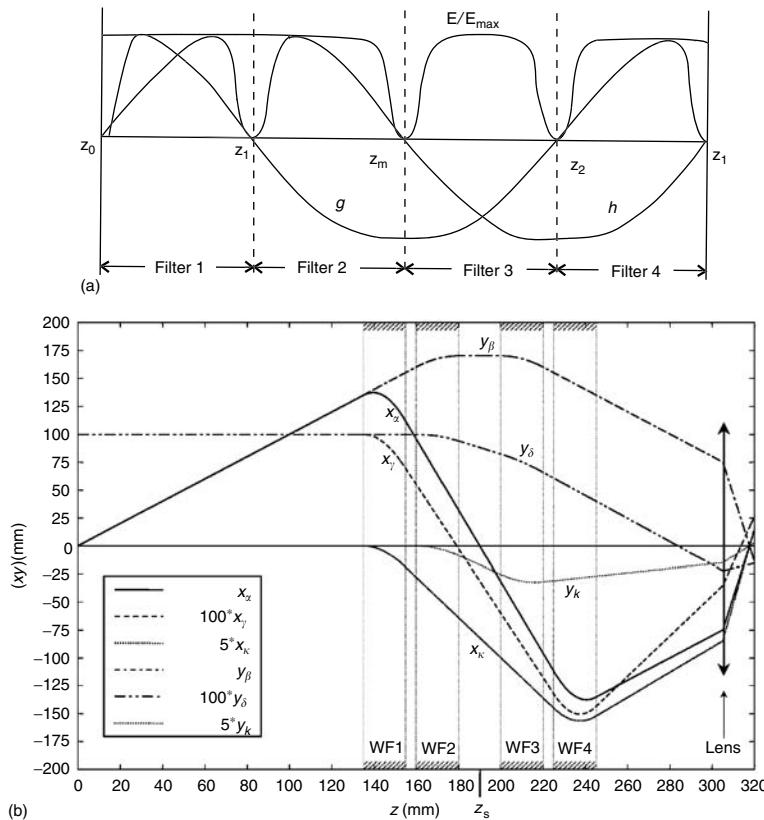
$$\begin{aligned} A_{11} &= - \int_{z_o}^{z_i} (A h_x^2 + Ch_x'^2 + Dh_x h_x') dz \\ A_{22} &= - \int_{z_o}^{z_i} (B h_y^2 - Ch_y'^2 - Dh_y h_y') dz \\ A_{33} &= - \int_{z_o}^{z_i} (A g_x^2 + Cg_x'^2 + Dg_x g_x') dz \\ A_{44} &= - \int_{z_o}^{z_i} (B g_y^2 - Cg_y'^2 - Dg_y g_y') dz \\ A_{13} &= - \int_{z_o}^{z_i} \{2Ag_x h_x + 2Cg_x' h_x' + D(g_x h_x)' h_x\} dz \\ A_{24} &= - \int_{z_o}^{z_i} \{2Bg_y h_y - 2Cg_y' h_y' - D(g_y h_y)' h_y\} dz \\ B_{11} &= - \int_{z_o}^{z_i} \{Gg_x g_y + 2Cg_x' g_y' + D(g_x g_y)' h_y\} dz \end{aligned}$$

$$\begin{aligned}
B_{21} &= - \int_{z_o}^{z_i} \{Gg_y h_x + 2Cg'_y h'_x + D(g_y h_x)' \} h_y dz \\
B_{12} &= - \int_{z_o}^{z_i} \{Gg_x h_y + 2Cg'_x h'_y + D(g_x h_y)' \} h_y dz \\
B_{22} &= - \int_{z_o}^{z_i} \{Gh_x h_y + 2Ch'_x h'_y + D(h_x h_y)' \} h_y dz \\
C_{cx} &= - \int_{z_o}^{z_i} (E + F) h_x^2 dz \\
C_{tx} &= - \int_{z_o}^{z_i} (E + F) g_x h_x dz \\
C_{c3} &= \int_{z_o}^{z_i} E h_x dz \\
C_{cy} &= - \int_{z_o}^{z_i} H h_y^2 dz \\
C_{ty} &= - \int_{z_o}^{z_i} H g_y h_y dz
\end{aligned} \tag{6.214}$$

where

$$\begin{aligned}
A &= \frac{5F_1 p_2}{8\phi_0^2} - \frac{\eta F_1 Q_2}{2\phi_0^{1/2}} - \frac{5F_1^3}{16\phi_0^3} + \frac{F_1''}{8\phi_0} - \frac{p_3}{4\phi_0} + \frac{\eta Q_3}{2\phi_0^{1/2}} \\
B &= -\frac{F_1 p_2}{8\phi_0^2} - \frac{F_1''}{8\phi_0} + \frac{p_3}{4\phi_0} - \frac{\eta Q_3}{2\phi_0^{1/2}} \quad C = \frac{F_1}{4\phi_0} \\
D &= \frac{F_1'}{2\phi_0} \quad E = \frac{5F_1^2}{16\phi_0^2} \quad F = \frac{\eta Q_2}{2\phi_0^{1/2}} - \frac{p_2}{2\phi_0} + \frac{5F_1^2}{16\phi_0^2} \\
G &= \frac{F_1''}{4\phi_0} + \frac{p_3}{2\phi_0} - \frac{\eta Q_3}{\phi_0^{1/2}} + \frac{\eta F_1 Q_2}{2\phi_0^{3/2}} - \frac{F_1 p_2}{2\phi_0^2} \\
H &= \frac{p_2}{2\phi_0} - \frac{\eta Q_2}{2\phi_0^{1/2}}
\end{aligned} \tag{6.215}$$

Liu and Tang also give the terms of third rank to be added to the right-hand side of Equation 6.209. They analyze a device consisting of four Wien filters in series and show that all second-rank aberrations can be eliminated by exploiting the symmetry of the configuration. In view of the importance of symmetry in electron-optical design (see in particular Section 6.4.2), we reproduce the basic rays through the Wien quadruplet (Figure 6.5a); it can be seen that the *g*-ray is symmetric about the midplane of the quadruplet and antisymmetric about the midplanes of the first and last Wien doublets whereas the reverse is true for the *h*-ray. Numerous Wien quadruplets have been



**FIGURE 6.5** Rays through a quadruplet of Wien filters, showing how symmetry is exploited. (a) The Liu–Tang quadruplet, for aberration correction (After Liu, X.D. and Tang, T.-t., *Nucl. Instrum. Meth. Phys. Res. A.*, 363(1–2), 254–260, 1995. Courtesy of the authors and Elsevier.) and (b) the Plies–Bärtle quadruplet for use as a monochromator. (After Plies, E. and Bärtle, J., *Microsc. Microanal.*, 9(Suppl. 3), 28–29, 2003. Courtesy of the authors and Cambridge University Press.)

investigated by Plies and Bärtle, one of which is particularly attractive as a monochromator (Plies and Bärtle, 2003). Here, the focusing sequence is CNNC in one section and NCCN in the other, where C denotes *converging* and N, *neutral*; the fundamental rays are shown in Figure 6.5b. For recent progress see Marianowski and Plies (2008). For the different notations in use, see Table 6.2.

We now examine some other major contributions to Wien filter studies (Table 6.1). Rose (1987b, 1990a) has specialized the very general theory of Plies and Typke (1978) to this situation and shows how the aberration coefficients can be derived from the corresponding eikonal function (whereas Liu and Tang used variation of parameters). The aberration coefficients are not given explicitly by Rose but can be deduced from any of the forms of the eikonal function to be found in his articles. In one of these, all derivatives of the field functions and the paraxial solutions have been eliminated; we have extracted expressions corresponding to those of Liu and Tang (above) and list them here. Note that they are not immediately comparable with Equation 6.214 as not only is a certain amount of partial integration needed to make them identical, but the stigmatic focusing condition must also be imposed on the formulas of Liu and Tang:

$$A_{11} = 3 \int (K_1 + K_2) h^3 dz$$

$$A_{22} = \int (K_2 - 3K_1) h^3 dz$$

$$\begin{aligned}
A_{33} &= 3 \int (K_1 + K_2) g^2 h dz \\
A_{44} &= \int (K_2 - 3K_1) g^2 h dz \\
A_{13} &= 6 \int (K_1 + K_2) g h^2 dz \\
A_{24} &= 2 \int (K_2 - 3K_1) g h^2 dz \\
B_{22} &= 2 \int (K_2 - 3K_1) h^3 dz \\
B_{11} &= 2 \int (K_2 - 3K_1) g^2 h dz \\
B_{12} &= 2 \int (K_2 - 3K_1) g h^2 dz \\
B_{21} &= 2 \int (K_2 - 3K_1) g h^2 dz \\
C_{cx} &= 2 \int 3 \{(K_1 + K_2) h^2 w_\kappa + (K_3 + K_4) h^2\} dz \\
C_{cy} &= 2 \int (K_2 - 3K_1) h^2 w_\kappa + (K_4 - K_3) h^2 \} dz \\
C_{tx} &= 2 \int 3 \{(K_1 + K_2) g h w_\kappa + (K_3 + K_4) g h\} dz \\
C_{ty} &= 2 \int (K_2 - 3K_1) g h w_\kappa + (K_4 - K_3) g h \} dz
\end{aligned} \tag{6.216}$$

**TABLE 6.2**  
**Relations between the Notations Used Here (PWH), by**  
**Rose, and in the Articles Cited here by Tsuno, Scheinfein,**  
**Ioanoviciu and Liu and Tang (T-S-I-L and T)**

PWH	Rose	T-S-I-L and T
$\phi$	$\Phi_o$	
$F_1$	$-\Phi_1^{(r)}$	$E_1$
$F_2$		
$P_2$	$2\Phi_2^{(r)}$	$-2E_2$
$Q_2$		
$P_3$	$-6\Phi_3^{(r)}$	$6E_3$
$Q_3$		
$P_4$	$24\Phi_4^{(r)}$	$-24E_4$
$Q_4$		
$B$		
$B_1$		
$B_2$	$-\Psi_1^{(i)}$	$B_1$
$P_2$		
$Q_2$	$2\Psi_2^{(i)}$	$-2B_2$
$P_3$		
$Q_3$	$-6\Psi_3^{(i)}$	$6B_3$
$P_4$		
$Q_4$	$24\Psi_4^{(i)}$	$-24B_4$

in which

$$\begin{aligned}
 K_1 &= \frac{2\eta\phi^{1/2}Q_3 - p_3}{12\phi} + \frac{F_1p_2}{16\phi^2} - \frac{F_1^3}{64\phi^3} \\
 K_2 &= \frac{F_1}{16\phi} \left( \frac{p_2}{\phi} - \frac{F_1^2}{\phi^2} \right) \\
 K_3 &= \frac{1}{8} \left( \frac{p_2}{\phi} - \frac{F_1^2}{\phi^2} \right) \\
 K_4 &= \frac{F_1^2}{8\phi^2} + \frac{1}{16} \left( 3 \frac{\phi'^2}{\phi^2} + \frac{2\eta^2B^2}{\phi} \right)
 \end{aligned} \tag{6.217}$$

In his first article, Rose (1987b) lists the conditions for which the geometrical terms (second-order aberrations) vanish. The residual chromatic aberrations are then analyzed and practical designs proposed. In his second long article on the general subject of Wien filters, Rose (1990a) examines the possibility of selecting the design parameters in such a way that the device acts as a corrector of spherical and chromatic aberrations. Electrostatic and magnetic round lens terms are retained here, because the device is intended for use as a corrector, but these fields do not overlap those of the Wien filter. Rose then establishes the conditions in which the filter is nondispersing and free of second-rank aberrations. Finally, he investigates the third-rank aberrations; the integrand of the next higher-order contribution to the eikonal function is given, and the aberration integrals for the aperture aberrations are listed and evaluated in the SCOFF approximation. The design of a possible corrector is then examined. His suggestion is exploited in the low-voltage TEM design of Delong and Štěpán (2006).

The contribution of Scheinfein (1989) is interesting because not only are all the second-rank aberration coefficients given in the SCOFF approximation, but the contributions to these coefficients from the fringing fields are also estimated.

Finally, we examine the extensive work of Ioanoviciu, Tsuno and Martínez, much of which antedates the articles by Liu and Tang, Rose and Scheinfein cited above but which is considered last because the SCOFF approximation is employed throughout. Ioanoviciu (1973) considers the model fields that correspond to a cylindrical condenser and a (magnetic) wedge field. The *refractive index* is expanded for these fields and second-rank aberration coefficients are listed explicitly, including approximate formulas for the fringing fields. Turning now to the work of Tsuno, we draw attention to his important article of 1991, an extension of and improvement on the earlier studies of Tsuno et al. (1988–1989, 1990). This article is significant because the fact that, in real Wien filters, the Wien condition is unlikely to be satisfied exactly is recognized explicitly; see also Tsuno (1992, 1993). The paraxial equations and second-rank aberrations are calculated without assuming that the Wien and stigmatic focusing conditions are satisfied. The SCOFF approximation is adopted but the contribution from the fringing fields is also estimated. Later, Tsuno and Martínez, who had published articles on the numerical analysis of Wien filters (Martínez and Tsuno, 2002, 2004a,b), made thorough studies with Ioanoviciu of Wien filters (Ioanoviciu et al., 2004; Tsuno et al., 2003, 2005) and we now recapitulate these in some detail.

A compact notation, well suited to the SCOFF approximation, is introduced. The cyclotron radius  $R$  given by

$$R = \frac{\phi_0^{1/2}}{\eta B_2} \tag{6.218}$$

is used systematically as unit of length:

$$u = \frac{x}{R} \quad v = \frac{y}{R} \quad w = \frac{z}{R} \quad (6.219)$$

and the various field components are scaled with respect to the basic deflection functions,  $F_1$  and  $B_2$ :

$$\begin{aligned} e_2 &= -\frac{p_2 R}{2F_1} \\ b_2 &= -\frac{Q_2 R}{2B_2} \\ e_3 &= \frac{p_3 R^2}{6F_1} \\ b_3 &= \frac{Q_3 R^2}{6B_2} \\ e_4 &= -\frac{p_4 R^3}{24F_1} \\ b_4 &= -\frac{Q_4 R^3}{24B_2} \end{aligned} \quad (6.220)$$

The equations of motion take the form

$$\begin{aligned} u'' + k^2 u &= F_{r1} + F_{r2} + F_{r3} \\ v'' + p^2 v &= F_{v2} + F_{v3} \end{aligned} \quad (6.221)$$

in which

$$\begin{aligned} k^2 &= 2e_2 - 2b_2 + 1 \\ F_{r1} &= \kappa / 2 \\ F_{r2} &= r_1 u^2 + 2uu'' + u'^2/2 + r_2 u\kappa - u''\kappa - \kappa^2/8 + r_3 v^2 - v'^2/2 \\ F_{r3} &= r_4 u^3 + r_5 uu'' + r_6 uu'^2 + r_7 u^2 \kappa \\ &\quad + r_8 u\kappa^2 + \kappa^3/16 + 3u'^2\kappa/4 + r_9 uv^2 \\ &\quad + r_{10} uv'^2 + r_{11} u'vv' + u'v'v'' \\ &\quad + r_{11} u''v^2 - u''v'^2 + r_{12} v^2\kappa + 3v'^2\kappa/4 \end{aligned} \quad (6.222)$$

and

$$\begin{aligned}
 r_1 &= -3e_3 + 3b_3 - 2b_2 - e_2 - 1/2 \\
 r_2 &= b_2 + 1/2 \\
 r_3 &= e_2 + 3e_3 - 3b_3 \\
 r_4 &= -4e_4 + 4b_4 - 3b_3 - e_3 - b_2 - e_2 - 2b_2e_2 - 1/2 \\
 r_5 &= 2e_2 \\
 r_6 &= -2e_2 + 3b_2 - 3/2 \\
 r_7 &= 3b_3/3 + b_2 + e_2/2 + 3/4 \\
 r_8 &= -b_2/4 - 3/8 \\
 r_9 &= 12e_4 - 12b_4 + 3b_3 + 3e_3 + 2e_2b_2 + e_2 \\
 r_{10} &= 3b_2 - 4e_2 - 3/2 \\
 r_{11} &= -2e_2 \\
 r_{12} &= -3b_3/2 - e_2/2
 \end{aligned} \tag{6.223}$$

Similarly

$$\begin{aligned}
 p^2 &= -2e_2 + 2b_2 \\
 F_{v2} &= -h_2v\kappa - v''\kappa + h_1uv + 2uv'' + u'v' \\
 F_{v3} &= h_2v^3 + h_3vv'^2 + h_4v''v^2 + h_5v\kappa^2 + h_6uv\kappa \\
 &\quad + h_7u^2v - h_4u^2v'' - h_4uu'v' + h_8u'^2v - u'^2v'' + u'u''v'
 \end{aligned} \tag{6.224}$$

with

$$\begin{aligned}
 h_1 &= 6e_3 - 6b_3 + 2b_2 \\
 h_2 &= -4e_4 + 4b_4 - 2b_2e_2 \\
 h_3 &= 2e_2 - 3b_2 \\
 h_4 &= -2e_2 \\
 h_5 &= b_2/4 \\
 h_6 &= -3b_3 - b_2 \\
 h_7 &= 12e_4 - 12b_4 + 2b_2e_2 + b_2 + 6b_3 \\
 h_8 &= 4e_2 - 3b_2
 \end{aligned} \tag{6.225}$$

In the SCOFF approximation, we have

$$F_{r_2} = m_0 + m_1 \cos kw + m_2 \sin kw + m_3 \cos^2(kw) + m_4 \sin kw \cos kw \quad (6.226)$$

and

$$\begin{aligned} \Delta u &= p_0 + p_1 \cos kw + p_2 \sin kw + p_3 w \cos kw + p_4 w \sin kw + p_5 \cos^2 kw + p_6 \sin kw \cos kw \\ \Delta v &= q_0 + q_1 \cos kw + q_2 \sin kw + q_3 w \cos kw + q_4 w \sin kw + q_5 \cos^2 kw + q_6 \sin kw \cos kw \end{aligned} \quad (6.227)$$

in which

$$\begin{aligned} p_0 &= (m_0 + 2m_3/3)/k^2 \\ p_1 &= -(m_0 + m_3/3)/k^2 \\ p_2 &= (m_2/2 + m_4/3)/k^2 \\ p_3 &= -m_2/2k \\ p_4 &= m_1/2k \\ p_5 &= -m_3/3k^2 \\ p_6 &= -m_4/3k^2 \end{aligned} \quad (6.228)$$

where

$$\begin{aligned} m_0 &= (u_o - \kappa)^2 / 4 + 2\alpha^2(r_1 - 1) - v_o^2 / 4 + 2\beta^2 r_3 + \kappa^2 \left( r_1 + r_2 - \frac{1}{8} \right) \\ m_1 &= \kappa(u_o - \kappa) \left( 2r_1 + r_2 - \frac{1}{2} \right) \\ m_2 &= \alpha\kappa \left( 2r_1 + r_2 - \frac{1}{2} \right) / k \\ m_3 &= \left( r_1 - \frac{5}{4} \right) \{ (u_o - \kappa)^2 - 2\alpha^2 \} + \left( r_3 + \frac{1}{4} \right) (v_o^2 - 2\beta^2) \\ m_4 &= \left\{ \alpha(u_o - \kappa) \left( 2r_1 - \frac{5}{2} \right) + v_o \beta \left( 2r_3 + \frac{1}{2} \right) \right\} / k \end{aligned} \quad (6.229)$$

and

$$\begin{aligned} q_0 &= (n_0 + 2n_3/3)/k^2 \\ q_1 &= -(n_0 + n_3/3)/k^2 \\ q_2 &= (n_2/2 + n_4/3)/k^2 \\ q_3 &= -n_2/2k \\ q_4 &= n_1/2k \\ q_5 &= -n_3/3k^2 \\ q_6 &= -n_4/3k^2 \end{aligned} \quad (6.230)$$

where

$$\begin{aligned}
 n_0 &= (u_o - \kappa)v_o/2 + 2\alpha\beta(h_l - 1) \\
 n_1 &= v_o\kappa\left(h_l - b_2 - \frac{1}{2}\right) \\
 n_2 &= \beta\kappa\left(h_l - b_2 - \frac{1}{2}\right)/k \\
 n_3 &= \left(h_l - \frac{3}{2}\right)\{(u_o - \kappa)v_o - 2\alpha\beta\} \\
 n_4 &= \left(h_l - \frac{3}{2}\right)\{(u_o - \kappa)\beta + \alpha v_o\}/k
 \end{aligned} \tag{6.231}$$

and  $\alpha$  and  $\beta$  denote  $u'_o$  and  $v'_o$ , respectively.

For stigmatic focusing, we must have

$$e_2 - b_2 = -1/4 \tag{6.232}$$

which implies  $k = 1/\sqrt{2}$ . For  $kw = 2\pi$ , where there is no energy dispersion, we have

$$\begin{aligned}
 \Delta u &= p_0 + p_1 + 2\pi p_2/k + p_5 \\
 \Delta v &= q_0 + q_1 + 2\pi q_2/k + q_5
 \end{aligned} \tag{6.233}$$

But  $p_0 + p_1 + p_5$  and  $q_0 + q_1 + q_5$  vanish and hence all the second-rank aberrations vanish at  $w = 2\pi/k$  or  $z = 2\pi R/k$  if  $p_3$  and  $q_3$  also vanish, which implies that  $m_2$  and  $n_2$  must be zero. If we require that only the geometrical aberrations vanish, then it is sufficient to set

$$p_3 = q_3 \tag{6.234a}$$

or

$$12(e_3 - b_3) + 4b_2 + 2e_2 + \frac{1}{2} = 0 \tag{6.234b}$$

This is the second-order geometric-aberration-free condition. Stigmatic focusing requires that

$$e_2 = -\frac{m+2}{8} \quad b_2 = -\frac{m}{8} \tag{6.235}$$

where  $m$  is arbitrary, whereupon condition 6.234b becomes

$$e_3 - b_3 = \frac{m}{16} \tag{6.236}$$

Returning to expressions 6.233 for  $\Delta u$  and  $\Delta v$ , we find

$$\begin{aligned}
 \Delta u \left( z = \frac{2\pi R}{k} \right) &= -\frac{\alpha\kappa}{8k^2}(m-2) \\
 \Delta v \left( z = \frac{2\pi R}{k} \right) &= -\frac{\beta\kappa}{8k^2}(m-2)
 \end{aligned} \tag{6.237}$$

and so all the second-order terms vanish for

$$m = 2 \tag{6.238a}$$

which implies

$$b_3 - e_3 = -\frac{1}{8} \quad b_2 = -\frac{1}{4} \quad e_2 = -\frac{1}{2} \quad (6.238b)$$

Note that this is not the same as the value proposed by Rose (1987), which corresponds to  $m = 6$  ( $e_2 = -1$ ,  $b_2 = -3/4$ ,  $e_3 - b_3 = 3/8$ ).

In their later article (Tsuno et al., 2005), expressions for the third-rank aberrations are given for two planes: the first focus,  $z = z_i^{(1)}$ , at which  $z = \pi R$ , and the second focus,  $z = z_i^{(2)}$ , at which  $z = 2\pi R$ . There will be both second- and third-rank aberrations at the first plane  $z = z_i^{(1)}$  but only third-rank aberrations at the second plane,  $z = z_i^{(2)}$ . The results are as follows:

$$\begin{aligned} u(z_i^{(1)}) = & -u_0 + 2\kappa + A_{uu}u_0^2 + A_{ud}u_0 + A_{aa}\alpha^2 + A_{ad}\alpha\kappa + A_{dd}\kappa^2 + A_{vv}v_0^2 + A_{bb}\beta^2 \\ & + A_{uuu}u_0^3 + A_{uua}u_0^2\alpha + A_{uud}u_0^2\kappa + A_{uaa}u_0\alpha^2 + A_{uad}u_0\alpha\kappa + A_{udd}u_0\kappa^2 + A_{uvv}u_0v_0^2 \\ & + A_{uvb}u_0v_0\beta + A_{ubb}u_0\beta^2 + A_{aaa}\alpha^3 + A_{aad}\alpha^2\kappa + A_{add}\alpha\kappa^2 + A_{avv}\alpha v_0^2 + A_{avb}\alpha v_0\beta \\ & + A_{abb}\alpha\beta^2 + A_{ddd}\kappa^3 + A_{dvv}\kappa v_0^2 + A_{dvb}\kappa v_0\beta + A_{dbb}\kappa\beta^2 \end{aligned}$$

$$\begin{aligned} v(z_i^{(1)}) = & -v_0 + A_{vu}v_0u_0 + A_{vd}v_0\kappa + A_{ba}\beta\alpha + A_{bd}\beta\kappa + A_{vv}v_0^3 + A_{vvb}v_0^2\beta + A_{vbb}v_0\beta^2 \\ & + A_{vdd}v_0\kappa^2 + A_{vdu}v_0\kappa u_0 + A_{vda}v_0\kappa\alpha + A_{vuu}v_0u_0^2 + A_{vua}v_0u_0\alpha + A_{vaa}v_0\alpha^2 + A_{bbb}\beta^3 \\ & + A_{bdd}\beta\kappa^2 + A_{bdu}\beta\kappa u_0 + A_{bda}\beta\kappa\alpha + A_{buu}\beta u_0^2 + A_{bua}\beta u_0\alpha + A_{baa}\beta\alpha^2 \end{aligned} \quad (6.239)$$

$$\begin{aligned} u(z_i^{(2)}) = & u_0 - 2A_{ad}\alpha\kappa + A_{2uuu}u_0^2\alpha + A_{2uud}u_0^2\kappa + A_{2uad}u_0\alpha\kappa + A_{2udd}u_0\kappa^2 \\ & + A_{2uvb}u_0v_0\beta + A_{2aaa}\alpha^3 + A_{2aad}\alpha^2\kappa + A_{2add}\alpha\kappa^2 + A_{2avv}\alpha v_0^2 \\ & + A_{2abb}\alpha\beta^2 + A_{2ddd}\kappa^3 + A_{2dvv}\kappa v_0^2 + A_{2dvb}\kappa v_0\beta + A_{2dbb}\kappa\beta^2 \end{aligned}$$

$$\begin{aligned} v(z_i^{(2)}) = & v_0 - 2A_{bd}\beta\kappa + A_{2vvb}v_0^2\beta + A_{2vdd}v_0\kappa^2 + A_{2vdu}v_0\kappa u_0 \\ & + A_{2vda}v_0\kappa\alpha + A_{2vua}v_0u_0\alpha + A_{2bbb}\beta^3 + A_{2bdd}\beta\kappa^2 \\ & + A_{2bdu}\beta\kappa u_0 + A_{2bda}\beta\kappa\alpha + A_{2buu}\beta u_0^2 + A_{2baa}\beta\alpha^2 \end{aligned} \quad (6.240)$$

in which

### **$\pi$ -filter, $x$ -direction**

$$A_{uu} = (m - 4)/4$$

$$A_{ud} = -(m - 4)/2$$

$$A_{aa} = m - 6$$

$$A_{ad} = \pi k(m - 2)/2$$

$$A_{dd} = (m - 1)/2$$

$$A_{vv} = (m - 12)/12$$

$$\begin{aligned}
A_{bb} &= (m - 6) / 3 \\
A_{uuu} &= -(m - 4)^2 / 16 \\
A_{uua} &= -3(\pi/k)(b_{31} + t_1) + \pi(9m^2 - 64m + 64) / (256k) \\
A_{uud} &= -2(7b_{31} + 8t_1) + (3m^2 - 20m + 40) / 16 \\
A_{uaa} &= (m - 12)(m - 6) / 4 \\
A_{uad} &= 6(\pi/k)(b_{31} + t_1) + \pi(7m^2 - 96m + 192) / (128k) \\
A_{udd} &= \pi^2(m - 2)^2 / 32 + 4(7b_{31} + 8t_1) - (2m^2 - 9m + 12) / 8 \\
A_{uvv} &= (-5m^2 + 72m - 144) / 144 \\
A_{uvb} &= 3(\pi/k)(b_{31} + 2t_1) + 3\pi m^2 / (128k) \\
A_{ubb} &= -(m - 6)(m + 12) / 36 \\
A_{aaa} &= -6(\pi/k)(b_{31} + t_1) + \pi(9m^2 - 64m + 64) / (128k) \\
A_{aad} &= -8(7b_{31} + 8t_1) + (m^2 + 8m - 60) / 4 \\
A_{add} &= -3(\pi/k)(4b_{31} + t_1) - \pi(11m^2 - 240m + 480) / (256k) \\
A_{avv} &= (3\pi/2k)(b_{31} + 2t_1) + \pi(25m^2 - 192m + 192) / (768k) \\
A_{avb} &= m(m - 6) / 6 \\
A_{abb} &= 9(\pi/k)(b_{31} + 2t_1) + \pi(43m^2 - 192m + 192) / (384k) \\
A_{ddd} &= -(\pi^2/32)(m - 2)^2 - 8(3b_{31} + 4t_1) + (m^2 - 2m + 2) / 8 \\
A_{dvv} &= 2(b_{31} + 8t_1) + (17m^2 - 36m + 72) / 144 \\
A_{dvb} &= -3(\pi/k)(b_{31} + 2t_1) + \pi m(7m - 32) / (384k) \\
A_{dbb} &= 8(3b_{31} + 8t_1) + (19m^2 - 24m - 36) / 36
\end{aligned} \tag{6.241a}$$

### **$\pi$ -filter, y-direction**

$$\begin{aligned}
A_{vd} &= -m / 6 \\
A_{vu} &= m / 6 \\
A_{ba} &= 2(m - 6) / 3 \\
A_{bd} &= \pi k(m - 2) / 2 \\
A_{vvv} &= -m(m - 12) / 144 \\
A_{vvb} &= -3(\pi/k)t_1 - \pi(13m^2 + 192m - 192) / (768k) \\
A_{vbb} &= m(m - 6) / 36
\end{aligned}$$

$$\begin{aligned}
A_{vdd} &= \pi^2/32(m-2)^2 - 4(3b_{31} + 8t_1) - m(16m-3)/72 \\
A_{vdu} &= 4(3b_{31} + 8t_1) + m(17m-12)/72 \\
A_{vda} &= -(3\pi/k)(b_{31} + 2t_1) + \pi(7m^2 - 224m + 384)/(384k) \\
A_{vuu} &= -m(5m-12)/144 \\
A_{vua} &= 3(\pi/k)(b_{31} + 2t_1) + 3\pi m^2/(128k) \\
A_{vaa} &= -(m-6)(m+24)/36 \\
A_{bbb} &= -6(\pi/k)t_1 - \pi(13m^2 + 192m - 192)/(384k) \\
A_{bdd} &= 3(\pi/k)(3b_{31}/2 + 5t_1) + \pi(101m^2 + 208m - 672)/(768k) \\
A_{bdu} &= -3(\pi/k)(b_{31} + 2t_1) - \pi(9m^2 + 32m - 192)/(384k) \\
A_{bdu} &= 8(6b_{31} + 16t_1) + (5m^2 + 8m - 60)/6 \\
A_{buu} &= (3\pi/2k)(b_{31} + 2t_1) + \pi(25m^2 - 192m + 192)/(768k) \\
A_{buu} &= (m-12)(m-6)/6 \\
A_{baa} &= 9(\pi/k)(b_{31} + 2t_1) + \pi(43m^2 - 192m + 192)/(384k)
\end{aligned} \tag{6.241b}$$

### 2π - filter, x - direction

$$\begin{aligned}
A_{2uuu} &= (3\pi/k)(b_{31} + b_{32} + t_1 + t_2) + (\pi/k)(-9m^2/128 + m/2 - 1/2) \\
A_{2uud} &= 2\{7(b_{31} - b_{32}) + 8(t_1 - t_2)\} \\
A_{2uad} &= (\pi m/k)(9m/64 - 1) - (\pi/k)\{6(b_{31} + b_{32} + t_1 + t_2) - 1\} \\
A_{2udd} &= -(2\pi/8)(m-2)^2 - 4\{7(b_{31} - b_{32}) + 8(t_1 - t_2)\} \\
A_{2uvb} &= -3\pi m^2/(64k) - (3\pi/k)\{b_{31} + b_{32} + 2(t_1 + t_2)\} \\
A_{2aaa} &= (6\pi/k)(b_{31} + b_{32} + t_1 + t_2) + (\pi/k)(-9m^2/64 + m - 1) \\
A_{2aad} &= 8\{7(b_{31} - b_{32}) + 8(t_1 - t_2)\} \\
A_{2add} &= (3\pi/k)\{4(b_{31} + b_{32}) + 5(t_1 + t_2)\} + (\pi/k)(-21m^2/128 + 5m/8 - 1/4) \\
A_{2avv} &= \pi m/(2k)(-25m/192 + 1) - (\pi/k)\{(3/2)(b_{31} + b_{32}) + 3(t_1 + t_2) + 1/2\} \\
A_{2abb} &= -(9\pi/k)(b_{31} + b_{32} + 2(t_1 + t_2)) + \pi/k(-43m^2/192 + m - 1) \\
A_{2ddd} &= 8\{3(b_{31} - b_{32}) + 4(t_1 - t_2)\} + \pi^2(m-2)^2/8 \\
A_{2dvv} &= -2\{b_{31} - b_{32} + 8(t_1 - t_2)\} \\
A_{2dvv} &= 3\pi m^2/(64k) + (3\pi/k)\{b_{31} + b_{32} + 2(t_1 + t_2)\} \\
A_{2dbb} &= -8\{3(b_{31} - b_{32}) + 8(t_1 - t_2)\}
\end{aligned} \tag{6.242a}$$

### 2 $\pi$ -filter, y-direction

$$\begin{aligned}
A_{2vvb} &= \pi m / (2k)(13m/192 + 1) + (\pi/k)\{3(t_1 + t_2) - 1/2\} \\
A_{2vda} &= 4\{3(b_{31} - b_{32}) + 8(t_1 - t_2)\} - \pi^2(m-2)^2/8 \\
A_{2vdu} &= -4\{3(b_{31} - b_{32}) + 8(t_1 - t_2)\} \\
A_{2vda} &= 3\pi m^2 / (64k) + (3\pi/k)\{b_{31} + b_{32} + 2(t_1 + t_2)\} \\
A_{2vua} &= -3\pi m^2 / (64k) - (3\pi/k)\{b_{31} + b_{32} + 2(t_1 + t_2)\} \\
A_{2bbb} &= 6\pi/k(t_1 + t_2) + \pi/k(13m^2/192 + m - 1) \\
A_{2bdd} &= -(3\pi/k)\{3(b_{31} + b_{32})/2 + 5(t_1 + t_2)\} + (\pi/k)(-133m^2/384 + 5m/8 - 1/4) \\
A_{2bdu} &= (\pi m/k)(25m/192 - 1) + (3\pi/k)\{b_{31} + b_{32} + 2(t_1 + t_2) + 1/3\} \\
A_{2bda} &= -16\{3(b_{31} - b_{32}) + 8(t_1 - t_2)\} \\
A_{2buu} &= \pi m / (2k)(1 - 25m/192) - 3\pi / (2k)\{b_{31} + b_{32} + 2(t_1 + t_2) + 1/3\} \\
A_{2baa} &= -(9\pi/k)\{b_{31} + b_{32} + 2(t_1 + t_2)\} + \pi/k(-43m^2/192 + m - 1)
\end{aligned} \tag{6.242b}$$

We recall that  $k = \sqrt{(1/2)}$ .

The field component  $b_3$  is selected as an independent parameter and  $e_3$  has been replaced by  $b_3 + m/16$  (Equation 6.236). The sextupole and octopole components of the  $\pi$ -filter and the second half of the  $2\pi$ -filter are shown separately; thus,  $b_{31}$  means  $b_3$  of the  $\pi$ -filter and  $b_{32}$  means  $b_3$  of the second half of the  $2\pi$ -filter and likewise for the other terms. For the octopole components  $e_4$  and  $b_4$ , we write

$$\begin{aligned}
t_1 &= e_{41} - b_{41} \\
t_2 &= e_{42} - b_{42} \\
t &= t_1 + t_2
\end{aligned} \tag{6.243}$$

It is interesting to see that some coefficients vanish when  $b_{31} = b_{32}$  and  $t_1 = t_2$ , whereas other coefficients are canceled when  $b_{31} = -b_{32}$  or  $t_1 = -t_2$ .

These results can be exploited to find configurations with particularly desirable properties. For full discussion, the reader should consult the original papers; here, we simply point out some highlights. Consider the third-order axial aberration (analogous to spherical aberration) at the plane  $z = z_i^{(2)}$ ; for this to vanish, we require

$$A_{2aaa} = A_{2bbb} = 0 \tag{6.244a}$$

for which

$$b_3 = b_{31} + b_{32} = 5m^2/144 \tag{6.244b}$$

and

$$t = t_1 + t_2 = -\frac{13m^2/192 + m - 1}{6} \tag{6.244c}$$

Ioanoviciu et al. plot the corresponding aberration figures for five values of the free parameter  $m$ :  $-2$  (magnetic quadrupole only),  $0$  (electrostatic quadrupole only),  $2$  (all second-rank aberrations absent at  $z = z_i^{(2)}$ ),  $4$  (no particular properties), and  $6$  (all second-order aberrations vanish at  $z = z_i^{(1)}$ ).

To obtain a round beam, we need

$$A_{2aaa} = A_{2bbb} \quad (6.245)$$

which gives the conditions 6.244b and 6.244c, and to this we add

$$A_{2aaa} = A_{2abb} \quad (\equiv A_{2baa}) \quad (6.246a)$$

which ensures that the beam has the same radius in the planes midway between the  $x$ - and  $y$ -axes. This yields

$$t = t_1 + t_2 = -\frac{29}{1152} m^2 \quad (6.246b)$$

For what value of  $m$  does  $A_{2bbb}$  vanish? The answer is the solution to the quadratic equation

$$m^2 - 12m + 12 = 0 \quad (6.247a)$$

namely

$$m = 6 \pm 2\sqrt{6} \quad (6.247b)$$

Tsuno et al. point out that the Wien filter can be used as an aberration corrector, since negative values of the chromatic and aperture aberration coefficients are easily obtained; their design appears to be preferable to the earlier arrangements proposed by Mentink et al.(1999) and Steffen et al. (2000).

For other work on Wien filters, see the works of Ioanoviciu (1974), Ioanoviciu and Cuna (1974), Smith and Munro (1986), Tang (1986b), Kato and Tsuno (1990), Martínez and Tsuno (2004b, 2007), Sakurai et al. (1995), and Tsuno and Rouse (1996) as well as the earlier work of Andersen (1967) and Andersen and Le Poole (1970). We also remind the reader of the work of Plies and Bärtle on the design of monochromators based on Wien filters, already mentioned earlier. Numerous Wien quadruplets were studied and a most interesting configuration emerged, the properties of which are described in Plies and Bärtle (2003).

## 6.4 ABERRATION REPRESENTATION AND SYMMETRY

### 6.4.1 REPRESENTATION

In this section, we draw attention to two important aspects of aberration studies, namely, matrix representation and the effect of system symmetry. The use of matrices to represent aberrations goes back to the work of A.C.S. van Heel's group (see van Heel, 1949 and Brouwer and Walther, 1967) and especially to that of Brouwer (1957). These matrices were introduced into electron optics by Hawkes (for round lenses, see Hawkes, 1970b,d; for quadrupoles, Hawkes, 1970a,c; and for superimposed round lenses and deflection systems, see Hawkes, 1989, 1991). Here, we limit the discussion to round lenses to bring out the nature and purpose of this representation.

Let us introduce column vectors in a pair of conjugate planes,  $z_o$  and  $z_m$ , the subscript  $m$  indicating that the magnification between this pair of planes is  $\mathbf{M}$ . We write

$$\mathbf{u}_o = \begin{pmatrix} u_o \\ u'_o \\ u_o r_o^2 \\ u'_o r_o^2 \\ u_o V_o \\ u'_o V_o \\ u_o \theta_o^2 \\ u'_o \theta_o^2 \\ u_o v_o \\ u'_o v_o \end{pmatrix} \quad (6.248)$$

and likewise for  $\mathbf{u}_m$ , where  $\theta_o^2 = x_o'^2 + y_o'^2$  so that  $\mathbf{u}_o$  and  $\mathbf{u}_m$  are related by

$$\mathbf{u}_m = \mathbf{M}\mathbf{u}_o \quad (6.249)$$

where  $\mathbf{M}$  is a  $10 \times 10$  matrix of the form

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}_1 & \mathbf{M}_2 \\ \mathbf{M}_3 & \mathbf{M}_4 \end{pmatrix} \quad (6.250)$$

and  $\mathbf{M}_1$  is the  $2 \times 2$  paraxial matrix:

$$\mathbf{M}_1 = \begin{pmatrix} M & 0 \\ c & rm \end{pmatrix} \quad (6.251)$$

with

$$c = -1/f_i \quad r = f_o/f_i = (\phi_o/\phi_i)^{1/2} \quad (6.252)$$

The matrix  $\mathbf{M}_2$  has two rows and eight columns; the first row consists of the usual aberration coefficients whereas the second row consists of the set of coefficients that measure aberrations of gradient

$$\mathbf{M}_2 = \begin{pmatrix} Mm_{11} & Mm_{12} & Mm_{13} & Mm_{14} & Mm_{15} & Mm_{16} & Mm_{17} & Mm_{18} \\ m_{21} & m_{22} & m_{23} & m_{24} & m_{25} & m_{26} & m_{27} & m_{28} \end{pmatrix} \quad (6.253)$$

and

$$\begin{aligned} m_{11} &= D + id & m_{15} &= K + ik \\ m_{12} &= F - A & m_{16} &= C \\ m_{13} &= 2A + ia & m_{17} &= a \\ m_{14} &= 2K & m_{18} &= 2k \end{aligned} \quad (6.254)$$

$$\begin{aligned}
m_{21} &= c(D + id) - ri_1 m \\
m_{22} &= c(F - A) - (D - id)rm + c^2 rm / 2 \\
m_{23} &= c(2A + ia) - 2Dr m + c^2 rm \\
m_{24} &= 2cK - (2A - ia)rm + cr^2 m^2 \\
m_{25} &= c(K + ik) - (F - A)rm + cr^2 m^2 / 2 \\
m_{26} &= cC - (K - ik)rm - rm(1 - r^2 m^2) / 2 \\
m_{27} &= ca - 2dr m \\
m_{28} &= 2ck - arm
\end{aligned} \tag{6.255}$$

in which  $i_1$  is defined in Equations 6.111 and 6.112. The matrix  $\mathbf{M}_3$  is null and  $\mathbf{M}_4$ , generated by  $\mathbf{M}_1$ , encapsulates the rules needed for adding the aberration coefficients. It is easily seen that

$$\mathbf{M}_4 = \begin{pmatrix} M^3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ cM^2 & rM & 0 & 0 & 0 & 0 & 0 & 0 \\ cM^2 & 0 & rM & 0 & 0 & 0 & 0 & 0 \\ c^2M & rc & rc & r^2 m & 0 & 0 & 0 & 0 \\ c^2M & 0 & 2rc & 0 & r^2 m & 0 & 0 & 0 \\ c^3 & rc^2 M & 2rc^2 M & 2r^2 cm^2 & r^2 cm^2 & r^3 m^3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & rM & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & rc & r^2 m \end{pmatrix} \tag{6.256}$$

If a second lens (or lens system) is placed behind this first lens, the plane  $z = z_m$  will be conjugate to a further plane,  $z = z_{m'}$ , the magnification between  $z_m$  and  $z_{m'}$  being  $M'$ . Clearly

$$\mathbf{u}_{m'} = \mathbf{M}' \mathbf{u}_m = \mathbf{M}' \mathbf{M} \mathbf{u}_o \tag{6.257}$$

so that if we write

$$\mathbf{P} = \mathbf{M}' \mathbf{M} \quad P = M'M \quad \text{and} \quad p = P^{-1} \tag{6.258}$$

then

$$\mathbf{u}_p = \mathbf{u}_{m'} = \mathbf{P} \mathbf{u}_o \tag{6.259}$$

The matrix  $\mathbf{P}$  must be of the form

$$\mathbf{P} = \begin{pmatrix} \mathbf{P}_1 & \mathbf{P}_2 \\ \mathbf{P}_3 & \mathbf{P}_4 \end{pmatrix} \quad \text{with} \quad \mathbf{P}_1 = \begin{pmatrix} P & 0 \\ c_p & r_p p \end{pmatrix} \tag{6.260}$$

$\mathbf{P}_3$ , like  $\mathbf{M}_3$  is null, and  $\mathbf{P}_4$  has the same structure as  $\mathbf{M}_4$ . The elements of  $\mathbf{P}_2$  must have the same dependence on  $p$  as those of  $\mathbf{M}_2$  have on  $m$ . After a certain amount of calculation, the polynomial

coefficients of the combination are found to be related to those of each of the components as follows:

$$p_{11} = m_{11} + M^2 m'_{11} + Mc(m'_{12} + m'_{13}) + c^2(m'_{14} + m'_{15}) + c^3 mm'_{16}$$

or

$$D_p + id_p = D + id + M^2(D' + id') + Mc(F' + A' + ia') + c^2(3K' + ik') + c^3C'$$

$$p_{12} = m_{12} + rm'_{12} + rcm m'_{14} + 2rc^2 mm'_{16}$$

or

$$F_p = F + rF' + 4rcmK' + 2rc^2mC'$$

$$p_{13} = m_{13} + rm'_{13} + rcm(m'_{14} + 2m'_{15}) + 2rc^2mm'_{16}$$

or

$$2A_p + ia_p = 2A + ia + r(2A' + ia') + rcm(4K' + 2ik') + 2rc^2mC'$$

$$p_{14} = m_{14} + r^2m^2m'_{14} + 2r^2cm^3m'_{16}$$

or

$$K_p = K + r^2m^2K' + r^2cm^3C'$$

$$p_{15} = m_{15} + r^2m^2m'_{15} + 2r^2cm^3m'_{16}$$

or

$$K_p + ik_p = K + ik + r^2m^2(K' + ik') + r^2cm^3C'$$

$$p_{16} = m_{16} + r^3m^4m'_{16}$$

or

$$C_p = C + r^3m^4C'$$

$$p_{17} = m_{17} + rm'_{17} + rcm m'_{18}$$

or

$$a_p = a + ra' + 2rcmk'$$

$$p_{18} = m_{18} + r^2m^2m'_{18}$$

or

$$k_p = k + r^2m^2k' \quad (6.261)$$

For the integrals  $i_j, j = 1-9$ , first appearing in Equations 6.111 and 6.112, we find

$$\begin{aligned}
 i_1^{(p)} &= i_1 + \frac{c^4}{r} \left( i'_1 D_p^4 - 4i'_2 D_p^3 \frac{r'}{c} + 2(i'_3 + i'_4) D_p^2 \left( \frac{r'}{c'} \right)^2 - 4i'_5 D_p \left( \frac{r'}{c'} \right)^3 + i'_6 \left( \frac{r'}{c'} \right)^4 + \frac{1}{2} D_p^3 \left( c'^2 + \frac{1}{D_p^2} \right) \right) \\
 i_2^{(p)} &= \frac{c'}{r'} \left( i_2 D_p - \frac{i_1}{c} \right) + \frac{c^3}{r} \left( i'_2 D_p^3 - (i'_3 + i'_4) D_p^2 \frac{r'}{c'} + 3i'_5 D_p \left( \frac{r'}{c'} \right)^2 - i'_6 \left( \frac{r'}{c'} \right)^3 - \frac{1}{2} D_p \frac{c'}{r'} \right) \\
 i_3^{(p)} &= \left( \frac{c'}{r'} \right)^2 \left( i_3 D_p^2 - 2i_c \frac{D_p}{c} + \frac{i_1}{c^2} \right) + \frac{c^2}{r} \left( i'_3 D_p^2 - 2i'_5 D_p \frac{r'}{c'} + i'_6 \left( \frac{r'}{c'} \right)^2 + \frac{1}{2} D_p \left( \frac{c'}{r'} \right)^2 \right) \\
 i_4^{(p)} &= \left( \frac{c'}{r'} \right)^2 \left( i_4 D_p^2 - 4i_2 \frac{D_p}{c} + 2 \frac{i_1}{c^2} \right) + \frac{c^2}{r} \left( i'_4 D_p^2 - 4i'_5 D_p \frac{r'}{c'} + 2i'_6 \left( \frac{r'}{c'} \right)^2 + D_p \left( \frac{c'}{r'} \right)^2 \right) \\
 i_5^{(p)} &= \left( \frac{c'}{r'} \right)^3 \left( i_5 D_p^3 - (i_3 + i_4) \frac{D_p^2}{c} + 3i_2 \frac{D_p}{c^2} - \frac{i_1}{c^3} \right) + \frac{c}{r} \left( i'_5 D_p - i'_6 \frac{r'}{c'} - \frac{1}{2} D_p \left( \frac{c'}{r'} \right)^3 \right) \tag{6.262} \\
 i_6^{(p)} &= \left( \frac{c'}{r'} \right)^4 \left( i_6 D_p^4 - 4i_5 \frac{D_p^3}{c} + 2(i_3 + i_4) \frac{D_p^2}{c^2} - 4i_2 \frac{D_p}{c^3} + \frac{i_1}{c^4} \right) + \frac{1}{r} \left( i'_6 + \frac{1}{2} D_p \left( D_p^2 \frac{c^2}{r^2} + 1 \right) \right) \left( \frac{c'}{r'} \right)^4 \\
 i_7^{(p)} &= i_7 + c^2 \left( i'_7 D_p^2 - i'_8 D_p \frac{r'}{c'} + i'_9 \left( \frac{r'}{c'} \right)^2 \right) \\
 i_8^{(p)} &= \frac{c'}{r'} \left( -\frac{2i_7}{c} + i_8 D_p \right) + c \left( i'_8 D_p - 2i'_9 \frac{r'}{c'} \right) \\
 i_9^{(p)} &= \left( \frac{c'}{r'} \right)^2 \left( \frac{i_7}{c^2} - \frac{i_8 D_p}{c} + i_9 D_p^2 \right) + i'_9
 \end{aligned}$$

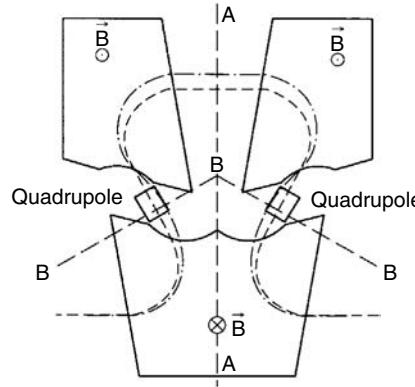
With the aid of these relations, it is a straightforward matter to calculate the aberrations of a system consisting of several round lenses and, more important, to establish the origin of undesirably large terms.

The corresponding addition rules have been established for quadrupole lenses (Hawkes, 1970d) superimposed round lenses, and deflection fields (Hawkes, 1991). The formulas are too space-consuming to be reproduced here. Addition rules for the chromatic aberration coefficients are given in Hawkes (2004).

#### 6.4.2 SYMMETRY

We saw at the beginning of this study of aberrations that the nature of the aberrations of a system is determined by the symmetry about the optic axis. Round electrostatic lenses are characterized by the five real aberration coefficients of glass lenses; round magnetic lenses have slightly lower symmetry with the result that three coefficients become complex. For quadrupoles, the number of coefficients increases considerably.

Symmetry *about* the optic axis is not the only variety, however. Systems consisting of several optical elements may exhibit symmetry or antisymmetry relative to planes *normal* to the optic axis. One of the first examples of the practical benefits to be drawn from studying such symmetry was the



**FIGURE 6.6** Symmetry and antisymmetry planes of an  $\Omega$ -filter.

*Russian quadruplet* of quadrupoles. In the quest for quadrupole multiplets having the same overall effect as a round lens, it was realized early on (Yavor, 1962; Dymnikov and Yavor, 1963; Dymnikov et al., 1964, 1965) that an antisymmetric sequence of quadrupoles would have the same focal lengths in the  $x$ - and  $y$ -directions. For a given geometry, therefore, the excitations could be varied until the foci coincided, whereupon the paraxial matrices in the two planes would be the same. Sets of curves showing the appropriate pairs of excitations for numerous quadrupole geometries have been established; these are known as *load characteristics*.

Another very interesting example is encountered in the study of imaging energy analyzers. In  $\Omega$ -filters, a plane of symmetry naturally arises (A–A in Figure 6.6) and planes of antisymmetry can also be imposed by suitable choice of design (B–B in Figure 6.6). By exploiting such symmetries, many of the aberrations of such filters can be made to vanish. A very detailed account of the design procedure is available (Rose and Krah, 1995), the culmination of many years of gradually improved understanding of these systems. See also Tsuno (1997, 1999, 2001, 2004). Another major contribution is the analysis of the effect of symmetries on aberrations by Hoffstätter (1999), who introduces a new form of transfer map  $\mathbf{M}$  which relates a starting vector  $\zeta_2$  in one plane of the system to its value in another plane  $\zeta_1$ :

$$\zeta_2 = \mathbf{M}\zeta_1 \quad (6.263)$$

The originality of Hoffstätter's work lies in the choice of the elements in the vector  $\zeta$ . He writes  $\zeta^T = (\gamma, \alpha, \delta, \beta, \kappa)$  in which  $\kappa$  is as usual the relative energy deviation. The meaning of the other quantities can be understood from the following equation

$$\begin{aligned} \begin{pmatrix} x(s) \\ y(s) \end{pmatrix} &= \gamma(\gamma_1, \alpha_1, \delta_1, \beta_1, \kappa, s)g_\gamma + \alpha(\gamma_1, \alpha_1, \delta_1, \beta_1, \kappa, s)h_\alpha \\ &\quad + \delta(\gamma_1, \alpha_1, \delta_1, \beta_1, \kappa, s)g_\delta + \beta(\gamma_1, \alpha_1, \delta_1, \beta_1, \kappa, s)h_\beta \\ \begin{pmatrix} p_x(s)/p_0 \\ p_y(s)/p_0 \end{pmatrix} &= \gamma(\gamma_1, \alpha_1, \delta_1, \beta_1, \kappa, s)g'_\gamma + \alpha(\gamma_1, \alpha_1, \delta_1, \beta_1, \kappa, s)h'_\alpha \\ &\quad + \delta(\gamma_1, \alpha_1, \delta_1, \beta_1, \kappa, s)g'_\delta + \beta(\gamma_1, \alpha_1, \delta_1, \beta_1, \kappa, s)h'_\beta \end{aligned} \quad (6.264)$$

Note that the subscripts attached to the fundamental solutions  $g$  and  $h$  have been chosen to remind us of the argument with which they are associated;  $s$  denotes arc-length along the optic axis. This

equation expresses the fact that at each point  $s$  on the optic axis, the transverse position and the normalized momenta  $p_x/p_0$  and  $p_y/p_0$  can be written as a linear combination of the fundamental rays  $g_\gamma$  and  $h_\alpha$  in one section containing the optic axis and  $g_\delta$  and  $h_\beta$  in the other section. As usual, the  $h$ -rays are the axial fundamental rays, those which intersect the axis at the object plane, and the  $g$ -rays are the field rays. Hoffstätter examines the form of the transfer map  $\mathbf{M}$  for a variety of configurations and is able to deduce many of the known findings concerning the cancellation of aberrations by means of system symmetry and also to derive new results.

## 6.5 PARASITIC ABERRATIONS

The aberrations of any real system are not confined to those permitted by its symmetry. In practice, small imperfections in construction and alignment will disturb the ideal symmetry, and these defects will generate additional aberrations; these are known as *parasitic aberrations*. The best known is the axial astigmatism, which arises in round lenses since it is impossible to machine perfectly circular openings. The dominant effect is the same as that of a very weak quadrupole, and this parasitic astigmatism is thus routinely canceled in electron microscopes by means of a stigmator, a device that itself produces a weak quadrupole field. In reality, it has more than four poles so that the orientation of the correcting field can be made to coincide with the astigmatism ellipticity.

The astigmatism is not, however, the only parasitic aberration of importance, particularly when the highest resolution is required. In the past, it has been usual to consider the effect of various kinds of electrode or polepiece imperfection on the potential or field distribution and then to calculate the effects of these field perturbations. In practice, however, the first step is not very helpful, except insofar as it allows us to estimate the precision needed during machining, since many parasitic aberrations are caused by misalignment or poor adjustment and are hence variable in time; they may even vary from day to day. It is more usual nowadays to analyze the aberrations that can arise and to devise ways of measuring and correcting them automatically if the microscope has an advanced computer control system. Here, therefore, we shall not discuss the relation between mechanical imperfections and field perturbations; for extensive analysis, see the works of Glaser (1942), Glaser and Schiske (1953), Sturrock (1949, 1951b), Bertein (1947–8, 1948), Bertein et al. (1947), Archard (1953), Der-Shvarts (1954), Stoyanov (1955a,b), Kanaya and Ishikawa (1959), Kanaya and Kawakatsu (1961) and the account in Kanaya (1985); see also Rose (1968), Amboss and Jennings (1970), Dvořák (2002), Edgcombe (1991), Franzen and Munro (1987), Greenfield et al. (2006), Kasper (1968/9b), Kurihara (1990), Lenz (1987), Munro (1988), Tong et al. (1987), Tsumagari et al. (1986, 1987), Wei and Yan (1999), Ximen and Cheng (1964), Ximen and Li (1990), Ximen and Xi (1964), Yavor (1993), Yavor and Berdnikov (1995), Yin et al. (2007), and Zhu and Liu (1987). In addition to these, see also the works of Hillier and Ramberg (1947), Rang (1949a,b), Recknagel and Haufe (1952/3), Leisegang (1953, 1954a,b, 1956, Ziff. 24), Haine and Mulvey (1954), and Stojanow (1958) on the stigmator, discussed at length by Riecke (1982, section 4.4.2), Glaser (1956, Ziff. 37α), and in Hawkes and Kasper (section 32.4).

The effect of parasitic aberrations is to superimpose on the ideal object–image relation, already perturbed by the geometric and chromatic aberrations of any perfect system, a whole series of additional terms. There is no generally accepted notation for these. In the widely used studies of Saxton (1995), Uhlemann and Haider (1998), and Krivanek et al. (1999a), the aberration coefficients are derived from the wave aberrations (equivalent to the use of the eikonal). Krivanek et al. (1999a) distinguish the aberration coefficients by means of two subscripts, the first denoting the order of the aberration and the second, the symmetry about the optic axis. A further label  $(a, b)$  is needed to separate orthogonal contributions to the same aberration when these are present. Table 6.3 shows the relation between the notations used by Uhlemann and Haider, Saxton, Krivanek et al., and Hawkes and Kasper. The relation between the notation of Uhlemann and Haider and that of Krivanek et al. is clearly seen from the following expressions for the corresponding phase shifts  $(2\pi/\lambda)\chi$ :

**TABLE 6.3**  
**Different Notations for the Parasitic Aberration Coefficients**

Uhlemann and Haider (1998)	Saxton (1995)	Krivanek et al. (1999a)	Hawkes and Kasper (1989)
$C_1$	$C_1$	$C_1$	
$A_1$	$A_1$	$C_{1,2}$	$b_1 + ib_2$
$B_2$	$\frac{1}{3}\bar{B}_2$	$\frac{1}{3}C_{2,1}^*$	$\frac{1}{4}(3A_{30} - 3iA_{03} + A_{12} - iA_{21})$
$A_2$	$A_2$	$C_{2,3}$	$\frac{3}{4}(3A_{30} - 3iA_{03} - A_{12} + iA_{21})$
$C_s = C_3$	$C_3$	$C_{3,0} = C_3$	
$S_3$	$\frac{1}{4}\bar{B}_3$	$\frac{1}{4}C_{3,2}^*$	
$A_3$	$A_3$	$C_{3,4}$	
$B_4$		$\frac{1}{5}C_{4,1}^*$	
$D_4$		$\frac{1}{5}C_{4,3}^*$	
$A_4$		$C_{4,5}$	
$C_5$		$C_{5,0} = C_5$	
$S_5$		$\frac{1}{6}C_{5,2}^*$	
$A_5$		$C_{5,6}$	
$R_5$		$\frac{1}{6}C_{5,4}^*$	

This table includes terms not present in the publications cited, supplied by the authors in question.

#### Uhlemann and Haider

$$\chi = \operatorname{Re} \left\{ \begin{array}{l} \frac{1}{2}w\bar{w}C_1 + \frac{1}{2}\bar{w}^2A_1 + w^2\bar{w}B_2 + \frac{1}{3}\bar{w}^3A_2 + \frac{1}{4}(w\bar{w})^2C_3 + w^3\bar{w}S_3 \\ + \frac{1}{4}\bar{w}^4A_3 + w^3\bar{w}^2B_4 + w^4\bar{w}D_4 + \frac{1}{5}\bar{w}^5A_4 + \frac{1}{6}(w\bar{w})^3C_5 + \frac{1}{6}\bar{w}^6A_5 \\ + w^5\bar{w}R_5 + w^4\bar{w}^2S_5 \end{array} \right\}$$

#### Krivanek et al.

$$\begin{aligned} \chi &= \operatorname{Re} \left\{ \sum \frac{C_{m,n}}{m+1} (w\bar{w})^{(m-n+1)/2} \bar{w}^n \right\} \\ &= \operatorname{Re} \left\{ \begin{array}{l} \frac{1}{2}w\bar{w}C_1 + \frac{1}{2}\bar{w}^2C_{1,2} + \frac{1}{3}w^2\bar{w}C_{2,1}^* + \frac{1}{3}\bar{w}^3C_{2,3} + \frac{1}{4}(w\bar{w})^2C_3 + \frac{1}{4}w^3\bar{w}C_{3,2}^* \\ + \frac{1}{4}\bar{w}^4C_{3,4} + \frac{1}{5}w^3\bar{w}^2C_{4,1}^* + \frac{1}{5}w^4\bar{w}C_{4,3}^* + \frac{1}{5}\bar{w}^5C_{4,5} + \frac{1}{6}(w\bar{w})^3C_{5,0} + \frac{1}{6}\bar{w}^6C_{5,6} \\ + \frac{1}{6}w^5\bar{w}C_{5,4}^* + \frac{1}{6}w^4\bar{w}^2C_{5,2}^* \end{array} \right\} \end{aligned}$$

in which

$$C_{m,n} = C_{m,n,a} + iC_{m,n,b}$$

in the notation of Krivanek et al., and  $w$  denotes  $\omega$  (Uhlemann and Haider) or  $\vartheta_x + i\vartheta_y$  (Krivanek et al.).

Both Uhlemann and Haider and Saxton concentrate on the axial parasitic aberrations (those that are independent of the position in the object plane) though off-axis aberrations are considered in the later part of their studies, by Saxton especially.

In the Uhlemann and Haider notation, the permissible aberrations take the following form:

$$\begin{aligned}
 \Delta x_i + i\Delta y_i = & C_1 w'_o + A_1 \bar{w}'_o \\
 & + w'_o (B_2 w'_o + 2\bar{B}_2 \bar{w}'_o) + A_2 \bar{w}'^2_o \\
 & + C_s w'^2_o \bar{w}'_o + w'_o (S_3 w'^2_o + 3\bar{S}_3 \bar{w}'^2_o) + A_3 \bar{w}'^3_o \\
 & + w'^2_o \bar{w}'_o (2B_4 w'_o + 3\bar{B}_4 \bar{w}'_o) + w'_o (D_4 w'^3_o + 4\bar{D}_4 \bar{w}'^3_o) + A_4 \bar{w}'^4_o \\
 & + C_5 w'^3_o \bar{w}'^2_o + A_5 \bar{w}'^5_o + w'_o (R_5 w'^4_o + 5\bar{R}_5 \bar{w}'^4_o) + 2w'^2_o \bar{w}'_o (S_5 w'^2_o + 2\bar{S}_5 \bar{w}'^2_o)
 \end{aligned} \tag{6.265}$$

The term in  $A_1$  creates an axial *twofold* astigmatism; this was the first parasitic aberration to be well understood and led to the invention of the stigmator. The quadratic terms ( $A_2$  and  $B_2$ ) give rise to axial coma ( $B_2$ ) and to an aberration figure with threefold symmetry ( $A_2$ ). The early article by Kanaya and Kawakatsu (1961) already mentioned that a stigmator with many poles and some degree of independent excitation would be needed to correct both the linear and the quadratic parasitic aberrations.  $C_s$  or  $C_3$  is the familiar spherical aberration and the other third-order terms are a four-fold astigmatism ( $A_3$ ) and a *star* aberration ( $S_3$ ). At fourth order, we have a fivefold astigmatism ( $A_4$ ), a coma ( $B_4$ ), and a three-lobed aberration ( $D_4$ ) whereas at fifth order, there is spherical aberration ( $C_5$ , real), a sixfold astigmatism ( $A_5$ ), and two others,  $S_5$  (star) and  $R_5$ .

There has been a revival of interest in these aberrations in recent years with the development of computer alignment, adjustment, and control of the electron microscope. The strategy to be adopted to recognize and measure the various aberrations and to cancel those that can be eliminated while minimizing any others has been investigated and numerous proposals have emerged. This aspect of aberration studies may be explored through the work of Krivanek, who reawakened awareness of the threefold aberration (Krivanek and Fan, 1992a,b; Krivanek and Leber, 1993, 1994; Krivanek, 1994), as well as the systematic studies of Saxton (Saxton, 1994, 1995; Saxton et al., 1994; Chand et al., 1995) and Uhlemann and Haider (1998) already cited, where references to the extensive earlier work are to be found. Considerable progress is being made in the diagnosis and cancellation of higher-order parasitic and residual aberrations (Hartel et al., 2004). Several ways of measuring the parasitic aberrations have been proposed. We draw attention to the work of Meyer et al. (2002, 2004) and the recent perfected method of Kirkland et al. (2006).

## 6.6 ABERRATION CORRECTION

### 6.6.1 INTRODUCTION

Ever since Scherzer showed in 1936 that the spherical and chromatic aberration coefficients of round lenses cannot be eliminated by skillful design, efforts have been made to find the lens geometry that gives the smallest coefficients and to design correctors of these aberrations. Scherzer (1947) himself listed the various ways of achieving correction, and especially spherical aberration correction, which involved abandoning one or other of the assumptions in the proof that the coefficients cannot be reduced to zero. Meanwhile, several approaches had been discussed by Gabor (1942–43, 1945c) and Zworykin et al. (1945). An excellent account of the earlier attempts has been prepared by Septier (1966) and further surveys are to be found in Hawkes (1980a, 2007) and Hawkes and Kasper (chapter 41). Among the various possibilities, departure from rotational symmetry, the use of electron mirrors, and the introduction of a field discontinuity have been pursued most assiduously. The other techniques, notably those involving the use of a cloud of space charge or of high-frequency fields, have attracted only sporadic attention.

In the 1990s, the situation changed dramatically. All the earlier attempts to correct  $C_s$  by means of quadrupole–octopole elements, some of which had been far reaching, had failed, owing essentially to the difficulty of aligning and controlling systems of such complexity (a quadrupole–octopole corrector typically contains at least four quadrupoles and three octopoles). Nevertheless, the system had been known to be workable in principle since the 1960s (Deltrap, 1964) and Zach and Haider (1994, 1995a,b) therefore attempted to use such a corrector for a less exacting task than the correction of the high-resolution objective of a TEM. Instead, they used the device to correct the spherical aberration of the probe-forming lens of a low-voltage SEM and did indeed succeed in reducing the effect of this aberration on the size of the probe or the current in a probe of given size. In the mid-1990s, Krivanek attacked the difficult task of using a quadrupole–octopole system to correct the spherical aberration of a high-resolution STEM. By this date, fast computer control had become possible and the difficulty that had defeated so many of his predecessors was no longer insurmountable. In 1997, Krivanek described a prototype corrector with which he had succeeded in reducing the effect of spherical aberration in a modified Vacuum Generators (VG) STEM (Krivanek et al., 1997a).

Quadrupoles and octopoles are not the only nonrotationally symmetric elements capable of combating  $C_s$ , however. It had been known since 1965 that sextupoles possess an aberration that could be used to cancel the spherical aberration of a round lens, but this idea was not pursued for many years, since sextupoles exhibit second-order effects and it thus seemed difficult to exploit their third-order spherical-type aberration for correction. Many years later, it was realized that the unwanted second-order effects can be canceled by combining two sextupoles in a suitable way, and this led to a large body of work on sextupole correctors; these should be less difficult to operate than quadrupole–octopole correctors since they have no linear effect. Moreover, they should be advantageous for microscopes of the TEM type in which a relatively large area is focused simultaneously (and not sequentially, as in scanning instruments) because the electron beam is not despatched so far from the optic axis as it is in a quadrupole–octopole corrector. That a sextupole corrector can correct the spherical aberration of the objective lens of a TEM was demonstrated in 1998 by Haider et al. (1998a–c).

The effect of chromatic aberration can be reduced in two ways: the aberration coefficient itself can be canceled by means of a corrector, as is done for the spherical aberration; alternatively, the energy spread of the electron beam can be narrowed so that  $C_c \Delta\phi/\phi$  is smaller. This second approach has generated a substantial literature on electron monochromators, which will not be examined here, though it is covered indirectly in Section 6.3.8—abundant information is to be found in the papers of Plies (1978), Tang (1986b), Tsuno et al. (1988–9, 1990, 2003, 2005), Tsuno (1991, 1992, 1993, 1999), Reimer (1995), Tsuno and Rouse (1996), Kahl and Rose (1998, 2000), Huber and Plies (1999, 2000), Mook and Kruit (1998, 1999a,b, 2000), Mook et al. (1999, 2000), Tiemeijer (1999a,b), Batson et al. (2000), Kahl and Voelkl (2001), Martínez and Tsuno (2002, 2004a,b), Tiemeijer et al. (2002), Benner et al. (2003a–c), Mukai et al. (2003a,b), Plies and Bärtle (2003), Ioanoviciu et al. (2004), Huber et al. (2004), Freitag et al. (2004a,b, 2005), Bärtle and Plies (2005, 2006, 2007), Browning et al. (2006), Walther and Stegmann (2006a,b), Walther et al. (2006), van Aken et al. (2007), Essers et al. (2007), Irsen et al. (2007), Matijevic et al. (2007), and Ringnalda et al. (2007).

Chromatic aberration correctors are an active area of research. As early as 1947, Scherzer showed that an electrostatic round lens could be combined with an electrostatic quadrupole in such a way that the overall chromatic aberration coefficient of the two would be negative. This idea was taken up by Archard (1955), a close student of Scherzer's thinking, and has recently been revived.

A second approach is based on the observation that the sign of the coefficient of chromatic aberration of a combined electrostatic–magnetic quadrupole lens can be positive or negative. Wien filters too can be used for chromatic compensation. A last possibility is to incorporate an electron mirror into the imaging sequence, since once again, the coefficient can have either sign.

### 6.6.2 MINIMIZATION

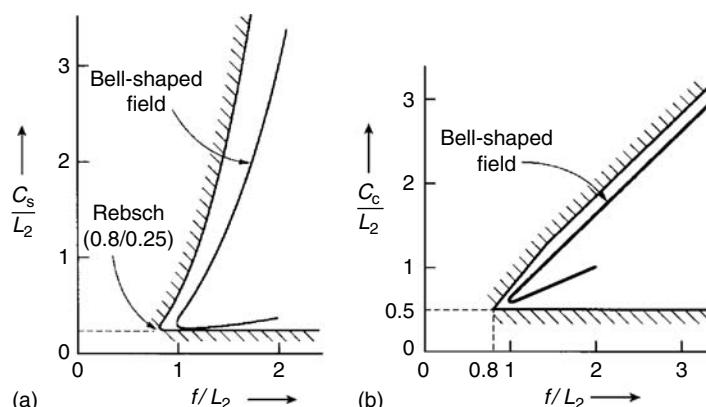
Before discussing correction in detail, we first mention attempts to optimize round lenses. The essential document here is an article by Tretner (1959), the culmination of a series of articles centered on the same topic, in which the domains within which the coefficients of spherical and chromatic aberrations must fall in real electrostatic and magnetic lenses are established. His findings are reproduced in full in Hawkes and Kasper (section 36.4.1) but for the derivation, based on the calculus of variations, the original publication (in German) must be consulted. Typical results are shown in Figure 6.7; Figure 6.7a shows the limiting value of  $C_s/L_2$  for magnetic lenses as a function of  $f/L_2$ , where  $f$  is the objective focal length. The length  $L_2$  is given by

$$L_2 = \frac{1}{|B'/B|_{\max}}$$

and is hence defined in terms of the maximum relative magnetic flux density gradient. Figure 6.7b shows  $C_c/L_2$ . Glaser's bell-shaped field is included, for comparison, and the corner points are the theoretical minima.

Several attempts have been made to establish the electrode or polepiece shape of the *best* lens. An early design was proposed by Sugiura and Suzuki (1943), and with the advent of computers, the dependence of lens quality on geometry could be explored systematically. We mention in particular the work of Moses (1973, 1974; Rose and Moses, 1973), who used the variational method to establish magnetic field distributions with minimum spherical aberration and no coma. Like Sugiura and Suzuki, Moses found that the field should fall off slowly on the image side. Later, Szilagyi used dynamic programming in the search for optimum field shapes; for a connected account of that work, see Szilagyi (1988, chapter 9).

These attempts to find optimized lenses result not in a lens design but in a field distribution. The final stage in which the electrode or polepiece shapes that will create such distributions are established is far from easy and has created much acrimonious dispute. We shall not comment on it here; it is now possible to calculate field distributions and lens properties so rapidly that the lens boundaries can be adjusted interactively until the desired distribution is attained (e.g., Hill and Smith, 1980, 1982; Tsuno and Smith, 1985, 1986; Taylor and Smith, 1986; Glatzel and Lenz, 1988; Szép and Szilagyi, 1988, 1990; Lencová and Wisselink, 1990; Lencová, 1995; Barth et al., 1995; Martínez and Sancho, 1995; Degenhardt, 1997; Edgcombe et al., 1999). A further step was taken by van der Stam and Kruit (1995, 1999) and van der Stam et al. (1993), who created a software tool for the predesign of optical systems.



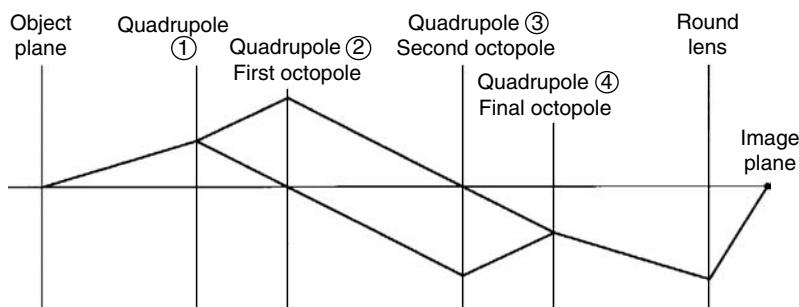
**FIGURE 6.7** The domains within which the (a) spherical and (b) chromatic aberration coefficients must lie.

All these studies are concerned with magnetic or electrostatic lenses but rarely with both simultaneously. Designers have no doubt been deterred by the awareness that  $C_s$  is apparently always worse when both types of field are present than when only one is active: the formula for  $C_s$  seems to show this clearly. However, this impression is misleading because the comparison is valid only if the paraxial solutions are the same in the two situations. This is, of course, extremely unlikely since the paraxial equations are not the same. In fact, mixed lenses can be superior to purely electrostatic or purely magnetic lenses, as the work of Yada (1986) and the much fuller studies of Frosien and Plies (1987), Plies and Schweizer (1987), Frosien et al. (1989), Plies and Elstner (1989), Preikszas (1990), Hordon et al. (1993), and Weimer and Martin (1994) show; see also Preikszas and Rose (1995), Zhukov et al. (2004), and the survey by Plies (1994).

### 6.6.3 CORRECTION

#### 6.6.3.1 Quadrupoles and Octopoles

The derivation of expressions 6.40b and 6.45 for  $C_s$ , which show that this coefficient cannot change sign, is valid only if certain conditions are satisfied: the lens must be static, rotationally symmetric, space-charge-free, and free of discontinuities in the axial potential  $\phi(z)$  and its derivative  $\phi'/\phi$ . Correction involves abandoning at least one of these conditions. The least traumatic involves departing from rotational symmetry and the early workers soon showed that spherical aberration could be eliminated with the aid of quadrupoles and octopoles. Either they could be combined into a complicated unit that would replace the objective lens altogether or they could be the basis for a corrector, to be used in conjunction with a traditional objective. Scherzer's (1947) early explanation of the principle remains the simplest. Consider a combination of four quadrupoles and three octopoles (Figure 6.8). Two of the octopoles are placed at line foci and hence modify two of the quadrupole aberration coefficients independently. The third octopole is used to correct the third coefficient after which some adjustment of the others will be required. In practice, it will not be necessary to separate all these different field components, and Burfoot (1953) showed that, in the electrostatic case, a single four-electrode lens has all the degrees of freedom needed for correction. The numerous attempts to incorporate such correction into practical instruments is charted in the work of Septier (1966), Hawkes (1980a), and in Hawkes and Kasper (chapter 41). The task of using such a device to correct the objective lens of a high-resolution TEM is formidable, for the quadrupoles introduce new aberrations into an already highly perfected system, and the role of the octopoles is to cancel the combination of these new aberrations and those of the objective lens. It is more reasonable to use a quadrupole-octopole corrector to improve the performance of a lens that is not operating at the limits of its performance and it was in such an attempt that Zach and Haider succeeded in improving the characteristics of a low-voltage SEM (Webster et al., 1988; Zach and Haider, 1994, 1995a,b; Zach, 2000). Soon after, a corrector of the same type but with more quadrupoles was designed for use with a STEM and, thanks



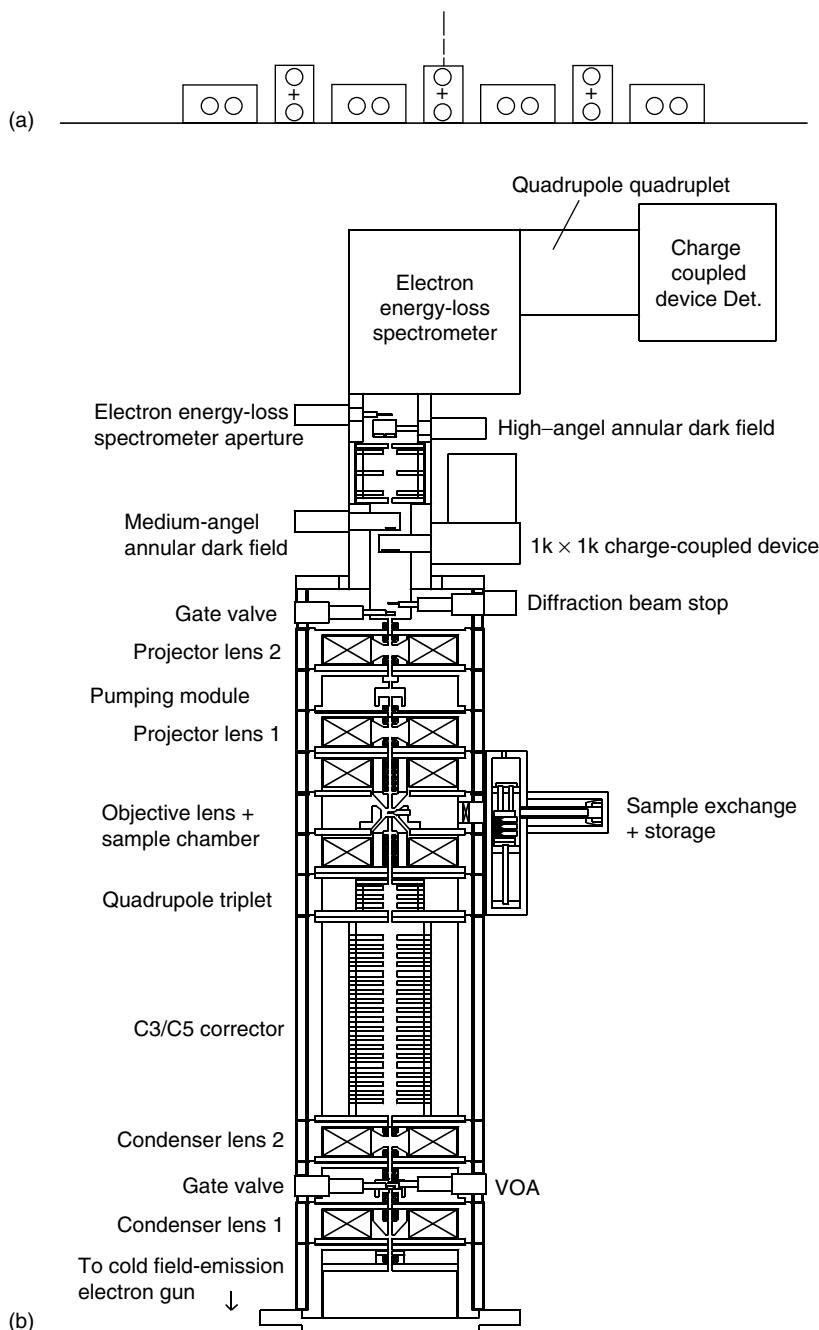
**FIGURE 6.8** The simplest arrangement for correcting spherical aberration by means of four quadrupoles and three octopoles.

to fast computer diagnostics and control, showed itself capable of reducing the probe size, or alternatively allowing more current into the original uncorrected probe (Krivanek et al., 1997a,b). Since then, many improvements have been made to the original design, and several VG STEMs have been equipped with this type of corrector. The Krivanek (Nion) corrector consists of the basic quadrupoles and octopoles, all under computer control, together with other multipole fields designed to compensate for misalignments and parasitic aberrations in general. In the second-generation Nion corrector, 16 quadrupoles as well as three combined quadrupole–octopole elements are used. An additional quadrupole triplet is situated between the corrector and the probe-forming lens. The corrector itself consists of an alternating sequence of quadrupole quadruplets and quadrupole–octopole elements (Figure 6.9). With this arrangement, the center planes of the quadrupole–octopole elements are all conjugates and are also conjugate to a plane close to the coma-free plane of the probe-forming lens. All the fifth-order geometrical aberrations of the combination of corrector and probe-forming lens can then be eliminated. It is not surprising to find that the optics of this configuration is similar to that of the Rose ultracorrector, described in more detail in Section 6.6.3.4 (Rose, 2003b, 2004, 2005, 2006). The excitations of the various components are adjusted systematically by the software. The evolution of the corrector, which has been fitted to many VG STEMs, to the SuperSTEMs at Daresbury and to the Nion STEM, can be studied in the following publications: Krivanek et al. (1997a,b, 1998, 1999a,b, 2000, 2001, 2002, 2003, 2004, 2005a,b, 2006, 2007a,b, 2008); Dellby et al. (2000, 2001, 2005, 2007), Batson et al. (2002), Batson (2003), Lupini et al. (2003), Pennycook et al. (2003, 2006a,b), Nellist et al. (2004a,b, 2006), Bacon et al. (2005), Bleloch et al. (2005, 2007), Varela et al. (2005) and Lupini and Pennycook (2007). The quadrupole–octopole corrector designed for an FEI STEM/TEM is described by Mentink et al. (2004). See also Joy (2008).

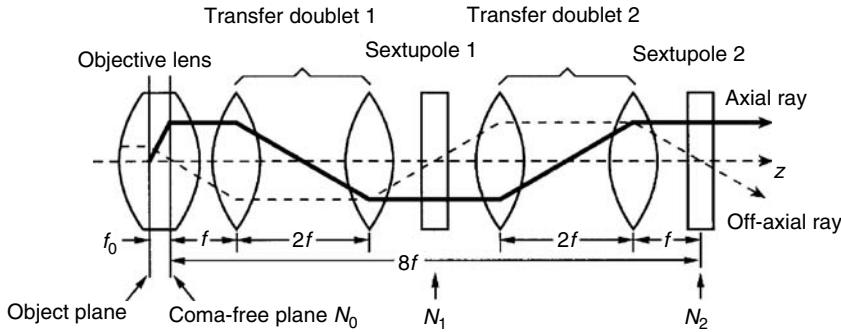
A newly designed STEM into which correction is incorporated from the outset is being produced by the Nion company (see Chapter 12 by Krivanek and Krivanek et al. 2007a,b, 2008a–c). For subsequent use of a quadrupole–octopole corrector in the SEM and related studies, see Zach (2006), Honda et al. (2004a,b), Uno et al. (2004a,b, 2005), Kazumori et al. (2004a,b) and Baranova et al. (2004); an extremely thorough treatment is to be found in (Matsuya and Nakagawa, 2004), where parasitic and residual aberrations are studied and the presence of retarding fields is also considered. Various model fields are used (the Glaser–Schiske model for symmetric electrostatic lenses and a tanh model for the retarding field).

### 6.6.3.2 Sextupoles

Another way of correcting  $C_s$  by abandoning rotational symmetry involves the use of sextupole lenses, which have long been known to have an aberration capable of canceling that of round lenses (Hawkes, 1965b). Combinations of sextupoles that should in theory provide correction have been studied by Beck (1979), Crewe and Kopf (1980a,b), Crewe (1980, 1982, 1984), Rose (1981, 1990b), Ximen (1983), Ximen and Crewe (1985), Shao (1988a,b), Shao et al. (1988), and Chen and Mu (1991). As we have seen, the second-order effect of a sextupole is characterized by four terms of the form  $\int H(z)h^{3-n}k^n dz$ ,  $n = 0, 1, 2$ , and 3, in which  $H(z)$  represents the field distribution in the (electrostatic and magnetic) sextupole and  $h(z)$  and  $k(z)$  are two linearly independent solutions of the standard paraxial equation for round lenses (these solutions collapse to straight lines in the absence of any round lens component). All four terms can be made to vanish by suitable choice of the symmetry of the configuration, the simplest of which is shown in Figure 6.3. When coupling such a device to a microscope objective, the coma-free condition must be satisfied. The (isotropic) coma-free plane of an objective is situated within the lens field and must hence be imaged onto the front focal plane of the round-lens doublet in the corrector by means of another doublet (Figure 6.10). If the anisotropic coma must be eliminated as well as the isotropic coma, an objective design in which two coils are used in tandem could be employed (Rose, 1971b). In the 1990s, a corrector of this type was incorporated into a Philips CM20 TEM (Uhlemann et al., 1994; Haider et al., 1994, 1995) and this led to the successful correction of  $C_s$  in a commercial instrument by Haider et al. (1998a–c). A large



**FIGURE 6.9** The Nion quadrupole–octopole corrector incorporated in scanning transmission electron microscopes (STEMs). (a) The sequence of quadrupole quadruplets and quadrupole–octopoles of which the corrector is composed, Q–Q: quadrupole quadruplet; Q–O: quadrupole–octopole element and (b) schematic view of the corrector incorporated in a STEM. (After Krivanek, O.L. et al., *Proc. 13th Eur. Microscopy Cong.*, Antwerp, Belgian Society for Microscopy, Liège, 2004. Courtesy of the authors and the Belgian Microscopy Society.)



**FIGURE 6.10** Correction of spherical aberration by means of sextupoles.

literature has grown up around these devices, manufactured by CEOS: sextupole correction may be traced in the articles (in addition to the early publications already cited) of Haider et al. (1982, 1994, 1995, 1998a–c, 2000, 2004, 2006a–d, 2008a–c), Rose (1990b, 2002a,b), Haider and Uhlemann (1997), Haider (1998, 2000, 2003), Rose et al. (1998), Foschepoth and Kohl (1998), Uhlemann et al. (1998), Urban et al. (1999), Müller et al. (2002, 2005, 2006, 2007), Kabius et al. (2002), Lentzen et al. (2002), Liu et al. (2002), Benner et al. (2003a–c, 2004a,b), Chang et al. (2003, 2006), Hosokawa and Sawada (2003), Hosokawa et al. (2003, 2006), Jia et al. (2003), Sawada et al. (2004a–c), Hartel et al. (2004, 2007a,b), Titchmarsh et al. (2004), Haider and Müller (2005), Hutchison et al. (2005), Thust et al. (2005), Blom et al. (2006), Kaneyama et al. (2006), Lentzen and Thust (2006), Mitsuishi et al. (2006a,b), Walther and Stegmann (2006a,b), Walther et al. (2006), Nakamura et al. (2006), Watanabe et al. (2006), Chen et al. (2007), Baranova et al. (2007), Freitag et al. (2007), Ringnalda et al. (2007), Kirkland (2007), Irsen et al. (2007), Sawada et al. (2007), and Mayer et al. (2007).

### 6.6.3.3 Correctors of Chromatic Aberration

#### 6.6.3.3.1 All-Electrostatic Correctors

The chromatic aberration coefficients of a system containing electrostatic round lens fields and electrostatic quadrupole fields can be written as

$$\begin{aligned} C_{cx} &= \sqrt{\phi_0} \int \frac{1}{\phi^{1/2}} \left( h_x'^2 - \frac{\phi'}{2\phi} h_x h_x' \right) dz \\ C_{cy} &= \sqrt{\phi_0} \int \frac{1}{\phi^{1/2}} \left( h_y'^2 - \frac{\phi'}{2\phi} h_y h_y' \right) dz \end{aligned} \quad (6.266)$$

The Picht transformation

$$H_x = \left( \frac{\phi}{\phi_0} \right)^{1/4} h_x \quad H_y = \left( \frac{\phi}{\phi_0} \right)^{1/4} h_y \quad (6.267)$$

transforms these to

$$\begin{aligned} C_{cx} &= \phi_0 \int \frac{1}{\phi} \left[ \left( H_x' - \frac{\phi'}{2\phi} H_x \right)^2 - \frac{1}{16} \frac{\phi'^2}{\phi^2} H_x^2 \right] dz \\ C_{cy} &= \phi_0 \int \frac{1}{\phi} \left[ \left( H_y' - \frac{\phi'}{2\phi} H_y \right)^2 - \frac{1}{16} \frac{\phi'^2}{\phi^2} H_y^2 \right] dz \end{aligned} \quad (6.268)$$

If the quadratic term could be reduced to zero (or at least made smaller than the absolute value of the negative term), the overall chromatic aberration coefficient would be negative, and the device would hence be suitable for use as a corrector. The quadratic term vanishes if the field distribution is such that the paraxial ray  $H_x(z) \propto \phi^{1/2}$  and similarly for  $H_y(z)$ . In general, this is not an acceptable form for  $H_x(z)$  and  $H_y(z)$ , which must vanish in the object (and image) planes but if the corrector is telescopic, such a form is permissible. By substituting this expression for  $H_x(z)$  and  $H_y(z)$ , and hence  $h_x(z)$  and  $h_y(z)$ , back into the paraxial equation, it is found that the field functions  $\phi(z)$  and  $p_2(z)$  must be related by the formula

$$p_2 = \phi'' - \frac{\phi'^2}{8\phi} \quad (6.269)$$

Configurations in which this condition (Scherzer, 1947) is closely satisfied have been found by Weißbäcker and Rose (2000, 2001, 2002) and Maas et al. (Maas et al., 2000, 2001, 2003; Henstra and Krijn, 2000). Weißbäcker and Rose have investigated several geometries in which the complexity increases with the practical usefulness of the corrector. In the simplest design, four distinct elements are employed and the corrector is capable of correcting both the chromatic and the spherical aberrations in a scanning instrument. The first component is a quadrupole, which renders the incoming beam astigmatic. This is followed by a three-element corrector consisting of quadrupole fields superimposed on a (round) einzel lens field. A second corrector unit cancels the chromatic aberration in the other plane. Octopoles are incorporated so that the spherical aberration can be corrected simultaneously.

Unfortunately, such a corrector introduces large off-axis aberrations and is hence not suitable for use with (fixed-beam) TEMs. Weißbäcker and Rose therefore considered first an extended version in which a third such element is added and then proposed a doubly symmetric electrostatic corrector (DECO) in which each correcting element is enclosed within two quadrupole doublets (Figure 6.11). The symmetry conditions can now be arranged in such a way that the chromatic aberration and the coma vanish, whereas the spherical aberration is corrected by means of octopoles as usual.

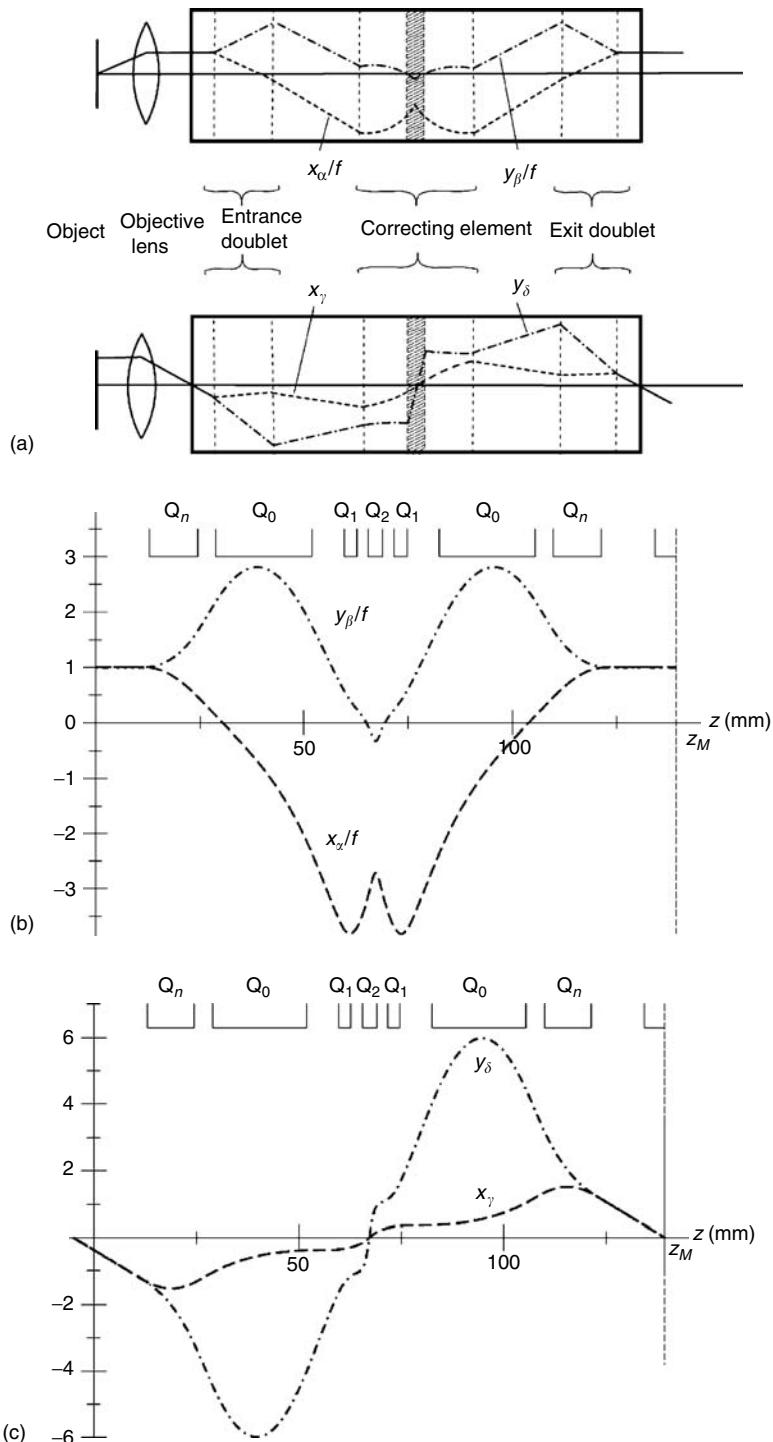
In the parallel investigations of Henstra, Maas, Mentink, and coworkers, a configuration consisting of nine elements (Figure 6.12) is explored. At the outer extremities are quadrupoles to create astigmatism and subsequently annihilate it. In the center are five combined round lens and quadrupole units. Two other quadrupoles are included to match the slowly decaying round lens field of the first and last combined elements. For further work on this approach, see Baranova et al. (2002, 2004).

#### 6.6.3.3.2 Mixed Quadrupole Correctors

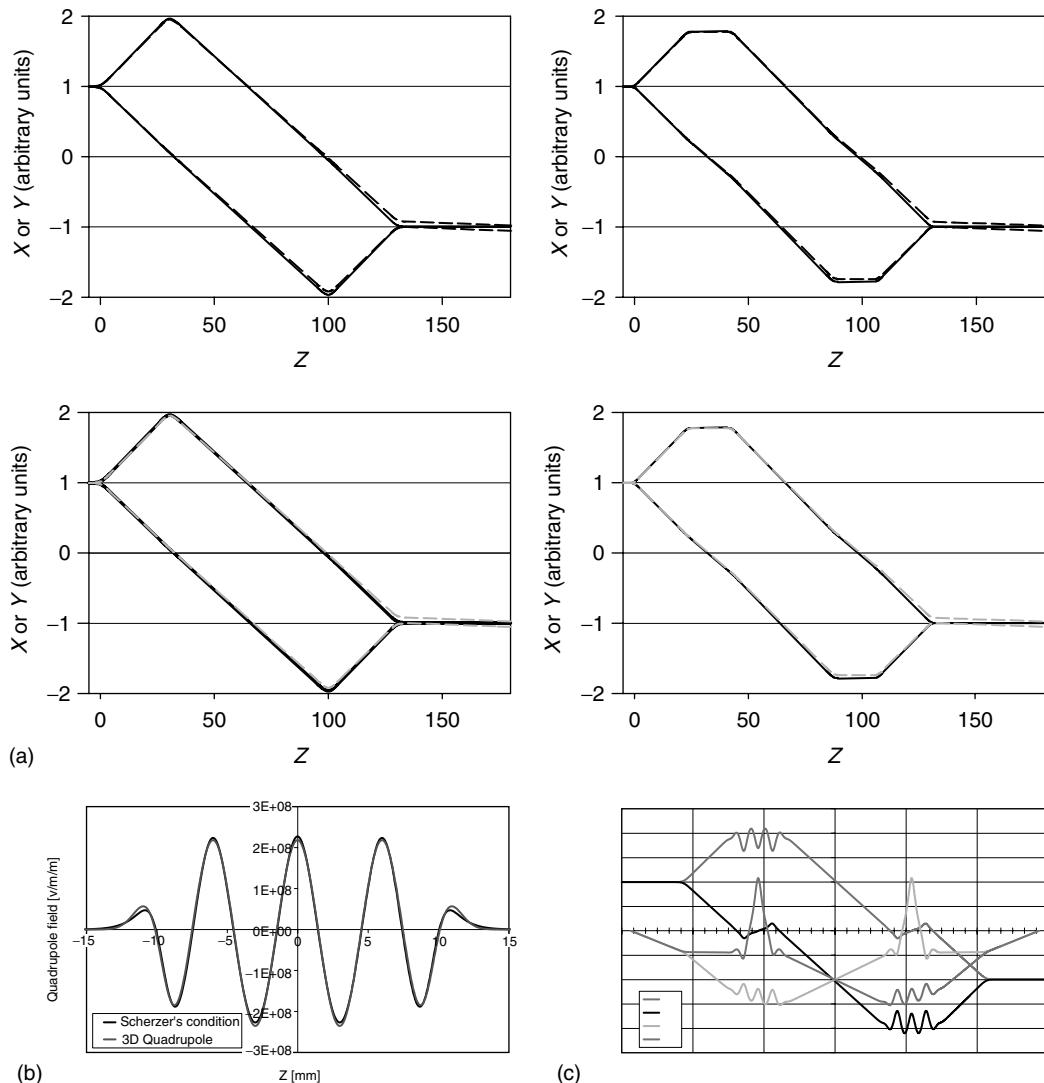
Quadrupole lenses consisting of four electrodes and four magnetic poles situated midway between the electrodes are capable of correcting the chromatic aberration of a round lens. They must of course be part of a suitable configuration and are currently incorporated in the complex superaplanator and ultracorrector described in Section 6.6.3.4. Such correctors used not to be seriously considered for correction of  $C_c$  in the TEM, for the simpler configurations increased the spherical (and other) aberrations unacceptably. However, they are being reconsidered (Haider and Müller, 2004), for there exist more elaborate arrangements that do not have this handicap (Haider et al., 2006b, 2007; Hartel et al., 2008).

#### 6.6.3.4 General Multipole Correctors

In 2000, Rose described a *hexapole planator* capable of correcting field curvature and (third-order) astigmatism. Such correction is essential for the needs of projection lithography systems. This proposal was explored in detail by Munro et al. (2001) who first examined the limitations of projection systems for charged particles at that date and then analyzed an electrostatic and a magnetic configuration free of all third-order geometrical aberrations, inspired by Rose's hexapole planator.

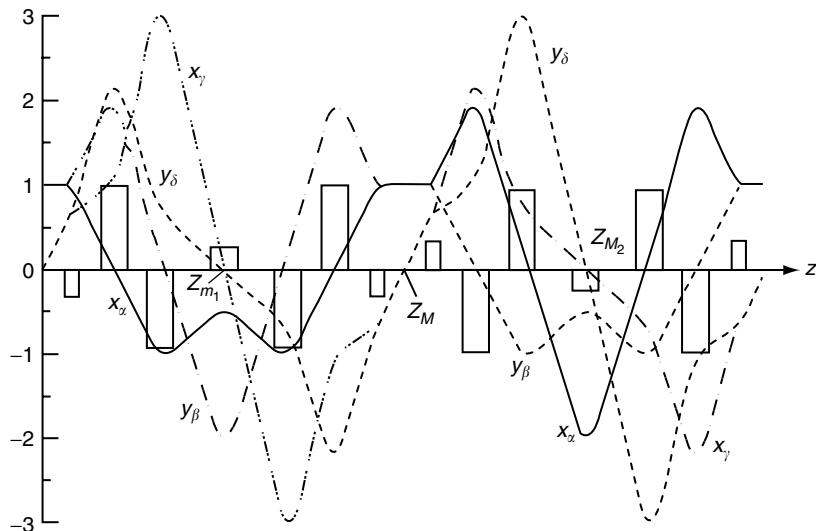


**FIGURE 6.11** Doubly symmetric electrostatic corrector (DECO) of spherical and chromatic aberrations. (a) Axial and field rays in one half of the DECO, each half of which consists of a corrector enclosed between quadrupole doublets; (b) detailed view of the *axial* rays in one unit of the DECO; and (c) detailed view of the *field* rays in the first half of the DECO. (After Weißbäcker, C. and Rose, H., *J. Electron Microsc.*, 51(1), 45–51, 2002. Courtesy of the authors and Oxford University Press.)



**FIGURE 6.12** (See color insert following page 340.) Scherzer's proposal for electrostatic correction of chromatic aberration. (a) Rays in a quadrupole quadruplet or sextuplet; (b) match between the potentials needed to satisfy condition 6.269 and those in the corrector; and (c) Gaussian rays in the sextuplet after correcting elements have been placed at the line foci. The chromatic aberration has been corrected and the residual chromatic aberration of magnification is small. (After Maas D.J. et al., *Proc. SPIE*, 4510, 205–217, 2001; Maas D. et al., *Microsc. Microanal.*, 9(Suppl. 3), 24–25, 2003. Courtesy of the authors, SPIE, and the Microscopy Society of America.)

Subsequently, two very complex correctors have been proposed by Rose, which are capable of correcting the spherical and chromatic aberrations as well as other primary aberrations that could be harmful once the axial aberrations have been brought under control (Rose, 2003, 2004, 2005, 2006, 2008a). These are known as the *superaplanator* and the *ultracorrector*. The first of these, which is suitable for a TEM, consists of two symmetrical quadrupole quintuplets and three (or more) octopoles and corrects spherical and chromatic aberrations. Very high stability of the excitations is essential (Rose, 2005), of the order of a few parts in  $10^7$ , but this is attainable today. Symmetry is exploited to make the device as easy to use as possible. Thus the quadrupole fields are symmetric



**FIGURE 6.13** The ultracorrector. (After Rose, H., *Ultramicroscopy*, 103(1), 1–6, 2005. Courtesy of the author and Elsevier.)

with respect to the center plane of each quintuplet; conversely, the whole (double-quintuplet) unit exhibits antisymmetry about its midplane (the plane midway between the two quintuplets). An octopole is placed in this midplane and at the center of each quintuplet (Figure 6.13).

The superaplanator is adequate for the TEM, where no particular attention needs be paid to third-order astigmatism and field curvature, the zone to be imaged being so small. In electron projection lithography, however, it is desirable to correct all the primary chromatic and geometrical aberrations. This can in principle be achieved with the configuration known as the ultracorrector. Here, two identical multipole multiplets are used. Each consists of seven quadrupoles and seven octopoles, themselves being disposed symmetrically about the center plane of the multiplet. Once again, the multiplet fields are antisymmetric with respect to the plane midway between two multiplets in which an additional octopole is situated. Detailed description of the modes of action of these complex structures is to be found in Rose's articles. A corrector of this type has been chosen for the transmission electron aberration-corrected microscope (TEAM) project (O'Keefe, 2003, 2004; O'Keefe et al., 2006; Kabius et al., 2004; Rose 2007c, Haider et al., 2008a,d).

The advent of correctors has obliged microscopists to re-examine the origins of contrast in the electron image, notably in the TEM. The need for some kind of phase plate to enhance image contrasts has been recognized, to replace the ‘virtual’ phase plate provided by the combination of spherical aberration and defocus. For some reflections on this question, see Matsumoto and Tonomura (1996), Lentzen (2004), Schultheiß et al. (2006), Rose (2007), Majorovits et al. (2007), Dries et al. (2008), Gamm et al. (2008a,b), Kaiser et al. (2008), Matijevic et al. (2008), Urban (2008) and Schröder et al. (2008).

### 6.6.3.5 Mirrors

Intermittent attempts have been made to exploit the fact that the spherical aberration of electron mirrors can change sign (Zworykin et al., 1945, section 17.2; Ramberg, 1949; Septier, 1966; Słowko, 1972, 1975; Kasper, 1968/9a; Rempfer and Mauck, 1985, 1986, 1992; Rempfer, 1990a,b; Shao and Wu, 1989, 1990a,b; Rempfer et al., 1997). The problem common to all these is that the incoming electrons, carrying information about the specimen, must somehow be separated from those emerging from the mirror, without introducing dispersion and other aberrations that would be at least as harmful as the aberration to be corrected. Among all the ingenious proposals that have been

made, we draw attention to three recent designs. In one of these designs (Rose and Preikszas, 1992; Preikszas and Rose, 1994, 1995, 1997; Wichtendahl et al., 1998; Preikszas et al., 2000), a mirror corrector is introduced into a low-energy electron microscope (which can, incidentally, operate in reflection or transmission). This bold design for the spectromicroscope for all relevant techniques (SMART) project uses a beam splitter with high symmetry; this is exploited, as explained earlier, to ensure that many of the aberrations of such a system of fact vanish. All four directions defined by the beam splitter are used. The source and condensers are above so that the electrons initially travel in a southerly direction; they emerge traveling west, where they encounter a field lens, the electrostatic objective, and the specimen. The backscattered electrons return to the beam splitter, emerge from the lower face, again traveling south, are reflected by the mirror, and return to the beam splitter, which now deflects them so that they finally emerge from the right face, traveling east. They then pass through the projector system and reach the recording plane (Figure 6.14). For further information about the SMART at BESSY II in Berlin, see the works of Fink et al. (1997), Müller et al. (1999), Hartel et al. (2000, 2002), Schmidt et al. (2002, 2007), Rose et al. (2004), and Bauer (2007). An analogous corrector is part of the PEEM-3 project at the Advanced Light Source in Berkeley (Wan et al., 2004, 2006; Wu et al., 2004; Feng et al., 2005; Schmid et al., 2005; Feng and Scholl, 2007).

The second project (Könenkamp et al., 2008) is a much improved version of the instrument developed by Rempfer et al. (1997). Here the optic axis is Y-shaped (Figure 6.15). The electron beam created by photoemission at the specimen is deflected twice, to bring it onto the axis of the mirror; here it is reflected and again reflected twice to bring it onto an axis parallel to the original axis. The symmetry is now favorable for aberration cancellation and simulations suggest that the spherical and chromatic aberrations can be controlled satisfactorily.

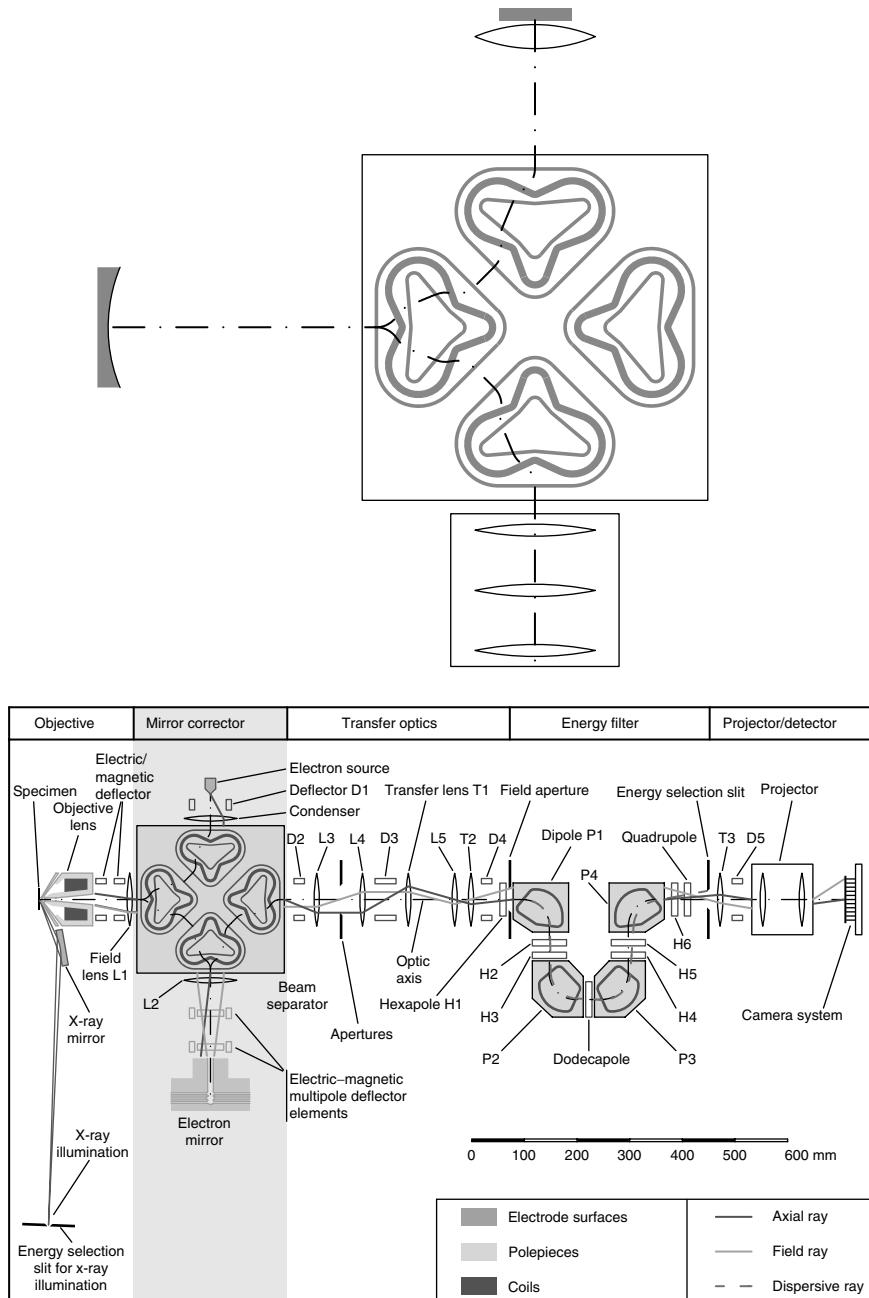
The most recent is a design described by Tromp (2008), whose aim is to find a simpler configuration than that of the SMART while at the same time retaining the ability to correct aberrations. Two prism assemblies are used here. The first directs the incident beam onto the specimen; the electrons emitted are redirected along the original optic axis by the second half of the first prism unit, passed through an electrostatic lens, and enter a second prism assembly, identical with the first. They are reflected at an electron mirror, just as in the original Castaing–Henry geometry, but the role of the mirror is now to provide aberration correction. The electrons return to the second prism unit and regain the original optic axis (Figure 6.16). In this arrangement, the second-order aberrations are canceled by symmetry, and alignment is relatively straightforward.

Another very original way of exploiting the ability of mirrors to cancel spherical and chromatic aberrations has been made by Crewe (1992, 1993a,b, 1995); in one practical design, the beam traverses the specimen twice, once in each direction. See also Crewe et al. (1995a,b, 2000) and Tsai (2000).

#### 6.6.3.6 Other Techniques

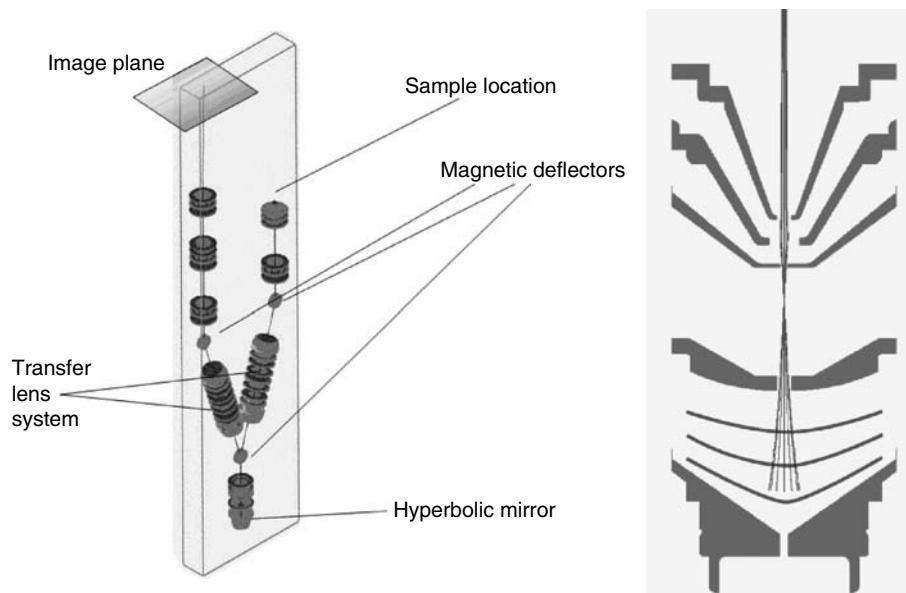
The introduction of a cloud of space charge into the path of the beam can in principle improve the aberrations, but in practice it has proved too difficult to control the cloud and prevent it from having a ‘clouded-glass’ effect. For descriptions of attempts to exploit this idea, see Gabor (1947), Ash and Gabor (1955), and Haufe (1958). The theory is discussed by Scherzer (1947) and the articles by Gabor (1945a,b) are relevant; see also Le Poole (1972). A thorough study of the technique was made by Typke (1968/9), which seemed to sound its death knell, but it is being revived for use with ion beams (Wang et al., 1995a–c; Tang et al. 1996a,b; Chao et al., 1997; Orloff, 1997).

The presence of a thin foil (or gauze), placed across the path of the beam in an electrostatic or combined electrostatic–magnetic lens, is much more promising and in one laboratory at least (Hanai et al., 1998) work continued on aberration correction in this way for many years. Scherzer remained attached to it (Scherzer, 1948, 1949, 1980). For other earlier work, we refer to Hawkes and Kasper (section 41.3.2). After completing his study of space charge, Typke (1972; erratum 1976) went on to study foil lenses and a practical design emerged. Curved foils were studied by

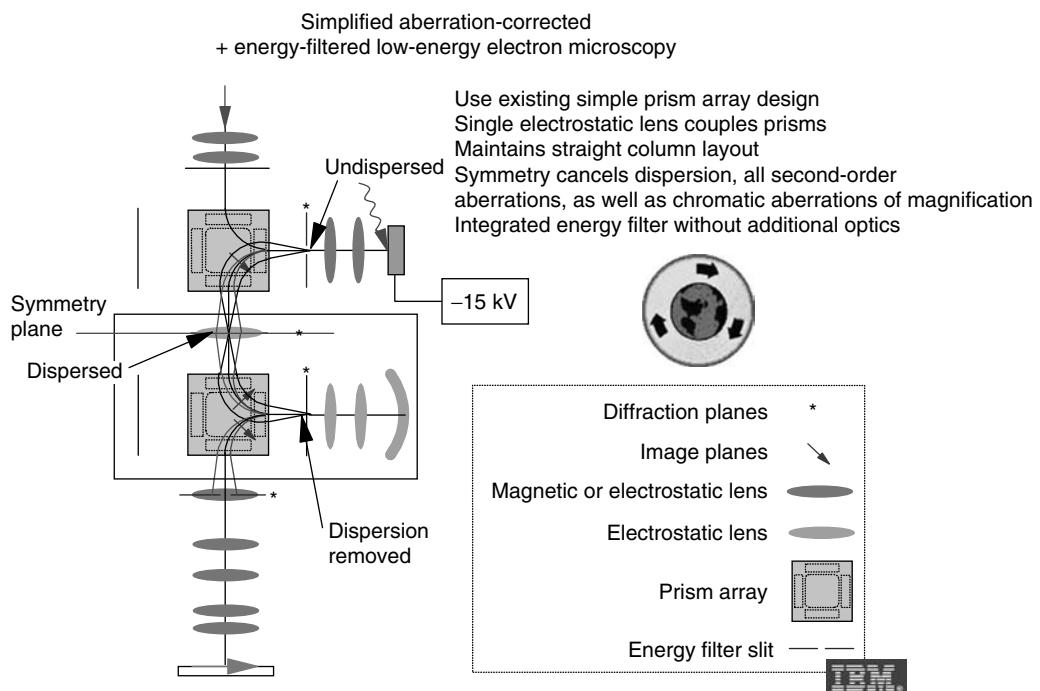


**FIGURE 6.14** The spectromicroscope for all relevant techniques (SMART) project incorporating a dispersion-free beam separator and a tetrode mirror. (a) Basic configuration of prisms and a tetrode mirror (After Preikszas, D. and Rose, H., *J. Electron Microsc.*, 46(1), 1–9, 1997. Courtesy of the authors and Oxford University Press.) and (b) the entire layout of the SMART. (After Hartel P. et al. *Adv. Imaging Electron Phys.*, 120, 41–133, 2002. Courtesy of the authors and Elsevier.)

Hoch et al. (1976). During the same period, work on aberration correction using this technique was also in progress in Japan and may be traced through the papers of Maruse et al. (1970a,b), Ichihashi and Maruse (1971, 1973), Hibino and Maruse (1976), Hibino et al. (1977, 1978, 1981), and Hanai et al. (1982, 1984, 1994, 1995, 1998). This has been revived for wide-angle correction by



**FIGURE 6.15** A Y-shaped photoemission electron microscopy design incorporating an aberration-correcting mirror. Left: general layout. Right: details of the hyperbolic mirror. (Courtesy of R. Könenkamp.)

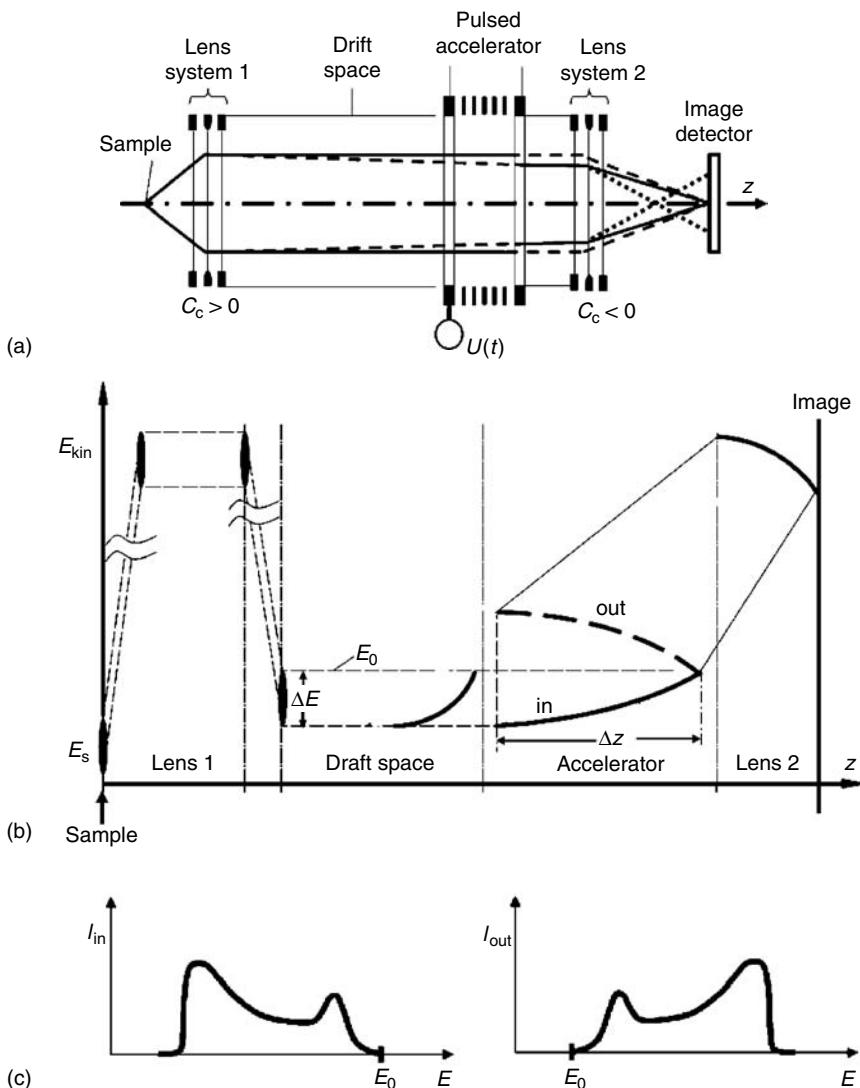


**FIGURE 6.16** (See color insert following page 340.) The disposition of the components of the IBM photoemission electron microscope showing the two prism units, the electron mirror, and the connecting electrostatic lens. The two sets of conjugate planes—image planes and diffraction planes—are identified by arrows and stars, respectively. (Courtesy of R. Tromp.)

Matsuda et al. (2005) following the ideas of Kato and Sekine (1995, 1996). Formulas for the aberration coefficients of round lenses containing foils are listed in Munro and Wittels (1977) and in van Aken et al. (2004), where expressions for the aberrations of gradient are also given.

A very different type of foil corrector has been investigated by van Aken (van Aken et al., 2002a,b, 2004; van Aken 2005), who attempted to exploit the marked increase in the mean free path of electrons in metal foils at very low voltages. If the latter is atomically flat, a mean free path of the order of 5 nm is expected for several metals at 5 eV above the Fermi level and a transmission of  $\sim 10\%$  is predicted for thin foils ( $\sim 5$  nm thick). By sandwiching the foil between a pair of electrodes, which first retard the incident electrons and then accelerate them again, correction of both chromatic and spherical aberrations is in principle possible. Another ingenious way of using the correcting property of a foil has been proposed by van Bruggen et al. (2006b) for use with a multibeam source (van Bruggen et al., 2006a).

Another way of combating spherical aberration that once seemed promising but has proved impractical involves the use of high-frequency excitation. The idea is simple and appealing: the electrons that travel far from the optic axis experience too strong a focusing force. By illuminating an einzel lens with a pulsed beam instead of a continuous beam and reducing the potential applied to the central electrode of the lens in such a way that the outer electrons, which have further to travel than the axial electrons and hence arrive a little later, encounter a weaker focusing field, it should be possible to cancel the effect of spherical aberration (Scherzer, 1946, 1947). The frequency required to reduce the central electrode potential rapidly enough is in the gigahertz range. The explanation given above then proves oversimplified since the electrons spend a substantial part of a cycle of the microwave field in the lens, but the principle remains valid. It is important to ensure that the *phase condition* (Nesslinger, 1939) is satisfied; this condition ensures that the lens is achromatic to a first approximation and hence reasonably insensitive to the microwave phase when the electron pulse arrives. Gabor (1950) too was intrigued by the possibilities of high-frequency operation. The fullest theoretical studies have been made by Oldfield (1971, 1973, 1974) and Matsuda and Ura (1974); experimental tests have been performed and further proposals made by Vaidya (Vaidya and Hawkes, 1970; Vaidya, 1972, 1975a,b; Garg and Vaidya, 1974; Vaidya and Garg, 1974; Pandey and Vaidya, 1975, 1977, 1978). The fullest experimental study is that of Oldfield, who built a microwave circuit and measured the optical properties of some cavities. This project was not pursued, however, and the problems of generating short enough pulses and, above all, dealing with the increased energy spread of the beam were clearly formidable. Recently, Calvo has reconsidered the technique along very different lines (Calvo, 2002, 2004; Calvo and Lazcano, 2002; Calvo and Laroze, 2002) but no clear conclusion can be drawn. Yet another way of using dynamic fields for spherical and chromatic aberration correction has been proposed by Schönhense and Spiecker (2002a,b, 2003, Schönhense et al., 2006), with photoemission and low-energy electron microscopes in mind. For chromatic aberration, an ingenious way of inverting the energy distribution of the electron beam is proposed. The electrons are generated at the sample by a pulsed beam and, after acceleration and collimation, they enter a drift space in which the faster electrons draw away from the slower ones, like athletes sprinting toward the finishing line in a race. Beyond the drift space is an accelerator, initially switched off. When the fastest electrons have emerged from it, the electric accelerating field is rapidly switched on, and by suitable choice of the accelerator field, the slower electrons can be accelerated to a higher energy than the fast electrons, thus effectively inverting the original energy distribution (Figure 6.17). The chromatic aberration of the final lens will then bring the whole beam to a smaller focus than in the absence of energy inversion. For spherical aberration correction, Schönhense and Spiecker consider an *electrostatic* lens in which the lens strength is abruptly altered when the (pulsed) beam reaches the center. The correcting effect can be understood by regarding the half-lens as a diverging lens (the converging effect of the second half of the lens having been suppressed by switching off the lens). The image of the object is then virtual and Scherzer's result is no longer applicable. Another way of understanding this is to note that the abrupt change in lens strength is equivalent to a discontinuity in the field and, as for the foil correctors mentioned



**FIGURE 6.17** Chromatic correction based on inversion of the electron energy distribution. (a) Schematic cross section of the system (greatly exaggerated in the radial direction). (b) Electron energy distribution as a function of the optical path;  $E_{kin}$  denotes the actual kinetic energy of the electron ensemble. The distributions  $E_{kin}$  as a function of  $z$  before and after passing the pulsed accelerating field are denoted “in” and “out,” respectively. (c) Schematic representation of the electron energy distribution before and after passing the accelerator,  $I_{in}$  and  $I_{out}$ . (After Schönhense et al. 2002a. Courtesy of the authors and the American Institute of Physics.)

earlier, this is known to offer a way of achieving correction. Related designs have been described by Khursheed (2002, 2005).

One final aspect of aberration correction must be mentioned here: holography. In all the foregoing approaches, we attempt to correct the aberration at or close to the objective lens. Holography was invented by Gabor (1948, 1949, 1951) in an attempt to eliminate the undesirable effect of spherical aberration, not by correcting the objective, but by introducing a second correction step at the image. The first tests failed because bright, coherent sources of light or electrons were not yet available, and many years passed before correction was achieved by (off-axis) electron holography. We refer to Tonomura (1998, 1999) and Lichte (1991, 1997, 2008) for extended accounts of the

subject and to those of Ade (1994), Ishizuka et al. (1994), Tonomura (1994, 2006), Tonomura et al. (1995), Kahl and Rose (1995), Matteucci et al. (1997), Lehmann (2000, 2004, 2005), Lehmann and Lichte (2002, 2005), Lichte and Lehmann (2002, 2008), Geiger et al. (2004, 2005, 2008), Lichte (2002, 2003, 2005, 2006), Lehmann et al. (2005) and Dunin-Borkowski et al. (2007) for subsequent developments.

## FOR FURTHER INFORMATION

All the textbooks on particle optics devote considerable space to aberrations. The work of Glaser is conveniently and magisterially presented in his textbook (1952) and more readably in his *Handbuch der Physik* article (1956). Other noteworthy contributions have been made by Sturrock (1951a, 1952, 1955), Hawkes (1965b), Rose (1968), and Rose and Petri (1971). The book of Wollnik (1987) is useful for sector fields, although it is necessary to consult the original papers by Wollnik, often in collaboration with Matsuo and Matsuda, for details; the study by Rose and Krahl (1995), cited in the main text, is also indispensable here as is a later chapter by Rose (2003a). *Science of Microscopy* (Hawkes and Spence, 2007) contains a chapter on aberration correctors by Hawkes (2007) and the chapters on high-resolution TEM by Kirkland and Hutchison (2007) and on STEM by Nellist (2007) also touch on the subject. A forthcoming book by Rose will be directly relevant and an entire volume of *Advances in Imaging and Electron Physics* will be devoted to aberration-corrected microscopy (Hawkes, 2008).

New developments are regularly reported in the *Proceedings of the International Conferences on Charged Particle Optics*: Giessen (Wollnik, 1981); Albuquerque (Schriber and Taylor, 1987); Toulouse (Hawkes, 1990); Tsukuba (Ura et al., 1995); Delft (Kruit and van Amersfoort, 1999); Greenbelt, Maryland (Dragt and Orloff, 2004); and Cambridge (Munro and Rouse, 2008); the next conference is to be held in Singapore, in 2010 (Khursheed). These are almost all published in *Nuclear Instruments and Methods in Physical Research*, Part A. The International and European Congresses on Electron Microscopy also reserve oral and poster sessions for electron optics in which aberration studies usually play a large part (full publishing details of the earlier congresses are given in Hawkes and Kasper, Vol. 3, 1893ff, brought closer to the present time in Hawkes, 2003). The main SPIE conference occasionally has sessions on charged particle optics (*Proc. SPIE*, 2014, 1993; 2522, 1995; 2858, 1996; 3155, 1997; 3777, 1999; 4510, 2001); the English translation of the Russian seminars on Problems in Theoretical and Applied Electron and Ion Optics, published in Russian in *Prikladnaya Fizika*, also appear in *Proc. SPIE* (see vols 4187, 5025, 5398, and 6278). At the biennial seminars on Recent Trends in Charged Particle Optics and Surface Physics Instrumentation, held in Skalský Dvůr near Brno, new work is often presented before it reaches the regular journals. Much theoretical work has been published in *Optik*, but it is now also to be found in other journals of microscopy or scientific instruments, notably *Ultramicroscopy*, *Journal of Microscopy*, and *Journal of Electron Microscopy*. Until recently, many papers on particle optics came from the former Soviet Union, notably from Leningrad/St Petersburg and from Alma Ata in Kazakhstan and were frequently published in *Zhurnal Tekhnicheskoi Fiziki* (*Soviet Physics: Technical Physics*, now *Technical Physics*) or *Radiotekhnika i Elektronika* (available in English translation under several different names). Section Zh of *Referativny Zhurnal* usefully complements the more familiar INSPEC abstracting services.

## ACKNOWLEDGMENTS

I am particularly grateful to those electron opticians who have been kind enough to send me lists of aberration coefficients as files that I could incorporate directly, thereby avoiding much tedious typing and above all, the risk of introducing errors. Their work is of course referred to in the text, but I also thank Katsushige Tsuno, Seitkerim and Raushan Bimurzaev, Zhixiong Liu and Xieqing Zhu here for being so very cooperative.

## REFERENCES\*

- Ade, G. 1994. Digital techniques in electron off axis holography. *Adv. Electron. Electron Phys.*, 89, 1–51.
- Aken, R.H. van 2005. Low-energy electron beams through ultra-thin foils, applications for electron microscopy. *Proefschrift*, Delft.
- Aken, R.H. van, Hagen, C.W., Barth, J.E. and Kruit, P. 2002a. Low-energy foil aberration corrector. *Ultramicroscopy*, 93(3–4), 321–330.
- Aken, R.H. van, Hagen, C.W., Barth, J.E. and Kruit, P. 2002b. Concept for a low energy foil corrector. In *Proc. 15th Int. Cong. on Electron Microscopy*, Durban, Cross, R., Engelbrecht, J., Witcomb, M. and Richards, P. Eds., Vol. 1, 1101–1102.
- Aken, R.H. van, Lenc, M. and Barth, J.E. 2004. Aberration integrals for the low-voltage foil corrector. *Nucl. Instrum. Meth. Phys. Res., A*, 519(1–2), 205–215.
- Aken, P.A. van, Koch, C.T., Sigle, W., Höschken, R., Rühle, M., Essers, E., Benner, G. and Matijevic, M. 2007. The sub-electron-volt-sub-angstrom-microscope (SESAM): pushing the limits in monochromated and energy-filtered TEM. *Microsc. Microanal.*, 13, Suppl. 2, 862–863 (CD).
- Amboss, K. and Jennings, J.C.E. 1970. Electron-optical investigation of the origin of asymmetries in air-cored solenoids. *J. Appl. Phys.*, 41(4), 1608–1616.
- Andersen, W.H.J. 1967. Optimum adjustment and correction of the Wien filter. *Brit. J. Appl. Phys.*, 18(11), 1573–1579.
- Andersen, W.H.J. and Le Poole, J.B. 1970. A double wienfilter as a high resolution, high transmission electron energy analyser. *J. Phys. E: Sci. Instrum.*, 3(2), 121–126.
- Archard, G.D. 1953. Magnetic electron lens aberrations due to mechanical defects, *J. Sci. Instrum.*, 30(10), 352–358.
- Archard, G.D. 1955. A possible chromatic correction system for electron lenses. *Proc. Phys. Soc. (London)*, B, 68(11), 817–829.
- Archard, G.D. 1960. Fifth order spherical aberration of magnetic lenses. *Brit. J. Appl. Phys.*, 11(11), 521–522.
- Artsimovich, L.A. 1944. [Electron optical properties of emission systems]. *Izv. Akad. Nauk SSSR (Ser. Fiz.)*, 6, 313–329.
- Ash, E.A. and Gabor, D. 1955. Experimental investigations on electron interaction. *Proc. R. Soc. London, A*, 228(1175), 477–490.
- Bacon, N.J., Corbin, G.J., Dellby, N., Hrmčík, P., Krivanek, O.L., McManama-Smith, A., Murfitt, M.F. and Szilagyi, Z.S. 2005. Nion UltraSTEM: an aberration-corrected STEM for imaging and analysis. *Microsc. Microanal.*, 11, Suppl. 2, 1422–1423 (CD).
- Baranova, L.A. and Yavor, S.Ya. 1986. *Elektrostaticeskie Elektronnye Linzy*, (Nauka, Moscow).
- Baranova, L.A. and Yavor, S.Ya. 1989. The optics of round and multipole electrostatic lenses. *Adv. Electron. Electron Phys.*, 76, 1–207.
- Baranova, L.A., Read, F.H. and Cubric, D. 2002. Computer simulation of an electrostatic aberration corrector for a low-voltage scanning electron microscope. In *Proc. 8th Seminar on Recent Trends in Charged Particle Optics and Surface Physics Instrumentation*, Skalský Dvůr, Frank, L. Ed., 74 (ISI, Brno).
- Baranova, L.A., Read, F.H. and Cubric, D. 2004. Computational simulation of an electrostatic aberration corrector for a low-voltage scanning electron microscope. *Nucl. Instrum. Meth. Phys. Res., A*, 519(1–2), 42–47.
- Baranova, L.A., Read, F.H. and Cubric, D. 2008. Computer simulations of hexapole aberration correctors. *Phys. Procedia*.
- Barth, J.E., van der Steen, H.W.G. and Chmelik, J. 1995. Improvements to the electrostatic lens optimization method SOEM. *Proc. SPIE*, 2522, 128–137.
- Bärtele, J. and Plies, E. 2005. Energy width measurements of an electrostatic  $\Omega$ -monochromator using thermal and field emitters. *Proc. Microscopy Conf. 2005*, 1 (Paul-Scherrer-Institut Proc. 05–01).
- Bärtele, J. and Plies, E. 2006. Energy width reduction using an electrostatic  $\Omega$ -monochromator. In *Proc. 16th Int. Microscopy Cong.*, Sapporo, Ichinose, H. and Sasaki, T. Eds., (IMC-16 Publishing Committee, Sapporo).

\* Despite the length of this list, it is by no means complete and the review articles cited should also be consulted by anyone interested in the history of aberrations and their correction. There is some redundancy in references to recent work, which may be reported in several congresses in the same year, but I felt that a complete list would be found useful when the time comes to reconsider the present very active decades.

- Bärtle, J. and Plies, E. 2008. Energy width reduction using an electrostatic  $\Omega$ -monochromator. *Phys. Procedia*.
- Batson, P.E. 2003. Aberration correction results in the IBM STEM instrument. *Ultramicroscopy*, 96(3–4), 239–249.
- Batson, P.E., Mook, H.W. and Kruit, P. 2000. High brightness monochromator for STEM. In *Microbeam Analysis 2000*, Williams, D.B. and Shimizu, R. Eds., 213–214 (Institute of Physics, Bristol and Philadelphia).
- Batson, P.E., Dellby, N. and Krivanek, O.L. 2002. Sub-ångstrom resolution using aberration corrected optics. *Nature*, 418(6898), 617–620.
- Bauer, E. 2007. LEEM and SPLEEM. In *Science of Microscopy*, Hawkes, P.W. and Spence, J.C.H. Eds., 605–656 (Springer, New York).
- Beck, V.D. 1979. A hexapole spherical aberration corrector. *Optik*, 53(4), 241–255.
- Benner, G., Essers, E., Huber, B. and Orchowski, A. 2003a. Design and first results of SESAM. *Microsc. Microanal.*, 9, Suppl. 3, 66–67.
- Benner, G., Matijevich, M., Orchowski, A., Schindler, B., Haider, M. and Hartel, P. 2003b. State of the first aberration-corrected, monochromatized 200 kV FEG-TEM. *Microsc. Microanal.*, 9, Suppl. 3, 38–39.
- Benner, G., Orchowski, A., Haider, M. and Hartel, P. 2003c. State of the first aberration-corrected, monochromatized 200 kV FEG-TEM. *Microsc. Microanal.*, 9, Suppl. 2, 938–939 (CD).
- Benner, G., Essers, E., Matijevic, M., Orchowski, A., Schlossmacher, P. and Thesen, A. 2004a. Performance of monochromized and aberration-corrected TEMs. *Microsc. Microanal.*, 10, Suppl. 2, 108–109.
- Benner, G., Matijevic, M., Orchowski, A., Schlossmacher, P., Thesen, A., Haider, M. and Hartel, P. 2004b. Sub-Ångstrom and sub-eV resolution with the analytical SATEM. In *Materials Research in an Aberration-free Environment*, *Microsc. Microanal.*, 10, Suppl. 3, 6–7 website.
- Berdnikov, A.S. and Stam, M.A.J. van der 1995. Algorithms for numerical calculation of aberration coefficients with an arbitrary optic axis. *Nucl. Instrum. Meth. Phys. Res.*, A, 363(1–2), 295–300.
- Bertein, F. 1947–8. Quelques défauts des instruments d'optique électronique et leur correction. *Ann. Radioélectr.*, 2(10), 379–408 and 3(1), 49–62.
- Bertein, F. 1948. Influence des défauts de forme d'une électrode simple. *J. Phys. Radium*, 9(3), 104–112.
- Bertein, F., Bruck, H. and Grivet, P. 1947. Influence of mechanical defect of the objectives on the resolving power of the electrostatic microscope, *Ann. Radioélectr.*, 2(9), 249–252.
- Berz, M. 1987. The method of power series tracking for the mathematical description of beam dynamics. *Nucl. Instrum. Meth. Phys. Res.*, A, 258(3), 431–436.
- Berz, M. 1989. Differential algebraic description of beam dynamics to very high orders. *Particle Acc.*, 24(2), 109–124.
- Berz, M. 1990. Arbitrary order description of arbitrary particle optical systems. *Nucl. Instrum. Meth. Phys. Res.*, A, 298(1–3), 436–440.
- Berz, M. 1995. Modern map methods for charged particle optics. *Nucl. Instrum. Meth. Phys. Res.*, A, 363(1–2), 100–104.
- Berz, M. 1999. Modern Map Methods in Particle Beam Physics. *Adv. Imaging & Electron Phys.*, Vol. 108.
- Berz, M. and Makino, K. 2004. New approaches for the validation of transfer maps using remainder-enhanced differential algebra. *Nucl. Instrum. Meth. Phys. Res.*, A, 519 (1–2), 53–62.
- Bimurzaev, S. and Yakushev, E.M. 2004. Methods of parametrization of exact electron trajectory equations. *Proc. SPIE*, 5398, 51–58.
- Bimurzaev, S.B., Daumenov, T., Sekunova, L.M. and Yakushev, E.M. 1983a. [Space-time focusing in a two-electrode electrostatic mirror.] *Zh. Tekh. Fiz.*, 53(3), 524–528; *Sov. Phys. Tech. Phys.*, 28(3), 326–329.
- Bimurzaev, S.B., Daumenov, T., Sekunova, L.M. and Yakushev, E.M. 1983b. [Space-time characteristics of a three-electrode electrostatic mirror with rotational symmetry.] *Zh. Tekh. Fiz.*, 53(6), 1151–1156; *Sov. Phys. Tech. Phys.*, 28(6), 696–699.
- Bimurzaev, S.B., Bimurzaeva, R.S. and Yakushev, E.M. 1999. Calculation of time-of-flight chromatic aberrations in the electron-optical systems possessing straight optical axes. *Nucl. Instrum. Meth. Phys. Res.*, A, 427(1–2), 271–274.
- Bimurzaev, S.B., Serikbaeva, G.S. and Yakushev, E.M. 2003. Electrostatic mirror objective with eliminated spherical and axial chromatic aberrations. *J. Electron Microsc.*, 52(4), 365–368.
- Bimurzaev, S.B., Serikbaeva, G.S. and Yakushev, E.M. 2004. Calculation of the focussing quality of the electrostatic mirror objective free of the third-order spherical aberration. *Nucl. Instrum. Meth. Phys. Res.*, A, 519(1–2), 70–75.
- Bleloch, A., Falke, U., Goodhew, P. and Weng, P. 2005. Surprises from aberration-corrected STEM. In *Proc. 7th Multinat. Cong. Electron Microscopy*, Portorož, Čeh, M., Dražič, G. and Fidler, S. Eds., 47–50.

- Bleloch, A.L., Brown, M.D., Brydson, R., Craven, A., Goodhew, P. and Tatlock, G. 2007. Performance of a fifth-order aberration corrected STEM. In *Proc. of EMAG 2007*, Brown, P.D., Moebus, G. and Baker, R.T. Eds., *J. Phys.: Conf Series*.
- Blom, D.A., Allard, L.F., Mishina, S. and O'Keefe, M.A. 2006. Early results from an aberration-corrected JEOL 2200FS STEM/TEM at Oak Ridge National Laboratory. *Microsc. Microanal.*, 12(6), 483–491.
- Brouwer, W. 1957. *The Use of Matrix Algebra in Geometrical Optics*, Dissertation, Delft.
- Brouwer, W. and Walther, A. 1967. Geometrical optics and design of optical instruments. In *Advanced Optical Techniques*, Heel, A.C.S. van Ed., 503–570 and 571–631 (North Holland, Amsterdam).
- Brown, K.L. 1968. A first- and second-order matrix theory for the design of beam transport systems and charged particle spectrometers. In *Advances in Particle Physics*, Vol. 1, Cool, R.L. and Marshak, R.E. Eds., 71–134 (Wiley-Interscience, New York). Earlier distributed as SLAC Report No. 75, 1967.
- Brown, K.L. and Servranckx, R.V. 1985. First- and second-order charged particle optics. In *Physics of High Energy Particle Accelerators*, Month, M., Dahl, P.F. and Dienes, M. Eds., 62–138 (American Institute of Physics, New York; AIP Proc. 127).
- Brown, K.L. and Servranckx, R.V. 1987. Optics modules for circular accelerator design. *Nucl. Instrum. Meth. Phys. Res.*, A, 258(3), 480–502.
- Browning, N.D., Arslan, I., Erni, R., Idrobo, J.C., Ziegler, A., Bradley, J., Dai, Z., Stach, E.A. and Bleloch, A. 2006. Monochromators and aberration correctors: taking EELS to new levels of energy and spatial resolution. *J. Phys.: Conf. Ser.*, 26, 59–64.
- Bruggen, M.J. van, Someren, B. van, Hagen, C.W. and Kruit, P. 2006a. Development of a multi-beam scanning electron microscope for sub-10 nm electron beam induced deposition. In *Proc. 16th Int. Microscopy Cong.*, Sapporo, Ichinose, H. and Sasaki, T. Eds., Vol. 2, 1130 (IMC-16 Publishing Committee, Sapporo).
- Bruggen, M.J. van, Someren, B. van and Kruit, P. 2006b. Use of a foil corrector in a multibeam source: correction of off-axis macrolens aberrations. In *Proc. 10th Seminar on Recent Trends in Charged Particle Optics and Surface Physics Instrumentation*, Skalský Dvůr, Müllerová, I. Ed., 69–70 (ISI, Brno).
- Burfoot, J.C. 1953. Correction of electrostatic lenses by departure from rotational symmetry. *Proc. Phys. Soc. (London)*, B, 66(9), 775–792.
- Calvo, M. 2002. Transfer theory and image formation in a time dependent aberrationless magnetic field lens. *Optik*, 113(6), 233–240.
- Calvo, M. 2004. Optimal resolution of a time-dependent aberrationless magnetic lens. *Ultramicroscopy*, 99(2–3), 179–187.
- Calvo, M. and Laroze, D. 2002. Aberrations theory of a time dependent magnetic field lens. *Optik*, 113(10), 429–434.
- Calvo, M. and Lazcano, P. 2002. Theory of a time dependent magnetic lens. *Optik*, 113(1), 31–34.
- Carey, D.C. 1987. *The Optics of Charged Particle Beams* (Harwood, Chur, Switzerland).
- Chand, G., Saxton, W.O. and Kirkland, A.I. 1995. Aberration measurement and automated alignment of the TEM. In *Electron Microscopy and Analysis 1995*, Cherns, D. Ed., 297–300 (Institute of Physics Conf. Ser. 147).
- Chang, L.Y., Chen, F.R., Kirkland, A.I. and Kai, J.J. 2003. Calculations of spherical aberration-corrected imaging behaviour. *J. Electron Microsc.*, 52(4), 359–364.
- Chang, L.Y., Kirkland, A.I. and Titchmarsh, J.M. 2006. On the importance of fifth-order spherical aberration for a fully corrected electron microscope. *Ultramicroscopy*, 106(4–5), 301–306.
- Chao, L.C., Orloff, J. and Wang, L. 1997. Spherical aberration corrector using space charge. *J. Vac. Sci. Technol.*, B, 15(6), 2732–2736.
- Chen, E. and Mu, C. 1991. New development in correction of spherical aberration of electro-magnetic round lens. In *Proc. Int. Symp. Electron Microscopy*, Kuo, K. and Yao, J. Eds., 28–35 (World Scientific, Singapore).
- Chen, W.-x., Xu, J., Chen, L. and Zhang, H.-z. 2007. Correction of spherical aberration of a medium voltage TEM. *J. Chinese Electron Microsc. Soc.*, 26(1), 1–7.
- Cheng, M., Tang, T.-t. and Yao, Z.-h. 2001a. Study on differential algebraic aberration method for electrostatic electron lenses. *Optik*, 112(6), 250–254.
- Cheng, M., Tang, T.-t. and Yao, Z.-h. 2001b. Study on differential algebraic chromatic aberration method for Glaser's bell-shaped magnetic lenses. *Optik*, 112(10), 483–486.
- Cheng, M., Tang, T.-t. and Yao, Z.-h. 2002a. High-order combined geometric–chromatic aberrations of magnetic electron lenses using differential algebra. *J. Phys. D: Appl. Phys.*, 35(4), 409–411.
- Cheng, M., Tang, T.-t. and Yao, Z.-h. 2002b. Study on time-of-flight property of electron optical systems by differential algebraic method. *J. Phys. D: Appl. Phys.*, 35(13), 1566–1599.
- Cheng, M., Tang, T.-t., Lu, Y.-l. and Yao, Z.-h. 2003. Differential algebraic method for arbitrary order curvilinear-axis geometric–chromatic aberration analysis. *J. Phys. D: Appl. Phys.*, 36(7), 923–928.

- Cheng, M., Zhou, S.-f., Lu, Y.-l. and Tang, T.-t. 2006. Differential algebraic method for arbitrary-order chromatic aberration analysis of electrostatic lenses. *Optik*, 117(4), 196–198.
- Chu, H.C. and Munro, E. 1982. Numerical analysis of electron beam lithography systems. Part III: Calculation of the optical properties of electron focusing systems and dual-channel deflection systems with combined magnetic and electrostatic fields. Part IV: Computerized optimization of the electron optical performance of electron beam lithography systems using the damped least squares method. *Optik*, 61(2), 121–145 and (3), 213–236.
- Cleaver, J.R.A. 1987. In-lens deflection for scanning ion beam systems. *Optik*, 75(2), 75–81.
- Crewe, A.V. 1980. Studies on sextupole correctors. *Optik*, 57(3), 313–327.
- Crewe, A.V. 1982. A system for the correction of axial aperture aberrations in electron lenses. *Optik*, 60(3), 271–281.
- Crewe, A.V. 1984. The sextupole corrector. I. Algebraic calculations. *Optik*, 69(1), 24–29.
- Crewe, A.V. 1991a. An algebraic approach to the symmetrical einzel lens. *Optik*, 88(3), 118–125.
- Crewe, A.V. 1991b. A new characterization of the magnetic lens. *Optik*, 89(2), 70–74.
- Crewe, A.V. 1991c. The three element electrostatic lens. *Optik*, 90(4), 151–157.
- Crewe, A.V. 1992. Electron motion in tuned fields. I. The algebra. II. Some applications. *Ultramicroscopy*, 41(4), 269–277 and 279–285.
- Crewe, A.V. 1993a. On the aberrations of uniform fields. *Optik*, 96(4), 180–184.
- Crewe, A.V. 1993b. The aberration problem in electron optics. *Proc. SPIE*, 2014, 77–83.
- Crewe, A.V. 1995. Limits of electron probe formation. *J. Microscopy*, 178(2), 93–100.
- Crewe, A.V. and Kopf, D. 1980a. A sextupole system for the correction of spherical aberration. *Optik*, 55(1), 1–10.
- Crewe, A.V. and Kopf, D. 1980b. Limitations of sextupole correctors. *Optik*, 56(4), 391–399.
- Crewe, A.V., Ruan, S., Tsai, F. and Korda, P. 1995a. The first test on a magnetically focused mirror corrector. In *Electron Microscopy and Analysis 1995*, Cherns, D. Ed., 301–304 (Institute of Physics Conf. Ser. 147).
- Crewe, A.V., Tsai, F., Korda, F. and Ruan, S. 1995b. The first test on a magnetically focused mirror corrector. In *Microscopy and Analysis 1995*, Bailey, G.W., Ellisman, M.H., Hennigar, R.A. and Zaluzec, N.J. Eds., 562–563 (Jones and Begell, New York).
- Crewe, A.V., Ruan, S., Korda, P. and Tsai, F.C. 2000. Studies of a magnetically focused electrostatic mirror. I. Experimental test of the first order properties. *J. Microscopy*, 197(2), 110–117 [See Tsai, 2000 for Part II].
- Dauben, J.W. 1998. *Abraham Robinson: The Creation of Nonstandard Analysis, a Mathematical Odyssey* (Princeton University Press, Princeton).
- Dauben, J.W. 2003. Abraham Robinson 1918–1974, A biographical memoir. *Biographical Memoirs National Academy of Science*, Vol. 82 (National Academy Press, Washington D.C.).
- Daumenov, T., Sapargaliev, A.A. and Yakushev, E.M. 1978. [New method of determining the point characteristic function for charged particles in an electron optical system with a straight optic axis.] *Zh. Tekh. Fiz.*, 48(12), 2447–2454; *Sov. Phys. Tech. Phys.*, 23(12), 1400–1404.
- Daumenov, T., Sapargaliev, A.A., Sekunova, L.M. and Yakushev, E.M. 1981. [Space-time focusing of charged particles in fields with rotational symmetry]. *Zh. Tekh. Fiz.*, 51(6), 1137–1145; *Sov. Phys. Tech. Phys.*, 26(6), 647–651.
- Degenhardt, R. 1997. Optimization of electrostatic lens systems: fast approximate calculation of imaging properties of einzel lenses. *Proc. SPIE*, 3155, 70–77.
- Dellby, N., Krivanek, O.L. and Lupini, A.R. 2000. Progress in aberration-corrected STEM. *Microsc. Microanal.*, 6, Suppl. 2, 100–101.
- Dellby, N., Krivanek, O.L., Nellist, P.D., Batson, P.E. and Lupini, A.R. 2001. Progress in aberration-corrected scanning transmission electron microscopy. *J. Electron Microsc.*, 50(3), 177–185.
- Dellby, N., Krivanek, O.L., Murfitt, M.F. and Nellist, P.D. 2005. Design and testing of a quadrupole/octupole  $C_3/C_5$  aberration corrector. *Microsc. Microanal.*, 11, Suppl. 2, 2130–2131(CD).
- Dellby, N., Krivanek, O.L. and Murfitt, M.F. 2008. Optimized quadrupole–octupole  $C_3/C_5$  aberration corrector for STEM. *Phys. Procedia*.
- Delong, A. and Štěpán, P. 2006. Low voltage electron microscope for biological objects. In *Proc. 10th Seminar on Recent Trends in Charged Particle Optics and Surface Physics Instrumentation*, Skalský Dvůr, Müllerová, I. Ed., 11–12 (ISI, Brno).
- Deltrap, J.M.H. 1964. Correction of spherical aberration with combined quadrupole–octopole units. In *Proc. 3rd Eur. Regional Conf. on Electron Microscopy*, Prague, Titlbach, M. Ed., Vol. A, 45–46 (Czechoslovak Academy of Sciences, Prague).
- Der-Shvarts, G.V. 1954. [Influence of imperfections of the rotational symmetry of the focusing fields on the resolution of the magnetic objectives of electron microscopes.] *Zh. Tekh. Fiz.*, 24, 859–870.

- Der-Shvarts, G.V. and Makarova, I.S. 1972. [Third-order field aberrations of axially symmetric magnetic lenses]; *Izv Akad. Nauk SSSR (Ser. Fiz.)*, 36(6), 1304–1311; *Bull. Acad. Sci. USSR (Phys. Ser.)*, 36(6), 1164–1170.
- Der-Shvarts, G.V. and Makarova, I.S. 1973. [Simple formulae for an approximate calculation of the third-order aberration coefficients of thin, axially symmetric single-gap magnetic lenses.] *Radiotekh. Elektron.*, 18, 2374–2378; *Radio Eng. Electron. Phys.*, 18, 1722–1725.
- Dowker, I.C., Read, F.H., Bowring, N.J. and Hammond, P. 1995. Aberrations of an electrostatic three-cylinder system used to simultaneously focus and deflect charged particles. *Nucl. Instrum. Meth. Phys. Res.*, A, 363 (1–2), 54–58.
- Dragt, A.J. 1987. Elementary and advanced Lie algebraic methods with applications to accelerator design, electron microscopes, and light optics. *Nucl. Instrum. Methods Phys. Res.*, A, 258(3), 339–354.
- Dragt, A.J. 1990. Numerical third-order transfer map for solenoid, *Nucl. Instrum. Methods Phys. Res.*, A, 298(1–3), 441–9.
- Dragt, A.J. and Forest, E. 1986. Lie algebraic theory of charged-particle optics and electron microscopes. *Adv. Electron. Electron Phys.*, 67, 65–120.
- Dragt, A. and Orloff, J. Eds., 2004. Proc. of the 6th Int. Conf. on Charged Particle Optics (CPO-6), College Park (MD), *Nucl. Instrum. Meth. Phys. Res.*, A, 519, 1–487.
- Dragt, A.J., Forest, E. and Wolf, K.B. 1986. Foundations of a Lie algebraic theory of geometrical optics. In *Lie Methods in Optics*, Mondragón, J.S. and Wolf, K.B. Eds., 105–157 (Springer, Berlin) *Lecture Notes in Physics*, Vol. 250.
- Dries, M., Schultheiß, K., Gamm, B., Störmer, H., Gerthsen, D., Barton, B. and Schröder, R.R. 2008. Application of a Hilbert phase plate in transmission electron microscopy of materials science samples. *Proc. European Microscopy Congress*, Aachen, Luysberg, M., Tillmann, K. and Weirich, T. Eds., Instrumentation and Methods volume, Springer, Berlin.
- Dunin-Borkowski, R.E., Kasama, T., McCartney, M.R. and Smith, D.J. 2007. Electron holography. In *Science of Microscopy*, Hawkes, P.W. and Spence, J.C.H. Eds., 1141–1195 (Springer, New York).
- Dušek, H. 1959. Über die elektronenoptische Abbildung durch Orthogonalsysteme unter Verwendung eines neuen Modellfeldes mit streng angebbaren Paraxialbahnen. *Optik*, 16(7), 419–445.
- Dvořák, D. 2002. Misalignment aberrations. In *Proc. 8th Seminar on Recent Trends in Charged Particle Optics and Surface Physics Instrumentation*, Skalský Dvůr, Frank, L. Ed., 17–18 (ISI, Brno).
- Dymnikov, A.D. 1995. The methods of shuttle-sums and of shuttle-integrals in ion optics. *Nucl. Instrum. Meth. Phys. Res.*, A, 363(1–2), 435–439.
- Dymnikov, A.D. 1999. Matrix methods in periodic focusing systems. *Nucl. Instrum. Meth. Phys. Res.*, A, 427(1–2), 6–11.
- Dymnikov, A.D. and Hellborg, R. 1993. Matrix theory of the motion of a charged particle in curvilinear space-time. I. General theory. II. Nonlinear theory. Formulae and algorithms. *Nucl. Instrum.*, A, 330(3), 323–342 and 343–362.
- Dymnikov, A.D. and Martínez, G. 1997. The matriza method for an optimal synthesis of nonlinear magnetic and electrostatic focusing systems for a given beam current. *Proc. SPIE*, 3155, 125–133.
- Dymnikov, A.D. and Yavor, S.Ya. 1963. [Four quadrupole lenses as an analogue of an axially symmetric system.] *Zh. Tekh. Fiz.*, 33(7), 851–858; *Sov. Phys. Tech. Phys.*, 8(7), 639–643.
- Dymnikov, A.D., Fishkova, T.Ya. and Yavor, S.Ya. 1964. System of revolution composed of four quadrupole lenses. In *Proc. 3rd Eur. Reg. Conf. Electron Microscopy*, Prague, A, 43–44.
- Dymnikov, A.D., Fishkova, T.Ya. and Yavor, S.Ya. 1965. [Effect of geometrical parameters on optical characteristics of a system of four quadrupole lenses similar to an axisymmetric lens.] *Zh. Tekh. Fiz.*, 35(3), 431–440; *Sov. Phys. Tech. Phys.*, 10(3), 340–346.
- Dymnikov, A.D., Martínez, G. and Azbaid, A.H. 1998. The matriza method for an optimal synthesis of electrostatic multiple cylinder lenses. *Nucl. Instrum. Meth. Phys. Res.*, A, 403(2–3), 195–204.
- Edgcombe, C.J. 1991. On the perturbation of trajectories by small displacement of electrodes. *Optik*, 86(4), 120–128.
- Edgcombe, C.J., Lupini, A.R. and Taylor, J.H. 1999. Robust optimization for magnetic lens design. *Nucl. Instrum. Meth. Phys. Res.*, A, 427(1–2), 306–309.
- El-Kareh, A.B. and El-Kareh, J.C.J. 1970. *Electron Beams, Lenses and Optics* (Academic Press, New York, 2 vols.).
- Essers, E., Matijevic, M., Benner, G., Höschens, R., Sigle, W. and Koch, C. 2007. Analytical performance of the SESAM microscope. *Microsc. Microanal.*, 13, Suppl. 3, 18–19.
- Essers, E., Mittmann, D., Mandler, T. and Benner, G. 2008. The MANDOLINE filter and its performance. *Proc. European Microscopy Congress*, Aachen, Luysberg, M., Tillmann, K. and Weirich, T. Eds., Instrumentation and Methods volume (Springer, Berlin).

- Feng, J. and Scholl, A. 2007. Photoemission electron microscopy (PEEM). In *Science of Microscopy*, Hawkes, P.W. and Spence, J.C.H. Eds., 657–695 (Springer, New York).
- Feng, J., Forest, E., Macdowell, A.A., Marcus, M., Padmore, H., Raoux, S., Robin, D., Scholl, A., Schlueter, R., Schmid, P., Stöhr, J., Wan, W., Wei, D.H. and Wu, Y. 2005. An x-ray photoemission electron microscope using an electron mirror aberration corrector for the study of complex materials. *J. Phys.: Cond. Matter*, 17, S1339–S1350.
- Fink, R., Weiss, M.R., Umbach, E., Prekszas, D., Rose, H., Spehr, R., Hartel, P., Engel, W., Degenhardt, R., Wichtendahl, R., Kuhlenbeck, H., Erlebach, W., Ihmann, K., Schlögl, R., Freund, H.-J., Bradshaw, A.M., Lilienkamp, G., Schmidt, Th., Bauer, E. and Benner, G. 1997. SMART: a planned ultrahigh-resolution spectromicroscope for BESSY II. *J. Electron Spectrosc. Relat. Phenom.*, 84(1–3), 231–250.
- Forest, E. 1998. *Beam Dynamics, a New Attitude and Framework* (Harwood, Chur, Switzerland).
- Foschepoth, M. and Kohl, H. 1998. Amplitude contrast—a way to obtain directly interpretable high-resolution images in a spherical-aberration-corrected transmission electron microscope. *Phys. Stat. Sol (a)*, 166(1), 357–366.
- Franzen, N. and Munro, E. 1987. Interactive computer programs for analyzing rotationally symmetric optics, together with misalignments, using an integral equation method. In *Proc. Int. Symp. Electron Optics*, Beijing 1986, Ximen, J.-y. Ed., 57–60 (Institute of Electronics, Academia Sinica, Beijing).
- Freitag, B., Kujawa, S., Mul, P.M. and Tiemeijer, P.C. 2004a. First experimental proof of spatial resolution improvement in a monochromized and  $C_s$ -corrected TEM. In *Materials Research in an Aberration-free Environment, Microsc. Microanal.*, 10, Suppl. 3, 4–5 website.
- Freitag, B., Kujawa, S., Mul, P.M., Tiemeijer, P.C. and Snoeck, E. 2004b. First experimental proof of spatial resolution improvement in a monochromized and  $C_s$ -corrected TEM. In *Proc. 8th Asia-Pacific Conf. on Electron Microscopy*, Kanazawa, Tanaka, N., Takano, Y., Mori, H., Seguchi, H., Iseki, S., Shimada, H. and Simamura, E. Eds., 18–19 CD (8APEM Publication Committee, Uchinada, Ishikawa).
- Freitag, B., Kujawa, S., Mul, P.M., Ringnalda, J. and Tiemeijer, P.C. 2005. Breaking the spherical and chromatic aberration barrier in transmission electron microscopy. *Ultramicroscopy*, 102(3), 209–214.
- Freitag, B., Stekelenburg, M., Ringnalda, J. and Hubert, D. 2007. Atomic resolution  $C_s$ -corrected HR-S/TEM from 80–300 kV: optimizing for technique and material. *Microsc. Microanal.*, 13, Suppl. 2, 1162–1163 (CD).
- Frosien, J. and Plies, E. 1987. High performance electron optical column for testing ICs with submicrometer design rules. *Microelectron. Eng.*, 7(2–4), 163–172.
- Frosien, J., Plies, E., and Anger, K. 1989. Compound magnetic and electrostatic lenses for low-voltage applications. *J. Vac. Sci. Technol. B*, 7(6), 1874–1877.
- Gabor, D. 1942–1943. Electron optics. *Electron. Eng.*, 15, 295–299, 328–331 and 337, 372–374.
- Gabor, D. 1945a. Stationary electron swarms in electromagnetic fields. *Proc. R. Soc. London, A*, 183(995), 436–453.
- Gabor, D. 1945b. Dynamics of electron beams. Applications of Hamiltonian dynamics to electronic problems. *Proc. Inst. Radio Eng.*, 33(11), 792–805.
- Gabor, D. 1945c. *The Electron Microscope* (Hulton Press, London, 1948).
- Gabor, D. 1947. A space-charge lens for the focusing of ion beams. *Nature*, 160, 89–90.
- Gabor, D. 1948. A new microscope principle. *Nature*, 161, 777–778.
- Gabor, D. 1949. Microscopy by reconstructed wave-fronts. *Proc. R. Soc. London, A*, 197(1051), 454–487.
- Gabor, D. 1950. Electron optics at high frequencies and at relativistic velocities. *Rev. Opt.*, 29(4), 209–231.
- Gabor, D. 1951. Microscopy by reconstructed wave fronts: II. *Proc. Phys. Soc. (London), B*, 64(6), 449–469.
- Gamm, B., Schultheiß, K., Gerthsen, D. and Schröder, R.R. 2008a. Optimal imaging parameters in  $C_s$ -corrected transmission electron microscopy with a physical phase plate. *Proc. European Microscopy Congress*, Aachen, Luysberg, M., Tillmann, K. and Weirich, T. Eds., Instrumentation and Methods volume (Springer, Berlin).
- Gamm, B., Schultheiß, K., Gerthsen, D. and Schröder, R.R. 2008b. Effect of a physical phase plate on contrast transfer in an aberration-corrected transmission electron microscope. *Ultramicroscopy*, 108(9), 878–884.
- Gao, J. and Mu, C.-j. 1999. Computer-aided analysis of seventh-order geometrical aberrations of a magnetic lens. *J. Yunnan Univ.* 21, 32–34.
- Garg, R.K. and Vaidya, N.C. 1974. Computer design and electron optics of synklysmotron lenses. In *Proc. 8th Int. Congr. Electron Microscopy*, Canberra, Vol. I, 154–155.
- Geiger, D., Lichte, H., Lehmann, M., Haider, M. and Freitag, B. 2004. Electron holography with a  $C_s$ -corrected TEM. In *Proc. 13th Eur. Microscopy Cong.* Antwerp, Dyck, D. van and Oostveldt, P. van Eds., Vol. I, 187–188 (Belgian Society for Microscopy, Liège).

- Geiger, D., Lichte, H., Lehmann, M., Freitag, B. and Haider, M. 2005. Electron holography with a  $C_s$ -corrected TEM at Triebenberg Laboratory. *Proc. Microscopy Conf. 2005*, Davos, 32 (Paul-Scherrer-Institut Proc. 05-01).
- Geiger, D., Lichte, H., Linck, M. and Lehmann, M. 2008. Electron holography with a  $C_s$ -corrected transmission electron microscope. *Microsc. Microanal.* 14(1), 68–81.
- Glaser, W. 1940. Über ein von sphärischer Aberration freies Magnetfeld. Über den Öffnungsfehler der Elektronenlinsen, Bemerkungen zu vorstehender Erwiderung. *Z. Phys.*, 116(1–2), 19–33 and (11–12) 734–735.
- Glaser, W. 1941. Strenge Berechnung magnetischer Linsen der Feldform  $\frac{H_0}{1 + (z/a)^2}$ . *Z. Phys.*, 117(5–6), 285–315.
- Glaser, W. 1942. Über elektronenoptische Abbildung bei gestörter Rotationssymmetrie. *Z. Phys.*, 120, 1–15.
- Glaser, W. 1949. Über die Theorie der elektrischen und magnetischen Ablenkung von Elektronenstrahlbündeln und ein ihr angepaßtes Störungsverfahren. *Ann. Physik*, 4(7), 389–408.
- Glaser, W. 1952. *Grundlagen der Elektronenoptik* (Springer, Vienna).
- Glaser, W. 1956. Elektronen- und Ionenoptik. *Handbuch der Physik*, 33, 123–395.
- Glaser, W. and Schiske, P. 1953. Bildstörungen durch Polschuhasymmetrien bei Elektronenlinsen. *Z. Angew. Phys.*, 5(9), 329–339.
- Glatzel, U. and Lenz, F. 1988. Optimization of electrostatic immersion lenses. *Optik*, 79(1), 15–18.
- Greenfield, D. and Monastyrskiy, M.A. 2009. Selected problems of computational charged-particle optics. *Adv. Imaging & Electron Phys.*, Vol. 155.
- Greenfield, D., Monastyrskiy, M.A. and Tarasov, V.A. 2006. Perturbation techniques in the problems of computational charged particle optics. *Proc. SPIE*, 6278, 627804: 1–15; *Prikladnaya Fizika* No., 3, 64–75.
- Haantjes, J. and Lubben, G.J. 1957. Errors of magnetic deflection, I. *Philips Res. Rep.*, 12, 46–68.
- Haantjes, J. and Lubben, G.J. 1959. Errors of magnetic deflection, II. *Philips Res. Rep.*, 14, 65–97.
- Haider, M. 1998. Correctors for electron microscopes: tools or toys for scientists? In *Proc. EUREM-11, the 11th Eur. Conf. on Electron Microscopy*, Dublin 1996 CESEM, Ed., Vol. I, 363–364 (CESEM, Brussels).
- Haider, M. 2000. Towards sub-Ångstrom point resolution by correction of spherical aberration. In *Proc. 12th Eur. Cong. on Electron Microscopy* Brno, Tománek, P. and Kolařík, R. Eds., Vol. III, I.145–I.148 (Czechoslovak Society for Electron Microscopy, Brno).
- Haider, M. 2003. Current and future developments in order to approach a point resolution of 0.5 Å. *Microsc. Microanal.*, 9, Suppl. 2, 930–931 (CD).
- Haider, M. and Müller, M. 2004. Design of an electron optical system for the correction of the chromatic aberration  $C_c$  of a TEM objective lens. In *Materials Research in an Aberration-free Environment, Microsc. Microanal.*, 10, Suppl. 3, 2–3 website.
- Haider, M. and Müller, M. 2005. Is there a road map of aberration correction towards ultra-high resolution in TEM and STEM? *Microsc. Microanal.*, 11, Suppl. 2, 546–547 (CD).
- Haider, M. and Uhlemann, S. 1997. Seeing is not believing: reduction of artefacts by an improved point resolution with a spherical aberration corrected 200 kV transmission electron microscope. *Microsc. Microanal.*, 3, Suppl. 2, 1179–1180.
- Haider, M., Bernhardt, W. and Rose, H. 1982. Design and test of an electric and magnetic dodecapole lens. *Optik*, 63(1), 9–23.
- Haider, M., Braunshausen, G. and Schwan, E. 1994. State of the development of a  $C_s$  corrected high resolution 200 kV TEM. In *Proc. 13th Int. Conf. Electron Microscopy*, Paris, Vol. 1, 195–196 (Editions de Physique, Les Ulis).
- Haider, M., Braunshausen, G. and Schwan, E. 1995. Correction of the spherical aberration of a 200 kV TEM by means of a hexapole corrector. *Optik*, 99(4), 167–179.
- Haider, M., Rose, H., Uhlemann, S., Kabius, B. and Urban, K. 1998a. Towards 0.1 nm resolution with the first spherically corrected transmission electron microscope. *J. Electron Microsc.*, 47(5), 395–405.
- Haider, M., Rose, H., Uhlemann, S., Schwan, E., Kabius, B. and Urban, K. 1998b. A spherical-aberration-corrected 200 kV transmission electron microscope. *Ultramicroscopy*, 75(1), 53–60.
- Haider, M., Uhlemann, S., Schwan, E., Rose, H., Kabius, B. and Urban, K. 1998c. Electron microscopy image enhanced. *Nature*, 392(6678), 768–769.
- Haider, M., Uhlemann, S. and Zach, J. 2000. Upper limits for the residual aberrations of a high-resolution aberration-corrected STEM. *Ultramicroscopy*, 81(3–4), 163–175.
- Haider, M., Müller, M. and Hartel, P. 2004. Present state and future trends of aberration correction. In *Proc. 8th Asia-Pacific Conf. on Electron Microscopy*, Kanazawa, Tanaka, N., Takano, Y., Mori, H., Seguchi, H., Iseki, S., Shimada, H. and Simamura, E. Eds., 16–17 CD (8APEM Publication Committee, Uchinada, Ishikawa).

- Haider, M., Müller, H., Hartel, P. and Uhlemann, S. 2006a.  $C_s$ -correction for 0.5 Å resolution of a monochromated STEM. Abstracts, Pre-meeting Congress on Materials Research in an Aberration-free Environment, 2 pp.
- Haider, M., Müller, H. and Uhlemann, S. 2006b. Improvement path for the hexapole  $C_s$ -corrector toward 0.5 Å resolution. *Microsc. Microanal.*, 12, Suppl. 2, 1468–1469 (CD).
- Haider, M., Müller, H. Uhlemann, S. and Zach, J. 2006c. Correction of spherical and chromatic aberration for TEM. Abstracts, Pre-meeting Congress on Materials Research in an Aberration-free Environment, 2 pp.
- Haider, M., Müller, M., Hartel, P. and Uhlemann, S. 2006d. Advancement of hexapole  $C_s$ -correctors for high resolution CTEM and STEM. In *Proc. 16th Int. Microscopy Cong.*, Sapporo, Ichinose, H. and Sasaki, T. Eds., Vol. 2, 614 (IMC-16 Publishing Committee, Sapporo).
- Haider, M., Loebau, U., Höschchen, R., Müller, H., Uhlemann, S. and Zach, J. 2007. State of the development of a  $C_c$  and  $C_s$  corrector for TEAM. *Microsc. Microanal.*, 13, Suppl. 2, 1156–1157 (CD).
- Haider, M., Müller, H., Uhlemann, S., Zach, J., Loebau, U. and Höschchen, R. 2008a. Prerequisites for a  $C_c/C_s$ -corrected ultrahigh-resolution TEM. *Ultramicroscopy* 108(3), 167–178.
- Haider, M., Müller, H. and Uhlemann, S. 2008b. Present and future hexapole aberration correctors for high resolution electron microscopy. *Adv. Imaging & Electron Phys.* 153.
- Haider, M., Müller, H., Uhlemann, S., Hartel, P. and Zach, J. 2008c. Developments of aberration correction systems for current and future requirements. *Proc. European Microscopy Congress*, Aachen, Luysberg, M., Tillmann, K. and Weirich, T. Eds., Instrumentation and Methods volume (Springer, Berlin).
- Haider, M., Hartel, P., Loebau, U., Höschchen, R., Müller, H., Uhlemann, S., Kahl, F. and Zach, J. 2008d. Progress on the development of a  $C_c/C_s$  corrector for TEAM. *Microsc. Microanal.*, 14, Suppl. 2, 800–801 (CD).
- Haine, M.E. and Mulvey, T. 1954. The application and limitations of the edge-diffraction test for astigmatism in the electron microscope. *J. Sci. Instrum.*, 31(September), 326–332.
- Hanai, T., Hibino, M. and Maruse, S. 1982. Improvement of an electron probe profile using the foil lens. *J. Electron Microsc.*, 31(4), 360–367.
- Hanai, T., Hibino, M. and Maruse, S. 1984. Reduction of the probe diameter through the correction of spherical aberration by the foil lens. *J. Electron Microsc.*, 33(4), 329–334.
- Hanai, T., Hibino, M. and Maruse, S. 1994. Effect of the three-fold astigmatism on the performance of the foil lens used for correction of the spherical aberration of a probe-forming lens. *Optik*, 97(2), 86–90.
- Hanai, T., Arai, S. and Hibino, M. 1995. Increase of current density of the electron probe by correction of the spherical aberration with a side-entry type foil lens. *J. Electron Microsc.*, 44(5), 301–306.
- Hanai, T., Yoshida, H. and Hibino, M. 1998. Characteristics and effectiveness of a foil lens for correction of spherical aberration in scanning transmission electron microscopy. *J. Electron Microsc.* 47(3), 185–192.
- Hartel, P., Preikszas, D., Spehr, R. and Rose, H. 2000. Performance of the mirror corrector for an ultrahigh-resolution spectromicroscope. In *Proc. 12th Eur. Congr. on Electron Microscopy* Brno, Tománek, P. and Kolařík, R. Eds., Vol. III, I.153–I.154 (Czechoslovak Society for Electron Microscopy, Brno).
- Hartel, P., Preikszas, D., Spehr, R., Müller, H. and Rose, H. 2002. Mirror corrector for low-voltage electron microscopes. *Adv. Imaging & Electron Phys.*, 120, 41–133.
- Hartel, P., Müller, H., Uhlemann, S. and Haider, M. 2004. Residual aberrations of hexapole-type  $C_s$  correctors. In *Proc. 13th Eur. Microscopy Cong.*, Antwerp, Dyck, D. van and Oostveldt, P. van Eds., Vol. I, 41–42 (Belgian Society for Microscopy, Liège).
- Hartel, P., Müller, H., Uhlemann, S. and Haider, M. 2007a. Experimental set-up of an advanced hexapole  $C_s$ -corrector. *Microsc. Microanal.*, 13, Suppl. 2, 1148–1149 (CD).
- Hartel, P., Müller, H., Uhlemann, S. and Haider, M. 2007b. First experimental results with advanced hexapole  $C_s$ -correctors. *Microsc. Microanal.*, 13, Suppl. 3, 4–5.
- Hartel, P., Müller, H., Zach, J., Löbau, U., Höschchen, R. and Haider, M. 2008. Demonstration of  $C_c/C_s$ -corection in HRTEM. *Proc. European Microscopy Congress*, Aachen, Luysberg, M., Tillmann, K. and Weirich, T. Eds., Instrumentation and Methods volume (Springer, Berlin).
- Hartmann, B. and Wollnik, H. 1995. The particle motion through the fringing field of a toroidal condenser in a third-order approximation. *Nucl. Instrum. Meth. Phys. Res.*, A, 363 (1–2), 458–462.
- Haufe, G. 1958. Korrektur des Öffnungsfehlers bei elektrostatischen Linsen durch Raumladung. *Optik*, 15(9), 521–537.
- Hawkes, P.W. 1964. The paraxial chromatic aberrations of quadrupole systems. In *Proc. 3rd Eur. Reg. Conf. Electron Microscopy*, Prague, A, 5–6.
- Hawkes, P.W. 1965a. The paraxial chromatic aberrations of rectilinear orthogonal systems. *Optik*, 22(8), 543–551.

- Hawkes, P.W. 1965b. The geometrical aberrations of general electron optical systems. I. The conditions imposed by symmetry. II. The primary (third order) aberrations of orthogonal systems, and the secondary (fifth order) aberrations of round systems. *Philos. Trans. R. Soc. London, A*, 257(1086), 479–552.
- Hawkes, P.W. 1966. *Quadrupole Optics* (Springer, Berlin).
- Hawkes, P.W. 1970a. Asymptotic aberration integrals for round lenses. *Optik*, 31(3), 213–219.
- Hawkes, P.W. 1970b. Asymptotic aberration integrals for quadrupole systems. *Optik*, 31(4), 302–314.
- Hawkes, P.W. 1970c. The addition of round electron lens aberrations. *Optik*, 31(6), 592–599.
- Hawkes, P.W. 1970d. The addition of quadrupole aberrations. *Optik*, 32(1), 50–60.
- Hawkes, P.W. 1970e. Quadrupoles in electron lens design. *Adv. Electron. Electron Phys.*, Suppl 7.
- Hawkes, P.W. 1980a. Methods of computing optical properties and combating aberrations for low-intensity beams. In *Applied Charged Particle Optics*, Septier, A. Ed., *Adv. Electron. Electron Phys.*, Suppl. 13A, 45–157.
- Hawkes, P.W. 1980b. Some approximate magnetic lens aberration formulae: a commentary. *Optik*, 56(3), 293–320.
- Hawkes, P.W. 1982. Magnetic lens theory. In *Magnetic Electron Lenses*, Hawkes, P.W. Ed., 1–56 (Springer, Berlin, *Topics in Current Physics*, Vol. 18).
- Hawkes, P.W. 1987. Comment on “Geometric aberrations in electrostatic lenses: II”. *J. Phys. E: Sci. Instrum.*, 20(2), 234–235.
- Hawkes, P.W. 1989. The structure of the asymptotic aberration coefficients of combined focusing and deflection systems. Part I. General theory and aberrations of position. Part II. Aberrations of gradient and overall structure. *Optik*, 83(3), 104–112 and (4), 122–128.
- Hawkes, P.W., Ed. 1990. Proceedings of the Third International Conference on Charged Particle Optics (CPO-3), Toulouse. *Nucl. Instrum. Meth. Phys. Res.*, A, 298, 1–508.
- Hawkes, P.W. 1991. Addition formulae for the coefficients of the asymptotic aberration polynomials of combined focusing and deflection systems. *Optik*, 87(2), 55–62.
- Hawkes, P.W. 1999. Rebuttal of “The relationship of lens aberrations to image aberrations in electron optics” by G.F. Rempfer, *Optik*, 110(8), 372–374.
- Hawkes, P.W. 2000. Clarification of my criticism of “The relationship of lens aberrations to image aberrations in electron optics” by G.F. Rempfer, *Optik*, 111(5), 227–228.
- Hawkes, P.W. 2003. Electron optics and electron microscopy: conference proceedings and abstracts as source material. *Adv. Imaging & Electron Phys.*, 127, 205–375.
- Hawkes, P.W. 2004. The addition of chromatic aberration coefficients. *Ultramicroscopy*, 99(2–3), 197–200.
- Hawkes, P.W. 2007. Aberration correctors. In *Science of Microscopy*, Hawkes, P.W. and Spence, J.C.H. Eds., 696–747 (Springer, New York).
- Hawkes, P.W. Ed. 2008. Aberration-corrected Electron Microscopy. *Adv. Imaging & Electron Phys.*, Vol. 153.
- Hawkes, P.W. and Kasper, E. 1989. *Principles of Electron Optics*, Vols. 1 and 2, (Academic Press, London).
- Hawkes, P.W. and Lencová, B. 2002. Studies on the asymptotic cardinal elements and aberration coefficients of symmetric electrostatic lenses. *Optik*, 113(2) 78–82.
- Hawkes, P.W. and Spence, J.C.H. Eds., 2007. *Science of Microscopy* (Springer, New York).
- Hawkes, P.W., Lencová, B. and Lenc, M. 1995. Approximate expressions for electron lens properties. *J. Microscopy*, 179(2), 145–157.
- Healy, L.M. and Dragt, A.J. 1989. Concatenation of Lie algebraic maps. In *Lie Methods in Optics II*, Wolf, K.B. Ed., 67–95 (Springer, Berlin; *Lecture Notes in Physics*, Vol. 352).
- Heel, A.C.S. van. 1949. Calcul des aberrations du cinquième ordre et projets tenant compte de ces aberrations. In *La Théorie des Images Optiques*, Fleury, P. Ed., 32–67 (Edns Revue d’Optique Théorique et Instrumentale, Paris).
- Henstra, A. and Krijn, M.P.C.M. 2000. An electrostatic achromat. In *Proc. 12th Eur. Cong. on Electron Microscopy*, Brno, Tománek, P. and Kolařík, R. Eds., Vol. III, I.155–I.156 (Czechoslovak Society for Electron Microscopy, Brno).
- Hibino, M. and Maruse, S. 1976. Correction of the spherical aberration of a magnetic lens with a foil lens. *J. Electron Microsc.*, 25(4), 229–236.
- Hibino, M., Sugiyama, S., Hanai, T. and Maruse, S. 1977. Correction of the spherical aberration of a magnetic lens using a foil lens. In *Proc. 5th Int. Conf. High-Voltage Electron Microscopy*, Kyoto, 49–52.
- Hibino, M., Sugiyama, S., Hanai, T. and Maruse, S. 1978. Characteristics of the foil lens for the correction of the spherical aberration of the strongly excited magnetic lens. *J. Electron Microsc.*, 27(4), 259–265.

- Hibino, M., Kuzuya, M. and Maruse, S. 1981. Correction of the spherical aberration of the objective lens in a conventional transmission electron microscope by means of a foil lens. *Jpn. J. Appl. Phys.*, 20(4), 781–782.
- Hill, R. and Smith, K.C.A. 1980. An interactive computer system for magnetic lens analysis. In *Proc. 7th Eur. Cong. Electron Microscopy*, The Hague, 1, 60–61.
- Hill, R. and Smith, K.C.A. 1982. Applications of the Cambridge interactive electron lens analysis system: CIELAS. In *EMAG-81*, Goringe, M.J. Ed., 71–74 (Institute of Physics Conf. Series 61).
- Hillier, J. and Ramberg, E.G. 1947. The magnetic electron microscope objective: contour phenomena and the attainment of high resolving power. *J. Appl. Phys.*, 18(1), 48–71.
- Hoch, H., Kasper, E. and Kern, D. 1976. Der Öffnungsfehler 3. Ordnung und der axiale Farbfehler von rotationssymmetrischen Elektronenlinsen mit gekrümmter geladener transparenter Folie. *Optik*, 46(4), 463–473.
- Hoffstätter, G.H. 1999. Comments on aberration correction in symmetric imaging energy filters. *Nucl. Instrum. Meth. Phys. Res.*, A, 427(1–2), 275–281.
- Honda, K., Uno, S., Nakamura, N., Matsuya, M., Achard, B. and Zach, J. 2004a. An automatic geometrical aberration correction system of scanning electron microscopes. In *Proc. 13th Eur. Microscopy Cong.*, Antwerp, Dyck, D. van and Oostveldt, P. van Eds., Vol. I, 43–44 (Belgian Society for Microscopy, Liège).
- Honda, K., Uno, S., Nakamura, N., Matsuya, M. and Zach, J. 2004b. An automatic geometrical aberration correction system of scanning electron microscopes. In *Proc. 8th Asia-Pacific Conf. on Electron Microscopy*, Kanazawa, Tanaka, N., Takano, Y., Mori, H., Seguchi, H., Iseki, S., Shimada, H. and Simamura, E. Eds., 44–45 (CD) (8APEM Publication Committee, Uchinada, Ishikawa).
- Hordon, L.S., Huang, Z., Maluf, N., Browning, R. and Pease, R.F.W. 1993. Limits of low-energy electron optics. *J. Vac. Sci. Technol. B*, 11(6), 2299–2303.
- Hosokawa, F. and Sawada, H. 2003. A  $C_s$ -corrected TEM. *Denshi Kenbikyo* 38, Suppl. 2, 55–58.
- Hosokawa, F., Tomita, T., Naruse, M., Honda, T., Hartel, P. and Haider, M. 2003. A spherical aberration-corrected 200 kV TEM. *J. Electron Microsc.*, 52(1), 3–10.
- Hosokawa, F., Sawada, H., Sannomiya, T., Kaneyama, T., Kondo, Y., Hori, M., Yuasa, S., Kawazoe, M., Nakamichi, T., Tanishiro, Y., Yamamoto, N. and Takayanagi, K. 2006. Design and development of  $C_s$  corrector for a 300 kV TEM and STEM. In *Proc. 16th Int. Microscopy Cong.*, Sapporo, Ichinose, H. and Sasaki, T. Eds., Vol. 2, 582 (IMC-16 Publishing Committee, Sapporo).
- Huber, A. and Plies, E. 1999. Numerical ray tracing of electrons in different 3D fringing fields of spherical deflectors. *Proc. SPIE*, 3777, 264–274.
- Huber, A. and Plies, E. 2000. Simulation of electron trajectories in the 3D field of two coupled hemispherical deflectors. In *Proc. 12th Eur. Cong. on Electron Microscopy*, Brno, Tománek P. and Kolařík, R. Eds., Vol. III, I.473–I.474 (Czechoslovak Society for Electron Microscopy, Brno).
- Huber, A., Bärtle, J. and Plies, E. 2004. Initial experiences with an electrostatic  $\Omega$ -monochromator for electrons. *Nucl. Instrum. Meth. Phys. Res.*, A, 519(1–2), 320–324.
- Hutchison, J.L., Titchmarsh, J.M., Cockayne, D.J.H., Doole, R.C., Hetherington, C.J.D., Kirkland, A.I. and Sawada, H. 2005. A versatile double aberration-corrected energy filtered HREM/STEM for materials science. *Ultramicroscopy*, 103(1), 7–15.
- Ichihashi, M. and Maruse, S. 1971. Correction of spherical aberration by means of a thin conducting foil. *J. Electron Microsc.*, 20(3), 167–171.
- Ichihashi, M. and Maruse, S. 1973. Optical properties of the foil lens for the correction of the spherical aberration. *J. Electron Microsc.*, 22(4), 321–328.
- Inada, H., Zhu, Y., Wall, J., Volkov, V., Nakamura, K., Konno, M., Kaji, K., Jarausch, K. and Twesten, R.D. 2008. The newly installed aberration corrected dedicated STEM (Hitachi HD2700C) at Brookhaven National Laboratory. *Proc. European Microscopy Congress*, Aachen, Luysberg, M., Tillmann, K. and Weirich, T. Eds., Instrumentation and Methods volume (Springer, Berlin).
- Ioanoviciu, D. 1973. Ion optics of a Wien filter with inhomogeneous fields. *Int. J. Mass Spectrom. Ion Phys.*, 11(2), 169–184.
- Ioanoviciu, D. 1974. Transfer matrices of real crossed electric and magnetic fields. *Int. J. Mass Spectrom. Ion Phys.*, 15(1), 89–103.
- Ioanoviciu, D. and Cuna, C. 1974. Electrostatic deflector-Wien filter double focusing mass spectrometers. *Int. J. Mass Spectrom. Ion Phys.*, 15(1), 79–87.
- Ioanoviciu, D., Yefchak, G.E. and Enke, C.G. 1989. Second-order space-time transfer matrix of the two-stage electron mirror. *Int. J. Mass. Spectrom. Ion. Proc.*, 94(3), 281–291.

- Ioanoviciu, D., Tsuno, K. and Martínez, G. 2004. Third order aberration theory of double Wien filters. *Rev. Sci. Instrum.*, 75, 4434–4441.
- Irsen, S., Kelm, K., Schillinger, R., Schmidt, H., Schmalzried, C., Telle, R., Mader, W. and Quandt, E. 2007. A new monochromated and illumination-C<sub>s</sub>-corrected HR(S)TEM: first applications in materials science. *Microsc. Microanal.*, 13, Suppl. 3, 6–7.
- Ishizuka, K., Tanji, T., Tonomura, A., Ohno, T. and Murayama, Y. 1994. Aberration correction using off-axis holography. I. Aberration assessment. II. Beyond the Scherzer limit. *Ultramicroscopy*, 53(4), 361–370 and 55(2), 197–207.
- Jia, C.L., Lentzen, M. and Urban, K. 2003. Atomic-resolution imaging of oxygen in perovskite ceramics. *Science*, 299(5608), 870–873.
- Joy, D.C. 2008. The aberration corrected SEM. In *Biological Low-Voltage Scanning Electron Microscopy*, Schatten, H. and Pawley, J.B., Eds., 107–127 (Springer, New York).
- Kaashoek, J. 1968. A study of magnetic-deflection errors. *Philips Res. Rep., Suppl.* 11, 1–114.
- Kabius, B., Haider, M., Uhlemann, S., Schwan, E., Urban, K. and Rose, H. 2002. First application of a spherical-aberration corrected transmission electron microscope in materials science. *J. Electron Microsc.*, 51, S51–S58.
- Kabius, B., Miller, D.J. and Zaluzec, N.J. 2004. Aberration correction for the TEAM project: design and applications. In *Proc. 13th Eur. Microscopy Cong.*, Antwerp, Dyck, D. van and Oostveldt, P. van Eds., Vol. I, 25–26 (Belgian Society for Microscopy, Liège).
- Kahl, F. and Rose, H. 1995. Theoretical concepts of electron holography. *Adv. Imaging & Electron Phys.*, 94, 197–257.
- Kahl, F. and Rose, H. 1998. Outline of an electron monochromator with small Boersch effect. In *Proc. 14th Int. Conf. on Electron Microscopy*, Cancún, Calderón, H.A. and Yacamán, M.J. Eds., Vol. 1, 71–72 (Institute of Physics, Bristol and Philadelphia).
- Kahl, F. and Rose, H. 2000. Design of a monochromator for electron sources. In *Proc. 12th Eur. Cong. on Electron Microscopy*, Brno, Tománek, P. and Kolařík, R. Eds., Vol. III, I.459–I.460 (Czechoslovak Society for Electron Microscopy, Brno).
- Kahl, F. and Voelkl, E. 2001. Present concepts and designs for gun monochromators. *Microsc. Microanal.*, 7, Suppl. 2, 922–923.
- Kaiser, U., Chuvilin, A., Schröder, R.R., Haider, M. and Rose, H. 2008. Sub-Ångström low-voltage electron microscopy—future reality for deciphering the structure of beam-sensitive nanoobjects? *Proc. European Microscopy Congress*, Aachen, Luysberg, M., Tillmann, K. and Weirich, T. Eds., Instrumentation and Methods volume (Springer, Berlin).
- Kanaya, K. 1985. Reminiscences of the development of electron optics and electron microscope instrumentation in Japan. *Adv. Electron. Electron Phys.*, Suppl. 16, 317–386.
- Kanaya, K. and Ishikawa, A. 1959. Misalignment defects on the electron microscope images. *J. Electronmicrosc.*, 7, 13–15.
- Kanaya, K. and Kawakatsu, H. 1961. Electrostatic stigmators used in correcting second and third order astigmatisms in the electron microscope. *Bull. Electrotech. Lab.*, 25, 641–656.
- Kaneyama, T., Tomita, T., Hosokawa, F., Sawada, H., Sannomiya, T., Deguchi, S., Kawazoe, M., Miyata, T., Kobayashi, E., Kondo, Y., Tanishiro, Y and Takayanagi, K. 2006. Design and development of a 300 kV super-high resolution FETEM R005. In *Proc. 16th Int. Microscopy Cong.*, Sapporo, Ichinose, H. and Sasaki, T. Eds., Vol. 2, 616 (IMC-16 Publishing Committee, Sapporo).
- Kang, Y.-f., Tang, T.-t., Ren, Y. and Guo, X.-l. 2007. Differential algebraic method for computing the high order aberrations of practical electron lenses. *Optik*, 118(4), 158–162.
- Kasper, E. 1968/9a. Die Korrektur des Öffnungs- und Farbfehlers im Elektronenmikroskop durch Verwendung eines Elektronenspiegels mit überlagertem Magnetfeld. *Optik*, 28(1), 54–64.
- Kasper, E. 1968/9b. Die Korrektur von Bildfehlern, die in elektronenoptischen Systemen mit Ringblende durch Störung der Rotationssymmetrie verursacht werden. *Optik*, 28(4), 407–416.
- Kasper, E. 1987a. An advanced method of field calculation in electron optical systems with axisymmetric boundaries. *Optik*, 77(1), 3–12.
- Kasper, E. 1987b. Computer simulation of electron optical systems. *Nucl. Instrum. Methods Phys. Res. A*, 258(3), 466–479.
- Kato, M. and Sekine, T. 1995. Spherical aberration correction of electrostatic lenses using spherical meshes. *J. Vac. Sci. Technol.*, A, 13(4), 2255–2260.
- Kato, M. and Sekine, T. 1996. Evaluation of the dispersive nature of meshes used for the spherical aberration correction of electrostatic lenses. *J. Vac. Sci. Technol.*, A, 14(2), 453–461.

- Kato, M. and Tsuno, K. 1990. Numerical analysis of trajectories and aberrations of a Wien filter including the effect of fringing fields, *Nucl. Instrum. Meth. Res., A*, 298(1–3), 296–320.
- Kazumori, H., Honda, K., Matuya, M. and Date, M. 2004a. Field emission SEM with a spherical and chromatic aberration corrector. In *Proc. 8th Asia-Pacific Conf. on Electron Microscopy*, Kanazawa, Tanaka, N., Takano, Y., Mori, H., Seguchi, H., Iseki, S., Shimada, H. and Simamura, E. Eds., 52–53 CD (8APEM Publication Committee, Uchinada, Ishikawa).
- Kazumori, H., Honda, K., Matsuya, M., Date, M. and Nielsen, C. 2004b. Field emission SEM with a spherical and chromatic aberration corrector. *Microsc. Microanal.*, 10, Suppl. 2, 1370–1371 (CD).
- Kel'man, V.M. and Yavor, S.Ya. 1961. [Achromatic quadrupole electron lenses.] *Zh. Tekh. Fiz.*, 31(12), 1439–1442; *Sov. Phys. Tech. Phys.*, 6(12), 1052–1054.
- Kel'man, V.M., Fedulina, L.V. and Yakushev, E.M. 1971a. [Combined cylindrical electron mirrors]. I. *Zh. Tekh. Fiz.*, 41(10), 2016–2022; *Sov. Phys. Tech. Phys.*, 16(10), 1598–1602.
- Kel'man, V.M., Fedulina, L.V. and Yakushev, E.M. 1971b. [Paraxial properties and aberrations of a two-electrode cylindrical mirror]. *Zh. Tekh. Fiz.*, 41(9), 1832–1838; *Sov. Phys. Tech. Phys.*, 16(9), 1449–1454.
- Kel'man, V.M., Fedulina, L.V. and Yakushev, E.M. 1971c. [Theory of electrostatic cylindrical electron mirrors]. *Zh. Tekh. Fiz.*, 41(7), 1489–1497; *Sov. Phys. Tech. Phys.*, 16(7), 1171–1176.
- Kel'man, V.M., Fedulina, L.V. and Yakushev, E.M. 1972a. [Theory of combined cylindrical electron mirrors], II. *Zh. Tekh. Fiz.*, 42(2), 297–302; *Sov. Phys. Tech. Phys.*, 17(2), 238–241.
- Kel'man, V.M., Sekunova, L.M. and Yakushev, E.M. 1972b. [New method for design of axially symmetric electron mirrors]. *Zh. Tekh. Fiz.*, 42(11), 2279–2287; *Sov. Phys. Tech. Phys.*, 17(11), 1786–1791.
- Kel'man, V.M., Sapargaliev, A.A. and Yakushev, E.M. 1972–4. [Theory of the cathode lens. I. Cylindrical cathode lens. II. Electrostatic cathode lens with rotational field symmetry. III. Derivation of equations for composite cathode lenses. IV. Aberrations of composite cathode lenses]. *Zh. Tekh. Fiz.*, 42(10), 2001–2009; 43(1), 52–60; 44(5), 928–937, 938–943; *Sov. Phys. Tech. Phys.*, 17(10), 1607–1611; 18(1), 33–37; 19(5), 592–596, 597–599.
- Kel'man, V.M., Sekunova, L.M. and Yakushev, E.M. 1973. [Theory of axisymmetric electron mirrors. I. Trajectory equations. II. Electron optical properties], *Zh. Tekh. Fiz.*, 43(9), 1799–1806 and 1807–1817; *Sov. Phys. Tech. Phys.*, 18(9), 1142–1146 and 1147–1152.
- Khursheed, A. 2002. A low-voltage time-of-flight electron emission microscope. *Optik*, 113(11), 505–509.
- Khursheed, A. 2005. A method of dynamic chromatic correction in low-voltage scanning electron microscopes. *Ultramicroscopy*, 103(4), 255–260.
- Kirkland, A. 2007. Aberration corrected TEM: current status and future prospects. In *Proc. of EMAG 2007*, Brown, P.D., Moebus, G. and Baker, R.T., Eds., *J. Phys: Conf Series*.
- Kirkland, A.I., Meyer, R.R. and Chang, L.-y. S. 2006. Local measurement and computational refinement of aberrations for HRTEM. *Microsc. Microanal.*, 12(6), 461–468.
- Kirkland, A.I., Chang, S.L.-y. and Hutchison, J.L. 2007. Atomic resolution transmission electron microscopy. In *Science of Microscopy*, Hawkes, P.W. and Spence, J.C.H. Eds., 3–64 (Springer, New York).
- Kolesnikov, S.V. and Monastyrskii, M.A. 1988. [General theory of spatial and temporal aberrations in cathode lenses with slightly broken axial symmetry], I, II. *Zh. Tekh. Fiz.*, 58(1), 3–10 and 11–19; *Sov. Phys. Tech. Phys.*, 33(1), 1–5 and 6–10.
- Könenkamp, R., Rempfer, G., Word, R., Almaraz, L., Dixon, T., Elstner, J., Jones, T., Nisenfeld, M. and Skoczyłas, W. 2008. First results obtained with an aberration-corrected PEEM employing a hyperbolic electron mirror. *Phys. Procedia*.
- Krivanek, O. 1994. Three-fold astigmatism in high-resolution transmission electron microscopy. *Ultramicroscopy*, 55(4), 419–433.
- Krivanek, O.L. and Fan, G.Y. 1992a. Application of slow-scan charge-coupled device (CCD) cameras to online microscope control. *Scanning Microsc.*, Suppl. 6, 105–114.
- Krivanek, O.L. and Fan, G.Y. 1992b. Complete HREM autotuning using automated diffractogram analysis. In *Proc. 50th Annu. Meeting EMSA*, 1, 96–97.
- Krivanek, O.L. and Leber, M.L. 1993. Three-fold astigmatism: an important TEM aberration. In *Proc. 51st Annu. Meeting Microscopy Society of America*, Bailey, G.W. and Rieder, C.L. Eds., 972–973 (San Francisco Press, San Francisco).
- Krivanek, O.L. and Leber, M.L. 1994. Auto tuning for 1 Å resolution. In *Proc. 13th Int. Cong. Electron Microscopy*, Paris, Vol. 1, 157–158 (Editions de Physique, Les Ulis).
- Krivanek, O.L., Dellby, N., Spence, A.J., Camps, A. and Brown, L.M. 1997a. Aberration correction in the STEM. In *Proc. EMAG 1997*, Rodenburg, J.M., Ed., 35–39 (Institute of Physics Conf. Ser. 153).

- Krivanek, O.L., Dellby, N., Spence, A.J., Camps, R.A. and Brown, L.M. 1997b. On-line aberration measurement and correction in STEM. *Microsc. Microanal.*, 3, Suppl. 2, 1171–1172.
- Krivanek, O.L., Dellby, N. and Brown, L.M. 1998. Spherical aberration corrector for a dedicated STEM. In *Proc. EUREM-11, the 11th Eur. Conf. on Electron Microscopy*, Dublin 1996, CESEM, Ed., Vol. I, 352–353 (CESEM, Brussels).
- Krivanek, O.L., Dellby, N. and Lupini, A.R. 1999a. Towards sub-Å electron beams. *Ultramicroscopy*, 78(1–4), 1–11.
- Krivanek, O.L., Dellby, N. and Lupini, A.R. 1999b. STEM without spherical aberration. *Microsc. Microanal.*, 5, Suppl. 2, 670–671.
- Krivanek, O.L., Dellby, N. and Lupini, A.R. 2000. Advances in  $C_s$ -corrected STEM. In *Proc. 12th Eur. Cong. on Electron Microscopy*, Brno, Tománek, P. and Kolařík, R. Eds., Vol. III, I.149–I.150 (Czechoslovak Society for Electron Microscopy, Brno).
- Krivanek, O.L., Dellby, N., Nellist, P.D., Batson, P.E. and Lupini, A.R. 2001. Aberration-corrected STEM: the present and the future. *Microsc. Microanal.*, 7, Suppl. 2, 896–897.
- Krivanek, O.L., Dellby, N., Murfitt, M., Nellist, P.D. and Szilagyi, Z. 2002. STEM aberration correction: where next? *Microsc. Microanal.*, 8, Suppl. 2, 20–21.
- Krivanek, O.L., Nellist, P.D., Dellby, N., Murfitt, M.F. and Szilagyi, Z. 2003. Towards sub-0.5 Å beams. *Ultramicroscopy*, 96(3–4) 229–237.
- Krivanek, O.L., Corbin, G.J., Dellby, N., Murfitt, M., Nagesha, K., Nellist, P.D. and Szilagyi, Z. 2004. Nion UltraSTEM: a new STEM for sub-0.5 Å imaging and sub-0.5 eV analysis. In *Proc. 13th Eur. Microscopy Cong.*, Antwerp, Dyck, D. van and Oostveldt, P. van Eds., Vol. I, 35–36 (Belgian Society for Microscopy, Liège).
- Krivanek, O.L., Bacon, N., Corbin, G.J., Dellby, N., Hrcicirik, P., Smith, A., Murfitt, M., Nellist, P., Own, C., Woodruff, J. and Szilagyi, Z. 2005a. Atomic-resolution imaging and EELS analysis by aberration-corrected STEM. In *Proc. 7th Multinat. Cong. Electron Microscopy*, Portorož, Čeh, Dražič, G. and Fidler, S. Eds., 31–36.
- Krivanek, O.L., Dellby, N., McManama-Smith, A., Murfitt, M., Nellist, P.D. and Own, C.S. 2005b. An aberration-corrected STEM for diffraction studies. *Microsc. Microanal.*, 11, Suppl. 2, 544–545 (CD).
- Krivanek, O.L., Bacon, N.J., Corbin, G.J., Dellby, N., Elston, B.F., Hrcicirik, P., Keyse, R.J., Murfitt, M.F., Own, C.S., Szilagyi, Z.S. and Woodruff, J.W. 2006. New approaches to instrumentation for high resolution STEM and TEM. In *Proc. 16th Int. Microscopy Cong.*, Sapporo, Ichinose, H. and Sasaki, T. Eds. Vol. 2, 615 (IMC-16 Publishing Committee, Sapporo).
- Křivánek, O.L., Bacon, N., Corbin, G.J., Dellby, N., Elston, B.F., Hrcicirik, P., Keyse, R.J., Murfitt, M.F., Own, C.S., Woodruff, J.W. and Szilagyi, Z.S. 2007a. Atomic resolution nanoanalysis. In *Proc. 8th Multinat. Conf. on Microscopy*, Prague, Nebesářová, J. and Hozák, P. Eds., 3–8 (Czechoslovak Microscopy Society, Brno).
- Krivanek, O.L., Corbin, G.J., Dellby, N., Elston, B., Keyse, R., Murfitt, M., Own, C.S. and Szilagyi, Z.S. 2007b. UltraSTEM progress: flexible electron optics, high-performance sample stage. *Microsc. Microanal.*, 13, Suppl. 2, 878–879 (CD).
- Krivanek, O.L., Corbin, G.J., Dellby, N., Elston, B.F., Keyse, R.J., Murfitt, M.F., Own, C.S., Szilagyi, Z.S. and Woodruff, J.W. 2008a. An electron microscope for the aberration-corrected era. *Ultramicroscopy*, 108(3), 179–195.
- Krivanek, O.L., Dellby, N., Murfitt, M., Own, C.S., and Szilagyi, Z.S. 2008b. STEM aberration correction: an integrated approach. *Proc. European Microscopy Congress*, Aachen, Luysberg, M., Tillmann, K. and Weirich, T. Eds., Instrumentation and Methods volume (Springer, Berlin).
- Krivanek, O.L., Dellby, N., Keyse, R.J., Murfitt, M.F., Own, C.S. and Szilagyi, Z.S. 2008c. Advances in aberration-corrected STEM and EELS. *Adv. Imaging & Electron Phys.* 153.
- Kruit, P. and Amersfoort, P.W. van Eds., 1999. Proc. of the Fifth Int. Conf. on Charged Particle Optics (CPO-5), Delft, *Nucl. Instrum. Meth. Phys. Res.*, A, 427, 1–422.
- Kujama, S., Erni, R., Freitag, B., Tiemeijer, P. and Stekelenburg, M. 2006. High resolution TEM at 80 kV acceleration voltage—is approaching 1 ångström possible? *Microsc. Microanal.*, 12, Suppl. 2, 1470–1471 (CD).
- Kulikov, Yu.V., Monastyrskii, M.A., Ushakov, V.K. and Shchelev, M.Ya. 1985. [Theory and calculation of the temporal aberrations of cathode lenses]. *Trudy Fiz. Inst. P.N. Lebedev*, 155, 147–158.
- Kurihara, K. 1990. Aberration due to misalignment in electrostatic lens and deflection systems. *J. Vac. Sci. Technol. B*, 8(3), 452–455.
- Lambrakis, E., Marai, F.Z. and Mulvey, T. 1977. Correction of spiral distortion in the transmission electron microscope. In *Proc. EMAG77*, Misell, D.L. Ed., 35–38 (Institute of Physics Conf. Ser. 36).

- Lehmann, M. 2000. Determination and correction of the coherent wave aberrations from a single off-axis electron hologram by means of a genetic algorithm. *Ultramicroscopy*, 85(3), 165–182.
- Lehmann, M. 2004. Quantitative evaluation of off-axis electron holograms with atomic resolution. In *Proc. 13th Eur. Microscopy Cong.*, Antwerp, Dyck, D. van and Oostveldt, P. van Eds., Vol. I, 183–184 (Belgian Society for Microscopy, Liège).
- Lehmann, M. 2005. Electron holography for quantitative measurements at atomic dimensions. *Proc. Microscopy Conf. 2005*, Davos, 20 (Paul-Scherrer-Institut Proc. 05-01).
- Lehmann, M. and Lichte, H. 2002. Tutorial on off-axis electron holography. *Microsc. Microanal.*, 8(6), 447–466.
- Lehmann, M. and Lichte, H. 2005. Electron holographic materials analysis at atomic dimensions. *Crystal Res. Technol.*, 40(1–2), 149–160.
- Lehmann, M., Geiger, D. and Lichte, H. 2005. First experiences using electron holography with aberration-corrected TEM. *Microsc. Microanal.*, 11, Suppl. 2, 2146–2147.
- Leisegang, S. 1953. Zum Astigmatismus von Elektronenlinsen. *Optik*, 10(1–3), 5–14.
- Leisegang, S. 1954a. Ein einfacher Stigmator für magnetische Elektronenlinsen. *Optik*, 11(2), 49–60.
- Leisegang, S. 1954b. Zur Zentrierung magnetischer Elektronenlinsen. *Optik*, 11(9), 397–406.
- Leisegang, S. 1956. Elektronenmikroskope. *Handbuch der Physik*, 33, 396–545.
- Lenc, M. and Lencová, B. 1988. On the deflection of electrons with homogeneous fields. *Optik*, 78(4), 127–131.
- Lenc, M. and Müllerová, I. 1992a. Electron optical properties of a cathode lens. *Ultramicroscopy*, 41(4), 411–417.
- Lenc, M. and Müllerová, I. 1992b. Optical properties and axial aberration coefficients of the cathode lens in combination with a focusing lens. *Ultramicroscopy*, 45(1), 159–162.
- Lencová, B. 1981. Deflection aberrations of multi-stage deflection systems. *Optik*, 58(1), 25–35.
- Lencová, B. 1988. On the design of electron beam deflection systems. *Optik*, 79(1), 1–12.
- Lencová, B. 1995. Computation of electrostatic lenses and multipoles by the first order finite element method. *Nucl. Instrum. Meth. Phys. Res., A*, 363(1–2), 190–197.
- Lencová, B. and Hawkes, P.W. 2004. Chromatic aberrations of electrostatic lens doublets. In *Electron Microscopy and Analysis 2003*, McVitie, S. and McComb, D. Eds., 191–194 (Institute of Physics Conf. Ser. 179).
- Lencová, B. and Lenc, M. 1990. Computation of multi-lens focusing systems. *Nucl. Instrum. Meth. Phys. Res. A*, 298(1–3), 45–55.
- Lencová, B. and Lenc, M. 1994. Computation of properties of electrostatic lenses. *Optik*, 97(3), 121–126.
- Lencová, B. and Lenc, M. 1997. Third order geometrical and first order chromatic aberrations of electrostatic lenses. *Optik*, 105(3), 121–128.
- Lencová, B. and Oral, M. 2006. Programs for electron optical design. In *Proc. 10th Seminar on Recent Trends in Charged Particle Optics and Surface Physics Instrumentation*, Skalský Dvůr, Müllerová, I. Ed., 41–44 (ISI, Brno).
- Lencová, B. and Wisselink, G. 1990. Program for the computation of lenses and deflectors. *Nucl. Instrum. Meth. Phys. Res., A*, 298(1–3), 56–66.
- Lencová, B. and Zlámal, J. 2005. EOD (Electron Optical Design) program for computations for electron microscopy. *Proc. Microscopy Conf. 2005*, Davos, 8 (Paul-Scherrer-Institut Proc. 05–01).
- Lencová, B. and Zlámal, J. 2006. EOD (Electron Optical Design) program features. In *Proc. 10th Seminar on Recent Trends in Charged Particle Optics and Surface Physics Instrumentation*, Skalský Dvůr, Müllerová, I. Ed., 41–44 (ISI, Brno).
- Lencová, B. and Zlámal, J. 2008. A new program for the design of electron microscopes. *Phys. Procedia*.
- Lencová, B., Lenc, M. and Hawkes, P.W. 2008. A pitfall in the calculation of higher order aberrations. *Ultramicroscopy*, 108(8), 737–740.
- Lentzen, M. 2004. The tuning of a Zernike phase plate with defocus and variable spherical aberration and its use in HRTEM imaging. *Ultramicroscopy* 99(4), 211–220.
- Lentzen, M. and Thust, A. 2006. Optimum aberration setting for sub-Ångström HRTEM and contrast theory for Wien-filter monochromators. In *Proc. 16th Int. Microscopy Cong.*, Sapporo, Ichinose, H. and Sasaki, T. Eds., Vol. 2, 638 (IMC-16 Publishing Committee, Sapporo).
- Lentzen, M., Jahnen, B., Jia, C.L., Thust, A., Tillmann, K. and Urban, K. 2002. High-resolution imaging with an aberration-corrected transmission electron microscope. *Ultramicroscopy*, 92(3–4), 233–242.
- Lenz, F. 1987. Preferential axis in electron optical systems with perturbed symmetry. In *Proc. Int. Symp. Electron Optics*, Beijing 1986, Ximen, J.-y. Ed., 122–126 (Institute of Electronics, Academia Sinica, Beijing).

- Le Poole, J.B. 1972. Space charge correction of spherical aberration. In *Proc. 5th Eur. Congr. Electron Microscopy*, Manchester, 130–131.
- Li, Y. 1986. On the relativistic fifth order geometrical aberration equation of a combined focusing–deflection system. *Optik*, 74(2), 65–70.
- Li, Y. 1993. A theory of relativistic asymptotic aberrations of a combined focusing-deflection system. *Sci. China*, A36(8), 957–968.
- Li, Y. 1994. A general theory of relativistic asymptotic aberrations of a system of combined focusing-deflection sub-systems. *J. Phys. D: Appl. Phys.*, 27(11), 2271–2283.
- Li, Y. and Ni, W.-x. 1988. Relativistic fifth order geometrical aberrations of a focusing system. *Optik*, 78(2), 45–47.
- Li, Y. and Ximen, J.-y. 1982. [On the relativistic aberration theory of a combined focusing-deflection system with multi-stage deflectors.] *Acta Phys. Sin.*, 31(5), 604–614.
- Li, C., Mu, C.-j. and Chen, E.-g. 1995. Fifth-order aberrations of rotationally symmetrical electron optics system. *J. Yunnan Univ. Nat. Sci.*, 17(2), 11–18.
- Li, C., Mu, C.-j., Zhang, Y.-m., Chen, X.-m. and Liu, T.-g. 2002. Normalized fifth-order spherical aberration coefficients derived from a Gaussian trajectory on rotationally symmetric magnetic lenses. *Micron*, 33(4), 353–356.
- Lichte, H. 1991. Electron image plane off-axis holography of atomic structures. *Adv. Opt. Electron Microsc.*, 12, 25–91.
- Lichte, H. 1997. Electron holography methods. In *Handbook of Microscopy, Methods I*, Amelinckx, S., Dyck, D. van, Landuyt, J. van. and Tendeloo, G. van Eds., 515–536 (VCH, Weinheim and New York); reprinted in *Electron Microscopy, Principles and Fundamentals*, Amelinckx, S., Dyck, D. van, Landuyt, J. van. and Tendeloo, G. van Eds., 271–292 (VCH, Weinheim and New York).
- Lichte, H. 2002. Electron holography: mystery and reality. *Phil. Trans. R. Soc. London*, A, 360, 897–920.
- Lichte, H. 2003. Which atoms are where? Which fields are around? *Microsc. Microanal.*, 9, Suppl. 2, 768–769.
- Lichte, H. 2005. Which atom is where? Which fields are around? Answers offered by electron holography. In *Proc. 7th Multinat. Cong. Electron Microscopy*, Portorož, Čeh, M., Dražič, G. and Fidler, S. Eds., 19–24.
- Lichte, H. 2006. Electron holography: present state in the Triebenberg Laboratory. In *Proc. 16th Int. Microscopy Cong.* Sapporo, Ichinose, H. and Sasaki, T. Eds., Vol. 2, 723 (IMC-16 Publishing Committee, Sapporo).
- Lichte, H. 2008. Performance limits of electron holography. *Ultramicroscopy* 108(3), 256–262.
- Lichte, H. and Lehmann, M. 2002. Electron holography: a powerful tool for the analysis of nanostructures. *Adv. Imaging & Electron Phys.*, 123, 225–255.
- Lichte, H. and Lehmann, M. 2003. Electron holography. *Microsc. Microanal.*, 9, Suppl. 3, 14–15.
- Lichte, H. and Lehmann, M. 2008. Electron holography—basics and applications. *Repts Prog. Phys.*, 71(1), 16102.
- Liu, H. 1999. Simulation of multipole systems including primary and secondary aberration analysis. *Proc. SPIE*, 3777, 92–106.
- Liu, Z. 2002. Fifth-order canonical geometric aberration analysis of electrostatic round lenses. *Nucl. Instrum. Meth. Phys. Res., A*, 488, 42–50.
- Liu, Z.-x. 2004a. Improved fifth-order geometric aberration coefficients of electron lenses. *J. Phys. D: Appl. Phys.*, 37(5), 653–659.
- Liu, Z.-x. 2004b. Differential algebraic description for third- and fifth-order aberrations of electromagnetic lenses. *Nucl. Instrum. Meth. Phys. Res., A*, 519 (1–2), 154–161.
- Liu, Z.-x. 2006. Differential algebraic method for aberration analysis of typical electrostatic lenses. *Ultramicroscopy*, 106(3), 220–232.
- Liu, X.D. and Tang, T.-t. 1995. Aberration analysis of Wien filters and design of an electron energy-selective imaging system. *Nucl. Instrum. Meth. Phys. Res., A*, 363(1–2), 254–260.
- Liu, Z. and Ximen, J.-y. 1992. Numerical calculation of higher order chromatic aberrations for rotationally symmetric electrostatic lenses. *J. Appl. Phys.*, 72(1), 28–32.
- Liu, H., Munro, E., Rouse, J. and Zhu, X. 2002. Simulation methods for multipole imaging systems and aberration correctors. *Ultramicroscopy*, 93(3–4), 271–291.
- Liu, H.-n., Rouse, J., Munro, E. and Wang, L.-p. 2008. Software for designing multipole aberration correctors. *Phys. Procedia*.
- Lupini, A.R. and Pennycook, S.J. 2007. Aberration corrected imaging in the STEM. *Microsc. Microanal.*, 13, Suppl. 2, 1146–1147 (CD).
- Lupini, A.R., Varela, M., Borisevich, A.Y., Travagliini, S.M. and Pennycook, S.J. 2003. Advances in aberration corrected STEM at ORNL. In *Proc. EMAG 2003*, McVitie, S. and McComb, D. Eds., 211–214 (Institute of Physics Conf Ser. 179).

- Maas, D.J., Henstra, A., Krijn, M.P.C.M. and Mentink, S.A.M. 2000. Electrostatic correction in LV-SEM. *Microsc. Microanal.*, 6, Suppl. 2, 746–747.
- Maas, D.J., Henstra, S., Krijn, M. and Mentink, S. 2001. Electrostatic aberration correction in low-voltage SEM. *Proc. SPIE*, 4510, 205–217.
- Maas, D., Mentink, S. and Henstra, S. 2003. Electrostatic aberration correction in low-voltage SEM. *Microsc. Microanal.*, 9, Suppl. 3, 24–25.
- Mahwin, J. 1994. Introduction. In *Advances in Analysis, Probability and Mathematical Physics: Contributions of Nonstandard Analysis*, Albeverio, S., Luxemburg, W.A.J. and Wolff, M.P.H. Eds., i–iii (Springer, Heidelberg).
- Majorovits, E., Barton, B., Schultheiß, K., Pérez-Willard, F., Gerthsen, D. and Schröder, R.R. 2006. Optimizing phase contrast in transmission electron microscopy with an electrostatic (Boersch) phase plate. *Ultramicroscopy*, 107(2–3), 213–226.
- Makino, K. and Berz, M. 1997. Arbitrary order aberrations for elements characterized by measured fields. *Proc. SPIE*, 3155, 221–227.
- Makino, K. and Berz, M. 1999. COSY INFINITY version 8. *Nucl. Instrum. Meth. Phys. Res.*, A, 427(1–2), 338–343.
- Marai, F.Z. and Mulvey, T. 1977. Scherzer's formula and the correction of spiral distortion in the electron microscope. *Ultramicroscopy*, 2(2–3), 187–192.
- Marianowski, K. and Plies, E. 2008. Third-rank computation of electron and ion optical systems with several and rotated Wien filters. *Proc. European Microscopy Congress*, Aachen, Luysberg, M., Tillmann, K. and Weirich, T. Eds., Instrumentation and methods volume (Springer, Berlin).
- Martínez, G. and Sancho, M. 1995. Accurate calculation of geometric aberrations in electrostatic lenses. *Nucl. Instrum. Meth. Phys. Res.*, A, 363(1–2), 198–204.
- Martínez, G. and Tsuno, K. 2002. BEM simulation of Wien filters. *Ultramicroscopy*, 93(3–4), 253–261.
- Martínez, G. and Tsuno, K. 2004a. Design of Wien filters with high resolution. *Ultramicroscopy*, 100(1–2), 105–114.
- Martínez, G. and Tsuno, K. 2004b. 2nd order aberration corrected Wien filter. In *Proc. 13th Eur. Microscopy Cong.*, Antwerp, Dyck, D. van and Oostveldt, P. van Eds., Vol. I, 63–64 (Belgian Society for Microscopy, Liège).
- Martínez, G. and Tsuno, K. 2008. The use of multipole fields for aberration correction in  $\pi/2$  Wien filters. *Phys. Procedia*.
- Maruse, S., Hiratake, S. and Ichihashi, M. 1970a. Korrektur vom Öffnungsfehler des Elektronenlinse. *Jpn. J. Appl. Phys.*, 9(12), 1549–1550.
- Maruse, S., Ichihashi, M. and Hiratake, S. 1970b. Korrektur des Öffnungsfehlers mittels einer durchstrahlten Folie. In *Proc. 7th. Int. Congr. Electron Microscopy*, Grenoble, Vol. II, 1–2.
- Matijevic, M., Stegmann, H., Essers, E. and Benner, G. 2007. Recent results of an in-column energy filter TEM with monochromator. *Microsc. Microanal.*, 13, Suppl. 2, 1160–1161 (CD).
- Matijevic, M., Lengweiler, S., Preikszas, D., Müller, H., Schröder, R.R. and Benner, G. 2008. Electron optical design of the phase aberration corrected electron microscope. *Proc. European Microscopy Congress*, Aachen, Luysberg, M., Tillmann, K. and Weirich, T. Eds., Instrumentation and Methods volume (Springer, Berlin).
- Matsuda, J.-i. and Ura, K. 1974. The aberration theory of the electron trajectory in the R-F fields. Part 1. The second-order aberration. Part 2. The third-order aberration. *Optik*, 40(2), 179–192 and (3), 284–295.
- Matsuda, H., Daimon, H., Kato, M. and Kudo, K. 2005. Approach for simultaneous measurement of two-dimensional angular distribution of charged particles: spherical aberration correction using ellipsoidal mesh. *Phys. Rev. E*, 71, 066503:1–8.
- Matsumoto, T. and Tonomura, A. 1996. The phase constancy of electron waves traveling through Boersch's electrostatic phase plate. *Ultramicroscopy* 63(1), 5–10.
- Matsu, T. 1975. Particle trajectories in a cylindrical condenser calculated in a third-order approximation. *Nucl. Instrum. Meth.*, 126(2), 273–279.
- Matsuya, M. and Nakagawa, S. 2004. Resolution estimation of low energy probe lens system. *Nucl. Instrum. Meth. Phys. Res. A*, 519(1–2), 264–279.
- Matsuya, M., Saito, M. and Nakagawa, S. 1995. A semi-analytical aberration calculation method using polynomials and Lie algebra. *Nucl. Instrum. Meth. Phys. Res.*, A, 363(1–2), 261–269.
- Matsuya, M., Saito, M. and Nakagawa, S. 1997. A study of the high-order aberration coefficients using Lie algebra. *Nucl. Instrum. Meth. Phys. Res.*, A, 385(3), 456–462.

- Matteucci, G., Missiroli, G.F. and Pozzi, G. 1998. Electron holography of electrostatic fields. *Adv. Imaging & Electron Phys.*, 99, 171–240.
- Mayer, J., Houben, L., Lopatin, S., Luysberg, M. and Thust, A. 2007. Performance and applications of the aberration corrected TEM and STEM instruments at the Ernst Ruska-Centre. *Microsc. Microanal.*, 13, Suppl. 2, 128–129.
- Melkitch, A. 1947. Ausgezeichnete astigmatische Systeme der Elektronenoptik, I, II. *Sitzungsber. Akad. Wiss. Wien, Mathemat.-Naturwiss. Kl.*, Abt. IIa, 155(9–10), 393–438 and 439–471.
- Mentink, S.A.M., Steffen, T., Tiemeijer, P.C. and Krijn, M.P.C.M. 1999. Simplified aberration corrector for low-voltage SEM. In *Proc. EMAG 1999*, Kiely, C.J. Ed., 83–86 (Institute of Physics Conf. Ser. 161).
- Mentink, S.A.M., Zande, M.J. van der, Kok, C. and Rooy, T.L. van 2004. Development of a  $C_s$  corrector for a Tecnai 20 FEG STEM/TEM. In *Proc. EMAG 2003*, McVitie, S. and McComb, D. Eds., 165–168 (Institute of Physics Conf. Ser. 179).
- Meyer, R., Kirkland, A.I. and Saxton, W.O. 2002. A new method for the determination of the wave aberration function for high resolution TEM. I. Measurement of the symmetric aberrations. *Ultramicroscopy*, 92(2), 89–109.
- Meyer, R., Kirkland, A.I. and Saxton, W.O. 2004. A new method for the determination of the wave aberration function for high resolution TEM. II. Measurement of the antisymmetric aberrations. *Ultramicroscopy*, 99(2–3), 115–123.
- Mitsuishi, K., Takeguchi, M., Kondo, Y., Hosokawa, F., Okamoto, K., Sanomiya, T., Hori, M., Iwama, T., Kawazoe, M. and Furuya, K. 2006a. Dual-hexapole type  $C_s$  corrector on ultra-high vacuum scanning transmission electron microscope. In *Proc. 16th Int. Microscopy Cong.*, Sapporo, Ichinose, H. and Sasaki, T. Eds., Vol. 2, 625 (IMC-16 Publishing Committee, Sapporo).
- Mitsuishi, K., Takeguchi, M., Kondo, Y., Hosokawa, F., Okamoto, K., Sanomiya, T., Hori, M., Iwama, T., Kawazoe, M. and Furuya, K. 2006b. Ultra-high vacuum third-order spherical aberration ( $C_3$ ) corrector for a scanning transmission electron microscope. *Microsc. Microanal.*, 12(6), 456–460.
- Monastyrskii, M.A. 1989. [Study of the aberrations of emission electron optical systems in regions with low potentials.] *Zh. Tekh. Fiz.*, 59(12), 49–56; *Sov. Phys. Tech. Phys.*, 34(12), 1398–1402.
- Mook, H.W. and Kruit, P. 1998. Fringe field monochromator for high brightness electron sources. In *Proc. 14th Int. Conf. on Electron Microscopy*, Cancún, Calderón, H.A. and Yacamán, M.J. Eds., Vol. 1, 73–74 (Institute of Physics, Bristol and Philadelphia).
- Mook, H.W. and Kruit, P. 1999a. On the monochromatisation of high brightness sources for electron microscopy. *Ultramicroscopy*, 78(1–4), 43–51.
- Mook, H.W. and Kruit, P. 1999b. Optics and design of the fringe-field monochromator for a Schottky field-emission gun. *Nucl. Instrum. Meth. Phys. Res.*, A, 427(1–2), 109–120.
- Mook, H.W. and Kruit, P. 2000. Construction and characterization of the fringe field monochromator for a field emission gun. *Ultramicroscopy*, 81(3–4), 129–139.
- Mook, H.W., Batson, P.E. and Kruit, P. 1999. Operation of the fringe field monochromator in field emission STEM. In *Proc. EMAG 1999*, Kiely, C.J. Ed., 223–226 (Institute of Physics Conf. Ser. 161).
- Mook, H.W., Batson, P.E. and Kruit, P. 2000. Monochromator for high brightness electron guns. In *Proc. 12th Eur. Cong. on Electron Microscopy*, Brno, Tománek, P. and Kolařík, R. Eds., Vol. III, I.315–I.316 (Czechoslovak Society for Electron Microscopy, Brno).
- Moses, R.W. 1973. Lens optimization by direct application of the calculus of variations. In *Image Processing and Computer-Aided Design in Electron Optics*, Hawkes, P.W. Ed., 250–272 (Academic Press, London).
- Moses, R.W. 1974. Aberration correction for high-voltage electron microscopy. *Proc. R. Soc. London*, A, 339(1619), 483–512.
- Mu, C.-j., Chen, E.-g., Gao, J. and Li, C. 1999. Analyses of fifth-order aberrations of rotationally symmetric electron optical system. *Acta Electron. Sin.*, 27(8), 67–69.
- Mukai, M., Kaneyama, T., Tomita, T., Tsuno, K., Terauchi, M., Tsuda, K., Saitoh, K., Naruse, M., Honda, T. and Tanaka, M. 2003a. MIRAI-21 electron microscope—performance of the monochromator. *Microsc. Microanal.*, 9, Suppl. 2, 954–955 (CD).
- Mukai, M., Kaneyama, T., Tomita, T., Tsuno, K., Terauchi, M., Tsuda, K., Saitoh, K., Kersker, M., Naruse, M., Honda, T. and Tanaka, M. 2003b. MIRAI-21 electron microscope—performance of the monochromator. In *Proc. 15th Int. Cong. on Electron Microscopy*, Durban, Cross, R., Engelbrecht, J. and Witcomb, M. Eds., Vol. 3, 321–322 (Microscopy Society of Southern Africa, Onderstepoort).
- Müller, H., Preikszas, D. and Rose, H. 1999. A beam separator with small aberrations. *J. Electron Microsc.*, 48(3), 191–204.

- Müller, H., Uhlemann, S. and Haider, M. 2002. Benefits and possibilities for  $C_c$ -correction in TEM/STEM. *Microsc. Microanal.*, 8, Suppl. 2, 12–13.
- Müller, H., Uhlemann, S., Hartel, P. and Haider, M. 2005. Optical design, simulation, and alignment of present-day and future aberration correctors for the transmission electron microscope. *Proc. Microscopy Conf. 2005*, Davos, 6 (Paul-Scherrer-Institut Proc., 05–01).
- Müller, H., Uhlemann, S., Hartel, P. and Haider, M. 2006. Advancing the hexapole  $C_s$ -corrector for the scanning transmission electron microscope. *Microsc. Microanal.*, 12(6), 442–455.
- Müller, H., Uhlemann, S., Hartel, P. and Haider, M. 2008. Aberration-corrected optics: from an idea to a device. *Phys. Procedia*.
- Munro, E. 1974. Calculation of the optical properties of combined magnetic lenses and deflection systems with superimposed fields. *Optik*, 39(4), 450–466.
- Munro, E. 1980. Electron beam lithography. In *Applied Charged Particle Optics*, Septier, A. Ed., *Adv. Electron. Electron Phys.*, Suppl. 13B, 73–131.
- Munro, E. 1988. Finite difference programs for computing tolerances for electrostatic lenses. *J. Vac. Sci. Technol. B*, 6(3), 941–948.
- Munro, E. and Chu, H.C. 1982. Numerical analysis of electron beam lithography systems. Part I: Computation of fields in magnetic deflectors. Part II: Computation of fields in electrostatic deflectors. *Optik*, 60(4), 371–390 and 61(1), 1–16.
- Munro, E. and Rouse, J.A. Eds., 2008. Proceedings of the Seventh International Conference on Charged Particle Optics (CPO-7) Cambridge, *Phys. Procedia*.
- Munro, E. and Wittels N.D. 1977. Calculation of the optical properties of electron lenses with electron-transparent foils. *Optik*, 47(1), 25–38.
- Munro, E., Zhu, X., Rouse, J.A., Liu, H. 2001. Aberration correction for charged particle lithography. *Proc. SPIE*, 4510, 218–224.
- Munro, E., Rouse, J., Liu, H., Wang, L. and Zhu, X. 2006a. Simulation software for designing electron and ion beam equipment. *Microelec. Eng.*, 83 (4–9), 994–1002.
- Munro, E., Wang, L., Rouse, J. and Liu, H. 2006b. Differential algebraic method for electron optical design. In *Proc. 10th Seminar on Recent Trends in Charged Particle Optics and Surface Physics Instrumentation*, Skalský Dvůr, Müllerová, I. Ed., 53–54 (ISI, Brno).
- Munro, E., Wang, L., Rouse, J. and Liu, H. 2006c. Theory and application of the differential algebraic method for computing the aberrations of electron optical systems. In *Proc. 16th Int. Microscopy Cong.*, Sapporo, Ichinose, H. and Sasaki, T. Eds., Vol. 2, 588 (IMC-16 Publishing Committee, Sapporo).
- Nakabushi, N., Sakurai, T. and Matsuda, H. 1983. Relativistic ion optics of crossed toroidal electric and inhomogeneous magnetic fields. *Int. J. Mass Spectrom. Ion Phys.*, 49(2), 89–101.
- Nakamura, K., Konno, M., Yaguchi, T., Kamino, T., Terauchi, D., Inada, H., Tanaka, H., Taniguchi, Y. and Isakozawa, S. 2006. Development of a  $C_s$ -corrected dedicated STEM. In *Proc. 16th Int. Microscopy Cong.*, Sapporo, Ichinose, H. and Sasaki, T. Eds., Vol. 2, 633 (IMC-16 Publishing Committee, Sapporo).
- Nellist, P.D. 2007. Scanning transmission electron microscopy. In *Science of Microscopy*, Hawkes, P.W. and Spence, J.C.H. Eds., 65–132 (Springer, New York).
- Nellist, P.D., Chisholm, M.F., Dellby, N., Krivanek, O.L., Murfitt, M.F., Szilagyi, Z.S., Lupini, A.R., Borisevich, A., Sides, W.H. and Pennycook, S.J. 2004a. Direct sub-Angstrom imaging of a crystal lattice. *Science*, 305, 1741.
- Nellist, P.D., Dellby, N., Krivanek, O.L., Murfitt, M.F., Szilagyi, Z., Lupini, A.R. and Pennycook, S.J. 2004b. Towards sub 0.5 angstrom beams through aberration corrected STEM. In *Proc. EMAG 2003*, McVitie, S. and McComb, D. Eds., 159–164 (Institute of Physics Conf. Ser. 179).
- Nellist, P.D., Chisholm, M.F., Lupini, A.R., Borisevich, A., Sides, W.H., Pennycook, S.J., Dellby, N., Keyse, R., Krivanek, O.L., Murfitt, M.F. and Szilagyi, Z.S. 2006. Aberration-corrected STEM: current performance and future directions. In *Proc. EMAG-NANO 2005*, Brown, P.D., Baker, R. and Hamilton, B., Eds., *J. Phys.: Conf. Ser.* 26, 7–12.
- Nesslinger, A. 1939. Über Achromasie von Elektronenlinsen, *Jahrbuch AEG Forschungs-Inst.*, 6, 83–85.
- Niimi, H., Chun, W.-J., Suzuki, S., Asakura, K. and Kato, M. 2007. Aberration-corrected multipole Wien filter for energy-filtered x-ray photoemission electron microscopy. *Rev. Sci. Instrum.*, 78(6), 063710:1–8.
- Ohiwa, H., Blackwell, R.J. and Siegel, B.M. 1981. Design of an electrostatic ion optical system for microfabrication with 100 Å resolution. *J. Vac. Sci. Technol.*, 19(4), 1074–1076.
- Ohye, T., Morita, C. and Shimoyama, H. 2006. Calculation of relativistic aberration coefficients of electrostatic accelerating lens. In *Proc. 16th Int. Microscopy Cong.*, Sapporo, Ichinose, H. and Sasaki, T. Eds., Vol. 2, 590 (IMC-16 Publishing Committee, Sapporo).

- O'Keefe, M.A. 2003. HRTEM at half-Ångström resolution: from 0 Å to TEAM. *Microsc. Microanal.*, 9, Suppl. 2, 936–937 (CD).
- O'Keefe, M.A. 2004. Sub-Ångström resolution with aberration-corrected TEM: present and future. In *Materials Research in an Aberration-free Environment*, *Microsc. Microanal.*, 10, Suppl. 3, 68–69 website.
- O'Keefe, M.A., Allard, L.F., Pennycook, S.J. and Blom, D.A. 2006. Transcending the one-ångström atomic resolution barrier in the TEM. *Microsc. Microanal.*, 12, Suppl. 2, 162–163.
- Oldfield, L. 1971. Focal properties of cylindrically symmetric microwave cavities. In *Proc. 25th Anniversary Meeting EMAG*, 94–96 (Institute of Physics Conf. Ser. 10).
- Oldfield, L. 1973. Computer design of high frequency electron-optical systems. In *Image Processing and Computer-Aided Design in Electron Optics*, Hawkes, P.W. Ed., 370–399 (Academic Press, London).
- Oldfield, L. 1974. The use of microwave cavities as electron lenses. In *Proc. 8th Int. Cong. Electron Microscopy*, Canberra, Vol. I, 152–153.
- Orloff, J. 1997. Spherical aberration correction of a focused ion beam with space charge. *Proc. SPIE*, 3155, 216–220.
- Ovsyannikova, L.P. and Yavor, S.Ya. 1965. [Third-order aberration of the width of a line image in a combined quadrupole lens.] *Zh. Tekh. Fiz.*, 35(5), 940–946; *Sov. Tech. Phys.*, 10(5), 723–726.
- Pandey, A.K. and Vaidya, N.C. 1975. Possible use of the synklysmotron lenses in high voltage electron microscopy. In *Proc. 4th Int. Conf. High-Voltage Electron Microscopy*, Toulouse, 51–54.
- Pandey, A.K. and Vaidya, N.C. 1977. Use of synklysmotron lenses in high voltage electron microscopy. In *Proc. 5th Int. Conf. High-Voltage Electron Microscopy*, Kyoto, 53–56.
- Pandey, A.K. and Vaidya, N.C. 1978. On improving the resolving power of high voltage electron microscopes. *Bull. Electron Microsc. Soc. India*, 2(2), 1–2.
- Pennycook, S.J., Lupini, A.R., Kadavanich, A., McBride, J.R., Rosenthal, S.J., Puettner, R.C., Yahil, A., Krivanek, O.L., Dellby, N. and Nellist, P.D. 2003. Aberration-corrected scanning transmission electron microscopy: the potential for nano- and interface science. *Z. Metallkde*, 94(4), 350–357.
- Pennycook, S.J., Chisholm, M.F., Lupini, A.R., Peng, Y., Varela, M., Bentheim, K. van, Borisevich, A., Jong, N. de. and Oxley, M.P. 2006a. Aberration-corrected STEM—more than just higher resolution. In *Proc. 16th Int. Microsc. Cong.*, Sapporo, Ichinose, H. and Sasaki, T. Eds., Vol. 2, 603 (IMC-16 Publishing Committee, Sapporo).
- Pennycook, S.J., Chisholm, M.F., Lupini, A.R., Peng, Y., Varela, M., Bentheim, K. van, Borisevich, A., Jong, N. de. and Oxley, M.P. 2006b. Aberration-corrected STEM—more than just higher resolution. *Microsc. Microanal.*, 12, Suppl. 2, 132–133.
- Plies, E. 1978. Proposal for an electrostatic energy filter and a monochromator. In *Proc. 9th Int. Cong. on Electron Microscopy*, Toronto, Sturgess, J.M. Ed. Vol. 1, 50–51 (Microscopical Society of Canada, Toronto).
- Plies, E. 1982. Berechnung zusammengesetzter elektronenoptischer Fokussier- und Ablenksysteme mit überlagerten Feldern. *Siemens Forsch.-u. Entwickl.-Ber.*, 11(1), 38–45 and (2) 83–90.
- Plies, E. 1994. Electron optics of low-voltage electron beam testing and inspection. *Adv. Opt. Electron. Microsc.*, 13, 123–242.
- Plies, E. 2002. Simplified computation of third-rank image aberrations of electron-optical systems with curved axis. *Ultramicroscopy*, 93(3–4), 305–319.
- Plies, E. and Bärtle, J. 2003. Proposal of a novel highly symmetric Wien filter monochromator. *Microsc. Microanal.*, 9, Suppl. 3, 28–29.
- Plies, E. and Elstner, R. 1989. Computation of first- and third-order properties of compound electron-optical lens and deflector systems. I, II. *Optik*, 82(2), 57–63 and 83(1), 11–18.
- Plies, E. and Schweizer, M. 1987. Calculation of fields and secondary electron trajectories in a new in-lens spectrometer for electron beam testing. *Siemens Forsch. u. Entwickl.-Ber.*, 16(1), 30–36.
- Plies, E. and Typke, D. 1978. Dreidimensional abbildende Elektronenmikroskope. II. Theorie elektronenoptischer Systeme mit gekrümmter Achse. *Z. Naturforsch.*, 33a, 1361–1377.
- Pozzi, G. 1995. Multislice approach to lens analysis. *Adv. Imaging & Electron Phys.*, 93, 173–218.
- Preikszas, D. 1990. Optimierung rotationssymmetrischer kombiniert elektromagnetischer Elektronenlinsen. Diploma-Thesis, Darmstadt.
- Preikszas, D. and Rose, H. 1994. Corrected low-energy electron microscope for multi-mode operation. In *Proc. 13th Int. Conf. Electron Microscopy*, Paris, Vol. 1, 197–198. (Editions de Physique, Les Ulis).
- Preikszas, D. and Rose, H. 1995. Procedures for minimizing the aberrations of electromagnetic compound lenses. *Optik*, 100(4), 179–187.
- Preikszas, D. and Rose, H. 1997. Correction properties of electron mirrors. *J. Electron Microsc.*, 46(1), 1–9.

- Preikszas, D., Hartel, P., Spehr, R. and Rose, H. 2000. SMART electron optics. In *Proc. 12th Eur. Cong. on Electron Microscopy*, Brno, Tománek, P. and Kolařík, R. Eds., Vol. III, I.181–I.184 (Czechoslovak Society for Electron Microscopy, Brno).
- Radlická, T. 2008. Lie algebraic methods in charged particle optics. *Adv. Imaging & Electron Phys.*, 151, 241–362.
- Ramberg, E.G. 1949. Aberration correction with electron mirrors. *J. Appl. Phys.*, 20(2), 183–187.
- Rang, O. 1949a. Der elektrostatische Stigmator, ein Korrektiv für astigmatische Elektronenlinsen. *Optik*, 5(8/9), 518–530.
- Rang, O. 1949b. Eine nach Brechkraft und Azimut einfach einstellbare astigmatische Elektronenlinse. *Phys. Blätter*, 5, 78–80.
- Recknagel, A. 1936. Zur Theorie des Elektronenspiegels. *Z. Tech. Phys.*, 17(12), 643–645.
- Recknagel, A. 1937. Zur Theorie des Elektronenspiegels. *Z. Phys.*, 104(5–6), 381–394.
- Recknagel, A. 1941a. Über die sphärische Aberration bei elektronenoptischer Abbildung. *Z. Phys.*, 117, 67–73.
- Recknagel, A. 1941b. Theorie des elektrischen Elektronenmikroskops für Selbststrahler. *Z. Phys.*, 117(11–12), 689–708.
- Recknagel, A. and Haufe, G. 1952/3. Über den Stigmator der Elektronenmikroskope. *Wiss. Z. Tech. Hochschule Dresden*, 2(1), 1–10.
- Reimer, L., Ed. 1995. *Energy-filtering Transmission Electron Microscopy* (Springer, Berlin and New York).
- Rempfer, G.F. 1985. Unipotential electrostatic lenses: paraxial properties and aberrations of focal length and focal point. *J. Appl. Phys.*, 57(7), 2385–2401.
- Rempfer, G. 1990a. A theoretical study of the hyperbolic electron mirror as a correcting element for spherical and chromatic aberration in electron optics. *J. Appl. Phys.*, 67(10), 6027–6040.
- Rempfer, G.F. 1990b. Comment on “A study on hyperbolic mirrors as correctors”. *Optik*, 86(2), 73–74.
- Rempfer, G.F. 1999a. The relationship of lens aberrations to image aberrations in electron optics. *Optik*, 110(1), 17–24.
- Rempfer, G.F. 1999b. Reply to “Rebuttal of ‘The relationship of lens aberrations to image aberrations in electron optics’” by P.W. Hawkes, *Optik*, 110(8), 375–377.
- Rempfer, G.F. 2000. Reply to “Clarification of my criticism of ‘The relationship of lens aberrations to image aberrations in electron optics’” by P.W. Hawkes, *Optik*, 111(5), 229–234.
- Rempfer, G.F. and Mauck, M.S. 1985. Aberration-correcting properties of the hyperbolic electron mirror. In *Proc. 43rd Annu. Meeting EMSA*, 132–133.
- Rempfer, G.F. and Mauck, M.S. 1986. An experimental study of the hyperbolic electron mirror. In *Proc. 44th Annu. Meeting EMSA*, 886–887.
- Rempfer, G.F. and Mauck, M.S. 1992. Correction of chromatic aberration with an electron mirror. *Optik*, 92(1), 3–8.
- Rempfer, G.F., Desloge, D.M., Skoczylas, W.P. and Griffith, O.H. 1997. Simultaneous correction of spherical and chromatic aberrations with an electron mirror. *Microsc. Microanal.*, 3, 14–27.
- Ren, Y., Tang, T.-t., Kang, Y.-f. and Guo, X.-l. 2007. The aberration theory of a combined electron focusing-deflection system with a rotating deflection field following the rotation of the electrons. *Optik*, 118(12), 569–574.
- Renau, A. and Heddle, D.W.O. 1986. “Geometric aberrations in electrostatic lenses: II”. Aberration coefficients and their interrelations. *J. Phys. E: Sci. Instrum.*, 19(4), 288–295.
- Renau, A. and Heddle, D.W.O. 1987. Reply to comment on “Geometric aberrations in electrostatic lenses: II”. *J. Phys. E: Sci. Instrum.*, 20(2), 235–236.
- Riecke, W.D. 1982. Practical lens design. In *Magnetic Electron Lenses*, Hawkes, P.W. Ed., 163–157 (Springer, Berlin; Topics in Current Physics, Vol. 18).
- Ringnalda, J., Freitag, B., Hubert, D. and Stekelenburg, M. 2007. Practical electron microscopy with correctors and monochromators. *Microsc. Microanal.*, 13, Suppl. 2, 860–861 (CD).
- Robinson, A. 1961. Nonstandard analysis. *Proc. K. Ned. Akad. Wetensch.*, A, 46, 432–440, also *Indagationes Mathematicae*, 23, 432–440.
- Robinson, A. 1966. *Nonstandard Analysis* (North-Holland, Amsterdam).
- Rose, H. 1967. Über die Korrigierbarkeit von Linsen für schnelle Elektronen. *Optik*, 26(3), 289–298.
- Rose, H. 1968. Über die Berechnung der Bildfehler elektronenoptischer Systeme mit gerader Achse. *Optik*, 27(7), 466–474 and (8)497–514.
- Rose, H. 1970. Berechnung eines elektronenoptischen Apochromaten. *Optik*, 32(2), 144–164.
- Rose, H. 1971a. Abbildungseigenschaften sphärisch korrigierter elektronenoptischer Achromate. *Optik*, 33(1), 1–24.

- Rose, H. 1971b. Elektronenoptische Aplanate. *Optik*, 34(3), 285–311.
- Rose, H. 1978. Aberration correction of homogeneous magnetic deflection systems. *Optik*, 51(1), 15–38.
- Rose, H. 1981. Correction of aperture aberrations in magnetic systems with threefold symmetry. *Nucl. Instrum. Meth. Phys. Res.*, 187(1), 187–199.
- Rose, H. 1987a. Hamiltonian magnetic optics. *Nucl. Instrum. Meth. Phys. Res. A*, 258(3), 374–401.
- Rose, H. 1987b. The retarding Wien filter as a high-performance imaging filter. *Optik*, 77(1), 26–34.
- Rose, H. 1990a. Inhomogeneous Wien filter as a corrector compensating for the chromatic and spherical aberration of low-voltage electron microscopes. *Optik*, 84(3), 91–107.
- Rose, H. 1990b. Outline of a spherically corrected semiplanatic medium-voltage transmission electron microscope. *Optik*, 85(1), 19–24.
- Rose, H. 1994. Correction of aberrations, a promising means for improving the spatial and energy resolution of energy-filtering electron microscopes. *Ultramicroscopy*, 56(1–3), 11–25.
- Rose, H. 2000. High-performance electron microscopes of the future. In *Proc. 12th Eur. Cong. on Electron Microscopy*, Brno, Tománek, P. and Kolařík, R. Eds., Vol. III, I.21–I.24 (Czechoslovak Society for Electron Microscopy, Brno).
- Rose, H. 2002a. Theory of electron-optical achromats and apochromats. *Ultramicroscopy*, 93(3–4), 293–303.
- Rose, H. 2002b. Correction of aberrations—past, present and future. *Microsc. Microanal.*, 8, Suppl. 2, 6–7.
- Rose, H. 2003a. Advances in electron optics. In *High-Resolution Imaging and Spectrometry of Materials*, Ernst, F. and Rühle, M. Eds., 189–270 (Springer, Berlin and New York; Springer Series in Materials Science, Vol. 50).
- Rose, H. 2003b. Outline of an ultracorrector compensating for all primary chromatic and geometrical aberrations of charged-particle lenses. *Microsc. Microanal.*, 9, Suppl. 3, 32–33.
- Rose, H. 2004. Outline of an ultracorrector compensating for all primary chromatic and geometrical aberrations of charged-particle lenses. *Nucl. Instrum. Meth. Phys. Res. A*, 519(1–2), 12–27.
- Rose, H. 2005. Prospects for aberration-free microscopy. *Ultramicroscopy*, 103(1), 1–6.
- Rose, H. 2006. Principles and prospects of aberration correction. In *Proc. 16th Int. Microscopy Cong.*, Sapporo, Ichinose, H. and Sasaki, T. Eds., Vol. 2, PL-3 (IMC-16 Publishing Committee, Sapporo).
- Rose, H. 2007. Criteria and prospects for realizing optimum electron microscopes. *Microsc. Microanal.*, 13, Suppl. 2, 134–135.
- Rose, H. 2008a. *Geometrical Electron Optics* (Springer, Berlin).
- Rose, H. 2008b. History of direct aberration correction. *Adv. Imaging & Electron Phys.* 153.
- Rose, H. and Krahl, D. 1995. Electron optics of imaging energy filters. In *Energy-Filtering Transmission Electron Microscopy*, Reimer, L. Ed., 43–149 (Springer, Berlin).
- Rose, H. and Moses, R.W. 1973. Minimaler Öffnungsfehler magnetischer Rund- und Zylinderlinsen bei feldfreiem Objektraum. *Optik*, 37(3), 316–336.
- Rose, H. and Pejas, W. 1979. Optimization of imaging magnetic filters free of second-order aberrations. *Optik*, 54(3), 235–250.
- Rose, H. and Petri, U. 1971. Zur systematischen Berechnung elektronenoptischer Bildfehler. *Optik*, 33(2), 151–165.
- Rose, H. and Preikszas, D. 1992. Outline of a versatile corrected LEEM. *Optik*, 92(1), 31–44.
- Rose, H. and Preikszas, D. 1995. Time-dependent perturbation formalism for calculating the aberrations of systems with large ray gradients. *Nucl. Instrum. Meth. Phys. Res., A*, 363(1–2), 301–315.
- Rose, H., Haider, M. and Urban, K. 1998. Elektronenmikroskopie mit atomarer Auflösung. *Phys. Blätter.*, 54(5), 411–416.
- Rose, H., Hartel, P. and Preikszas, D. 2004. Outline of the mirror corrector for SMART and PEEM3. In *Microsc. Microanal.*, 10, Suppl. 3, 28–29.
- Rouse, J., Munro, E. and Tsuno, K. 1997. Computer analysis of imaging energy filters. *Proc. SPIE*, 3155, 193–204.
- Sakurai, T., Toyoda, M., Hayashibara, H. and Matsuo, T. 1995. Third-order trajectory calculations of Wien filters. *Int. J. Mass Spectrom. Ion Proc.*, 146/147, 217–222.
- Sang, G.-h. and Mu, C.-j. 2005. The computer-aided analysis of ninth-order geometrical aberrations of magnetic lens. *J. Yunnan Univ. Nat. Sci.*, 27(1), 40–43.
- Sawada, H., Tomita, T., Kaneyama, T., Hosokawa, F., Naruse, M., Honda, T., Hartel, P., Haider, M., Tanaka, N., Hetherington, C.J.D., Doole, R.C., Kirkland, A.I., Hutchison, J.L., Titchmarsh, J.M. and Cockayne, D.J.H. 2004a. C<sub>s</sub> corrector for imaging. *Microsc. Microanal.*, 10, Suppl. 2, 978–979 (CD).

- Sawada, H., Tomita, T., Naruse, M., Honda, T., Hartel, P., Haider, M., Hetherington, C.J.D., Doole, R.C., Kirkland, A.I., Hutchison, J.L., Titchmarsh, J.M. and Cockayne, D.J.H. 2004b.  $C_s$  corrector for illumination. *Microsc. Microanal.*, 10, Suppl. 2, 1004–1005 (CD).
- Sawada, H., Tomita, T., Naruse, M., Honda, T., Hartel, P., Haider, M., Hetherington, C.J.D., Doole, R.C., Kirkland, A.I., Hutchison, J.L., Titchmarsh, J.M. and Cockayne, D.J.H. 2004c. 200 kV TEM with  $C_s$  correctors for illumination and imaging. In *Proc. 8th Asia-Pacific Conf. on Electron Microscopy*, Kanazawa, Tanaka, N., Takano, Y., Mori, H., Seguchi, H., Iseki, S., Shimada, H. and Simamura, E. Eds., 20–21 CD (8APEM Publication Committee, Uchinada, Ishikawa).
- Sawada, H., Hosokawa, F., Sannomiya, T., Kaneyama, T., Tomita, T., Kondo, Y., Tanaka, T., Oshima, Y., Tanishiro, Y. and Takayanagi, K. 2007. Development of spherical aberration corrected 300 kV FETEM. *Microsc. Microanal.*, 13, Suppl. 2, 880–881 (CD).
- Saxton, W.O. 1994. Tilt-shift analysis for TEM auto-adjustment: a better solution to the data-fitting problem. *J. Comput.-Assist. Microsc.*, 6(2), 61–76.
- Saxton, W.O. 1995. Observation of lens aberrations for very high resolution electron microscopy. I. Theory. *J. Microsc.*, 179(2), 201–213.
- Saxton, W.O., Chand, G. and Kirkland, A.I. 1994. Accurate determination and compensation of lens aberrations in high resolution EM. In *Proc. 13th Int. Conf. Electron Microscopy*, Paris, Vol. 1, 203–204. (Editions de Physique, Les Ulis).
- Scheinfein, M.R. 1989. Second order transfer matrices for inhomogeneous field Wien filters including spin-precession. *Optik*, 82(3), 99–113.
- Scherzer, O. 1936. Über einige Fehler von Elektronenlinsen. *Z. Phys.*, 101(9–10), 593–603.
- Scherzer, O. 1946. Zur Korrigierbarkeit von Elektronenlinsen. *Phys. Blätter*, 2, 110.
- Scherzer, O. 1947. Sphärische und chromatische Korrektur von Elektronen-Linsen. *Optik*, 2, 114–132.
- Scherzer, O. 1948. Ein elektronenoptischer Apochromat. *Z. Naturforsch.*, 3a, 544–545.
- Scherzer, O. 1949. The theoretical resolution limit of the electron microscope. *J. Appl. Phys.*, 20(1), 20–29.
- Scherzer, O. 1980. Eine sphärisch korrigierte Folien-Linse für die Phasenmikroskopie mit Elektronen. *Optik*, 56(2), 133–147.
- Schmid, P., Feng, J., Padmore, H., Robin, D., Rose, H., Schlueter, R. and Wan, W. 2005. Correction and alignment strategies for the beam separator of the photoemission electron microscope 3 (PEEM3). *Rev. Sci. Instrum.*, 76, 023302:1–14.
- Schmidt, Th., Groh, U., Fink, R., Umbach, E., Schaff, O., Engel, W., Richter, R., Kuhlenbeck, H., Schlögl, R., Freund, H.-J., Bradshaw, A.M., Preikszas, D., Hartel, P., Spehr, R., Rose, H., Lilienkamp, G., Bauer, E. and Benner, G. 2002. XPEEM with energy-filtering: advantages and first results from the SMART project. *Surface Rev. Lett.*, 9(1), 223–232.
- Schmidt, Th., Marchetto, H., Groh, U., Lévesque, P.L., Maier, F.C., Skála, T., Kuhlenbeck, H., Hartel, P., Spehr, R., Preikszas, D., Lilienkamp, G., Bauer, E., Fink, R., Rose, H., Freund, H.-J. and Umbach, E. 2007. SMART: spectro-microscopy with aberration correction for many relevant techniques. Paper presented at LEEM-PEEM V, to appear in *Surface Sci.*
- Schönhense, G. and Spiecker, H. 2002a. Correction of chromatic and spherical aberration in electron microscopy utilizing the time structure of pulsed electron sources. *J. Vac. Sci. Technol.*, B, 20(6), 2526–2534.
- Schönhense, G. and Spiecker, H. 2002b. Chromatic and spherical aberration correction using time-dependent acceleration and lens fields. In *Proc. 8th Seminar on Recent Trends in Charged Particle Optics and Surface Physics Instrumentation*, Skalský Dvůr, Frank, L. Ed., 71–73 (ISI, Brno).
- Schönhense, G. and Spiecker, H. 2003. Chromatic and spherical correction using time-dependent acceleration and lens fields. *Microsc. Microanal.*, 9, Suppl. 3, 34–35.
- Schönhense, G., Elmers, H.J., Nepijko, S.A. and Schneider, C.M. 2006. Time-resolved photoemission electron microscopy. *Adv. Imaging & Electron Phys.*, 142, 160–323.
- Schriber, S.O. and Taylor, L.S. Eds., 1987. Proc. of the 2nd Int. Conf. on Charged Particle Optics, Albuquerque. *Nucl. Instrum. Meth. Phys. Res.*, A, 258(3), 289–598.
- Schröder, R.R., Barton, B., Schultheiß, K., Gamm, B. and Gerthsen, D. 2008. In-focus phase contrast: present state and future developments. *Proc. European Microscopy Congress*, Aachen, Luysberg, M., Tillmann, K. and Weirich, T. Eds., Instrumentation and Methods volume (Springer, Berlin).
- Schultheiß, K., Pérez-Willard, F., Barton, B. and Gerthsen, D. 2006. Fabrication of a Boersch phase plate for phase contrast imaging in a transmission electron microscope. *Rev. Sci. Instrum.*, 77(3), 033701.
- Sebastiano, A. di. and Pozzi, G. 1996. A multislice approach to the spherical aberration theory in electron optics. *J. Mod. Opt.*, 43(11), 2367–2379.
- Sekunova, L.M. 1977. [Paraxial properties and aberrations of a two-electrode axisymmetric mirror.] *Zh. Tekh. Fiz.*, 47(10), 2030–2042; *Sov. Phys. Tech. Phys.*, 22(10), 1181–1187.

- Sekunova, L.M. and Yakushev, E.M. 1975. [Electron optical properties of a mirror in the shadow mode of a mirror electron microscope.] I, II. *Zh. Tekh. Fiz.*, 45(4), 723–731 and 732–740; *Sov. Phys. Tech. Phys.*, 20(4), 458–462 and 463–467.
- Septier, A. 1963. Lentille quadrupolaire magnéto-électrique corrigée de l'aberration chromatique. Aberration d'ouverture de ce type de lentilles. *C. R. Acad. Sci. Paris*, 256, 2325–2328.
- Septier, A. 1966. The struggle to overcome spherical aberration in electron optics. *Adv. Opt. Electron Microsc.*, 1, 204–274.
- Shao, Z. 1988a. Correction of spherical aberration in the transmission electron microscope. *Optik*, 80(2), 61–75.
- Shao, Z. 1988b. On the fifth order aberration in a sextupole corrected probe forming system. *Rev. Sci. Instrum.*, 59(11), 2429–2437.
- Shao, Z. and Crewe, A. V. 1987. Spherical aberration of multipoles. *J. Appl. Phys.*, 62(4), 1149–1153.
- Shao, Z. and Wu, X.D. 1989. Adjustable four-electrode electron mirror as an aberration corrector. *Appl. Phys. Lett.*, 55(26), 2696–2697.
- Shao, Z. and Wu, X.D. 1990a. Properties of a four-electrode adjustable electron mirror as an aberration corrector. *Rev. Sci. Instrum.*, 61(4), 1230–1235.
- Shao, Z. and Wu, X.D. 1990b. A study on hyperbolic mirrors as correctors. *Optik*, 84(2), 51–54.
- Shao, Z. and Ximen, J.-y. 1992. Canonical aberration theory of multipoles: a critical reanalysis. *J. Appl. Phys.*, 71(4), 1588–1593.
- Shao, Z., Beck, V. and Crewe, A.V. 1988. A study of octupoles as correctors. *J. Appl. Phys.*, 64(4), 1646–1651.
- Shimoyama, H. 1982. The dependence of the asymptotic aberration coefficients of electron lenses on the magnification. *J. Electron Microsc.*, 31(1), 8–17.
- Słowko, W. 1972. Parametry elektronooptyczne zwierciadła dwururowego; Ograniczenia zakresu kompensacji aberracji sferycznej w dwusoczewkowym układzie odwzorowującym. *Prace Nauk. Inst. Technol. Elektron. Politech. Wrocław*, 6, 89–113 and 10, 71–73.
- Słowko, W. 1975. On compensation of spherical aberration in electron lenses by using an electron mirror, *Opt. Appl.*, 5(2), 27–32.
- Smith, M.R. and Munro, E. 1986. A unified aberration theory for electrostatic round lenses, multipole lenses and deflectors. *Optik*, 74(1), 7–16.
- Stam, M.A.J. van der. and Kruit, P. 1995. Computer assistance for the pre-design of optical instruments. *Nucl. Instrum. Meth. Phys. Res.*, A, 363(1–2), 205–210.
- Stam, M.A.J. van der. and Kruit, P. 1999. A general framework for numerical calculation of properties of charged-particle optical systems. *Nucl. Instrum. Meth. Phys. Res.*, A, 427(1–2), 368–374.
- Stam, M.A.J. van der, Barth, J.E. and Kruit, P. 1993. Design of a multi-mode transport lens using optimization program SOEM. *Proc. SPIE*, 2014, 45–46.
- Steffen, T., Tiemeijer, P.C., Krijn, M.P.C.M. and Mentink, S.A.M. 2000. Correction of chromatic and spherical aberration using a Wien filter. In *Proc. 12th Eur. Cong. on Electron Microscopy*, Brno, Tománek, P. and Kolařík, R. Eds., Vol. III, I.151–I.152 (Czechoslovak Society for Electron Microscopy, Brno).
- Stojanow, P.A. 1958. Stigmator für Mehrlinsen-Mikroskop. In *Proc. 4th Int. Cong. Electron Microscopy*, Berlin, Vol. 1, 61–66.
- Stoyanov, P.A. 1955a. [The effect of departures of the geometrical shape of the polepieces of an objective from circular symmetry on the resolving power of the electron microscope.] *Zh. Tekh. Fiz.*, 25, 625–635.
- Stoyanov, P.A. 1955b. [Astigmatism for the electron microscope]. *Zh. Tekh. Fiz.*, 25, 2537–2541.
- Sturrock, P.A. 1949. The aberrations of magnetic electron lenses due to asymmetries. In *Proc. Conf. Electron Microscopy*, Delft, 89–93.
- Sturrock, P.A. 1951a. Perturbation characteristic functions and their application to electron optics. *Proc. R. Soc. London*, A, 210(1101), 269–289.
- Sturrock, P.A. 1951b. The aberrations of magnetic electron lenses due to asymmetries. *Philos. Trans. R. Soc. London*, A, 243(868), 387–429.
- Sturrock, P.A. 1952. The imaging properties of electron beams in arbitrary static electromagnetic fields. *Philos. Trans. R. Soc. London*, A, 245(894), 155–187.
- Sturrock, P.A. 1955. *Static and Dynamic Electron Optics* (Cambridge University Press, Cambridge).
- Sugiura, Y. and Suzuki, S. 1943. On the magnetic electron lens of minimum spherical aberration. *Proc. Imp. Acad. Jpn*, 19, 293–302.
- Szép, J. and Szilagyi, M. 1988. A novel approach to the synthesis of electrostatic lenses with minimized aberrations. *IEEE Trans.*, ED-35(7), 1181–1183.
- Szép, J. and Szilagyi, M. 1990. Improved automated lens design for LMI sources. *J. Vac. Sci. Technol.*, A, 8(5), 3758–3762.

- Szilagyi, M. 1988. *Electron and Ion Optics*, (Plenum, New York).
- Takaoka, A. 1995. Theory of cathode lenses for probe beams. *Nucl. Instrum. Meth. Phys. Res.*, A, 363(1–2), 243–248.
- Takaoka, A. and Ura, K. 1984. Paraxial rays and aberrations in field emission gun. *Optik*, 68(2), 107–119.
- Takaoka, A. and Ura, K. 1989. Application of Liouville's theorem to accuracy estimation of paraxial and aberration theory. *Optik*, 83(3), 101–103.
- Takaoka, A. and Ura, K. 1990. Trajectory theory of cathode lens with weak surface electric field. *Optik*, 84(5), 176–182.
- Tang, T.-t. 1986a. Asymptotic aberrations of combined electrostatic focusing-deflection systems. *Optik*, 74(2), 43–47.
- Tang, T.-t. 1986b. Aberration analysis of a crossed-field analyzer. *Optik*, 74(2), 51–56.
- Tang, T.-t., Orloff, J. and Wang, L. 1996a. Modeling and design of space charge lenses/aberration correctors for focused ion beam systems. *J. Vac. Sci. Technol.*, B, 14(1), 80–84.
- Tang, T.-t., Orloff, J. and Wang, L. 1996b. Monte-Carlo simulation on discrete space charge effects of space charge lenses. *Optik*, 102(4), 141–146.
- Taylor, J.H. and Smith, K.C.A. 1986. The application of the complex optimization technique to the design of magnetic electron lenses. In *Proc. 11th Int. Cong. Electron Microscopy*, Kyoto Vol. I, 287–298.
- Thust, A., Barthel, J., Houben, L., Jia, C.L., Lentzen, M., Tillmann, K. and Urban, K. 2005. Strategies for aberration control in sub-ångström HRTEM. *Microsc. Microanal.*, 11, Suppl. 2, 58–59.
- Tiemeijer, P.C. 1999a. Operation modes of a TEM monochromator. In *Proc. EMAG 1999*, Kiely, C.J. Ed., 191–194 (Institute of Physics Conf. Ser. 161).
- Tiemeijer, P.C. 1999b. Measurement of Coulomb interactions in an electron beam monochromator. *Ultramicroscopy*, 78(1–4), 53–62.
- Tiemeijer, P.C., Lin, J.H.A. van, Freitag, B.H. and Jong, A.F. de 2002. Monochromized 200 kV (S)TEM. *Microsc. Microanal.*, 8, Suppl. 2, 70–71.
- Tillmann, K., Thust, A. and Urban, K. 2004. Spherical aberration correction in tandem with exit-plane function reconstruction: interlocking tools for the atomic scale imaging of lattice defects in GaAs. *Microsc. Microanal.*, 10(2), 185–198.
- Titchmarsh, J.M., Cockayne, D.J.H., Doole, R.C., Hetherington, C.J.D., Hutchison, J.L., Kirkland, A.I. and Sawada, H. 2004. A versatile double aberration-corrected, energy-filtered HREM/STEM for materials science. In *Proc. 13th Eur. Microscopy Cong.*, Antwerp, Dyck, D. van and Oostveldt, P. van Eds., Vol. I, 27–28 (Belgian Society for Microscopy, Liège).
- Tong, L., Wu, C. and Luo, F. 1987. The influence of misalignment and/or misorientation of electron lenses on image quality. In *Proc. Int. Symp. Electron Optics*, Beijing, 1986, Ximen, J.-y. Ed., 249–251 (Institute of Electronics, Academia Sinica, Beijing).
- Tonomura, A. 1994. Recent advances in electron-holographic interference microscopy. In *Proc. 13th Int. Conf. Electron Microscopy*, Paris, Vol. 1, 281–284. (Editions de Physique, Les Ulis).
- Tonomura, A. 1998. *The Quantum World Unveiled by Electron Waves* (World Scientific, Singapore).
- Tonomura, A. 1999. *Electron Holography* (Springer, Berlin; Springer Series in Optical Sciences, Vol. 70).
- Tonomura, A. 2006. Electron holography. In *Proc. 16th Int. Microscopy Cong.*, Sapporo, Ichinose, H. and Sasaki, T. Eds., Vol. 2, 718 (IMC-16 Publishing Committee, Sapporo).
- Tonomura, A., Allard, L.F., Pozzi, G., Joy, D.C. and Ono, Y.A. Eds., 1995. *Electron Holography* (North-Holland, Amsterdam).
- Toyoda, M., Hayashibara, H., Sakurai, T. and Matsuo, T. 1995. Third-order transfer matrix for crossed electric and magnetic fields calculated algebraically using a symbolic computation program. *Int. J. Mass Spectrom. Ion Proc.*, 146/147, 195–216.
- Tretner, W. 1959. Existenzbereiche rotationssymmetrischer Elektronenlinsen. *Optik*, 16(3), 155–184.
- Tromp, R.M. 2007. A new design for an energy-filtered, aberration-corrected LEEM/PEEM instrument. Paper presented at LEEM-PEEM V, to appear in *Surface Sci.*
- Tsai, F.C. 2000. Studies of a magnetically focused electrostatic mirror. II. Aberration corrections. *J. Microscopy*, 197(2), 118–135 [for Part I, see Crewe et al., 2000].
- Tsumagari, T., Murakami, J., Ohiwa, H. and Noda, T. 1986. Aberrations of electrostatic systems with machining error. *J. Vac. Sci. Technol.*, B, 4(1), 140–142.
- Tsumagari, T., Murakami, J., Ohiwa, H. and Noda, T. 1987. Analysis of mechanical aberrations in the electrostatic deflector, *Japan. J. Appl. Phys.*, 26(10), 1772–1776.
- Tsuno, K. 1991. Aberration analysis of a Wien filter for electrons, *Optik*, 89(1), 31–40.
- Tsuno, K. 1992. Electron optical analysis of a high-resolution electron energy loss spectrometer with a retarding Wien filter. *Rev. Sci. Instrum.*, 63(9), 4112–4121.

- Tsuno, K. 1993. Electron optical analysis of a retarding Wien filter for electron spectroscopic imaging. *Rev. Sci. Instrum.*, 64(3), 659–666.
- Tsuno, K. 1997. Design of an omega filter for a 200 kV electron microscope. *Rev. Sci. Instrum.*, 68(1), 109–115.
- Tsuno, K. 1999. Optical design of electron microscope lenses and energy filters. *J. Electron Microsc.*, 48(6), 801–820.
- Tsuno, K. 2001. S filter, a compact, high dispersion in-column energy filter. *Proc. SPIE*, 4510, 71–79.
- Tsuno, K. 2004. Evaluation of in-column energy filters for analytical electron microscopes. *Nucl. Instrum. Meth. Phys. Res., A*, 519(1–2), 286–296.
- Tsuno, K. and Harada, Y. 1981a. Minimisation of radial and spiral distortion in electron microscopy through the use of a triple pole-piece lens. *J. Phys. E: Sci. Instrum.*, 14(3), 313–320.
- Tsuno, K. and Harada, Y. 1981b. Elimination of spiral distortion in electron microscopy using an asymmetrical triple pole-piece lens. *J. Phys. E: Sci. Instrum.*, 14(8), 955–960.
- Tsuno, K. and Rouse, J. 1996. Simulation of electron trajectories of Wien filters for high-resolution EELS installed in TEM. *J. Electron Microsc.*, 45(5), 417–427.
- Tsuno, K. and Smith, K.C.A. 1985. Computation of lens shape with minimum spherical aberration coefficient. In *Proc. EMAG 85*, Tatlock, G.J. Ed., 121–124 (Institute of Physics Conf. Ser. 78).
- Tsuno, K. and Smith, K.C.A. 1986. Optimization of objective lens shape for high resolution electron microscopes. In *Proc. 11th Int. Cong. Electron Microscopy*, Kyoto, Vol. I, 295–296.
- Tsuno, K., Terauchi, M. and Tanaka, M. 1988–9. Electron trajectory calculation of a stigmatic-focus Wien filter for electron energy loss spectroscopy I, II, III. *Optik*, 78(2), 71–77; 80(4), 149–154; 83(3), 77–84.
- Tsuno, K., Terauchi, M. and Tanaka, M. 1990. Design and performance of a retarding Wien filter installed in a TEM for EELS. In *Proc. EMAG–MICRO 89*, Goodhew, P.J. and Elder, H.Y. Eds., Vol. 1, 71–74 (Institute of Physics Conf. Ser. 98).
- Tsuno, K., Munro, E. and Rouse, J. 1995. A beam separator of a close packed prism array for low energy electron microscopy. *Nucl. Instrum. Meth. Phys. Res., A*, 363(1–2), 276–283.
- Tsuno, K., Kaneyama, T., Honda, T. and Ishida, Y. 1999. Design of omega-mode imaging energy filters. *Nucl. Instrum. Meth. Phys. Res., A*, 427(1–2), 187–196.
- Tsuno, K., Ioanoviciu, D. and Martínez, G. 2003. Aberration corrected Wien filter as a monochromator of high spatial and high energy resolution electron microscopes. *Microsc. Microanal.*, 9, Suppl. 2, 944–945 (CD).
- Tsuno, K., Ioanoviciu, D. and Martínez, G. 2005. Third-order aberration theory of Wien filters for monochromators and aberration correctors. *J. Microscopy*, 217(3), 205–215.
- Typke, D. 1968/9. Der Öffnungsfehler von rotationssymmetrischen Elektronenlinsen mit Raumladung. *Optik*, 28(5), 488–494.
- Typke, D. 1972. Sphärische und chromatische Korrektur von Elektronenlinsen mit Hilfe von Influenzladungen, I, II. *Optik*, 34(5), 573–594 and 36(1), 124–138; erratum 44(4), 1976, 509.
- Uhlemann, S. and Haider, M. 1998. Residual wave aberrations in the first spherical aberration corrected transmission electron microscope. *Ultramicroscopy*, 72(3–4), 109–119.
- Uhlemann, S., Haider, M. and Rose, H. 1994. Procedures for adjusting and controlling the alignment of a spherically corrected electron microscope. In *Proc. 13th Int. Conf. Electron Microscopy*, Paris, Vol. 1, 193–194 (Editions de Physique, Les Ulis).
- Uhlemann, S., Haider, M., Schwan, E. and Rose, H. 1998. Towards a resolution enhancement in the corrected TEM. In *Proc. EUREM-11, the 11th Eur. Conf. on Electron Microscopy*, Dublin 1996, CESEM, Ed., Vol. I, 365–366 (CESEM, Brussels).
- Uno, Y., Morita, H., Simazu, N. and Hosokawa, T. 1995. Fifth-order aberration analysis of a combined electrostatic–magnetic focusing–deflection system. *Nucl. Instrum. Meth. Phys. Res., A*, 363(1–2), 10–17.
- Uno, S., Honda, K., Nakamura, N., Matsuya, M., Achard, B. and Zach, J. 2004a. An automated aberration correction method in scanning electron microscopes. In *Proc. 13th Eur. Microscopy Cong.*, Antwerp, Dyck, D. van and Oostveldt, P. van Eds., Vol. I, 37–38 (Belgian Society for Microscopy, Liège).
- Uno, S., Honda, K., Nakamura, N., Matsuya, M. and Zach, J. 2004b. Probe shape extraction and automatic aberration correction in scanning electron microscopes. In *Proc. 8th Asia-Pacific Conf. on Electron Microscopy*, Kanazawa, Tanaka, N., Takano, Y., Mori, H., Seguchi, H., Iseki, S., Shimada, H. and Simamura, E. Eds., 46–47 CD (8APEM Publication Committee, Uchinada, Ishikawa).
- Uno, S., Honda, K., Nakamura, N., Matsuya, M. and Zach, J. 2005. Aberration correction and its automatic control in scanning electron microscopes. *Optik*, 116(9), 438–448.
- Ura, K. 1994. Magnification dependence of aberration coefficients in electrostatic bipotential lens. *J. Electron Microsc.*, 43(6), 406–407.

- Ura, K., Hibino, M., Komuro, M., Kurashige, M., Kurokawa, S., Matsuo, T., Okayama, S., Shimoyama, H. and Tsuno, K. Eds., 1995. Proc. 4th Int. Conf. on Charged Particle Optics (CPO4), Tsukuba, *Nucl. Instrum. Meth. Phys. Res., A*, 363(1–2), 1–496.
- Urban, K. 2008. Studying atomic structures by aberration-corrected transmission electron microscopy. *Science*, 321, 506–510.
- Urban, K., Kabius, B., Haider, M. and Rose, H. 1999. A way to higher resolution: spherical-aberration correction in a 200 kV transmission electron microscope. *J. Electron Microsc.*, 48(6), 821–826.
- Vaidya, N.C. 1972. Synklysmotron lenses—a new electron-optical correcting system. *Proc. IEEE*, 60(2), 245–247.
- Vaidya, N.C. 1975a. The design and properties of synklysmotron lenses. *J. Phys. D: Appl. Phys.*, 8(4), 368–373.
- Vaidya, N.C. 1975b. Correction of spherical aberration by a new system of high frequency lenses (synklysmotron lenses). *Optik*, 42(2), 129–134.
- Vaidya, N.S. and Garg, R.K. 1974. Correction of spherical aberration in electron lenses by time-varying fields. In *Proc. 8th Int. Cong. Electron Microscopy*, Canberra, I, 150–151.
- Vaidya, N.C. and Hawkes, P.W. 1970. Correction of spherical aberration by means of microwave lenses. In *Proc. 7th. Int. Cong. Electron Microscopy*, Grenoble, II, 19–20.
- Varela, M., Lupini, A.R., Benthem, K. van, Borisevich, A.Y., Chisholm, M.F., Shibata, N., Abe, E. and Pennycook, S.J. 2005. Materials characterization in the aberration-corrected scanning transmission electron microscope. *Annu. Rev. Materials Sci.*, 35, 539–569.
- Walther, T. and Stegmann, H. 2006a. Performance evaluation of a new monochromatic and aberration-corrected 200 kV field-emission scanning transmission electron microscope. In *Proc. 16th Int. Microscopy Cong.*, Sapporo, Ichinose, H. and Sasaki, T. Eds., Vol. 2, 607 (IMC-16 Publishing Committee, Sapporo).
- Walther, T. and Stegmann, H. 2006b. Preliminary results from the first monochromated and aberration-corrected 200 kV field-emission scanning transmission electron microscope. *Microsc. Microanal.*, 12(6), 498–505.
- Walther, T., Quandt, E., Stegmann, H., Thesen, A. and Benner, G. 2006. First experimental test of a new monochromated and aberration-corrected 200kV field-emission scanning transmission electron microscope. *Ultramicroscopy*, 106(11–12), 963–969.
- Wan, W., Feng, J., Padmore, H.A. and Robin, D.S. 2004. Simulation of a mirror corrector for PEEM. *Nucl. Instrum. Meth. Phys. Res. A*, 519, 222–229.
- Wan, W., Feng, J. and Padmore, H.A. 2006. A new separator design for aberration-corrected photoemission electron microscopes. *Nucl. Instrum. Meth. Phys. Res., A*, 564(1), 537–543.
- Wang, L., Orloff, J., Book, D. and Tang, T.-t. 1995a. A new method for calculating the axial potential due to space charge in electrostatic optics. *J. Phys. D: Appl. Phys.*, 28(9), 1791–1801.
- Wang, L., Orloff, J. and Tang, T.-t. 1995b. Study of space-charge devices for focused ion beam systems. *J. Vac. Sci. Technol. B*, 13(6), 2414–2418.
- Wang, L., Tang, T.-t. and Orloff, J. 1995c. Calculation of aberrations of a space charge lens and its possible applications. *Proc. SPIE*, 2522, 363–374.
- Wang, L.-p., Tang, T.-t. and Cai, J. 1999a. Study on the object oriented implementation of automatic differentiation technique. *Optik*, 110(10), 449–454.
- Wang, L.-p., Tang, T.-t., Cheng, B.-j. and Cai, J. 1999b. Automatic differentiation method for the aberration analysis of electron optical systems. *Optik*, 110(9), 408–414; also published in *J. Xi'an Jiaotong Univ.*, 33, 1999, 101–105.
- Wang, L.-p., Tang, T.-t. and Cheng, B.-j. 2000. Differential algebraic method for arbitrary-order aberration analysis of electron diffraction systems. *Optik*, 111(7), 290–296.
- Wang, L.-p., Tang, T.-t., Cheng, B.-j. and Cleaver, J.R.A. 2002a. Application of differential algebraic method to the aberration analysis of curvilinear-axis electron optical systems. *Optik*, 113(4), 171–176.
- Wang, L.-p., Tang, T.-t., Cheng, B.-j. and Cleaver, J.R.A. 2002b. Differential algebraic method for arbitrary-order aberration analysis of combined electron beam focusing–deflection systems. *Optik*, 113(4), 181–187.
- Wang, L.-p., Rouse, J., Liu, H., Munro, E. and Zhu, X. 2004. Simulation of electron optical systems by differential algebraic method combined with Hermite fitting for practical lens fields. *Microelectron. Eng.*, 73–74, 90–96.
- Wang, L.-p., Munro, E., Rouse, J. and Liu, H.-n. 2008a. Simulation of electron mirrors by the differential algebra method. *Phys. Procedia*.
- Wang, L.-p., Rouse, J., Munro, E., Liu, H.-n. and Zhu, X.-q. 2008b. Aberration analysis of electron mirrors by a differential algebraic method. *Optik*, 119(2), 90–96.

- Watanabe, M., Ackland, D.W., Kiely, C.J., Williams, D.B., Kanno, M., Hynes, R. and Sawada, H. 2006. Optimization of a spherical-aberration-corrected scanning transmission electron microscope for atomic-resolution annular dark-field imaging and electron energy-loss spectrometry. In *Proc. 16th Int. Microscopy Cong.*, Sapporo, Ichinose, H. and Sasaki, T. Eds., Vol. 2, 606 (IMC-16 Publishing Committee, Sapporo).
- Webster, N.E., Haider, M. and Houf, H. 1988. Design and construction of a multipole-element control unit. *Rev. Sci. Instrum.*, 59(6), 999–1001.
- Wei, L. and Yan, T. 1999. Determination of the acceptable tolerance in the manufacture of electron optical system. *Nucl. Instrum. Meth. Phys. Res., A*, 427(1–2), 393–398.
- Weimer, E. and Martin, J.-P. 1994. Development of a new ultra-high performance scanning electron microscope. In *Proc. 13th. Int. Conf. Electron Microscopy*, Paris, Vol. 1, 67–68 (Editions de Physique, Les Ulis).
- Weißbäcker, C. and Rose, H. 2000. Electrostatic correction of the chromatic and spherical aberration of charged particle lenses. In *Proc. 12th Eur. Cong. on Electron Microscopy*, Brno, Tománek, P. and Kolařík, R. Eds., Vol. III, I.157–I.158 (Czechoslovak Society for Electron Microscopy, Brno).
- Weißbäcker, C. and Rose, H. 2001. Electrostatic correction of the chromatic and the spherical aberration of charged-particle lenses. I. *J. Electron Microsc.*, 50(5), 383–390.
- Weißbäcker, C. and Rose, H. 2002. Electrostatic correction of the chromatic and the spherical aberration of charged-particle lenses, II. *J. Electron Microsc.*, 51(1), 45–51.
- Wichtendahl, R., Fink, R., Kuhlenbeck, H., Preikszas, D., Rose, H., Spehr, R., Hartel, P., Engel, W., Schlögl, R., Freund, H.-J., Bradshaw, A.M., Lilienkamp, G., Schmidt, T., Bauer, E., Benner, G. and Umbach, E. 1998. SMART: an aberration-corrected XPEEM/LEEM with energy filter. *Surface Rev. Lett.*, 5, 1249–1256.
- Wien, W. 1897/8, Die elektrostatischen Eigenschaften der Kathodenstrahlen. *Verhandl. Deutsch. Phys. Ges.*, 16, 165–172.
- Wien, W. 1898. Untersuchungen über die elektrische Entladung in verdünnten Gasen. *Ann. Physik*, 301(6), 440–452.
- Wollnik, H. 1967. Second order transfer matrices of real magnetic and electrostatic sector fields. *Nucl. Instrum. Methods*, 52(2), 250–272.
- Wollnik, H., Ed. 1981. Proc. First Conf. on Charged Particle Optics, Gießen, *Nucl. Instrum. Methods*, 187, 1–314.
- Wollnik, H. 1987. *Optics of Charged Particles* (Academic Press, Orlando, FL).
- Wollnik, H. and Berz, M. 1985. Relations between elements of transfer matrices due to the conditions of symplecticity, *Nucl. Instrum. Methods, A*, 238(1), 127–140.
- Wu, M.-zh. 1957. [Study of the fifth-order aberrations of rotationally symmetric electron optical systems]. *Scientia Sinica*, 6, 833–846 [in Russian, author's name transliterated as U M.-zh.]; *Acta Phys. Sin.*, 13(3), 181–206 [in Chinese].
- Wu, Y.K., Robin, D.S., Forest, E., Schleuter, R., Anders, S., Feng, J., Padmore, H. and Wei, D.H. 2004. Design and analysis of beam separator magnets for third generation aberration compensated PEEMs. *Nucl. Instrum. Meth. Phys. Res., A*, 519(1–2), 230–241.
- Ximen, J.-y. 1957. [On the electron optical properties and aberration theory of combined immersion objectives.] *Acta Phys. Sin.*, 13(4), 339–356.
- Ximen, J.-y. 1983. The aberration theory of a combined magnetic round lens and sextupoles system. *Optik*, 65(4), 295–309.
- Ximen, J.-y. 1986. Aberration theory in electron and ion optics. *Adv. Electron. Electron Phys.*, Suppl. 17.
- Ximen, J.-y. 1988. Aberration theory for combined electro-magnetic focusing-deflection systems with curvilinear axes; supplement. *Optik*, 80(2), 55–60 and (3), 98–102.
- Ximen, J.-y. 1990a. Canonical theory and matrix representations in electron optical systems possessing aberrations. *Optik*, 84(3), 83–90.
- Ximen, J.-y. 1990b. Canonical aberration theory in electron optics. *J. Appl. Phys.*, 68(12), 5963–5967.
- Ximen, J.-y. 1991a. Canonical aberration theory for calculating higher order chromatic aberration. *J. Appl. Phys.*, 69(4), 1962–1967.
- Ximen, J.-y. 1991b. Canonical theory in electron optics. *Adv. Electron. Electron Phys.*, 81, 231–277.
- Ximen, J.-y. 1995. Canonical aberration theory in electron optics up to ultra-high order approximation. *Adv. Imaging & Electron Phys.*, 91, 1–36.
- Ximen, J.-y. 1996. Ultra-high order canonical aberration calculation and integration transformation in rotationally symmetric magnetic and electrostatic lenses. *Adv. Imaging & Electron Phys.*, 97, 359–407.
- Ximen, J.-y. and Cheng, C.-s. 1964. [Analytical solution of the potential distribution and electron optics in magnetic lens with pole-pieces of elliptic distortion]. *Acta Electron. Sin.*, 4(12), 72–84.

- Ximen, J.-y. and Crewe, A.V. 1985. Correction of spherical and coma aberrations with a sextupole-round lens–sextupole system. *Optik*, 69(4), 141–146.
- Ximen, J.-y. and Li, D. 1990. Three-dimensional boundary-element method for analyzing nonrotationally symmetric perturbations in electrostatic lenses. *J. Appl. Phys.*, 67(4), 1643–1649.
- Ximen, J.-y. and Liu, Z. 1996. Analytical analysis and numerical calculation of ultra-high order aberrations in an ideal hyperbolic electrostatic lens. *Optik*, 102(1), 24–30.
- Ximen, J.-y. and Liu, Z. 1997. A theorem of higher-order spheric aberrations in Glaser's bell-shaped magnetic lens. *Optik*, 107(1), 17–25.
- Ximen, J.-y. and Liu, Z. 1998. Analytical analysis and numerical calculation of first- and third-order chromatic aberrations in Glaser's bell-shaped magnetic lens. *Optik*, 109(2), 68–76.
- Ximen, J.-y. and Liu, Z. 2000a. Analysis and calculation of third- and fifth-order aberrations in combined bell-shaped electromagnetic lens—a new theoretical model in electron optics. *Optik*, 111(2), 75–84.
- Ximen, J.-y. and Liu, Z. 2000b. Third-order geometric aberrations in Glaser's bell-shaped magnetic lens for object magnetic immersion. *Optik*, 111(8), 355–358.
- Ximen, J.-y. and Xi, Z.-x. 1964. [Theoretical study of field distribution and aberrations in a magnetic lens with pole pieces of inclined end-faces.] *Acta Electron. Sin.*, 3(9), 24–35.
- Ximen, J.-y., Zhou, L.-w. and Ai, K.-c. 1983. Variational theory of aberrations in cathode lenses. *Optik*, 66(1), 19–34; *Acta Phys. Sin.*, 32(12), 1536–1546.
- Ximen, J.-y., Liang, J. and Liu, Z. 1997. Analytical analysis and numerical calculation of third- and fifth-order aberrations in Glaser's bell-shaped magnetic lens. *Optik*, 105(4), 155–164.
- Yada, K. 1986. Effect of concave-lens action of a mini static lens on the spherical aberration coefficient of a magnetic lens. In *Proc. 11th Int. Conf. Electron Microscopy*, Kyoto, Vol. 1, 305–306.
- Yakushev, E.M. and Sekunova, L.M. 1986. Theory of electron mirrors and cathode lenses. *Adv. Electron. Electron Phys.*, 68, 337–416.
- Yavor, S.Ya. 1962. [Electron optical properties of combined electric and magnetic quadrupole lenses.] In *Proc. Symp. Electron Vacuum Physics*, Budapest, 125–137.
- Yavor, S.Ya. 1968. *Fokusirovka Zaryazhennykh Chastits Kvadrupol'nymi Linzami* (Atomizdat, Moscow).
- Yavor, M.I. 1993. Methods for calculation of parasitic aberrations and machining tolerances in electron optical systems. *Adv. Electron. Electron Phys.*, 86, 225–281.
- Yavor, S.Ya., Dymnikov, A.D. and Ovsyannikova, L.P. 1964. Achromatic quadrupole lenses. *Nucl. Instrum. Meth.*, 26(1), 13–17.
- Yavor, M.I. and Berdnikov, A.S. 1995. ISISOS: a program to calculate imperfect static charged particle optical systems. *Nucl. Instrum. Meth. Phys. Res.*, A, 363(1–2), 416–422.
- Yin, H.-c., Qin, S.-l. and Zhong, Z.-f. 2008. Software for simulation of electron optical systems. *Phys. Procedia*.
- Zach, J. 2000. Aspects of aberration correction in LVSEM. In *Proc. 12th Eur. Cong. on Electron Microscopy*, Brno, Tománek, P. and Kolařík, R. Eds., Vol. III, I.169–I.172 (Czechoslovak Society for Electron Microscopy, Brno).
- Zach, J. 2006. Aberration correction in SEM and FIB—the state of the art. In *Proc. 16th Int. Microscopy Cong.*, Sapporo, Ichinose, H. and Sasaki, T. Eds., Vol. 2, 662 (IMC-16 Publishing Committee, Sapporo).
- Zach, J. and Haider, M. 1994. Correction of spherical and chromatic aberrations in a LVSEM. In *Proc. 13th Int. Conf. Electron Microscopy*, Paris, Vol. 1, 199–200 (Editions de Physiques, Les Ulis).
- Zach, J. and Haider, M. 1995a. Aberration correction in a low voltage SEM by a multipole corrector. *Nucl. Instrum. Meth. Phys. Res.*, A, 363(1–2), 316–325.
- Zach, J. and Haider, M. 1995b. Correction of spherical and chromatic aberration in a low voltage SEM. *Optik*, 98(3), 112–118.
- Zeitler, E. 1990. Analysis of an imaging magnetic energy filter. *Nucl. Instrum. Meth. Phys. Res.*, A, 298 (1–3), 234–246.
- Zhang, T. and Cleaver, J.R.A. 1996. Computation of optical properties off the Gaussian image plane for electron-beam systems. *Optik*, 103(4), 151–166.
- Zhou, L.-w., Li, Y., Zhang, Z.-q., Monastyrski, M.A. and Schelev, M.Ya. 2005a. On the theory of temporal aberrations for cathode lenses. *Optik*, 116(4), 175–184.
- Zhou, L.-w., Li, Y., Zhang, Z.-q., Monastyrski, M.A. and Schelev, M.Ya. 2005b. [Theory of temporal aberrations for electron optical imaging systems by direct integral method.] *Acta Phys. Sin.*, 54(8), 3596–3602.
- Zhou, L.-w., Li, Y., Zhang, Z.-q., Monastyrski, M.A. and Schelev, M.Ya. 2005c. Test and verification of temporal aberration theory for electron optical imaging systems by an electrostatic concentric spherical system model. *Acta Phys. Sin.*, 54(8), 3603–3608.

- Zhou, L.-w., Monastyrski, M.A., Schelev, M.Ya., Zhang, Z.-q. and Li, Y. 2006. [On the temporal aberration theory of electron optical imaging systems by tau-variation method.] *Acta Electron. Sin.*, 34(2), 193–197.
- Zhu, X. and Liu, H. 1987. Numerical computation of the misalignment effect of lenses and deflectors in electron beam focusing and deflection systems. In *Proc. Int. Symp. Electron. Optics*, Beijing, 1986, Ximen, J.-y. Ed., 309–313 (Institute of Electronics, Academia Sinica, Beijing).
- Zhu, X.-q., Munro, E., Liu, H. and Rouse, J. 1999. Aberration compensation in charged particle projection lithography. *Nucl. Instrum. Meth. Phys. Res.*, A, 427(1–2), 292–298.
- Zhukov, V.A. and Zavyalova, A.V. 2006. Possibility and example of the axial aberration correction in the combined axially symmetric electromagnetic mirror. *Proc. SPIE*, 6278, 627808: 1–12.
- Zhukov, V.A., Berdnikov, Y.A. and Zhurkin, E. 2004. Extreme aberration properties of combined immersion lenses and their prospects in an ion nanolithography. *Proc. SPIE*, 5398, 63–70.
- Zworykin, V.K., Morton, G.A., Ramberg, E.G., Hillier, J. and Vance, A.W. 1945. *Electron Optics and the Electron Microscope* (Wiley, New York, and Chapman & Hall, London).

## A.1 APPENDIX: A BRIEF INTRODUCTION TO DIFFERENTIAL ALGEBRA

Here we can give no more than the flavor of the method. In the two centuries after the creation of calculus by Newton and Leibniz, two paths of development were pursued: the *method of limits* and the *infinitesimal method*. In the former, the derivative of a function is the *limit* of the ratio  $\{f(x + \Delta x) - f(x)\}/\Delta x$  as  $\Delta x \rightarrow 0$  whereas in the second method, the derivative is the *value* of this ratio for an infinitely small increase  $\Delta x$ . By the second half of the nineteenth century, thanks to Weierstrass, the method of limits had become dominant, but in 1960, Abraham Robinson showed that “the concepts and methods of contemporary Mathematical Logic are capable of providing a suitable framework for the development of the Differential and Integral Calculus by means of infinitely small and infinitely large numbers.... The resulting subject was called by me Non-standard Analysis, since it involves and was, in part, inspired by the so-called Non-standard models of Arithmetic whose existence was first pointed out by T. Skolem” (Robinson, 1966). Thus, in the words of Mahwin (1994), “Robinson achieves, for the infinitesimal method, what Weierstrass had achieved, a century earlier, for the method of limits.” A very thorough study of the evolution of this nonstandard analysis is to be found in J.W. Dauben’s biography of Robinson (1995) and in his biographical memoir (2003).

In the simplest of the differential algebras, ordered pairs of real numbers  $(a_0, a_1)$ , are considered, with the usual rules for addition and scalar multiplication:

$$(a_0, a_1) + (b_0, b_1) = (a_0 + b_0, a_1 + b_1)$$

$$t \cdot (a_0, a_1) = (t \cdot a_0, t \cdot a_1)$$

in which  $t$  is again a real number. A vector product is also defined:

$$(a_0, a_1) \cdot (b_0, b_1) = (a_0 \cdot b_0, a_1 b_0 + a_0 b_1)$$

The resulting algebra is denoted by  $_1D_1$ . Clearly,  $(a_0, 0)$  behaves just like  $a_0$  as  $(a_0, 0) \cdot (b_0, b_1) = (a_0 \cdot b_0, a_0 \cdot b_1) = a_0 \cdot (b_0, b_1)$ ; hence  $(1, 0)$  behaves like the usual unity 1 for multiplication. The number  $(0, 1)$  is less familiar, for

$$(0, 1) \cdot (a_0, a_1) = (0, a_0)$$

and so

$$(0, 1) \cdot (0, 1) = (0, 0)$$

$(0, 1)$  is thus in some sense the square root of zero and just as  $\sqrt{(-1)}$  is given the symbol  $i$ , so we reserve the symbol  $d$  for  $(0, 1)$  and refer to it as the *differential unit*.

Does  $(a_0, a_1)$  have a multiplicative inverse  $(b_0, b_1)$ , such that  $(a_0, a_1) \cdot (b_0, b_1) = (1, 0)$ ? The answer is that

$$(a_0, a_1) \cdot \left( \frac{1}{a_0}, -\frac{a_1}{a_0^2} \right) = (1, 0)$$

and so an inverse exists if and only if  $a_0 \neq 0$ . Hence the elements of  ${}_1D_1$  form a ring, not a field. The term  $a_0$  is called the *real part* of  $(a_0, a_1)$  and  $a_1$ , the *differential part*.

We now return to the expression for the derivative of a function. For an arbitrary small value of  $\Delta x$ , the real part must be independent of the choice of  $\Delta x$ . Let us choose  $\Delta x = d = (0, 1)$ . It can then be shown that

$$f'(x) = D[f(x + d)]$$

in which  $D[\dots]$  denotes the differential part of the argument. How does this work? Consider the simple function  $f(x) = x^n$  ( $n$  is an integer). Then

$$\begin{aligned} f(x + d) &= \{(x, 0) + (0, 1)\}^n \\ &= x^n + nx^{n-1}(0, 1) + \frac{n(n-1)}{2}x^{n-2}(0, 1)^2 + \dots \end{aligned}$$

But  $(0, 1)^2 = (0, 0)$  and so  $D[f(x + d)] = nx^{n-1}$  as expected ( $x^n$  has no differential part). As another example, consider  $f(x) = \sin x$ . Then

$$\begin{aligned} f(x + d) &= \sin\{(x, 0) + (0, 1)\} = \sin x \cos(0, 1) + \cos x \sin(0, 1) \\ \sin(0, 1) &= (0, 1) - \frac{1}{6}(0, 1)^3 + \dots = (0, 1) \\ \cos(0, 1) &= 1 - \frac{1}{2}(0, 1)^2 + \dots = 1 \end{aligned}$$

and so  $D[f(x + d)] = D[\sin x + \cos x(0, 1)] = \cos x$ , again as expected.

The real interest of these algebras resides in the more general algebras, denoted by  ${}_nD_v$ . Just as  ${}_1D_1$  enabled us to calculate the first derivative of a function of one variable by an ingenious substitution, these general algebras furnish the derivatives of a function of several variables up to a given order. For a function of two variables  $f(x, y)$ , for example, the algebra  ${}_2D_2$  gives

$$(\partial f / \partial x \quad \partial f / \partial y \quad \partial^2 f / \partial x^2 \quad \partial^2 f / \partial x \partial y \quad \partial^2 f / \partial y^2)$$

For further examples of the use of this tool in electron optics, see Wang et al. (1999a,b, 2000) and Cheng et al. (2001a,b, 2002a,b, 2003). A very helpful explanation of the procedure for obtaining aberration coefficients numerically is given by Radlická (2008).

---

# 7 Space Charge and Statistical Coulomb Effects

*Pieter Kruit and Guus H. Jansen*

## CONTENTS

7.1	Introduction.....	342
7.2	Analytical Approach to Space Charge Defocus and Aberrations.....	344
7.2.1	Laminar Flow and Space Charge Defocus.....	344
7.2.2	Equations for Space Charge Effects.....	344
7.2.2.1	The Ray Equation.....	344
7.2.2.2	Laminar Flow.....	345
7.2.2.3	Space Charge Defocus .....	346
7.2.2.4	Space Charge Aberrations .....	347
7.2.3	Numerical Examples .....	348
7.3	Analytical Approach to Statistical Coulomb Effects.....	349
7.3.1	General Formulation of the Problem.....	349
7.3.2	Reduction of the $N$ -Particle Problem.....	352
7.3.3	Two-Particle Dynamics .....	353
7.3.4	Overview of Approximate Solutions.....	354
7.3.4.1	Models Derived from Plasma Physics.....	354
7.3.4.2	First-Order Perturbation Models.....	355
7.3.4.3	Closest Encounter Approximation .....	357
7.3.4.4	Mean Square Field Fluctuation Approximation .....	357
7.3.4.5	Extended Two-Particle Approximation .....	358
7.3.5	Displacement Distribution in the Extended Two-Particle Approximation .....	359
7.3.6	Addition of Displacement in Several Beam Segments.....	364
7.4	Analytical Expressions for Trajectory Displacement .....	365
7.4.1	Parameter Dependencies When the Collisions Are Weak and Incomplete .....	365
7.4.2	Summary of Equations for Trajectory Displacement.....	368
7.4.3	Numerical Examples .....	369
7.5	Analytical Expressions for the Boersch Effect .....	370
7.5.1	Parameter Dependencies When the Collisions Are Weak .....	370
7.5.2	Summary of Equations for the Boersch Effect .....	372
7.5.3	Thermodynamic Limits.....	373
7.5.4	Numerical Examples .....	374
7.6	Monte Carlo Approach.....	375
7.7	Statistical Coulomb Effects in the Design of Microbeam Columns.....	377
7.7.1	Combination of Trajectory Displacement and Other Contributions to the Probe Size.....	377
7.7.2	Inclusion of the Boersch Effect .....	381
7.7.3	Design Rules for the Minimization of Statistical Interactions.....	381
7.7.4	A Strategy for the Calculation of Interaction Effects.....	381

7.8 Nonmicrobeam Instruments .....	382
7.9 Discussion .....	385
List of Symbols .....	386
Acknowledgment .....	387
References .....	387

## 7.1 INTRODUCTION

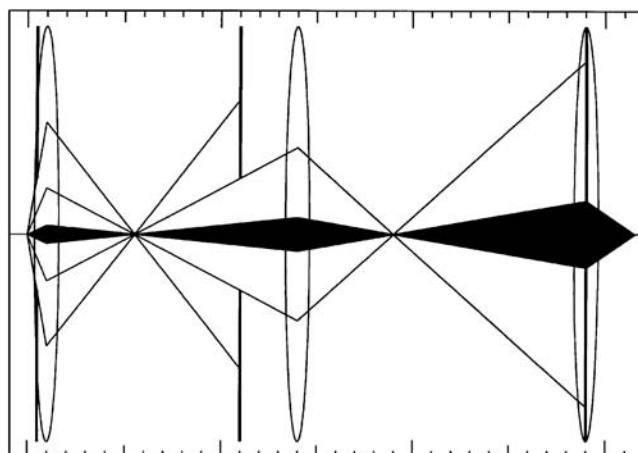
The essence of charged particle optics is that the trajectories of particles can be influenced by the electrostatic and magnetic fields of lenses, deflectors, etc. It is then implicit that the trajectories are also affected by the fields of neighboring charged particles in the beam. With the tendency to use brighter electron and ion sources, often at lower energies, the distance between neighbors decreases and the time of flight through the system increases. The particle interactions may then become dominant in determining the quality of the beam. The interaction manifests itself in three ways: the space charge effect, the trajectory displacement effect, and the Boersch effect or energy broadening effect. The space charge effect stands for the deflection by the total, averaged charge of all other particles in the beam. For a uniform charge distribution in a round beam, the deflection is proportional to the distance to the axis, causing primarily a defocus of the beam, which can be compensated by the system's lenses. For nonuniform distributions the effect can cause aberrations. The trajectory displacement and Boersch effect represent the so-called statistical effects, which are caused by the fluctuations in particle density related to the stochastic nature of the beam. The energy broadening, first investigated by Boersch (1954), represents the change in axial velocity resulting from the same individual interactions. It affects the system resolution via the chromatic aberration of lenses and deflectors. The trajectory displacement effect, first investigated by Loeffler (1964), represents the lateral shift in the particles' positions and the change in the velocity component perpendicular to the beam axis. It causes a direct deterioration of the system resolution. Because of the statistical nature of the trajectory displacement and the Boersch effect, it cannot be corrected for. The trajectory displacement decreases the beam brightness, which is usually considered to be a constant in the conservative fields of the optical elements.

The space charge effect increases linearly with the beam current. In wide beams of high currents (milliamperes or more) for low-brightness sources it is this effect that dominates. The dependency of the statistical effects on the density of the beam depends on the geometry of the beam as well as the particle density, taking on a square root dependency on the beam current for large densities. The statistical effects become dominant in narrow beams of low and moderate densities (beam currents of tens of microamperes and less), particularly when the space charge defocus is compensated for by adjusting the lenses. Typical beam parameters where the statistical interactions can be important are: for electrons, 1 keV acceleration from field emission sources or Schottky emitters at currents above 1 nA; for ions, 30 kV acceleration from liquid metal ion sources at currents above 0.1 nA. This means that low-energy scanning electron microscopes (SEM), e-beam testers, and lithography machines can be limited by statistical effects (Meisburger et al., 1992; Thomson, 1994; Beck et al., 1995). Modern low-energy instruments often have an intermediate section in which the electrons are accelerated to minimize the interactions. Electron beam lithography instruments, with their need for high currents, are also limited at higher-acceleration voltages. Focused ion beam instruments are particularly sensitive to the interactions because the particle density is higher and the time of flight through the system is longer than in electron beams due to the lower velocities. The properties of the liquid metal ion source are partly determined by statistical interactions near the emitter (Ward, 1985). Shaped beam ion lithography, once thought to be a promising lithography technique, is probably not a viable option because of coulomb effects (Vijgen and Kruit, 1992; Vijgen, 1994).

Space charge effects in low-current beams can be described by relatively simple equations. These will be reproduced in Section 7.2. The physics of statistical interactions is more complicated. Various analytical and semianalytical models have been proposed over the past 50 years or

so, which differ widely in underlying assumptions and, if comparable at all, predict significantly different results. Most models consider a single interaction phenomenon in a particular kind of beam geometry. Moreover, their application is usually restricted to a rather limited range of operating conditions, often without explicitly stating so. The most extensive study so far is the work by Jansen (1990), who developed the so-called extended two-particle approximation. This model provides sets of analytical expressions, both for the trajectory displacement and the Boersch effect, for different beam geometries and a wide range of operating conditions. Section 7.3 reviews the basic aspects of an analytical model for the statistical effects, reviews the various contributions, and discusses the extended two-particle approach in some detail. Sections 7.4 and 7.5 summarize the results of the extended two-particle model for probe-forming instruments, in which the interest is limited to the effects of the interactions on the size of the crossovers and the size of the final microprobe. These sections contain the essential equations of this chapter. A brute-force numerical technique, called Monte Carlo (MC) simulation, is sometimes used to avoid the difficulties of an analytical theory. The essence and possible errors of this method are discussed in Section 7.6. The equations available for the trajectory displacement and Boersch effect pertain to single segments between the lenses in a column. To evaluate how a whole instrument is affected, one must first choose a way of characterizing the quality of the instrument and then investigate the influence of the interactions on that quality. This is done in Section 7.7, which also contains design rules for the minimization of the effects.

The content of this chapter should allow the reader to obtain an understanding of the manifestation of coulomb effects, an appreciation for the difficulties in the theoretical description, and the knowledge necessary to estimate and then minimize the effects in a microbeam instrument. Compared with Jansen (1990), we deliberately limit ourselves in several ways: (1) The effects are only analyzed in the plane of the crossover, so instruments with Köhler illumination or lithography machines with shaped beam cannot be analyzed with the equations given here. (2) The equations presented in Sections 7.4 and 7.5 are only valid for narrow crossovers, that is, when the diameter of the beam in the crossover is small compared with the size at the entrance or exit of the beam segment. (3) Only beams with uniform current distributions are considered. (4) Only values for full width 50%, or the full width at half maximum (*FWHM*) of the distribution in the crossover are given. However, we give a more extensive description of how to use the knowledge of coulomb effects in the design of a practical instrument. A typical column for which the present overview is meant is given in Figure 7.1: a source, a few lenses to image the source on a target, and a few apertures.



**FIGURE 7.1** Schematic representation of a typical electron or ion beam instrument treated in this chapter.

## 7.2 ANALYTICAL APPROACH TO SPACE CHARGE DEFOCUS AND ABERRATIONS

### 7.2.1 LAMINAR FLOW AND SPACE CHARGE DEFOCUS

It is useful to make a clear distinction between the effects of the smoothed-out charge distribution in a beam and the effect of the statistical distribution of the charged particles. In a rotationally symmetric beam, for instance, the smoothed-out charge will not affect particles traveling on-axis, while the statistical interactions has almost the same effect on on-axis particles as on off-axis particles. We refer to the effects of the smoothed-out charge space charge effects. For high-density, high-current beams, they dominate the beam behavior. The effects can be important for beams in oscilloscope tubes or cathode ray tubes with currents from several microamperes to milliamperes, especially close to the gun. In ion beams, the effect is important at even smaller currents because the charge density is inversely proportional to the particle velocity. Thus, it is no wonder that there is an extensive literature on the subject; see, for example, Pierce (1954), Kirstein et al. (1967), and Nagy and Szilagyi (1974) or the standard textbooks on electron optics, for instance, Glaser (1952), Klemperer (1953), El-Kareh and El-Kareh (1970), or Hutter (1967). However, most of the works are not relevant to the space charge effect in the relatively low-current beams that are the subject of this book. This is because in high-current beams, a laminar flow condition occurs, in which the trajectories of the particles do not cross. In other words, a particle that starts off at the edge of a beam stays at the edge, always feeling the space charge force from one direction. In a low-current beam, a particle can first feel the force from one direction, then travel through the beam, cross the axis, and feel the space charge force from the other direction. Therefore, a different approach is necessary. The problem can be solved analytically again, if the deviations from the unperturbed trajectories are small. This regime has been studied by Meyer (1985), Zvorykin et al. (1961), Crewe (1978), Massey et al. (1981), De Chambost (1982a,b), Sasaki (1982), Van den Broek (1984, 1986), and Jansen (1990). In a round beam with a uniform current density distribution, the space charge effect corresponds to the action of an ideal negative lens and can thus be expressed in terms of a defocusing and a change in magnification. For nonuniform current density distributions, the *space charge lens* will not be ideal, an effect that can be described in terms of geometric aberrations.

For practical problems of complex geometries, it is best to use a computer program that calculates the particle trajectories under the influence of the space charge force. For large effects this calculation must necessarily be recursive—see, for example, Van den Broek (1986), Renau et al. (1982), Herrmannsfeldt et al. (1990), and Rouse et al. (1995).

### 7.2.2 EQUATIONS FOR SPACE CHARGE EFFECTS

#### 7.2.2.1 The Ray Equation

We shall restrict ourselves to the case of a rotationally symmetric beam in a region where no external forces act on the particles. The electrostatic force produced by the space charge can be found from Gauss's theorem, which states that the total flux of the force  $\vec{F}$  through the surface  $S$  depends only on the total charge within the volume  $V$  contained by  $S$ . If we consider a cylinder of length  $\Delta z$  and radius  $r$  and ignore any variation in the axial component of the space charge force ( $F_{\parallel} = \text{constant}$  or  $F_{\parallel} = 0$ ), Gauss's theorem states

$$\vec{F}_{\perp}(r, z) = \frac{e}{\epsilon_0 r} \int_0^r \rho(r_l, z) \cdot r_l dr_l \quad (7.1)$$

where  $\rho(r, z)$  is the smoothed-out charge distribution in the beam. To study the space charge effects without making *a priori* assumptions on the exact distribution of charge in the beam, it is advantageous to write

$$\rho(r, z) = \begin{cases} \frac{1}{\pi r_o(z)^2} \left[ a(z) + b(z) \left( \frac{r}{r_o(z)} \right)^2 + c(z) \left( \frac{r}{r_o(z)} \right)^4 + \dots \right] & \text{for } r \leq r_m(z) \\ 0 & \text{for } r > r_m(z) \end{cases} \quad (7.2)$$

where  $r_o(z)$  is a characteristic measure for the beam radius and  $r_m(z)$  is the outer radius. In this representation, a uniform distribution is characterized by

$$r_m(z) = r_o(z) \quad a(z) = 1 \quad b(z) = c(z) = \dots = 0 \quad (7.3)$$

A Gaussian current distribution is characterized by

$$r_m(z) = \infty \quad r_o(z) = \sqrt{2}\sigma(z) \quad a(z) = 1 \quad b(z) = \frac{-1}{2!} \quad c(z) = \frac{1}{3!} \dots \quad (7.4)$$

with  $\sigma(z)$  the sigma value of the Gaussian distribution being

$$\sigma(z) = \left[ \langle r(z)^2 \rangle / 2 \right]^{1/2} \quad (7.5)$$

A parabolic distribution is characterized by

$$r_m(z) = r_o(z) \quad a(z) = 2 \quad b(z) = -2 \quad c(z) = \dots = 0 \quad (7.6)$$

With these equations, a ray equation follows:

$$\frac{d^2r}{dz^2} = \frac{1}{4\pi\varepsilon_o} \left( \frac{m}{2e} \right)^{1/2} \frac{1}{V^{3/2}} \left[ a(z) \frac{r}{r_o(z)^2} + b(z) \frac{r^3}{2r_o(z)^4} + c(z) \frac{r^5}{3r_o(z)^6} + \dots \right] \quad (7.7)$$

### 7.2.2.2 Laminar Flow

Although it is not our purpose to study the precise effects in the laminar flow regime, it is necessary to estimate the conditions for which the laminar flow model is valid. For this, we consider a beam with a uniform current density distribution, starting at  $z_i$  with a beam semiangle  $r'_i = \alpha_0$ . By taking the ray equation for the particles moving along the edge of the beam and rewriting it, we obtain the following laminar flow equation:

$$\left( \frac{dr}{dz} \right)^2 = \frac{1}{2\pi\varepsilon_o} \left( \frac{m}{2e} \right)^{1/2} \frac{1}{V^{3/2}} \ln \left( \frac{r}{r_i} \right) + \alpha_0^2 \quad (7.8)$$

The minimum beam radius  $r_l$  follows by taking  $dr/dz = 0$ :

$$r_l = r_i \exp \left[ -2\pi\varepsilon_o \left( \frac{2e}{m} \right)^{1/2} \frac{\alpha_o^2 V^{3/2}}{I} \right] \quad (7.9)$$

The  $z$ -coordinate of the waist follows from integrating the ray equation, which cannot be done analytically. Schwartz (1957) studied how to obtain a minimum sort size under laminar flow conditions.

From this derivation it may seem that every beam, however small the current, would be described by the laminar flow ray equation; in other words, charged particles never seem to be able to cross the axis. This would be consistent with the conclusion from Gauss's theorem that the space charge force goes to infinity in a zero-size crossover. However, we have made two assumptions in the derivation. The first assumption was that all particles would cross the axis if not affected by space charge. This would occur only in a beam with zero-size crossover, so effectively the assumption was that the unaffected crossover size would be much smaller than the waist  $r_l$  as found from Equation 7.9. We shall call this the *focus size condition* for laminar flow. If this condition is not fulfilled, the space charge force cannot become large enough to keep an edge particle at the same edge of the beam; thus, it will start to cross the beam. The second assumption was that even close to the waist, the charge could be considered to be smoothed-out. This is justified only if the number of particles simultaneously exerting a force on the edge particle is much larger than one; this requirement is known as the *continuum condition*.

If we apply this condition to the region around the waist of the beam and take the number of particles in a cylinder of length  $r_l$ , it reads

$$\left( \frac{m}{2e} \right)^{1/2} \cdot \frac{r_l I}{e V^{1/2}} \gg 1 \quad (7.10)$$

If this equation is not satisfied, we may not assume that the laminar flow model can be applied.

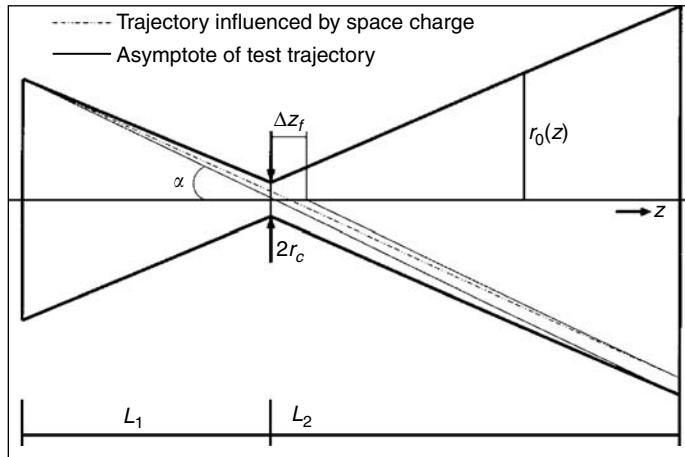
### 7.2.2.3 Space Charge Defocus

The other condition under which we can find analytical expressions for the space charge effect is only fulfilled at low currents. We assume that the current density distribution in the beam is only slightly disturbed and calculate the trajectory of a test particle in first-order perturbation theory. Consider a beam section as in Figure 7.2 with characteristic radius  $r_o(z) = r_c + \alpha_o |z - z_o|$  and current density distribution as described by Equation 7.2. The undisturbed trajectory  $r(z) = \alpha(z_c - z)$  of a test particle is now substituted in the ray equation (Equation 7.7) and the perturbation  $\Delta r(z)$  is calculated. The first term in Equation 7.3 yields an effect proportional to  $\alpha$ , as with an ideal lens. If the effect is expressed as a defocus distance  $\Delta z_f$ , one finds

$$\Delta z_f = \frac{1}{4\pi\varepsilon_o} \left( \frac{m}{2e} \right)^{1/2} \cdot \frac{aI(L_1 + L_2)}{\alpha_o^2 V^{3/2}} \cdot \frac{K_1 \cdot Z(K_1) + K_2 \cdot Z(K_2)}{(K_1 + K_2)} \quad (7.11)$$

with

$$Z(K) = 1 + \frac{1}{1+K} - \frac{2}{K} \ln(1+K) \quad (7.12)$$



**FIGURE 7.2** Beam section with space charge defocusing effect.

$$K_1 = \frac{\alpha_o L_1}{r_c} \quad K_2 = \frac{\alpha_o L_2}{r_c} \quad (7.13)$$

and  $a = 1$  for uniform or Gaussian current distribution.

For  $K > 100$ , one may approximate  $Z(K) = 1$  within 10% accuracy, and the  $K$ -dependent term in the equation disappears. For  $K < 100$ , this term is always smaller than 1. The physical interpretation of this term being the correction for the region in which the test particle travels through the center of the beam. For the derivation of an expression for the effect on magnification and consideration of a more general test trajectory, see Jansen (1990).

When we consider a cylindrical beam segment of radius  $r_o$  and length  $L$  with a test particle traveling parallel to the axis, one has to assume that this segment is followed by a lens with focal distance  $f$  to get a useful expression for defocus:

$$\Delta z_f = \frac{1}{4\pi\epsilon_o} \left( \frac{m}{2e} \right)^{1/2} \cdot \frac{aIL}{r_o^{2V^{3/2}}} f^2 \quad (7.14)$$

Obviously, if the defocus values found from these equations become appreciable, as compared with  $L_1$  and  $L_2$ , the approximations in deriving the equations are not valid any more.

Space charge defocusing can be compensated by adjusting the strength of one or more lenses in the system, so for the operator it might not even be noticeable. However, if the beam current is not stable, it leads to blurring or a focus drift. In a shaped beam lithography systems, where the beam current depends on the selected shape, a dynamic focus correction is necessary. (Sturans et al., 1998; Sohda et al., 1995).

#### 7.2.2.4 Space Charge Aberrations

The second term in the ray equation (Equation 7.2) gives rise to third-order geometric aberrations. If we limit our analysis again to a particle which crosses the axis at the position of the waist (see Figure 7.2), the only effect is spherical aberration. For other aberrations, see Jansen (1990). The effect can be expressed as a coefficient of spherical aberration of the space charge lens

$$C_s = \frac{1}{4\pi\epsilon_o} \left( \frac{m}{2e} \right)^{1/2} \frac{-bI(L_1 + L_2)}{\alpha_o^4 V^{3/2}} \frac{S(K_1) + S(K_2)}{K_1 + K_2} \quad (7.15)$$

with

$$S(K) = \frac{1}{2}K - 2\ln(1+K) + \frac{9K + 21K^2 + 13K^3}{6(1+K)^3} \quad (7.16)$$

We recall that  $b$  is  $-1/2$  for a Gaussian current density distribution and  $-2$  for a parabolic distribution, so the spherical aberration coefficient is positive, as it is for all regular charged particle lenses. For large  $K$  the function  $S(K)$  approaches  $1/2K$ , so the  $K$ -containing term in Equation 7.15 approaches half. For  $K$ -values  $< 10$ , this term decreases rapidly.

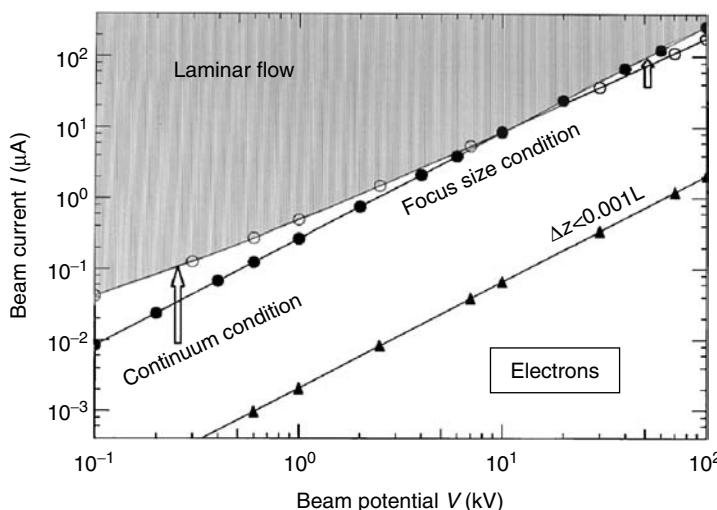
In a cylindrical beam segment, radius  $r_o$ , length  $L$ , followed by a lens with focal distance  $f$ ,  $C_s$  is given by

$$C_s = \frac{1}{4\pi\varepsilon_0} \left(\frac{m}{2e}\right)^{1/2} \frac{-bIL}{2r_o^4 V^{3/2}} f^4 \quad (7.17)$$

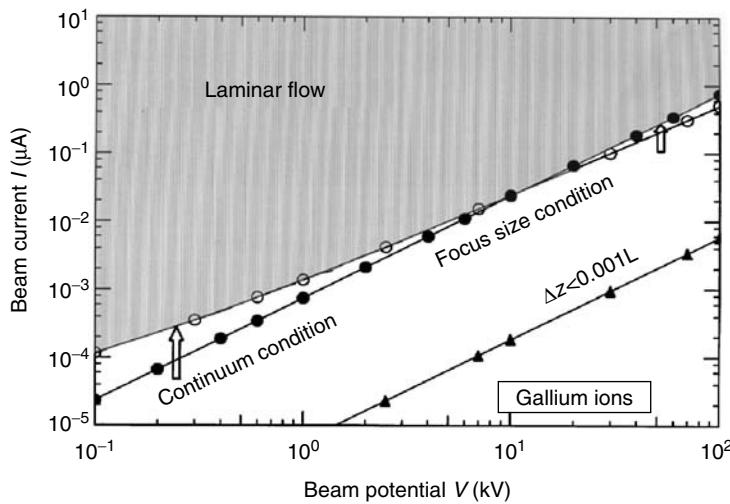
When we compare the space charge effects in a cylindrical beam segment with those in a segment of the same size but having a narrow crossover in the middle, we find that they are identical.

### 7.2.3 NUMERICAL EXAMPLES

The currents and acceleration voltages for which the conditions of laminar flow are satisfied are given in Figures 7.3 and 7.4 for electrons and Ga ions, respectively. The beam angle in this example is  $\alpha_o = 1$  mrad, typical for a long segment ( $L = 200$  mm,  $L_1 = 50$  mm, and  $L_2 = 150$  mm) in a microbeam system. For the continuum we have substituted Equation 7.9 into Equation 7.10 with  $r_i = \alpha_o L_1$ . The focus size condition is expressed as  $r_l > r_u$ , with  $r_u$  the unaffected crossover size, in the numerical example taken as  $1$   $\mu\text{m}$ . For laminar flow, both conditions must be satisfied. The same figures also show the  $I$ ,  $V$  values for which the space charge defocus is  $1/1000$  of the segment length (in this example,  $0.2$  mm). As the defocus is proportional to  $I$ , it is easy to find the  $I$ ,  $V$  values for different defocus. At  $\sim 20$  mm defocus we run into the laminar flow situation, but by then the approximations for the defocus equation are not satisfied anymore.



**FIGURE 7.3** Laminar flow conditions (shaded area) for an electron beam segment of 200 mm length with an unperturbed crossover position at 50 mm and crossover size 1  $\mu\text{m}$ , half angle 1 mrad. The  $I$  and  $V$  values for which the space charge defocus is 0.2 mm are also indicated.



**FIGURE 7.4** Laminar flow conditions (shaded area) for a beam segment of gallium ions of 200 mm length with an unperturbed crossover position at 50 mm and crossover size 1  $\mu\text{m}$ , half angle 1 mrad. The  $I$  and  $V$  values for which the space charge defocus is 0.2 mm are also indicated.

## 7.3 ANALYTICAL APPROACH TO STATISTICAL COULOMB EFFECTS

### 7.3.1 GENERAL FORMULATION OF THE PROBLEM

For the analysis of the statistical coulomb effect, the beam is usually schematized as a succession of beam segments, separated by optical components. The impact of the interaction between the particles is then studied for a single beam segment, assuming that the total effect generated in the entire beam can be represented as a sum of the effects generated in the individual beam segments. This is not a trivial assumption, as will be discussed in more detail at the end of this section. The objective of the theory is now to evaluate the statistical effects as a function of the experimental parameters  $I$ ,  $V$ ,  $L$ ,  $r_c$ ,  $\alpha_o$ , and the crossover location parameter  $S_c$  defined as

$$S_c = \frac{L_1}{L} \quad 0 \leq S_c \leq 1 \quad (7.18)$$

where  $L_1$  is the distance from the start of the beam segment to the crossover. Thus,  $S_c$  is equal to zero for a diverging beam and equal to one for a converging beam. The segments are assumed to be rotationally symmetric. Furthermore, the acceleration region in the gun as well as the lens and deflector areas are usually treated as infinitely thin. The unperturbed trajectories of the particles then are visualized as straight lines which are broken at the location of the lenses and deflectors.

The basic approach of the analytical model is to consider the coulomb interaction of a single test particle with all other particles in the beam, referred to as field particles. Owing to this interaction, the test particle will experience a deviation from its unperturbed trajectory. The term *displacement* will be used to denote any deviation in position or velocity from the unperturbed values. The displacement of the test particle is fully determined by the initial coordinates of the field particles, relative to the test particle. Owing to the stochastic nature of the beam, another test particle running along the same trajectory some time later will be surrounded by a different configuration of field particles and consequently, experience a different displacement. The objective is to determine the distribution of displacements of a large set of test particles, running successively with identical velocity along a specific trajectory in the beam called the reference trajectory. It is assumed that the

test particles are well separated and can be considered as independent. The problem now consists of two parts. The first is called the *dynamical part of the problem*. This determines the total influence of a given set of field particles on the path of a specific test particle. The second is called the *statistical part of the problem*. This determines the distribution of displacements in velocity and position of the entire set of test particles successively arriving at the end of the beam segment.

As we are interested in the principal aspects of the calculation, we will introduce two approximations, which can be avoided in a more generic calculation, but which help to simplify the problem while retaining its key features:

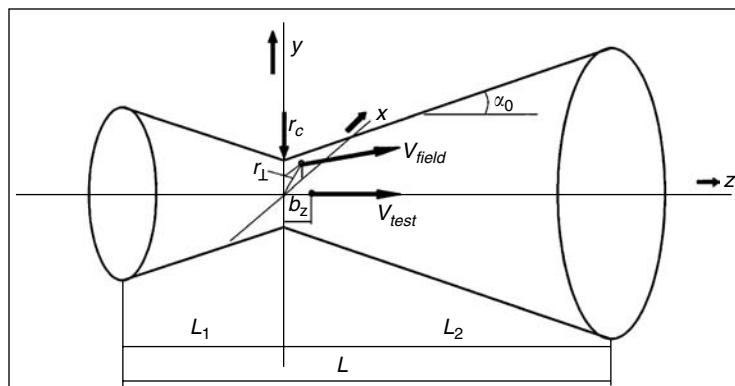
- We will assume that the beam is monochromatic at the start of the beam segment. This implies that the initial axial velocity is realistic as the axial velocity spread is, in general, several orders of magnitude smaller than the lateral velocity spread due to acceleration of the beam near the source (Zimmerman, 1970; Jansen, 1990).
- We will take the beam axis as reference trajectory. Owing to the rotational symmetry of the beam, no systematic forces act on the test particles moving along this central trajectory. Accordingly, the displacements experienced by the particles are purely stochastic, which means that their average displacement is zero. Particles moving along other trajectories in the beam will be subject to the combined action of space charge and statistical effects. The evaluation of the statistical effects experienced along such off-axis trajectories is more involved.

It can be shown that the result obtained for a central reference trajectory in a monochromatic beam is representative for other particle trajectories and nonmonochromatic beams as well (Jansen, 1990).

Figure 7.5 shows the unperturbed trajectories of a test particle and a field particle at the moment the field particle passes the  $x$ - and  $y$ -plane of a fixed coordinate system within the beam segment. The  $x$ - and  $y$ -plane are taken equal to the crossover plane. The  $z$ -axis of this system coincides with the beam axis. The particles run in the positive  $z$ -direction. For convenience of notation, the set of relative coordinates of a single field particle, with respect to the test particle, is abbreviated as  $\vec{\xi}$  where

$$\vec{\xi} = \{x, y, b_z, v_x, v_y\} \quad \text{or} \quad \vec{\xi} \{r_{\perp}, \phi, b_z, v, \psi\} \quad (7.19)$$

using rectangular or cylindrical coordinates, respectively. The quantity  $r_{\perp}$  is the modulus of the projection of the relative position vector  $\vec{r} = \vec{r}_{\text{field}} - \vec{r}_{\text{test}}$  on the crossover plane and  $b_z$  is the impact



**FIGURE 7.5** Test particle and field particle in a beam segment.

parameter. The vector  $\vec{\xi}$  gives a complete specification of the unperturbed trajectory of the field particle relative to that of the test particle. By specifying the set  $\vec{\xi}_1, \vec{\xi}_2, \dots, \vec{\xi}_{N-1}$ , of all  $N - 1$  field particles, the unperturbed coordinates of all the particles can be determined at any moment.

Given the configuration of field particles, specified by  $\vec{\xi}_1, \vec{\xi}_2, \dots, \vec{\xi}_{N-1}$ , the test particle will experience a certain displacement from its unperturbed path. The displacement in velocity is expressed in terms of its components  $\Delta v_x$ ,  $\Delta v_y$ , and  $\Delta v_z$ . The spatial displacement can best be expressed in terms of a virtual displacement in the crossover plane. For the probe-forming systems considered here, this plane will be optically conjugate to the target plane. The virtual displacement is determined by extrapolating the final position of the test particle toward the crossover plane by a straight line along its final velocity. The quantities  $\Delta x$  and  $\Delta y$  are the  $x$ - and  $y$ -components of the virtual radial displacement  $\Delta r$  in the crossover plane. The trajectory displacement effect is related to the generation of these displacements. The Boersch effect corresponds to the generation of a spread in axial velocities and the relevant displacement therefore is  $\Delta v_z$ . In general, we will denote the set of relevant displacements by the vector  $\Delta \vec{\eta}$ . The quantity  $\Delta \vec{\eta}$  may represent a simple scalar (as in case of the Boersch effect), but is in general a vector.

Now, the dynamical part of the problem is to evaluate the displacement of the test particle  $\Delta \vec{\eta}$  as a function of the configuration field particles  $\vec{\xi}_1, \vec{\xi}_2, \dots, \vec{\xi}_{N-1}$

$$\Delta \vec{\eta} = \Delta \vec{\eta}(\vec{\xi}_1, \vec{\xi}_2, \dots, \vec{\xi}_{N-1}) \quad (7.20)$$

The statistical part of the problem is to find the probability  $P_N(\vec{\xi}_1, \vec{\xi}_2, \dots, \vec{\xi}_{N-1})$  of the configuration  $\vec{\xi}_1, \vec{\xi}_2, \dots, \vec{\xi}_{N-1}$  and to evaluate the distribution of displacements from

$$\rho(\Delta \vec{\eta}) = \int d\vec{\xi}_1 d\vec{\xi}_2 \dots d\vec{\xi}_{N-1} P_N(\vec{\xi}_1, \vec{\xi}_2, \dots, \vec{\xi}_{N-1}) \delta[\Delta \vec{\eta} - \Delta \vec{\eta}(\vec{\xi}_1, \vec{\xi}_2, \dots, \vec{\xi}_{N-1})] \quad (7.21)$$

in which  $\delta[\vec{x}]$  is the (multidimensional) delta-Dirac function. The distribution  $\rho(\Delta \vec{\eta})$  contains all the desired information.

The evaluation of the root mean square (rms) width of the distribution is, in general, less involved than the calculation of the full displacement distribution  $\rho(\Delta \vec{\eta})$ .

It would be sufficient to consider the rms when the displacement distribution would be Gaussian for all practical conditions, as is assumed too often. In that case the rms value can directly be related through a proportionality constant to experimentally more meaningful width measures such as the *FWHM* and the full width median ( $FW_{50}$ ). The latter is defined as the width that contains 50% of the distribution. Unfortunately, as will be shown later, Gaussian behavior is the exception rather than the rule: the displacement distribution takes on the Gaussian form for relatively high particle densities only when the total displacement of a test particle is the result of a large number of independent scattering events with the different field particles. Non-Gaussian behavior prevails for most conditions. This implies that all models which evaluate an rms value only are essentially incomplete and even misleading. The rms is dominated by the tails of the distribution and may show a different dependency on the experimental parameters than the *FWHM* and  $FW_{50}$  which are determined by the core of the distribution. We, therefore, have to consider the calculation of the full distribution  $\rho(\Delta \vec{\eta})$  from Equation 7.21.

An explicit analytical calculation of  $\rho(\Delta \vec{\eta})$  from Equation 7.21 in terms of the experimental parameters does not appear to be feasible, and one is forced to introduce some further simplifying assumptions. The different solutions that have been presented in literature involve different assumptions, unfortunately, leading to significantly different predictions of the magnitude of the interaction effects as a function of the experimental parameters. A substantial number

of articles have been published over the past 40 years to deal with the problem specified by Equations 7.20 and 7.21.

The different approaches will be reviewed in Section 7.3.4. To facilitate that review, a couple of related issues will be addressed first. Section 7.3.2 deals with two approximations, common to most theories, which serve to reduce the  $N$ -particle problem to a two-particle problem. Section 7.3.3 reviews the approximations required to solve the two-particle dynamical problem. The reader is referred to Jansen (1990) for a more detailed historical overview of the various models presented in the literature, as well as for a quantitative comparison of the results.

### 7.3.2 REDUCTION OF THE $N$ -PARTICLE PROBLEM

All analytical approaches have in common that the initial coordinates of the field particles are assumed to be identical statistically independent quantities, which implies

$$P_N\left(\vec{\xi}_1, \vec{\xi}_2, \dots, \vec{\xi}_{N-1}\right) = \prod_{i=1}^{N-1} P_2\left(\vec{\xi}_i\right) \quad (7.22)$$

where  $P_2(\vec{\xi})$  is the probability that a field particle has coordinates  $\vec{\xi}$  relative to the test particle. This can be directly expressed in terms of the beam geometry parameters. For instance, for the on-axis reference trajectory in a rotationally symmetric beam with a uniform spatial and angular distribution, one may write

$$\int d\vec{\xi} P_2\left(\vec{\xi}\right) = \int_0^{v_o} dv \frac{2v}{v_o^2} \int_0^{2\pi} d\Phi \int_0^{r_c} dr \frac{2r_\perp}{r_c^2} \int_{-S_c L}^{(1-S_c)L} \frac{db_z}{L} \quad (7.23)$$

where  $v_o = \alpha_o v_z$ . It was anticipated that the magnitude of the displacement  $\Delta\eta = |\Delta\vec{\eta}|$  will depend on the angles  $\phi$  and  $\psi$  only through the relative angle  $\Phi = \psi - \phi$ .

Equation 7.22 ignores all correlations in the coordinates of the field particles. In practical systems, statistical correlations between the particle coordinates are determined both by the initial probability distribution of the particles leaving the cathode and the dynamics under the influence of the coulomb force during the time of flight. The last source of correlation is closely related to the concept of Debye screening, first discussed by Debye and Hückel (1923). Screening of the test particle requires that the interaction takes place over a sufficiently long time period to approach thermodynamic equilibrium. It corresponds to a redistribution of the field particles surrounding the test particle, under influence of the electrostatic field of the test particle. The screening, built up during the flight through the beam sections preceding the one considered, corresponds to correlations within the configuration  $\vec{\xi}_1, \vec{\xi}_2, \dots, \vec{\xi}_{N-1}$ . Equation 7.22 ignores these correlations as well as those related to the emission process. This simplification may lead to an overestimation of the probability of large displacement  $\Delta\vec{\eta}$  and the tails of the resulting displacement distribution  $\rho(\Delta\vec{\eta})$  should be viewed with some caution.

A second type of assumption made in most analytical models is that the  $N$ -particle problem of calculating the displacement  $\Delta\vec{\eta}(\xi_1, \xi_2, \dots, \xi_{N-1})$  can somehow be expressed in terms of the two-particle problem of calculating the displacement  $\Delta\vec{\eta}_2(\vec{\xi})$  of the test particle caused by the interaction with a single field particle with coordinates  $\vec{\xi}$ . One implementation of this is to assume that the  $N$ -particle displacement can be expressed as a sum of all two-particle effects:

$$\Delta\vec{\eta}\left(\vec{\xi}_1, \vec{\xi}_2, \dots, \vec{\xi}_{N-1}\right) \approx \sum_{i=1}^{N-1} \Delta\vec{\eta}_2\left(\vec{\xi}_i\right) \quad (7.24)$$

The dynamical part of the problem is now to evaluate the two-particle displacement  $\Delta\vec{\eta}_2(\vec{\xi})$  as a function of the relative coordinates  $\vec{\xi}$ , and the statistical part is to perform some kind of averaging over the distribution of relative coordinates  $P_2(\vec{\xi})$ .

### 7.3.3 TWO-PARTICLE DYNAMICS

Despite the substantial simplification of the problem by reducing it from an  $N$ -particle problem to a two-particle problem, a general analytical solution can still not be obtained without introducing further approximations. The most significant difficulty is related to the two-particle dynamical problem of evaluating  $\Delta\vec{\eta}_2(\vec{\xi})$ . The calculation of the motion of two particles through a force which is, as the coulomb force, proportional to the inverse square of the relative distance between the particles, is called the Kepler problem. It was solved by Newton (1687). The classic solution of the Kepler problem specifies the relative trajectory of the particles in the so-called orbital plane, where one finds the well-known hyperbolic orbit. However, this solution is useful only if the particles effectively come from infinity and recede to infinity.

Our problem requires a more detailed analysis, which takes the time of flight through the beam segment into account. To find the coordinates of the test particle at the exit plane, one needs to derive the coordinates in the orbital plane as a function of time  $t$ . As it turns out, one can obtain the time  $t$  explicitly as a function of the relative distance  $r$ , thus  $t = t(r)$ , but the function involved cannot be inverted. This prohibits an exact explicit analytical calculation of the particle trajectories as a function of time  $t$ .

A good approximation of the exact final coordinates of the particles involved in a two-particle collision can be obtained for two types of collisions. One is a *complete collision* (or nearly complete collision), in which the particles effectively do come from infinity and recede to infinity. For this type of collision, it is sufficient to consider the asymptotic values of the coordinates. A (nearly) complete collision can occur only if the particles are far apart at the entrance and exit of the beam segment, and close near the crossover. Thus, the quantities  $K_1$  and  $K_2$  defined in Equation 7.13 must be large and the average axial distance between the particles must be smaller than the maximum beam diameter. The latter condition is called the extended beam condition, as opposed to the pencil beam condition. The pencil beam condition is defined as

$$\chi_c = \lambda\alpha_o L = \left( \frac{m}{2e^3} \right)^{1/2} \frac{1\alpha_o L}{V^{1/2}} \ll 1 \quad (7.25)$$

where  $\lambda$  is the linear particle density  $I/ev_z$ . The other type is a *weak collision*, in which the deviations from the unperturbed trajectories are small. First-order perturbation dynamics can be used to describe such a collision. In this approach, one approximates the actual interaction force by the force which would occur when the particles would follow their unperturbed trajectories. Weak collisions occur at low particle densities, characterized by a small *scaled linear particle density*

$$\bar{\lambda} = \frac{m^{1/2}}{2^{7/2}\pi\varepsilon_o e^{1/2}} \frac{1}{\alpha_o^2 I^{3/2}} < 10^{-3} \quad (7.26)$$

For the definition in a cylindrical beam,  $\alpha_o$  is replaced by  $r_o/L$ . For the remaining collisions, those which are both *strong and incomplete*, one has to perform the inversion of the function  $t(r)$  by numerical means. An explicit analytical expression for the resulting displacement can therefore not be obtained. See Jansen (1990) for the mathematical definition of this classification.

### 7.3.4 OVERVIEW OF APPROXIMATE SOLUTIONS

We now return to the original problem of evaluating the dynamical and the statistical part of the calculation, as described in Sections 7.3.1 and 7.3.2. The main approaches published to date can be grouped as follows.

#### 7.3.4.1 Models Derived from Plasma Physics

This approach exploits the formalism developed in the fields of plasma physics and stellar dynamics centered around the so-called Fokker–Planck equation (Trubnikov, 1965; Sivukhin, 1966; van Kampen, 1981). It treats the statistical interactions as a diffusion process in velocity space and focuses on the calculation of the mean square velocity shifts. The collision between two particles is treated as *complete*. The velocity change of a test particle experienced during its flight through a thin slice of the beam is calculated by assuming that it results from many complete collisions with different field particles. The total effect in the beam segment follows by integration over all slices.

Along these lines, Zimmerman (1968, 1969, 1970) calculated the Boersch effect generated in a cylindrical beam segment, thereby establishing the first extensive theory on the Boersch effect. Zimmerman introduced the concept of relaxation of internal energy. The internal energy of a particle is defined as its kinetic energy in the frame of reference moving with the beam. The internal energy can be expressed in terms of a beam temperature  $T$ , using

$$\frac{1}{2}m\langle\Delta v^2\rangle = \frac{3}{2}k_B T \quad (7.27)$$

where  $\langle\Delta v^2\rangle = \langle v^2 \rangle - \langle v \rangle^2$ . The essential observation is that the axial internal energy spread of the particles is significantly reduced during acceleration, namely, by a factor  $k_B T_o/eV_o$ , in which  $T_o$  is the beam temperature at the emission surface of the source and  $V_o$  the acceleration voltage. Although the axial internal energy spread is reduced, the lateral internal energy spread is not affected and a nonequilibrium situation is generated. One therefore has to distinguish between two beam temperatures, an axial beam temperature  $T_{||}$ , and a lateral beam temperature  $T_{\perp}$ , defined as

$$T_{||} = \frac{m}{k_B}\langle\Delta v_{||}^2\rangle \quad T_{\perp} = \frac{m}{2k_B}\langle\Delta v_{\perp}^2\rangle \quad (7.28)$$

The nonequilibrium situation caused by the acceleration implies that  $T_{||} \ll T_{\perp}$ . Owing to the coulomb interaction between the particles, relaxation will occur toward a more isotropic temperature distribution. Consequently the axial beam temperature will increase which can be expressed as an increase in energy spread. The rate at which the axial energy spread increases can be estimated from the average gain in axial velocity spread per unit of length due to complete collisions experienced by the test particle with its surrounding field particles, which are assumed to be many. Zimmerman found that the rms energy spread increases with the square root of the current density.

Zimmerman's work was extended by Knauer (1979a,b) to include converging and diverging beams. For these nonparallel beams one has to account for the fact that the lateral beam temperature will vary with the beam diameter. This can be understood from the principle of conservation of emittance, which says that—in the absence of interactions—the four-dimensional volume  $\Delta V_4$  in the phase space of lateral coordinates is constant along the beam:

$$\Delta V_4 = \Delta v_{\perp}^2 \Delta r_{\perp}^2 = \pi v_{||}^2 d\Omega dS = \text{constant} \quad (7.29)$$

where  $dS$  is the surface of a cross section of the beam and  $d\Omega$  the corresponding solid angle of rays through this surface. Conservation of emittance implies that a reduction of  $dS$  coincides with an increase of the lateral velocity spread, which is equivalent to an increase of the lateral beam temperature  $T_{\perp}$ . In a beam segment with a crossover with a uniform angular distribution, the lateral beam temperature can be expressed as

$$T_{\perp} \approx \frac{r_c^2}{r_o(z)^2} T_{\perp c} \quad \text{where} \quad T_{\perp c} = \frac{eV}{2k_B} \alpha_o^2 \quad (7.30)$$

as can be understood from Equations 7.28 and 7.29. Knauer modeled a nonparallel beam as a succession of thin slices each with a different lateral beam temperature and applied Zimmermann's result for a cylindrical beam to each slice. Implicit in this approach is the assumption that a test particle will experience many independent complete collisions in each of these slices.

The conditions required to justify the approach of Zimmermann and Knauer are, in general, not met in practical beams, as was shown by Jansen (1990) in a more detailed discussion of the Fokker–Planck approach. However, the conditions of many complete collisions per slice can, to a certain extent, be fulfilled in the crossover area. This is the *hottest* and most dense part of the beam and shows therefore the closest approximation to plasma. Zimmermann's and Knauer's approach may therefore give a reasonable approximation for the Boersch effect generated in the vicinity of a crossover for relatively high-density beams of moderate brightness.

Tang (1987) took a more elaborate approach along the same lines by considering the full Fokker–Planck equation. He evaluated the evolution of the single-particle distribution function in position and velocity along the beam axis under the influence of both the average space charge force and the fluctuating component of the interaction force. It should be emphasized that this approach assumes the same mechanism as the models of Zimmermann and Knauer, and is therefore not suited to describe the trajectory displacement effect, or the Boersch effect generated in the dilute parts of the beam.

#### 7.3.4.2 First-Order Perturbation Models

In this approximation it is assumed that the deviations from the unperturbed trajectories are small. The actual interaction force acting on the test particle is now approximated by the interaction force which would act on the test particle when all particles would follow their unperturbed trajectories. Contrary to the models based on the Fokker–Planck approach, this model starts from the actual unperturbed particle trajectories and does not assume the occurrence of many random scattering events in each slice of the beam.

To understand the essence of this approach, it is helpful to make the relation between the force  $\vec{F}$  acting on the test particle and its displacement  $\Delta\vec{\eta}$  explicit. In general, one can write the displacement  $\Delta\vec{\eta}$  as an integral over time as

$$\Delta\vec{\eta}\left(\vec{\xi}_1, \vec{\xi}_2, \dots, \vec{\xi}_{N-1}\right) = \int_{t_i}^{t_f} dt \vec{G}_{\Delta\eta}\left(\vec{\xi}_1, \vec{\xi}_2, \dots, \vec{\xi}_{N-1}, t\right) \quad (7.31)$$

where  $t_i$  and  $t_f$  correspond to the moments that the test particle enters and leaves the beam segment, respectively. The quantity  $\vec{G}_{\Delta\eta}$  depends on the displacement considered:

$$\begin{aligned} \vec{G}_{\Delta v_z} &= \frac{1}{m} \vec{F}_{\parallel}\left(\vec{\xi}_1, \vec{\xi}_2, \dots, \vec{\xi}_{N-1}, t\right) \\ \vec{G}_{\Delta r_{\perp}} &= \frac{1}{m} (t_f - t) \vec{F}_{\perp}\left(\vec{\xi}_1, \vec{\xi}_2, \dots, \vec{\xi}_{N-1}, t\right) \end{aligned} \quad (7.32)$$

where  $\vec{F}(\vec{\xi}_1, \vec{\xi}_2, \dots, \vec{\xi}_{N-1}, t)$  is the force experienced by the test particle at time  $t$  within the configuration of field particles  $\vec{\xi}_1, \vec{\xi}_2, \dots, \vec{\xi}_{N-1}$ . The displacements  $\Delta v_z$  and  $\Delta \vec{r}_\perp$  refer to the deviation from the unperturbed trajectory at  $t_f$ . The first-order perturbation approximation can be expressed as

$$\vec{F}(\vec{\xi}_1, \vec{\xi}_2, \dots, \vec{\xi}_{N-1}, t) \approx \vec{F}^*(\vec{\xi}_1, \vec{\xi}_2, \dots, \vec{\xi}_{N-1}, t) = \sum_{i=1}^{N-1} \vec{F}_i^*(\vec{\xi}_i, t) \quad (7.33)$$

where the asterisk refers to the force experienced along the unperturbed trajectories. The unperturbed trajectories are known in advance and can directly be expressed in terms of the coordinates  $\vec{\xi}_1, \vec{\xi}_2, \dots, \vec{\xi}_{N-1}$ . Notice that the second equality in Equation 7.33 is exact as a consequence of the fact that forces are additive.

Substitution of Equations 7.32 and 7.33 into Equation 7.31 leads back to Equation 7.24, which expresses that the total displacement can be calculated as the sum of two-particle effects. Thus, the decomposition of the  $N$ -particle problem comes out naturally within the first-order perturbation approximation on the argument that forces are additive.

The first model along these lines was developed by Loeffler (1969). Loeffler and Hudgin (1970) extended the model to evaluate the impact of statistical interactions on the lateral properties of the beam. They distinguished three different phenomena of statistical interactions in a beam with a narrow crossover, corresponding to each of the three components of the coulomb force exerted by a field particle on a test particle running along the beam axis. The Boersch effect, which corresponds to an energy change, is associated with the longitudinal component. The lateral force is separated in a component in the direction of the transverse velocity of the field particle and a component perpendicular to this. This latter causes an *angular change* of the test particle. The former does not produce angular deflection. Its sign, however, changes when the particles pass the point of closest approach, which will, for a narrow crossover, roughly coincide with the passage through the crossover. The angular deflection, experienced before reaching the crossover, is thus compensated in the trajectory following the crossover. This force component causes, therefore, only a lateral shift, which Loeffler and Hudgin termed *trajectory displacement*. The median width measures of the distributions of energy changes, angular changes, and trajectory displacements are obtained by numerical integration over all possible trajectories of the field particles relative to the test particle, employing a Monte Carlo method. Loeffler and Hudgin presented normalized curves, from which the magnitude of the different phenomena can be determined. All three curves clearly indicate the existence of two separate regimes, which show a different dependency on the experimental parameters. For high particle densities, all effects ultimately increase with approximately the square root of the beam current, but for low particle densities, different powers (varying between 1 and 3) are found for the separate effects.

It should be emphasized that the results of Loeffler and Hudgin depend critically on the following two assumptions:

- The first-order perturbation approximation requires that all interactions are weak causing only small deviations from the unperturbed trajectories.
- The crossover diameter is very small compared with the diameter of the beam at the beam segment entrance and exit.

Accordingly, their results are expected to break down with the occurrence of strong collisions at higher particle densities, as well as for beam geometries in which the crossover cannot be considered as narrow.

Finally, it is important to note that a straightforward calculation of the rms value from Equation 7.24 based on weak collisions would lead to divergent results. Loeffler and Hudgin avoided this problem by using a specific statistical procedure in which the  $n$ th largest displacement

is computed from the total set of  $N - 1$  displacements of the test particle caused by the interaction with each field particle. From this ordered set the median width of the displacement distributions is then computed, rather than an rms value.

#### 7.3.4.3 Closest Encounter Approximation

This approach assumes that the displacement of the test particle is dominated by the single strong interaction with its nearest-neighbor field particle, which is defined as the particle with the smallest impact parameter with respect to the test particle. The impact parameter  $b$  of a two-particle collision is defined as the closest distance of approach in the absence of interaction. Let  $b_n$  be the impact parameter of the nearest neighbor and  $\vec{\xi}_n$  the coordinates of this particle relative to the test particle. The closest encounter approximation reduces the  $N$ -particle dynamical problem to a two-particle problem by assuming

$$\Delta\vec{\eta}\left(\vec{\xi}_1, \vec{\xi}_2, \dots, \vec{\xi}_{N-1}\right) \approx \Delta\vec{\eta}_2\left(\vec{\xi}_n\right) \quad (7.34)$$

and then using complete collision dynamics to derive the displacement  $\Delta\vec{\eta}_2(\vec{\xi}_n)$  caused by the nearest neighbor. A model of this type was proposed by Rose and Spehr (1980) for the Boersch effect and Weidenhausen et al. (1985) to evaluate stochastic angular deflections. This seems a reasonable model provided that the collision can indeed be considered as nearly complete which relies on the conditions given in Section 7.3.3.

For the statistical part of the problem, this model evaluates the probability  $W_2(\vec{\xi}_n, b_n)$  that an arbitrary field particle, with coordinates  $\vec{\xi}_n$ , is the nearest neighbor of the test particle. The rms displacement is then evaluated as

$$\langle\Delta\eta^2\rangle = \int db_n \int d\xi W_2\left(\vec{\xi}, b_n\right) \Delta\eta_z\left(\vec{\xi}_n\right)^2 \quad (7.35)$$

The closest encounter approach presupposes that the main contribution to the rms value comes from a single collision with its nearest-neighbor field particle and that this collision can be treated as nearly complete. It should be noticed that this approach is, in general, not suited to evaluate other width measures than the rms, such as the  $FW_{50}$  and  $FWHM$  values. This can be understood from the fact that the central part of the displacement distribution effect is related to weak collisions which are not necessarily complete and result from many weak interactions rather than a single strong interaction.

#### 7.3.4.4 Mean Square Field Fluctuation Approximation

In all approaches described so far, the displacement of the test particle was first calculated for a specific configuration of field particles, which is the dynamical part of the calculation, and then averaged over all possible configurations, which is the statistical part of the calculation. In the mean square field fluctuation approximation this order is interchanged. The basic assumption of this approach is that the time development of the distribution of field particles surrounding the test particle can be described as a succession of independent states. The electrostatic force, acting on the test particle, is separated into a systematic (space charge) component, corresponding to the smoothed-out distribution of charge, and a stochastic fluctuating component, depending on the instantaneous distribution of field particles. The former will be zero for a test particle which is surrounded by a spherical symmetric distribution of field particles. The latter varies randomly, and its instantaneous value during two successive independent states is assumed to show no correlation. The total effect of the fluctuations on the displacement distribution of the test particle is found by adding statistically the displacements experienced during the subsequent independent states.

In mathematical terms this approach can be expressed as follows. Apply Equation 7.31 to state number  $s$  by taking the time integral boundaries equal to the start and finish of this state. The displacement of the test particle in state number  $s$  due to the interaction with a configuration of field particles specified by  $\vec{\xi}_1, \vec{\xi}_2, \dots, \vec{\xi}_{N-1,s}$  then follows as

$$\Delta\vec{\eta}_s(\vec{\xi}_{1,s}, \vec{\xi}_{2,s}, \dots, \vec{\xi}_{N-1,s}) \approx \Delta t_s \vec{G}_{\Delta\eta}(\vec{\xi}_{1,s}, \vec{\xi}_{2,s}, \dots, \vec{\xi}_{N-1,s}) \quad (7.36)$$

where  $\Delta t_s$  is the duration of state number  $s$ , which is assumed to be short. The quantity  $\vec{G}_{\Delta\eta}$  depends on the displacement considered and can directly be related to the interaction force

$$\vec{G}_{\Delta v_z} = \frac{1}{m} \vec{F}_{||}(\vec{\xi}_{1,s}, \vec{\xi}_{2,s}, \dots, \vec{\xi}_{N-1,s}) \quad \vec{G}_{\Delta r_{\perp}} = \frac{\Delta t_s}{m} F_{\perp}(\vec{\xi}_{1,s}, \vec{\xi}_{2,s}, \dots, \vec{\xi}_{N-1,s}) \quad (7.37)$$

similar to Equation 7.32. The mean square displacement follows as

$$\langle \Delta\vec{\eta}^2 \rangle = \sum_{s=1}^{N_s} \Delta t_s^2 \sqrt{\langle \vec{G}_{\Delta\eta}^2 \rangle_s} \quad (7.38)$$

based on the assumption that the configurations of the successive states are statistically independent. The quantity  $\langle \vec{G}_{\Delta\eta}^2 \rangle_s$  is directly related to the mean square fluctuation of the electrostatic force in state number  $s$ , as can be seen from Equation 7.34.

Models along these lines for both the Boersch effect and the trajectory displacement effect were published by De Chambost and Hennion (1979), De Chambost (1982a,b), Sasaki (1984, 1986) and Massey et al. (1981). The latter model deviates from the others in the sense that the successive states are not considered as independent, but rather the opposite as fully correlated. Accordingly, Equation 7.38 is replaced by

$$\sqrt{\langle \Delta\vec{\eta}^2 \rangle} = \sum_{s=1}^{N_s} \Delta t_s \sqrt{\langle \vec{G}_{\Delta\eta}^2 \rangle_s} \quad (7.39)$$

The models of this type encounter two inherent problems. First, one is forced to introduce a lower limit for the distance between the test particle and its surrounding field particles; otherwise the mean square of the fluctuations in the electrostatic force, acting on the test particle, would diverge. The result depends critically on the value of this lower limit, but its choice seems arbitrary. Second, when using Equation 7.38 rather than Equation 7.39, additional assumptions are required to estimate the speed of the fluctuations or, in other words, the sizes of the time intervals  $\Delta t_s$ . The approximations used to solve these problems vary for the three models mentioned, leading to significant differences in the results. The models of Massey et al. and Sasaki, predict a square root dependency on the beam current, whereas the result of De Chambost and Hennion shows a 3/8 power dependency.

### 7.3.4.5 Extended Two-Particle Approximation

The extended two-particle model was first presented by van Leeuwen and Jansen (1983) and has been described in detail by Jansen (1990). We shall describe this model more extensively than the other models, as we use its results in Sections 7.4 and 7.5. Its main features are as follows:

- The  $N$ -particle problem is reduced to a two-particle problem by using the approximations expressed by Equations 7.22 and 7.23. A statistical scheme is utilized which produces the

full displacement distribution  $\rho(\Delta\vec{\eta})$  from which any width measure (e.g., rms, FW<sub>50</sub>, or FWHM) can be derived.

- The two-particle dynamical problem of evaluating  $\Delta\vec{\eta}_2(\vec{\xi})$  is treated in an exact manner. Weak collision dynamics is used for those collisions that fulfill the conditions described in Section 7.3.3. For the remaining collisions, complete collision dynamics is used when the conditions appropriate for this regime are fulfilled. A numerical calculation is carried out for the remaining collisions that are neither weak nor complete, that is, strong and incomplete. The general problem is solved numerically, while explicit analytical expressions are obtained for the limiting cases that the contribution of strong incomplete collisions can be ignored. This approach leads to results which are valid over a wide range of beam geometries and particle densities, covering all different classes of beam segments.

The most surprising result of these investigations is probably that one has to distinguish a variety of regimes, depending on the particle density as well as on the geometry of the beam. Each regime corresponds to a different type of displacement distribution. Gaussian behavior occurs for high particle densities only. For low particle densities, the interaction distributions may show distinct non-Gaussian features, such as a narrow core and long tails.

The extended two-particle approximation can, in a way, be regarded as a synthesis of the methods outlined in Sections 7.3.4.2. and 7.3.4.3:

- In case all interactions are weak, it produces the results of the first-order perturbation approximation: in both models, the  $N$ -particle problem is solved through Equations 7.22 and 7.24, while, for weak interactions, the first-order perturbation model yields identical results as the more generic two-particle dynamics of the extended two-particle model.
- In case the displacement of the test particle is dominated by a single complete collision, it resembles the closest encounter approximation: for (nearly) complete collisions, the more generic two-particle dynamics of the extended two-particle model yields identical results as the closest encounter approximation. The handling of the  $N$ -particle problem differs, as the extended two-particle model accounts for the (weak) interaction with other particles than the nearest neighbor as well. When the interaction with the nearest neighbor is strong, the calculation of the other interactions will not be accurate, as the test particle will deviate significantly from its unperturbed trajectory. However, as the total displacement of the test particle is dominated by that single strong interaction, (as assumed in the closest encounter approximation), the exact contribution of the weak collision taking place simultaneously is immaterial. Thus, for the conditions where the closest encounter approximation is justified, both methods yield identical results.

The exact two-particle analysis provides accurate results even for strong but incomplete collisions, which are not covered by the other models. The extended two-particle approximation breaks down when the test particle is involved in two or more strong collisions simultaneously or successively in the same beam segment. For these conditions the reduction of the  $N$ -particle dynamical problem of Equation 7.24 becomes inaccurate.

### 7.3.5 DISPLACEMENT DISTRIBUTION IN THE EXTENDED TWO-PARTICLE APPROXIMATION

The equations for the interaction effects which are given in this chapter are based on the extended two-particle approximation. Therefore, we shall, in this section, follow the general lines of the derivation of a displacement distribution function in this theory. This naturally leads to a discussion of the different types of distribution functions that can be expected. The reduction of the  $N$ -particle

problem in the extended two-particle approximation is based on Equations 7.22 and 7.24. Substitution in Equation 7.21 yields

$$\rho(\Delta\vec{\eta}) = \int d\vec{\xi}_1 d\vec{\xi}_2 \dots d\vec{\xi}_{N-1} \prod_{i=1}^{N-1} P_2(\vec{\xi}_i) \delta\left(\Delta\vec{\eta} - \sum_{i=1}^{N-1} \Delta\vec{\eta}_2(\vec{\xi}_i)\right) \quad (7.40)$$

which is the  $N$ -particle displacement distribution on the binary interaction approximation. To simplify the computation, we would like to separate the two-particle problem of calculating the displacement caused by a single field particle from the full  $N$ -particle statistics. For this purpose, we define the two-particle displacement distribution as

$$\rho_2(\vec{\eta}) = \int d\vec{\xi} p(\vec{\xi}) \delta\left(\Delta\vec{\eta} - \Delta\vec{\eta}_2(\vec{\xi})\right) \quad (7.41)$$

where

$$p(\vec{\xi}) = LP_2(\vec{\xi}) = \frac{N}{\lambda} P_2(\vec{\xi}) \quad (7.42)$$

with the linear particle density  $\lambda = I/eV_z$ . The function  $\rho(\Delta\vec{\eta})$  represents the displacement distribution for a system containing only the test particle and one field particle. It is normalized to  $L$  rather than to 1 for the purpose of isolating the dependency on  $N$ .

One might anticipate that the full  $N$ -particle displacement distribution  $\rho(\Delta\vec{\eta})$  can somehow be represented as a convolution of the two-particle distributions  $\rho(\Delta\vec{\eta})$  of all field particles. The usual method to handle convolutions is to multiply the Fourier transforms and evaluate the inverse Fourier transform of the resulting expression. A rigorous treatment of the problem, as given by Jansen (1990), shows that the appropriate Fourier type of transformation to the spatial frequency space, or  $k$ -space, has the following form:

$$P(\vec{k}) = \int d(\Delta\vec{\eta}) \rho_2(\Delta\vec{\eta}) (1 - e^{-i\vec{k} \cdot \Delta\vec{\eta}}) \quad (7.43)$$

which differs slightly from the usual Fourier transform. An essential feature of this form is that the term between brackets vanishes for small displacement  $\Delta\vec{\eta}$ . Similarly, the Fourier transform of the  $N$ -particle displacement distribution  $P_N(\vec{k})$  is not simply equal to the power  $N - 1$  of the two-particle transform  $p(k)$ , as would be the case for a straightforward convolution of  $N - 1$  identical distributions. But, rather

$$P_N(\vec{k}) = \left[ 1 - \frac{\lambda}{N} p(\vec{k}) \right]^{N-1} \approx e^{-\lambda p(\vec{k})} \quad (7.44)$$

The last equality in Equation 7.44 requires  $N$  to be large. The inverse Fourier transform of  $P_N(\vec{k})$  yields the  $N$ -particle displacement distribution  $\rho(\Delta\vec{\eta})$ :

$$\rho(\Delta\vec{\eta}) = \frac{1}{(2\pi)^n} \int d\vec{k} e^{-\lambda p(\vec{k})} e^{i\vec{k} \cdot \Delta\vec{\eta}} \quad (7.45)$$

where  $n$  is the dimension (or number of components) of the displacement  $\Delta\vec{\eta}$ . Thus  $n = 1$  for the Boersch effect ( $\Delta\eta = \Delta v_z$ ), and  $n = 2$  for the trajectory displacement effect ( $\Delta\eta = \Delta r$ ).

The key relations of the model have been expressed in vector notation so far, using the coordinates vector  $\vec{\xi}$ , displacement vector  $\Delta\vec{\eta}$ , and the spatial frequency vector  $\vec{k}$ . We will now give

the equivalent scalar form for  $n = 1$  and  $n = 2$ , which are the two most important cases. For the Boersch effect ( $n = 1$ ,  $\Delta\eta = \Delta v_z$ ), Equations 7.41, 7.43, and 7.45 can be rewritten as

$$\begin{aligned}\rho_2(\Delta v_z) &= \int_0^{v_0} dv \frac{2v}{v_0^2} \int_0^{2\pi} d\Phi \int_0^{r_c} dr \frac{2r_\perp}{r_c^2} \int_{-S_c L}^{(1-S_c)L} db_z \delta[\Delta v_z - \Delta v_z(v, \Phi, r_\perp, b_z)] \\ p(k) &= 2 \int_{-\infty}^{\infty} d\Delta v_z \rho_2(\Delta v_z) \sin^2(k\Delta v_z/2) \\ \rho(\Delta v_z) &= \frac{1}{\pi} \int_{-\infty}^{\infty} dk \cos(k\Delta v_z) e^{-\lambda p(k)}\end{aligned}\quad (7.46)$$

where the integral over  $\vec{\xi}$  was expressed in terms of the experimental parameters using Equations 7.23 and 7.42, which pertain to a beam segment with a crossover with uniform angular and spatial distribution. In the derivation of the expression for  $p(k)$  we used that the average displacement  $\langle \Delta v_z \rangle$  will vanish for the central reference trajectory. Similarly, one finds for the trajectory displacement effect ( $n = 2$ ,  $\Delta\eta = \Delta r$ ),

$$\begin{aligned}\rho_2(\Delta r) &= 2\pi\Delta r \rho_2(\Delta \vec{r}) = \int_0^{v_0} dv \frac{2v}{v_0^2} \int_0^{2\pi} d\Phi \int_0^{r_c} dr \frac{2r_\perp}{r_c^2} \int_{-S_c L}^{(1-S_c)L} db_z \delta[\Delta r - \Delta r(v, \Phi, r_\perp, b_z)] \\ p(k) &= \int_0^x d\Delta r \rho_2(\Delta r) [1 - J_0(k\Delta r)] \\ \rho(\Delta r) &= 2\pi\Delta r \rho(\Delta \vec{r}) = \Delta r \int_0^{\infty} k dk J_0(k\Delta r) e^{-\lambda p(k)}\end{aligned}\quad (7.47)$$

where  $J_0$  is the Bessel function of zero order. The derivation of Equation 7.47 utilizes the integral representation

$$J_0(x) = \frac{1}{2\pi} \int_0^{2\pi} d\phi e^{ix\cos(\phi)}\quad (7.48)$$

Performing this three-step approach with the full collision dynamics is not possible analytically. When it is done numerically, one finds that there are parameter regimes for which the function  $\lambda p(k)$  in Equations 7.46 and 7.47 can be approximated by the simple relationship

$$\lambda p(k) = A_\gamma k^\gamma\quad (7.49)$$

where the exponent  $\gamma$  is a numerical constant and  $A_\gamma$  represents some function of the experimental parameter. For small  $k$ -values,  $p(k)$  is always quadratic in  $k$ . This can be understood from its definition in terms of  $\rho_z(\Delta \vec{\eta})$ , specified in Equations 7.43, 7.46, or 7.47, by taking a Taylor expansion of the integrand for small  $k$  and using the fact that the average displacement  $\langle \Delta \vec{\eta} \rangle$  vanishes for the central reference trajectory; the  $k^2$ -term is the first nonvanishing term in the Taylor expansion in  $k$ . Small  $k$ -values correspond to large displacement  $\Delta \vec{\eta}$  and are thus relevant both for the tails of the displacement distribution, whatever the average displacement is, and for the regime in which large displacements are dominant. Analyzing which kind of collisions contribute to the quadratic part of the  $p(k)$  transform, it turns out that all types (weak incomplete, weak complete, and strong)

**TABLE 7.1**  
**Different Regimes and Displacement Distributions**

Beam Geometry		Dominant Collision Type			Dimension 1( $\Delta E$ )/2( $\Delta \alpha, \Delta r$ )	$\gamma$	Type of Distribution
Pencil (p)/ Extended (e)	Crossover (x)/ Cylindrical (c)	Weak (w)/ Strong (s)	Complete (c)/ Incomplete (i)				
p	x, c	w	I	1( $\Delta E$ )	1/2	1 D Pencil	
p	x, c	w	I	2( $\Delta \alpha, \Delta r$ )	1/3	2 D Pencil	
e	x, c	w	I	1, 2( $\Delta E, \Delta \alpha, \Delta r$ )	3/2	Holtsmark	
e	x	w	c	1( $\Delta E$ )	1	Lorentz	
e	x	w	c	2( $\Delta \alpha$ )	1/2	Weak complete	
e	x, c	w, s	c, i	1, 2( $\Delta E, \Delta \alpha, \Delta r$ )	2	Gauss	

contribute, and that strong collisions are significant. When  $p(k)$  is quadratic for the full  $k$ -range, the beam must be of high density and of the three-dimensional type (*extended*); in that situation there are many independent collisions.

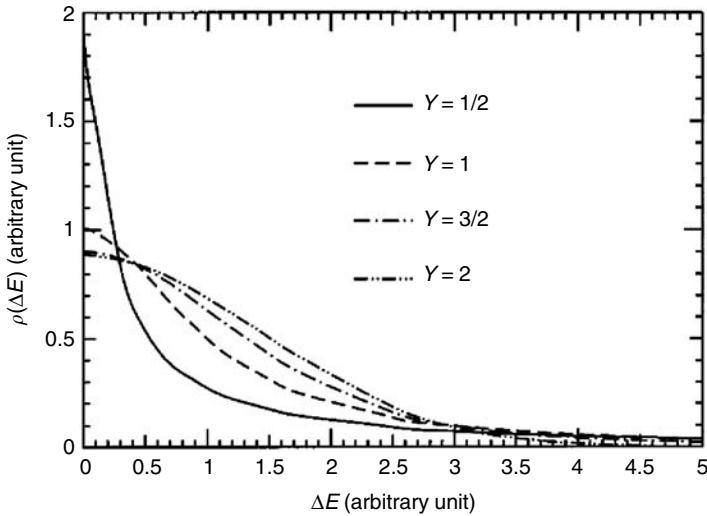
In the regimes where  $\gamma \neq 2$ , it is possible to connect a different interaction mechanism to each  $\gamma$  value. Then we can also reason the other way around: if we specify one statistical situation and approximate the dynamical problem within one type of collision, Equations 7.46 and 7.47 can be partly solved analytically and the solution shows that the function  $\lambda p(k)$  can indeed be dominated by the form of Equation 7.49. The statistical problem is specified by the choice between a cylindrical (c) segment or a segment with a crossover (x) and by the density of the particles which yield either a pencil beam (p) or an extended beam (e). The dynamical problem is specified by the distinction between weak (w) and strong (s) collisions and between complete (c) and incomplete (i) collisions. Finally, the three-step approach distinguishes between mathematics for dimension 1 (Equation 7.48) and dimension 2 (Equation 7.47). Strong collisions are always related to a  $k^2$  dependence, leaving us with 16 perturbations of the other four choices. However, some combinations of these four parameters do not occur in reality and quite a few lead to the same value of  $\gamma$ . For many of these cases, the analytical distribution function was derived by Jansen (1990). Here we indicate each combination of parameters by a five digit code  $\nabla$  which stands for either of the possibilities of that digit.

In a pencil beam, the dominant type of collision is always weak and incomplete, thus only the combinations pxwil, pcwil, pxwi2, and pcwi2 exist. (p $\nabla$ wil) leads to  $\gamma = 1/2$ , (p $\nabla$ wi2) leads to  $\gamma = 1/3$ . In an extended beam, when the dominant type of collision is weak and incomplete, we find  $\gamma = 3/2$  both for cylindrical and crossover beams, both for the Boersch effect and the trajectory displacement, thus for all combinations (e $\nabla$ wi $\nabla$ ).

In an extended beam, the displacement of a weak complete collision behind the crossover compensates the displacement before the crossover. Thus, one shall only find Boersch effect and angular deflection, but no displacement at the position of the crossover in the regime. For the Boersch effect (exwc1),  $\gamma = 1$ . For the angular deflection (exwc2),  $\gamma = 1/2$ . Table 7.1 gives an overview of the occurrence of different  $\gamma$  values.

The representation of the distributions in Equations 7.46 and 7.47 now provides a convenient means to calculate the shape of the displacement function. For instance, for  $\gamma = 2$  one finds for the distributions  $\rho(\Delta v_z)$  and  $\rho(\Delta \vec{r})$  by integration of the corresponding Equations 7.46 and 7.47,

$$\begin{aligned} \rho(\Delta v_z) &= \frac{1}{\sqrt{4\pi A_2}} e^{-\Delta v_z^2/4A_2} \\ \rho(\Delta \vec{r}) &= \frac{1}{4\pi A_2} e^{-\Delta \vec{r}^2/4A_2} \end{aligned} \quad (7.50)$$



**FIGURE 7.6** Distribution functions for the Boersch effect.

which are the one- and two-dimensional Gauss distributions, respectively, with  $\sigma = \sqrt{2A_2}$ . Similarly, one finds for  $\gamma = 1$ ,

$$\begin{aligned}\rho(\Delta v_z) &= \frac{1}{\pi A_i \left[ 1 + (\Delta v_z^2 / A_i)^2 \right]} \\ \rho(\Delta \vec{r}) &= \frac{1}{2\pi A_i^2 \left[ 1 + (\Delta r^2 / A_i)^2 \right]^{3/2}}\end{aligned}\quad (7.51)$$

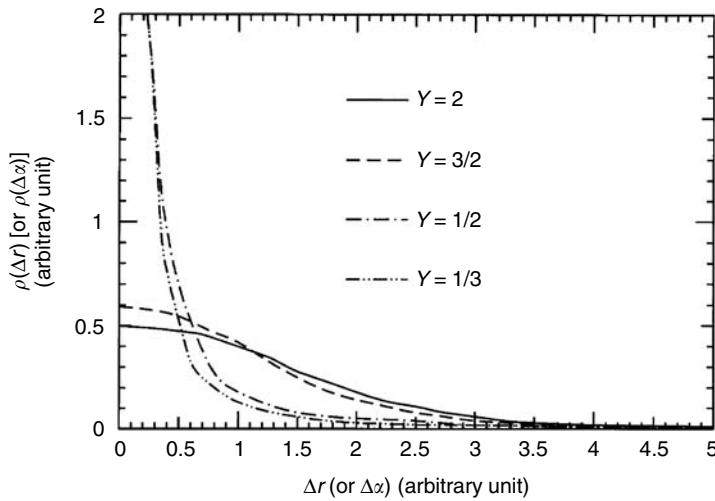
which are the one- and two-dimensional Lorentz distributions, respectively. A Holtsmark distribution corresponds to  $\gamma = 3/2$ . For  $\gamma$  values other than 1 and 2 it is not possible to give an explicit distribution function as in Equations 7.50 and 7.51, so one has to fall back on the integral expressions of Equations 7.46 and 7.47.

The different types of displacement distributions are depicted in Figure 7.6, for the one-dimensional case ( $n = 1$ ), and in Figure 7.7, for the two-dimensional case ( $n = 2$ ), which pertain to the Boersch effect and the trajectory displacement effect, respectively. The figures show that a small  $\gamma$ -value corresponds to a distribution with a narrow core and long tails.

In the approximation of Equation 7.49, the  $FWHM$  and  $FW_{50}$  of the displacement distribution  $\rho(\Delta \vec{r})$  are given by

$$\begin{aligned}FWHM_{\Delta \eta} &= A_{\gamma}^{1/\gamma} FWHM_{\gamma} \\ FW_{50\Delta \eta} &= A_{\gamma}^{1/\gamma} FW_{50\gamma}\end{aligned}\quad (7.52)$$

where  $FWHM_{\gamma}$  and  $FW_{50\gamma}$  represent numerical factors depending on  $\gamma$  only. These values are given in Table 7.2 for  $n = 1$  and  $n = 2$ . Equations 7.52 express the  $FWHM$  and  $FW_{50}$  values in terms of the experimental parameters, represented by  $A_{\gamma}$ , and the parameter  $\gamma$ . The parameter  $A_{\gamma}$  is for all conditions directly proportional to the beam current  $I$ , as can be seen from Equation 7.49, using that  $\lambda$  is directly proportional to  $I$ . Accordingly, the  $FWHM$  and  $FW_{50}$  values are, for both the Boersch effect and the trajectory displacement effect, proportional to  $I^{1/2}$  in the Gaussian regime



**FIGURE 7.7** Distribution functions for the trajectory displacement and angular deflection effect.

**TABLE 7.2**  
**Full Width at Half Maximum ( $FWHM$ ) and Full Width Median ( $FW_{50}$ ) Proportionality Constants as a Function of  $\gamma$  for the One-Dimensional ( $n = 1$ ) and Two-Dimensional ( $n = 2$ ) Displacement Distributions**

$\gamma$	$FWHM_\gamma$		$FW_{50\gamma}$	
	$n = 1$	$n = 2$	$n = 1$	$n = 2$
1/3	0.063581	0.013013	3.5751	6.8096
1/2	0.44711	0.17997	2.5677	4.7478
1	2	1.5328	2	4.5289
3/2	2.8775	2.6554	1.9379	3.3206
2	3.3302	3.3302	1.9078	3.3302

( $\gamma = 2$ ), proportional to  $I^{2/3}$  in the Holtsmark regime ( $\gamma = 3/2$ ), and linearly proportional to  $I$  in the Lorentzian regime ( $\gamma = 1$ ). For pencil beams the Boersch effect increases with  $I^2$  ( $\gamma = 1/2$ ) and the trajectory displacement effect increases with  $I^3$  ( $\gamma = 1/3$ ). Notice that, vice versa, one may derive the shape of the distribution from the observed current dependency.

### 7.3.6 ADDITION OF DISPLACEMENT IN SEVERAL BEAM SEGMENTS

To derive the properties of the probe at the target, the problem of adding the statistical interaction effects generated in the individual beam segments needs to be considered. A key issue is the degree of correlation between the interaction distributions generated in the successive beam segments. We will consider the following two extreme cases:

- *No correlation.* The optical elements between the beam segments effectively redistribute the configuration of particles destroying any correlation between them. This is expected to be true for beam segments with narrow crossovers and relatively high particle densities.

- *Full correlation.* The configuration of particles remains the same in the passage to the next beam segment. This is expected to be true for nearly cylindrical beam segments, especially at low particle densities.

For the case of *no correlation* it can be proved that the total interaction distribution generated in  $N_b$  successive beam segments can be represented in the  $k$ -domain as (Jansen, 1990)

$$\left[ \lambda p(\vec{k}) \right]_T = \sum_{j=1}^{N_b} \lambda_j p_j(\vec{k}) \quad (7.53)$$

where  $\lambda_j$  and  $p_j(\vec{k})$  refer to beam segment number  $j$ . Effectively, the distribution functions are convoluted. When the individual displacement distributions are all of the same type (same  $\gamma$ -value), one finds

$$FW_T = \left[ \sum_{j=1}^{N_b} FW_j^\gamma \right]^{1/\gamma} \quad (7.54)$$

which is referred to as the  $\gamma$ -rule for the summation of beam segments. The quantity  $FW_j$  denotes an arbitrary width measure (e.g., the  $FWHM$  or  $FW_{50}$ ) of the displacement distribution generated in the beam segment  $j$  ( $j = 1, 2, \dots, N_b$ ). If the distributions are *not all of the same type*, one can use either Equation 7.53 when the  $p(k)$  functions are known or an effective  $\gamma$ -value in Equation 7.54 calculated as an average of the individual  $\gamma$ -values weighted with the size of the corresponding contributions  $FW_j$ .

For the case of *full correlation* the total effect should be calculated as a linear sum of the individual contributions

$$FW_T = \sum_{j=1}^{N_b} FW_j \quad (7.55)$$

which follows from the fact that the displacements of the test particle generated by each field particle in the successive beam segments must be added together. See also the discussion in Mkrtchyan et al. (1995).

In reality, the distributions generated in different beam segments may be partially correlated and an exact solution cannot be obtained, as this would require knowledge of the detailed particle configurations. In that case, Equations 7.54 and 7.55 may serve as a lower and upper boundary of the total effect. Note that the two summation rules would yield identical results in case of a Lorenz distribution ( $\gamma = 1$ ). The  $\gamma$ -summation rule will yield larger values than the linear summation rule for  $\gamma < 1$ , while the opposite applies for  $\gamma > 1$ . MC simulations may serve to verify the results and to determine the degree of correlation between the displacements generated in the individual beam segments (see Section 7.6).

## 7.4 ANALYTICAL EXPRESSIONS FOR TRAJECTORY DISPLACEMENT

### 7.4.1 PARAMETER DEPENDENCIES WHEN THE COLLISIONS ARE WEAK AND INCOMPLETE

The dependency of the trajectory displacement on the experimental parameters can be predicted on the basis of some elementary physical arguments. The resulting equations are the same as those found by the rigorous approach of Section 7.3, except for some missing numerical prefactors and, of

course, not obtaining a distribution function. However, having derived these equations in this manner provides insight into when to apply the different equations for the different regimes and when a certain approximation may be applied.

We shall first derive equations for a cylindrical beam and then approach a general segment as a sequence of short cylinders. The experimental parameters are beam current  $I$ , acceleration  $V$ , particle mass  $m$ , charge  $e$  and the beam geometry. First, assume that the particle density is low enough to assure that collisions are weak and incomplete, and thus that the deviations from the unperturbed trajectories are small. Second, assume that the particles are initially at rest in the moving reference frame. The displacement of a test particle in a cylindrical beam can then be computed from first-order perturbation dynamics and follows from the induced transverse velocity

$$\Delta v_{\perp} = \frac{F_{\perp} T}{m} \quad (7.56)$$

where  $F_{\perp}$  is the combined force from all other particles in the beam perpendicular to the axis and  $T$  the flight time through the segment under consideration. If the reference particle is taken to be on the axis of a rotational symmetric beam, the space charge force is zero and the calculated displacement is purely statistical. The angular displacement is found from

$$\Delta\alpha = \frac{\Delta v_{\perp}}{v_z} \quad (7.57)$$

in which  $v_z$  is the particle velocity along the axis. The spatial displacement  $\Delta r$  in the image plane is given by

$$\Delta r = \Delta\alpha |z - z_i| \quad (7.58)$$

in which  $|z - z_i|$  is the distance between the location where the deflection  $\Delta\alpha$  occurs and the image plane.

Take the radius and length of the cylindrical beam segment to be  $r_o$  and  $L$ , respectively. The flight time is now  $T = L/v_z$ . The force component  $F_{\perp}$  will scale with  $C_o d_r/d^3$ , where  $d_r$  and  $d$  denote the average radial distance and the averaged distance between neighbor particles, and  $C_o = e^2/4\pi\varepsilon_o$ . For the angular displacement  $\Delta\alpha$ , one finds

$$\Delta\alpha \sim \frac{C_o}{m} \frac{L d_r}{v_z^2 d^3} \quad (7.59)$$

For an extended beam (where the lateral dimension  $r_o$  is large relative to the separation of the particles), both  $d_r$  and  $d$  will scale with  $n_o^{-1/3}$ , where  $n_d$  is the particle density given by

$$n_d = \frac{1}{e\pi r_o^2 v_z} \quad (7.60)$$

Accordingly, one finds for the conditions of the Holtsmark regime using that  $v_z = (2eV/m)^{1/2}$ ,

$$\Delta\alpha_e C_{HA} \cdot \frac{m^{1/3}}{\epsilon_o} \cdot \frac{I^{2/3}L}{V^{4/3}r_o^{4/3}} \quad (7.61)$$

The full calculations outlined in Section 7.3 yields  $C_{HA} = 0.102$  for the FWHM value of the distribution. To find the spatial displacement for a segment which is really a cylindrical beam, one should realize that a cylindrical beam segment is usually followed by a lens which focuses the beam in its back focal plane. Let  $f$  be the focal distance of the lens; then

$$\Delta r = f\Delta\alpha \quad (7.62)$$

For a pencil beam (where the lateral dimension  $r_o$  is small compared with the separation of the particles),  $d$  scales with  $\lambda^{-1}$ , where  $\lambda$  is the linear particle density ( $\lambda = 1v_z/e$ ). However, the radial distance  $d_r$  scales with the beam radius  $r_o$ , resulting in

$$\Delta\alpha_p = C_{PA} \cdot \frac{m^{3/2}}{\epsilon_o e^{7/2}} \cdot \frac{I^3 L r_o}{V^{5/2}} \quad (7.63)$$

The full calculation yields  $C_{PA} = 8.31 \times 10^{-4}$  for the FWHM value. In the extreme case in which all particles are on a line ( $r_o = 0$ ), there is no transverse component of the interaction force and  $\Delta\alpha$  becomes zero.

For the Gaussian regime, it is more difficult to derive a simple equation for the angular displacement in a cylindrical beam. Note that in the Holtsmark regime and the pencil beam regime, the angular displacement is proportional to the length of the cylindrical segment. Thus, one can subdivide a beam of complicated shape into thin cylindrical slices and calculate the total displacement as a linear sum of the contributions of the individual slices. This is the basis for Jiang's program ANALIC (Jiang, 1996).

For a beam segment with a narrow crossover of radius  $r_c$  with a beam semiangle  $\alpha_o$  and length  $L$ , the slice method works as follows. Define the axial coordinate of the crossover as  $z = 0$ . The beam radius at position  $z$  is then approximately given by

$$r_o(z) \approx r_c + \alpha_o |z| \quad (7.64)$$

The virtual displacement  $\Delta r$  in the crossover (i.e., the displacement as seen when this crossover is imaged downstream from the segment under consideration) is, for an extended beam,

$$\Delta r_e = \int_{z=-L_1}^{z=L-L_1} \frac{m^{1/3}}{\epsilon_o} \cdot \frac{I^{2/3} dz}{V^{4/3} (r_c + \alpha_o |z|)^{4/3}} \cdot |z| \quad (7.65)$$

If we take  $r_c \ll \alpha_o L_1$  and  $r_c \ll \alpha_o (L - L_1)$ , this simplifies to Equation 7.69 in the next section. For the situation in which the crossover is not relatively narrow, we refer to Jansen (1990). Similarly, one finds for a pencil beam Equation 7.70.

Note that the dependency on the transverse dimension (now  $\alpha_o$ ) reverses going from an extended beam to a pencil beam. In an extended beam, the trajectory displacement is minimized by increasing the distance between particles. In a pencil beam, the displacement is minimized by bringing the particles exactly on one line.

### 7.4.2 SUMMARY OF EQUATIONS FOR TRAJECTORY DISPLACEMENT

The theory gives the characteristic width values within each regime. To arrive at an expression for the entire range of parameters, one can choose the exponents in an interpolation equation. Numerical calculations show that a good fit is obtained with

$$\frac{1}{FW_{50}} = \left[ \left[ \left( \frac{1}{FW_{50G}} \right)^6 + \left( \frac{1}{FW_{50H}} \right)^6 \right]^{1/7} + \left( \frac{1}{FW_{50P}} \right)^{6/7} \right]^{7/6} \quad (7.66)$$

This type of interpolation relies on the assumption that the borders of the individual regimes coincide with the parameters for which the  $FW_{50}$  values of two adjoining regimes are equal. The interpolation between the pencil beam regime and the Holtsmark regime is approximated quite well by

$$\frac{1}{FW_{50}} = \frac{1}{FW_{50H}} + \frac{1}{FW_{50P}} \quad (7.67)$$

while the interpolation between the Holtsmark and the Gaussian regime is correctly dealt with by simply taking the smallest value of the  $FW_{50}$ .

The  $FW_{50}$  values in the Gaussian, Holtsmark and pencil beam regimes are, respectively, given by

$$FW_{50G} = 0.158 \frac{m^{1/4} e^{1/12}}{\varepsilon_o^{5/6}} \cdot \left[ \frac{1.27 S_c^{6/5} - 0.0025}{\left( 1 + \frac{2e^{4/9}}{(2\pi\varepsilon_o)^{4/9} \alpha_o^{4/3} V^{4/9} L^{4/9}} \right)^{1/2}} + \frac{0.95 - S_c}{1 + \frac{4e^{1/3}}{(2\pi\varepsilon_o)^{1/3} \alpha_o V^{1/3} L^{1/3}}} \right] \cdot \frac{I^{1/2} L^{2/3}}{\alpha_o V^{13/12}} \quad (7.68)$$

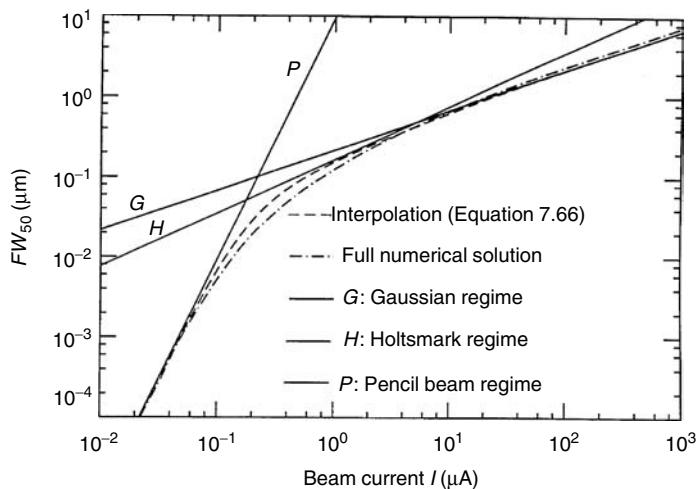
$$FW_{50H} = 0.172 \frac{m^{1/3}}{\varepsilon_o} \cdot [S_c^{2/3} + (1 - S_c)^{2/3}] \cdot \frac{I^{2/3} L^{2/3}}{V^{4/3} \alpha_o^{4/3}} \quad (7.69)$$

$$FW_{50P} = 0.145 \frac{m^{3/2}}{e^{7/2} \varepsilon_o} \cdot [S_c^3 + (1 - S_c)^3] \cdot \frac{I^3 L^3}{V^{5/2}} \quad (7.70)$$

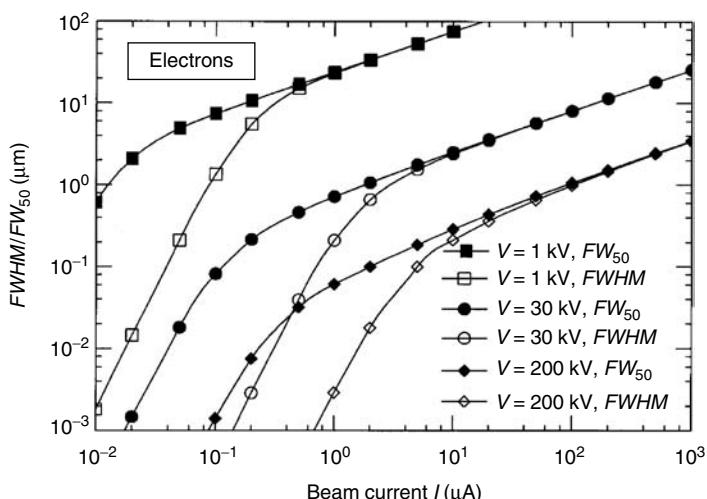
The numerical prefactors for the  $FWHM$  values are 0.158, 0.137 and  $2.77 \times 10^{-4}$ , respectively, for the Gaussian, Holtsmark and pencil beam equations, as tabulated in Table 7.2. Note that the pencil beam value is dramatically different. These equations are valid only for the displacement in the crossover of a beam with narrow crossover ( $K_1$  and  $K_2 \gg 1$ ) with a homogeneous current distribution. Equations 7.61 and 7.63 are the equivalent equations for a cylindrical beam segment. See Jansen (1990) for other configurations.

### 7.4.3 NUMERICAL EXAMPLES

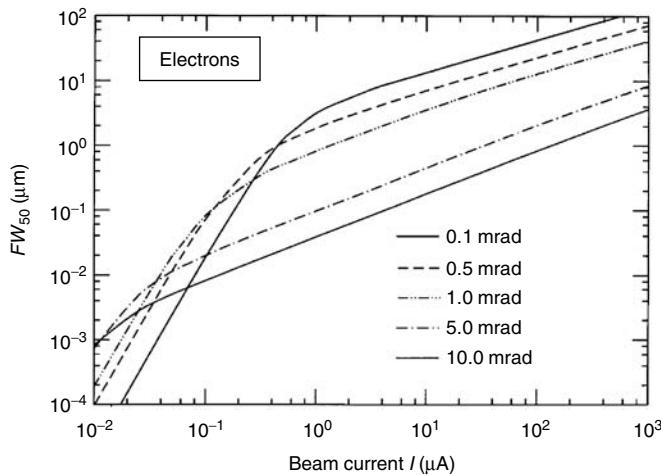
To give an impression of the order of magnitude of the trajectory displacement, we have calculated the effect for a segment with a narrow crossover as one would typically find in an electron microscope or a focused ion beam instrument. Those instruments usually have a short focal distance lens to form a probe on the sample. This lens demagnifies an intermediate crossover by a factor between 3 and 20. The aperture angle at the sample is in the order of 10 mrad, limited by the lens aberrations. Before that final lens is a section  $\sim$ 10 to 20 cm long, often with a crossover and an aperture angle  $\sim$ 1 mrad. It is this section that is taken as an example. Figure 7.8 shows how well the interpolation of Equation 7.66 approximates the  $FW_{50}$  values found from the full numerical solution of the distribution functions. Figures 7.9 and 7.10 show the trajectory displacement according



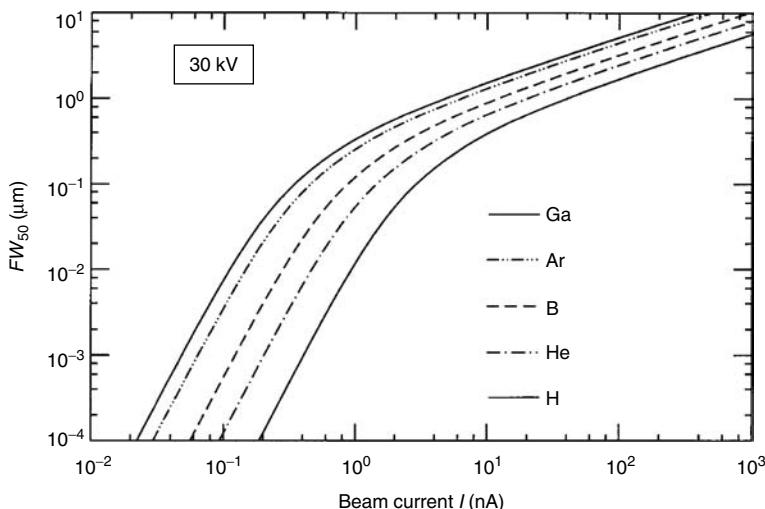
**FIGURE 7.8** Interpolation between  $FW_{50}$  values of trajectory displacement in different regimes. Numerical values are given for an electron beam segment of length 200 mm with a narrow crossover at 50 mm where the half angle is 1 mrad. The acceleration voltage is 100 kV.



**FIGURE 7.9** Trajectory displacement for electrons in a 200-mm long segment with a narrow crossover at 50 mm, half angle 1 mrad for acceleration voltages 1, 30, and 200 kV.



**FIGURE 7.10** Trajectory displacement for different values of the beam half angle in a 200-mm long segment with a narrow crossover at 50 mm, at an acceleration voltage of 30 kV.



**FIGURE 7.11** Trajectory displacement for ions in a 200-mm long segment with a narrow crossover at 50 mm, half angle 1 mrad. Ion masses 1(H), 4(He), 11(B), 40(Ar) and 70(Ga) are shown.

to Equations 7.66, 7.68, 7.69, and 7.70 for different electron energies. Figure 7.11 shows the trajectory displacement for ions. The broadening is given in the plane of the crossover inside the segment. So, the influence of the part of the segment behind the crossover is traced back to that plane.

## 7.5 ANALYTICAL EXPRESSIONS FOR THE BOERSCH EFFECT

### 7.5.1 PARAMETER DEPENDENCIES WHEN THE COLLISIONS ARE WEAK

The energy broadening dependency on the experimental parameters  $I$ ,  $V$ ,  $m$ ,  $e$ , and the beam geometry can be predicted through the same reasoning that we followed for the trajectory displacement

in the previous section. The change in forward velocity of a reference particle in first-order perturbation dynamics is, under the conditions stated in Section 7.4.1,

$$\Delta v_z = \frac{F_z \cdot T}{m} \quad (7.71)$$

where  $F_z$  is the combined force from all other particles in the beam, and  $T$  is the flight time given by  $T = L/v_z$ . Using this change in velocity as characteristic for the induced energy spread presupposes that the particles are initially at rest relative to each other. In the full theory of the Boersch effect this assumption has proved to give errors  $< 10\%$ . Consider a beam segment in which the particles interact weakly. The force component  $F_z$  will scale with  $C_o d_z / d^3$  where  $d_z$  and  $d$  denote the average axial distance and the average distance between neighbor particles and  $C_o = e^2 / 4\pi\epsilon_o$ . In a pencil beam, where the distance between particles is large compared with the radius, both  $d_z$  and  $d$  are equal to

$$d = \frac{e}{I} \cdot v_z \quad (7.72)$$

The energy spread then follows from

$$\frac{\Delta E_P}{E} = 2 \frac{\Delta v_z}{v_z} = C_{PE} \cdot \frac{m}{\epsilon_o e^2} \frac{I^2 L}{V^2} \quad (7.73)$$

The full calculation yields  $C_{PE} = 0.112$  for the *FWHM* value. As  $r_o$  does not appear in this equation, it holds both for a cylindrical beam and for a beam with a crossover. For an extended beam, both  $d$  and  $d_z$  scale with  $n_d^{-1/3}$ , where  $n_d$  is the particle density given by

$$n_d = \frac{1}{e\pi r_o^2 v_z} \quad (7.74)$$

Consider a cylindrical beam segment with beam radius  $r_o(z)$  and length  $L$ . The contribution to the energy spread from this segment is found directly from Equation 7.71:

$$\frac{\Delta E_H}{E} = C_{HE} \cdot \frac{m^{1/2}}{\epsilon_o} \frac{I^{2/3} L}{V^{4/3} r_o(z)^{4/3}} \quad (7.75)$$

The full calculation yields  $C_{HE} = 0.221$  for the *FWHM* value. For a beam segment with a narrow crossover,  $r_o(z)$  can be expressed as  $r_o(z) = r_c + \alpha_o |z|$  and the full energy spread is found by integrating over  $z$ . For  $r_c \ll \alpha_o L$  the integral yields Equation 7.80.

In the following analysis we consider complete collisions occurring in the crossover at particle distances that are large enough that the trajectories can still be considered essentially undisturbed. We assume that the  $z$ -component of the impact parameter  $b$  is large compared with the perpendicular distance  $r_\perp$  between the test particle and the field particle. The physical model is that the trajectory of a field particle crosses the axis inside the crossover and in passing the axis, the particle pushes the test particle which travels on the axis. This gives a velocity change,

$$\Delta v_z \cong \frac{2C_o}{m \times v_z b_z} \quad (7.76)$$

where  $\alpha$  is the angle between the field particle and the axis, and will scale with the beam semiangle  $\alpha_o$ . Accordingly, one finds

$$\frac{\Delta E_L}{E} = C_{LE} \cdot \frac{m^{1/2}}{\varepsilon_o e^{1/2}} \frac{I}{\alpha_o V^{3/2}} \quad (7.77)$$

The full calculation yields  $C_{LE} = 1.41$  for the FWHM value. As for the angular displacement, it is more difficult to derive the parameter dependency for the Gaussian regime. In the Holtsmark and pencil beam regimes the energy spread is proportional to the length of the cylindrical beam segment, giving the possibility of applying the slice method.

### 7.5.2 SUMMARY OF EQUATIONS FOR THE BOERSCH EFFECT

The theory which results in equations for the distribution of forward velocities ( $\rho(v_z)$ ) is similar to the theory for trajectory displacements, as outlined in Section 7.3. Again, it is impossible to find complete analytical solutions for all values of the beam parameters. Hence, in many situations an analytical function is fitted to the results of numerically solved integrals. From the full expressions for the distribution, one can derive the values of the  $FWHM$  and the  $FW_{50}$ , which are presented here. As discussed in Section 7.3, it is necessary to distinguish between four regimes: Gaussian, Holtsmark, Lorentzian, and pencil beam. For the calculation, it is not necessary to know in which regime the beam parameters are. By combining the results from the individual regimes, one obtains an expression which is valid for all regimes:

$$\frac{E}{\Delta E_{FW_{50}}} = \left[ \frac{E^4}{\Delta E_{FW_{50G}}^4} + \frac{E^4}{\Delta E_{FW_{50H}}^4} + \frac{E^4}{\Delta E_{FW_{50L}}^4} + \frac{E^4}{\Delta E_{FW_{50P}}^4} \right]^{1/4} \quad (7.78)$$

For very narrow crossovers, the  $FW_{50}$  equations are relatively simple:

$$\frac{\Delta E_{FW_{50G}}}{E} = 0.802 \frac{m^{1/4}}{\varepsilon_o^{1/2} e^{1/4}} \frac{I^{1/2}}{V^{3/4}} \quad \left( r_e \leq \frac{e}{8\varepsilon_o} \cdot \frac{1}{\alpha_o^2 V} \right) \quad (7.79)$$

$$\frac{\Delta E_{FW_{50H}}}{E} = 0.891 \frac{m^{1/3}}{\varepsilon_o} \frac{I^{2/3}}{r_c^{1/3} \alpha_o V^{4/3}} \quad (r_c \ll \alpha_o L) \quad (7.80)$$

$$\frac{\Delta E_{FW_{50L}}}{E} = 1.41 \frac{m^{1/2}}{\varepsilon_o e^{1/2}} \frac{I}{\alpha_o V^{3/2}} \quad (7.81)$$

$$\frac{\Delta E_{FW_{50P}}}{E} = 0.642 \frac{m}{\varepsilon_o e^2} \frac{I^2 L}{V^2} \quad (7.82)$$

For the  $FWHM$  values the respective numerical prefactors are 1.40, 1.32, 1.41, and 0.112. In all regimes except the pencil beam regime, most of the contribution to the energy spread comes predominantly from the crossover. If the crossover is not extremely narrow, the energy broadening

in the Gaussian and Holtsmark regimes is less than that given in Equations 7.79 and 7.80. For the Gaussian regime an extended equation can be given, still assuming  $K_1$  and  $K_2 \gg 1$  as

$$\frac{\Delta E_{FW50G}}{E} = 0.742 \frac{m^{1/4}}{E_o^{1/2} e^{1/4}} \left[ 1 + \frac{\pi \vec{r}_c}{\left[ 2 \ln(0.8673(114.6 + \vec{r}_c)) \right]^2} \right]^{-1/2} \frac{I^{1/2}}{V^{3/4}} \quad (7.83)$$

with

$$\vec{r}_c = \frac{8\pi\epsilon_o}{e} \alpha_o^2 V r_c$$

Equations 7.73, 7.75, and 7.77 are the equivalent equations for cylindrical beam segments. A simple interpretation of the addition rule is that, if one calculates the energy spread with every available equation, one can use the smallest numerical value as the correct answer.

### 7.5.3 THERMODYNAMIC LIMITS

There is a maximum of the energy spread that can be induced. This can be understood by distinguishing two principally different mechanisms as drivers of statistical interactions. The two mechanisms are as follows:

- *Relaxation of kinetic energy.* This mechanism is dominant in generating the Boersch effect in a beam segment with a narrow crossover, particularly at high-current densities. In a beam, an anisotropic distribution of internal kinetic energy results from the reduction of the axial velocity spread during acceleration. Relaxation toward an isotropic distribution of the internal kinetic energy will occur under the influence of the coulomb interactions, which corresponds to an increase in energy spread. The maximum energy spread is generated when the relaxation is complete. The corresponding equilibrium temperature  $T_e$  is, for a beam segment with a crossover and a uniform angular distribution, given by

$$T_{e,k} = \frac{2}{3} T_{\perp,c} = \frac{eV}{3k_B} \alpha_o^2 \quad (7.84)$$

The corresponding rms energy spread is equal to

$$\frac{\sqrt{\langle \Delta E^2 \rangle}}{E} = \frac{\sqrt{2k_B T_{e,k}}}{E} = \sqrt{\frac{2}{3}} \alpha_o \quad (7.85)$$

If one assumes that the distribution is Gaussian (and relaxation typically involves many collisions per particle, so this is realistic), the *FWHM* is

$$\frac{\Delta E_{FWHM}}{e} = 1.92 \alpha_o \quad (7.86)$$

- *Relaxation of potential energy.* This mechanism is dominant in (nearly) cylindrical beams, particularly at low current densities. It corresponds to a conversion of the initial potential energy of the randomly distributed cloud of particles into kinetic energy by the generation of microscopic velocities. The corresponding velocity distribution will be Holtsmarkian, provided that the beam can be considered as a three-dimensional beam. The upper limit is

reached when the microscopic density distribution obeys Maxwell–Boltzmann statistics. For a cylindrical beam  $T_e$  is given by

$$T_{e,p} = \frac{(4\pi a^2)^{1/3} C_o}{k_B} n^{1/3} \quad (7.87)$$

as was shown by Jansen (1990). The numerical constant  $a$  is equal to 0.08702, correcting a misprint in the original publication. The corresponding rms energy spread is equal to

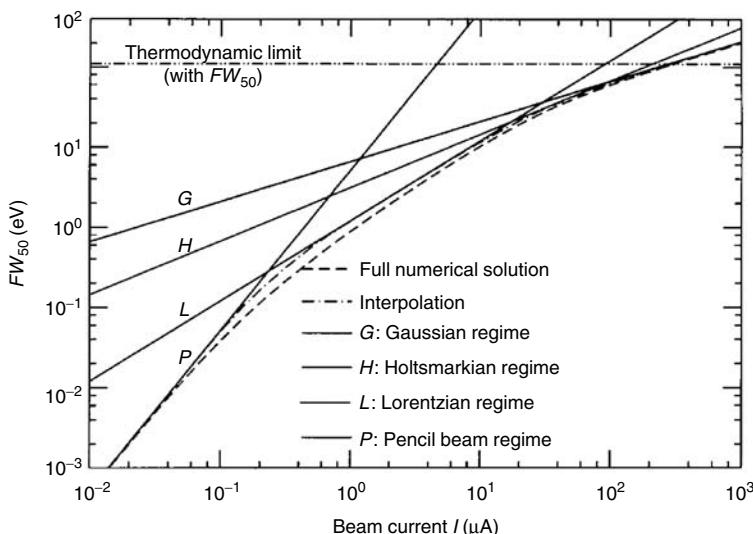
$$\sqrt{\frac{\langle \Delta E^2 \rangle}{E}} = 0.254 \frac{m^{1/12} e^{1/4}}{\varepsilon_o^{1/2}} \frac{J^{1/6}}{V^{7/12}} \quad (7.88)$$

where  $J = I/\pi r_o^2$  is the current density in the beam. For  $V = 10$  kV and  $J = 100$  A/cm<sup>2</sup> one finds  $\sqrt{\langle \Delta E^2 \rangle} = 0.54$  eV.

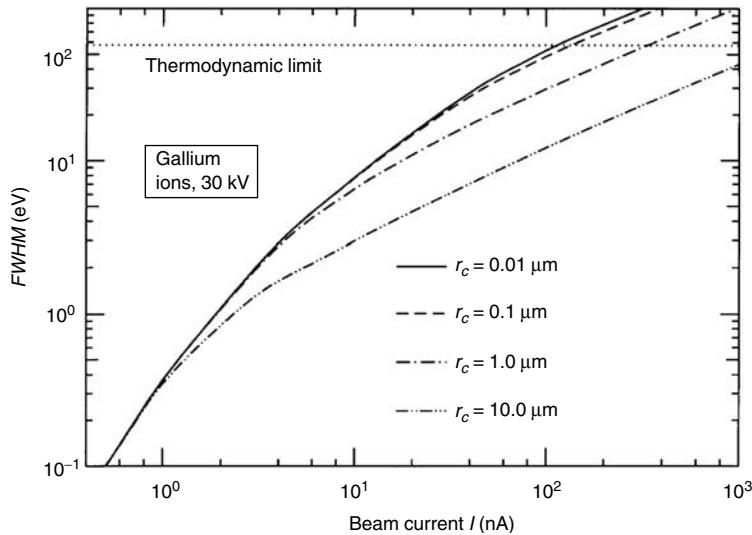
The occurrence of different mechanisms is closely related to the contribution of different types of collisions. When complete collisions are dominant, the Boersch effect will primarily be generated by conversion of kinetic energy. When the contribution of incomplete collisions becomes significant, relaxation of potential energy will play a role.

#### 7.5.4 NUMERICAL EXAMPLES

For the same example segment discussed in Section 7.4.3, the Boersch effect was calculated. Figure 7.12 compares the interpolation of Equation 7.78 with the values found from the full numerical calculation of the distribution functions. The thermodynamic limit is also indicated. In the Gaussian regime, where this limit is relevant, the  $FW_{50}$  is only slightly smaller than the  $FWHM$  value (see Table 7.2). Figure 7.13 gives the energy broadening in the segment for 30 kV gallium ions at different values of the crossover diameter.



**FIGURE 7.12** Interpolation between  $FW_{50}$  values of the energy broadening in different regimes. Numerical values are given for an electron beam segment of length 200 mm, with a crossover of radius 0.05  $\mu\text{m}$  at 50 mm, where the beam half angle is 1 mrad. Acceleration voltage is 100 kV.



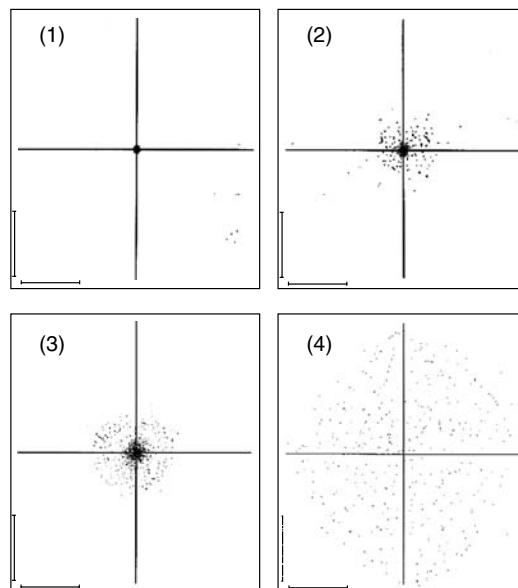
**FIGURE 7.13** Boersch effect for gallium ions in a 200 mm long segment with a crossover, radius  $r_c$ , at 50 mm, half angle 1 mrad. Curves given are for  $r_c = 0.01, 0.1, 1.0, 10.0 \mu\text{m}$ . The thermodynamic limit according to Equation 7.86 is indicated.

## 7.6 MONTE CARLO APPROACH

The particle-particle interaction in a beam is a classic  $N$ -body problem, for which MC simulation techniques are very useful. In the simulation, an ensemble of particles is created with random initial conditions within the source parameters. These particles are traced through the system, while taking into account both the total force exerted by the other particles in the ensemble and the forces, or directional changes, caused by optical elements. The advantage of the MC method is its potential accuracy, even in a complex beam geometry, possibly even nonrotationally symmetric, with many segments between lenses. The disadvantage of the MC method is the time consumption, which can be appreciable even for a run of only one set of experimental parameters. Also, it is difficult to gain insight from MC results into the dependency of the effects on experimental parameters and thus to design or optimize a column. Simulation programs for this purpose were developed by Groves et al. (1979), Groves (1994), Tang (1983), Jansen (1990), Jansen et al. (1983), Yau et al. (1983), Munro (1987), Sasaki (1979, 1982), El-Kareh and Smith (1979), Dayan and Jones (1981) and Jones et al. (1983, 1985). Two of these are commercially available (Munro and Jansen). All MC programs for the present work are essentially the same. At the emission surface, a pseudorandom number generator is used to assign a position and a velocity to a particle, in such a way that the current distribution at the source, the angular distribution, the total current, and the energy spread resemble the macroscopic properties of the beam. Each particle is also assigned a random time  $T$  to start at the emission surface. The number of particles in a sample is typically between 100 and 100,000, which also determines the length of the sample. The exact choice is very important for both the calculation time and the accuracy of the calculation. The trajectories are determined by updating the position and velocity of every particle at regular time intervals. Within each interval, trajectories are calculated in a numerical integration routine, which is of third order in  $\Delta t$  in most programs, but can be of higher order to permit larger time steps. Optical elements, such as lenses, deflectors, and so on, can be simulated most easily in thin lens approximation by adding transverse velocities to the velocity of the particle. The output of these actions is the final positions and velocities of the particles. To obtain sufficient statistics, the calculation is repeated a number of times with new random initial conditions. In terms of time consumption this is more efficient than choosing more particles per sample. Finally, a data analysis program extracts characteristic beam properties from the set of final coordinates.

The accuracy of the results is limited by several types of errors. Some errors are related to the simplified physical model, for example, the simplification of the electrostatic fields near the source or the magnetic and electrostatic fields in the optical components. But also the possible occurrence of two particles that are very close together, which a correct emission model would forbid, is a model error. Ray-tracing integration errors can usually be estimated by the program and thus kept within acceptable limits. Another error occurs when interaction effects near the edge of the sample are calculated improperly. This has several causes. A particle at the edge does not have enough neighbors to represent the total beam. Also, the space charge force acting on a particle near the edge is unbalanced, resulting, for example, in an acceleration of particles at the front. These *finite size* effects can be minimized by choosing large samples; however, the calculation time goes up roughly with the square of the number of particles in a sample. Jansen corrects for unbalanced space charge acceleration, which partly solves the problem. Jones et al. (1983) have *ghost* charges traveling with the sample, each at one side, located on the axis. Groves (1994) subtracts the space charge effect of the finite sample by calculating the effect analytically and then adds the effect from a continuum beam of infinite length. Finally, errors occur in the data analysis and interpretation. These errors are partly of a statistical nature when estimating, for instance, an  $FW_{50}$  value from a distribution, but more generally are of a systematic nature. To fit a distribution function, the coordinates of the particles are gathered in a distribution histogram. The distribution function that the program tries to fit through a least squares routine will often not have the same form as the histogram, which leads to errors larger than expected on the basis of statistical estimates only.

Several modifications to the general concept of the MC simulations have been proposed and implemented. Jansen (1990) has replaced the time-consuming  $N$ -body interaction in the drift space by  $N$ -two-body interactions, which are then evaluated from analytical expressions. This speeds up the calculation by a factor of 10–100. As his programs still also contain the full  $N$ -body approach the results of the fast variant can easily be checked for selected points. Groves (1994) has reduced the finite size errors, which allows the use of smaller samples, thus also increasing the speed of calculations.



**FIGURE 7.14** Trajectory displacement in the focus of a two-lens ion beam system as determined by Monte Carlo simulation. The column has a typical gallium liquid metal ion source. The beam is apertured to the final current at the source. Column length is 260 mm and lens-object distance is 40 mm. The plot areas are  $4 \times 4 \mu\text{m}$ . The beam currents (at 30 kV) are 0.0, 0.1, 1.0, and 10.0 nA, respectively.

The MC approach is the only way to estimate the effects close to the emitter, where the particles are accelerated in a very nonhomogeneous field. Calculations of the Boersch effect were performed by Shimoyama et al. (1993), Thomson (1994) and Elswijk et al. (1995) for field emission electron sources. Calculations of trajectory displacement in a liquid metal ion source are reported by Ward (1985). Figure 7.14 shows a typical output of an MC program.

## 7.7 STATISTICAL COULOMB EFFECTS IN THE DESIGN OF MICROBEAM COLUMNS

### 7.7.1 COMBINATION OF TRAJECTORY DISPLACEMENT AND OTHER CONTRIBUTIONS TO THE PROBE SIZE

The best, most concise characterization of a probe-forming instrument is the relation between probe size and probe current for the optimized setting of the instrument. Without coulomb interactions there are four contributions to the probe size: the diffraction disk, the chromatic aberration disk  $d_c$ , the spherical aberration disk  $d_s$  and the geometric image of the source  $d_g$ . The total size is determined by the aperture angles in the lenses and the magnification of the system. The probe current is determined by the size of the geometric image of the source, the aperture angle at the probe and the brightness of the source. For an optimized instrument, limited by chromatic aberration, the current  $I_c$  in a total probe of  $FW_{50} = d_t$  is (Kruit et al., 1995)

$$I_c = 5.4d_t^4 B_r E^3 / (C_c^2 \Delta E^2) \quad (7.89)$$

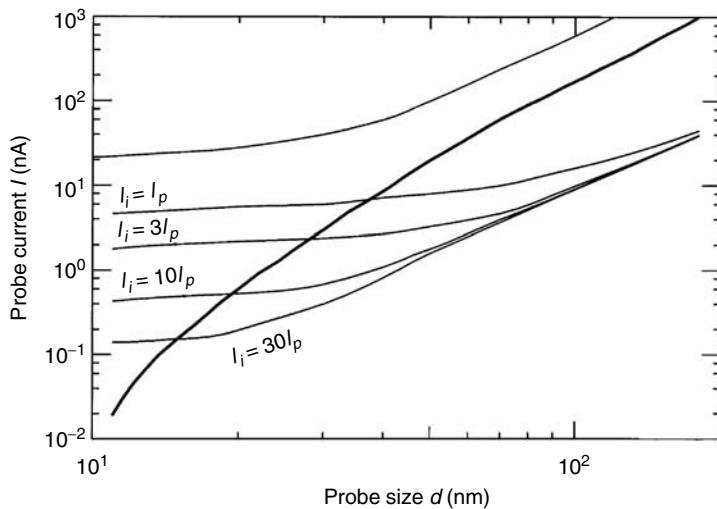
in which  $B_r$  is the reduced brightness (brightness divided by acceleration voltage),  $\Delta E$  the FWHM of a Gaussian energy distribution, and  $C_c$  the total chromatic aberration coefficient of the column, usually equal to the aberration of just the probe-forming lens. Note that if the  $FW_{50}$  of the energy distribution is used, the prefactor is 1.7 instead of 5.4. The current in a probe limited by spherical aberration is

$$I_s = 2.44d_t^{8/3} B_r / C_s^{2/3} \quad (7.90)$$

with  $C_s$  the total spherical aberration coefficient.

Several authors have analyzed the increase of the probe size due to coulomb effects as a function of the current (Brodie and Meisburger, 1992; Hirohata et al., 1992; Stickel, 1995), usually with the help of MC simulations. Thomson (1994) has analyzed how much current can be obtained in an electron probe of 25 nm diameter for different column configurations, optimizing the aperture angle for each configuration. If this is done for a range of diameters, one obtains a full characterization of the instrument performance under the influence of coulomb interactions. In the following, we summarize some studies (Kruit et al., 1995; Kruit and Jiang, 1996a,b; Jiang et al., 1996; Jiang and Kruit, 1996) with this approach, using Jansen's analytical equations in a program (Jiang et al., 1996) which combines the interaction effects with aberrations in a multi-segment column. An earlier study of this type is reported by Venables and Cox (1987).

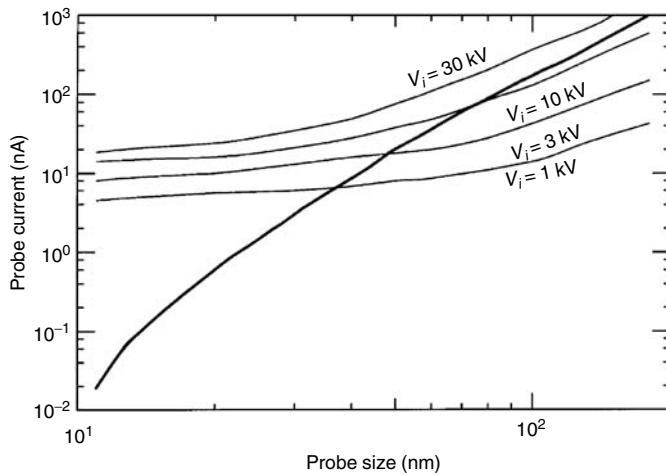
After having calculated the optimized  $I-d$  curve of a column, without coulomb effects, we check how much current  $I_{max}$  we can really afford in each spot size before the coulomb effects give an  $FW_{50}$  disk equal to  $d_t$ . The current is increased by increasing the brightness of the source, as (until the coulomb effects dominate) the optimized setting of aperture angle and magnification are independent of the brightness. In this fashion we obtain an  $I_{max}-d_t$  curve separating the area in the  $I-d$  plane where the column can be operated without influence of coulomb interactions from the area where the interactions dominate the performance.



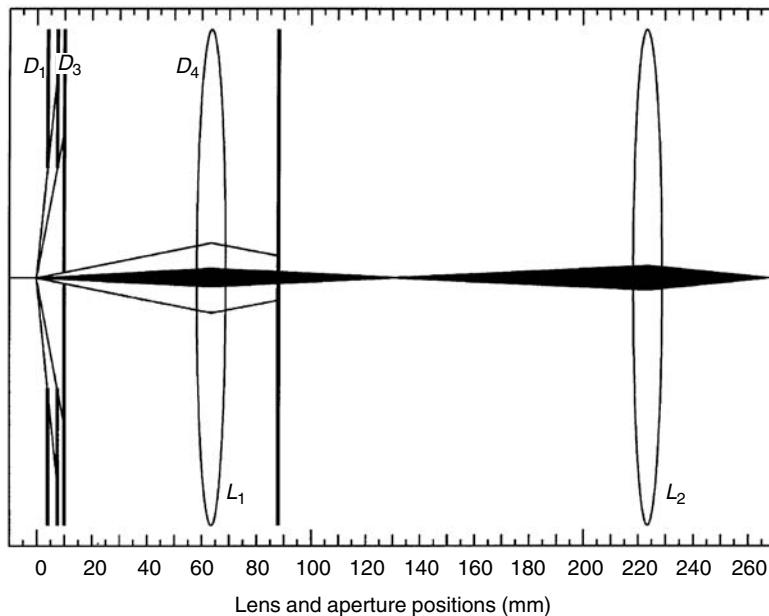
**FIGURE 7.15** Trajectory displacement limitation of the probe current in a two-lens scanning electron microscope. The thicker line indicates the brightness-limited current in the absence of coulomb interactions. The thin lines show the maximum current that can be allowed before the coulomb interactions dominate the performance for different currents in the intermediate section of the column.

Figure 7.15 shows the  $I$ - $d$ , curve of a 1-kV electron beam instrument, limited by the aberrations of the final lens ( $C_c = C_s = 1$  cm), with a source of reduced brightness  $2 \times 10^7 \text{ Am}^{-2}\text{sr}^{-1}\text{V}^{-1}$  and  $\Delta E = 0.5$  eV, typical for a Schottky emitter. The top  $I_{max}$ - $d_t$  curve shows how much current could be afforded in the probe before the coulomb interactions in only the final 1 cm of the column start to dominate. This would indicate the ultimate current obtainable with a 1-kV final beam energy in a lens of  $C_c = C_s = W = 1$  cm. Note that the Schottky emitter could not deliver sufficient brightness anyway, but the result serves to indicate that a higher brightness of the emitter, which would shift the  $I$ - $d$ , curve upward, is not by definition useful. It also shows the  $I_{max}$ - $d_t$  curves for a column with 1 cm distance between the final lens and the probe and an intermediate section of 20 cm with a crossover 8 cm in front of the final lens. The trajectory displacement was calculated in these two sections. The top of the four  $I_{max}$ - $d_t$  curves assumes the same current in both sections; the other curves assume an aperture in the final lens, cutting the current by factors of 3, 10, and 30. An analysis of this result shows that the horizontal part of the  $I_{max}$ - $d_t$  curve indicates a pencil beam regime and the sloping part a Holtsmark regime. Clearly, with this design of a column, the current at probe size over  $\sim 40$  nm would be limited by trajectory displacement. Figure 7.16 shows one of the solutions: acceleration of the electrons in the intermediate section.

Jiang et al. (1996) describe the full optimization procedure for a low-voltage SEM, including a recalculation of optimum magnifications and taking into account realistic aperture sizes and positions. So far, we have only checked if the trajectory displacement will influence the performance of a system. When it is concluded that there is an influence, we might ask if the instrument settings can be reoptimized so as to minimize the effect. Without interaction effects, the optimized setting gives a balance between the size of the source image and the size of the aberration disks, a balance obtained by the choice of magnification and aperture size. For each aperture size one can vary the magnification and calculate, or measure, the probe size. As an example we take the column of Figure 7.17 with a gallium liquid metal ion source ( $B_r = 2 \times 10^6 \text{ A/m}^2 \text{ srV}$ ,  $d_s = 50$  nm,  $\Delta E = 5$  eV, and  $E = 15$  kV), two lenses  $\sim 60$  and  $220$  mm from the source and a target at  $270$  mm. Figure 7.18 shows the probe size as a function of magnification. For small magnification the probe size is dominated by the aberrations, for large magnifications by the source image or by the trajectory displacement. Including the interaction effect, one finds the optimized situation at a different magnification.

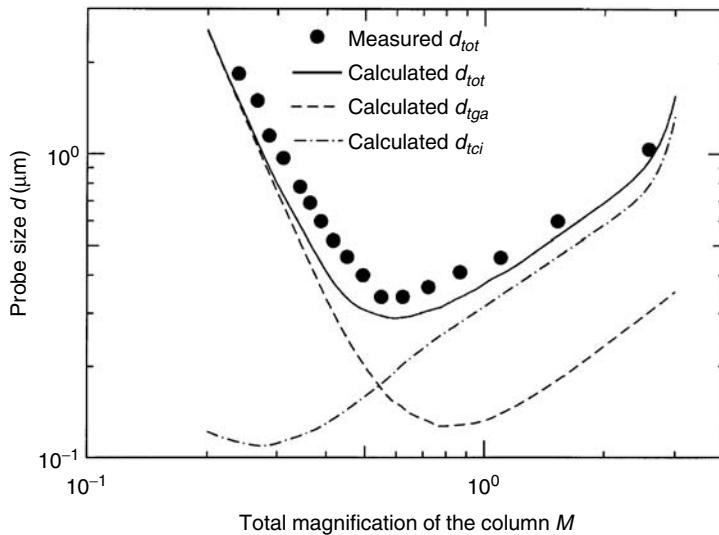


**FIGURE 7.16** Trajectory displacement limitation of the probe current in a two-lens scanning microscope for different values of acceleration in the intermediate section (1, 3, 10, and 30 kV, respectively). The current in the intermediate section is equal to the current in the probe.

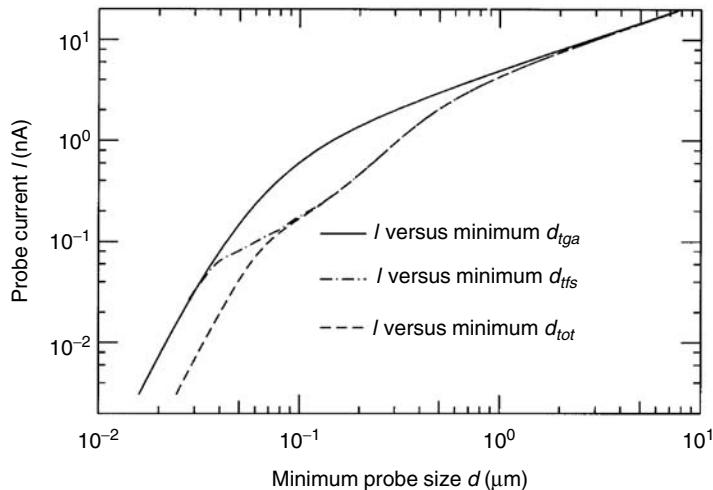


**FIGURE 7.17** Ion beam column configuration used for the data of Figures 7.18 and 7.19.

A measurement series as reported in Bi et al. (1997) is also shown. In the calculation, the interactions in all column segments are taken into account. In such a way, the smallest possible probe size can be determined for each probe current, leading to an optimized  $I-d$ , curve as shown in Figure 7.19. The performance at small beam currents is limited by the aberrations of the probe lens  $L_2$ , at large beam currents by the aberrations of both gun lens  $L_1$  and probe lens  $L_2$ , and at intermediate currents by the trajectory displacement. If the current is not limited by a variable aperture close to the source, but instead by an aperture further in the column, the performance at small probe currents is also limited by the trajectory displacement.



**FIGURE 7.18** Probe size versus magnification for the column of Figure 7.17. There is no crossover between the lenses. The current is limited by a 50- $\mu\text{m}$   $D_3$  aperture at 10 mm, allowing 0.8 nA into the column.  $d_{tga}$  is the probe size without interactions,  $d_{tci}$  is the interaction contribution to probe size.



**FIGURE 7.19** Probe current versus probe size  $d_{tot}$  for the column of Figure 7.17, when the current is reduced to its final values at  $D_3$ .  $d_{tga}$  is the probe size without interactions.  $d_{tfs}$  is the probe size if the current is reduced to its final value very close to the source.

In none of these calculations have we treated the region close to the source correctly because we assumed that the beam energy and beam current obtained their final value right at the emitting surface. The trajectory displacement effect in the accelerating region of electron emitters is expected to be small, even if the deflection in the collisions is appreciable, because the distance between the emitter and the collision location is small. This assumption was verified by Thomson (1994). The trajectory displacement in liquid metal ion sources might be the origin of the experimentally observed virtual source size (Ward, 1985).

### 7.7.2 INCLUSION OF THE BOERSCH EFFECT

For microbeam columns the importance of the Boersch effect is found in the possible increase of the chromatic aberration disk. For an estimate of the effect in a whole column, it is then necessary to first find which lenses contribute to the chromatic aberration disk. For each separate contribution only the energy broadening that occurred upstream from the contributing lens has to be taken into account. For the two example columns in the previous section (a 1-kV electron beam and a 15-kV ion beam column), the energy spread was calculated using the analytical approximations. Only for the highest currents the induced spread approaches the intrinsic energy spread of the source.

The energy broadening generated near the emitter, in the acceleration region, cannot be estimated from the equations in Section 7.5.2, but can be found from MC simulations or dedicated analytical models (Shimoyama et al., 1993; Thomson, 1994; Elswijk et al., 1995) or experiments (Bell and Swanson, 1979; Troyon, 1988). Most of the Boersch effect takes place close to the emitter where the velocity of the particles is low and the density very high. This means that it is impossible to limit the effect by aperturing the beam, so the energy spread caused by the interactions must be accepted as an intrinsic property of the source. Thus, it can be taken into account by using the appropriate experimental values of source energy spread. If the source is operated at a variable extraction voltage, this can influence the energy spread. For liquid metal ion sources, the energy spread increases dramatically with the total emission current.

### 7.7.3 DESIGN RULES FOR THE MINIMIZATION OF STATISTICAL INTERACTIONS

The analytical equations for the interaction effects and the numerical examples presented in the previous sections may be translated to some general rules for the designer and operator of a micro-beam instrument:

1. A shorter system usually decreases the Coulomb effects.
2. Fewer crossovers or avoiding crossovers altogether is especially useful if the system is limited by the chromatic aberration caused by energy broadening. However, the trajectory displacement may also be somewhat reduced and, in general, the spherical aberration disk is also smaller in a system with fewer crossovers as the lenses are less strongly excited. (Brodie and Meisburger, 1992; Jiang and Kruit, 1996).
3. A higher-acceleration voltage always decreases the Coulomb effects. If the energy of the particles at the target must be low, it is useful to accelerate in the column and decelerate close to the target.
4. A beam-limiting aperture close to the source, preferably reducing the current to the final probe current, always reduces the Boersch effect and the trajectory displacement effect in the pencil beam regime but not in the Holtsmark or Gaussian regime. This implies that the complication of positioning the aperture close to the source is only useful for the high-resolution, low-current range of operating conditions.
5. A very high brightness source is not useful in a column which is dominated by the trajectory displacement effect. If such a source is used, the magnification and aperture angle must be carefully chosen not to find a larger probe size than would be obtained with a lower-brightness source.

### 7.7.4 A STRATEGY FOR THE CALCULATION OF INTERACTION EFFECTS

The most common and most important question about Coulomb interactions is “Should I worry that my instrument’s performance is affected, or can I forget about it?” Only after this question is

answered one might want to know how large the effects really are and whether they can be reduced by appropriate actions. In this line of priorities, the following could be a strategy:

1. Schematize the instrument: thin lenses, sections with constant acceleration voltage. Find the appropriate parameters: source size, angular current density or brightness, semiangles in the sections between lenses, positions of crossovers, etc.
2. Find the probe current–probe size ( $I-d$ ) relation. Decide what to use for  $d$ : the  $FWHM$  or the  $FW_{50}$  values. The  $I-d$ , curve is often a combination of curves for individual aperture sizes.
3. Calculate for a few different  $(I, d)$  values found in step 2 (and the corresponding instrument settings) and how large the trajectory displacement is. Of course it is easiest to do this with an appropriate computer program (Jansen, 1990; Jiang, 1996), but an estimate can also be obtained by using the equations of Section 7.4. Often it is sufficient to use only the equations for the pencil beam regime and the Holtsmark regime. The addition rule for the  $FW_{50}$  and  $FWHM$  values amounts to taking the smallest value of the values found with the different equations. The contributions from the individual segments must be multiplied by the appropriate magnification to find the effect on the probe size.
4. Calculate the energy broadening and the related size of the chromatic aberration disk for a few  $(I, d)$  values.
5. Decide whether or not there is a problem. Is the probe size increased anywhere on the optimized  $I-d$ , curve? If there is a problem, continue with step 6.
6. To get a better assessment of the column's performance under the influence of the interaction effects, it is useful to determine the  $I_{max}-d$ , curve as shown in Figure 7.15. For this, the trajectory displacement and, if step 3 indicated that the Boersch effect is important, the energy broadening must be calculated as a function of the current in the column. As the column setting should not be changed as a function of current, it is the angular current density at the source that has to be varied (which however is not feasible for a liquid metal ion source).
7. Now, using the design rules of Section 7.7.3, one can try to push the  $I_{max}-d$ , curve to higher values of  $I$ , while at the same time making sure the brightness-limited  $I-d$ , curve of step 2 does not fall to lower values of  $I$ .
8. If it is impossible to decrease the interaction effects sufficiently, it becomes necessary to reoptimize the column setting, as in step 2, now including the interaction effects. The parameters for the optimized situation (aperture size and magnification) now depend on the source brightness. In the optimized probe, the contribution from the interaction effects will be of comparable size to the geometric source image or the aberration disk. The consequence is that a slight change of source brightness will influence the size of the probe. The current density distribution of the contribution from the interaction effect might be very different from the other contributions. Especially in the pencil beam regime it has long tails. Depending on the application, this might determine how large a relative contribution of the interaction effects is acceptable.
9. As the analytical equations are approximations, it is wise to check the results with an MC simulation and fine-tune the optimization of step 8 according to the results. It is important to look at the current distributions found by the simulation and not only the  $FWHM$  or  $FW_{50}$  values found from a fit to these distributions.

## 7.8 NONMICROBEAM INSTRUMENTS

The analysis in this chapter has concentrated on the effect of coulomb interactions on the size of the probe in microbeam instruments. For other applications of charged particle beams, the effects might be different. For example, in shaped beam lithography instruments or projection

lithography machines, the image of the shape does not coincide with the crossover and the crossover diameter may be substantial as compared to the beam size in the lenses. In that case there is a contribution to the edge unsharpness from the angular deflections occurring inside the crossover. The exact calculation of the trajectory displacement for these instruments has led to a substantial confusion. The limitations that trajectory displacement sets to the maximum allowed current in high-resolution lithography machines ultimately led to the failure of projects in shaped beam ion lithography (Vijgen, 1994), ion projection lithography (Chalupka et al., 1994; Hammel et al. 1994) and electron beam projection lithography (Berger et al. 1991; Mkrtchyan et al., 1995; Pfeiffer and Stikel, 1995). Mkrtchyan et al. (1995) pointed out that the equations in Jansen (1990) were not valid for typical projection systems and proposed a novel theory based on a nearest-neighbor approximation. Jansen (1998) responded that the inaccuracies were not caused by a fundamental shortcoming of the theoretical model of Jansen (1990), but rather by the use of multidimensional fit functions beyond the range of experimental conditions they were originally designed for. He presented modified equations to extend the applicability to broad beams used in projection systems. In contrast to the equations in Section 7.4, here it is necessary to include parameters  $S_i$  and  $K$  which describe the position of the image in each segment and the size of the crossover. For a segment of length  $L$ , where the crossover is at a distance  $L_1$  from the entrance, the image at a distance  $L_i$  from the entrance and  $\alpha$  and  $r_c$  defined as in Figure 7.2:

$$S_c = \frac{L_1}{L} \quad S_i = \frac{L_i}{L} \quad K = \frac{\alpha L}{2r_c} \quad (7.91)$$

The image may be outside of the segment, so  $-\infty < S_i < \infty$ .

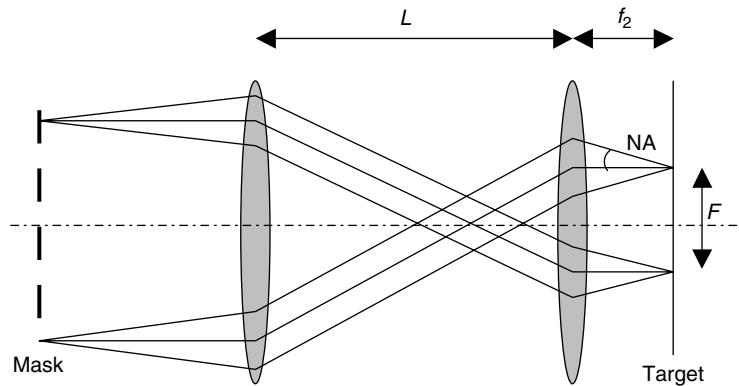
The trajectory displacement equations for the different regimes (Gaussian, weak complete collision, Holtsmark and pencil beam) become fairly complicated. As practical projection lithography machines have operated mainly in the Holtsmark regime, we shall only give that equation here:

$$FW_{50H} = 0.215 \frac{m^{1/3}}{\varepsilon_o} \cdot S_{HT}(S_c, S_i, K) \cdot \frac{I^{2/3} L^{2/3}}{V^{4/3} \alpha_o^{4/3}} \quad (7.92)$$

with

$$S_{HT}(S_c, S_i, K) = \left[ \left\{ \left( \frac{\frac{11S_c S_i K}{1+20S_c} + \frac{S_c |S_i - \frac{1}{2}| K^{4/3}}{1 - \frac{1}{2} S_c^{1/3}}}{1 - \frac{1}{2} S_c^{1/3}} \right)^{-1} + \left[ \left( \frac{1}{6} + 2|S_c - S_i| \right) K^{1/3} \right]^{-1} \right\}^2 + \left( \left| \frac{3S_c - 2S_i}{(2S_c)^{1/3}} + \frac{1 - 3S_c + 2S_i}{[2(1 - S_c)]^{1/3}} \right| + 600 \frac{|S_c - S_i|}{K^{1/3}} \right)^2 \right]^{-1/2} \quad (7.93)$$

This must be considered as the generalized Equation 7.69, given earlier. When this is applied to a projection system as shown in Figure 7.20, it is found that the contribution of the segment with the crossover dominates.  $S_i$  in that segment approaches infinity which enables simplification of the  $S_{HT}$



**FIGURE 7.20** Schematic representation of the projection lithography system.

expression. It is illustrative to express the result in the experimental parameters  $F$  (field size),  $NA$  (numerical aperture at the image), and  $f_2$  (focal distance of the final lens):

$$FW_{50H} = 0.683 \frac{m^{1/3}}{\varepsilon_o} \cdot T(S_c, K) \cdot \frac{I^{2/3} \cdot f_2^{5/3}}{V^{4/3} \cdot F \cdot NA^{1/3}} \quad (7.94)$$

where  $T(S_c, K)$  is a parameter that depends only weakly on  $S_c$  and  $K$ . For  $K < 10^3$ , it simplifies to

$$T(S_c, K) = \left( 1 + \frac{1}{(5.5S_cK^{2/3}/(1+20S_c^2)) + (S_cK/(2-S_c^{1/3}))} \right)^{-1} \quad (7.95)$$

Note that this is the result for a round, uniformly filled mask and a top-hat distribution in the crossover. For a square mask with uniform illumination and a Gaussian distribution in the crossover, the factor 0.683 should be replaced by 0.570. For very small crossovers, very small currents, and very large currents, the approximations break down. For a fuller account, see Jansen (1998), where the theory is also compared with MC results from the author's program.

Stickel (1998) compares the theories of Mkrtchyan et al. (1995) and Jansen (1998) with the results from two MC programs and concludes that the discrepancies between the results (up to factor 3 in  $FWHM$ ) are so significant that experimental verification is needed to determine the validity of the different approaches. Although some of the discrepancies may be attributed to the assumption that the rms value calculated in one of the simulations can simply be translated to an  $FW_{50}$  value through a numerical factor, the recommendation to compare with experiment is valid. The precise measurement of statistical Coulomb interactions is notoriously difficult. de Jager et al. (1999) report experimental results of interaction in ion projection lithography. Unfortunately, the exact experimental conditions are not given, so a direct comparison with theory is impossible. However, the results are extrapolated to a system for which the system parameters are given. If these extrapolated numbers are compared to the numbers given in Jansen (1998), there seems to be correspondence within a factor of two. Very precise measurements of interaction blur in electron projection lithography have been performed by Liddle et al. (2001) and Yahiro et al. (2001) in systems with exactly the parameters for which Jansen (1998) applied his theory and performed MC calculations. Experiments, MC simulations, and Jansen's analytical equations are in excellent agreement: all deviations are less than 20%.

In some instruments it is the energy spread itself that is important, rather than its effect on the probe size, or in addition to its effect on probe size. This is the case, for example, for e-beam testers where very short pulses must be produced. The spread in energies and thus flight times ultimately determines the timing resolution of the measurement. The energy spread in the beam is also important in energy-loss experiments in transmission electron microscopes or dedicated spectroscopy instruments. For a high-energy resolution, the primary beam must be filtered. The correct treatment of the space charge effect is essential for the design of highly monochromatic beams from moderately bright electron sources (Ibach, 1991). In the monochromator design for field emission sources, the statistical interactions play a dominant role (Mook and Kruit 1999; Tiemeijer, 1999).

## 7.9 DISCUSSION

A comparison between results of the analytical theory and MC simulations (Jansen, 1990, 1998; Vijgen, 1994) shows that the differences are in general <20% with a maximum usually occurring in the overlap region between regimes. The analytical theory usually predicts a larger effect than the MC simulation. Possible causes of errors in the MC simulation were discussed in Section 7.6. The inaccuracies in the analytical equations can have several causes:

- The equations were found by analyzing the displacements of a particle traveling on the beam axis. The statistical displacements of particles on the edge of the beam might be somewhat smaller (Jansen, 1990).
- The model assumes that the beam is initially monochromatic. The influence of this assumption, however, was carefully checked (Jansen, 1990).
- Within the analytical theory it is often necessary to use fit functions to describe certain dependencies that lead to some deviations from the numerical calculations.
- The derivation of  $FWHM$  and  $FW_{50}$  values relies on the assumption that the distribution functions in each regime are purely Gaussian, Lorentzian, etc. This is not necessarily the case, especially far from the center (in the tails). Obviously, in the transition between regions with different distributions the relation between  $FWHM$  and  $FW_{50}$  cannot be as simple as in Table 7.2. The latter can also be concluded from the fact that the transition point between regimes depends on the width measure chosen (see Figure 7.9).
- For very high currents, a thermodynamic equilibrium is established. This was taken into account for the Boersch effect by setting a maximum to the energy spread that can be expected. The limitation of the *extended two-particle model* (Equation 7.24) on which the analytical equations are based is more general. At high particle densities, a particle can experience strong collisions with more than one other particle. In that regime, the equations for trajectory displacement are no longer accurate anymore either and will generally overestimate the effect. However, we do not have indications for the maximum displacement similar to the thermodynamic limit of the Boersch effect. To stay away from this regime,

$$\bar{\lambda} = \frac{m^{1/2}}{2^{7/2}\pi\epsilon_o e^{1/2}} \frac{I}{\alpha_o^2 V^{3/2}} \quad (7.96)$$

should be  $\ll 1$  for beams with a crossover and

$$\lambda^* = \frac{m^{1/2}}{2^{7/2}\pi\epsilon_o e^{1/2}} \frac{IL^2}{r_o^2 V^{3/2}} \quad (7.97)$$

should be  $\ll 1$  for cylindrical beams.

Trajectory displacement and Boersch effect are never the only contributions to resolution and energy spread. The manner in which the different contributions must be added depends very much on the distribution functions of the different contributions (see Equation 7.54 and Barth and Kruit, 1996). If this is not done correctly, one may find theoretical values for resolution and energy spread that are far from the experimental values, or perhaps even worse: one might *optimize* the instrument design or the settings of parameters such as magnification and aperture angle to a point that is not really the optimum.

## LIST OF SYMBOLS

$A_\gamma$	function of experimental parameters, describing $(\vec{k})$
$\alpha_o$	beam semiangle
$b$	impact parameter
$C_c$	chromatic aberration coefficient
$C_o$	$e^2/4\pi\epsilon_o$
$C_s$	spherical aberration coefficient
$\gamma$	numerical constant, characterizing a distribution function
$d_r$	average radial distance between particles
$d_z$	average $z$ -distance between particles
$d$	average distance between particles
$\Delta\alpha$	angular displacement
$\Delta\vec{\eta}$	displacement of the test particle
$\Delta\vec{\eta}_2(\vec{\xi})$	displacement of the test particle as a result of the interaction with one field particle
$\Delta r$	displacement of test particle in the $r$ -direction
$\Delta v_\perp$	displacement of the test particle's velocity $\perp z$
$\Delta v_x, \Delta v_y, \Delta v_z$	displacement of the test particle velocity
$\Delta Z_f$	space charge defocus
$E$	particle kinetic energy
$e$	electron charge
$\epsilon_o$	dielectric constant
$\vec{\xi}(xyb_zv_xv_y)$	relative coordinates of field particle with respect to test particle before interaction
$F$	focal distance of lens
$I$	beam current
$k_B$	Boltzmann constant
$K_1, K_2$	size of beam at side of beam segment, relative to size of crossover
$L$	length of beam segment
$L_1, L_2$	distance from side of beam segment to crossover
$\lambda$	linear particle density, $I/e v_z$
$m$	electron or ion mass
$n_d$	particle density
$N$	number of particles in a beam segment
$P_2(\xi)$	probability of one field particle to have coordinates $(\xi)$
$P_N(\xi_1, \dots, \xi_{N-1})$	probability of configuration $(\xi_1, \dots, \xi_{N-1})$
$P_n(\vec{k}), P(\vec{k})$	Fourier transform of $\rho(\Delta\vec{\eta})$
$P_2(\vec{k})$	Fourier transform of $p_2(\Delta\vec{\eta})$
$r_c$	minimum beam radius in crossover beam
$r_o(z)$	radius of cylindrical beam
$r_\lambda$	minimum beam radius in laminar flow condition
$r_\perp$	modulus of the projection of the relative position vector on the crossover plane
$\rho(\Delta\vec{\eta})$	distribution of displacements

$\rho_2(\Delta\eta)$	displacement distribution for a system containing only the test particle and one field particle
$\rho(r, z)$	charge distribution
$S_c$	crossover location parameter = $L_1/L$
$T$	flight time through segment
$v_z$	$z$ -component of particle's velocity
$V, V_e$	particle's acceleration voltage
$W$	distance between lens and target

## ACKNOWLEDGMENT

The authors thank X. R. Jiang for valuable comments and assistance in producing the figures.

## REFERENCES

- Barth, J. E. and Kruit, P. 1996. *Optik* (Stuttgart), 101, 101–109.
- Beck, S., Plies, E., and Schiebel, S. 1995. *Nucl. Instrum. Methods Phys. Res. A*, 363, 31–42.
- Bell, A. E. and Swanson, L. W. 1979. *Phys. Rev. B*, 19, 3353.
- Berger, S. D., Gibson, G. M., Camarda, R. M., Farrow, R. C., Huggins, H. A., Kraus, J. S., and Liddle, J. A. 1991. *J. Vac. Sci. Technol. Eng.*, 35, 249–252.
- Bi, J., Jiang, X. R., and Kruit, P. 1997. *Microelectron. Eng.*
- Boersch, H. 1954, *Zeitschr. Phys.*, 139, 115.
- Brodie, A. D. and Meisburger, W. D. 1992. *Microelectron. Eng.*, 17, 399–404.
- Chalupka, A., Stengl, G., Buschbeck, H., Lammer, G., Vonach, H., Fischer, R., Hammel, E., Loeschner, H., Nowak, R., Wolf, P., Finkelstein, W., Hill, R. W., Berry, I. L., Harriott, L. R., Melngalis, J., Randall, J. N., Wolfe, J. C., Stroh, H., Wollnik, H., Mondelli, A. A., Petillo, J. J., and Leung, K. 1994. *J. Vac. Sci. Technol. B12(6)*, 3513–3517.
- Crewe, A. V. 1978. *Optik*, 52, 337.
- Dayan, P. S. and Jones, G. A. C. 1981. *J. Vac. Sci. Technol.*, 19, 1094.
- De Chambost, E. 1982. *Optik*, 55, 357.
- De Chambost, E. 1982. *Optik*, 62, 189.
- De Chambost, E. and Hennion, C. 1979. *Optik*, 55, 357.
- Debye, P. and Hückel, E. 1923. *Zeitschr. Phys.*, 24, 185.
- El-Kareh, A. B. and El-Kareh, J. C. J. 1970. *Electron Beams, Lenses and Optics*. Academic Press, New York, p. 250.
- El-Kareh, A. B. and Smither, M. A. 1979. *J. Appl. Phys.*, 50, 5596.
- Elswijk, H. B., van Rooy, T., and Schiller, C. 1995. *J. Vac. Sci. Technol.*, B13, 1037–1043.
- Glaser, W. 1952. *Grundlagen der Elektronenoptik*. Springer-Verlag, Vienna, p. 66.
- Groves, T. R. 1994. *J. Vac. Sci. Technol.*, B12, 3483–3488.
- Groves, T. R., Hammond, D. L., and Kuo, H. 1979. *J. Vac. Sci. Technol.*, 16, 1680.
- Hammel, E., Chalupka, A., Fegerl, J., Fischer, R., Lammer, G., Löschner, H., Malek, L., Nowak, R., Stengl, G., Vonach, H., Wolf, P., Brünger, W. H., Buchmann, L.-M., Torkler, M. T., Cekan, E., Fallmann, W., Pascke, F., Stangl, G., Thalinger, F., Berry, I. L., Harriott, L. R., Finkelstein, W., and Hill, R. W. 1994. *J. Vac. Sci. Technol. B12*, 3538.
- Herrmannsfeldt, W. B., Becker, R., Brodie, I., Rosengreen, A., and Spindt, C. A. 1990. *Nucl. Instrum. Methods Phys. Res. A*, 298, 39–44.
- Hirohata, S., Kosugi, T., Sawaragi, H., Aihara, R., and Gamo, K. 1992. *J. Vac. Sci. Technol.*, B10, 2814–2818.
- Hutter, R. 1967. In *Focusing of Charged Particle*, A. Septier, Ed., Vol. 2. Academic Press, New York, p. 3.
- Ibach, H. 1991. *Electron Energy Loss Spectrometers*, Springer Series in Optical Sciences, Vol. 63.
- Jager, P. W. H. de, Derksen, G., Mertens, B., Cekan, E., Lammer, G., Vonach, H., Buschbeck, H., Zeininger, M., Horner, C., Löschner, H., Stengl, G., Bleeker, A. J., Benschop, J., Shi, F., Volland, B., Hudek, P., Heerlein, H., and Rangelow, I. W., Kaesmaier, R. 1999. *J. Vac. Sci. Technol. Jager*, B17, 3098–3106.

- Jansen, G. H. 1990. Coulomb interactions in particle beams, *Adv. Electron. Electron Phys.*, Suppl. 21, Academic Press, New York.
- Jansen, G. H. 1998. *J. Appl. Phys.*, 84(8), 4549–4567.
- Jansen, G. H., van Leeuwen, J. M. J., and van der Mast, K. D. 1983. *Proc Microcirc. Eng.*, 99.
- Jiang, X. R. 1996. Ph.D. thesis, Delft University of Technology.
- Jiang, X. R., Barth, J. E., and Kruit, P. 1996. *J. Vac. Sci. Technol.*, B14, 3747–3752.
- Jiang, X. R. and Kruit, P. 1996. *Microelectr. Eng.*, 30, 249–252.
- Jones, G. A. C., Rao, V. R. M., Sun, H. T., and Ahmed, H. 1983. *J. Vac. Sci. Technol.*, B1, 1298–1302.
- Jones, G. A. C., Sargent, P. M., Norris, T. S., and Ahmed, H. 1985. *J. Vac. Sci. Technol.*, B3, 124–127.
- Kirstein, P. T., Kino, G. S., and Walters, W. E. 1967. *Space Charge Flow*. McGraw-Hill, New York.
- Klemperer, O. 1953. *Electron Optics*. Cambridge University Press, Cambridge, MA.
- Knauer, W. 1979. *Optik*, 54, 211.
- Knauer, W. 1979. *J. Vac. Sci. Technol.*, 16, 1676.
- Kruit, P. and Jiang, X. R. 1996. *J. Vac. Sci. Technol.*, B14, 1635–1641.
- Kruit, P. and Jiang, X. R. 1996. *Proc. SPIE* 96, 2, 858, 166–176.
- Kruit, P., Vijgen, L. J., and Jiang, X. R. 1995. *Nucl. Instrum. Methods Phys. Res. A*, 363, 220–224.
- Liddle, J. A., Blakey, M. I., Bolan, K., Farrow, R. C., Gallatin, G. M., Kasica, R., Katsap, V., Knurek, C. S., Li, J., Mkrtchyan, M., Novembre, A. E., Ocola, L., Orphanos, P. A., Peabody, M. L., Stanton, S. T., Teffeu, K., and Waskiewicz, W. K., and Munro, E. 2001. *J. Vac. Sci. Technol.*, B19, 476–481.
- Loeffler, K. H. 1964. Ph.D. thesis, University of Berlin.
- Loeffler, K. H. 1969. *Zeitschr. Angew. Phys.*, 27, 145.
- Loeffler, K. H. and Hudgin, R. H. 1970. *Proc Int. Conf. Electron Microsc.*, 7, 67.
- Massey, G. A., Jones, M. D., and Plummer, B. P. 1981. *J. Appl. Phys.*, 52, 3780.
- Meisburger, W. D., Brodie, A. D., and Desai, A. A. 1992. *J. Vac. Sci. Technol.*, B10, 2804–2808.
- Meyer, W. E. 1958. *Optik*, 15, 398.
- Mkrtchyan, M. M., Liddle, J. A., Berger, S. D., Harriott, L. R., Gibson, J. M., and Schwartz, A. M. 1995. *J. Appl. Phys.*, 78, 6888.
- Mook, H. W. and Kruit, P. 1999. *Ultramicroscopy*, 78, 43–51.
- Munro, E. 1987. *Nucl. Instrum. Methods Phys. Res. A*, 258, 443.
- Nagy, G. A. and Szilagyi, M. 1974. *Introduction to the Theory of Space Charge Optics*. Macmillan Press, London.
- Newton, I. 1687. In *Philosophia Naturalis Principia Mathematica*, A. Koyre and I. B. Cohen, Eds. Harvard University Press, Cambridge, MA, 1972, Reprint of the 3rd ed. (1726).
- Pfeiffer, H. C. and Stickel, W. 1995. *Microelectron. Eng.* 27, 143–146.
- Pierce, J. R. 1954. *Theory and Design of Electron Beam*. D. van Nostrand, New York.
- Renau, A., Read, F. H., and Brunt, J. N. H. 1982. *J. Phys. E: Sci. Instrum.*, 15, 347–354.
- Rose, H. and Spehr, R. 1980. *Optik*, 57, 339.
- Rouse, J., Zhu, X., and Munro, E. 1995. *Proc. SPIE*, 2522, 375–386.
- Sasaki, T. 1979. *Conf. VLSI: Architecture, Design, Fabrication*, Cal. Inst. Technol.
- Sasaki, T. 1982. *J. Vac. Sci. Technol.*, 21, 695.
- Sasaki, T. 1984. *J. Vac. Sci. Technol.*, A2, 1352–1360.
- Sasaki, T. 1986. *J. Vac. Sci. Technol.*, B4, 135–139.
- Schwartz, J. W. 1957. *RCA Rev.*, 18, 3.
- Shimoyama, H., Shimazaki, Y., and Tanaka, A. 1993. *Proc. SPIE*, 2014, 99–103.
- Sivukhin, D. V. 1966. In *Reviews of Plasma Physics*, M. A. Leontovich, Ed., Vol. 4. Consultants Bureau Enterprises, New York, 93.
- Sohda, Y., Satoh, H., Someda, Y., Saitou, N., Itoh, H., and Sasaki, M. 1995. *Nucl. Instrum. Methods Phys. Res.*, A363, 18–24.
- Stickel, W. 1995. *Proc. SPIE*, 2522, 340–350.
- Stickel, W. 1998. *J. Vac. Sci. Technol.*, B16, 3211–3214.
- Sturans, M. A., Pfeiffer, H. C., Stickel, W., and Groves, T. R. 1998. *J. Vac. Sci. Technol.*, B6, 1995.
- Tang, T. T. 1983. *Optik*, 64, 237.
- Tang, T. T. 1987. *Optik*, 76, 38.
- Thomson, M. G. R. 1994. *J. Vac. Sci. Technol.*, B12, 3498–3502.
- Tiemeijer, P. C. 1999 *Ultramicroscopy*, 78, 53–62.
- Troyon, M. 1988. *J. Microsc. Spectrosc. Electr.*, 13, 49.

- Trubnikov, B. A. 1965. In *Reviews of Plasma Physics*, M. A. Leontovich, Ed., Vol. 1. Consultants Bureau Enterprises, New York, 105.
- Van den Broek, M. H. L. M. 1984. *Optik*, 67, 69.
- Van den Broek, M. H. L. M. 1986. *J. Appl. Phys.*, 59, 3923.
- Van Kampen, N. G. 1981. *Stochastic Processes in Physics and Chemistry*, North-Holland, Amsterdam.
- Van Leeuwen, J. M. J. and Jansen, G. H. 1983. *Optik*, 65, 179.
- Venables, J. A. and Cox, G. 1987. *Ultramicroscopy*, 21, 33–45.
- Vijgen, L. J. 1994. Coulomb interactions in focused ion beam system, Ph.D. thesis, Delft University of Technology, Delft University Press.
- Vijgen, L. J. and Kruit, P. 1992. *J. Vac. Sci. Technol.*, B10, 2809–2813.
- Ward, J. W. 1985. *J. Vac. Sci. Technol.*, B3, 207–213.
- Weidenhausen, A., Spehr, R., and Rose, H. 1985. *Optik*, 69 126.
- Yahiro, T., Suzuki, S., Irita, T., Hirayanagi, N., Shimizu, H., Kojima, S., Morita, K., Kawata, S., Okino, T., and Suzuki, K. 2001. *J. Vac. Sci. Technol.*, B19, 2468–2473.
- Yau, Y. W., Groves, T. R., and Pease, R. F. W. 1983. *J. Vac. Sci. Technol.*, B1, 1141–1144.
- Zimmermann, B. 1968. Ph.D. thesis, University Karlsruhe.
- Zimmermann, B. 1969. *Rec. Symp. Electron Ion Laser Beam Technol.*, 10, 297.
- Zimmermann, B. 1970. *Advances in Electronics and Electron Physics*, Vol. 29. Academic Press, New York, 257.
- Zvorykin, V. K., Morton, G. A., Ramber, E. G., Hillier, J., and Vance, A. W. 1961. *Electron Optics and the Electron Microscope*, 5th ed., sect. 16.9. John Wiley & Sons, New York.



---

# 8 Resolution

*Mitsugu Sato*

## CONTENTS

8.1	Preface.....	392
8.2	Introduction.....	393
8.3	General Concept of Resolution and Conventional Criterions.....	395
8.3.1	Origin of the Concept of Resolution and Rayleigh's Criterion .....	395
8.3.2	Image Formation for an Optical System by Means of Charged Particle Beams .....	396
8.3.3	Conventional Definitions of the Resolution .....	398
8.3.3.1	Resolution Defined by Beam Size .....	398
8.3.3.2	Definition of Resolution Based on the Contrast Performance of an Optical System.....	399
8.4	A New Definition of the Resolution Based on Image Quality.....	402
8.4.1	A Concept of the Resolution in Terms of the Quality of Optical Image.....	402
8.4.2	Figure of Merit for the Quality of an Optical Image .....	403
8.4.3	Relation Between Rayleigh's Criterion and the Density-of-Information Passing Capacity .....	405
8.4.4	Limitations of the Application .....	406
8.5	Calculation of The Density-of-Information Passing Capacity .....	406
8.5.1	Spatial Frequency Response for an Electron Beam .....	406
8.5.1.1	Two-Dimensional Fourier Transform of Source Intensity Distribution .....	406
8.5.1.2	Optical Transfer Function for a Monoenergetic Beam .....	407
8.5.1.3	Optical Transfer Function with Chromatic Aberration Included .....	409
8.5.2	Spatial Frequency Response for Ion Beams.....	411
8.5.2.1	Mathematical Model for Current Density of Monoenergetic Beams .....	411
8.5.2.2	Current Density of Monoenergetic Beams for a Rotationally Symmetric System .....	412
8.5.3	Properties of the Best Focus Position.....	415
8.5.3.1	Best Focus Position for Electron Beams.....	415
8.5.3.2	Best Focus Position for Ion Beams .....	416
8.5.4	Approximation Method for the Density-of-Information Passing Capacity .....	416
8.5.4.1	Basis of the Concept for the IPC Approximation .....	417
8.5.4.2	Synthesis of Wave and Geometric Optics .....	418
8.5.4.3	The Fitting Functions for the IPC.....	422
8.5.4.4	Comparison between Fitting Method and Numerical Computation .....	423
8.5.5	Comparison between Calculated Resolution and Experimental Results .....	423
8.6	Optimum Condition and Attainable Resolution .....	425
8.6.1	Optimum Condition for a Diffraction-Limited System .....	426
8.6.2	Optimum Condition for a Source-Limited System.....	428
8.7	Measurement of the Resolution.....	431

8.8 Other Factors Limiting the Resolution .....	433
8.8.1 Effect of the Information Size in the Specimen.....	433
8.8.2 Effect of Vibration .....	433
8.8.3 Effect of a Magnetic Stray Field .....	434
Acknowledgment .....	434
References .....	434

## 8.1 PREFACE

The past few years have seen remarkable advances in the technologies of scanning electron microscopy (SEM) and scanning ion microscopy. Until relatively recently, scanning electron microscopes (and certainly scanning ion microscopes) were capable of producing only focused probes at least several nanometers in size (where size means approximately the full width at half maximum [FWHM] of the current density distribution). In these cases the resolution of the instrument could be reasonably understood using classical theory: the point spread function (PSF) of the instrument—the distribution of current—could be calculated and from this the resolution was predicted. In the perfect case where the source is a point and there are no aberrations, one ends up with the classical Rayleigh definition of resolution for an optical instrument, given in Section 8.2, which says the resolution is encompassed by a single number  $x_r$  determined by the diffraction of the radiation in the instrument. However, there are complicating details in scanning electron (or ion) microscopy. The source is not a point and there are aberrations, although their effects can be calculated. The calculated PSF is then approximate, mainly due to the source, but it could still be used to estimate the resolution (in the case of the SEM, an additional complication is that the image detail obtainable is often limited by the interaction volume of the electron beam in the sample that generates detectable secondary electrons).

More fundamental problems in defining resolution are the noise in the image and instrumental defects such as contamination, vibration, and so on. The signal-to-noise ratio (S/N) of the image is perhaps most fundamental. It has to be large enough to be able to detect fine detail but it cannot be made indefinitely large. Indeed, in the case of the scanning ion microscope (focused ion beam [FIB]) with heavy ions, the limit on the S/N is sputtering: the volume from which a signal is generated has to be large enough to provide an acceptable signal before it disappears. S/N limits the amount of information that can be gathered from the beam–specimen interaction. However, by taking into account how S/N limits the visible contrast in an image, the resolution could still be estimated from the classical PSF.

Now, however, another type of problem has arisen as the probe size of electron and ion scanning microscopes approaches atomic dimensions: the interaction of the charged particles with the atomic structure of the specimen needs to be taken into account. The information obtained from a sample when it interacts with the focused beam depends on more than just the size of the beam. It depends critically on the interaction between the beam and the atoms in the sample and the way the atoms in the sample are arranged among themselves. Also, the information in an image depends not only on the fineness of the detail in it, but also on the contrast, and both are limited by the noise in the image. Thus it is necessary to rethink what resolution means in terms of the beam, the noise, and the structure of the sample—a problem faced earlier in transmission electron microscopy (TEM, see, e.g., the review of resolution concepts by den Dekker and van den Bos,<sup>1</sup> and the article on coherent and incoherent imaging systems by Van Aert, Van Dyck, and den Dekker<sup>2</sup> for detailed discussions of this problem. See also the article by Van Aert, den Dekker, Van Dyck, and van den Bos<sup>3</sup>).

It appears the way things are evolving in scanning microscopy is similar to what happened in transmission microscopy. To characterize an image—to be able to state with some confidence what the true distribution of atoms in it really is, which is what I think a practical definition of resolution really involves—it will probably be necessary to make a best estimate of what this distribution is using an appropriate physical model and to convolve that with the known PSF of the instrument (including the source), to predict what the image should look like. Then the

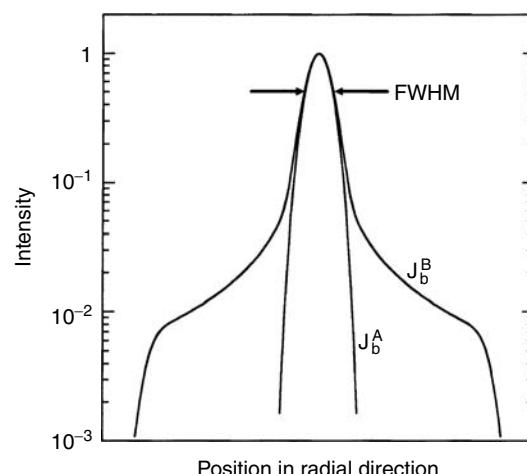
difference between the predicted and the measured image will be used to refine the model of the target until some convergence between measurement and prediction is achieved. One hopes that the final result corresponds to reality. This must be done in the presence of noise and instrumental/environmental defects such as vibration, acoustical and electromagnetic noise, drift, human error, and so on. In going through this exercise it will be necessary to pay close attention to how the signal is generated in the sample (whether secondary electrons, e.g., are generated far from the impact point of the focused beam due to scattering and how to take these into account; how an ion beam changes the sample structure). It will be fraught with difficulties and there needs to be a well-defined way of doing all these so that different observers can agree that they are seeing, or doing, the same thing.

Parenthetically, there is also the issue—if it still has meaning—of the relationship between the electron or ion beam probe size and the resolution that will be seen in an image at the nanometer or subnanometer level (it is the opinion of this writer that such a relationship is rather tenuous). There are a number of ways of defining beam size such as the diameter of a beam containing 50% of the beam current, the rise-distance of a beam swept over a knife-edge (a dicey proposition when the beam is of atomic size, although the number can be calculated if not measured), the method based on the information passing capacity (IPC) of an optical system, which is described in the following chapter, the contrast-to-gradient method, and the derivative method. This is of interest because beam size is a number used to compare different instruments.

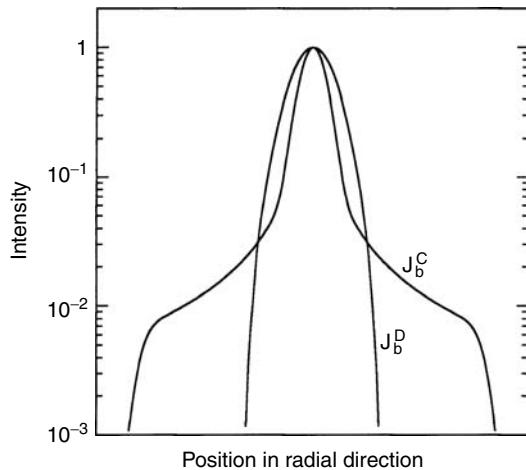
## 8.2 INTRODUCTION

In electron and ion probe-forming systems, the spatial resolution is usually the most important characteristic of the optical system. The resolution measured when a beam is used to produce an image by scanning it across a specimen has a close connection with the beam size. Thus, a simple concept for such resolution of an optical system would be to estimate it from the beam size produced by the optical system. The beam size could be defined by the disk at which the beam intensity has fallen to half (FWHM), or  $1/e$ , or 10%, etc., of the maximum, for example, depending on application.

One may, however, meet a situation where the beam intensities are given by those shown in Figure 8.1, for example. The FWHM of the distributions  $J_b^A$  and  $J_b^B$  shown in Figure 8.1 are the same, but they have differently shaped beam tails. In this case, if the contrast in the images produced by these two beams is high enough to ignore the noise fluctuations, the distributions  $J_b^A$  and  $J_b^B$  would yield the same resolution. If this is not the case, because a long tail on the beam intensity



**FIGURE 8.1** Intensity distribution of beams with different tails.



**FIGURE 8.2** Intensity distribution of beams with different tails and different full width at half maximum.

distribution makes the contrast worse,  $J_b^B$  would give poorer resolution than  $J_b^A$ . A more difficult situation is shown in Figure 8.2. Because the distribution  $J_b^C$  in Figure 8.2 has smaller FWHM than that of  $J_b^D$ , but has a longer tail, it is difficult to say generally which one gives better resolution. In such a case, the best distribution would depend on the S/N of the image.

Let us consider an actual process in which one chooses the best optical system from among several different ones. One will, at first, take images of the same specimen using each optical system. Next, one would compare the quality of each image, which is estimated from the resolution and the S/N. Finally, one will choose the best one. The best optical system is thus the system which gives the best-quality image. The concept of resolution which we present in this chapter is determined from a process similar to the one described earlier: it is found by introducing a quantitative value that represents the quality of an optical image.

A suitable concept for the quantitative analysis of the quality of image was proposed by Shannon<sup>4</sup> in terms of the mean information content of an optical image. The concept has been investigated by Gabor,<sup>5</sup> Elias et al.,<sup>6</sup> and Blanc-Lapierre.<sup>7</sup> It has been applied to the assessment of optical images by Fellgett,<sup>8</sup> Fellgett and Linfoot,<sup>9</sup> and Linfoot.<sup>10</sup> Black and Linfoot<sup>11</sup> have determined the optimum condition of an optical system in terms of the information content of an optical image suffering from spherical aberration and diffraction. The same concept has also been applied to SEM by Simon<sup>12</sup> based on Gaussian distribution beams.

The information content of an optical image is characterized by the resolution and the S/N of the image, which are the same quantities as those characterizing the quality of an image. Thus, it is possible to represent the quality of image in terms of the information content of an optical image. According to this concept, an optical system can be regarded as the system which transfers information from a specimen to the final optical image. Optical performance can thus be represented by the IPC of an optical system. As the IPC depends only on the resolution if the S/N is fixed, the optical system which gives better resolution can easily be found by comparing its IPC with those of other systems at a given S/N. The resolution of an optical system can also be determined from the IPC at each S/N.

A major part of this chapter presents how the IPC is applied to the determination of resolution, and how it is calculated exactly or approximated in practical applications for both electron and ion beam systems. Determination of the optimum condition in terms of the IPC is also presented for diffraction-limited optical systems and for source size-limited systems, where the resolution depends also on the beam current, as it is important when the optical system is designed. Practical methods for measuring the resolution are also introduced in this chapter.

### 8.3 GENERAL CONCEPT OF RESOLUTION AND CONVENTIONAL CRITERIONS

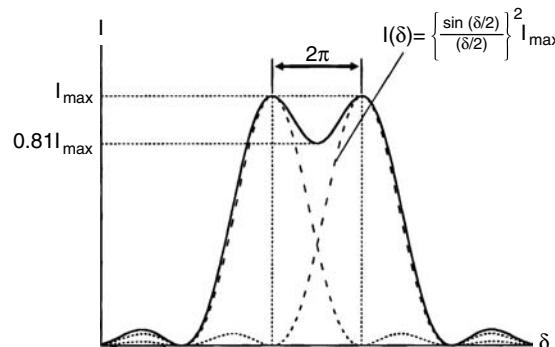
#### 8.3.1 ORIGIN OF THE CONCEPT OF RESOLUTION AND RAYLEIGH'S CRITERION

The origin of the concept of resolution is found in a spectrum analysis by means of an interferometer. Rayleigh<sup>13</sup> proposed that the two monochromatic spectra should be regarded as being just resolved when the maximum of one spectrum coincides with the minimum of the other, as shown in Figure 8.3. This criterion is called *Rayleigh's criterion*.

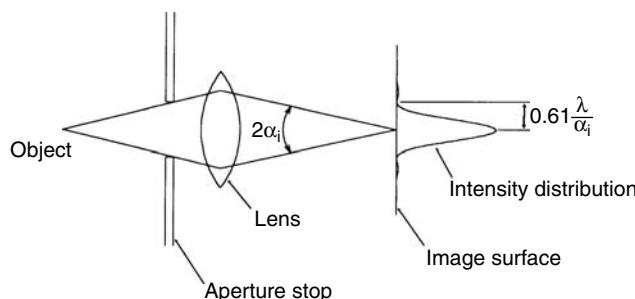
A primitive idea for the resolution of an optical system is to define it as the minimum distance between separate pointlike objects that can be distinguished in the image. As Rayleigh's criterion is well adapted to the concept of resolution for a diffraction-limited system, it can be applied to an image-forming system such as an optical microscope. Figure 8.4 shows a simple model of an image-forming system. The intensity of a pointlike object is spread out in the image plane due to diffraction. Such an intensity distribution in the image for a pointlike object is called the *spread function*, and it corresponds to the image blurring. When using a circular aperture, the spread function  $w(r_i)$  at the image surface has been given by Airy<sup>14</sup> as

$$w(r_i) = \left( \frac{2J_1(x)}{x} \right)^2 \quad \text{with } x = 2\pi r_i \left( \frac{\alpha_i}{\lambda} \right) \quad (8.1)$$

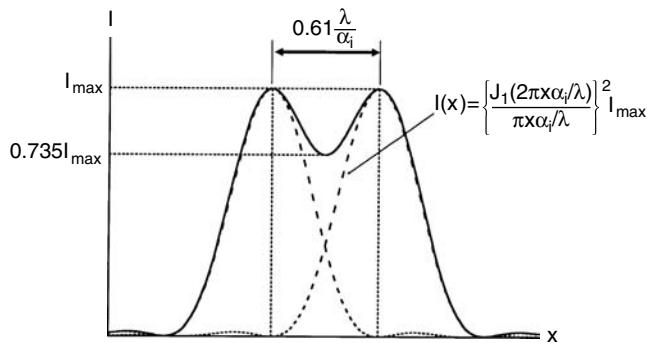
Here  $J_1(x)$  denotes the first-order Bessel function, and  $\lambda$  and  $\alpha_i$  are the wavelength and aperture half angle at the image surface, respectively. The spread function given by Equation 8.1 is called the *Airy pattern*, and it has its first zero at  $r_i = 0.61\lambda/\alpha_i$ . The disk having radius  $r_i$  is called the *Airy disk*. If we suppose an incoherent system, the intensity in the image for two neighboring pointlike objects



**FIGURE 8.3** Overlapped two monochromatic spectra.



**FIGURE 8.4** A simple model of an image-forming system.



**FIGURE 8.5** Overlapped Airy patterns showing Rayleigh's criterion.

is the sum of the spread functions for each object, as shown in Figure 8.5. By applying Rayleigh's criterion to such a system, two pointlike objects can be just resolved in the image when the distance  $x_i$  between the two objects measured in the image surface is

$$x_i = 0.61 \frac{\lambda}{\alpha_i} \quad (8.2)$$

Thus, Rayleigh's criterion for a diffraction-limited optical system corresponds to the radius of the Airy disk.

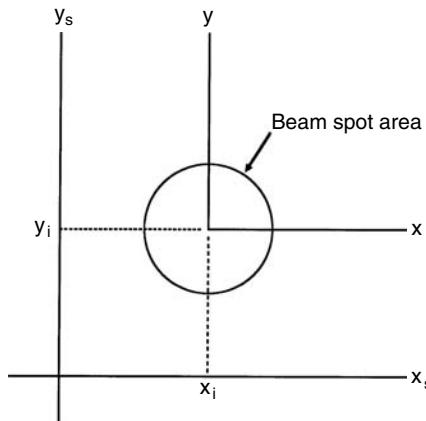
In actuality, however, it is difficult to determine the resolution clearly because it is ultimately limited by the contrast in the image which is governed by the noise (random fluctuations in intensity). Nevertheless, Rayleigh's criterion, as well as the other criteria of the resolution, is useful when comparing the optical performance between different optical systems or different optical conditions. Thus, the important matter for such a criterion of resolution is to represent the optical performance appropriately for an image-forming system.

### 8.3.2 IMAGE FORMATION FOR AN OPTICAL SYSTEM BY MEANS OF CHARGED PARTICLE BEAMS

There are several image-forming methods possible by means of charged particle beams. The most common methods are TEM and SEM. We choose for our example in this chapter the image-forming method of SEM to determine the resolution of probe-forming systems. As the same principle of image forming used in SEM is also employed in an FIB system, the results can also be applied to such a system.

An electron beam is focused into a small spot on a specimen and is scanned across the specimen in a raster, like the spot of a cathode ray tube (CRT). As secondary electrons and backscattered electrons are generated from a specimen due to interactions between the primary electrons and the specimen, they carry information about the specimen structure. Thus, it is possible to reproduce an image of the specimen structure by displaying the intensity of the signal from the specimen at corresponding positions on a CRT screen. When primary electrons hit a point  $(x_o, y_o)$  on a specimen, secondary electrons and backscattered electrons are generated not only from that point but also from the other positions neighboring the point due to the scattering of primary electrons inside the specimen. Thus, the intensity of the signal at each point  $(x_s, y_s)$  on a specimen can be written as  $s_{sp}(x_s, y_s; x_o, y_o)$  and the total intensity  $s_o(x_o, y_o)$  is given by

$$s_o(x_o, y_o) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s_{sp}(x_s, y_s; x_o, y_o) dx_s dy_s \quad (8.3)$$



**FIGURE 8.6** A model of an imaging signal produced by a beam scan.

Primary electrons are spread out in an intensity distribution  $J_b(x, y)$  on a specimen due to lens aberrations, source size, diffraction, etc. The intensity contribution  $\Delta s(x_i, y_i)$  in the area  $dx dy$  at the position  $(x_i, y_i)$ , where  $(x_i, y_i)$  corresponds to the beam center on the specimen, is given by (see Figure 8.6)

$$\Delta s(x_i, y_i) = J_b(x, y)s_o(x_i + x, y_i + y)dx dy \quad (8.4)$$

The image signal  $s(x_i, y_i)$  is then the sum of each contribution  $\Delta s(x_i, y_i)$  in the irradiated area.

Let us consider a small field A on a specimen where the beam intensity distribution  $J_b$  can be regarded as constant over A. Such a field A is called an *isoplanatism patch* of the system. The intensity  $s(x_i, y_i)$  in the isoplanatism patch is then given by

$$s(x_i, y_i) = \iint_A J_b(x, y)s_o(x_i + x, y_i + y)dx dy \quad (8.5)$$

If we suppose a rotationally symmetric distribution for  $J_b$ , that is,  $J_b(-x, -y) = J_b(x, y)$ , Equation 8.5 becomes

$$s(x_i, y_i) = \iint_A J_b(-x, -y)s_o(x_i + x, y_i + y)dx dy = \iint_A J_b(x, y)s_o(x_i - x, y_i - y)dx dy \quad (8.6)$$

Equation 8.6 shows that the intensities  $s(x_i, y_i)$  in the image are given by the convolution of the object structure  $s_o(x_o, y_o)$  with beam intensity distribution  $J_b(x, y)$ .

If the signal  $s_o(x_o, y_o)$  is taken to be a delta function, which corresponds to a pointlike object, the intensity  $s(x_i, y_i)$  corresponds to the spread function  $w(x_i, y_i)$ . Thus, we obtain

$$w(x_i, y_i) = \iint J_b(x, y)\delta(x_i - x, y_i - y)dx dy = J_b(x_i, y_i) \quad (8.7)$$

Rayleigh's criterion can thus be applied to the definition of resolution and generalized by using the intensity distribution  $J_b(x_i, y_i)$  of a beam instead of the Airy pattern. Note that a pointlike object intensity does not exist for an actual specimen because the signal from a specimen has already been spread out by scattering of the primary beam inside the specimen. Nevertheless, the pointlike object intensity is a useful assumption when estimating the optical performance.

### 8.3.3 CONVENTIONAL DEFINITIONS OF THE RESOLUTION

There are two important approaches which are widely used to determine the resolution of an optical system. One is to define it by the beam size, and the other is to define it from the spatial frequency response of an optical system calculated from the Fourier transform of the beam intensity distribution.

#### 8.3.3.1 Resolution Defined by Beam Size

This definition is based on the spread function that is assumed to be given by a Gaussian distribution. Let us define a beam radius  $r_b$  at which the intensity has fallen to  $1/e$  of the maximum. If two Gaussian distributions  $J_b(r) = J_b(0)\exp[-(r/r_b)^2]$  overlap with a separation  $2r_b$ , as shown in Figure 8.7, the overlapped profile is similar to that of the Airy patterns of Figure 8.5. Thus, the resolution corresponding to Rayleigh's criterion can be defined by the beam diameter  $2r_b$ .

This concept has been applied to the estimation of the resolution for an optical system by the method of addition of errors in quadrature,<sup>15,16</sup> and the method is widely used. In the method, diffraction, lens aberrations, source size, etc. are regarded as the independent contributions of several Gaussian distributions on the final beam intensity. The contributions of spherical aberration  $d_s$  and chromatic aberration  $d_c$  are determined from the worst-case trajectories at the disk of least confusion. Each contribution upon the beam diameter is then defined as

$$d_s = k_1 \frac{1}{2} C_{si} \alpha_i^3 \quad (8.8)$$

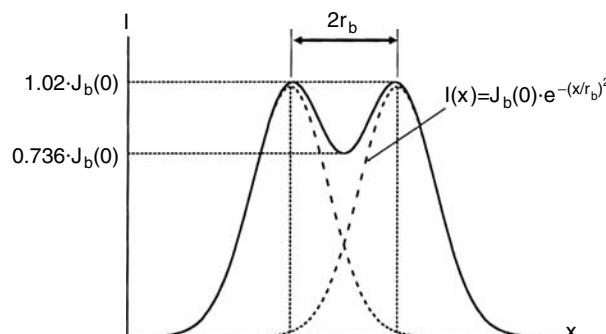
$$d_c = k_2 C_{ci} \left( \frac{\Delta V}{V_i} \right) \alpha_i \quad (8.9)$$

and

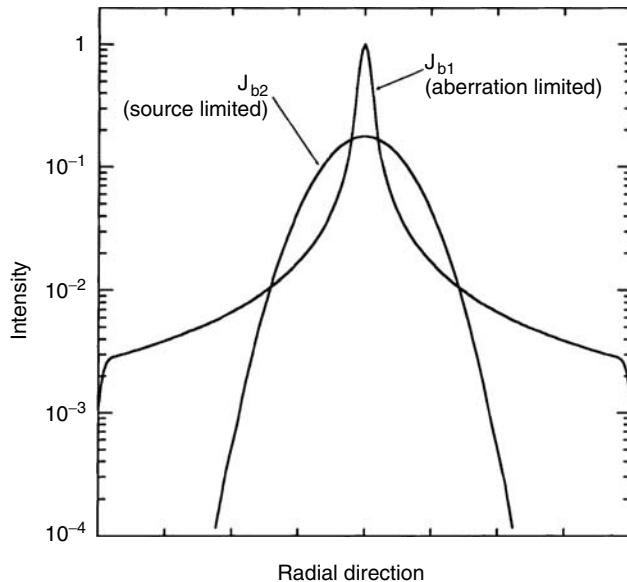
$$d_d = 0.61 \frac{\lambda}{\alpha_i} \quad (8.10)$$

respectively. Here  $C_{si}$  and  $C_{ci}$  are spherical and chromatic aberration coefficients defined at the image plane, respectively.  $V_i$  and  $\Delta V$  are the final beam energy at the specimen and energy spread, respectively. In Equations 8.8 and 8.9,  $k_1$  and  $k_2$  are constants, and the usual choice for them would be  $k_1 = k_2 = 0.5$ , for example. Then, the beam diameter  $d$  is calculated by

$$d = \sqrt{d_{ss}^2 + d_d^2 + d_s^2 + d_c^2} \quad (8.11)$$



**FIGURE 8.7** Overlapped two Gaussian distributions neighboring with distance  $2r_b$ .



**FIGURE 8.8** Beam intensity distributions having the same value of Equation 8.11 for beam diameter.

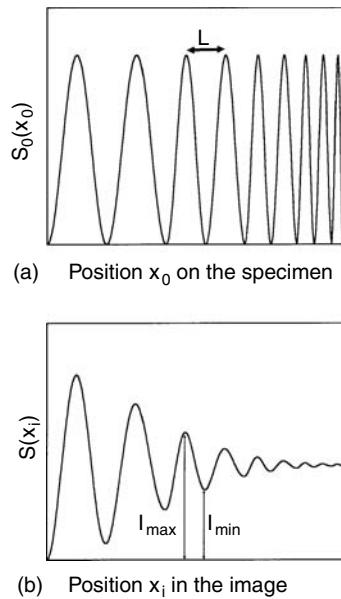
Here  $d_{ss}$  is the source diameter referred to the image (demagnified by the optical system). The resolution  $R_{es}$  of an optical system is then given by the beam diameter  $d$ .

Though the resolution can be simply calculated by Equation 8.11, one must carefully note that the results are often misleading because actually the contributions of aberrations and diffraction are not Gaussian distributions. Figure 8.8 shows the beam intensity distributions, where the beam diameters given by Equation 8.11 are the same in each case, calculated for the same beam currents. The beam intensity  $J_{b1}$  suffers from significant spherical aberration and  $J_{b2}$  suffers from significant source size. If we compare  $J_{b1}$  with  $J_{b2}$ , we see that  $J_{b1}$  will realize better resolution because the central peak of  $J_{b1}$  is much higher and sharper than that of  $J_{b2}$ . This shows that Equation 8.11 cannot be used for an accurate prediction of the resolution if the optical system suffers from aberrations.

To derive the properties of a beam which have more physical meaning than those of the worst-case trajectories, theories for aberration mixing<sup>17-19</sup> and averaging of aberrations<sup>20</sup> have been developed based on moments of the current distribution of beams. However, these methods cannot treat actual wave optical effects, as pointed out by Crewe.<sup>21</sup>

### 8.3.3.2 Definition of Resolution Based on the Contrast Performance of an Optical System

Rayleigh's criterion is determined from contrast in the image for neighboring pointlike objects. Thus, a possible way to determine the resolution of an actual optical system would be to define it in terms of the contrast performance of an optical system. Consider a signal which is derived from a beam scan over a specimen which we suppose to have an increasing sine-wave-like object structure  $s_o(x_o)$ , as shown in Figure 8.9a. The contrast  $(I_{\max} - I_{\min})/(I_{\max} + I_{\min})$  of the signal  $s(x_i)$  in the image will decrease as the distance  $L$  between lines becomes shorter until finally the separate lines cannot be resolved, because of the finite beam size as shown in Figure 8.9b. As the reciprocal of the distance between lines corresponds to the spatial frequency, Figure 8.9b shows that the contrast of an optical image decreases with increasing spatial frequency of the object structure.



**FIGURE 8.9** (a) Object structure with sine-wave amplitude of increasing spatial frequency. (b) Image structure of object in (a) showing the decreasing amplitude with increasing frequency.

The contrast in the image which depends on spatial frequency is represented by the two-dimensional Fourier transform of the intensity distribution  $s(x_i, y_i)$  in the image. As  $s(x_i, y_i)$  can be written as the convolution of the object structure  $(x_o, y_o)$  with the beam intensity distribution (see Equation 8.6), the Fourier transform of  $s(x_i, y_i)$  can be written as

$$F[s(x_i, y_i); v_x, v_y] = F[s_o(x_o, y_o); v_x, v_y] \cdot F[J_b(x, y); v_x, v_y] \quad (8.12)$$

Here  $F[f(x, y); v_x, v_y]$  denotes the two-dimensional Fourier transform of  $f(x, y)$ , and the variables  $v_x$  and  $v_y$  represent the spatial frequencies for the  $x$ - and  $y$ -directions, respectively. Therefore,  $F[f(x, y); v_x, v_y]$  is, by definition

$$F[f(x, y); v_x, v_y] = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-2\pi i(v_x x + v_y y)} dx dy \quad (8.13)$$

The second term, in the right side of Equation 8.12, corresponds to a spatial frequency response of an optical system, and it represents the contrast degradation in the image due to spread function.

For a rotationally symmetric optical system, the spatial frequency response  $\tau_b$  is defined by

$$\tau_b(v) = \frac{F[J_b(r); v]}{F[J_b(r); v]_{v=0}} \quad (8.14)$$

where  $v$  is written as

$$v = \sqrt{v_x^2 + v_y^2} \quad (8.15)$$

As the operator  $F[f(r); v]$  now represents the two-dimensional Fourier transform of a rotationally symmetric function  $f(r)$ , it can be written as

$$F[f(r); v] = \int_0^\infty f(r) J_0(2\pi v r) r dr \quad (8.16)$$

Here  $J_0(r)$  denotes the zeroth-order Bessel function.

If the intensity  $J_b(r)$  is produced by an optical system with a point source, the spatial frequency response corresponds to the *optical transfer function* (OTF). The spatial frequency response  $\tau_b(v)$  including the effect of source size is then written in terms of the OTF  $\tau(v)$  as

$$\tau_b(v) = \tau_s(v)\tau(v) \quad (8.17)$$

Here  $\tau_s(v)$  corresponds to the two-dimensional Fourier transform of the source intensity distribution  $J_s(r)$  referred to the image plane (i.e., measured at the specimen), and it is defined by

$$\tau_s(v) = \frac{F[J_s(r); v]}{F[J_s(r); v]_{v=0}} \quad (8.18)$$

Let us consider the OTF  $\tau^{\text{ideal}}(v)$  for an ideal aberration-free optical system, where  $J_b(r)$  is given by the Airy pattern, to determine the resolution corresponding to Rayleigh's criterion. For convenience, we introduce a dimensionless spatial frequency  $\bar{v}$  defined by

$$\bar{v} = \frac{\lambda}{\alpha_i} v \quad (8.19)$$

The OTF  $\tau^{\text{ideal}}(\bar{v})$  is written as (see Section 8.4.1)

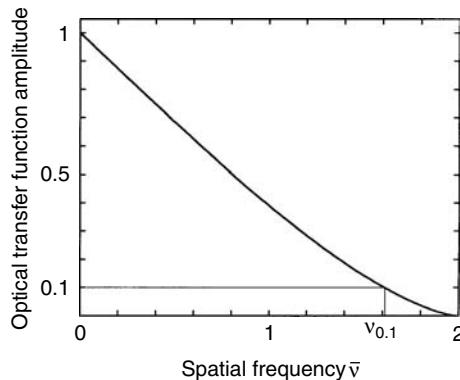
$$\tau^{\text{ideal}}(\bar{v}) = \frac{1}{\pi} (2\theta - \sin 2\theta) \quad (8.20)$$

where

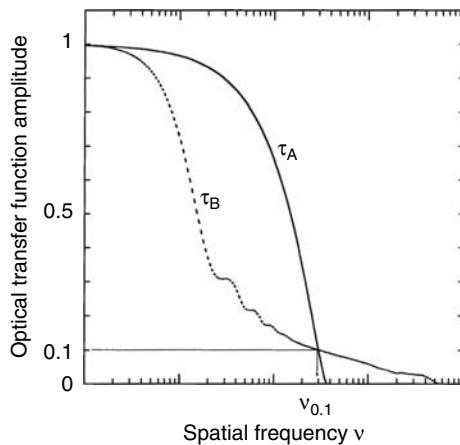
$$\theta = \cos^{-1}\left(\frac{\bar{v}}{2}\right) \quad \text{for } \bar{v} \leq 2 \quad (8.21)$$

By substituting the reciprocal of Rayleigh's criterion ( $0.61\lambda/\alpha_i$ ) for the spatial frequency  $v$  in Equation 8.19, that is,  $\bar{v} = 0.61^{-1}$ , we obtain  $\tau^{\text{ideal}}(0.61^{-1}) \cong 0.1$ . Thus, the reciprocal  $v_{0.1}^{-1}$  of the spatial frequency when  $\tau_b(v_{0.1}) = 0.1$  would be a possible definition of the resolution corresponding to Rayleigh's criterion, as shown in Figure 8.10.

This criterion seems to be used for determining the resolution of an optical system under general conditions. Nevertheless, we can encounter difficulties when determining the resolution using this criterion. Consider spatial frequency responses  $\tau_A(v)$  and  $\tau_B(v)$ , as shown in Figure 8.11, for different optical systems. The response  $\tau_A(v)$  is given by an ideal aberration-free optical system and  $\tau_B(v)$  is of a system suffering from significant spherical aberration. These responses have the same value of  $v_{0.1}^{-1}$ . Because a large number of spherical aberration produces a long tail in the beam intensity distribution, the amplitude of the response  $\tau_B(v)$  has a very low level over a wide range of spatial frequencies. Therefore, the contrast of the image produced by a system with such a large amount of spherical aberration would also be very low. In such a low-contrast image whether or not the detail of the object structure can be resolved depends on the S/N of the image. If the S/N of the



**FIGURE 8.10** Spatial frequency response of an aberration-free optical system showing the criterion of the resolution corresponding to Rayleigh's criterion.



**FIGURE 8.11** Spatial frequency responses for different optical systems having the same criterion of the resolution.

image is poor,  $\tau_A(v)$  would give better resolution than  $\tau_B(v)$  as the noise fluctuation limits the lower limit of the contrast. On the other hand, if the S/N is good enough to reproduce the detail of a low-contrast image, the resolution given by  $\tau_B(v)$  would be better because the response  $\tau_B(v)$  has nonzero amplitude at a spatial which is more than 10 times higher than that of  $\tau_A(v)$ . This suggests that the S/N must be taken into account in the definition of actual resolution.

## 8.4 A NEW DEFINITION OF THE RESOLUTION BASED ON IMAGE QUALITY

As the effect of S/N on the resolution is not taken into account in the conventional definitions presented in the previous sections, it is difficult to predict the resolution rigorously for an actual optical system. In the following paragraphs, a new concept of the resolution<sup>22</sup> taking into account the S/N is introduced in terms of the quality of an optical image.

### 8.4.1 A CONCEPT OF THE RESOLUTION IN TERMS OF THE QUALITY OF OPTICAL IMAGE

In an actual situation, one usually determines the resolution of an optical system by estimating the quality of the final image including the effect of noise fluctuations. To determine the better optical

system, one may also compare the quality of an image which is determined by the resolution and the S/N with that produced by other optical systems. Therefore, it is more accurate to estimate the resolution in terms of the quality of the final image produced by an optical system.

Let us consider a quantity  $q_{\text{image}}$  which represents the quality of an optical image. If the image is that of a pointlike object, it is characterized by resolution  $R_{\text{es}}$  of an optical system and by an S/N  $k_{\text{sn}}$  in the image, that is,

$$q_{\text{image}} = q_{\text{image}}(R_{\text{es}}, k_{\text{sn}}) \quad (8.22)$$

Now, consider two optical images for a pointlike object: one is the image produced by a supposed ideal aberration-free optical system where the spread function  $w(r_i)$  is given by an Airy pattern with radius  $0.61\lambda/\alpha_i$ . The other is the image produced by an actual system including the effects of aberrations, source size, and so on. If the quality of each image is the same at a given S/N  $k_{\text{sn}}$ , we can write

$$q_{\text{image}}(R_{\text{es}}^{\text{ideal}}, k_{\text{sn}}) = q_{\text{image}}(R_{\text{es}}, k_{\text{sn}}) \quad (8.23)$$

Here  $R_{\text{cs}}^{\text{ideal}}$  and  $R_{\text{es}}$  are the resolutions for an ideal system and for the actual system, respectively. If Rayleigh's criterion is applied to the ideal system to define the resolution, then  $R_{\text{cs}}^{\text{ideal}}$  is given by  $0.61\lambda/\alpha_i$ . The resolution  $R_{\text{es}}$  for an actual system can then be determined by

$$R_{\text{es}} = R_{\text{es}}^{\text{ideal}} = 0.61 \frac{\lambda}{\alpha_i} \quad (8.24)$$

at S/N  $k_{\text{sn}}$ . The resolution  $R_{\text{es}}$  given by Equation 8.24 means that the quality of the image produced by the optical system suffering from aberrations, etc. is the same as that produced by an ideal system with resolution (defined by Rayleigh's criterion)  $R_{\text{cs}}^{\text{ideal}}$ .

#### 8.4.2 FIGURE OF MERIT FOR THE QUALITY OF AN OPTICAL IMAGE

The *mean information content* of an optical image has been proposed by Shannon<sup>4</sup> as a figure of merit suitable for representing image quality. The details of the theory of how Shannon's concept of quantity of information can be applied to the assessment of optical images can be found in Refs 8 and 9.

According to the theory, if an optical image is regarded as a received message, an estimation of the quality of the optical system then becomes a problem in communication theory. Consider a digital imaging system where the image is produced by  $M$  pixels and  $m$  brightness levels. The system can produce  $m^M$  different images. If each image is regarded as a *message state*, and each message state is supposed to be taken as having equal prior probabilities, the information content of each image (a number corresponding to the image states represented by binary bits) is  $M \log_2 m$ . If each pixel is either black or white, for example, with probability half in each case, then the amount of information in each image is  $M$  bits. In an optical image, as the change in intensity of the image is negligibly small as the object point moves about an isoplanatic patch of the optical system, such a patch in the image can be regarded as a pixel. Also, because noise fluctuation limits the smallest detectable change of brightness, the number of meaningful brightness levels in the image takes a finite value that depends on the S/N of the image. Therefore, the information content of an optical image has a finite value which is determined by the resolution and the S/N. Thus, the parameters determining the information content are the same as those determining the quality of optical image as defined earlier.

From consideration of a digital imaging system, it is found that the resolution  $R_{es}$  is proportional to the pixel size  $r_{pix}$ . As the density of pixels is proportional to  $r_{pix}^{-2}$ , the density-of-information content of an optical image is proportional to  $R_{es}^{-2}$ .

Applying the theory of Linfoot<sup>10</sup> to charged particle optics, the mean information content  $H_A$  of an optical image in the isoplanatic patch A having an area of  $|A|$  is given by

$$H_A = \frac{|A|}{2 \ln 2} \int \int \ln \left[ 1 + \left| \frac{S(v_x, v_y)}{S(0,0)} \right|^2 \cdot \frac{P}{N} \right] dv_x dv_y \quad (8.25)$$

Here  $S(v_x, v_y)$  is the two-dimensional Fourier transform of the intensity distribution in the image. In Equation 8.25, P and N denote the statistical mean (total) power of the signal and that of the noise, respectively, measured in the isoplanatic patch A, that is, the ratio P/N corresponds to S/N of the image. As the area  $|A|$  may vary with optical system, it is more convenient to define the density  $\rho_H$  of the information content as

$$\rho_H = \frac{H_A}{|A|} = \frac{1}{2 \ln 2} \int \int \left[ 1 + \left| \frac{S(v_x, v_y)}{S(0,0)} \right|^2 \cdot \frac{P}{N} \right] dv_x dv_y \quad (8.26)$$

Let us consider an optical image for a pointlike object to estimate the resolution of an optical system. The intensity of the image then coincides with the spread function. In a primitive model of image formation, the spread function is given only by the beam intensity distribution at a specimen. In actuality, however, as the final image is displayed on a CRT screen, it is also blurred by the raster size of the CRT. If the intensity of the raster is given by a Gaussian distribution with radius  $r_{raster}$ , the spread function  $w_a(r_i)$ , referred to the specimen, due to the raster size is written as

$$w_a(r_i) = e^{-(r_i/r_a)^2} \quad (8.27)$$

where

$$r_a = \frac{r_{raster}}{M_{CRT}} \quad (8.28)$$

Here  $M_{CRT}$  is the magnification of the image on the CRT. The two-dimensional Fourier transform  $\tau_a(v)$  of the spread function  $w_a(r_i)$  produced by a receiving surface is called *the acceptance factor*. The Fourier transform of the final image for a pointlike object is then represented by multiplying the acceptance factor  $\tau_a(v)$  with the spatial frequency response  $\tau_b(v)$  of the optical system. Thus, we obtain

$$\frac{S(v_x, v_y)}{S(0,0)} = \tau_a(v)\tau_b(v) \quad \text{with } v = \sqrt{v_x^2 + v_y^2} \quad (8.29)$$

If  $w_a(r_i)$  is given by Equation 8.27,  $\tau_a(v)$  is written as

$$\tau_a(v) = e^{-(\pi r_a v)^2} \quad (8.30)$$

By introducing a new variable  $\phi$  defined by

$$v_x = v \cos \phi \quad v_y = v \sin \phi \quad (8.31)$$

Equation 8.26 for a rotationally symmetric system becomes

$$\rho_H = \frac{\pi}{\ln 2} \int_0^{+\infty} \ln \left[ 1 + |\tau_a(v)\tau_b(v)|^2 \cdot \frac{P}{N} \right] v dv \quad (8.32)$$

As a pointlike object can be regarded as an infinite information source, Equation 8.32 represents the density-of-information passing capacity (density-of-IPC) of an optical system.

In the case of a diffraction-limited optical system, it is more convenient to normalize each variable to the diffraction unit of length  $\lambda/\alpha_i$ . By replacing the variable  $v$  in Equation 8.32 with  $\bar{v} = v\lambda/\alpha_i$  as defined by Equation 8.19, we obtain

$$\rho_H = \frac{\pi}{(\lambda/\alpha_i)^2 \ln 2} \int_0^2 \ln \left[ 1 + |\tau_a(\bar{v})\tau_b(\bar{v})|^2 \cdot \frac{P}{N} \right] \bar{v} d\bar{v} \quad (8.33)$$

The density  $\rho_H$  of the IPC given by Equation 8.32 or 8.33 can thus form the figure of merit representing the quality of an optical image.

#### 8.4.3 RELATION BETWEEN RAYLEIGH'S CRITERION AND THE DENSITY-OF-INFORMATION PASSING CAPACITY

Rayleigh's criterion is determined from the Airy pattern which is produced by an ideal aberration-free optical system (for a point source). The density  $\rho_H^{\text{ideal}}$  of the IPC for an ideal system can then be calculated by replacing  $\tau_a, \tau_b$  in Equation 8.33 with  $\tau^{\text{ideal}}$  (the Fourier transform of the Airy pattern). It is given by

$$\rho_H^{\text{ideal}} = \frac{H^{\text{ideal}}}{\pi(\lambda/\alpha_i)^2} \quad (8.34)$$

Here  $H^{\text{ideal}}$  is defined by

$$H^{\text{ideal}} = \frac{\pi^2}{\ln 2} \int_0^2 \ln \left[ 1 + |\tau^{\text{ideal}}(\bar{v})|^2 \cdot \frac{P}{N} \right] \bar{v} d\bar{v} \quad (8.35)$$

$H^{\text{ideal}}$  denotes the IPC for an ideal system per area of a circle with radius  $\lambda/\alpha_i$ . As  $\tau^{\text{ideal}}$  is determined only by  $\bar{v}$  (see Equation 8.20),  $H^{\text{ideal}}$  depends only on the S/N (P/N). We see from Equation 8.34 that the resolution  $R_{es}$  corresponding to Rayleigh's criterion ( $0.6\lambda/\alpha_i$ ) is proportional to  $\rho_H^{1/2}$ . If the resolution  $R_{es}$  of an actual system is defined so as to coincide with  $0.6\lambda/\alpha_i$  when  $\rho_H = \rho_H^{\text{ideal}}$ , then it will be given by

$$R_{es} = 0.61 \frac{\lambda}{\alpha_i} \sqrt{\frac{\rho_H^{\text{ideal}}}{\rho_H}} \quad (8.36)$$

By substituting Equation 8.34 into Equation 8.36, we obtain

$$R_{es} = 0.61 \sqrt{\frac{H^{\text{ideal}}(P/N)}{\pi \rho_H}} \quad (8.37)$$

As  $\rho_H$  can be determined in general, Equation 8.37 can also be applied to the determination of the resolution in the limit of geometric optics ( $\lambda \rightarrow 0$ ), such as for an ion beam system.

#### 8.4.4 LIMITATIONS OF THE APPLICATION

The resolution given by Equation 8.37 is equivalent to Rayleigh's criterion for an ideal aberration-free optical system, so the limitation for the application of the IPC method corresponds to that of Rayleigh's criterion in an aberration-free optical system. As the S/N is not taken into account in Rayleigh's criterion by itself, it would be difficult to apply the IPC method to the determination of the resolution in most cases because a certain amount of noise is unavoidable in most observations.

### 8.5 CALCULATION OF THE DENSITY-OF- INFORMATION PASSING CAPACITY

To estimate the resolution of an optical system by using the IPC method, it is necessary to calculate the spatial frequency response  $\tau_b(v)$  of the optical system. In the following paragraphs, the derivations of the theoretical expressions of  $\tau_b(v)$  for an electron beam and for an ion beam are outlined. Though the density  $\rho_H$  of the IPC can be determined by integrating  $\tau_b(v)$  over the spatial frequency  $v$ , the actual calculation is usually very time-consuming. To avoid such a computational problem, in an actual application an approximation method for finding  $\rho_H$  at the best focus position is also presented.

#### 8.5.1 SPATIAL FREQUENCY RESPONSE FOR AN ELECTRON BEAM

The spatial frequency response  $\tau_b(v)$  of an optical system is given by the two-dimensional Fourier transform of the beam intensity distribution. As was shown in Equation 8.17,  $\tau_b(v)$  is found by multiplying the OTF  $\tau(v)$  of an optical system by the Fourier transform  $\tau_s(v)$  of the source intensity distribution (referred to the image plane); that is,  $\tau_b(v) = \tau_s(v)\tau(v)$ . The OTF  $\tau(v)$  including the effects of chromatic aberration is determined by integrating the OTF  $\tau_{\text{mono}}(v)$  for a monoenergetic beam over the energy distribution of the beam.

##### 8.5.1.1 Two-Dimensional Fourier Transform of Source Intensity Distribution

In most applications, the source intensity distribution  $B_s(r)$  is well represented by a Gaussian distribution:

$$B_s(r) = B_s(0) \cdot e^{-(r/r_{ss})^2} \quad (8.38)$$

Here  $r_{ss}$  denotes the source radius at which the intensity has fallen to  $1/e$  of the maximum. The source intensity  $J_s(r)$  referred to the image plane is written as

$$J_s(r) = J_s(0) \cdot e^{-(r/\sigma_s)^2} \quad \text{with } \sigma_s = M \cdot r_{ss} \quad (8.39)$$

Here  $M$  is the linear magnification of the optical system. The two-dimensional Fourier transform  $\tau_s(v)$ , which is normalized to  $\tau_s(0)$ , for a source intensity distribution  $J_s(r)$  is then written as

$$\tau_s(v) = e^{-(\pi\sigma_s v)^2} \quad (8.40)$$

For convenience, we define a dimensionless source size  $\bar{\sigma}_s$  by

$$\bar{\sigma}_s = \frac{\sigma_s}{D} \quad (8.41)$$

where  $D$  is a diffraction unit of length  $D = \lambda/\alpha_i \cdot \tau_s$  can be represented in terms of the dimensionless spatial frequency  $\bar{v} = vD$  as

$$\tau_s(\bar{v}) = e^{-(\pi\bar{\sigma}_s\bar{v})^2} \quad (8.42)$$

### 8.5.1.2 Optical Transfer Function for a Monoenergetic Beam

According to diffraction theory, the OTF  $\tau_{\text{mono}}$  for monoenergetic beams (beams with no chromatic aberration) is given by the autocorrelation function of the pupil function  $G$ , which is defined by<sup>23</sup>

$$G(\rho) = S_d(\rho) \exp[2\pi i \phi(\rho)] \quad (8.43)$$

Here  $i$  is the imaginary unit and  $\rho$  the radial coordinate in the pupil plane, which is normalized to the pupil radius. The domain function  $S_d(\rho)$  is defined by

$$S_d(\rho) = \begin{cases} 1 & \text{at } \rho \leq 1 \\ 0 & \text{at } \rho > 1 \end{cases} \quad (8.44)$$

In Equation 8.43,  $\phi(\rho)$  is the aberration function for the wave front. In the presence of spherical aberration and defocusing,  $\phi(\rho)$  is written as<sup>24</sup>

$$\phi(\rho) = -Ap^2 - B\rho^4 \quad (8.45)$$

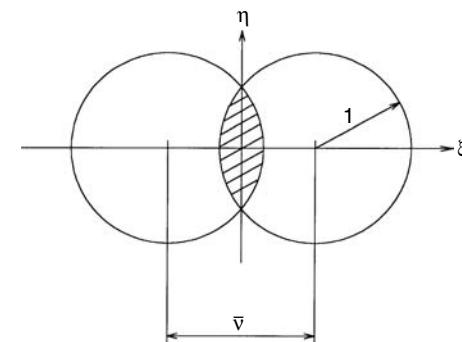
Here  $A$  and  $B$  are defined by

$$A = \frac{1}{2} \frac{z\alpha_i}{D} \quad B = \frac{1}{4} \frac{C_{si}\alpha_i^3}{D} \quad \text{with } D = \frac{\lambda}{\alpha_i} \quad (8.46)$$

In Equation 8.46,  $z$  is the axial position of the target plane measured from the Gaussian image plane. The distance between pupil functions for calculating the autocorrelation function, as shown in Figure 8.12, corresponds to the dimensionless spatial frequency  $\bar{v}$ .

Let us define an orthogonal coordinate system  $(\xi, \eta)$  in the pupil plane by  $\rho^2 = \xi^2 + \eta^2$ . The OTF  $\tau_{\text{mono}}(\bar{v})$  can then be written as

$$\tau_{\text{mono}}(\bar{v}) = C_o \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} G[(\xi + \bar{v}/2)^2 + \eta^2] \bar{G}[(\xi - \bar{v}/2)^2 + \eta^2] d\xi d\eta \quad (8.47)$$



**FIGURE 8.12** The integration domain of Equation 8.47.

where  $C_o$  is a normalizing constant that ensures  $\tau_{\text{mono}}(0) = 1$ , and  $\bar{G}(\xi, \eta)$  denotes the complex conjugate of  $G(\xi, \eta)$ . As  $S_d(\rho)$  has nonzero value only in the area of the pupil circle, the integration in Equation 8.47 can be carried out only in the hatched area shown in Figure 8.12. By substituting Equations 8.43 and 8.45 into Equation 8.47, we find that  $G(\xi, \eta)$  is odd in  $\xi$  and even in  $\eta$ . Thus, the imaginary part in the integrand of Equation 8.47 is eliminated by integration over  $\xi$ . In addition, from symmetry of the pupil function  $G$ , we obtain

$$\tau_{\text{mono}}(\bar{v}) = 4C_o \int_0^{1-\bar{v}/2} d\xi \int_{-\sqrt{1-(\xi+\bar{v}/2)^2}}^{\sqrt{1-(\xi+\bar{v}/2)^2}} \cos(2\pi\{\phi((\xi + \bar{v}/2)^2 + \eta^2) - \phi((\xi - \bar{v}/2)^2 + \eta^2)\}) d\eta \quad (8.48)$$

By substituting Equation 8.45 into Equation 8.48 and introducing the Fresnel function  $S(x)$  and  $C(x)$  defined by

$$S(x) = \int_0^x \sin\left(\frac{\pi}{2} t^2\right) dt \quad \text{and} \quad C(x) = \int_0^x \cos\left(\frac{\pi}{2} t^2\right) dt \quad (8.49)$$

the integration over  $\eta$  in Equation 8.48 can be carried out analytically, so  $\tau_{\text{mono}}(\bar{v})$  is written as

$$\tau_{\text{mono}}(\bar{v}) = 4C_o \int_0^{1-\bar{v}/2} W(\xi) d\xi \quad (8.50)$$

The function  $W(\xi)$  is written in two forms, for  $C_{si} = 0$  and for  $C_{si} \neq 0$ :

1. In the case of  $B\bar{v}\xi = 0$  ( $C_{si} = 0$ ),

$$W(\xi) = \sqrt{1 - \left(\xi + \frac{\bar{v}}{2}\right)^2} \cos(2\pi\phi_1(\xi)) \quad (8.51)$$

2. In the case of  $B\bar{v}\xi \neq 0$  ( $C_{si} \neq 0$ ),

$$W(\xi) = \frac{1}{4\sqrt{B\bar{v}\xi}} \{ \cos(2\pi\phi_1(\xi))C(\phi_2(\xi)) - \sin(2\pi\phi_1(\xi))S(\phi_2(\xi)) \} \quad (8.52)$$

The functions  $\phi_1$  and  $\phi_2$  in Equations 8.51 and 8.52 are defined by

$$\phi_1(\xi) = \bar{v}\xi(2A + B\bar{v}^2 + 4B\xi^2) \quad (8.53)$$

and

$$\phi_2(\xi) = 4\sqrt{B\bar{v}\xi(1 - (\xi + \bar{v}/2)^2)} \quad (8.54)$$

respectively. In the numerical calculations of the OTF including spherical aberration and defocusing, it is more convenient to determine the Fresnel functions  $S(x)$  and  $C(x)$  from a

power series<sup>23</sup> instead of the numerical integrations in Equation 8.49. The power series for  $S(x)$  and  $C(x)$  are as follows:

1. In the case where  $x \leq 0.7$ ,

$$S(x) = x \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!(4n+3)} \left( \frac{\pi}{2} x^2 \right)^{2n+1} \quad (8.55)$$

$$C(x) = x \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!(4n+1)} \left( \frac{\pi}{2} x^2 \right)^{2n} \quad (8.56)$$

2. In the case where  $x > 0.7$ ,

$$S(x) = \frac{1}{2} - \frac{1}{\pi x} \left[ P(x) \sin \left( \frac{\pi}{2} x^2 \right) + Q(x) \cos \left( \frac{\pi}{2} x^2 \right) \right] \quad (8.57)$$

$$C(x) = \frac{1}{2} - \frac{1}{\pi x} \left[ P(x) \cos \left( \frac{\pi}{2} x^2 \right) - Q(x) \sin \left( \frac{\pi}{2} x^2 \right) \right] \quad (8.58)$$

$Q(x)$  and  $P(x)$  are given by

$$Q(x) = \sum_{n=0}^{\infty} \frac{(-1)^n (4n-1)!!}{(\pi x^2)^{2n}} \quad (8.59)$$

$$P(x) = \sum_{n=0}^{\infty} \frac{(-1)^n (4n+1)!!}{(\pi x^2)^{2n+1}} \quad (8.60)$$

where  $(2n+1)!! = (2n+1) \times (2n-1) \cdots \times 3 \times 1$  and  $1!! = (-1)!! = 1$ .

We find  $\tau_{\text{mono}}(0) = \pi C_0$  from Equations 8.50, 8.51, and 8.53, so the normalizing constant  $C_0$  becomes  $C_0 = 1/\pi$ , and  $\tau_{\text{mono}}(\bar{v})$  can be written as

$$\tau_{\text{mono}}(\bar{v}) = \frac{4}{\pi} \int_0^{1-\bar{v}/2} W(\xi) d\xi \quad (8.61)$$

If  $\phi_1 = 0$ , that is,  $A = B = 0$ , Equation 8.61 represents the OTF  $\tau^{\text{ideal}}(\bar{v})$  for an ideal aberration-free optical system; it is

$$\tau^{\text{ideal}}(\bar{v}) = \frac{1}{\pi} (2\theta - \sin(2\theta)) \quad \text{where } \theta = \cos^{-1} \frac{\bar{v}}{2} \quad 0 \leq \bar{v} \leq 2 \quad (8.62a)$$

$$\tau^{\text{ideal}}(\bar{v}) = 0 \quad \text{for } \bar{v} > 2 \quad (8.62b)$$

### 8.5.1.3 Optical Transfer Function with Chromatic Aberration Included

In the presence of chromatic aberration, the focus position of an electron will vary from the mean focus position due to the variation of the energy from the mean. The variation  $\delta f$  of the focus position

caused by the variation of the energy  $\delta V = V - V_i$  is expressed in terms of the chromatic aberration coefficient  $C_{ci}$  as

$$\delta f = C_{ci} \frac{\delta V}{V_i} \quad (8.63)$$

where  $V$  and  $V_i$  are the energy of an electron and the mean energy of the electrons in the beam, respectively.

Let us consider a function  $F_{es}(\delta V)$  which represents the energy spread (e.g., this could be the FWHM of the energy distribution), satisfying

$$\int_{-\infty}^{\infty} F_{es}(\delta V) d\delta V = 1 \quad (8.64)$$

Energy distributions for various electron sources have been published in Refs 25 through 27.

If the optical system is incoherent, the beam intensity  $J_b(r)$  is then given by

$$J_b(r) = \int_{-\infty}^{+\infty} F_{es}(\delta V) J_{mono}(r, \delta V) d\delta V \quad (8.65)$$

$J_{mono}(r, \delta V)$  represents the beam intensity for electrons with an energy  $V = V_i + \delta V$ . By changing the order of integration over  $\delta V$  with that for the Fourier transform, the OTF  $\tau(\bar{v})$  including the effect of chromatic aberration can be written as

$$\tau(\bar{v}) = \int_{-\infty}^{+\infty} \tau_{mono}(\bar{v}, \delta V) F_{es}(\delta V) d\delta V \quad (8.66)$$

As the variation  $\delta f$  of the focus position is equivalent to the variation of axial position at the target, the parameter  $A$  can be replaced with a new parameter  $A_c$  defined by

$$A_c = \frac{1}{2} \frac{\alpha_i(z - \delta f)}{D} = \frac{1}{2} \frac{\alpha_i}{D} \left( z - C_{ci} \frac{\delta V}{V_i} \right) \quad (8.67)$$

Let us define a parameter  $\Delta A$ , which represents the effect of chromatic aberration:

$$\Delta A = \frac{1}{2} \frac{\Delta f_c \alpha_i}{D} \quad \text{with } \Delta f_c = \frac{\Delta V}{V_i} C_{ci} \quad (8.68)$$

where  $\Delta V$  denotes the FWHM of the energy spread. The parameter  $A_c$  can then be written as

$$A_c = A - \frac{\delta V}{\Delta V} \Delta A \quad (8.69)$$

The OTF  $\rho(\bar{v})$  can thus be characterized by the parameter  $A$  (defocus),  $B$  (spherical aberration), and  $\Delta A$  (chromatic aberration).

### 8.5.2 SPATIAL FREQUENCY RESPONSE FOR ION BEAMS

In the case of ion beams, as the wave effect is generally negligible as a result of the heavy mass of the ions, the intensity distribution of ion beams can thus be determined by using geometric optics. The spatial frequency response of the optical system is then obtained by calculating the two-dimensional Fourier transform of the beam intensity distribution.

The intensity distribution for an ion beam at a target is determined not only by the lens aberrations but also by the coulomb interactions in the beam. When the ion beam is used to fabricate a specimen, which is one of the important applications, the coulomb interaction is important because high current beam is required to reduce the machining time. In the case of high-resolution imaging however, as a small beam current is usually used to reduce the beam size, the coulomb interaction through the optical column need not be a dominant phenomenon. In such a case, the major effects on the probe formation are the lens aberrations and source properties, such as the source size and energy spread.

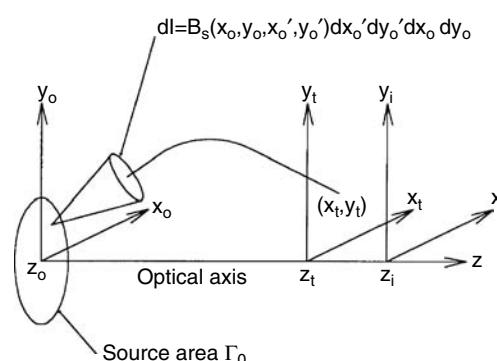
A general approach for finding the current density of beams has been given by Glaser,<sup>28</sup> but the numerical calculations involved are difficult when aberrations are present. More practical methods have been developed. Kawanami et al.<sup>29</sup> and Tuggle et al.<sup>30</sup> have investigated the beam current distribution including the effect of the energy distribution of a source. Rempfer and Mauck<sup>31</sup> have determined the current density of beams taking into account the effects of source intensity distribution and spherical and chromatic aberrations by using the analytical results of calculating the trajectories given by geometric optics. In the following paragraphs, a calculation method for beam current densities presented in Ref. 32 is outlined.

#### 8.5.2.1 Mathematical Model for Current Density of Monoenergetic Beams

Let us introduce a set of coordinates  $(x, y, z)$ , where  $z$  corresponds to the optical axis and  $(x_o, y_o, z_o)$ ,  $(x_i, y_i, z_i)$ , and  $(x_t, y_t, z_t)$  which refer to the object plane (at the source), the Gaussian image plane, and the target plane, respectively, as shown in Figure 8.13. We define the origin  $z = 0$  to be at the Gaussian image plane.

Consider the beam current  $dI$  emitted from an incremental source  $dx_o dy_o$  located at the object plane into the solid angle  $dx'_o dy'_o$ , where the prime denotes differentiation with respect to  $z$ . The beam current  $dI$  can then be written as

$$dI = B_s(x_o, y_o, x'_o, y'_o) dx'_o dy'_o \quad (8.70)$$



**FIGURE 8.13** A schematic diagram showing the mathematical model for current density of monoenergetic beams.

Here,  $B_s$  denotes the current distribution of the source which is, in general, a function of the position and the direction of trajectories at the object plane. The total beam current emitted from the source can be written as an integral over the area  $\Gamma_o$ , where  $\Gamma_o$  denotes the boundary of the source:

$$I_p = \iint_{(x_o, y_o) \in \Gamma_o} \iint_{(x'_o, y'_o) \in \Omega_o} B_s(x_o, y_o, x'_o, y'_o) dx'_o dy'_o dx_o dy_o \quad (8.71)$$

where  $\Omega_o$  denotes the solid angle subtended at the object plane by the aperture stop. To calculate the current density at the target plane, we specify coordinates  $x_t$  and  $y_t$  and then determine all the trajectories with initial coordinates  $x_o$  and  $y_o$  at the object plane which contribute to the current at  $x_t$  and  $y_t$ . The position of the rays at the object plane is given by

$$\begin{aligned} x_o &= x_o(x_t, y_t, x'_o, y'_o) \\ y_o &= y_o(x_t, y_t, x'_o, y'_o) \end{aligned} \quad (8.72)$$

Equation 8.71 can then be written as

$$I_p = \iint_{(x_t, y_t) \in \Gamma_t} \iint_{(x'_o, y'_o) \in \Omega_o} |K| B_s(x_o, y_o, x'_o, y'_o) dx'_o dy'_o dx_t dy_t \quad (8.73)$$

where  $|K|$  is the absolute value of the Jacobian which is defined by

$$K = \begin{vmatrix} \frac{dx_o}{dx_t} & \frac{dx_o}{dy_t} \\ \frac{dy_o}{dx_t} & \frac{dy_o}{dy_t} \end{vmatrix} \quad (8.74)$$

In Equation 8.73, the integral area  $\Gamma_t$  is a set of solutions  $(x_t, y_t)$  satisfying

$$(x_o(x_t, y_t, x'_o, y'_o), y_o(x_t, y_t, x'_o, y'_o)) \in \Gamma_o \quad (8.75)$$

Now that the position  $(x_o, y_o)$  in Equation 8.73 is now determined by  $(x_t, y_t)$  and  $(x'_o, y'_o)$ . The current density of the monoenergetic beam  $J_{\text{mono}}(x_t, y_t)$  at the target can then be obtained from

$$J_{\text{mono}}(x_t, y_t) = \iint_{(x'_o, y'_o) \in \Omega'_o} |K| B_s(x_o(x_t, y_t, x'_o, y'_o), y_o(x_t, y_t, x'_o, y'_o), x'_o, y'_o) dx'_o dy'_o \quad (8.76)$$

Here the integral area  $\Omega'_o$  denotes a set of values  $(x'_o, y'_o)$ , satisfying

$$(x_o(x_t, y_t, x'_o, y'_o), y_o(x_t, y_t, x'_o, y'_o)) \in \Gamma_o \quad \text{and} \quad (x'_o, y'_o) \in \Omega'_o \quad (8.77)$$

Thus, the integration indicated in Equation 8.76 can be carried out if the relations in Equation 8.72 are given analytically.

### 8.5.2.2 Current Density of Monoenergetic Beams for a Rotationally Symmetric System

$$\begin{aligned} x'_o &= \alpha_o \rho \cos \theta \\ y'_o &= \alpha_o \rho \sin \theta \end{aligned} \quad (8.78)$$

where  $\alpha_o$  is the beam divergence half angle at the source. According to Ref. 32, the relation Equation 8.72 becomes

$$\begin{aligned} x_o &= \frac{x_t}{M} - \left( \frac{zM_a \alpha_o}{M} \rho + C_{so} \alpha_o^3 \rho^3 \right) \cos \theta \\ y_o &= \frac{y_t}{M} - \left( \frac{zM_a \alpha_o}{M} \rho + C_{so} \alpha_o^3 \rho^3 \right) \sin \theta \end{aligned} \quad (8.79)$$

Here  $M$  and  $M_a$  represent the linear and the angular magnifications of the optical system, respectively, and  $C_{so}$  denotes the spherical aberration coefficient defined at the object plane. The Jacobian is then given by

$$K = \left( \frac{\alpha_o}{M} \right)^2 \rho \quad (8.80)$$

Let us assume a homogeneous source angular distribution. The source distribution function  $B_s$  can then be written as

$$B_s = \begin{cases} B_s(r_o) & \text{for } r_o \leq R_o \\ 0 & \text{for } r_o > R_o \end{cases} \quad (8.81)$$

where  $r_o = (\sqrt{x_o^2 + y_o^2})$  denotes the radial coordinate in the object plane and  $R_o$  is the physical radius of the source.

We can replace  $\alpha_o$  at the source by corresponding angle  $\alpha_i$  at Gaussian image plane using the relation  $\alpha_i = M_a \alpha_o$ . Then, the spherical aberration at the object can be represented in terms of the parameters at the image by

$$C_{so} \alpha_o^3 = \frac{C_{si} \alpha_i^3}{M} \quad (8.82)$$

The radial coordinate  $r_t$  in the target plane can be written as  $r_t = x_t$  when  $y_t = 0$ . Then, from Equation 8.79, the trajectories at the object plane, normalized to the radius of the source, can be written as

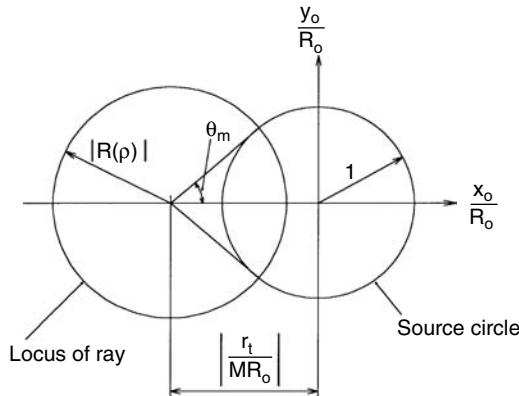
$$\left( \frac{r_o}{R_o} \right)^2 = R^2(\rho) + \left( \frac{r_t}{MR_o} \right)^2 - 2R(\rho) \left( \frac{r_t}{MR_o} \right) \cos \theta \quad (8.83)$$

with

$$R(\rho) = \frac{z\alpha_i}{MR_o} \rho + \frac{C_{si} \alpha_i^3}{MR_o} \rho^3 \quad (8.84)$$

The current density for a monoenergetic beam at the target can then be written as

$$J_{mono}(r_t, z) = \left( \frac{\alpha_o}{M} \right)^2 \int_0^{1/2\pi} \int_0^R B_s \left[ r_o \left( \frac{r_t}{MR_o}, \frac{z\alpha_i}{MR_o}, \frac{C_{si}}{MR_o}, \rho, \theta \right) \right] \rho d\theta d\rho \quad (8.85)$$



**FIGURE 8.14** A relation between a locus of the charged particle rays at the object plane and the normalized source circle.

However, in actuality, the source physical area imposes limits on the ranges of  $\rho$  and  $\theta$  because  $B_s \neq 0$  only within the area  $\Gamma_o$ . Therefore, the numerical calculation indicated in Equation 8.85 is carried out only in the domain  $B_s \neq 0$ , that is,

$$\left| \frac{r_o}{R_o} \right| < 1 \quad (8.86)$$

The domains for  $\rho$  and  $\theta$  which satisfy Equation 8.86 can be obtained by using the relation between the locus of Equation 8.83 (which is defined by  $\rho = \text{constant}$  at a given value of  $r_t$ ) and the source area as shown in Figure 8.14. We find, from Figure 8.14, that the integral boundary  $\Gamma_p$  for  $\rho$  is given by the solution of

$$\left| \frac{r_t}{MR_o} \right| - \leq |R(\rho)| \leq \left| \frac{r_t}{MR_o} \right| + 1 \quad \text{and} \quad 0 \leq \rho \leq 1 \quad (8.87)$$

The boundary for the parameter  $\theta$  for a given value of  $\rho$  can also be determined from Figure 8.14:

$$-\theta_m \leq \theta_m \quad \text{where } \theta_m = \cos^{-1} \frac{\left( \frac{r_t}{MR_o} \right)^2 + R^2(\rho) - 1}{\left| R(\rho) \frac{2r_t}{MR_o} \right|} \quad (8.88)$$

Equation 8.85 can then be written as

$$J_{\text{mono}}(r_t, z) = 2 \left( \frac{\alpha_o}{M} \right)^2 \int_{\Gamma_p} \int_0^{\theta_m} B_s[r_o(r_t, z, \rho, \theta)] \rho d\theta d\rho \quad (8.89)$$

When chromatic aberration is present, the beam current density  $J_b$  is obtained in terms of  $J_{\text{mono}}$  and the energy distribution  $F_{es}(\delta V)$  in the same way as in the calculation of the OTF. Thus, the beam

current density  $J_b(r_t, z_t)$  at the target plane ( $z = z_t$ ) including the effects of chromatic aberration is determined by

$$J_b(r_t) = \int_{-\infty}^{+\infty} F_{es}(\delta V) J_{mono} \left[ \frac{r_t}{MR_o}, \frac{C_{si}}{MR_o}, \frac{\alpha_i}{MR_o} \left( z_t - \frac{\Delta V}{\Delta V} \Delta f_c \right) \right] d\delta V \quad (8.90)$$

with

$$\Delta f_c = C_{ci} \frac{\Delta V}{V_i} \quad (8.91)$$

The nature of the energy distribution  $F_{es}(\delta V)$  depends on the source type and total emission current. Distributions for the gallium liquid metal ion source have been published by Swanson et al.<sup>33</sup> and Ishitani et al.,<sup>34</sup> for example. In a numerical calculation, the region of integration in Equation 8.90 would, of course, be replaced by finite integration boundaries as determined by application.

Let us assume a Gaussian distribution with a source radius  $r_{ss}$  and a source area with radius  $R_o$  of  $2r_{ss}$ . The source distribution function  $B_s(r_o)$  is then described by the truncated Gaussian

$$B_s(r_o) = \begin{cases} B_s(0) e^{-(r_o/r_{ss})^2} & r_o \leq 2r_{ss} \\ 0 & r_o > 2r_{ss} \end{cases} \quad (8.92)$$

We see from Equation 8.90 that the beam current distribution  $J_b$  for a rotationally symmetric system can be characterized by the terms  $A_g$ ,  $B_g$ , and  $\Delta A_g$  defined analogously to wave optics as

$$A_g = \frac{1}{2} \frac{z\alpha_i}{\sigma_s} \quad B_g = \frac{1}{4} \frac{C_{si}\alpha}{\sigma_s} \quad \text{and} \quad \Delta A_g = \frac{1}{2} \frac{\Delta f_c \alpha_i}{\sigma_s} \quad (8.93)$$

The quantity  $\sigma_s$  denotes the source radius referred to the image plane, that is,

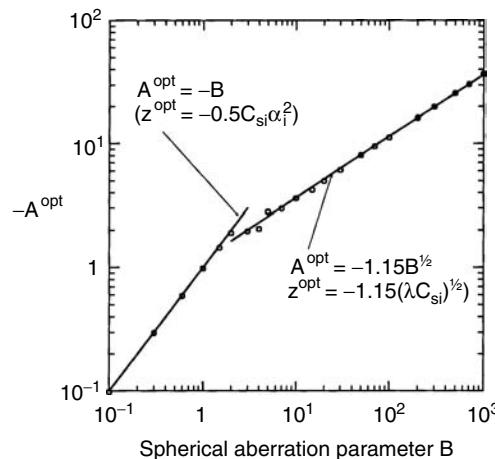
$$\sigma_s = M r_{ss} \quad (8.94)$$

### 8.5.3 PROPERTIES OF THE BEST FOCUS POSITION

The determination of the best focus position ( $z^{opt}$ ) is essential whenever estimating the resolution of an optical system, because all the results depend on the focus position. Thus, the resolution should be estimated at the best focus position. Consider a focus position when the focus is optimally adjusted by observing the image. The focus position can be determined by where the highest quality image is obtained. Such a focus position is given at the point where the density-of-IPC is a maximum. We discuss here the behavior of the best focus positions for electron and ion beams.

#### 8.5.3.1 Best Focus Position for Electron Beams

In a diffraction-limited system, the focus position is given by the parameter  $A$  defined by Equation 8.46. Thus, the best focus position  $z^{opt}$  is determined by  $A^{opt}$  which is determined in such a way that the density-of-IPC has a maximum value. Figure 8.15 shows the behavior of  $A^{opt}$  ( $= 0.5z^{opt}\alpha_i/D$ ) for an optical system with a point source for various conditions of spherical aberration (see Equation 8.46). When spherical aberration is present, the best focus position shifts toward the underfocus



**FIGURE 8.15** Behavior of the best focus condition  $A^{\text{opt}}$  for a diffraction-limited optical system as a function of the spherical aberration B.

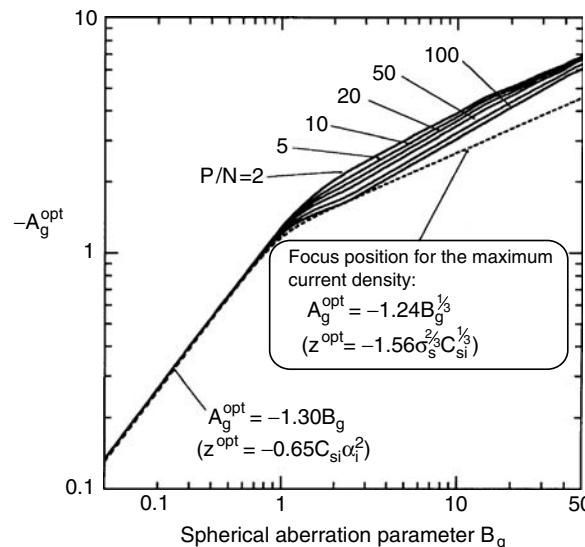
side ( $z < 0$ ) relative to the Gaussian image plane, as shown in Figure 8.15. If the spherical aberration is not dominant in the optical system ( $B \ll 1$ ),  $z^{\text{opt}}$  is found to be  $z^{\text{opt}} = 0.5C_{\text{si}}\alpha_i^2$ . This focus position is close to the position for the disk of least confusion, which is given by  $z = -0.75C_{\text{si}}\alpha_i^2$ . When spherical aberration becomes dominant in the optical system ( $B \gg 1$ ), on the other hand, the best focus position becomes independent of the beam convergence half angle  $\alpha_i$  but depends on the wavelength  $\lambda$ , as shown in the figure. Thus, the divergence of the best focus position from that for the disk of least confusion becomes larger with increasing spherical aberration ( $C_{\text{si}}\alpha_i^3$ ).

### 8.5.3.2 Best Focus Position for Ion Beams

In the case of ion beams, the best focus position for the maximum density-of-IPC almost coincides with the position at which the beam intensity is maximum. The focus position for ion beams can be represented by the parameter  $A_g^{\text{opt}}$  ( $= 0.5z^{\text{opt}}\alpha_i/\sigma_s$ ) corresponding to the best focus position for various values of  $B_g$  ( $= 0.25C_{\text{si}}\alpha_i^3/\sigma_s$ ) at various S/Ns. In this case, the best focus position depends on the source size when the spherical aberration is dominant in the optical system. Comparing Figure 8.15 with Figure 8.16, the behavior of the best focus position for ion beams (given by geometric optics) is similar to that for electron beams (given by wave optics), and we find that the source radius  $\sigma_s$  in geometric optics has a meaning similar to the diffraction unit of length  $D = \lambda/\alpha_i$  in wave optics.

### 8.5.4 APPROXIMATION METHOD FOR THE DENSITY-OF-INFORMATION PASSING CAPACITY

Although the spatial resolution of an optical system with a given S/N can be evaluated accurately by calculating the density-of-IPC at the best focus position, the numerical calculations are usually very time-consuming. To avoid computational problems, an approximation method is very useful for practical applications. A method for approximating the density-of-IPC at the best focus position has been developed<sup>35</sup> by means of functional fitting. The method can be applied to the calculation of the IPC at various S/Ns (P/N) (2, 5, 10, 20, 50, and 100 were chosen). The properties of beams given by geometric and wave optics are synthesized<sup>36</sup> to calculate the IPC in situations ranging from the pure geometric condition ( $\lambda = 0$ ) to the pure diffraction condition ( $\sigma_s = 0$ ). We outline here the approximation method presented in Refs 35 and 36 and how the properties of both wave and geometric optics are synthesized.



**FIGURE 8.16** Characteristic of the best focus condition  $A_g^{\text{opt}}$  (solid lines), where the density-of-information passing capacity is a maximum, as a function of the spherical aberration parameter  $B_g$  at selected values of signal-to-noise ratio. The dotted line shows the focus condition  $A_g$  for the maximum current density at the beam center.

#### 8.5.4.1 Basis of the Concept for the IPC Approximation

The purpose of method is to find analytical functions which can fit the IPC at the best focus position, accurately, in terms of all the parameters which characterize the IPC. The fitting functions can be found by using the results of the IPC obtained by carrying out the numerical calculations at various values for the parameters. In general, the IPC is characterized by the parameters A (defocus), B (spherical aberration),  $\Delta a$  (chromatic aberration), D (diffraction unit of length),  $\sigma_s$  (source radius),  $\tau_a$  (acceptance factor), and P/N (S/N). However, as we calculate the IPC only at the best focus position, the parameter A can be omitted from independent variables because it is determined by B,  $\Delta A$ , D,  $\sigma_s$ , and  $\tau_a$ . Also, if the source intensity and acceptance factor are both modeled by a Gaussian distribution, the effect of  $\tau_a$  can be subsumed into  $\sigma_s$  because the multiplication of two Gaussian distributions results in a Gaussian distribution.

By using the similarity between D and  $\sigma_s$  in wave and geometric optics, as presented in the previous section, the effect of source size ( $\sigma_s$ ) can be subsumed into the quantities  $D'$ ,  $B'$ , and  $\Delta A'$  corresponding to D, B, and  $\Delta A$ , respectively. The method of synthesis between wave and geometric optics is to derive parameters such as  $D'$ ,  $B'$ , and  $\Delta A'$  in which the effect of  $\sigma_s$  is subsumed. Thus, only  $D'$ ,  $B'$ , and  $\Delta A'$  finally remain as independent variables, and the density of the IPC  $\rho_H$  is given by

$$\rho_H = \frac{H_p(B', \Delta A', P/N)}{\pi D'^2} \quad (8.95)$$

where  $H_p$  is the IPC within a circular area with radius  $D'$ , which is given by

$$H_p(B', \Delta A', P/N) = \frac{\pi^2}{\ln 2} \int_0^2 \ln \left[ 1 + |\tau(B', \Delta A', A^{\text{opt}}, \bar{v})|^2 \cdot \frac{P}{N} \right] \bar{v} d\bar{v} \quad (8.96)$$

Fitting functions can then be determined separately for each S/N by using the numerically calculated results of the IPC given by Equation 8.96 at various values of  $B'$  and  $\Delta A'$ .

### 8.5.4.2 Synthesis of Wave and Geometric Optics

#### 8.5.4.2.1 Equivalent Optical System in Wave and Geometric Optics

Let us consider two optical systems: one is an aberration-free optical system for a point source, in wave optics (different limited), where the optical system is characterized by the diffraction unit of length  $D$ . The other is an aberration-free optical system, in geometric optics, with a finite source size, where the source radius referred to the specimen is given by  $\sigma_g$  that can be determined so that the density  $\rho_H$  of the IPC in geometric optics is the same as that of the system in wave optics at a given S/N (P/N). The source radius  $\sigma_g$  can then be written as

$$\sigma_g = \bar{\sigma}_g \cdot D \quad (8.97)$$

The coefficient  $\bar{\sigma}_g$  depends slightly on the S/N, as shown in Figure 8.17.

Next, consider the case where spherical and chromatic aberrations are present. As the beam intensity distribution, in geometric optics, is characterized by the parameters defined by Equation 8.93, the parameters  $B_g^{\text{eff}}$  and  $\Delta A_g^{\text{eff}}$  in geometric optics, defined by

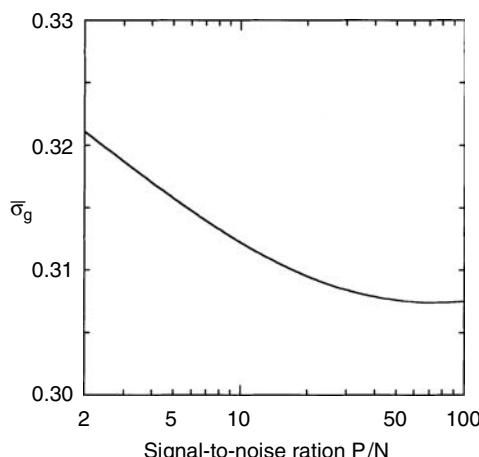
$$B_g^{\text{eff}} = \frac{1}{4} \frac{C_{si}\alpha_i^3}{\sigma_g/\bar{\sigma}_g} \quad (8.98)$$

$$\Delta A_g^{\text{eff}} = \frac{1}{2} \frac{\Delta f_c \alpha_i}{\sigma_g/\bar{\sigma}_g} \quad (8.99)$$

also characterize  $\rho_H$  at a given S/N. Thus,  $B_g^{\text{eff}}$  and  $\Delta A_g^{\text{eff}}$  can be determined as

$$B_g^{\text{eff}} = f_b(B) \quad \text{and} \quad \Delta A_g^{\text{eff}} = f_a(\Delta A) \quad (8.100)$$

so that both optical systems, including the effects of aberrations, result in the same value of  $\rho_H$ . Note that the aberrations such as  $C_{si}\alpha_i^3$  or  $\Delta f_c \alpha_i$  in geometric optics may differ from those in wave optics.



**FIGURE 8.17** Behavior of  $\bar{\sigma}_g$  showing a dependency on the signal-to-noise ratio P/N.

Here, only the values of  $B_g^{\text{eff}}$  and  $\Delta A_e^{\text{eff}}$  are important with respect to the density-of-IPC. Obviously, because the phenomena in wave optics are well described by geometric optics when the aberrations are large compared with the diffraction unit of length, we see

$$f_b(B) \cong B \quad \text{for } B \gg 1 \quad (8.101)$$

and

$$f_b(\Delta A) \cong \Delta A \quad \text{for } \Delta A \gg 1 \quad (8.102)$$

By using the relations given by Equations 8.97 and 8.100,  $\rho_H$  for the system in geometric optics can be obtained from

$$\rho_H = \frac{H_p(f_b^{-1}(B_g^{\text{eff}}), f_a^{-1}(\Delta A_e^{\text{eff}}), P/N)}{\pi(\sigma_g / \sigma_g)^2} \quad (8.103)$$

Now, consider the case where a source radius  $\sigma_s$  is present in the diffraction-limited optical system. The corresponding source radius  $\sigma_g$  in geometric optics, for which there results the same  $\rho_H$  as that in wave optics, then becomes

$$\sigma_g = \sqrt{(\bar{\sigma}_g D)^2 + (\varepsilon_d(\bar{\sigma}_s) \sigma_s)^2} \quad (8.104)$$

Here  $\varepsilon_d$  is an equivalency function with respect to  $\bar{\sigma}_s$ , and  $\bar{\sigma}_s$  is defined by

$$\bar{\sigma}_s = \frac{\sigma_s}{D} \quad (8.105)$$

As the equivalency represents the difference in behavior between wave and geometric optics, it takes the value 1 when both theories yield the same result, that is, when the phenomenon is well described by geometric optics.

The parameters  $B_g^{\text{eff}}$  and  $\Delta A_g^{\text{eff}}$ , in geometric optics, are also influenced by the source size  $\sigma_s$ . They are found to be

$$B_g^{\text{eff}} = \frac{D}{\sqrt{D^2 + (\varepsilon_b^{\text{eff}} \sigma_s / \bar{\sigma}_g)^2}} f_b(B) \quad (8.106)$$

and

$$\Delta A_g^{\text{eff}} = \frac{D}{\sqrt{D^2 + (\varepsilon_a^{\text{eff}} \sigma_s / \bar{\sigma}_g)^2}} f_a(\Delta A) \quad (8.107)$$

respectively. Here  $\varepsilon_a^{\text{eff}}$  and  $\varepsilon_b^{\text{eff}}$  are other equivalency functions, with respect to  $B$ ,  $\Delta A$ , and  $\sigma_s$ . These equivalency functions can be written in the following forms:

$$\varepsilon_a^{\text{eff}} = \frac{\varepsilon_a(T)}{\varepsilon_d(\bar{\sigma}_s)} \quad (8.108)$$

$$\varepsilon_b^{\text{eff}} = \varepsilon_d(\bar{\sigma}_s) \varepsilon_b(T) \quad (8.109)$$

Here T is defined by

$$T = \sqrt{B^2 + \Delta A^4} \quad (8.110)$$

The optical system in geometric optics, with the parameters  $\sigma_g$ ,  $B_g^{eff}$ , and  $\Delta A_g^{eff}$  now results in the same density  $\rho_H$  of the IPC as that of the system in wave optics, including the effects of  $\sigma_s$ , D, B, and  $\Delta A$ . Thus, the parameters  $D'$ ,  $B'$ , and  $\Delta A'$ , in which the effect of source size  $\sigma_s$  is subsumed, can be determined by

$$D' = \sigma_g / \bar{\sigma}_g \quad (8.111)$$

$$B' = f_b^{-1}(B_g^{eff}) \quad \text{and} \quad \Delta A' = f_a^{-1}(\Delta A_g^{eff}) \quad (8.112)$$

The algorithm for the approximation of  $\rho_H$  presented in this section is shown in Figure 8.18. From Equations 8.101 and 8.102, we see that

$$\lim_{\lambda \rightarrow 0} \frac{Df_b(B)}{\sqrt{D^2 + (\epsilon_b^{eff} \sigma_s / \bar{\sigma}_g)^2}} = \frac{1}{4} \frac{C_{si} \alpha_i^3}{\sigma_s / \bar{\sigma}_g} \quad (8.113)$$

and

$$\lim_{\lambda \rightarrow 0} \frac{Df_b(\Delta A)}{\sqrt{D^2 + (\epsilon_a^{eff} \sigma_s / \bar{\sigma}_g)^2}} = \frac{1}{2} \frac{\Delta f_c \alpha_i}{\sigma_s / \bar{\sigma}_g} \quad (8.114)$$

This shows that the algorithm shown in Figure 8.18 yields  $\rho_H$  under the continuum of conditions ranging from wave optics to geometric optics.

#### 8.5.4.2.2 Fitting Functions for the Equivalency Functions

The fitting functions for all equivalency functions have been found<sup>35</sup> in numerical calculations of the IPC for six values of the S/N: 2, 5, 10, 20, 50, and 100. The parameter  $\bar{\sigma}_g$  and those used in the following fitting functions are listed in Table 8.1.

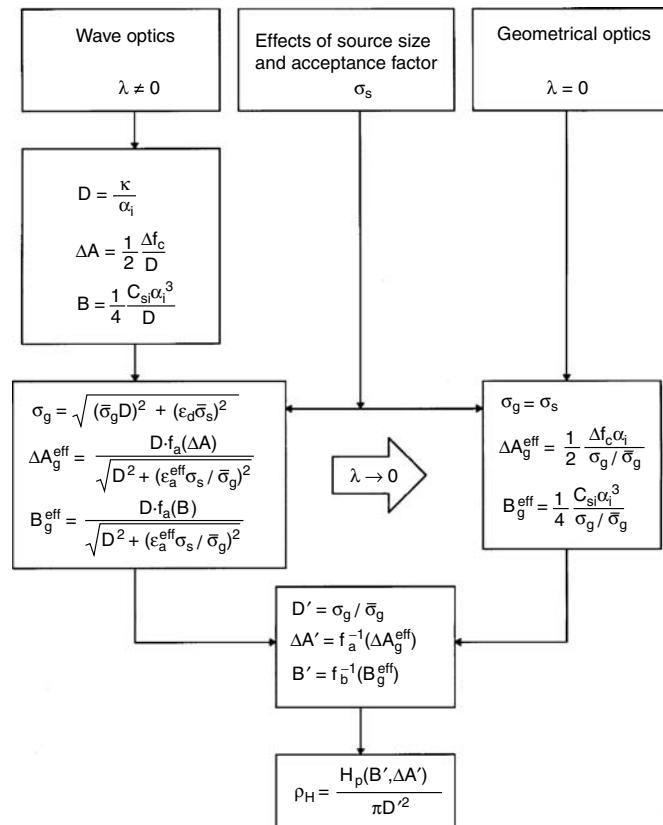
1. The fitting functions for  $f_a(\Delta A)$ ,  $f_b(B)$ ,  $f_a^{-1}(\Delta A_g^{eff})$ , and  $f_a^{-1}(B_g^{eff})$  are

$$f_a(\Delta A) = \frac{-[f_{a0}(f_{a1} + f_{a2}) - \Delta A] + \sqrt{(f_{a0}(f_{a1} + f_{a2}) - \Delta A)^2 + 4f_{a0}f_{a2}\Delta A}}{2f_{a0}} \quad (8.115)$$

$$f_a(B) = \frac{-[f_{b0}(f_{b1} + f_{b2}) - B] + \sqrt{(f_{b0}(f_{b1} + f_{b2}) - B)^2 + 4f_{b0}f_{b2}B}}{2f_{b0}} \quad (8.116)$$

$$f_a^{-1}(\Delta A_g^{eff}) = f_{a0} \left( 1 + \frac{f_{a1}}{f_{a2} + \Delta A_g^{eff}} \right) \Delta A_g^{eff} \quad (8.117)$$

$$f_a^{-1}(B_g^{eff}) = f_{a0} \left( 1 + \frac{f_{a1}}{f_{a2} + B_g^{eff}} \right) B_g^{eff} \quad (8.118)$$



**FIGURE 8.18** Algorithm for the approximation of  $\rho_H$ .

**TABLE 8.1**  
**List of Fitting Parameters for Synthesizing Properties**

	Signal-to-Noise Ratio (P/N)					
	2	5	10	20	50	100
$\bar{\sigma}_g$	0.321	0.316	0.312	0.310	0.308	0.308
$\epsilon_{d0}$	$4.72 \times 10^{-2}$	$4.86 \times 10^{-2}$	$5.06 \times 10^{-2}$	$5.57 \times 10^{-2}$	$6.77 \times 10^{-2}$	$7.38 \times 10^{-2}$
$\epsilon_{d1}$	1.27	1.27	1.25	1.32	1.52	1.47
$\epsilon_{d2}$	-1.10	-0.97	-0.87	-0.73	-0.48	-0.39
$\epsilon_{d3}$	1.04	0.95	0.80	0.62	1.60	1.40
$\epsilon_{d4}$	1.11	1.08	1.06	1.03	1.00	0.97
$f_{a0}$	1.14	1.16	1.17	1.18	1.21	1.25
$f_{a1}$	0.59	0.58	0.63	0.71	0.83	0.93
$f_{a2}$	0.54	0.40	0.36	0.33	0.32	0.33
$f_{b0}$	1.13	1.11	1.10	1.05	1.02	0.99
$f_{b1}$	1.65	1.75	1.79	2.09	2.25	2.35
$f_{b2}$	0.74	0.65	0.56	0.55	0.48	0.42
$\epsilon_{a1}$	2.5	2.9	3.0	2.2	1.5	1.0
$\epsilon_{a2}$	2.0	1.1	0.8	0.5	0.4	0.4
$\epsilon_{b1}$	0.53	0.44	0.37	0.29	0.21	0.17
$\epsilon_{b2}$	0.73	0.83	0.93	1.08	1.29	1.74

2. The fitting function for  $\varepsilon_d(\bar{\sigma}_s)$  is

$$\varepsilon_d(B_g^{\text{eff}}) = f_1(\bar{\sigma}_s) + f_2(\bar{\sigma}_s) \quad (8.119)$$

where the functions  $f_1(\bar{\sigma}_s)$  and  $f_2(\bar{\sigma}_s)$  are defined by

$$f_1(\bar{\sigma}_s) = 1 + \varepsilon_{d0} \cdot e \exp \left[ - \left( \frac{\ln \bar{\sigma}_s - \varepsilon_{d2}}{\varepsilon_{d1}} \right)^2 \right] \quad (8.120)$$

$$f_2(\bar{\sigma}_s) = \frac{2(\varepsilon_{d4} - 1)}{1 + \exp(\varepsilon_{d3} \cdot \bar{\sigma}_s^{1/2})} \quad (8.121)$$

3. The fitting functions for  $\varepsilon_a(T)$  and  $\varepsilon_b(T)$  are

$$\varepsilon_a(T) = \frac{\varepsilon_{a1} \cdot T^{\varepsilon_{a2}}}{1 + \varepsilon_{a1} \cdot T^{\varepsilon_{a2}}} \quad (8.122)$$

$$\varepsilon_b(T) = \frac{\varepsilon_{b1} \cdot T^{\varepsilon_{b2}}}{1 + \varepsilon_{b1} \cdot T^{\varepsilon_{b2}}} \quad (8.123)$$

#### 8.5.4.3 The Fitting Functions for the IPC

The fitting functions that represent  $H_p$  can be determined using the results obtained by carrying out the numerical integration indicated in Equation 8.96 for various values of  $B$ ,  $\Delta A$ , and P/N. As was shown in Ref. 35,  $H_p$  can be fitted accurately by the following expressions:

$$H_p^{-1} = \left[ \left\{ (H_{00}^{-1})^{n_{H_b}} + (H_b^{-1})^{n_{H_b}} \right\}^{n_{H_a}/n_{H_b}} + (H_a^{-1})^{n_{H_a}} \right]^{1/n_{H_a}} + H_{ab}^{-1} \quad (8.124)$$

The terms  $H_a^{-1}$ ,  $H_b^{-1}$ , and  $H_{ab}^{-1}$  represent the effects of chromatic and spherical aberrations and their mutual effects on the IPC, respectively. These terms are given by

$$H_a^{-1} = k_{H_a} \cdot \Delta A^{\gamma_{H_a}} \quad (8.125)$$

$$H_b^{-1} = k_{H_b} \cdot B^{\gamma_{H_b}} \quad (8.126)$$

and

$$H_{ab}^{-1} = k_{H_{ab}} \cdot B^{\gamma_{H_{ab}}} \cdot \Delta A^{\gamma_{H_{ab}}} \quad (8.127)$$

As  $H_p = H_{oo}$  if  $B = \Delta A = 0$ , we see that  $H_{oo}$  in Equation 8.124 represents the IPC of an ideal system, that is,  $H_{oo} = H^{\text{ideal}}$ . The resolution  $R_{es}$  of an optical system can then be determined from Equations 8.95 and 8.37.

The values of the fitting parameters of  $H_p$  for six values of the S/N are listed in Table 8.2.

**TABLE 8.2**  
**List of Fitting Parameters for the IPC for a Point Source**

	Signal-to-Noise Ratio (P/N)					
	2	5	10	20	50	100
$H_{o_0}$	4.99	9.89	15.46	22.81	35.23	46.53
$K_{H_b}$	0.235	0.115	$6.49 \times 10^{-2}$	$4.09 \times 10^{-2}$	$2.40 \times 10^{-2}$	$1.64 \times 10^{-2}$
$\gamma_{H_b}$	1.59	1.53	1.5	1.44	1.35	1.29
$n_{H_b}$	2.5	2.5	2.5	2.5	2.5	2.5
$K_{H_a}$	0.159	$6.78 \times 10^{-2}$	$3.64 \times 10^{-2}$	$1.99 \times 10^{-2}$	$9.21 \times 10^{-3}$	$5.25 \times 10^{-3}$
$\gamma_{H_a}$	1.92	1.92	1.92	1.92	1.92	1.92
$n_{H_a}$	0.7	0.7	0.7	0.7	0.7	0.7
$K_{H_{ab}}$	$5.21 \times 10^{-2}$	$1.82 \times 10^{-2}$	$1.23 \times 10^{-2}$	$6.58 \times 10^{-3}$	$3.69 \times 10^{-3}$	$3.01 \times 10^{-3}$
$\gamma_{H_{b1}}$	1.3	1.3	1.2	1.2	1.1	1.0
$\gamma_{H_{a1}}$	1.1	1.2	1.2	1.2	1.2	1.1

#### 8.5.4.4 Comparison between Fitting Method and Numerical Computation

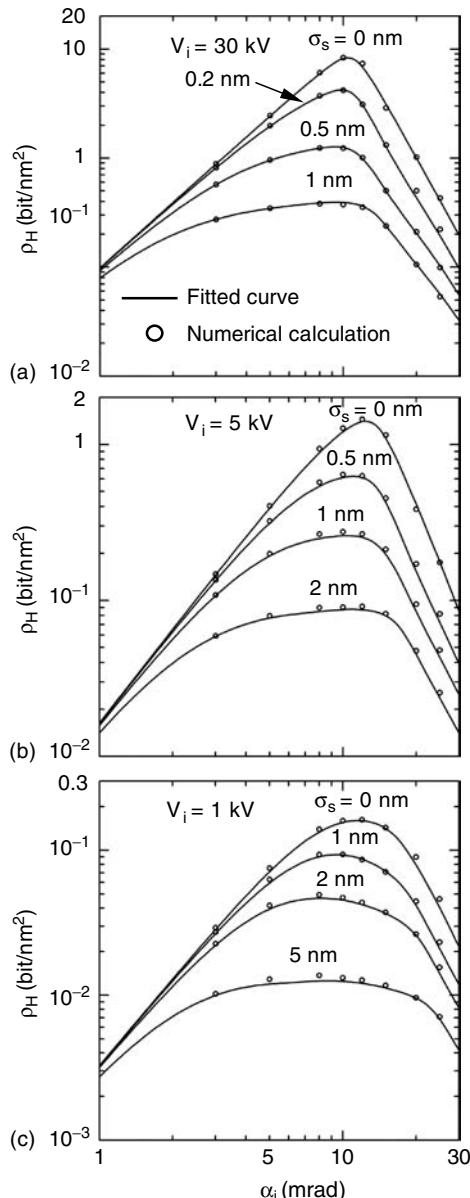
It is important to verify the accuracy of the approximation method. Let us compare the approximate values of  $\rho_H$  for an actual optical system (an Hitachi S-900<sup>37</sup> in-lens-type field emission SEM) obtained using the fitting functions, with direct calculations done by integrating Equation 8.33. Aberration coefficients  $C_{si}$  and  $C_{ci}$  and the energy spread  $\Delta V$  used in the calculations are as follows:  $C_{si} = 1.9$  mm,  $C_{ci} = 2.6$  mm, and  $\Delta V = 3$  eV.

Calculations were made under the following conditions of the optical system: (1) spherical aberration was dominant (at an accelerating voltage of 30 kV), (2) both spherical and chromatic aberrations contributed equally (at 5 kV), and (3) chromatic aberration was dominant (at 1 KV). The results obtained at various values for the beam half angle  $\alpha_i$  and the source radius  $\sigma_s$  (in which the effect of the acceptance factor is subsumed) are shown in Figures 8.19a through 8.19c. In the figures, the circles represent exact numerical calculations and the solid lines are the fitted results. These figures show that the fitted values are fairly precise over a wide range of conditions.

#### 8.5.5 COMPARISON BETWEEN CALCULATED RESOLUTION AND EXPERIMENTAL RESULTS

It is interesting to compare the calculated resolution and the experimental results for various beam half angle  $\alpha_i$ . Figure 8.20 shows several micrographs taken using the S-900 SEM at an accelerating voltage of 1 kV. The magnifications and recording time for all the micrographs were 150,000 times and 80 s, respectively. Aperture radii of 10, 15, and 25  $\mu\text{m}$  were used in this experiment, resulting in beam half angles  $\alpha_i$  of 6.7, 10, and 16.7 mrad, respectively. Beam currents at the specimen corresponding to the three aperture angles were 2.2, 5, and 14 pA, respectively. The small particles shown in the micrographs are gold which were sputtered onto a carbon plate. These micrographs show that  $\alpha_i = 10$  mrad gives a higher-quality image than those taken at the other conditions. However, the resolution in the image is not changed very much when  $\alpha_i$  varies from 6.7 to 16.7 mrad. If we determine the resolution from the minimum gap distance between particles found in the micrograph for  $\alpha_i = 10$  mrad, a value of 3.0 nm would be reasonable (see arrow in the figure).

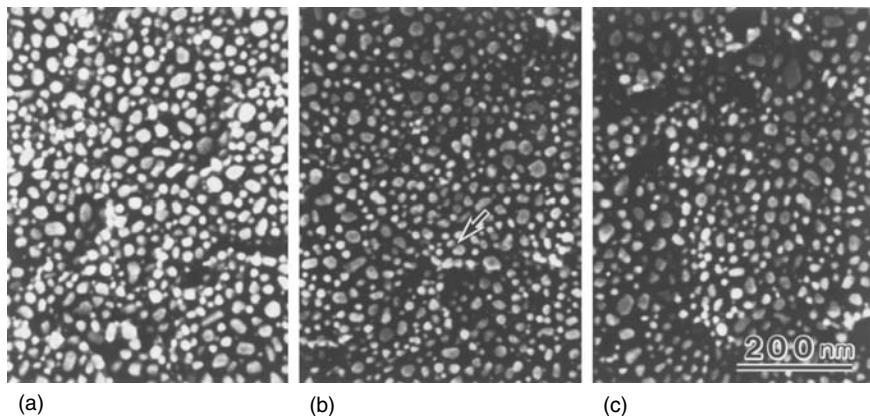
To calculate the resolution in terms of the density-of-IPC, it is necessary to determine the S/N of the image. The S/N of the image is determined by beam current, secondary electron yields (depending on specimen), detector efficiency, etc. To simplify the calculation, we use a constant value for the S/N that was roughly estimated from the micrographs. As the noise in the micrographs shown



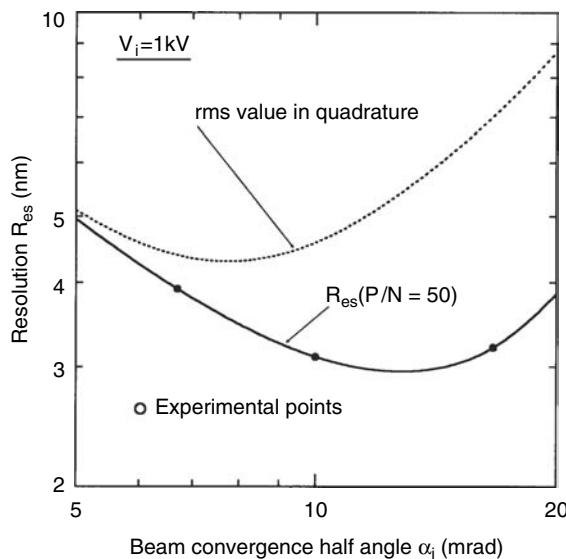
**FIGURE 8.19** Behavior of the density-of-information passing capacity as a function of beam convergence half angle  $\alpha_i$  in the case where (a) spherical aberration is dominant, (b) both spherical and chromatic aberrations influence equally, and (c) chromatic aberration is dominant.

in Figure 8.20 is fairly small, a reasonable and conservative value of such an S/N would be 50, for example. The calculations can then be made at P/N = 50.

The results obtained at various beam-convergence angles  $\alpha_i$  are shown in Figure 8.21. The dotted line in the figure represents the behavior of the beam diameter obtained using the method of addition of aberrations in quadrature. The results of the IPC method show that a resolution of  $\sim 3 \text{ nm}$  is obtained at  $\alpha_i = 13 \text{ mrad}$ . If we compare the calculated resolution under the experimental conditions with the micrographs, we see that the calculated results well represent the actual behavior of the resolution as a function of  $\alpha_i$ . On the other hand, the beam diameter obtained using the method of addition of



**FIGURE 8.20** Scanning electron microscopy (SEM) micrographs taken at different beam-convergence angle at beam voltage 1 kV and magnification 150,000 $\times$ , obtained using a Hitachi S-900 field emission SEM.<sup>37</sup> An arrow indicated the smallest gap showing the resolution (gap resolution) of  $\sim$ 3.0 nm: (a)  $\alpha = 6.7$  mrad; (b)  $\alpha = 10$  mrad; (c)  $\alpha = 16.7$  mrad.



**FIGURE 8.21** Calculated resolution  $R_{es}$  (solid line) as a function of beam convergence half angle  $\alpha_i$  with signal-to-noise ratio  $P/N = 50$ . The dotted line is the root mean square value in quadrature. The circles show the experimental points for the beam-convergence angle.

aberrations in quadrature (dotted line) significantly overestimates the resolution compared with actual images. This discrepancy is important especially when the beam-convergence angle is large.

## 8.6 OPTIMUM CONDITION AND ATTAINABLE RESOLUTION

The optimum condition of an image-forming system can be determined at which the highest quality image is obtained. Such a condition corresponds to the working point or distance where the density  $\rho_H$  of the IPC is maximum. As the resolution of the optical system is proportional to  $\rho_H^{-1/2}$ , such an optimum point also corresponds to the condition where the highest (attainable) resolution can be obtained.

As aberration decreases with decreasing beam-convergence angle ( $\alpha_i$ ), the resolution of an optical system is finally limited either by diffraction ( $\lambda/\alpha_i$ ) or by source size ( $\sigma_s$ ), depending on the application. In an SEM employing a field emission cathode, for example, as the source brightness is too high (~1000 times higher than that of the tungsten hairpin cathode), the source size can be demagnified sufficiently by the optical system so that the attainable resolution is limited by diffraction. If the source brightness is not high enough to realize sufficient demagnification while maintaining the necessary beam current, on the other hand, the source size will limit the attainable resolution of the optical system. In such a case, the optimum working point should be determined taking into account the beam current.

### 8.6.1 OPTIMUM CONDITION FOR A DIFFRACTION-LIMITED SYSTEM

In a diffraction-limited system with given coefficients of aberrations and energy spread, the optimum condition corresponds to the beam convergence half angle  $\alpha_i$  at which the highest resolution is obtained. Thus, a simple way to determine the optimum condition is to calculate the resolution  $R_{es}$  as a function of  $\alpha_i$  at a given S/N. The optimum value  $\alpha_i^{\text{opt}}$  is that at which  $R_{es}$  is minimum, and the resolution at this value corresponds to the best attainable resolution.

A more general approach has been proposed<sup>22</sup> for determining the optimum condition by using the behavior of the IPC. Let us consider a diffraction-limited system with a point source. The density  $\rho_H$  of the IPC at a given S/N is now written as

$$\rho_H = \frac{H_p(B, \Delta A)}{\pi(\lambda/\alpha_i)^2} \quad (8.128)$$

Now we define new parameters  $T$  and  $\Pi$ , which represent the working condition, by

$$T = \sqrt{B^2 + \Delta A^2} = \frac{\alpha_i^4}{4\lambda^2} \sqrt{(\lambda C_{si}) + (\Delta f_c^4)} \quad (8.129)$$

$$\Pi = \frac{\Delta A^2}{B + \Delta A^2} = \frac{\Delta f_c^2}{\lambda C_{si} + \Delta f_c^2} \quad (8.130)$$

The working point ( $T, \Pi$ ) moves along the line of constant  $\Pi$  when  $\alpha_i$  changes with the aberration coefficients and the wavelength is held constant.

Equation 8.128 can be written in terms of  $T$  and  $\Pi$  as

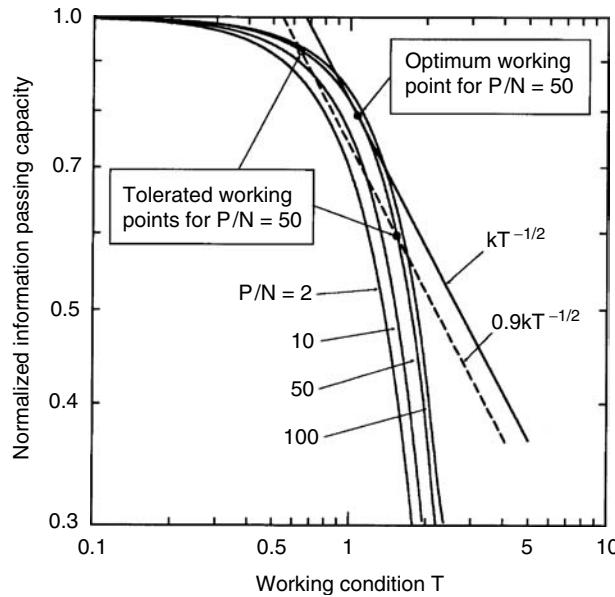
$$\rho_H = C_1 H_p(T, \Pi) T^{1/2} \quad (8.131)$$

where  $C_1$  is defined by

$$C_1 = \frac{2}{\pi \{(\lambda C_{si})^2 + \Delta f_c^4\}^{1/4}} \quad (8.132)$$

Therefore, the optimum condition  $T^{\text{opt}}$  which determines  $\alpha_i^{\text{opt}}$  is found as the point at which the curve  $H_p(T)$  with  $\Pi$  held constant, is also tangent to the line  $kT^{-1/2}$ , where  $k$  is a constant which is determined so that  $kT^{-1/2}$  is just tangent to  $H_p$ , as shown in Figure 8.22. The optimum value  $\alpha_i^{\text{opt}}$  can then be determined from

$$\alpha_i^{\text{opt}} = \frac{(2\lambda)^{1/2} (T^{\text{opt}})^{1/4}}{[(\lambda C_{si})^2 + \Delta f_c^4]^{1/8}} \quad (8.133)$$



**FIGURE 8.22** Behavior of the normalized information passing capacity  $H_p$  as a function of the parameter  $T$  at selected values of signal-to-noise ratio  $P/N$ .

If we define the tolerance on the working condition  $T$  to be the points at which there is a 10% drop in the density-of-IPC, the tolerance ( $T^-$  and  $T^+$ ) of the working condition is given by the intersection of  $H_p$  and the line  $0.9kT^{-1/2}$ , as shown in Figure 8.22.

The parameter  $\Pi$  represents the ratio between the spherical and chromatic aberrations, so  $\Pi = 0$  and  $\Pi = 1$  correspond to the conditions for pure spherical and pure chromatic aberrations, respectively. The range  $0 < \Pi < 1$  corresponds to conditions which include both spherical and chromatic aberrations. Figures 8.23a through 8.23f show the characteristics of  $T^{opt}$  and the tolerance range  $T^-$  and  $T^+$  obtained at S/Ns of 2, 5, 10, 20, 5, and 100. We see from these figures that a large S/N permits a wide tolerance range in the working condition, and the optimum condition also shifts toward larger convergence angles of the beam. The phenomenon is important especially in the case where a chromatic aberration is dominant.

To represent the attainable resolution of an optical system for a point source, it is more convenient to define the dimensionless resolution  $\bar{R}_{es}$  as

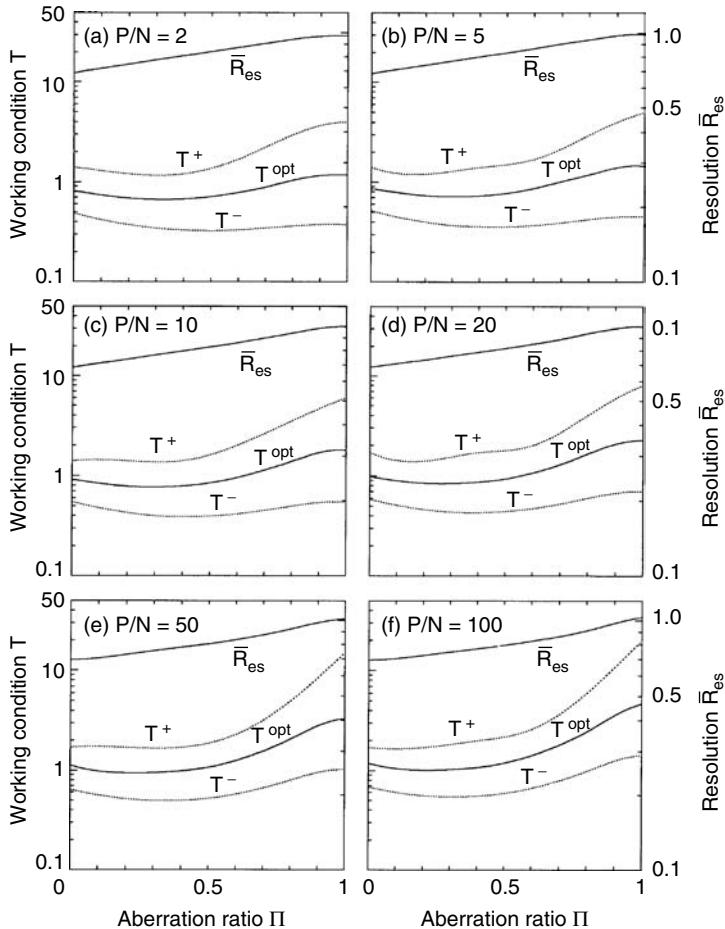
$$\bar{R}_{es} = \frac{R_{es}}{D} \quad (8.134)$$

From Equations 8.37 and 8.128,  $\bar{R}_{es}$  can be written as

$$\bar{R}_{es} = 0.61 \sqrt{\frac{H_{ideal}}{H_p(B, \Delta A)}} \quad (8.135)$$

As  $B$  and  $\Delta A$  at the optimum point are determined from  $T^{opt}$  and  $\Pi$ ,  $\bar{R}_{es}$  can also be represented as a function of  $P$ , as shown in Figures 8.23a through 8.23f. We can find the attainable resolution  $R_{es}$  from the figure as

$$R_{es} = \frac{\lambda}{\alpha_i^{opt}} \bar{R}_{es} \quad (8.136)$$



**FIGURE 8.23** Characteristics of the optimum working condition  $T^{opt}$  (solid line), the tolerance  $T^+$ , and  $T^-$  (dotted lines), and an attainable dimensionless resolution  $\bar{R}_{es}$  as a function of the aberration ratio  $P$ , at different signal-to-noise ratios.

### 8.6.2 OPTIMUM CONDITION FOR A SOURCE-LIMITED SYSTEM

The beam current  $I_p$  measured on a specimen is determined by the source brightness  $B_s$  ( $A/cm^2/sr$ ), that is,

$$I_p = (\pi \beta_i \sigma_s \alpha_i)^2 B_s \quad (8.137)$$

where  $\beta_i = (V_i/V_o)^{1/2}$  and  $\sigma_s$  is the source radius referred to the specimen, that is,  $\sigma_s = M r_{ss}$ . If  $I_p$  and  $B_s$  are given,  $\sigma_s$  is then determined from Equation 8.137 to be

$$\sigma_s = \frac{k_s}{\alpha_i} \quad \text{with } k_s = \frac{1}{\pi \beta_i} \sqrt{\frac{I_p}{B_s}} \quad (8.138)$$

Thus, the behavior of  $\sigma_s$  (with constant  $I_p$ ), depends on  $\alpha_i$  in the same way as diffraction. If  $k_s$  is of same order as the wavelength, the optimum value of  $\alpha_i$  ( $= \alpha_i^{opt}$ ) and the attainable resolution can be

found from the properties of the resolution that depend on  $\alpha_i$ , taking into account diffraction and the source size  $\sigma_s$ . The optimum magnification  $M^{\text{opt}}$  of the optical system can then be determined as

$$M^{\text{opt}} = \frac{\sigma_s(\alpha_i^{\text{opt}})}{r_{\text{ss}}} = \frac{k}{\alpha_i^{\text{opt}} r_{\text{ss}}} \quad (8.139)$$

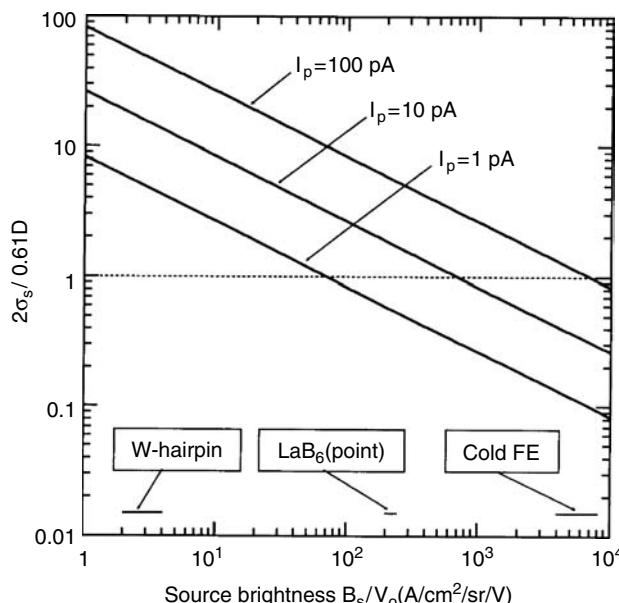
As the contribution of diameter  $2\sigma_s$  in the Gaussian distribution beam on the resolution is similar to that of the radius of Airy disk ( $0.61D$ ) in a diffraction-limited system, as was presented in Section 8.2, the ratio of these quantities represents whether or not the source size limits the attainable resolution. From Equation 8.138, we find

$$\frac{2\sigma_s}{0.61D} \approx \frac{1}{\lambda\sqrt{V_i}} \sqrt{\frac{I_p}{B_s V_o}} \quad (8.140)$$

Because the wavelength  $\lambda$  is proportional to  $V_i^{-1/2}$ ,  $\lambda\sqrt{V_i}$  becomes constant, so Equation 8.140 is independent of the final beam energy ( $V_i$ ). In Equation 8.140,  $B_s/V_o$  is also determined by the emitter type such as a field emitter,  $\text{LaB}_6$ , and so on. Thus, whether or not the source size is dominant in the optical system depends only on the emitter type and the final beam current. A condition where the optical system is diffraction limited is given by

$$\frac{2\sigma_s}{0.61D} \ll 1 \quad (8.141)$$

Figure 8.24 shows a relation between  $2\sigma_s/0.61D$  and  $B_s/V_o$  ( $\text{A}/\text{cm}^2/\text{sr}/\text{V}$ ) at selected beam currents of 1, 10, and 100 pA. Typical values of  $B_s/V_o$  for different types of emitter<sup>38</sup> are also shown in the figure. We see from the figure that the source size for a cold field emitter (cold FE) does not contribute on the resolution over a wide range of beam current, so the optical system can be operated under diffraction-limited conditions.



**FIGURE 8.24** Relation between source brightness and contribution of the source size relative to the diffraction limit.

In the case of an ion beam system, on the other hand, as no wave effects appear in the optical system, the optimum condition is determined from the source brightness and the aberrations. If the effects of coulomb interactions can be ignored, such as for an application using a small beam current, the optimum condition and attainable resolution can be found by applying the IPC method presented in the previous section.

The density  $\rho_H$  of the IPC at a given S/N in geometric optics can be determined by using the synthesized parameters as follows (see Section 8.4.4):

$$\rho_H = \frac{1}{\pi D'^2} H_p(B', \Delta A') \quad (8.142)$$

where  $D'$ ,  $B'$ , and  $\Delta A'$  are given by

$$D' = \frac{\sigma_s}{\sigma_g} = \frac{(k_s / \bar{\sigma}_g)}{\alpha_i} \quad B' = f_b^{-1}(B_g^{\text{eff}}) \quad \Delta A' = f_a^{-1}(\Delta A_g^{\text{eff}}) \quad (8.143)$$

with

$$B_g^{\text{eff}} = \frac{1}{4} \frac{C_{si} \alpha_i^3}{D'} \quad \text{and} \quad \Delta A_g^{\text{eff}} = \frac{1}{2} \frac{\Delta f_c \alpha_i}{D'} \quad (8.144)$$

As was done for a diffraction-limited system, we can define the parameters  $T_g$  and  $\Pi_g$ , which represent the working condition, by

$$T_g = \sqrt{(B_g^{\text{eff}})^2 + (\Delta A_g^{\text{eff}})^4} = \frac{\alpha_i^4}{4(k_s / \bar{\sigma}_g)} \sqrt{[(k_s / \bar{\sigma}_g) C_{si}]^2 + \Delta f_c^4} \quad (8.145)$$

$$\Pi_g = \frac{(\Delta A_g^{\text{eff}})}{B_g^{\text{eff}} (\Delta A_g^{\text{eff}})^2} = \frac{\Delta f_c^2}{(k_s / \bar{\sigma}_g) C_{si} + \Delta f_c^2} \quad (8.146)$$

The optimum condition  $T_g^{\text{opt}}$  and the tolerance range ( $T_g^-$ ,  $T_g^+$ ) at which there is a 10% drop in the density-of-IPC can be found as a function of  $P_g$  by using the same technique used for a diffraction-limited system. The optimum beam convergence half angle  $\alpha_i^{\text{opt}}$  can thus be determined by

$$\alpha_i^{\text{opt}} = \frac{(2k_s / \bar{\sigma}_g)^{1/2} (T_g^{\text{opt}})^{1/4}}{\left[ \{(2k_s / \bar{\sigma}_g) C_{si}\}^2 + \Delta f_c^4 \right]^{1/8}} \quad (8.147)$$

If we define the dimensionless resolution  $\bar{R}_{es}$  in geometric optics as

$$\bar{R}_{es} = \frac{R_{es}}{D'} = \frac{R_{es}}{(k_s / \bar{\sigma}_g) / \alpha_i} \quad (8.148)$$

then, from Equations 8.37, 8.142, and 8.143,  $\bar{R}_{es}$  is given by

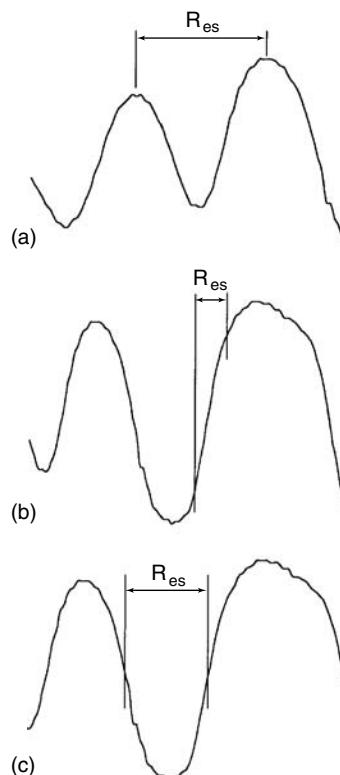
$$\bar{R}_{es} = 0.61 \sqrt{\frac{H^{\text{ideal}}}{H_p(B', \Delta A')}} \quad (8.149)$$

$\bar{R}_{es}$  at the optimum condition can also be found as function of  $\Pi_g$  in the same way as shown in Figures 8.23a through 8.23f.

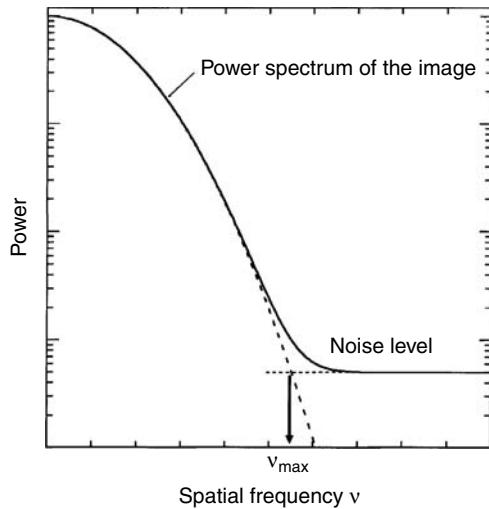
## 8.7 MEASUREMENT OF THE RESOLUTION

Because a pointlike object intensity does not exist in an actual specimen, the resolution of an optical system is usually measured from the image by the following methods: (1) point-to-point resolution, (2) edge resolution, and (3) gap resolution, as shown in Figures 8.25a through 8.25c. It is also possible to measure the resolution from the maximum spatial frequency  $v_{\max}$  by taking a Fourier transform of the image, as shown in Figure 8.26, where the resolution is determined by the reciprocal of  $v_{\max}$ . We see from the definitions that the results from measuring the edge resolution and gap resolution may give smaller value (better resolution) than that calculated based on Rayleigh's criterion. The point-to-point resolution has a close connection with Rayleigh's criterion. Such a result, however, would be worse than Rayleigh's criterion because of the finite size of the object. In any method, as the resolution is measured indirectly from the image of a specimen, the results depend not only on the optical performance but also on the specimen structure. Therefore, it is important when measuring the resolution of an optical system to use an appropriate specimen which is constructed with sufficient fine structure and sufficient contrast.

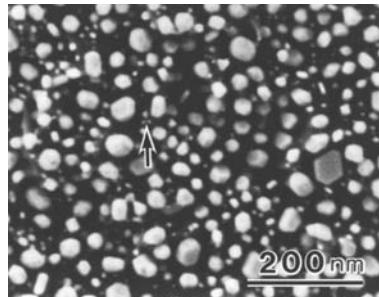
In an SEM, for example, the gap resolution is usually used to measure the resolution. In this method, one finds the shortest distance (gap) between fine particles in the image, as shown in Figure 8.27. A specimen used for the measurement of the resolution typically consists of fine particles of gold or platinum prepared by sputtering or evaporation on a carbon plate (or a magnetic tape). As the gap resolution (as well as the edge resolution) depends on the contrast in the signal, an adequate gray level in the image should be attained.



**FIGURE 8.25** Methods for measuring the resolution of an image, showing (a) the point-to-point resolution, (b) the edge resolution, and (c) the gap resolution.



**FIGURE 8.26** A model showing the maximum spatial frequency in an image.

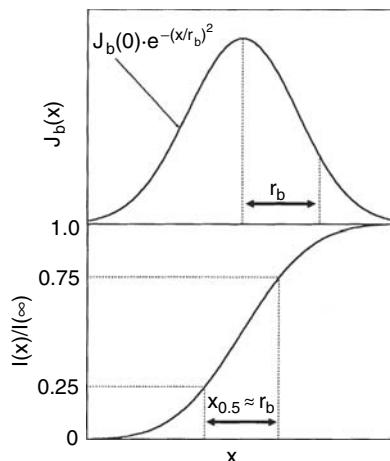


**FIGURE 8.27** Scanning electron microscopy micrograph of gold particles on a carbon plate showing the resolution of an optical system. An arrow shows the smallest gap found in the micrograph where the gap resolution of 2.5–3.0 nm can be measured from the gap distance.

When an electron or an ion beam scans across a sharp knife-edge, the transmitted intensity  $I(x)$  becomes

$$I(x) = \int_{-\infty}^x dx \int_{-\infty}^{+\infty} J_b(\sqrt{x^2 + y^2}) dy \quad (8.150)$$

If the beam intensity  $J_b(r)$  is given by a Gaussian distribution, the rise-distance  $x_{0.5}$  between points corresponding to 25 and 75% of  $I(x)$ , as shown in Figure 8.28, corresponds to the beam radius at which the intensity has fallen to  $1/e$  of the maximum. This definition corresponds to edge resolution. This method is useful for measuring beam size if the beam is relatively large, of the order of submicron or more. One must, however, deal carefully with the experimental results because the rise-distance  $x_{0.5}$  mainly depends on the tail of the distribution  $J_b(r)$ , so it tells little about the central peak when  $J_b(r)$  has long tail.



**FIGURE 8.28** An estimation of the beam size from the current profiles as the beam is swept over a knife-edge. The rise-distance  $x_{0.5}$  coincides with the beam radius at which the intensity has fallen to  $1/e$  of the maximum when the intensity is a Gaussian distribution.

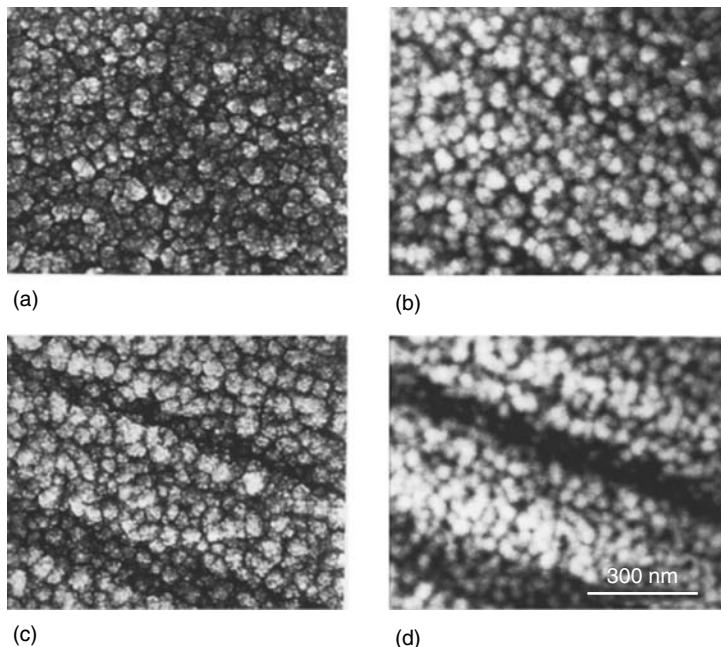
## 8.8 OTHER FACTORS LIMITING THE RESOLUTION

### 8.8.1 EFFECT OF THE INFORMATION SIZE IN THE SPECIMEN

An image signal, in SEM applications, is produced by secondary electrons and backscattered electrons generated from a specimen. As primary electrons penetrate inside the specimen, the signal is generated from finite volume in the specimen. The size of the volume which generates the imaging signal can be called the information size, and it depends on the density of the material in the specimen, the specimen thickness, and the primary electron energy. The information size for secondary electrons is usually smaller than that for backscattered electrons. Figure 8.29 shows the difference in the information size between secondary electrons and backscattered electrons taken using S-4500 field emission SEM.<sup>39</sup> Figures 8.29a and 8.29b are SEM micrographs of a hard disk surface produced by secondary electrons and backscattered electrons, respectively, taken at 2 kV beam voltage. We easily find from the micrographs that the resolution of the backscattered electron image (Figure 8.29b) is worse than that of the secondary electron image (Figure 8.29a). Figures 8.29c and 8.29d show the micrographs produced by secondary electron signal and backscattered electron signal, respectively, taken at 5 kV. Although the resolution of an optical system should improve with the increased beam voltage, the micrograph for backscattered electron image at 5 kV has poorer resolution compared with that taken at 2 kV, as shown in Figures 8.29b and 8.29d, because the information size increased with the increased beam voltage.

### 8.8.2 EFFECT OF VIBRATION

If the optical column vibrates because of any reason such as the vibration of a floor, acoustical input, and so on, the position of the primary beam moves relative to the specimen. The degree of influence of the vibration upon the image usually depends on the frequency. If the frequency of vibration coincides with a resonance point of the antivibration devices for the instrument, which is 2–3 Hz, typically, the image becomes most sensitive to the vibration. When the image is produced by summing or averaging each frame, which is obtained by high-speed scanning of the primary beam specimen, the vibration causes the image to be blurred.



**FIGURE 8.29** Scanning electron microscopy micrographs of hard disk surface showing the difference in information size between secondary electrons and backscattered electrons: (a) Secondary electron image taken at 2 kV, (b) backscattered electron image taken at 2 kV, (c) secondary electron image taken at 5 kV, and (d) backscattered electron image taken at 5 kV.

### 8.8.3 EFFECT OF A MAGNETIC STRAY FIELD

If a stray magnetic field near the instrument fluctuates, the primary beam also moves due to the field fluctuation. When the fluctuation of the magnetic field is synchronized with the AC power frequency, the magnetic field distorts the image in the case where the primary beam is scanned over a specimen synchronized with the power frequency. On the other hand, if the fluctuation of the magnetic field has a component which is not synchronized with the power frequency, the effect of such a fluctuation appears in the image in a way similar to that of vibration. The effect of magnetic stray field also depends on the optical conditions, such as primary beam energy, focal length of the objective lens, or the positions of a crossover for the primary beam.

### ACKNOWLEDGMENT

The author is grateful to J. Orloff of the University of Maryland for valuable discussions and suggestions.

### REFERENCES

1. A. J. den Dekker and A. van den Bos, Resolution: A survey, *J. Opt. Soc. Am.* **14A**, 547–557, 1997.
2. S. Van Aert, D. Van Dyck, and A. J. den Dekker, Resolution of coherent and incoherent imaging systems reconsidered—Classical criteria and a statistical alternative, *Opt. Express* **9**, 3830–3839, 2006.
3. S. Van Aert, A. J. den Dekker, D. Van Dyck, and A. van den Bos, *Science of Microscopy* (P.W. Hawkes and J.C.H. Spence, Eds), Vol. 2, pp. 1228–1265 (Springer, New York, 2007).
4. C. E. Shannon, *Bell Syst. Tech. J.*, 27, 379–423, 1948.
5. D. Gabor, *Tech. Rep. Electron. Mass. Inst. Technol.*, No. 238, 1952.

6. P. Elias, D. S. Grey, and D. Z. Robinson, *J. Opt. Soc. Am.*, 42, 127, 1952.
7. A. Blanc-Lapierre, *McGill Symp. Microwave Optics*, 2(46), 1953.
8. P. B. Fellgett, *J. Opt. Soc. Am.*, 43, 271, 1953.
9. P. B. Fellgett and E. H. Linfoot, *Philos. Trans.*, A247, 369, 1955.
10. E. H. Linfoot, *J. Opt. Soc. Am.*, 45, 808, 1955.
11. G. Black and E. H. Linfoot, *Proc. R. Soc. (London)*, A239, 522, 1957.
12. R. Simon, *J. Appl. Phys.*, 41, 4632, 1970.
13. L. Rayleigh, *Philos. Mag.*, 8(5), 261, 1879.
14. G. B. Airy, *Trans. Camb. Philos. Soc.*, 5, 283, 1835.
15. M. Von Ardenne, *Z. Tech. Phys.*, 19, 407, 1988.
16. M. Von Ardenne, *Z. Tech. Phys.*, 20, 238, 1940.
17. O. C. Wells, *IEEE Trans. Electron Dev.*, Ed-12, 224, 1965.
18. K. J. Harte, *J. Vac. Sci. Technol.*, 10, 1098, 1973.
19. J. L. Mauer, H. C. Pheiffer, and W. Stickel, *IBM J. Res. Dev.*, 21, 514, 1977.
20. J. Lm Mauer, *J. Vac. Technol.*, 15, 853, 1978.
21. A. V. Crewe, *Ultramicroscopy*, 23, 159, 1987.
22. M. Sato and J. Orloff, *Ultramicroscopy*, 41, 181, 1992.
23. M. Born and E. Wolf, *Principles of Optics*, Macmillan, New York, 1959.
24. Z. Shao and A. V. Crewe, *Ultramicroscopy*, 23, 169, 1987.
25. D. W. Tuggle, J.-Z. Li, and L. W. Swanson, *J. Microsc.*, 140, 293, 1985.
26. M. Troyon and P. Zinzindohouse, *Spectrosc. Electron.*, 12, 431, 1987.
27. M. Troyon, *Spectrosc. Electron.*, 13, 431, 1988.
28. Glaser, *Handb. Phys.*, 33, 123, 1956.
29. Y. Kawanami, T. Ishitani, K. Umemura, and S. Shukuri, *J. Vac. Sci. Technol.*, B5, 1364, 1987.
30. D. W. Tuggle, L. W. Swanson, and M. A. Gesley, *J. Vac. Sci. Technol.*, B4, 131, 1986.
31. G. F. Rempfer and M. S. Mauck, *J. Appl. Phys.*, 63, 2187, 1988.
32. M. Sato and J. Orloff, *J. Vac. Sci. Technol.*, B9, 26.2, 1991.
33. L. W. Swanson, G. A. Schwind, A. E. Bell, and J. E. Brady, *J. Vac. Sci. Technol.*, 16(6), 1864, 1979.
34. T. Ishitani, Y. Kawanami, T. Ohnishi, and K. Umemura, *Appl. Phys.*, A44, 233, 1987.
35. M. Sato and J. Orloff, *Optik*, 89, 44, 1991.
36. M. Sato and J. Orloff, *Optik*, 89, 82, 1991.
37. T. Nagatani, S. Satio, M. Sato, and M. Yamada, *Scanning Microsc.*, 1, 901, 1987.
38. O. C. Wells, Ed., *Scanning Electron Microscopy*, McGraw-Hill, New York, 1974, p. 73.
39. M. Sato, H. Todokoro, and K. Kageyama, *SPIE Proc.*, 2014, 17, 1993.



---

# 9 The Scanning Electron Microscope\*

*András E. Vladár and Michael T. Postek*

## CONTENTS

9.1	Introduction.....	438
9.2	Scanning Electron Microscope Architecture.....	438
9.2.1	Electron Source Types.....	440
9.2.1.1	Point Cathode Electron Source Types.....	440
9.2.2	Lens Design.....	443
9.2.2.1	Immersion Lens Technology.....	443
9.2.2.2	Extended Field Lens Technology.....	444
9.2.3	Scanning Electron Microscope Electronics and Digital Image Storage and Image Analysis .....	446
9.2.3.1	TV Rate Scanning.....	448
9.2.3.2	Digital Image Storage .....	449
9.2.3.3	Digital Image Transmission.....	449
9.2.3.4	Real-Time Image Analysis and Processing.....	449
9.2.3.5	Optimization of Operating Conditions .....	449
9.3	The Scanning Electron Microscope Sample.....	449
9.3.1	Nondestructive Inspection.....	450
9.3.1.1	Total Electron Emission.....	451
9.4	Electron Beam—Specimen Interactions.....	452
9.4.1	Electron Range .....	452
9.4.2	Modeling of the Scanning Electron Microscope Signal .....	455
9.5	Scanning Electron Microscope Signals .....	456
9.5.1	Secondary Electrons.....	456
9.5.1.1	Collection of Secondary Electrons .....	457
9.5.2	Backscattered Electrons .....	459
9.5.2.1	Collection of Backscattered Electrons .....	460
9.5.3	X-Rays .....	461
9.5.3.1	Collection of X-Rays .....	461
9.5.4	Absorbed Electrons .....	463
9.5.5	Cathodoluminescence (Light) .....	464
9.5.5.1	Detecting Cathodoluminescence .....	464
9.5.6	Auger Electrons .....	464
9.5.7	Transmitted Electrons .....	465

---

\* This work is a contribution of the National Institute of Standards and Technology. The work was supported (in part) by the National Semiconductor Metrology Program at the National Institute of Standards and Technology and is not subject to copyright. Certain commercial equipment is identified in this report to adequately describe the experimental procedure. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the equipment identified is necessarily the best available for the purpose.

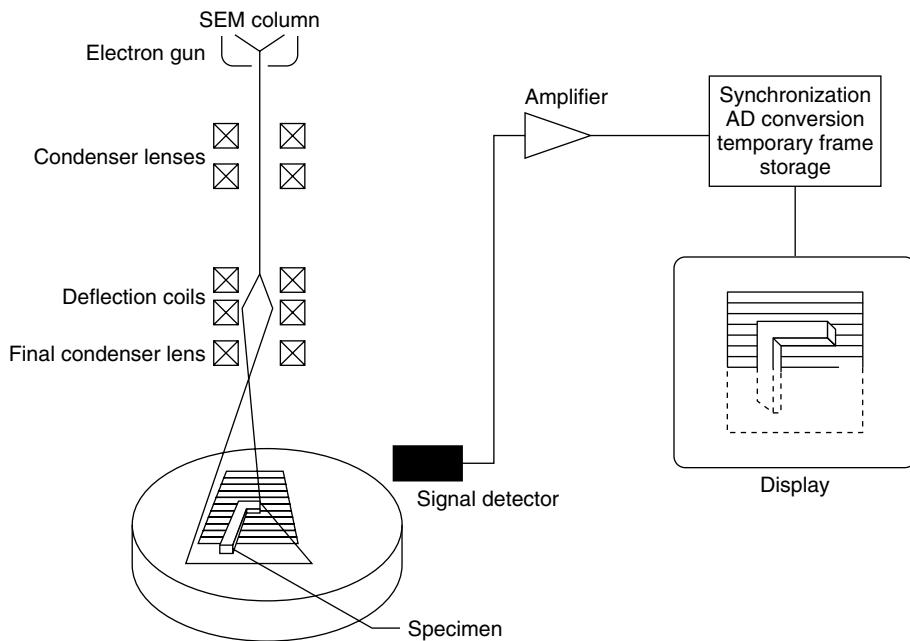
9.6 Applications .....	466
9.6.1 Low Accelerating Voltage Inspection .....	466
9.6.1.1 Linewidth Measurements .....	466
9.6.1.2 Particle Metrology .....	467
9.6.2 Voltage Contrast .....	468
9.6.2.1 Qualitative Voltage Contrast .....	468
9.6.2.2 Quantitative Voltage Contrast .....	470
9.6.3 Electron Beam Induced Current .....	470
9.6.4 Failure Analysis .....	470
9.6.5 Variable Vacuum Scanning Electron Microscope .....	472
9.6.5.1 Variable Vacuum Scanning Electron Microscope Column .....	473
9.6.5.2 Effect of the Gas on the Primary Electron Beam .....	474
9.6.5.3 Secondary Electron Signal Detection in a Gas .....	475
9.6.5.4 Gas Pressure and Path Optimization .....	477
9.6.5.5 Charge Control .....	478
9.6.5.6 Contamination Reduction .....	478
9.7 Scanning Electron Microscope Imaging and Accurate Measurements .....	478
9.7.1 Resolution Measurements for the Scanning Electron Microscope .....	478
9.7.2 Resolution of the Scanning Electron Microscope Image .....	480
9.7.3 Electron Beam Landing Position .....	483
9.7.4 Electron Beam Induced Contamination .....	485
9.8 Helium Ion Microscopy .....	488
9.9 Extra High-Resolution SEMs .....	492
9.10 Conclusions .....	493
References .....	494

## 9.1 INTRODUCTION

The scanning electron microscope (SEM) is extensively used in many fields of research and industrial production throughout the world. The instrument's popularity stems from its capability of quickly and easily providing shape, size, and other information on structures ranging in size from a few centimeters to a few nanometers. The SEM provides higher spatial resolution analysis and inspection than that afforded by current techniques using the optical microscopes. SEMs can deliver information on individual structures or over large areas of the samples. Unlike the optical microscope, the SEM offers a wide variety of analytical modes, each contributing unique information regarding the physical, chemical, and electrical properties of a particular specimen, device, or circuit. Owing to recent developments, especially in variable vacuum instrumentation, the SEM is now used in research, development, and production where previously excessive sample charging and damage limited its use.

## 9.2 SCANNING ELECTRON MICROSCOPE ARCHITECTURE

The SEM is named in this way because in it a finely focused beam of electrons is moved or scanned from point-to-point on the specimen surface in a precise, generally rectangular pattern called a raster pattern (Figure 9.1). The primary beam of electrons originate from an electron source and is accelerated toward the specimen by a voltage usually between 100 V (0.1 kV) and 30,000 V (30 kV). The electron beam travels down the column where it undergoes a multistep electron optical demagnification by one or more condenser lenses. This demagnification reduces the diameter of the electron beam from as large as several micrometers to nanometer dimensions. Depending on the application, magnification range, resolution required, and specimen nature, the operator optimizes the image by choosing the proper accelerating voltage and amount of condenser lens demagnification.



**FIGURE 9.1** Electron beam scanning and image formation.

**TABLE 9.1**  
**Achievable Scanning Electron Microscope Resolution**

Detector Position	Accelerating Voltage (kV)	In-Lens Field Emission (FE) (nm)	Extended FE (nm)	Postlens FE (nm)	Postlens LaB <sub>6</sub> (nm)	Postlens Tungsten (nm)
Upper	20	0.4	0.8			
	1.0	1.5	1.8			
Lower	30			1.5	2.5	3.5
	1.0			4.0	7.5	10.0

Generally, depending on the type of the SEM and the operating conditions, the electron beam diameter is in between 0.3 and 10 nm when it impinges on the sample.

Scanning coils or deflection plates driven by a digital or analog scan generator circuit of the microscope precisely deflect the electron beam in a raster pattern (Figure 9.1). This deflection is synchronized with the digital image storage circuit or with the deflection of a visual monitor, so there is a point-by-point visual representation of the signal being generated by the specimen as it is scanned. The smaller the area scanned in the raster pattern relative to the fixed size of the final image on the monitor, the higher the magnification. The proper calibration of the scale applied to this raster pattern (i.e., magnification) is essential.<sup>1</sup> As shown in Table 9.1, the SEM is capable of extremely high, over 500,000 times, magnification and resolution better than 0.4 nm, depending on the instrument design. In this regard all SEMs surpass all (far-field) optical microscopes, which depending on their illumination source, have their best resolution limited by diffraction to  $\sim 0.2 \mu\text{m}$  and their highest magnification to  $\sim 1000$  times.

One of the major characteristics of the SEM, in contrast to the optical microscope, is its great depth-of-field, which is 100–500 times greater than that of an optical microscope. This characteristic allows the SEM to produce completely in-focus micrographs of relatively rough surfaces even at

**TABLE 9.2**  
**Comparison of Pertinent Electron Source Characteristics**

	Units	Tungsten Filament	LaB <sub>6</sub> Emitter	CeB <sub>6</sub> Emitter	Cold Field Emitter	Schottky Field Emitter
References		76	5.8,50	7	51,76,89,92,102	51,76,85,86,89–91,94
Angular current intensity	mA/str	n/a	n/a	n/a	<0.1	0.1–1.0
Source brightness	A/(cm <sup>2</sup> · str)	10 <sup>6</sup>	10 <sup>7</sup>	10 <sup>7</sup>	10 <sup>9</sup>	10 <sup>8</sup> –10 <sup>9</sup>
Emitting surface area	μm <sup>2</sup>	≥>1	>1	>1	0.02	0.2
Crossover or virtual source diameter	nm	>10 <sup>4</sup>	>10 <sup>3</sup>	>10 <sup>3</sup>	3–5	15–25
Energy spread	eV	1–3	1–1.5	1–1.5	0.2–0.3	0.3–1.0
Source temperature	K	25–2900	1800	1800	300	1800
Work function	eV	4.5	2.6	2.4	4.5	2.8
Operating vacuum	Pa	10 <sup>-4</sup>	10 <sup>-6</sup>	10 <sup>-6</sup>	10 <sup>-9</sup> –10 <sup>-11</sup>	10 <sup>-8</sup> –10 <sup>-9</sup>
Short-term beam current stability	%RMS	<1	<1	<1	4–6	<1
Typical service life	h	40–100	1000	>1500	>2000	>2000

high magnifications. The SEM is also parfocal; it can be focused and the astigmatism corrected at a magnification higher than that needed for work, then the magnification can be reduced without loss of sharpness.

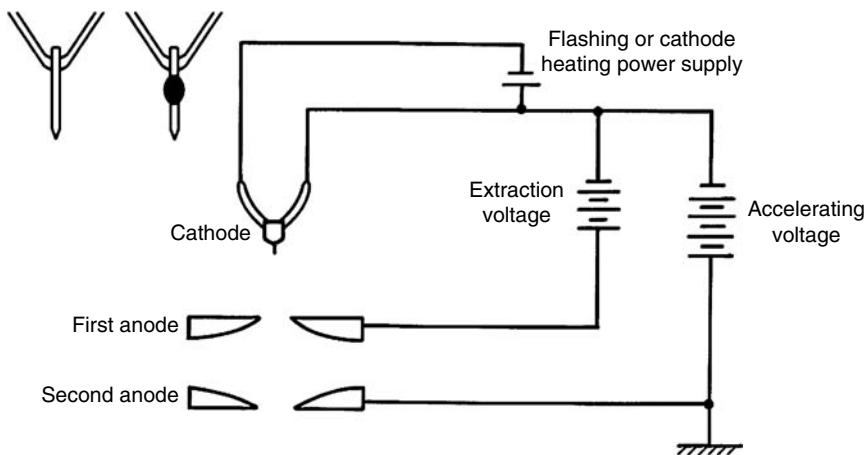
The proper operation of an SEM requires maintaining the electron microscope column under high vacuum as electrons cannot travel for any appreciable distance in air. The requirement for vacuum does tend to limit the type of sample that can be viewed in the SEM (see Sections 9.3 and 9.6.5). The mean free path or the average distance an electron travels before it encounters an air or gas molecule must be not much shorter than the SEM column; therefore, depending on the type of instrument design, ion pumps, diffusion pumps, or turbomolecular pumps are to be utilized to achieve the level of vacuum needed for the particular instrument design.

### 9.2.1 ELECTRON SOURCE TYPES

In the past the predominant electron sources in use were the thermionic emission type cathodes, especially tungsten and lanthanum hexaboride (LaB<sub>6</sub>).<sup>2</sup> Lanthanum hexaboride cathodes became more prevalent for low accelerating voltage applications because of their higher brightness and smaller source diameter in comparison to tungsten filaments (Table 9.2). Cerium hexaboride (CeB<sub>6</sub>) cathode is similar in operation and performance to the lanthanum hexaboride cathode.<sup>3</sup> Point cathode electron sources or field emission instrumentation have been available for some time, and today a wide variety of both laboratory type and process type instruments are commonly used with field emission technology (see Chapter 1). For many industrial applications such as those found on the semiconductor processing lines, only the field emission instruments provide the high resolution necessary for this type of work (Table 9.2). This is especially true at the low accelerating voltages needed for nondestructive inspections.<sup>4</sup> In the near future, nanometer-sized field emission tips with literally atomic proportions with one or only a few atoms at the very tip may also become available as practical electron and ion sources.<sup>5,6</sup>

#### 9.2.1.1 Point Cathode Electron Source Types

There are two basic categories of point cathode electron source types used in the current SEMs: cold field emission (CFE) and thermally assisted field emission (TFE) cathodes. The design of these is



**FIGURE 9.2** Field emission electron gun (schematic representation).

schematically shown in Figure 9.2. Although the concept of field emission can be traced to the early work of Wood<sup>7</sup> and was used in early instrumentation by Zworykin et al.<sup>8</sup> it was not until the late 1960s that Crewe et al.<sup>9,10</sup> developed a successful cold cathode field emission source that was later introduced in commercial instrumentation.

For many applications, such as analytical microscopy and microfabrication, the CFE emitter (Figure 9.2 upper left corner left tip) was not capable of producing the high currents and long-term electron emission stability needed.<sup>11,12</sup> Work to develop a high-current TFE cathode with relaxed vacuum and environment requirements began in the 1960s.<sup>13–15</sup> There have been several designs of TFE cathodes developed. The two major types are the tungsten ⟨100⟩ built-up emitter and the ZrO/W ⟨100⟩ cathode Schottky emitter. Currently, ZrO/W is the more commonly used TFE source in modern laboratory and SEM-based metrology instruments. The ZrO/W cathode provides higher electron emission than W (see Chapter 1).

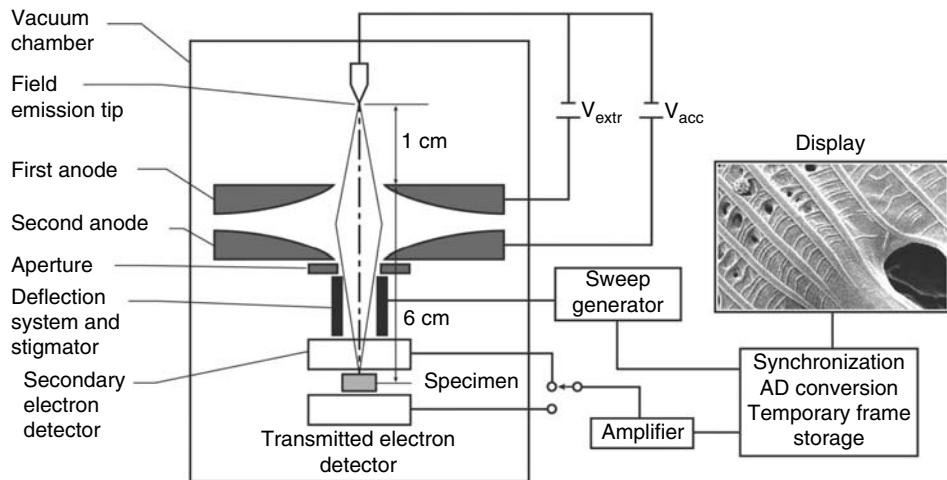
Instruments utilizing either CFE or Schottky emitter currently populate the SEM instrument field. Each type has its advantages and disadvantages. The characteristics of the various electron sources, as they are currently understood, including CFE and Schottky emitter, are summarized in Table 9.2 and are briefly discussed in the following sections.

#### 9.2.1.1.1 Cold Field Emission Cathodes

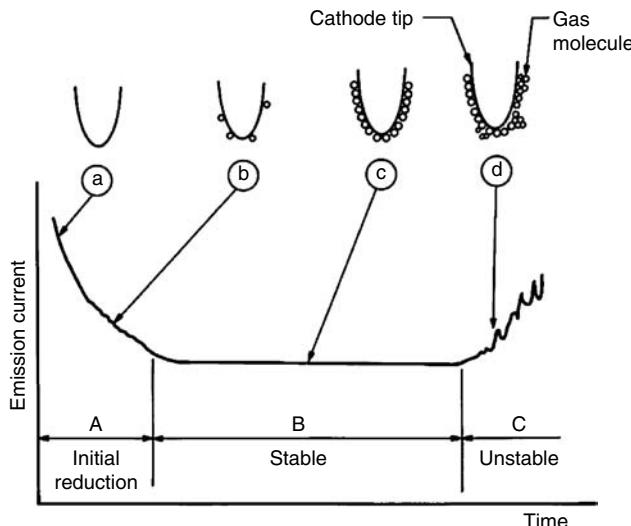
CFE cathodes, developed for use in the SEM by Crewe et al.,<sup>9,16</sup> have an advantage of providing a relatively high-current electron probe having a low energy spread, high brightness, and a small virtual source diameter, especially at low accelerating voltages. The CFE source diameter is sufficiently small that the electron gun alone (as shown in Figure 9.3) without any additional condenser lenses is capable of producing a 10 nm probe<sup>9,17</sup> (see also Chapter 10). From Table 9.2 it can be observed that depending on the type of instrument design, better than 1 nm resolution may be reached with an instrument equipped with a field emission electron source.

The overall advantages of the CFE are offset somewhat by the stringent requirements for ultra-high gun vacuum (Table 9.2) and some fluctuation (flicker) in the emission current. This fluctuation is due to the molecules adhering to the emitter (Figure 9.4).

Larger changes of the emission current can be avoided by regularly heating the emitter to high temperatures in a so-called flashing process, which resets the emission to its A region. The emission current fluctuation in the stable B region for hours, and the small changes are readily compensated for by constant beam monitoring and feedback control, and also through fast digital frame averaging, and in general, is not an issue of concern in modern instruments.



**FIGURE 9.3** Cold field emission electron microscope with no condenser lens.



**FIGURE 9.4** Cold field emission gun stability over time.

#### 9.2.1.1.2 Thermally Assisted Field Emission

The second basic category of point cathode electron sources is the TFE cathode. In this mode of operation, the cathode is heated and thus vacuum requirements are reduced and the emission current is relatively stable.<sup>18</sup> Because of its lower work function, the use of the Schottky point emitter, such as the zirconiated/tungsten ⟨100⟩ point cathode, is preferred. This source can produce a high-current electron beam with a slightly poorer energy spread. This differs from CFE by an amount as small as ~8 to 10% depending on how evaluation criteria are established.<sup>18,19</sup> As this source is currently being used for a number of different applications, the operational characteristics and parameters of the source are fairly varied. Thus, it is quite difficult to tabulate a direct comparison of source characteristics. For many SEM applications the Schottky emitter source is generally operated with conditions resulting in the lowest energy spread (0.3 eV) possible for that type of source. Under these conditions using test samples, resolution comparable to similarly equipped CFE

instruments has been obtained.<sup>20</sup> Unlike the CFE, the larger source diameter characteristic of this type of electron source requires the use of an extra condenser lens in the electron microscope column to increase the source demagnification. The need for increased demagnification also provides a positive secondary effect, as it also results in increased demagnification of external noise such as vibration and fields affecting the source.

### 9.2.2 LENS DESIGN

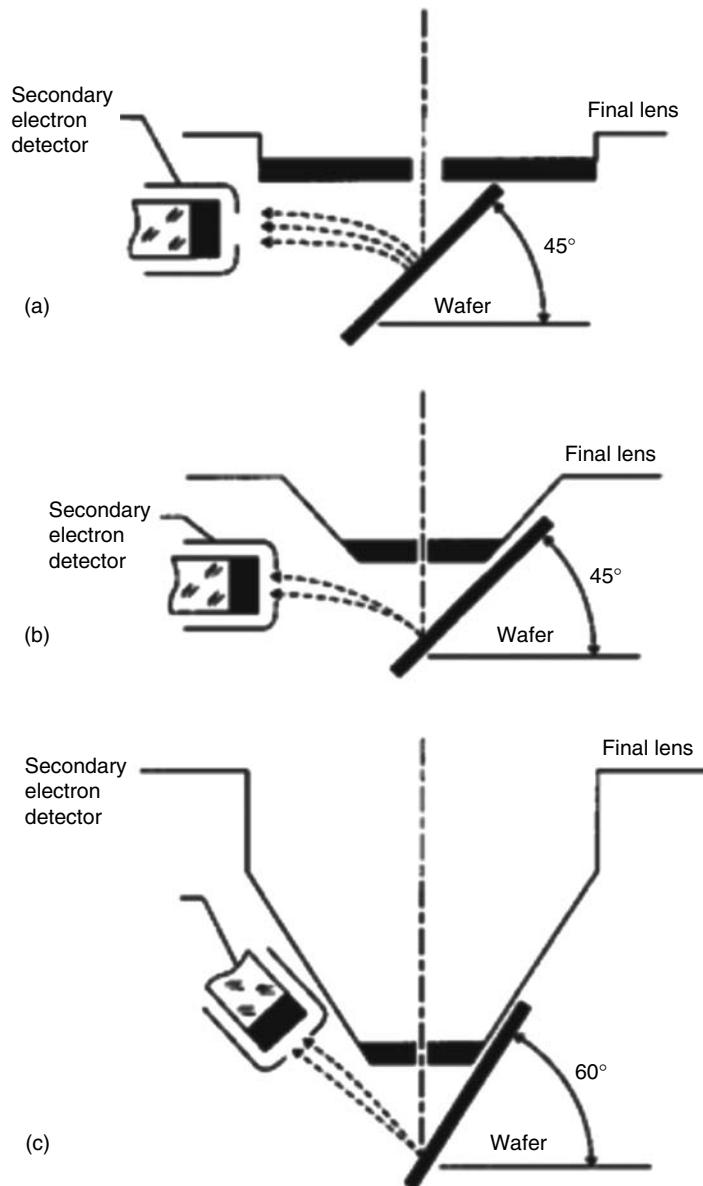
The samples viewed in the SEM are incredibly diverse. They range from very small biological and particulate specimens to specimens as large as jet turbine blades. The basic design of the SEM readily supports these applications with few instrumental changes except for the specimen chamber, vacuum system (to accommodate the larger volume), and final lens design. As many of the specimens being viewed in the SEMs are by their nature quite large, it is now common for instruments to be designed to accommodate up to 200 mm diameter size samples. Other instruments are being designed for specific industrial applications requiring the observation of even larger samples. Moving samples of such large dimensions rapidly within the specimen chamber, in vacuum, has been a difficult engineering problem. Not only did specimen chambers and stages need to be increased in size and travel, but also final lens technology required improvement.

Until a few years ago, flat *pinhole* lens technology predominated (Figure 9.5a). This was the state-of-the-art of the instrumentation. More recently, 45° and 60° conical lens technology with improved low accelerating voltage performance began to improve the manipulation and viewing of the sample within the specimen chamber (Figures 9.5b and 9.5c). However, these were still pin-hole type lenses and limitations imposed by the sample/lens geometry on the instrument resolution remained. For example, even a 60° conical lens having a broad front face would still be restricted to rather long working distances with highly tilted samples. Over the years, two innovative improvements in lens design were introduced. The first improvement was immersion lenses with through-the-lens electron detection and the second was extended field lens technology.

#### 9.2.2.1 Immersion Lens Technology

Pinhole lenses have always been restricted in that the space between the lens and the sample had to be shared by the electron detector (Figure 9.6a). Therefore, in typical SEM applications, some open working distance between the final lens and the sample surface is required to permit electron collection. Scanning transmission electron microscopes (STEMs) have, for many years, been able to place specimens directly into the bore of the objective lens, effectively immersing the sample into the lens field at essentially very short working distances. Unfortunately, the space in the lens is quite small and restricts the size of the specimen to be viewed to a few millimeters. The immersion lens concept and the through-the-lens electron collection technique were adapted into ultrahigh-resolution SEMs (Figure 9.6b). In this configuration, secondary electrons (SEs) were caught in the field of the lens and drawn upward to be collected by the detector placed above the lens. The term *through-the-lens detection* describes this operation. The fact that signal electrons can be transported back through the lens that focused the primary electron beam on the sample is a concept that has been reviewed by Kruit.<sup>21</sup> Unfortunately, the sample size restriction remained in the early instruments. Opening up the bore of the final lens and placement of the electron detector into the space above the lens also improved this geometry for normal SEM operation. In some instances, a small sample could even be carefully raised into the final lens bore for higher resolution (Figure 9.6c).

These solutions allowed shorter working distances, even for larger samples, and thus higher source demagnification and attainable instrument resolution. This approach proved to be very successful for instruments that do not require tilting, such as in-line wafer metrology instruments where low electron landing energy and high beam current with excellent resolution are indispensable. In these instruments the wafers could be viewed at short, close to 1 mm working distance with

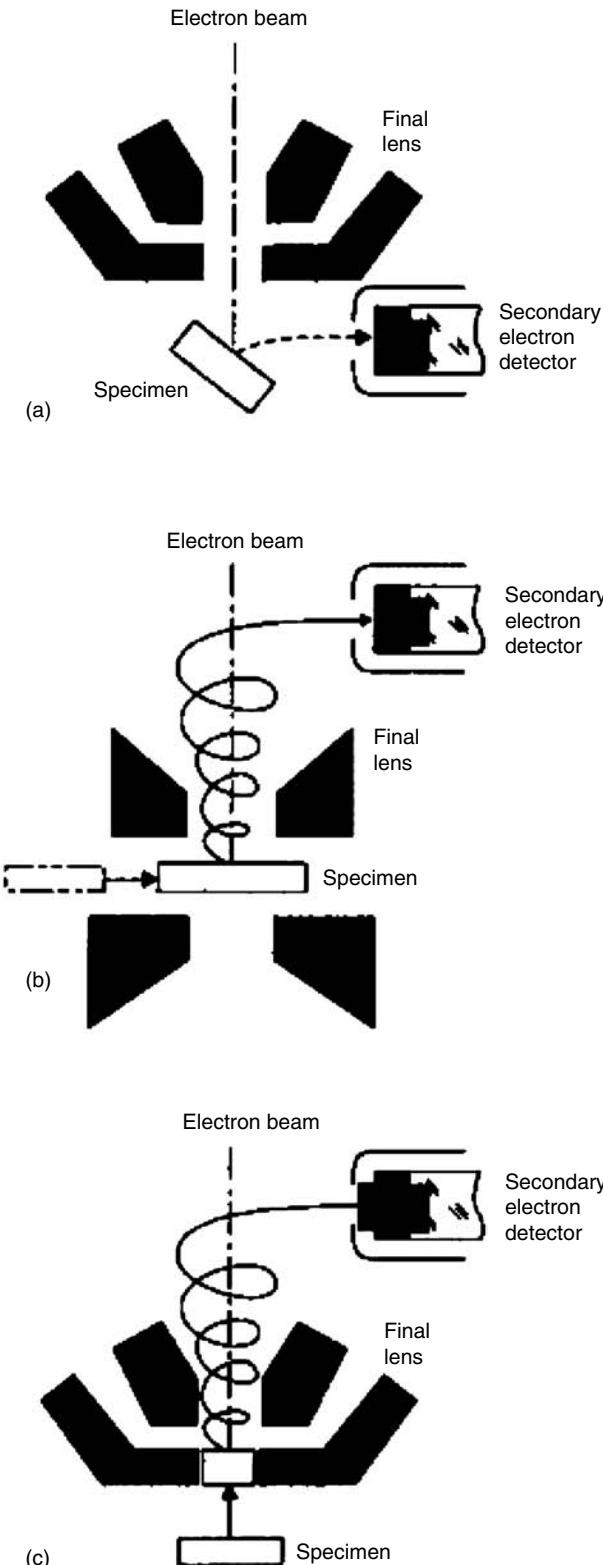


**FIGURE 9.5** The various scanning electron microscope (SEM) final lens designs: the flat pinhole lens (a) and 45° (b) and 60° (c) conical lenses allow for more or less sample tilt and area to be investigated.

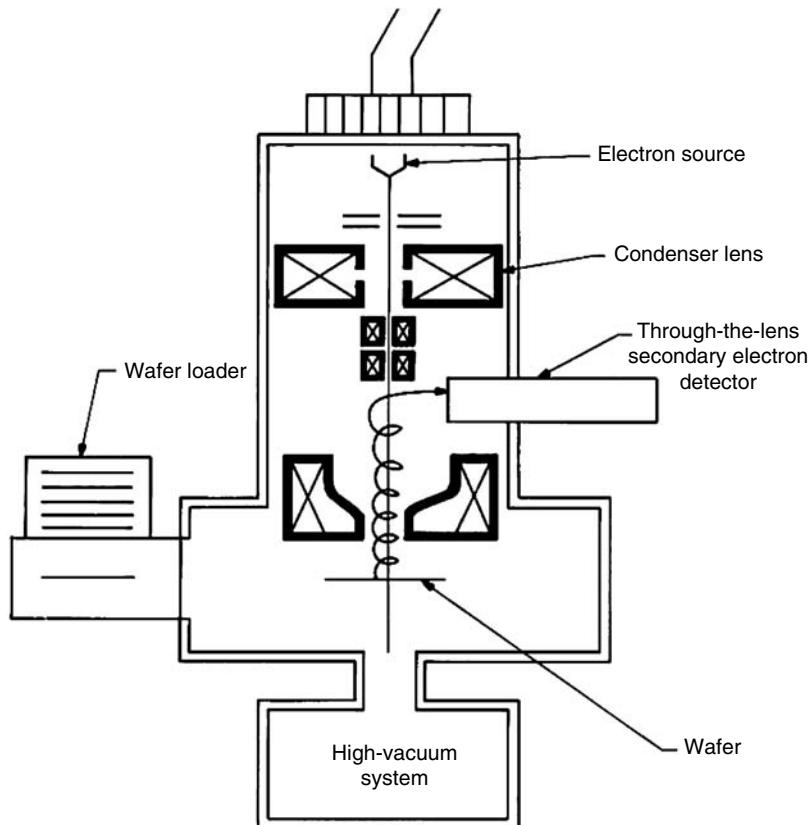
excellent signal collection efficiency and high throughput of 50 or more large wafers measured at five sites per hour (Figure 9.7).

### 9.2.2.2 Extended Field Lens Technology

It is well known that to obtain the highest resolution in the SEM, short working distance is required. Placement of the sample into the bore of the final lens near the principal plane of the lens is one alternative (as discussed earlier), but such an approach is limited to very small samples. Mulvey<sup>22</sup> proposed the design of a new type of lens, referred to as a snorkel lens, where the imaging magnetic field of the lens extends entirely beyond the lens structure. In essence, rather than placing the specimen



**FIGURE 9.6** Lower and through-the-lens secondary electron detection.



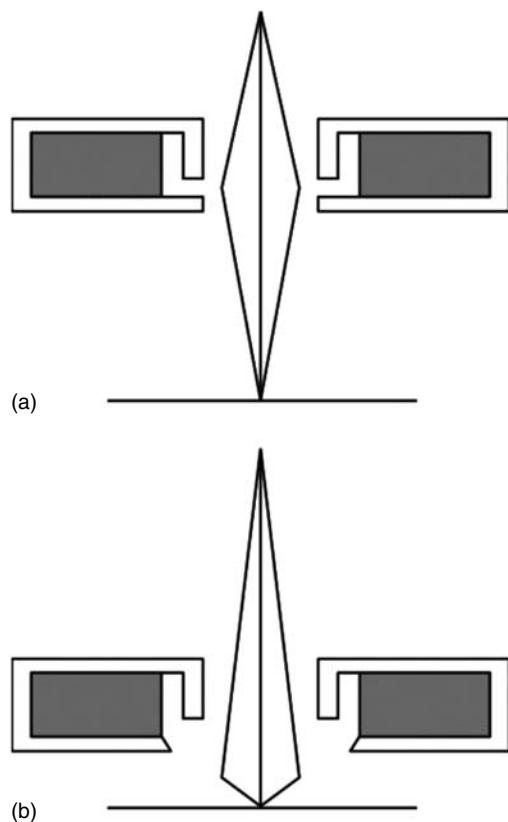
**FIGURE 9.7** Wafer metrology scanning electron microscope.

into the bore of the lens, the lens extends the field toward the sample. Employing an inverted snorkel or extended field type lens as the final lens of the SEM enables a large sample to be essentially immersed in the field of the lens external to the bulk of the lens (Figure 9.8).

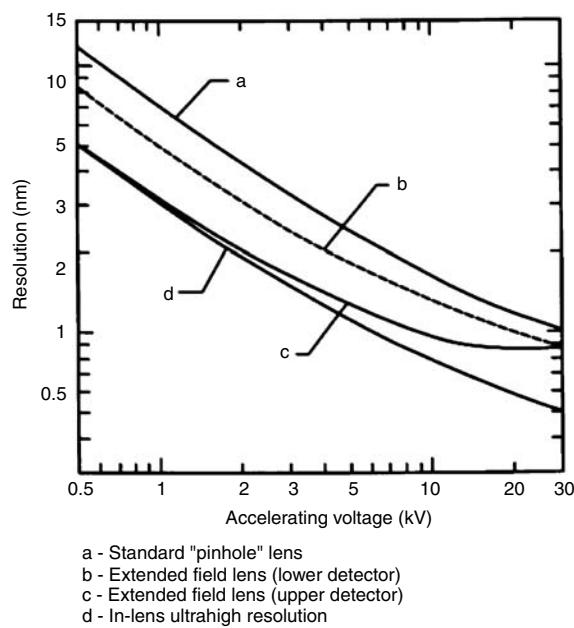
Because of the very short working distance resulting from this lens concept, high resolution is possible, especially if through-the-lens electron collection is also employed (Table 9.1). For laboratory and inspection instruments, an SE detector could be placed in the sample chamber as in the conventional design and a second detector placed above the lens for extremely short working distance, high-resolution operation. Either detector could be used depending on the need. Figure 9.9 shows a graphical comparison of a field emission laboratory type instrument with *pinhole* lens technology to one having extended field lens technology (with two detectors, upper and lower as described previously). Apparent in the graph is the substantial improvement in resolution possible with this technology, which approaches that of the ultrahigh resolution in lens instrument. Effectively, instruments with this technology can now resolve as well at low accelerating voltage (1 kV) as instruments equipped with lanthanum hexaboride that can resolve at high accelerating voltage.

### 9.2.3 SCANNING ELECTRON MICROSCOPE ELECTRONICS AND DIGITAL IMAGE STORAGE AND IMAGE ANALYSIS

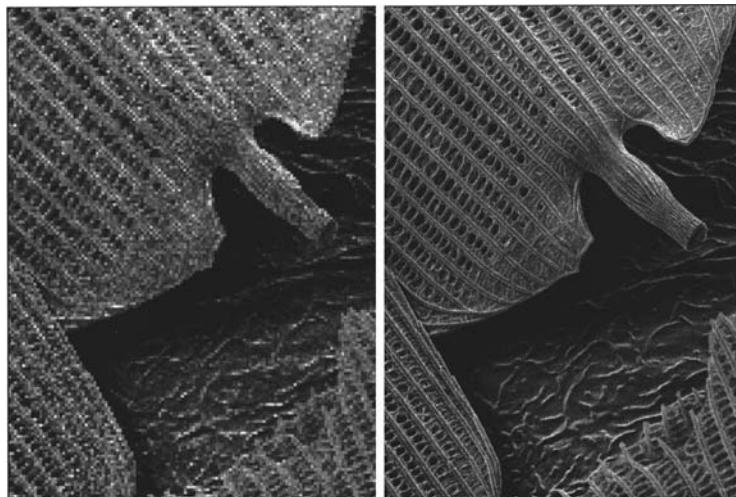
The electronics play an extremely important role in the proper functioning of an SEM. The SEM electronics have undergone radical changes over the past 15 years. One of the major advancements applied to SEM has been the incorporation of digital imaging technology. Early SEMs were only analog in nature with the display cathode ray tubes (CRTs) only slow-scan in nature. Slow-scan



**FIGURE 9.8** Regular (a) and Snorkel (b) lenses.



**FIGURE 9.9** Comparison of calculated resolution capabilities of various scanning electron microscope column and detector designs.



**FIGURE 9.10** Scanning electron microscope images of a butterfly wing acquired at  $512 \times 512$  pixel (left) and at  $4096 \times 4096$  pixel resolution (right); both images were digitally magnified eight times. The horizontal field-of-view is  $35 \mu\text{m}$  for both images.

operation is often very difficult and tedious for the operator, and proper focusing and astigmatism correction techniques are very difficult to learn. Advancements in semiconductor technology, notably the availability of less-expensive, high-density memory chips, and the development of inexpensive high-speed analog-to-digital converters, mass storage, and high-performance central processing units have fostered a revolution in the SEM field. Today, most modern SEMs have factory-installed digital electronics as a standard feature. These instruments generally have 8 or 256 bit gray level resolution with at least  $512 \times 512$  pixel images acquired at TV rate. Many of the more modern instruments operate at either  $2048 \times 2048$  or higher pixel images and at 10 bit or higher gray levels.<sup>23,24</sup> In addition, current slow-scan commercial frame grabber cards, directly applicable to the SEM, can have upward of 12–14 bit resolution, which permits image acquisition and measurement at  $4096 \times 4096$  pixel resolution or greater.<sup>23,24</sup> The value of high pixel density acquisition is shown in Figure 9.10. The left image in Figure 9.10 is an image of a butterfly wing that was acquired as a  $512 \times 512$  pixel image then digitally magnified eight times in the computer. The right image of the same specimen was originally acquired at  $4096 \times 4096$  pixel resolution and also digitally magnified eight times. Note the lack of detail in the first image.

Predigital electronic SEMs were plagued by the general problem of SEMs having a poor signal-to-noise ratio, especially at low accelerating voltages and fast scan rates. Developments in field emission cathode technology improved that situation, but parallel development of the modern digital imaging technology brought both these technologies together into an extremely powerful tool with exceptional flexibility. Some of the advantages afforded by digital imaging are given in Sections 9.2.3.1 through 9.2.3.5.

### 9.2.3.1 TV Rate Scanning

TV rate scanning is not new to SEM, but previously this type of operation had to be done at increased beam currents and thus reduced resolution in many instruments. Today, essentially the *slow-scan* presentation of the SEM is gone and is replaced with a flicker-free, real-time computer image. Digital integration of poorer signal-to-noise ratio images is transparently accomplished by frame buffering and frame averaging of the video signal. TV rate scanning has also been shown to be useful in the reduction of charging on many samples. New SEMs are built with completely separate acquisition and display techniques, and they take full advantage of the modern computer's ability to show

high-resolution images independently from the actual acquisition scheme of the SEM. With these instruments, images can be seen as they develop at any speed; comparisons and all image processing techniques can be applied during and after the collection of the images.

#### 9.2.3.2 Digital Image Storage

Image archiving of the digital images, either to optical or magnetic disks, provides a permanent record that is inexpensive, transportable, and easy to retrieve. Image quality is identical to the originally stored image. Standardized file storage such as the tagged image file format (tiff or tif) or other file formats can enable importation of the images into desktop computers, particularly into statistical analysis and word processing programs. It is important to point out that only nonlossy image compression type files are recommended to be used, as these do not alter the original information content of the collected images. Only for publication or printing purposes it is acceptable to use lossy compression types, such as jpg and jpeg (developed by Joint Photographic Experts Group).

#### 9.2.3.3 Digital Image Transmission

The image and measurement data can be transmitted via data lines to remote locations. It is readily possible to view the SEM image from a remote location and actually operate the SEM from that location in real time.

#### 9.2.3.4 Real-Time Image Analysis and Processing

Digital enhancement of the image can be done as the image is acquired, and the image and data can be processed at the SEM console. It should be noted that in many laboratory and metrology instruments the signal undergoes some processing as it is transported through the video chain. The operator may not even be provided with the ability to view the *raw* data. Blindly allowing the image or data to be processed should be approached with caution, and unaltered raw data should always be able to be obtained from an instrument.

#### 9.2.3.5 Optimization of Operating Conditions

The incorporation of fast computers allows digital SEMs to automatically optimize the operating conditions, such as the brightness, contrast, focus, and astigmatism correction. The operator can save optimum operating conditions for a particular sample set, then reload them as needed. Many of the instrument parameters that need to be changed when instrument conditions are altered can be changed automatically through look-up tables.

Until ~15 years ago, digital imaging was limited by the power, memory, and cost of the computer systems available, and, therefore, much of the digital imaging was done externally through interfacing to a powerful minicomputer coupled to an x-ray microanalysis system. Today, many desktop computers have computing capabilities surpassing these early minicomputers. Computer systems are now small and inexpensive enough to be directly incorporated into the SEM electronics console as a standard component by the SEM manufacturer. This concept presents a major advantage, because the digital architecture of modern SEMs now permits the application of a whole host of peripheral technologies associated with and being developed for the personal computer industry to be readily applied to the SEM.

### 9.3 THE SCANNING ELECTRON MICROSCOPE SAMPLE

A wide variety of specimens and samples are viewed in the SEM, and as stated earlier many of these samples are quite large. Thus, specimen size has become one major SEM design driver. The major requirement for any sample is that it must be compatible with the vacuum of the specimen

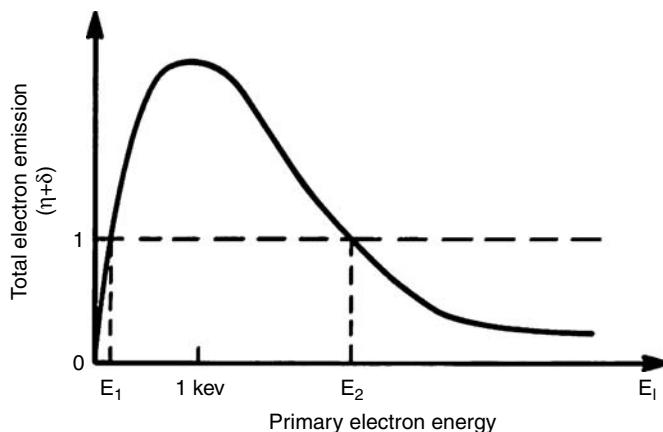
chamber—although specialized instruments have recently been developed that modify that requirement (see Section 9.6.5). Many samples require no preparation other than temporary placement on or attachment to the instrument stage. Other samples may necessitate some degree of specimen preparation depending on their nature.<sup>25</sup>

The design of modern SEMs makes it possible to rapidly load and quickly survey specimens, such as complete in-process wafers at extremely high throughput. Turbomolecular pumping and computer-controlled motorized sample stages enhance measurement throughput. For example, the present technology of semiconductors requires specimen chamber dimensions and stage motions to be sufficiently large to inspect and measure sub-50 nm structures for up to entire 300 mm semiconductor wafers, and larger specimen/wafer requirements are on the horizon. Special voltage contrast applications may require even larger chambers and stages for viewing items, such as completed printed circuit boards and flat panel displays. The flexibility of the SEM architecture permits such a vast array of specimens to be viewed and pertinent information obtained from them.

### 9.3.1 NONDESTRUCTIVE INSPECTION

Nondestructive inspection in an SEM implies that the specimen is not altered before insertion into the SEM or that the inspection in the SEM itself with its vacuum or the electron irradiation does not modify the function of the sample. Historically, scanning electron microscopy was done at relatively high accelerating voltages (typically 20–30 kV) to obtain both the best signal-to-noise ratio and image resolution. At high accelerating voltages, nonconducting or semiconducting samples require a thin layer of gold or a similar material to provide conduction to ground for the electrons and to improve the signal generation from the sample. Further, early instruments were designed to accept only relatively small samples so that a large sample such as a full wafer, typical of the modern semiconductor industry, needed to be broken prior to inspection. This was not cost-effective because for process monitoring it was necessary to sacrifice several rather expensive wafers during each processing run. Modern on-line inspection during the production process of semiconductor devices and integrated circuits (ICs) is designed to be nondestructive, which requires that the specimen be viewed in the SEM uncoated and intact. This required substantial overhaul of the SEM design, thus driving the incorporation of field emission electron sources for improved low landing energy performance, large chamber capability, clean vacuum systems, and digital frame storage. Therefore, the semiconductor industry is and has been a primary driver for many of the modern improvements to the SEM.

Nondestructive inspection in an SEM has another concern: high-energy electrons interacting with the sample can also damage the devices.<sup>26–28</sup> Low accelerating voltage (low electron landing energy) inspection is thought to eliminate, or at least to minimize, such device damage. (The landing energy is generally given in electron volts.) To accomplish this in the SEM, the sample is typically viewed at low landing energies. Low landing energy operation is generally defined as work below  $\sim 2.5$  keV—generally within a range of  $\sim 0.2$  to 1.2 keV. Many of these modern wafer and photo mask measurement SEMs use large sample biasing potentials, so the landing energy of the electrons is defined by both the sample biasing and accelerating voltages. Sample biasing greatly improves the SE signal from deep holes and trenches and can be effectively used for focusing the primary electron beam as well. Further advantages derived from operating the SEM at low landing energies are that the electrons impinging on the surface of the sample penetrate into the sample a shorter distance and have a higher cross section for the production of SEs near the surface where they can more readily escape and thus be collected. Thus, in this context, nondestructive evaluation requires that the sample not be broken and that it be viewed in an instrument at an accelerating voltage below the point where electron beam damage will become a problem. Hence, an understanding of the sample's electrical characteristics is useful prior to examination.



**FIGURE 9.11** Total emission as a function of primary electron energy.

For low landing energy operation, it is imperative to keep the primary electron beam accelerating voltage at the minimum practical values. This may necessitate variable accelerating voltages in fully compensated steps of as low as 100 V within the practical operating range for photoresist samples of 100 eV to 2.5 keV. In this way, ideal beam energy usually occurs where the SE emission volume results in a near-zero SE build-up on the surface of the sample (see Section 9.3.1.1). This ideal beam energy can vary from sample to sample depending on the incident current, nature of the substrate, or the type and thickness of the photoresist. Variation of only 100 V in accelerating voltage or of a couple of degrees of specimen tilt may change a useless charging image to one yielding useful sample information.

### 9.3.1.1 Total Electron Emission

The behavior of the total electrons emitted from a sample per unit beam electron is shown in Figure 9.11. This graph is extremely significant to nondestructive low accelerating voltage operation; the principles demonstrated are primary to the successful achievement of micrographs of insulating specimens such as shown in Figure 9.10. The points where this curve crosses unity (i.e.,  $E_1$  and  $E_2$ ) are the points where no net electrical charging of the sample will occur (i.e., the number of emitted electrons equals the incident electrons' number).

During irradiation with the electron beam, an insulating sample such as photoresist or silicon dioxide can collect beam electrons and develop a negative charge, causing a reduction in the primary electron beam energy incident on the sample. If the primary electron beam energy is 2.5 keV and the particular sample has an  $E_2$  point at 2 keV, then the sample will develop a negative charge to about  $-0.5$  kV so as to reduce the effective incident energy to 2 keV and bring the yield to unity. This charging can have detrimental effects on the electron beam and degrade the observed image. If the primary electron beam energy is chosen between  $E_1$  and  $E_2$ , there will be more electrons emitted than are incident in the primary beam, and the sample will charge positively. Positive charging is not as detrimental as negative charging, as positive charging is thought to be only limited to a few electron volts. Positive charging does, however, present a barrier to the continued emission of the low-energy SEs. This reduction in the escape of the SEs limits the surface potential but reduces signal as these electrons are now lost to the detector. The closer the operating point is to the unity yield points  $E_1$  and  $E_2$ , the lesser the charging effects are. Each material component of the specimen being observed has its own total emitted electron beam energy curve, and so it is possible that to completely eliminate sample charging a compromise must be made to adjust the voltage for all materials. For most materials an accelerating voltage in the range of  $\sim 1$  kV is sufficient to reduce

charging and minimize device damage. Specimen tilt also has an effect on the total electron emission, and it has been reported that increasing tilt shifts the E<sub>2</sub> point to higher accelerating voltages.<sup>4,29</sup>

Operation of the SEM at low beam energies is useful for the inspection of semiconductor and other classes of samples with a minimum of sample damage and charging. However, a detrimental result is a reduction in the beam current available from the electron source (as compared with high-voltage operation); thus, the signal-to-noise ratio is poorer. This results in a loss in apparent sample detail. High-brightness cathodes and digital frame storage techniques for multiscan signal integration or slow-scan rates coupled with photographic or electronic integration help to overcome this problem. Further, the more abiding problem with low accelerating voltage operation is the lower resolution (as compared to the higher beam energy operation) characteristic of this mode of operation. As discussed earlier, if an instrument is equipped with a high-brightness lanthanum hexaboride cathode capable of 2.5–4 nm resolution at 30 kV accelerating voltage, it may be only able to resolve ~7.5 to 12.5 nm resolution (under similar conditions) at 1 kV. This limitation is balanced by the fact that the electron beam penetrates much less into the sample at low accelerating voltage, and thus, the surface detail is enhanced. This resolution limitation must be understood and made a consideration of any work done in the SEM.

## 9.4 ELECTRON BEAM—SPECIMEN INTERACTIONS

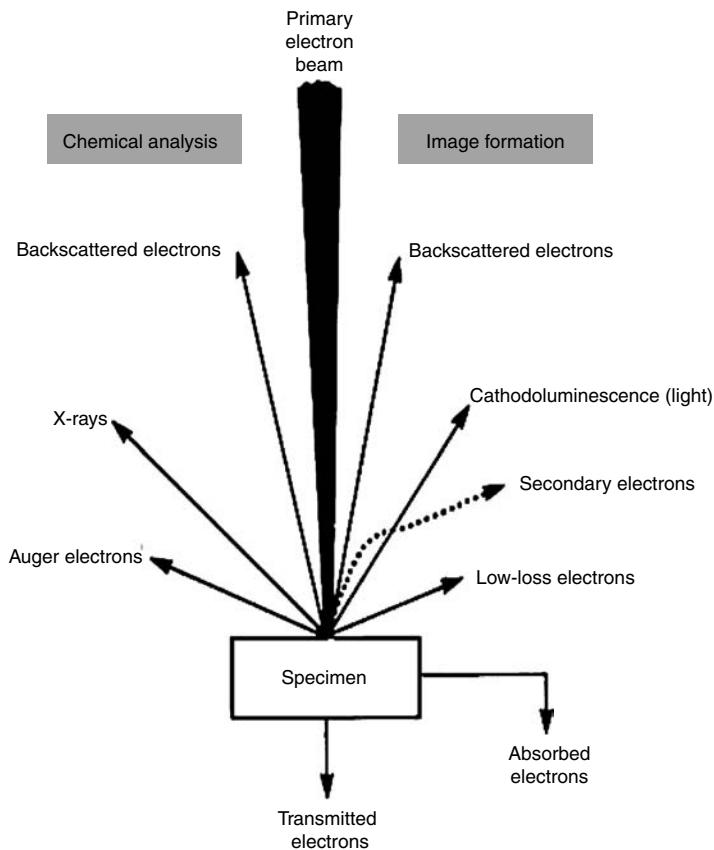
It is well understood that the incident electron beam enters into and interacts directly with the sample as it is scanned. This results in a variety of signals (see Section 9.5) being generated (Figure 9.12) from an interaction region (Figure 9.13) whose size and shape are dependent on the landing energy of the electron beam and the sample composition. The magnitude of this interaction is discussed in Section 9.4.1. For historical and practical reasons, the two major signals commonly used in SEM imaging are divided into two major groups: backscattered and SEs. However, it must be understood that this distinction is arbitrary, especially at low beam energies. Other commonly used signals include the collection and analysis of the x-rays, Auger electrons, transmitted electrons (TEs), cathodoluminescence (CL, light) and absorbed electrons.

### 9.4.1 ELECTRON RANGE

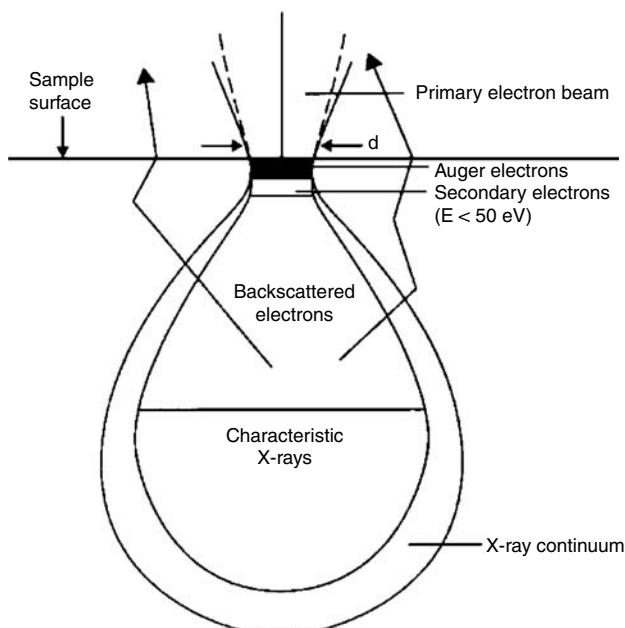
The primary electron beam can enter into the sample for some distance, even at low accelerating voltage; thus, it is important to understand and define this interaction volume. The maximum range of electrons can be approximated by several ways.<sup>30</sup> Unfortunately, as shown in Table 9.3, due to a lack of understanding of the basic physics underlying the interaction of an electron beam in a sample, especially at low accelerating voltages, there are no equations that accurately predict electron travel in a given sample.<sup>30,31</sup> However, one straightforward expression derived by Kanaya and Okayama<sup>32</sup> has been reported to be the most accurate presently available for approximating the range in low atomic weight elements and at low accelerating voltages. The Kanaya and Okayama range R<sub>KO</sub> is shown in the following equation:

$$R_{KO} = 0.0276AE_o^{1.67}/Z^{0.89}\rho$$

where E<sub>o</sub> is the electron energy (eV), A the atomic weight, ρ the density, and Z the atomic number. Using this equation, the calculated range of electrons in pure silicon can be shown to vary from 0.03 μm at 1 keV to 4.72 μm at 20 keV (Table 9.3). If it is considered that this calculated range approximates the boundaries of the electron trajectories as a region centered on the beam impact point (Figure 9.13), then it can be seen that, for example, the high-energy backscattered electrons (BSEs) (see Section 9.5.2) emerging from the surface area of this region do not, in general, carry much information about the high-resolution details making up the surface topography of the specimen at the point of beam incidence. Conversely, the SEs (see Section 9.5.1), due to their



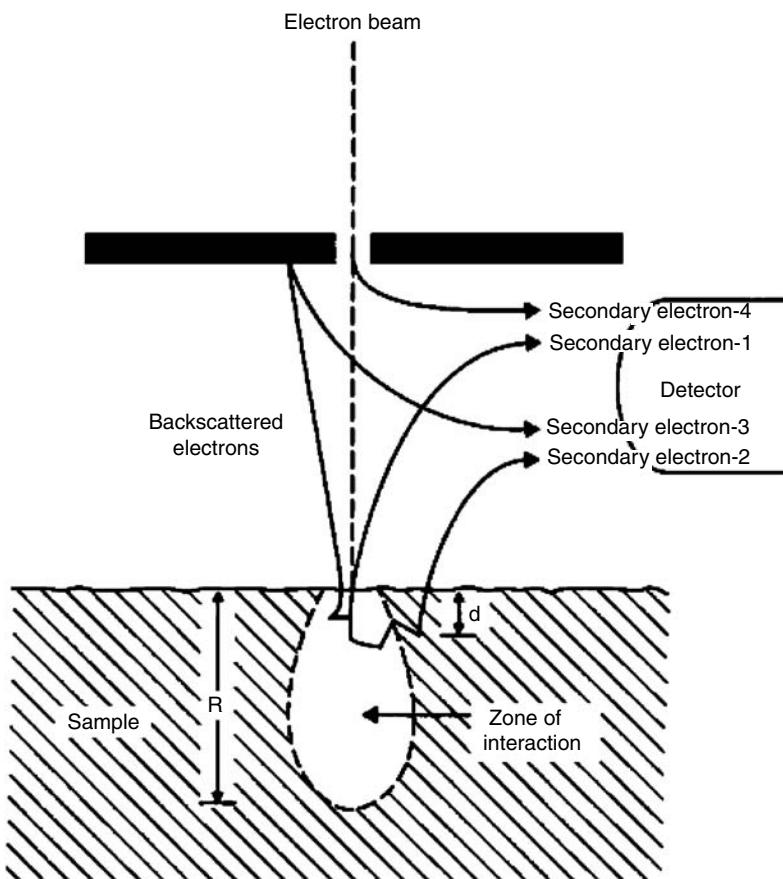
**FIGURE 9.12** Signals generated by the primary electron beam and their typical use.



**FIGURE 9.13** The excited volume and some signals generated within it.

**TABLE 9.3**  
**Comparison of Electron Range in Pure Silicon**

$E_0$ (keV)	Range ( $\mu\text{m}$ )			
	Browne	Reuter	Kanaya/Okayama	Grün
1	0.011	0.054	0.032	0.017
2	0.052	0.138	0.101	0.058
5	0.40	0.47	0.47	0.29
10	1.97	1.17	1.49	0.97
15	4.89	1.99	2.93	1.96
20	9.50	2.90	4.72	3.25



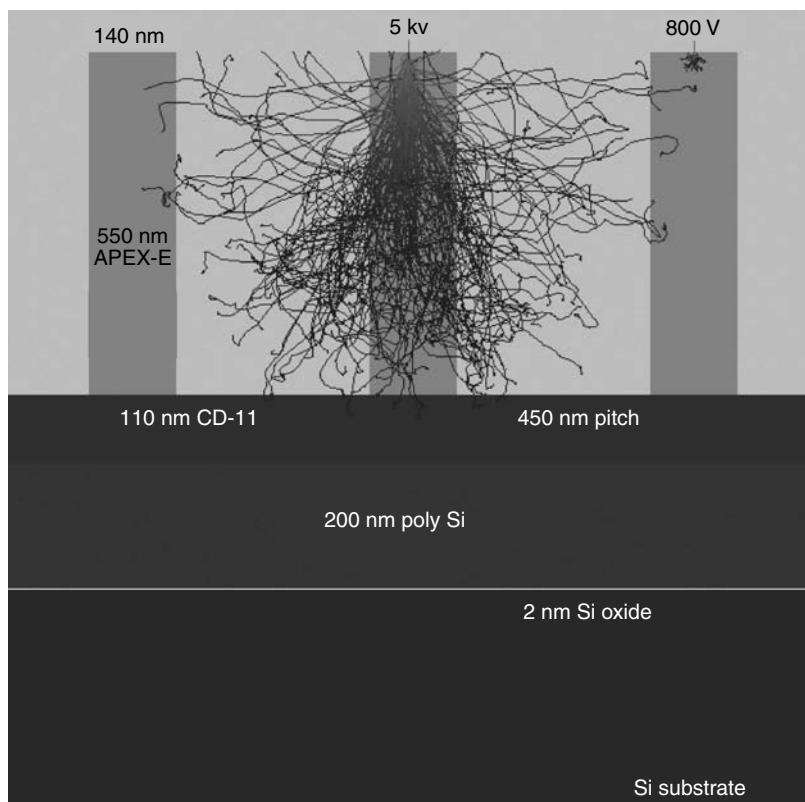
**FIGURE 9.14** Secondary and backscattered electrons.

inherently low energy, cannot reach the surface from deep in the specimen, so typically they escape from only a thin 5–10 nm region beneath the surface. Therefore, they do carry surface-specific information in the area directly interacting with the primary electron beam. But, it should be clearly understood that because of electron scatter in a sample, SEs can originate from points other than the point of impact of the primary electron beam.<sup>33–35</sup> Those that do originate from the point of impact are referred to as SE-1 electrons (Figure 9.14). The SE-1 electrons are the most desirable for

high-resolution imaging.<sup>34,35</sup> It would be convenient if it were possible to collect only the SE-1 electrons; unfortunately, SEs are also created by reemergent BSEs at the sample surface (SE-2) and at the pole-piece of the instrument (SE-3). Other contributions to the collected SE signal include line-of-sight BSEs and other sources particular to each instrument (SE-4). The effects of these four types of contributions to the actual image have not been fully evaluated.

#### 9.4.2 MODELING OF THE SCANNING ELECTRON MICROSCOPE SIGNAL

The appearance of a scanning electron micrograph is such that its interpretation seems simple. It should be understood that a simple interpretation may not always be correct. Care must always be taken so as not to become confused by *obvious* interpretations. The human brain often sees what it expects to see and not always what the SEM presents. The interaction of electrons with a solid is an extremely complex affair (e.g., each electron may scatter several thousand times before escaping or losing its energy, and a million or more electrons per second may interact with the sample); statistical techniques are an appropriate means for attempting to mathematically model this situation. The most adaptable tool, at the present time, is the *Monte Carlo* simulation technique. In this technique, the interactions are modeled and the trajectories of individual electrons are tracked through all regions of the sample (Figure 9.15). Because many different scattering events may occur and because there is no *a priori* reason to choose one over another, algorithms involving random numbers are used to select the sequence of interactions followed by any electron (hence the name, Monte Carlo). By repeating this process for a sufficiently large number of incident electrons (usually 5000 or more), the effect of the interactions is averaged,



**FIGURE 9.15** An integrated circuit manufacturing sample and overlaid 5 keV and 800 eV energy electron trajectories simulated for bulk resist material.

thus giving a useful idea of the way in which electrons will behave in the solid. The Monte Carlo modeling techniques were initially applied to x-ray analysis to understand the generation of this signal. Today, the Monte Carlo technique is being applied to the modeling of all the signal-generating mechanisms of the SEM.

The Monte Carlo technique provides many benefits to the understanding of the SEM image. Using this technique, each electron is individually followed; everything about it (its position, energy, direction of travel, etc.) is known at all times. Therefore, it is possible to take into account the sample geometry, the position and size of detectors, the size, shape, and intensity distribution of the electron beam, and other relevant experimental parameters of the beam. The computer required for these Monte Carlo simulations is modest, and so even current high-performance desktop personal computers can produce useful data in a reasonable time.

In its simplest form,<sup>22,36-38</sup> the Monte Carlo simulation allows the backscattered signal to be computed, as this only requires the program to count that fraction of the incident electrons that subsequently reemerge from the sample for any given position of the incident beam. By further dividing these BSEs on the basis of their energy and direction of travel as they leave the sample, the effect of the detection geometry and detector efficiency on the signal profile can also be studied. However, while this information regarding the BSEs is a valuable first step under most practical conditions, it is the SE signal that is most often used for imaging in the SEM, and recent work to improve this model is being done. Simulating the SE image is a more difficult problem because multiple electron trajectories must be computed and followed. While this is possible in the simplest cases, it is a more difficult and time-consuming approach when complex geometries are involved.

The importance of being able to model signal profiles for a given sample geometry is that it provides a quantitative way of examining the effect of various experimental variables (such as beam energy, probe diameter, choice of signal used) on the profile produced and gives a way of assessing how to deal with these profiles and determine a criterion of line edge detection for the given edge geometries and thus a linewidth.

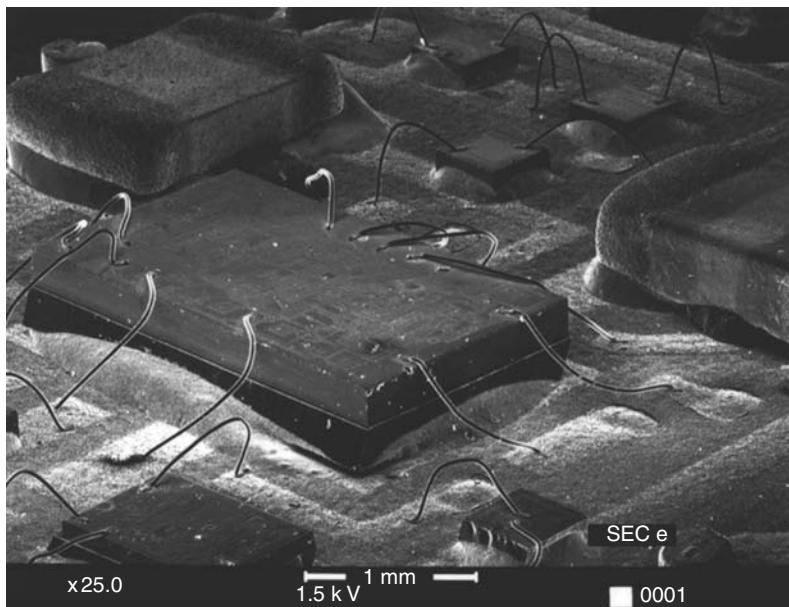
## 9.5 SCANNING ELECTRON MICROSCOPE SIGNALS

The interaction of an energetic electron beam with a solid results in a variety of *signals* being generated from a finite interaction region of the sample. A signal, as defined here, is something that can be collected, used, or displayed on the SEM, and the most commonly used SEM signals are shown in Figure 9.12. The size of the interaction region, as discussed previously is directly related to the accelerating voltage of the primary electron beam, the sample composition, and the sample geometry. All of the signals which are produced within the interaction zone (Figure 9.13) and leave the sample surface can be potentially used for imaging if the instrument is properly equipped to collect, display, and utilize them.

### 9.5.1 SECONDARY ELECTRONS

The most commonly collected signal in the SEM is the SE; most micrographs readily associated with the SEM are mainly composed of SEs (Figure 9.16). Some of the SEs, as discussed previously, are generated by the primary electron beam within a thin layer of the specimen surface, their escape depth varying with the accelerating voltage and the atomic number of the specimen. Typically, this depth ranges from 2 to 10 nm for metals and 5 to 50 nm for nonconductors. SEs also result from BSEs as they leave the specimen surface or collide with inside surfaces of the specimen chamber. The number of SEs emitted from a particular sample relates to the SE coefficient of the materials comprising the sample and other factors such as surface contamination.<sup>1,39-41</sup>

The SEs are arbitrarily defined as those electrons generated at the sample that are  $\leq 50$  eV in energy. The SEs are the most commonly detected for low accelerating voltage inspection due to



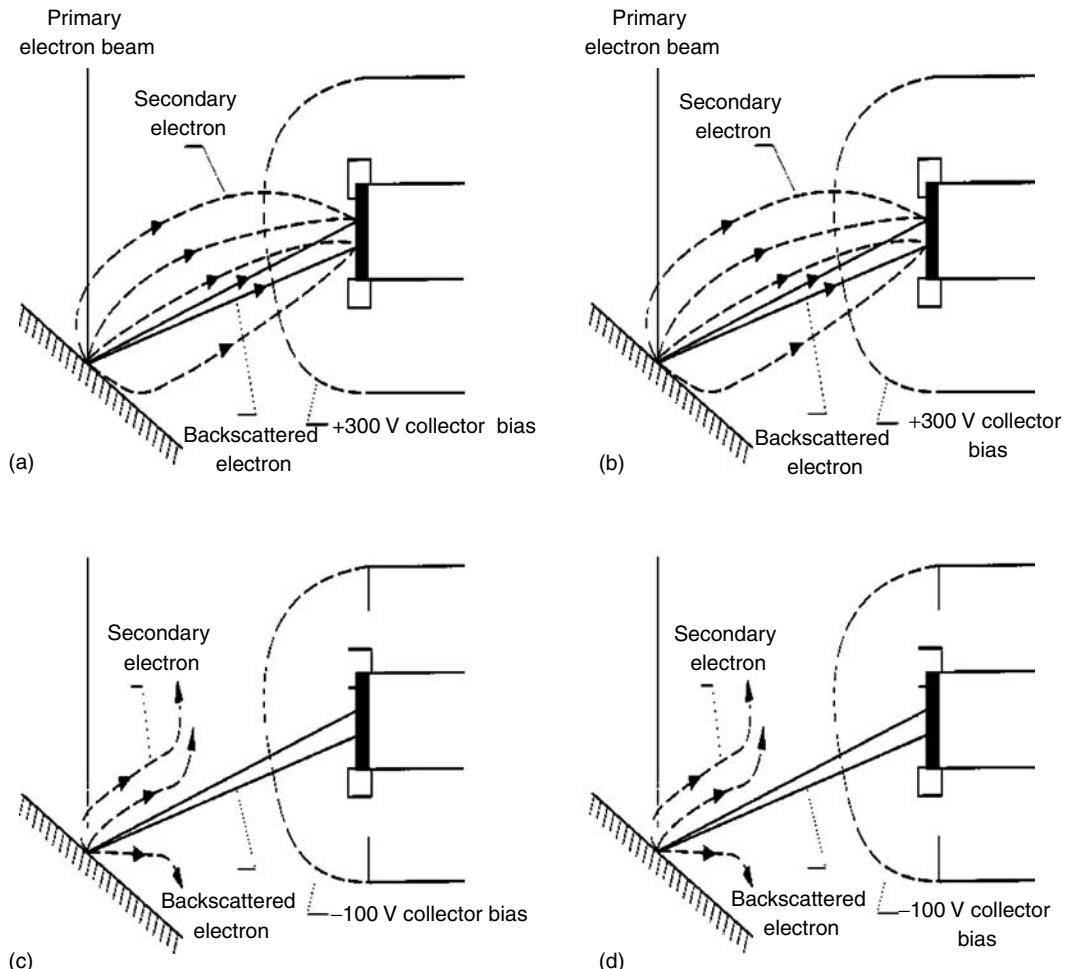
**FIGURE 9.16** Secondary electron image of an integrated circuit sample. The horizontal field-of-view is 5 mm.

their relative ease of collection, and because their signal is much stronger than any of the other types of electrons available for collection. But, due to the low energy of the SEs, they cannot escape from very deep in the sample, and thus their information content is generally surface specific. Consequently, the information carried by SEs potentially contains the high-resolution sample information of interest.

SEs do not originate only from the location of the primary electron beam, as discussed in Section 9.4.1. It has been calculated that the number of remotely generated SE (i.e., energy <50 eV) is larger than those generated from the primary electron beam interaction by a factor greater than three.<sup>42</sup> Peters<sup>43,44</sup> has measured the components of the SE signal from gold crystals and has found that, depending on the sample viewed, for the total SE image the contribution of the SE-2 is ~30% and that of the SE-3 electrons is ~60% as compared to ~10% of contribution by the SE-1 derived signal. The standard SE detector (discussed in Section 9.5.1.1) does not discriminate between these variously generated electrons, and thus the collected and measured SE signal is composed of a combination of all these signal forming mechanisms. The difficulties in interpreting this composite signal can lead to interpretation errors that can be highly variable and have a strong dependence on sample composition, sample geometry, and to a lesser or greater extent (depending on instrument design) on other physical factors, such as the instrument's internal geometry that induces anomalies in the detector collection field (i.e., stage position). Furthermore, as this signal is highly variable and often instrument specific, it is extremely difficult to model.

### 9.5.1.1 Collection of Secondary Electrons

In most SEMs, SEs are generally collected by the use of a scintillator type detector of the design of Everhart and Thornley<sup>45</sup> or a modification of that design. Owing to the low energy of the SE signal, the electron paths are easily influenced by any local electric or magnetic fields; therefore, this detector is equipped with a positively biased collector to attract the SEs. A collector screen or electrode at the front of the Everhart–Thornley (ET) detector operates at approximately +300 V, applying a field that influences and effectively attracts the SEs toward the detector (Figure 9.17).



**FIGURE 9.17** Secondary and backscattered electron detection with an Everhart–Thornley detector.

The SEs thus drawn are accelerated further by a +10 kV potential to a scintillator that (when hit by the electron) converts the electrons to light photons. The light is carried along a light-pipe outside the vacuum chamber where it is converted into current and amplified by a photomultiplier tube and then displayed as a modulating video signal. This type of electron detector with its excellent amplification and low-noise characteristics made scanning electron microscopy a viable technique. The collection efficiency of an SE detector relates directly to its position and potential. Detectors that have a location at some off-axis angle, as in many instruments designed to accept detectors for x-ray microanalysis, show preferentiality of detection with respect to the orientation of feature edges.

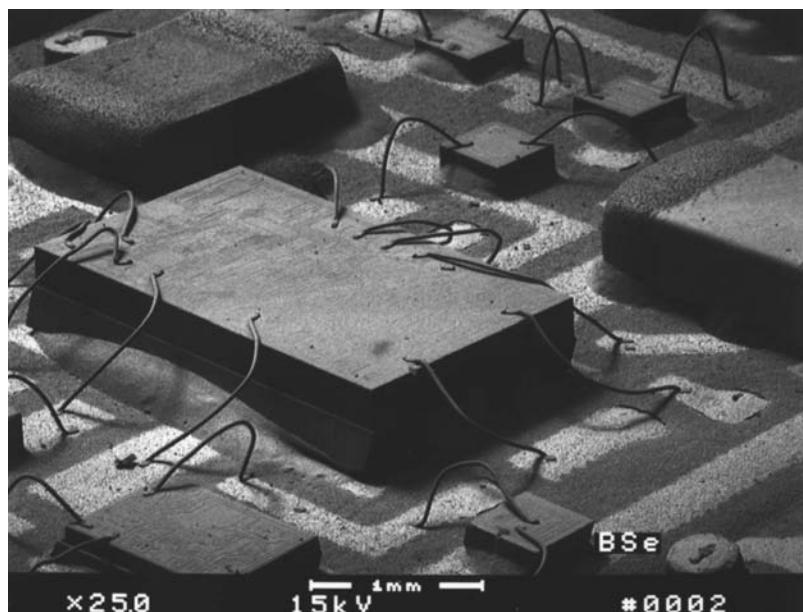
In these cases, it is not possible to achieve symmetrical video profiles. Furthermore, it is not easily determined if the video asymmetry demonstrated is derived from the position of the detector, from other problems introduced by the instrument's electronics, by specimen–electron beam interactions, or by a random summing of all possible problems. To compensate for an off-axis position of the SE detector, the sample can be physically tilted toward it until the video profile of the line becomes symmetrical; then the structure can be aligned as desired on the display by adjusting the raster pattern with digital raster rotation. As an error can be introduced using this technique (i.e., if the sample is tilted or the raster rotation is not linear), it is much more desirable to have an on-axis detector<sup>46</sup> or two similar detectors on either side of the sample and the signals properly balanced

and summed.<sup>7</sup> Some authors<sup>47–50</sup> proposed using microchannel plate (MCP) type detectors. The MCP can be placed above the sample in the specimen chamber or even in the microscope column as in a lens detector.

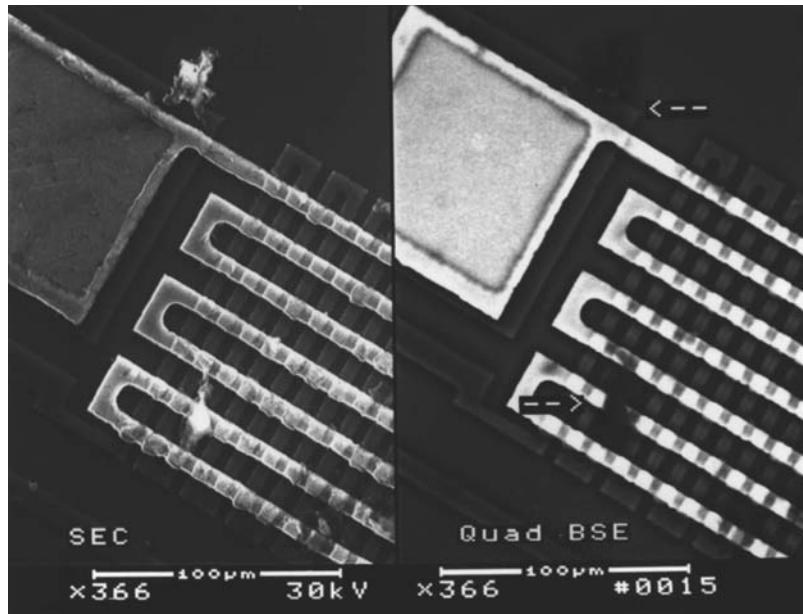
### 9.5.2 BACKSCATTERED ELECTRONS

BSEs are those electrons that have undergone either elastic or inelastic collisions with the sample and are reemitted with an energy that is a significant fraction (generally 50–80%) of the incident beam energy. This means that a 30 keV primary beam electron can produce a BSE up to 30 keV (or less) energy that can be collected and imaged or can interact further with the sample, instrument, or specimen chamber (see Section 9.5.1). The measured BSE yield varies with the sample, detector geometry, and chemical composition of the specimen. Because of their high energy, BSEs are very directional in their trajectories and are not easily influenced by applied electrostatic fields. Line-of-sight BSEs striking the ET detector contribute to all SE micrographs (Figure 9.17b). By applying a –100 V bias to the collector screen of the ET detector, the screen rejects the SEs and collects only the BSE having a small exit solid angle from the specimen, thus providing a highly topographical image of the sample. This concept is clearly demonstrated in Figure 9.18 where, fundamentally the same sample and instrumental conditions are being used as in Figure 9.16, the only difference is in the selection of signal to be recorded.

The nature of the imaging mechanism has a direct effect on the SE image. This is an extremely important point, because this means that the SE signal, usually collected in the SEM, is a composite of a number of signal mechanisms. The SE signal is composed not only of those SEs generated from initial interaction of the electron beam as it enters the sample (SE-1), but also from SEs generated by the escape of elastically and inelastically scattered electrons when they leave the sample surface (SE-2 and SE-3). The emitted BSEs can interact singly or multiply with other structures on the sample or other internal instrument components and generate more SEs (SE-4), and they can also be collected as a component of the SE image if their trajectory falls within the solid angle of collection of the electron detector.



**FIGURE 9.18** Backscattered electron image of an integrated circuit sample. The horizontal field-of-view is 5 mm.



**FIGURE 9.19** Secondary (left) and backscattered (right) electron images showing material contrast of an integrated circuit sample. The horizontal field-of-view for each is 150  $\mu\text{m}$ .

BSEs have a high energy relative to SEs, and thus the BSEs are not affected as greatly by surface charging as the SE. Thus, optimization of collection using tilt and collector bias can often enable observation of uncoated, charging samples. This is shown in Figure 9.19 where on the left-hand side SE image two dust particles are bright, indicating charging (in this case). The BSE image of Figure 9.19 shows the particles (arrows) dark in contrast and with no charge build-up. This figure also demonstrates another characteristic of the collection of the BSE signal. BSEs can relay chemical information about a sample. The angular distribution of BSE as they emerge from the sample surface follows a cosine function when the primary beam is incident normally on the surface. With a solid state detector (discussed in another section) mounted at the pole-piece for improved BSE collection efficiency, areas of higher atomic number (bright) can be differentiated from areas of lower atomic number (dark).<sup>51–53</sup> This is clearly shown in Figure 9.19 where the carbonaceous dust particles were recorded as being dark and the heavier aluminum lines of the circuit are bright. The ability to discern differences in atomic number relates to the BSE coefficient for each element.

### 9.5.2.1 Collection of Backscattered Electrons

BSEs are reemitted from the sample surface in approximately straight lines, and consequently they must be detected by placing a detector in their path. This may be accomplished by the use of either a solid-state diode detector,<sup>54</sup> a channel electron-multiplier detector,<sup>47–50</sup> or a scintillator detector positioned for this purpose.<sup>55</sup> The size and position of the detector affects the image and thus affects any measurements made from it. Therefore, the particular characteristics of the detector must be taken into account when analyzing the observed BSE signal for any application.

BSEs can also be collected through the use of energy filtering detectors or low-loss detectors.<sup>31,56</sup> Energy filtration has the advantage of detecting those electrons that have undergone fewer sample interactions (low loss) and thus have entered and interacted with the sample to a lesser degree (i.e., over a smaller volume of the specimen) and thus carry higher resolution information. This type of detector has been used successfully at low accelerating voltages, although it does suffer from signal-to-noise limitations.<sup>56</sup> Discriminating the BSEs from the SEs created at the sample is

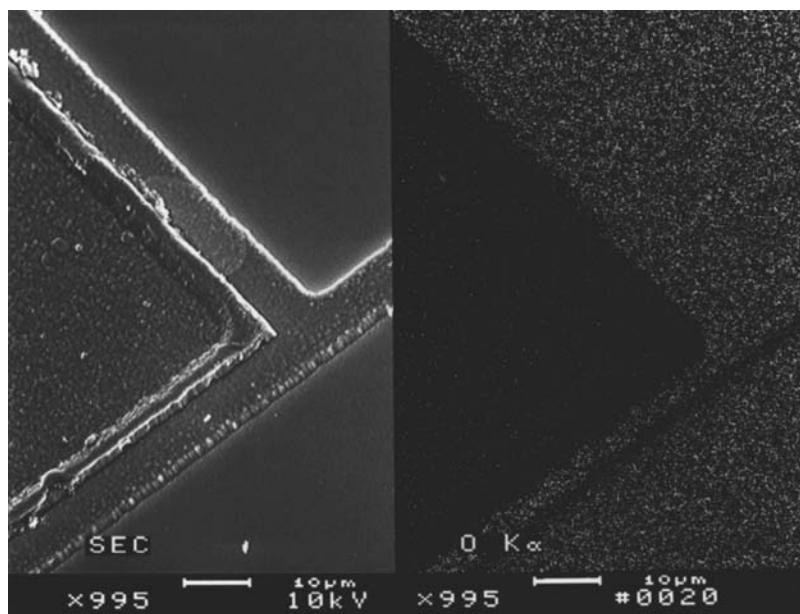
also possible by suppressing the collection of the SEs generated at the sample, allowing the BSEs to escape and create SEs (see Section 9.4.1) through interaction with internal column components (i.e., the final lens pole-piece), and collecting these converted BSEs. This technique is referred to as converted backscattered secondary electron (CBSE) collection.<sup>43,44,57,58</sup> The CBSE technique has also been successfully used at low accelerating voltages,<sup>59</sup> but also suffers from lack of signal and therefore low signal-to-noise ratio and other problems that presently make it unsuitable for other than research environments.

### 9.5.3 X-RAYS

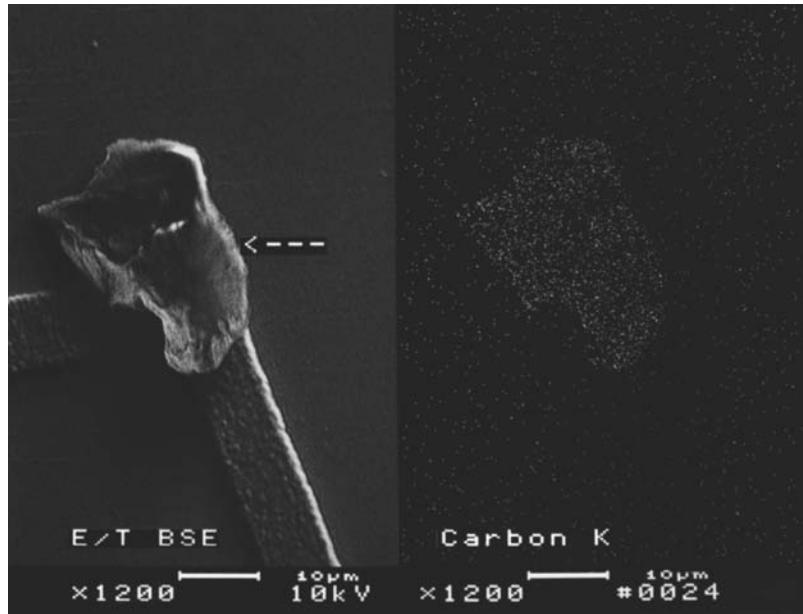
Primary beam electrons interacting with a sample can strike and eject an electron from a component element's inner electron shell. Rearrangement of the remaining electrons in the ionized atom as they seek lower energies results in x-rays characteristic of the element from which they originated. SEMs equipped with either (or both) an energy dispersive x-ray spectrometer (EDS) or wavelength dispersive x-ray spectrometer (WDS) can collect and qualitatively or quantitatively characterize elemental information about the bulk sample or its surface contaminants. Used in conjunction with an SEM, x-ray microanalysis cannot only identify the presence of a particular element or elements, but it can also identify the approximate location of the element in the sample through area mapping or linescan display. The collection of x-rays can take place simultaneously with other SEM modes and is essentially nondestructive to the sample (see Figure 9.20).

#### 9.5.3.1 Collection of X-Rays

The generation of x-rays occurs continually as the electron beam scans and interacts with the specimen. The collection and analysis of the characteristic x-rays generated, however, requires the installation of a special detector, which may be of two types. A detector system that separates the x-rays according to their wavelength is a wavelength dispersive spectrometer. Whereas, one which measures the energies of the incoming X-radiation is referred to as an energy dispersive spectrometer. An SEM can be equipped with either or both types of spectrometer depending on the instrument application.



**FIGURE 9.20** Backscattered electron image and oxygen elemental map on an integrated circuit. The horizontal field-of-view is 60  $\mu\text{m}$ .

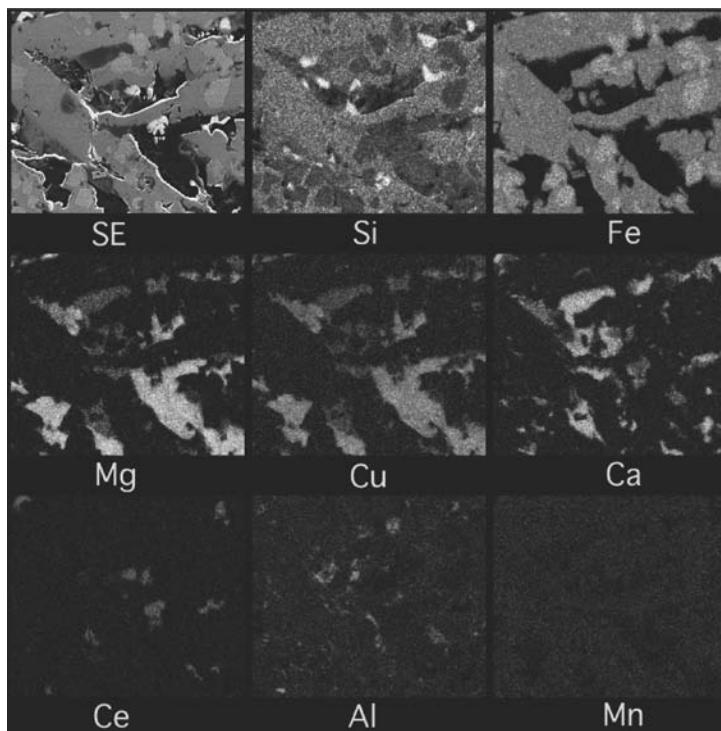


**FIGURE 9.21** Backscattered electron image and carbon elemental map on an integrated circuit. The horizontal field-of-view is 50  $\mu\text{m}$ .

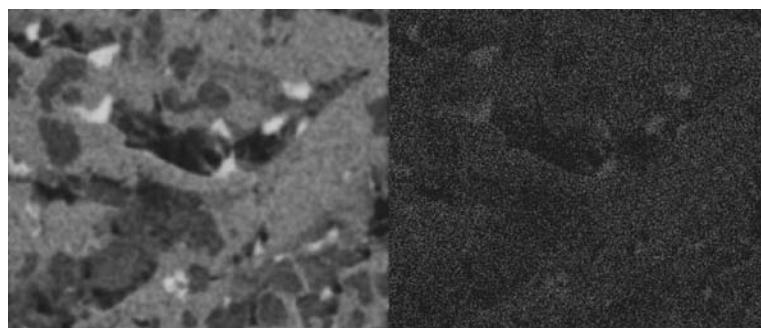
The processing of the collected x-ray information provides a set of peaks that are superimposed upon background radiation. This background is referred to as the continuum or bremsstrahlung radiation. The characteristic energies or wavelengths displayed (depending on the system used) relate directly to the elements present in the sample. For example, if an electron beam interacts with a sample and displaces a K shell electron, it is possible that an L shell electron can fill the vacancy. The transition of the electron from the L shell to the K shell results in the generation of a  $K_{\alpha}$  x-ray specific for the element present in the sample. Other transitions from other shells can also generate characteristic x-radiation (e.g., the transition from the M shell to the K shell yields a  $K_{\beta}$  x-ray). Depending on the particular element, several different characteristic x-rays can be emitted. The x-ray signal, like the other signals of the SEM, can be displayed to obtain information regarding the specimen of interest. Information concerning the elemental distribution can be obtained by allowing the x-ray signal from that element to modulate the image brightness as the electron beam scans the sample. Modulation of the Y-axis as the electron beam makes a single sweep across the region of analysis in the X-direction results in a line profile or linescan with the vertical (Y) displacement corresponding to the intensity of the x-rays from the selected element. Modulation of the brightness of the image by the x-ray signal results in an x-ray distribution map for the chosen element of interest as the electron beam traverses the raster pattern across the region of analysis (Figures 9.20 and 9.21). Modern digital electronics have facilitated the collection and display of the x-ray signal permitting collection of multiple-element maps as shown in Figure 9.22.

When collecting elements of main constituents of a sample for longer times, it is possible to generate gray-scale x-ray images that look very much like electron images as it is illustrated in Figure 9.23. Now, all these x-ray images can be fully quantitative especially with new high-speed x-ray detectors, like the new silicon drift detector (SDD) used in modern energy dispersive x-ray spectrometry.

The spatial resolution of the x-ray signal from a specimen is determined by a number of factors. It should be clearly understood that the characteristic x-rays are detectable from a far greater volume than that of the SEs (Figure 9.13), which greatly reduces the spatial resolution of the x-ray signal as compared to either the secondary or the BSE signal. The composition of the specimen also contributes to the spatial resolution attainable. A specimen of high atomic number is more electron-dense and



**FIGURE 9.22** Multiple-element x-ray map and the secondary electron image (upper right corner) of a complex alloy sample. The field-of-view is 200  $\mu\text{m}$ .

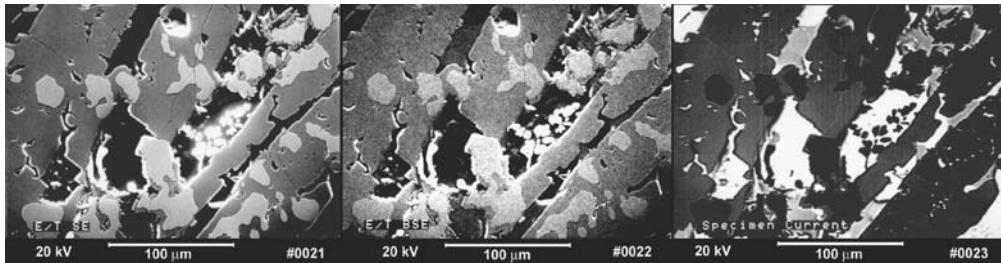


**FIGURE 9.23** Grayscale (left) and dot type Si x-ray (right) maps of the alloy sample of Figure 9.22. The field-of-view is 200  $\mu\text{m}$ .

yields more x-rays from a smaller excited volume for a given electron landing energy. The spectral resolution (measured in electron volts) is a function of the particular x-ray detector. Spectral resolution is the ability of the instrument to separate the wavelengths or the energies that are nearly equal to each other in magnitude. The property of the spatial resolution is represented by the width of a specific peak generated in the SEM by a standard specimen of known composition under standard conditions.

#### 9.5.4 ABSORBED ELECTRONS

The electron beam-specimen interaction typically produces a flow of electrons between specimen and ground. Collected and amplified, this electron beam induced specimen current signal provides useful information about the changes in current from point to point. Unlike other forms of scanning



**FIGURE 9.24** Secondary (left), backscattered (center), and absorbed electron (right) images of an alloy sample. The field-of-view is 250  $\mu\text{m}$ .

electron microscopy, this technique uses the sample as the signal detector. Observation of absorbed electrons provides the basis for techniques such as specimen current imaging and electron beam induced current (EBIC) imaging.

The relationship of the absorbed current ( $I_{SC}$ ) and the other generated signals ( $I_{BSE} + I_{SE}$ ) can be shown as follows:

$$\text{Primary electron beam current} = \text{Total signal current} (I_{BSE} + I_{SE}) + \text{Absorbed current} (I_{SC})$$

The total signal current is essentially equal to the algebraic sum of currents from SEs and BSE. Simply speaking, if we consider an element such as iron that has a BSE coefficient of  $\sim 0.3$  of the primary electron beam current and an SE coefficient of  $\sim 0.4$  of the initial beam current, to prevent the sample from accumulating a charge, a current of  $\sim 0.3$  of the initial beam current must flow out of the specimen to ground. This current flowing to ground can be collected, measured, or imaged (Figure 9.24).

### 9.5.5 CATHODOLUMINESCENCE (LIGHT)

The interaction of the electron beam with certain classes of materials may result in the radiation of visible, ultraviolet, or infrared radiation often referred to as CL or electron-induced fluorescence.<sup>60</sup> Luminescence of organic and inorganic materials is an inherent property. There are many semiconductors, nonconductive materials, and organic substances that emit characteristic CL. The emission of light by a specimen is a very complicated process, which is currently under study but has been reviewed by Yacobi and Holt<sup>58</sup> and Hough.<sup>61</sup> CL results from the recombination of electrons with the holes formed by the sample interaction with the primary electron beam. CL is used in many fields including biology, medicine, compound semiconductor technology, and mineralogy.

#### 9.5.5.1 Detecting Cathodoluminescence

Removing the scintillator from the front of the ET detector facilitates the direct collection of light, although higher-efficiency CL detectors using sensitive photomultipliers are available as accessory detectors. CL imaging is proving to be a powerful method for assessing the effects of crystal growth and processing procedures on material and device quality.<sup>62</sup>

### 9.5.6 AUGER ELECTRONS

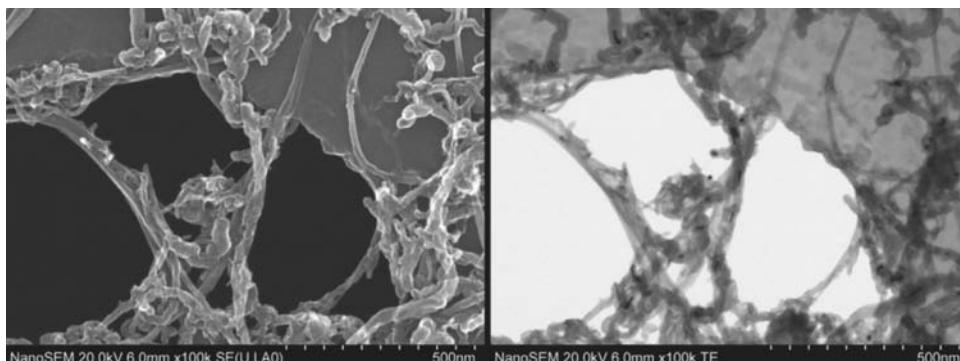
Auger electrons resulting from the interaction between an electron beam and a specimen originate within the first 1–5 nm of the surface and are highly characteristic of the specimen's surface chemistry. Analysis of these electrons enables the determination of elemental composition of films and surface contamination; the addition of ion beam etching facilities depth profiling. Since the sensitivity of this analysis to surface contamination requires the operation of an SEM at very high vacuum at the sample, only certain ultrahigh-vacuum dry-pumped SEMs with extremely low residual hydrocarbon partial

pressure can serve to detect Auger electrons. Auger spectroscopy is useful in the characterization of interfaces and surface contaminant levels down to a fraction of a monolayer.<sup>32</sup>

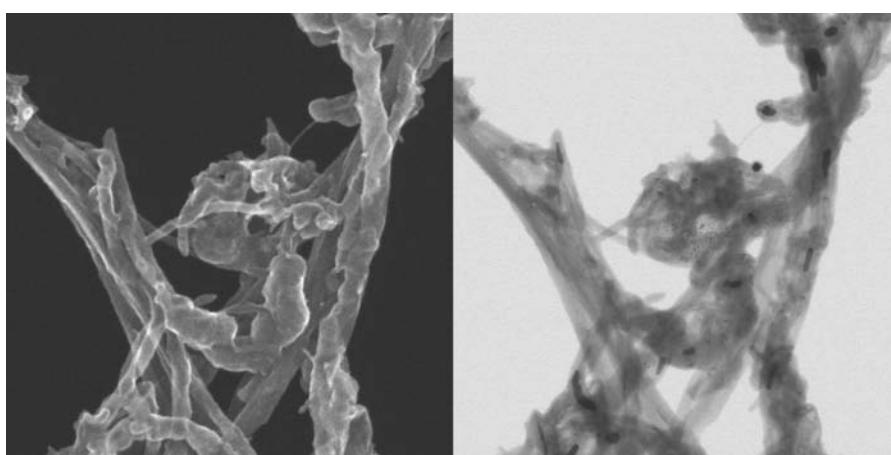
### 9.5.7 TRANSMITTED ELECTRONS

The primary electron beam can pass through a specimen when the range of the electrons is greater than the thickness of the sample. When this occurs, the electrons are said to be TE. This is similar to the function of the transmission electron microscope (TEM) when it is used in the scanning mode of operation or STEM on thin samples (see Chapter 10). This mode of operation is usually facilitated by preparing thin (<1 µm) TEM quality samples so that the electrons will be transmitted through the sample with some level of energy loss and interaction, that is, scattering. With samples of nanometer-scale structures it is also possible to prepare suitable samples by depositing the samples on electron transparent thin substrates (Figure 9.25). This mode of instrument operation enables the correlation of surface morphology to that of subsurface microstructure since simultaneous SE imaging and TE imaging can be done much like that done in the STEM. TE detection was one of the first modes of operation of the SEM,<sup>8,63</sup> and recently it has been used in the accurate metrology of x-ray masks.<sup>64</sup>

Digital images high in image pixels and spatial resolutions make studying of fine details possible (Figure 9.26). Digitally magnified sections of images of Figure 9.25 reveal the presence of few nanometer size Fe residues and single-wall carbon nanotubes.



**FIGURE 9.25** Secondary electron (left) and scanning transmission electron (right) images of a carbon nanotube sample. The field-of-view is 750 nm.



**FIGURE 9.26** Digitally magnified section of secondary electron (left) and scanning transmission electron (right) images of a carbon nanotube sample. The field-of-view is 100 nm.

## 9.6 APPLICATIONS

The SEM is currently being used in a staggering array of applications, and it would be impossible to cover all of them in this chapter. However, a number of examples related to semiconductor applications follow in Sections 9.6.1 through 9.6.5. It should be remembered that the applications for the different types of SEM signals vary depending on the environment in which the SEM operates. Production line applications vary from applications in research and failure analysis. When several SEMs are available for use, some may be specialized and serve for one or only a few of these applications; others may out of necessity be multipurpose.

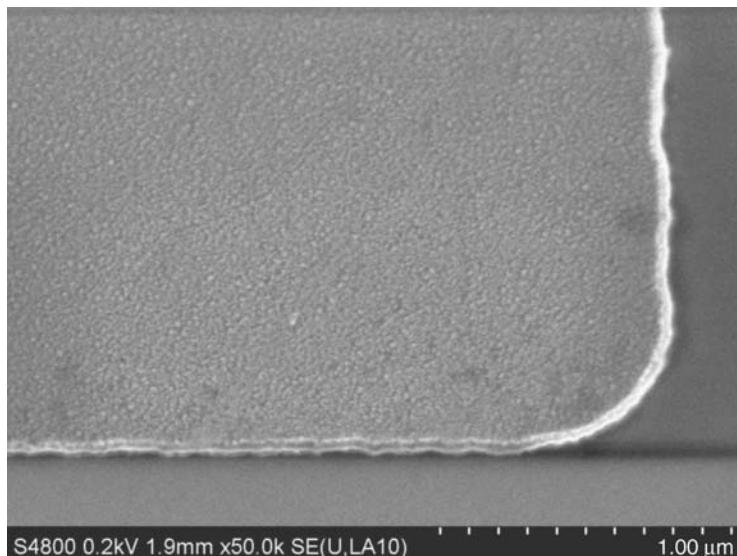
### 9.6.1 LOW ACCELERATING VOLTAGE INSPECTION

Low acceleration voltage operation for production and fabrication is presently of extreme interest to the semiconductor industry.<sup>14,24,26,55,65,66</sup> At low accelerating voltages (100 to 2.5 kV) it is possible to inspect in-process wafers in a nondestructive manner; with the advent of submicrometer geometries, it is imperative that on-line inspection takes place for nearly all processing steps. Modern, clean vacuum technology using turbomolecular and ion pumping enables contamination-free inspection to screen wafers for proper development, registration, etching, resist removal, and the absence of visual contaminants before the next processing step.

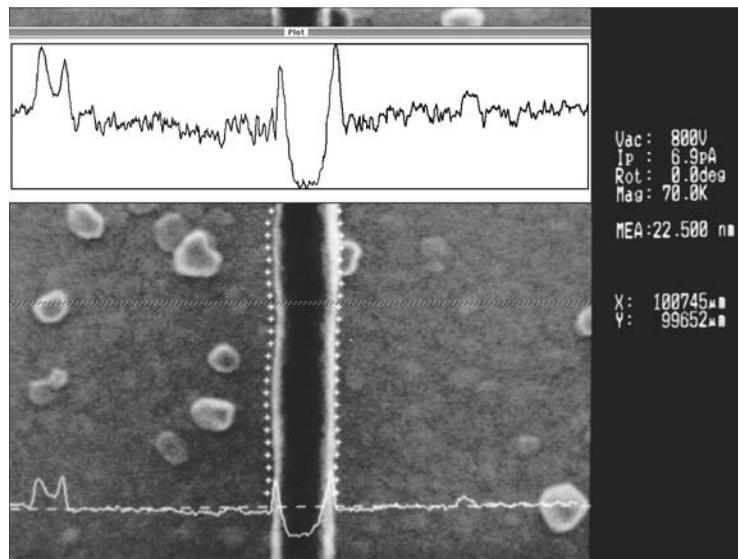
Low accelerating voltage operation is not only restricted to wafer fabrication; photomask inspection can also take place in the SEM (Figure 9.27). Defects in the photomask, either random or repeating, are sources of yield loss in device fabrication. Defects may occur in the glass, photoresist, or chrome and appear as pinholes, bridges, glass fractures, protrusions, solvent spots, intrusions, or even missing geometrical features. Many of the techniques developed for the semiconductor industry are being applied elsewhere, such as in the polymer industry.

#### 9.6.1.1 Linewidth Measurements

Linewidths and other critical dimensions (CDs) of IC device structures must be accurately controlled to ensure that the IC performance matches design specifications. However, traditional light-optical methods for the linewidth measurement of modern IC geometries are not able to attain the



**FIGURE 9.27** A 200 V accelerating voltage image of a Cr on quartz photomask sample. The field-of-view is 2.5  $\mu\text{m}$ .



**FIGURE 9.28** Linewidth measurement with critical dimension scanning electron microscope (CD-SEM).

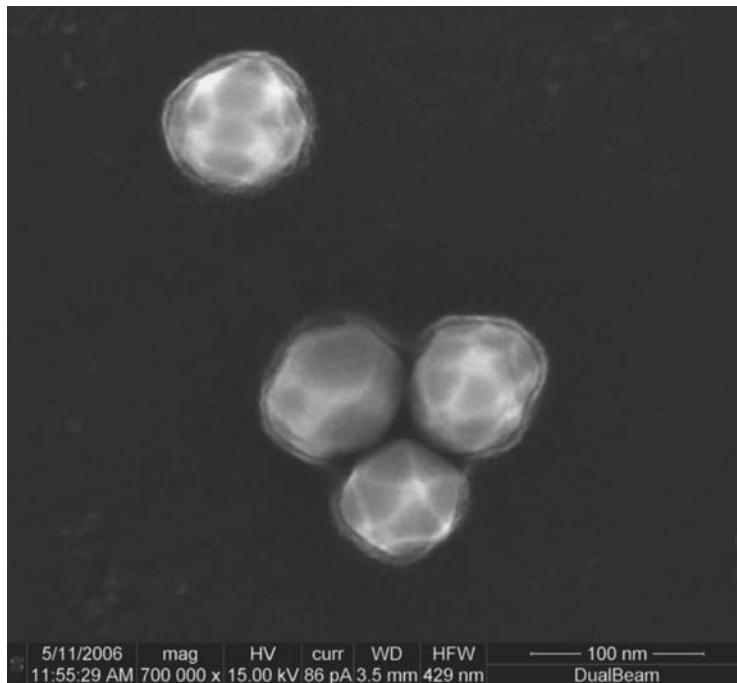
accuracy or precision necessary for inspection of these devices. Since present wafer fabrication methods employ resist exposure radiation of very short wavelength, such as 193 nm ultraviolet light and electron beams, it follows that testing and measuring the fabricated structures reaching sizes down to 10 nm would involve instrumentation capable of viewing and measuring at better than 1 nm resolution.

Two measurements critical to the semiconductor industry are *linewidth* and *displacement*. Linewidth is the size of an individual structure along a particular axis (Figure 9.28), and displacement is the measurement of the separation between the same position on two or more structures.<sup>23,27,54</sup>

Unlike the traditional optical microscope, the range in SEM magnification continuously spans more than four orders of magnitude. All SEM linewidth measurement systems rely on the accuracy of this magnification, computed from numerous internal instrument operational factors including working distance and acceleration voltage. Although acceptable for most applications, the magnification accuracy of many typical SEMs may be inadequate for critical measurement work since the long-term magnification may vary with time.<sup>1</sup> For critical, reproducible linewidth measurement, all sources of magnification instability must be minimized to achieve long-term repeatability and accuracy. Other sources of error in image formation and linewidth measurement have been outlined by Jensen and Swyt<sup>53</sup> and Postek<sup>12,34,46,67,68</sup> and must be corrected before the SEM can make accurate, reproducible measurements.

### 9.6.1.2 Particle Metrology

Particle metrology and characterization is now a growing field. Particles are a bane of semiconductor processing,<sup>69</sup> and many are induced by the processing as well as by the inspection process. The SEM has numerous moving parts. Each can generate particles through wear mechanisms. As the wafer is transferred into and out of the system, particles can be generated from contact with the wafer transfer machinery. Movement of the wafer into and out of the vacuum causes some degree of turbulence, which can mobilize particles, possibly depositing them on the wafer surface. Detecting these particles is especially important in keeping the production yield high. Various particles can also be used for many purposes where the size and size distribution of the particles are important parameters for the measurements and processes. Colloidal gold particles for example, have many



**FIGURE 9.29** Scanning electron microscope (SEM) images of 80 nm size colloidal gold particles. The field-of-view is 429 nm.

uses ranging from glass coloring to biomedical applications. High-resolution SEMs can be very effective instruments to quantify these and other similar particles (Figure 9.29).

### 9.6.2 VOLTAGE CONTRAST

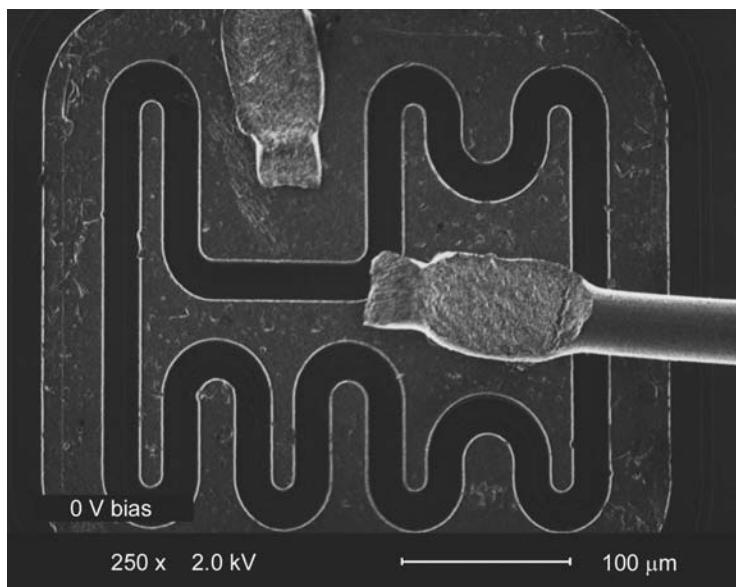
Voltage contrast in the SEM currently serves in routine semiconductor quality control and failure analysis. This application of the SEM is also undergoing continual development as a useful probing and measurement technique in circuit development.

The increasing miniaturization of IC components demands nanometer resolution techniques to verify the geometry and thus the speed, quality, and reliability of the circuits. Typically, the analysis of many ICs is accomplished through mechanical probing. However, as circuits become smaller, mechanical probes present difficulties in alignment and can easily damage metalized lines. Mechanical probes can also introduce capacitive loads to the device under test, possibly resulting in internal damage or compromising the operation of the circuit.

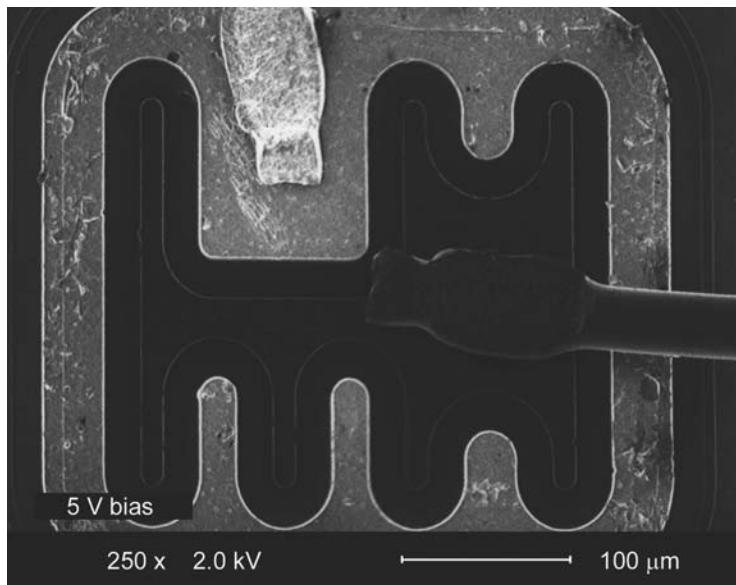
Used nondestructively in the voltage contrast mode, the electron beam of the SEM can electrically test circuits, permit the measurement of voltages, and allow for the creation of voltage distribution micrographs. As this technique became more sophisticated, two basic categories of voltage contrast have evolved: qualitative voltage contrast or the visualization of different voltages of an IC and quantitative voltage contrast that permits the precise measurement of voltages from an IC (requiring additional, sophisticated equipment).

#### 9.6.2.1 Qualitative Voltage Contrast

Qualitative voltage contrast relies on the existence of retarding electric fields above positively biased parts of the sample.<sup>70</sup> A crude form of energy selection of emitted SEs occurs based on the geometry of the specimen, the collector, and the applied voltages. Sections of the device receiving



**FIGURE 9.30** Si transistor with no bias voltage. The field-of-view is 4.3  $\mu\text{m}$ .



**FIGURE 9.31** Secondary electron image of Si transistor (shown in Figure 9.30) with 5 V bias voltage. The field-of-view is 4.3  $\mu\text{m}$ .

a positive voltage appear dark, and areas of negative voltage appear bright; parts at zero voltage remain at the same gray level as viewed in a typical secondary micrograph (Figure 9.30).

Engineers studying a device can visually observe voltages and voltage distributions that were previously only mentally envisioned. In this manner, the voltage contrast technique can troubleshoot device fabrication line problems on partially processed wafers or analyze failed devices even if they are only partially operational.

Static voltage contrast is one of the simpler techniques carried out in an SEM (Figure 9.31). The only elements required are the device, a wired socket, electrical vacuum feedthrough, and

power supplies. The voltage distribution images and micrographs provide information about individual logic states, voltage drops on resistors, and shorts and open connections in the device. Non-destructive testing of metal oxide semiconductor (MOS) devices requires acceleration voltages of <2.5 kV for nonpassivated devices. This technique may also serve for the analysis of passivated devices, but they usually require a higher primary beam energy, which could result in damage to the device.

The distribution of DC voltages provides a static representation of a chip. To study real-time high-speed operating device functions requires connecting the chip to a clock generator or function generator. Observation of the chip at TV scanning rates while it is being pulsed at clock rates below the scan frame rate allows observation of the voltage contrast effect in each part of the circuit structure. Pulsing the chip at a specific harmonic of the scanning line frequency generates strips of contrast in the display. Further adjustments very close to this harmonic frequency permit the measurement of the beat frequency of the chip.

### 9.6.2.2 Quantitative Voltage Contrast

Quantitative voltage contrast employs a somewhat more complex test setup than for qualitative voltage contrast. Often a dedicated instrument is used for this technique. Through the application of the electron beam to the area of interest, this technique can accurately measure the voltage of an IC through the collection and analysis of the SEs extracted from the sample.<sup>70–73</sup>

Quantitative voltage contrast requires the use of high-speed beam blanking units on the SEM column along with an electron beam sampling system. The technique enables the measurement of waveforms at circuit nodes with low distributed capacitance in ICs and thus allows the comparison of the actual device function with the electrical value calculated by computer simulation during the development phase. Fixing the beam on a point of interest and sending the signal through a logic state analyzer or oscilloscope can result in real-time measurements of better than 0.1 V potential sensitivity and 10 nm spatial resolution.

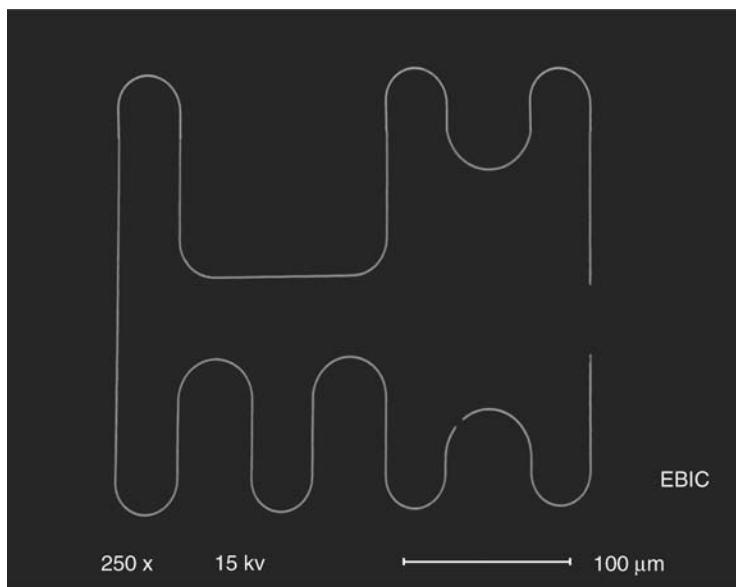
### 9.6.3 ELECTRON BEAM INDUCED CURRENT

The charge collection mode of the SEM has wide applications in the study of electrically active defects in semiconductor materials and devices. When an electron beam strikes a semiconductor specimen, it generates electron–hole pairs. If these mobile carriers reach or come from within the depletion region of a p–n junction, they are swept out by the existing junction potential, thus giving rise to an external reverse current. Inhomogeneities in the material alter the efficiency of the charge collection of this internal field. With the sample isolated from ground, collecting and amplifying the current provides information about the electrical nature of the junction or variations in the characteristics of the junction (Figure 9.32).

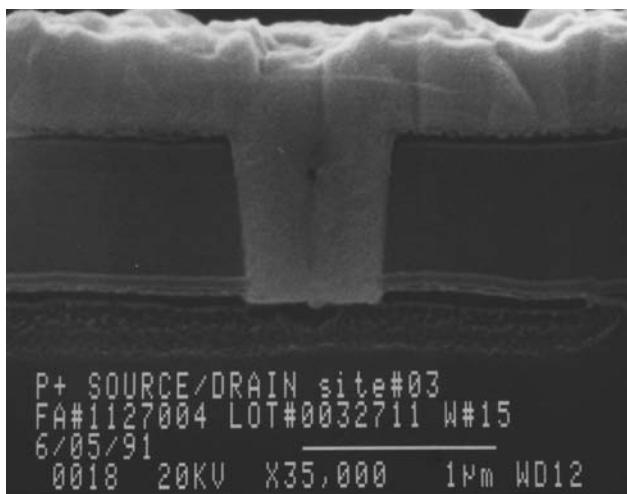
EBIC measurements are useful for the probing of semiconductor devices too small for mechanical probing. The EBIC method can locate leakage paths, defective p–n junctions, faulty diffusions, uncontacted junctions, shorted devices, electrically opened conductors, dopant inhomogeneities, or localized defects. However, knowledge of the device layout and the electrical parameters is essential for interpreting the EBIC analysis. The resolution obtained with this technique depends on the size of the volume excited by the electron beam; resolution in the order of 100 nm is commonly reported.<sup>74</sup>

### 9.6.4 FAILURE ANALYSIS

The SEM can also serve as a supplement to the light microscope with magnifications as low as five times and with a large depth-of-field. Using the SEM, engineers can observe and photograph IC interiors and exteriors more easily than with conventional light-optical systems. The SEM can rapidly resolve device failures since it does not suffer from illumination difficulties—such as specular



**FIGURE 9.32** Electron beam induced current image of a Si transistor (shown in Figure 9.30). The field-of-view is 4.3  $\mu\text{m}$ .



**FIGURE 9.33** SEM image of an integrated circuit cross section. The field-of-view is 3  $\mu\text{m}$ .

reflection—like those found in optical microscopes. Flexibility in the SEM goniometer stage permits tilting from  $-90$  to  $+90^\circ$ , continuous,  $360^\circ$  rotation, and X, Y, and Z motions during observation of the specimen at scanning or TV rates.

SEM failure analysis requires opening the device and properly attaching it to the stage. Wafers may be observed intact or cleaved and the fragments viewed in cross section for film profile evaluations (Figure 9.33). As failure analysis studies are often considered destructive, sample coating is often permitted but not always necessary; adjustment of the accelerating voltage can permit observation without sample coating. For passivated devices, removal of the protective glass coating is recommended. Evaluation of process control samples and failed devices in this manner

provides clues to how defects arise and often suggests subsequent corrective action. Areas of special interest are substrate contouring, metal processing, photomask alignment, and metallization failures.<sup>21,39</sup>

The use of the SEM in screening devices for potential defects is a required action for high-reliability semiconductor devices used in military and space applications. Screening procedures GSFC S 311 PI2A16<sup>63</sup> and MIL STD 883A17<sup>75</sup> reduce the risk of potential field failures due to slight step coverage defects, even though these devices may test electrically acceptable. With an SEM, the failure analyst can make use of all of the techniques involving the signals generated by electron beam-specimen interactions. Exploiting these signals individually or collectively permits the analyst to reach conclusions about the causes of and possible corrective actions for device failure. This information can lead to product improvement, increased yield, increased reliability, and more consistent quality.

### 9.6.5 VARIABLE VACUUM SCANNING ELECTRON MICROSCOPE

For many years, scanning electron microscopy has routinely been done exclusively at high vacuum in the specimen chamber but innovations in variable pressure SEM (VPSEM) have changed the rules. Following the seminal work of Danilatos<sup>9,17</sup> a rapid development of variable pressure instruments and electron detectors characterized the trend for years. These types of SEMs are now common, and it is expected that in the future the market share of high-vacuum-only instruments will further diminish, because they offer only a part of the possible working modes, and sample charging and contamination can be serious problems with these SEMs.

Low-vacuum microscopy was originally proposed in the early development of the SEM as a technique for the imaging of wet samples under close-to-natural conditions. It was recognized that the presence of low-pressure gas surrounding the sample reduced charging of insulating samples. Because of the level of technology and the type of electron detector employed at that time, these early low-vacuum instruments were only capable of producing an image using BSEs. This provided useful imaging at low magnifications and with good contrast from different materials. However, the high-magnification images were poor as compared with those obtained using high-vacuum technology.

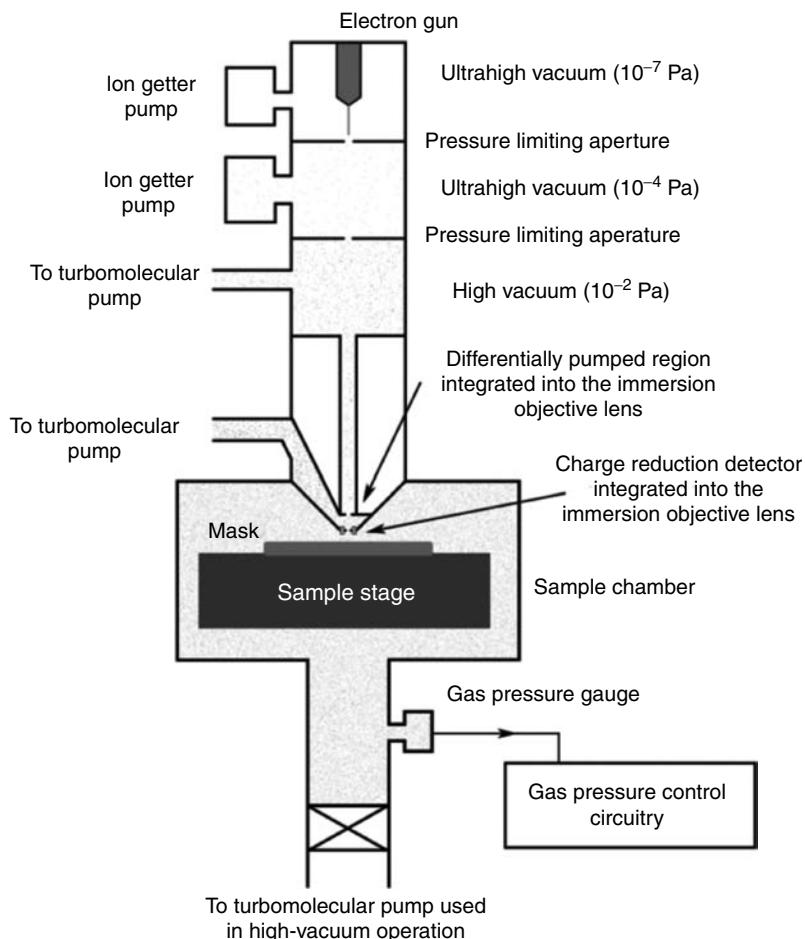
Research to understand the physics governing the imaging in low-vacuum conditions in the SEM was conducted by Danilatos and others. By the late 1980s, fundamental instrument technology and methods of imaging at low vacuum with not only BSEs but also SEs were developed.<sup>76</sup> This detection process allowed the high-resolution imaging of insulating samples at high beam landing energies without any form of conductive coating. A further advantage was that the high beam energies allowed x-ray analysis to be carried out on unprepared, insulating samples. It was also found that this technique proved to be less prone to electron beam induced contamination effects from hydrocarbons.

The advantages of this new SE detection method led to the development of a commercial instrument called the environmental scanning electron microscope™ (ESEM).<sup>77</sup> This instrument has been used mainly for imaging biological, food, chemical, pharmaceutical, and various materials type samples that are incompatible with high vacuum. With the ESEM, the gas atmosphere around the sample could be customized to produce *natural* conditions for the sample (e.g., fully hydrated biological material) or can be used as part of some reactive experiment (e.g., high-temperature corrosion). There is extensive literature on these applications, and some complete editions of journals have been dedicated to the subject.<sup>61</sup> In the mid-1990s researchers at Nikon Inc. attempted to apply ESEM technology to low accelerating voltage semiconductor manufacturing applications<sup>78</sup> but significant technical limitations of the technology at that time inhibited progress. Other designs of variable pressure or *nature* SEMs were built with limited vacuum ranges and signal detection typically restricted to BSEs.

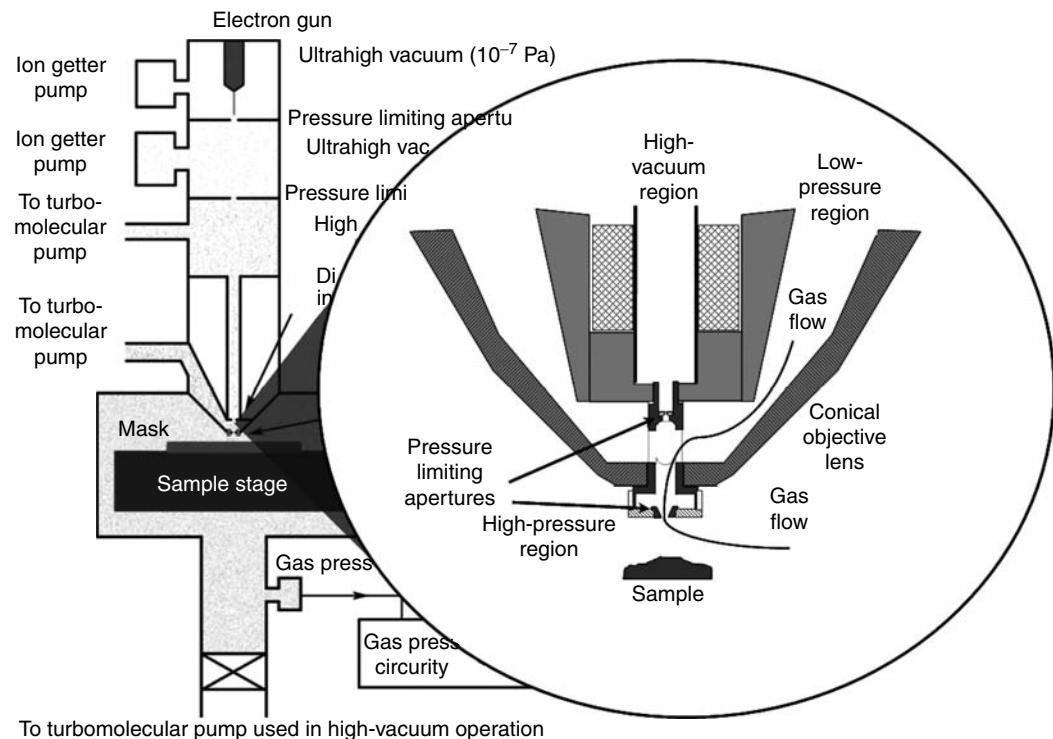
### 9.6.5.1 Variable Vacuum Scanning Electron Microscope Column

The VPSEM is unlike the typical high-vacuum SEM. As shown in Figures 9.34 and 9.35, the difference is mainly found in a column design that facilitates specialized vacuum technology. The configuration of the low-vacuum SEM column is about maintaining the electron gun and column at high vacuum and the specimen region at low-vacuum values while providing a transition region between them. The gas pressure in the low vacuum—depending on the type of SEM and detectors used—can go beyond 1000 Pa, but it is usually kept in the few ten to hundred pascals range. The gas and primary electron beam interactions depend on the gas pressure and electron energy result in an average electron path length, which should be sufficient to produce images. This is factored into the design of the instrument column. There are generally three pressure zones, connected by pressure limiting apertures of 100–200  $\mu\text{m}$  in diameter. These zones are the electron column zone, the differentially pumped zone, and the specimen chamber zone.

The electron column pressure is essentially maintained at high vacuum appropriate for the particular electron source and is generally in the low  $10^{-3}$  Pa region or better. The instruments equipped with field emission electron gun typically have more differential pumping zones so that the field emission tip can be kept at typically  $10^{-7}$  Pa vacuum levels.



**FIGURE 9.34** Schematic drawing of a variable vacuum scanning electron microscope (SEM).

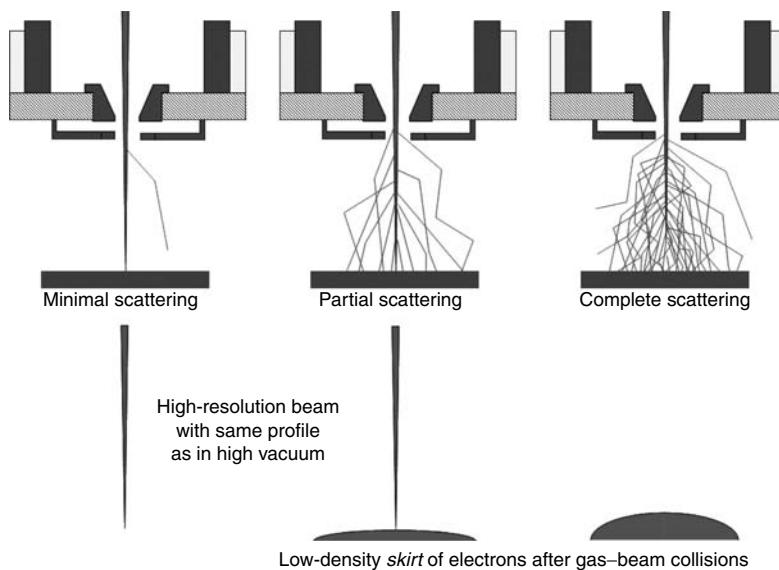


**FIGURE 9.35** Schematic drawing of a variable vacuum scanning electron microscope's (SEM's) objective lens area.

The differentially pumped zones are connected only by suitably sized apertures that allow certain amounts of gas or vapor molecules to get through. Pumps via appropriately designed pipes maintain the needed vacuum levels. These result in a gradual change in the vacuum levels in the vertical direction. For most of the electrons the electron path length is sufficient to reach the sample. The mean electron path length—depending on gas type and temperature—in high vacuum is several kilometers; at 100 Pa it is  $\sim 1$  cm and is close to 1 mm in 1000 Pa. The pressure in the specimen chamber zone is under the control of the operator and depending on need, can be varied from high-vacuum mode of operation to low-vacuum mode. In the low-pressure air or water vapor various biological samples survive for long periods of time under observation. Many biological samples do not lose much water and collapse during the pump-down of the specimen chamber.

#### 9.6.5.2 Effect of the Gas on the Primary Electron Beam

Incident, backscattered, and SEs interact with the gas and produce positive and negative ions. This interaction has both advantageous and disadvantageous effects. In the presence of an external field this interaction can produce a fast ion avalanche (a cascade of electron-ion pair generation), which in the presence of an electromagnetic field can be used for signal amplification. The amount of gas needed for the efficient amplification is enough to cause elastic and inelastic collisions between the primary electron beam and the gas. This interaction results in the nonnegligible scattering of the primary electrons. Because of this, some of the primary electron beam electrons will strike the sample at a considerable distance from the intended beam impact (focal) point. The resulting beam profile is illustrated in Figure 9.36. As shown in the figure, the beam is essentially split into two components. There is a central high-resolution component that has the same profile as the beam in



**FIGURE 9.36** Scattering of the primary electron beam in low-pressure gas (top) and the intensity distribution of the electrons that reach the sample (bottom).

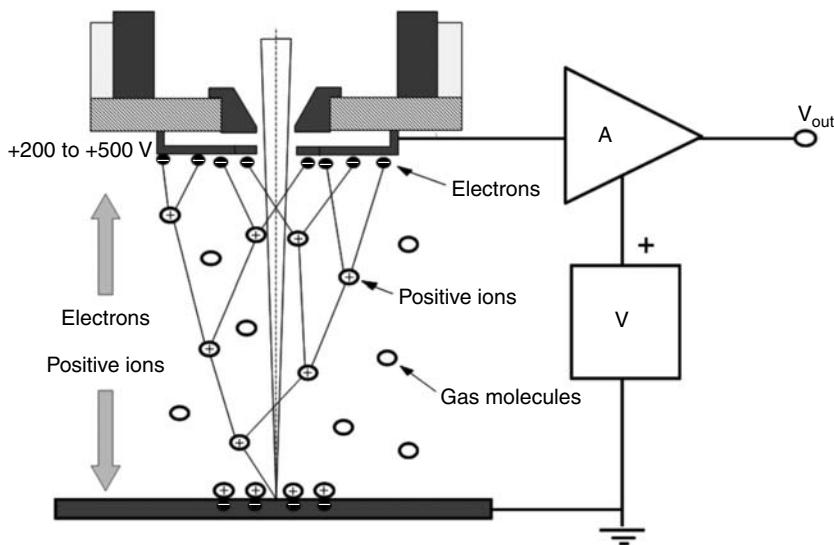
high vacuum and generates a high-resolution image. The other part of beam is scattered over a very broad area (called the *skirt*) that may be 100  $\mu\text{m}$  or more in diameter. This skirt has been characterized by various authors.<sup>15,79,80</sup>

The scattered electrons of the skirt generate an out-of-focus background, essentially a relatively even gray level offset in the image signal. This offset can easily be removed. However, the major effect of the skirt on the image is to degrade the signal-to-noise ratio, particularly at low beam landing energies. Another significant effect is that the skirt electrons generate x-rays that are coming from the overall skirt area and not the exact area being imaged. The percentage of the beam that is split into the skirt varies with many parameters (e.g., beam energy, gas material, pressure, temperature, and gas path length) but can be reduced to a few percent at high beam energies by careful design of the instrument and suitable selection of operating parameters.

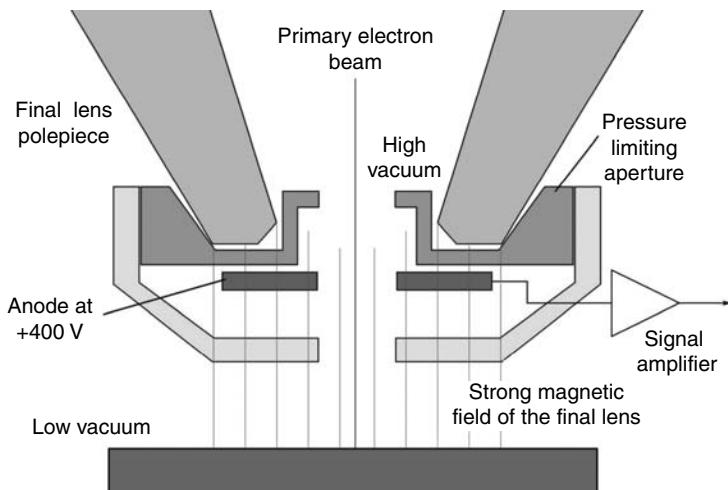
#### 9.6.5.3 Secondary Electron Signal Detection in a Gas

The imaging methods developed by Farley and Shah, and Mancuso et al.<sup>81,82</sup> use similar ideas of detecting and amplifying the SE signal in the gas. The general structure of this detector is shown in Figure 9.37. An electrode (detector plate) with a central hole is placed between the sample and electron column.

The electron beam passes through the hole and generates signal-carrying electrons at the beam impact point. As with any SEM, the electrons generated by the primary beam are both SEs and BSEs. The relatively high positive potential on the electrode strongly attracts the SEs and accelerates them toward the detector plate. Depending on the energy of the BSEs, they may also be collected if their trajectory leads them to impinge upon the electrode/detector plate. In this mode of operation, any of the incident or exiting electrons with sufficient energy (typically  $>20$  eV) have the potential to ionize any gas molecules they encounter as they are traveling toward the sample and accelerated toward the electrode. These ionization events generate both positive ions and more signal-carrying electrons, thus, signal amplification occurs. The electrons are then further accelerated, and potentially undergo more collisions. A controlled *cascade* of electrons is thus generated from



**FIGURE 9.37** Gaseous secondary electron detector.



**FIGURE 9.38** Helix-type gaseous secondary electron detector.

each initial signal electron. A more detailed explanation of the cascade process can be found in Thiel et al.<sup>80</sup> Eventually, the high-voltage electrode collects all the signal-carrying electrons as well as *background* electrons. The *gas cascade cell* between the sample and a positively biased electrode is fundamental to SE signal detection in this design of the high-pressure SEM. A low-noise, high-speed, high-gain amplifier is then employed to amplify the collected signal.

Alternatively, the light produced by the ionization of the gas may be detected directly using a photomultiplier. With this type of detection also, the BSE signal can be collected either independently or in concert with the SE signal. In some instances the detector/electrode may be placed in other positions that are not symmetrical about the beam, but the cascade and detection process are the same. The gas cascade cell is usually in the range 10–40 mm in size.

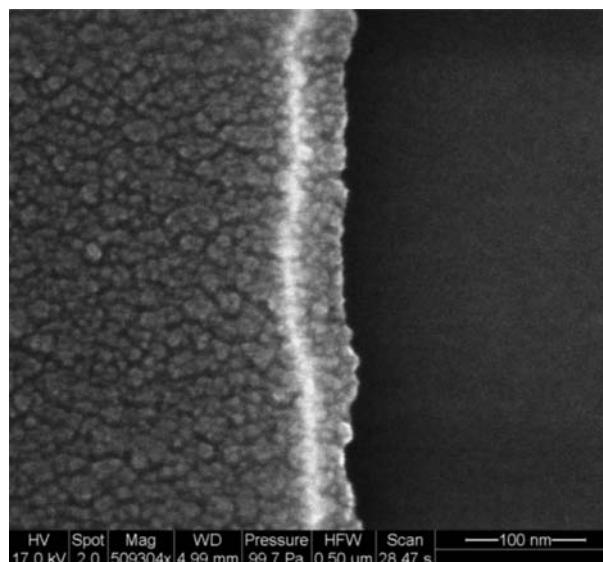
A new detector, called Helix, is allowing for high gains and excellent collection efficiency with short working distances (Thiel et al.). Figure 9.38 schematically shows the Helix type gaseous SE detector.

On the primary electron beam axis the electric and magnetic fields are parallel, but close to the anode the electric field is orthogonal to the magnetic field. These fields confine SEs to axial and radial oscillations within the detector. The SEs leaving the sample at a low energy start moving upward in a spiraling motion, similar to the case in in-lens detection. In this case as they enter the region of the detector, they are forced to follow a long, spiraling path not just in the vertical but also in the horizontal direction before they eventually hit the anode and get detected. Along their traveled path they loose all of their kinetic energy through inelastic collisions with gas molecules, which is essentially the gas amplification process. The size of this detector is small, and it is located at a short working distance; the vertical and horizontal spiraling motions of the SEs create a long electron gas amplification length, which in turn allows for ultrahigh-resolution imaging even at low electron beam energies. Similar to other gaseous detectors, charging of insulating specimens can be stabilized by positive ions produced in the gas ionization cascade. The bandwidth and signal-to-background level of this detector are superior to those of other types of low-vacuum SE detectors. Another advantage is that the SE detection is annular (central to the primary electron beam). All these make this detector superior to others in several regards and especially advantageous for the semiconductor metrology industry in CD measurements on photolithographic masks.

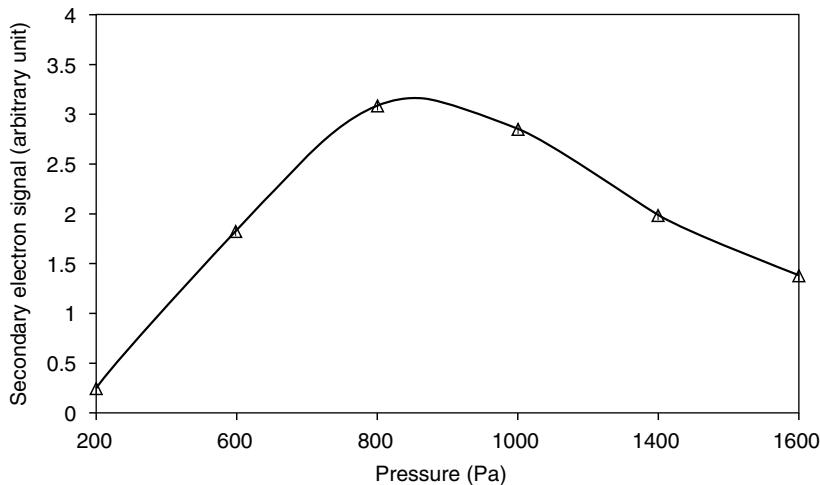
Figure 9.39 shows an SE image of a photomask. This mask is built with 70 nm Cr structures on a 6 mm thick quartz substrate. Despite the high kilovolts landing energy, the charging is well controlled, and the high resolution reveals the grain structure of the Cr layer.

#### 9.6.5.4 Gas Pressure and Path Optimization

Work by Thiel et al. has shown that there are optimum values of gas pressure for a given electromagnetic field strength. If the pressure is too high, then the electrons gain insufficient energy between collisions to cause ionization. If the pressure is too low, then the electrons move too far in the gas before collision and the number of amplification *stages* is reduced. A typical curve of this variation is shown in Figure 9.40. In this case, the gas used was water vapor, because it has good ionization characteristics, but a wide range of gases such as helium, argon, and nitrogen could also be used. Some of the gases are more easily ionized and thus can reduce the charging more effectively, but at the highest pressures there is no significant difference due to the choice of gases.



**FIGURE 9.39** Secondary electron image of a photomask. The field-of-view is 500 nm. (Courtesy FEI Co.)



**FIGURE 9.40** Gas pressure versus secondary electron signal.

### 9.6.5.5 Charge Control

The *cascade* process that amplifies the SE signal, as stated previously, also generates positive gas ions that drift to negatively charged regions of the sample surface under the influence of the field from the electrode. These ions mitigate the sample charging caused by the primary beam interaction with insulating samples. They also compensate the negative charge. In fact, there are typically many more ions generated by the cascade process than are required to counteract sample charging due to electron irradiation. An overabundance of ions can cause the sample surface to acquire a positive potential.<sup>79,83</sup> Fortunately, this positive potential is spread over a large area and does not cause the typical charging effects that are common in high-vacuum SEM operation (bright or dark regions and significant primary beam deflection). The main effect of the surface potential is to reduce the field strength of the gas cascade cell electrode, which then reduces the amplification in the gas. Thus, a natural balancing occurs. On a small sample the effect is minimal because of the proximity of the grounded sample holder. One solution is to place a grounded conductor close above the sample to absorb the excess ions.<sup>13,84</sup> This effect needs to be explored further for the large photo-mask and semiconductor samples.

### 9.6.5.6 Contamination Reduction

The charge control of the cascade process changes the local electromagnetic field and results in very mild plasma, which alters the beam chemistry that otherwise may cause hydrocarbons to adhere to the sample surface at the irradiated area. This effect greatly reduces the deposition of carbon contamination on the scanned areas, even at high magnifications and beam intensities, where the dose of the electrons is high. It is not uncommon to observe actual cleaning of the sample surface, which advantageously results in higher SE yield.

## 9.7 SCANNING ELECTRON MICROSCOPE IMAGING AND ACCURATE MEASUREMENTS

### 9.7.1 RESOLUTION MEASUREMENTS FOR THE SCANNING ELECTRON MICROSCOPE

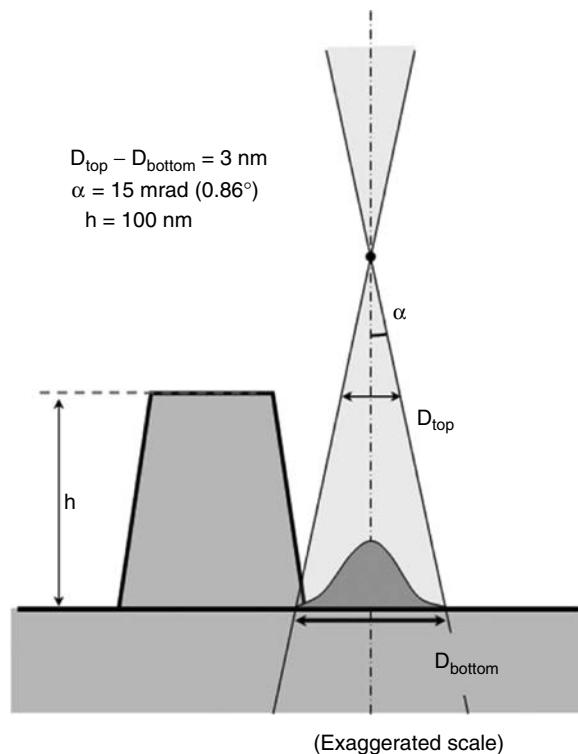
SEMs are primarily used for generating images and for x-ray microanalysis, but specialized instruments are used for high-speed dimensional measurements in production of ICs. SEM images appear

to be realistic, rich in details, and might provide a good depth-of-field, so many times people have some sense of three-dimensional (3-D) representation of the artifacts. The results can be obtained relatively easily and quickly on large and diverse sets of samples and present a lot of useful information, which are excellent for qualitative work.

At the same time SEMs are far from perfect for measurements. Quantitative work can be very difficult for charging, contamination, and other reasons. Truly high spatial resolution, high degree of repeatability, and accuracy are difficult to achieve and maintain. Also, many parameters needed for highly accurate quantitative measurements are unknown or unaccounted for. Much better control over the performance of the SEM is needed and possible. For highly accurate measurements all important parameters of the measuring system have to be known and kept under good control. These include the electron beam with its shape, size, and electron intensity distribution, signal-to-noise ratio, and the linearity of the X- and Y-direction scans and the signal chain. For correct quantitative interpretation and analysis of the results an accurate model of the signal generation is also necessary.

The electron probe has a 3-D distribution of electrons along and across the electron beam. It can be characterized by an asymmetric hourglass-like, double-conical shape, by beam convergence angles with the best focus at the disk of least confusion, etc. Figure 9.41 shows an idealized representation of the electron beam.

The SEM signal is a complex product of the interaction of the electron probe, the sample geometry and material, and the electromagnetic field present around the sample either from the SEM itself or from sample biasing and charging. Two 3-D entities, electron probe, and the sample define the 3-D information volume, which is that part of the 3-D excited volume where the SEM



**FIGURE 9.41** Idealized electron beam with double-conical shape and Gaussian electron intensity distribution across the beam. In this case there is a 3 nm difference in the beam diameter at the top and at the bottom of the 100 nm structure.

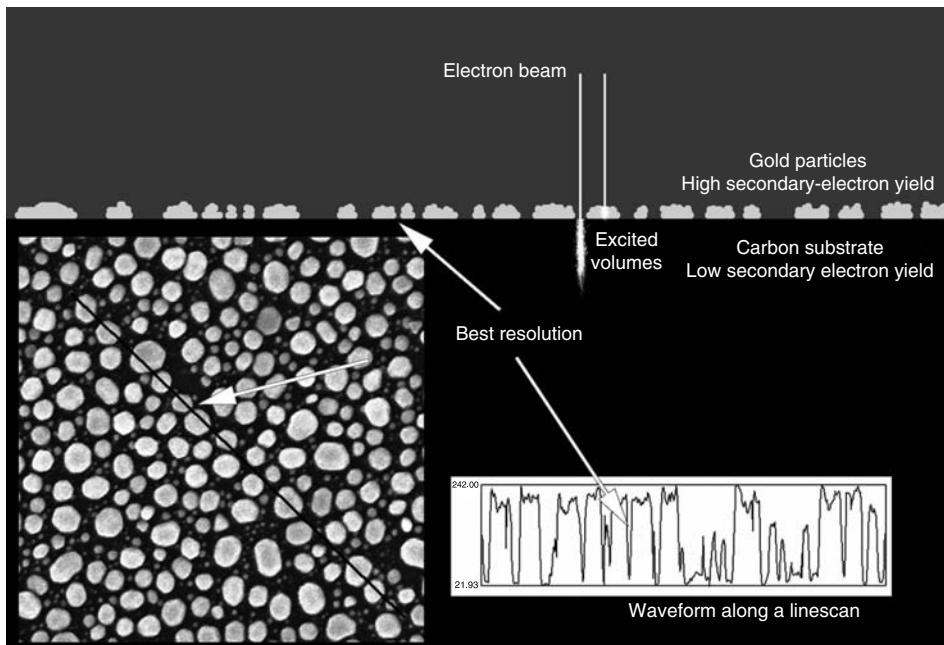
signals come from. The excited volume is approximately the convolution of the shape of the electron beam and the shape of the sample. In general, to be able to accurately measure the shape and size of the sample, one has to know and use in the measurements all significant parameters of the electron beam. This can be done with excellent fidelity by using sophisticated modeling methods.

### 9.7.2 RESOLUTION OF THE SCANNING ELECTRON MICROSCOPE IMAGE

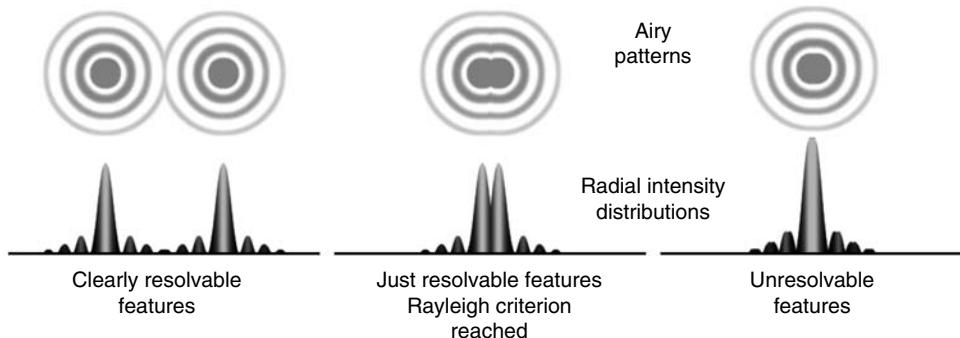
In practice one of the most important SEM parameters is the spatial resolution of the image (see Chapter 8). All SEMs are characterized by some kind of best attainable resolution, which is one of the key parameters for instrument acceptance after installation and for tracking the performance of the instrument. Yet, it is one of most elusive parameters of the SEM. The measured *spatial resolution of the SEM* is a highly simplified account for the performance of the complex imaging system and imaging process of the SEM. Figure 9.42 shows the traditional gap-based resolution measurement. Usually gold-on-carbon or gold-evaporated magnetic tape samples are used for this measurement and once the magnification or scale is correctly calibrated at a suitably high magnification, an image is taken, and finally, the width of the smallest gap is found and presented as the best resolution of the SEM.

This type of resolution measurement is only moderately objective, partly because it strongly depends on focusing and astigmatism, which is difficult to set exactly at truly high resolution, and the noise is not taken into account. This measurement is more or less following an old, limitedly useful method, which was developed for optical microscopes by Rayleigh.<sup>85</sup> Figure 9.43 shows the three important distinctive resolution cases with their corresponding Airy disks and radial intensity distributions.

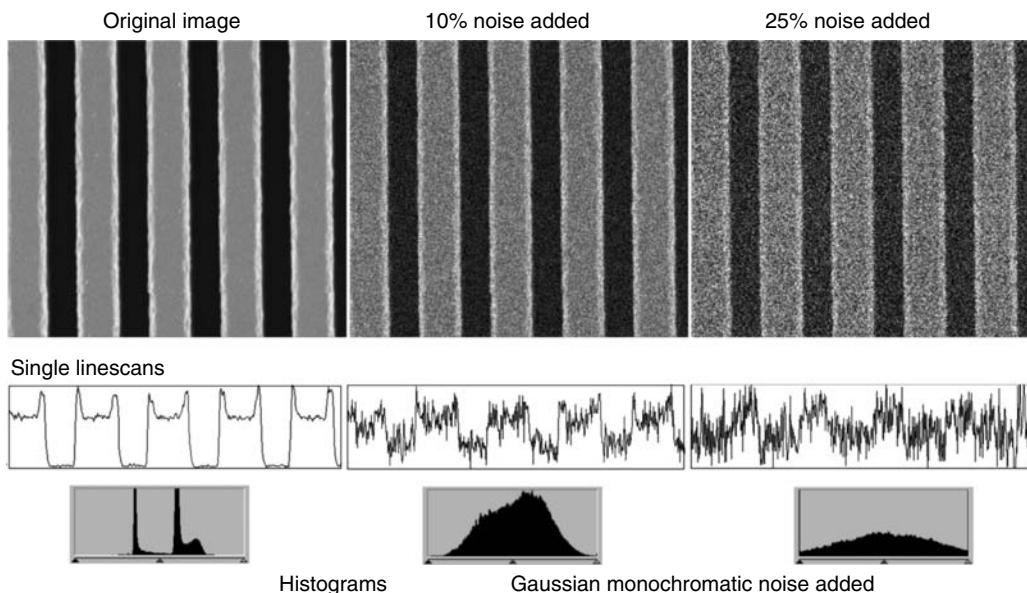
In the case of an optical system that can be characterized by the left diagram, two features close to each other can be clearly seen apart. If two particles get closer (center diagram) and fulfill



**FIGURE 9.42** The traditional gap-based resolution measurement on gold-on-carbon sample. The field-of-view of the insert is 1  $\mu\text{m}$ .



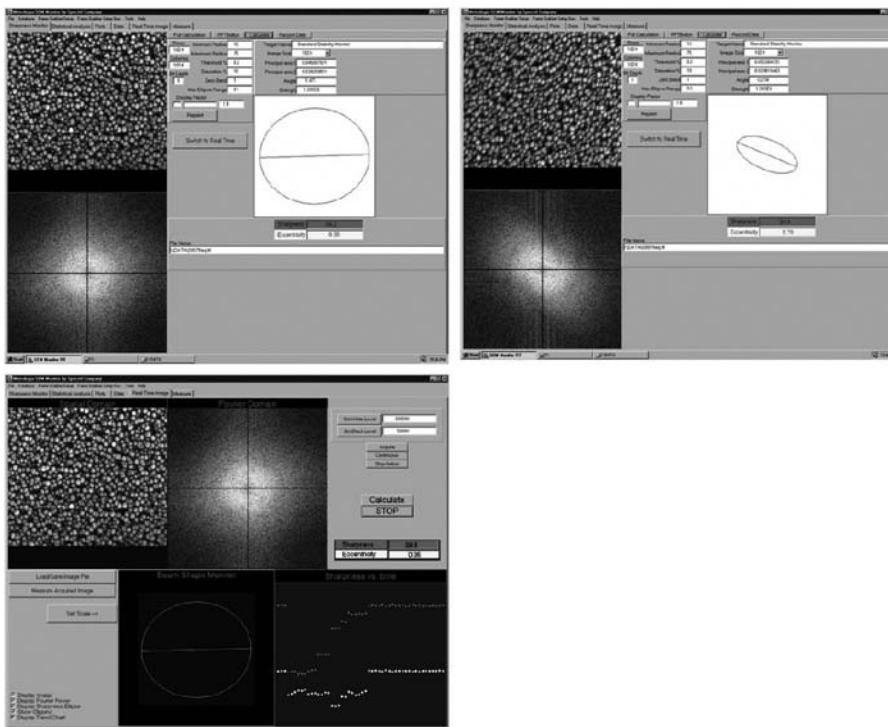
**FIGURE 9.43** The three important distinctive resolution cases with their corresponding Airy disks and radial intensity distributions.



**FIGURE 9.44** Degradation of image quality caused by increasing noise. Integrated circuit (IC) resist lines. The field-of-view is 2  $\mu\text{m}$ .

the Rayleigh criterion, they still can be distinguished, but if they get any closer the distinction is no longer possible. One will see them as one large particle with diagram on the right of Figure 9.43. The Rayleigh criterion underestimates the attainable resolution, and it does not take the noise into account. Clearly, a better, more sophisticated approach is needed, especially for SEMs that compete at  $\sim 1 \text{ nm}$  resolution.

Noise limits the information contained by and retrievable from an image. For this reason an exact resolution measurement must account for it. Figure 9.44 illustrates the degradation of the possibility to separate features from the background and loss of details caused by noise. On the left image the IC resist lines are clearly visible, even small details are resolved. The histogram (distribution of the frequencies of various gray level pixels across the image) shows two well-separable peaks and the single linescan is well formed. These allow for reliable linewidth measurements. As the noise increases (center and left), first smaller and then larger details are no longer discernible, and the histogram and the linescans do not provide ways for easy separation any more.



**FIGURE 9.45** Scanning electron microscope monitor system measuring image sharpness and eccentricity (astigmatism) values on correctly set (upper left) stigmatic (upper right) images and helping the operator in real time and optimizing the SEM (lower left).

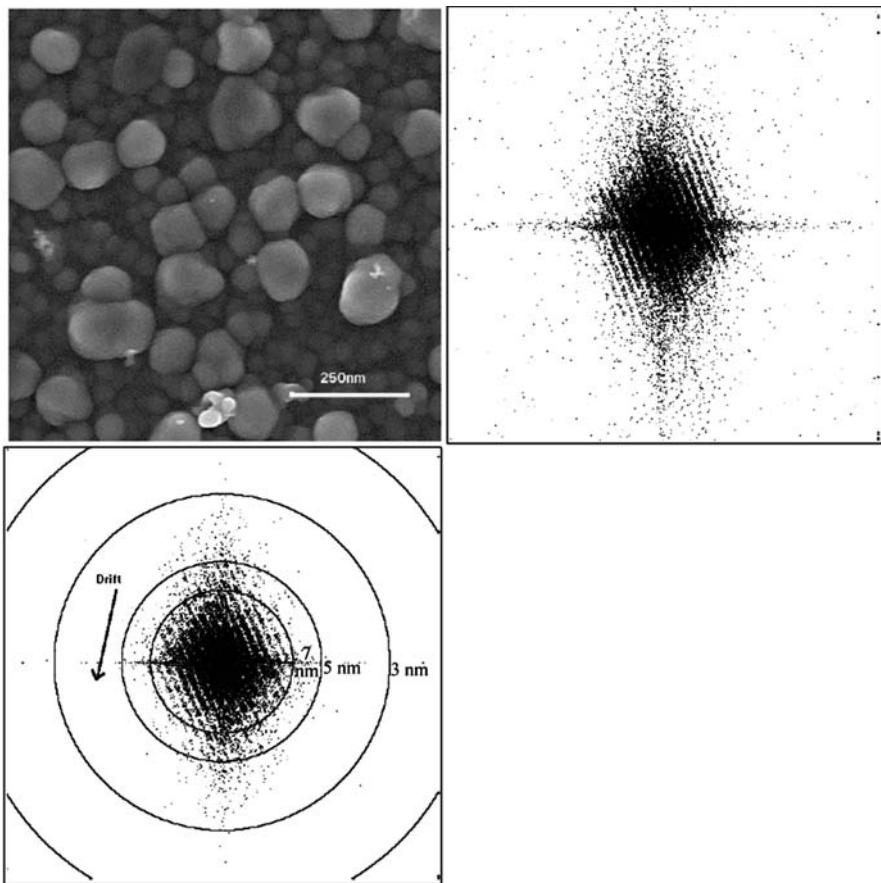
A single number as the resolution of the SEM is valid only for cases where the fully 3-D information volume formed by the sample and the electron beam is very small (negligible). Clearly, with high-performance SEMs that can readily focus the electrons into a  $<1$  nm spot, this is never the case; so this resolution number cannot be accurate. Nevertheless, it can be used for tracking the resolution performance of the SEM by using sophisticated algorithms and a suitable sample. If the sample and measuring conditions remain the same, these methods report a consistent number related to the focusing ability of the electron optical column.

Various efforts were aimed at the development of such methods. One is based on calculations in the Fourier domain. Some of these use two images and cross-correlation calculations to help the distinction of signal from noise. SEM monitor system, developed by Spectel Research,<sup>86</sup> is capable of measuring image sharpness and eccentricity (astigmatism) values in real time and help the operator optimize these parameters of the SEM consistently (Figure 9.45). This is especially important for process control purposes in the semiconductor industry.

In this case a NIST RM 8091 SEM sharpness measurement sample was used. This has small 10–30 nm size Si structures (*Si grass*) on the Si substrate.<sup>87</sup>

Another similar method called SMART was developed by D. Joy, which also works with frequency domain calculations and reports a resolution number (Figure 9.46). The SMART algorithm works as a macro with NIH and Scion Image, and it is available for free from the Internet.<sup>88</sup>

Resolution calculation methods that work in the spatial domain, for example, FEI Image software could also be very useful (see Figure 9.47). This software measures the transitions of many thousands of edges to arrive at an overall resolution value. It also reports a general beam shape,

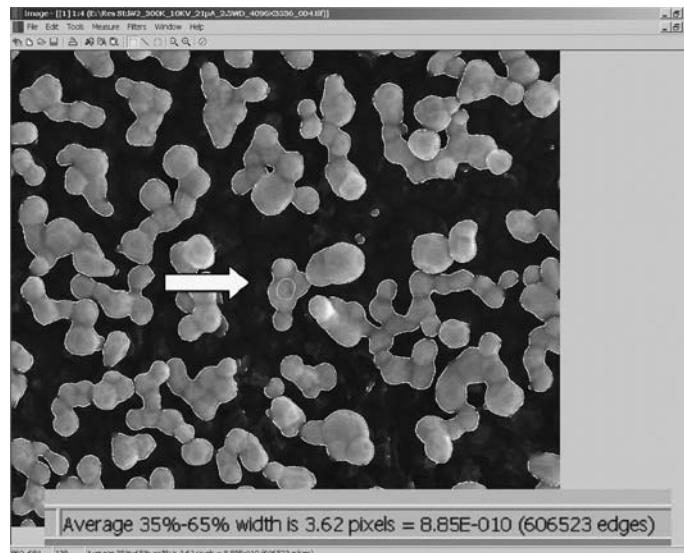


**FIGURE 9.46** Si oxide image (upper left) in SMART image resolution algorithm that calculates a cross-correlation diffractogram (upper right). A superimposed scale (lower left) helps to determine the 8 nm resolution in this case. The image offset was horizontal but the fringes are seen to be rotated by  $\sim 25^\circ$ .

which is preferably a circle in cross section as shown in the middle of the figure (marked with an arrow). This software also reports a signal-to-noise ratio, image quality, and clipping information which are useful parameters in comparisons and tracking of the performance of the SEM.

### 9.7.3 ELECTRON BEAM LANDING POSITION

Image formation in the SEM relies on sequential and synchronized recording of the image. The electron beam is scanned over the sample with signals generated by analog or digital X- and Y-direction scan generator and electromagnetic lenses or deflection plates that force the electron beam to follow a certain pattern. The signal recording is also controlled by the same signals so that a two-dimensional intensity distribution is recorded. This distribution can be viewed as an image. As long as the beam exactly follows the desired pattern (customarily starting at the upper left corner and line by line going from left to right) and the sample is motionless, the process results in an ideal image. This is never the case, because the landing position of the electron beam—beyond the intended X- and Y-direction motions—also changes unintentionally during the time of imaging. This happens because there are unwanted and inevitable motions of the electron beam caused by disturbing, nonstationary electromagnetic fields around the SEM and by imperfections of the electronics of the SEM, and also because of the mechanical vibration and drift of the sample stage. This



**FIGURE 9.47** (See color insert following page 340.) The 0.9 nm spatial resolution performance of the environmental scanning electron microscope (ESEM). Note the good, circular beam shape in the center of the image. The field-of-view of the gold-on-carbon sample is 200 nm.

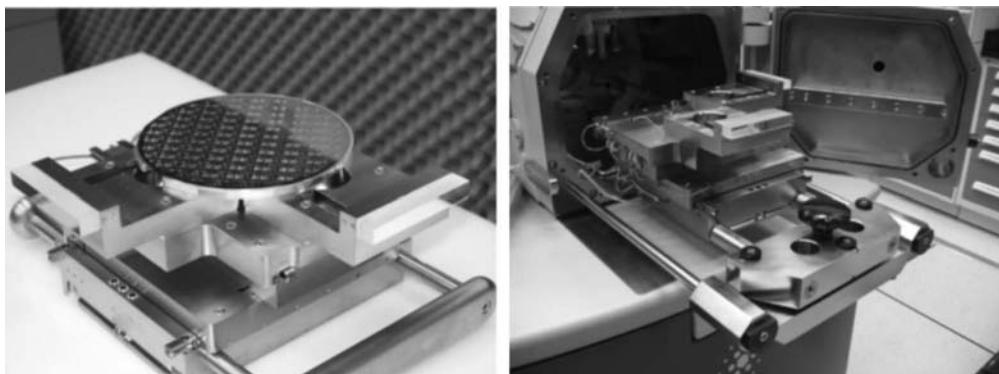
is one of the most bothersome problems of SEM imaging and metrology and a great deal of design work, and carefully chosen and maintained environmental conditions are needed to be able to take advantage of the focusing ability of the electron optical column. Essentially all SEM imaging takes place at a moving target with an imperfectly aimed electron beam.

In the case of high-resolution imaging with the best SEMs currently commercially available the landing position must happen with subnanometer accuracy. It is very difficult to fulfill this demanding requirement, but it is surmountable. Generally, it is easier to make sure that the electron beam follows the desired pattern, but it is much more difficult to deal with the problems of the unwanted motions of the sample stage. With small samples it is possible to use special SEMs that have the sample within the focusing lens and operate the SEM similarly to STEMs. For large samples special sample stages equipped with high-accuracy laser interferometer can be used.

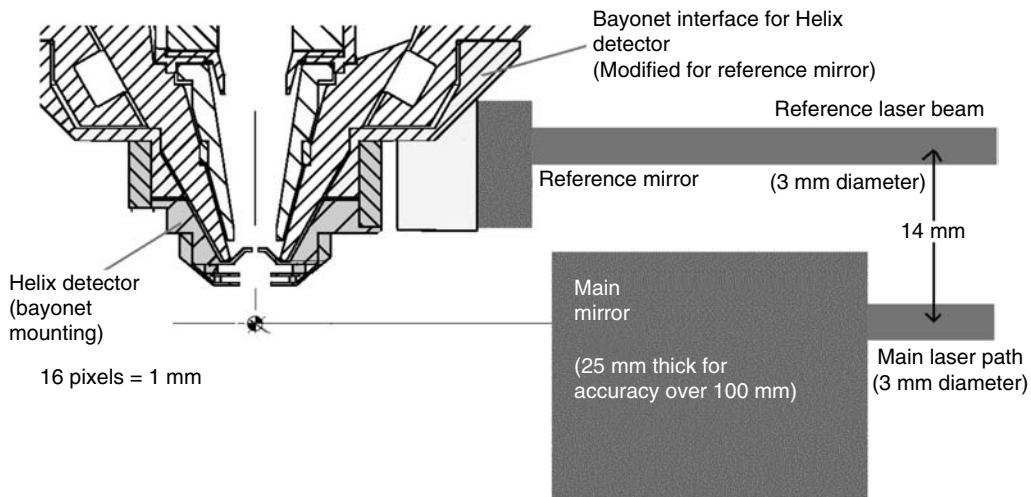
NIST is currently developing this technique with a laser interferometer sample stage (Figure 9.48) that can operate with 200 (as shown on the left-hand side of Figure 9.48) and 300 mm wafers and 6 and 9 in. photomasks. The laser interferometer sample stage (right-hand side of Figure 9.48) was designed to be exchanged easily with the original sample stage of the SEM. This is facilitated by a high-precision mechanism, which ensures a three-point support and excellent prealignment for the X- and Y-direction lasers. The interferometers and the accompanying electronics have 38 pm laser position resolution and a 10 MHz readout rate.

The beam path of the laser interferometer sample stage was designed to minimize Abbe errors. The reference mirrors are securely bolted to the final (objective or focusing) lens, and they do not degrade the performance of the Helix electron detector, which is used in low-vacuum mode. The system is optimized at 4 mm working distance, in which short working distance allows for excellent spatial resolution (Figure 9.49).

Figure 9.50 shows the most important components of the laser interferometer. The Renishaw make system works with fiber optical beam delivery, which gives good flexibility in the design and assembly of the whole measurement system. The laser interferometer head has multiple beams that beyond displacement in the X- or Y-direction also measure pitch and yaw in both directions. The electronics is a homodyne system that has an array detector as its sensor.



**FIGURE 9.48** (See color insert following page 340.) Laser interferometer sample stage with a 200 mm wafer (left). Laser interferometer sample stage before insertion into the environmental scanning electron microscope (ESEM) (right).



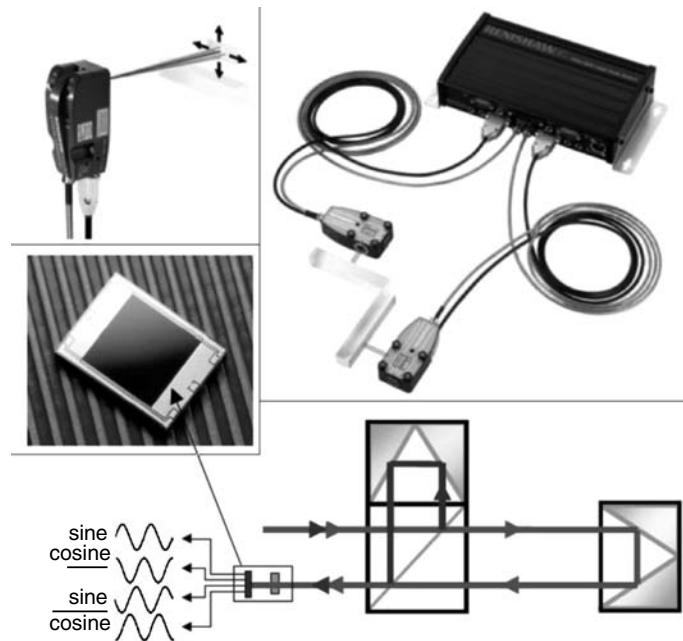
**FIGURE 9.49** (See color insert following page 340.) Schematic drawing of the beam path of the laser interferometer sample stage. The system is optimized at 4 mm working distance.

Figure 9.51 illustrates the motion of the sample stage followed by a high-resolution laser interferometer.

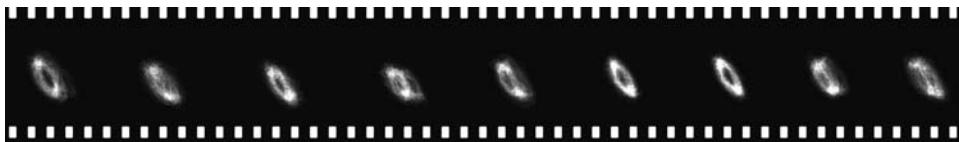
Both slow, low-frequency stage motions (*stage drift*) and higher ones (*vibration*) follow a perturbed pendulum motion, that is, the stage displacement eventually turns back (Figure 9.51). It is evident that even the stage drift that appears to move the image in only one direction follows this type of pattern, because the stage never hits the wall of the sample chamber. The complex sample stage motion that varies rapidly during the collection of the image is one of the reasons for the limited usefulness of simple deconvolution techniques in SEM image sharpening.

#### 9.7.4 ELECTRON BEAM INDUCED CONTAMINATION

Beyond the spatial resolution or focusing ability of the system it is also essential to avoid electron beam induced contamination, which until now is one of the most bothersome problems of the SEM. Even in clean vacuum instrument working with oil-free low-vacuum and turbomolecular high-vacuum pumps a polymerized hydrocarbon layer with low SE yield can be deposited. This



**FIGURE 9.50** Key components of the laser interferometer: interferometer head with pitch and yaw capability (upper left), optical beam delivery (upper right), array detector (center left), and beam schematic laser beam path (bottom).

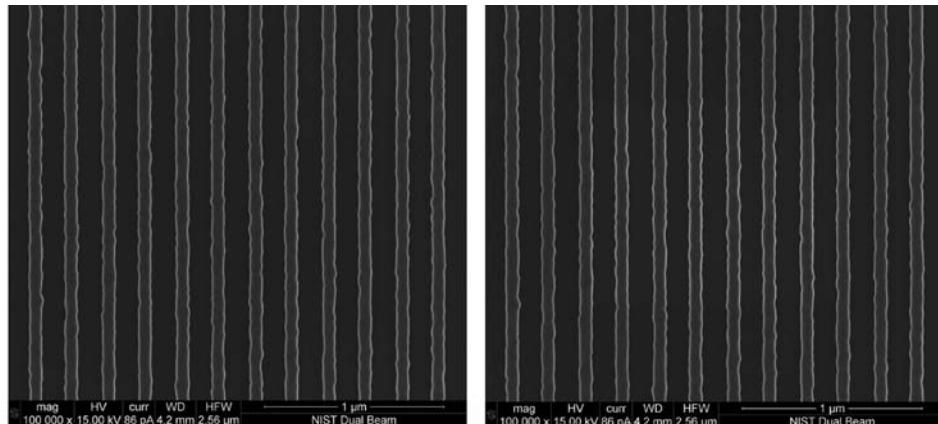


**FIGURE 9.51** Motion of the sample stage during a 60 s slow-scan imaging followed by a high-resolution laser interferometer.

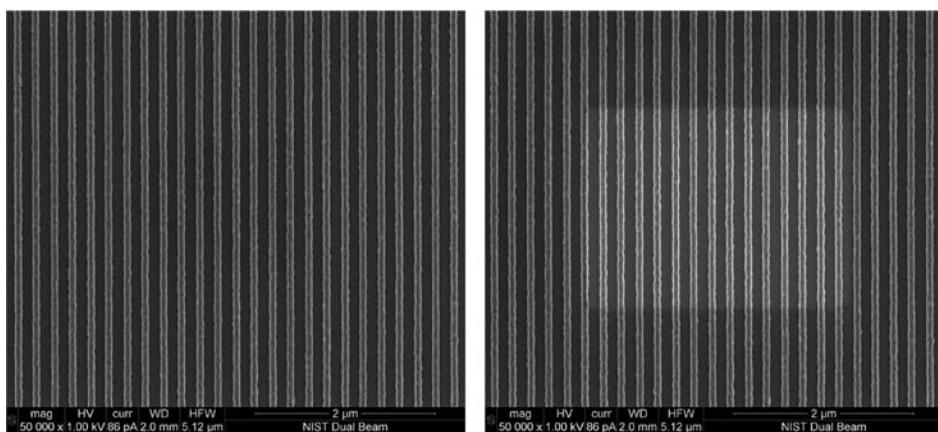
can change the size and shape of the small structures of current and future state-of-the-art ICs and nanotechnology samples, and produces erroneous, misleading results on almost all other samples. To remedy this problem, a new procedure was successfully developed at NIST. This allows for essentially contamination-free investigation of various nanotechnology and semiconductor technology samples. This is very important for wafer, photomask, and nanomanufacturing samples as well, where the structures to measure are very small, literally a few or a few tens of nanometers; even a small amount of deposited material can significantly alter the results and cause large measurement errors. The NIST cleaning procedures and a viable contamination specification offer an effective solution for this problem. This contamination specification calls for the following steps: on amorphous Si patterns on a Si wafer, using the best resolution imaging parameters, take one image at 100,000 times magnification. Save the image, and go up to 200,000 times magnification and continuously image for 10 min, and finally, go back to 100,000 times magnification and take another image. If there is any visible darkening, frame, any structure beyond the sample itself in the middle of the last image, the instrument fails to meet this specification. Figures 9.52 and 9.53 illustrate the

obtainable performance for high- and low-accelerating voltages, respectively. For both, images on the left-hand side were taken before the 10 min continuous electron irradiation, and the right ones were taken after. There is now a sign of the common darkening due to carbonaceous contamination formed under the electron beam. The irradiated locations (in the center) apparently got even cleaner, especially at low accelerating voltage irradiation. This is probably due to the removal of the molecular layer of water and carbon dioxide, which are naturally present on any surface exposed to air.

If the specification is not met, first cleaning the sample in the mixture of 3:1 concentrated sulfuric acid to 30% hydrogen peroxide solution (acid piranha solution) takes place. This ferocious oxidizer will clean all hydrocarbon residues from the sample in <30 min. If the instrument—with the clean sample—fails the test again, it needs to get cleaned with an effective low-energy plasma cleaning process. For this XIE Scientific Evactron<sup>89</sup> was used successfully.



**FIGURE 9.52** High accelerating voltage images taken on a Si wafer sample before (left) and after (right) 10 min continuous electron irradiation. Note the absence of common darkening due to contamination.



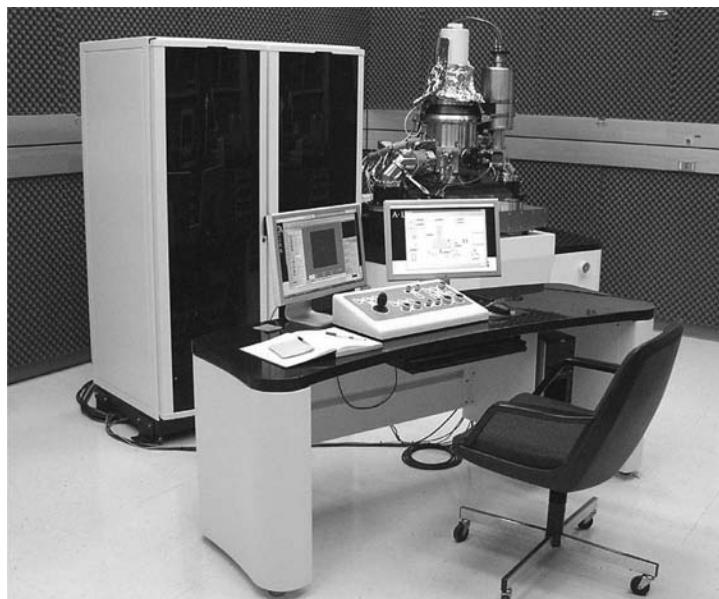
**FIGURE 9.53** Low accelerating voltage images taken on a Si wafer sample before (left) and after (right) 10 min continuous electron irradiation. Note the brighter irradiated area and the absence of common darkening due to contamination.

## 9.8 HELIUM ION MICROSCOPY

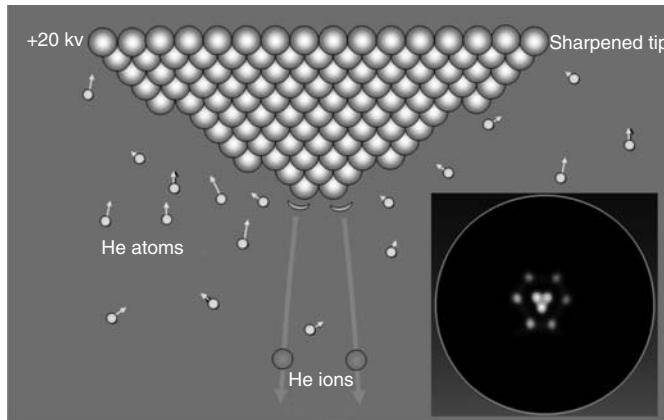
It is rare that a new type of microscopy emerges. This is the case with the recent introduction of the helium (He) ion microscope (HIM). It is a new and promising imaging and measurement technique. It is likely that the researchers will go through many of the possible uses and applications that SEM have been tried and used for. It is also probable that it will find its place among the viable techniques for material and biological research. Figure 9.54 shows the first-in-the-world commercial HIM made by ALIS/Zeiss Orion that was installed at NIST in 2007.

The HIM is expected to be an excellent tool for nanotechnology, nanomanufacturing, and general high-resolution microscopy. For many applications it has advantages over both the traditional scanning electron and existing focused ion beam microscopes that use gallium (Ga) ions. The He ion source is a gas phase field ionization source (GFIS) that has very high brightness, low energy spread, and a stable emission current, which can, in theory, be brought to a focus with a probe size of  $<0.25$  nm in diameter (see Chapter 3 for a detailed account of GFIS optics). The He ion technology offers several unique advantages over an SEM. First, the interaction volume of the He beam and the sample can be considerably smaller than that of an SEM, so sample details at the surface can be better investigated. Upon penetrating the sample, the He beam produces both scattered ions and large quantities of SEs, making imaging with low probe currents feasible. The depth-of-field is also large, which can be very advantageous for high-resolution investigation of biological and other samples characterized with fine surface details and large depths. The SE signal produces images with strong and familiar topographic contrast, very similar to the SE images obtained from an SEM.

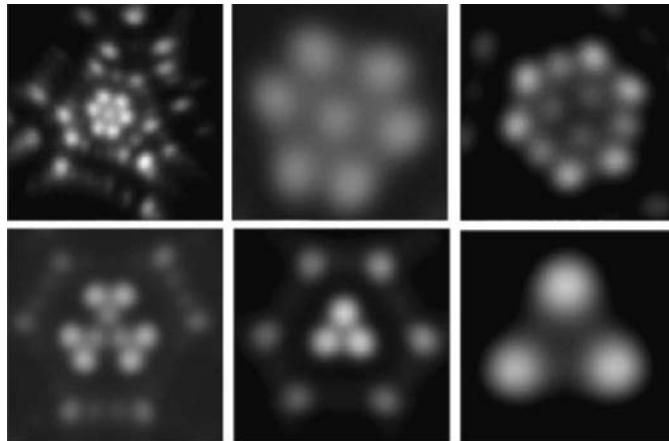
In the HIM, the helium ionization takes place at the atomically sharp tip, which is cryogenically cooled in a low-pressure helium gas environment. The arriving helium atoms adhere everywhere to the tip, but only at the very tip of the pyramid is the electric field high enough for the atoms to ionize. With a proprietary automatic process the tip of the emitter is formed in such a way as to have only three atoms or a trimer, schematically shown in Figure 9.55. Under normal operating conditions the emission from a single atom of a three-atom cluster (trimer) is selected by tilting the



**FIGURE 9.54** The first commercial helium ion microscope (ca. 2007).



**FIGURE 9.55** Helium ion generation at the tip of the emitter formed by three atoms (trimer). The insert shows a field ion microscope image of the helium ion emission. (Courtesy Carl Zeiss, Inc.)



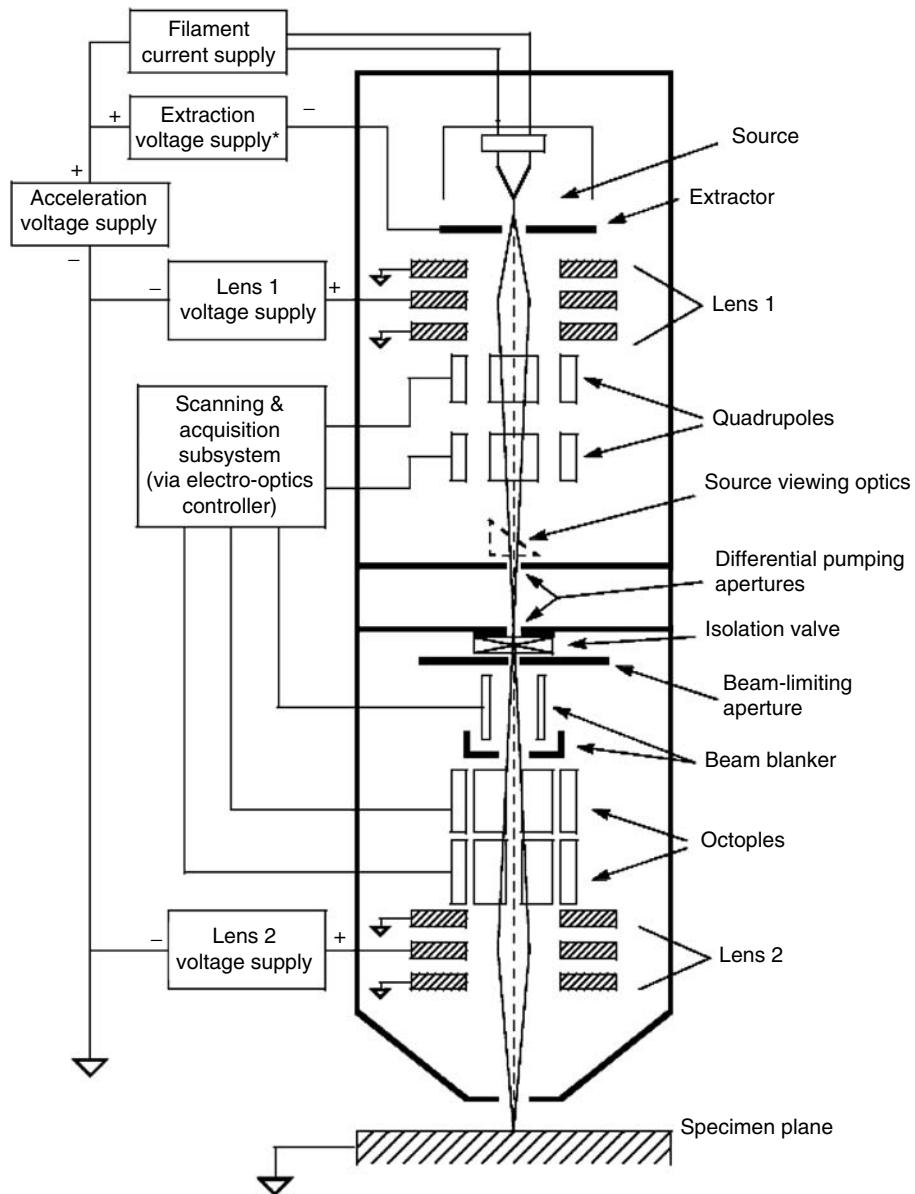
**FIGURE 9.56** Tip fabrication process viewed in field ion mode by the removal of one atomic shelf at a time. Note the formation of the desired trimer in the last image. (Courtesy Carl Zeiss, Inc.)

gun and using a beam limiting aperture. This permits a smaller source size but at the expense of less beam current. Fortunately, due to the high-brightness helium ion source, the beam is still very bright relative to electron sources.

The tip fabrication process allows for *in situ* formation of the tip apex shape by adding and removing atoms to reach the desired configuration. This process can be easily monitored in the field ion microscope (FIM) imaging mode. Figure 9.56 shows a series of FIM images that were obtained during the process of fabricating a tip with three atoms at the end of the emitter.

The ion-optical column (see Figure 9.57) uses electrostatic lenses and has an easy-to-use graphical control interface. Almost all functions are computerized, including the renewal processes. The tip could last for many days, and it can be reformed many times to maintain the original performance.

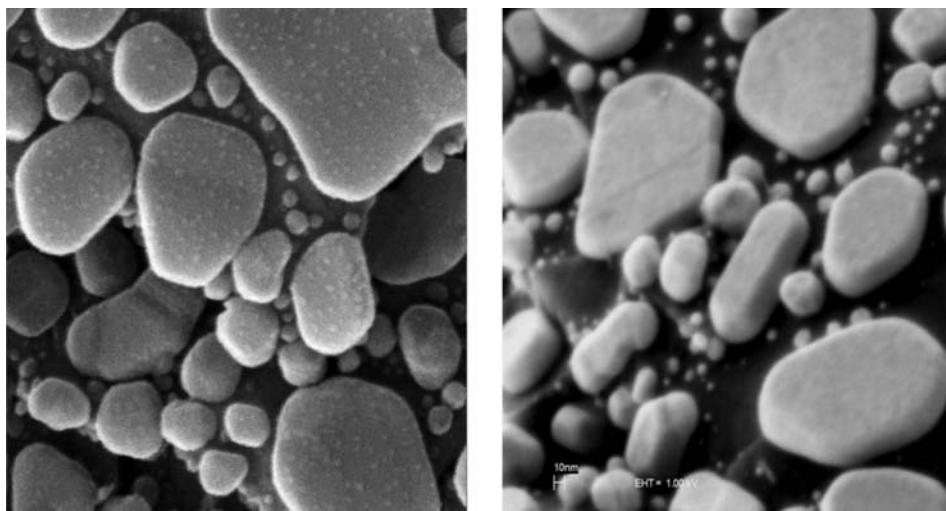
The first results show very promising performance and allow for favorable comparisons. Figure 9.58 presents a comparison between the HIM and the SEM on a gold-on-carbon sample. The SEs generated by the HIM at 30 keV landing energy show better definition of the small gold surface structural detail in comparison to the low-accelerating voltage SEM image.



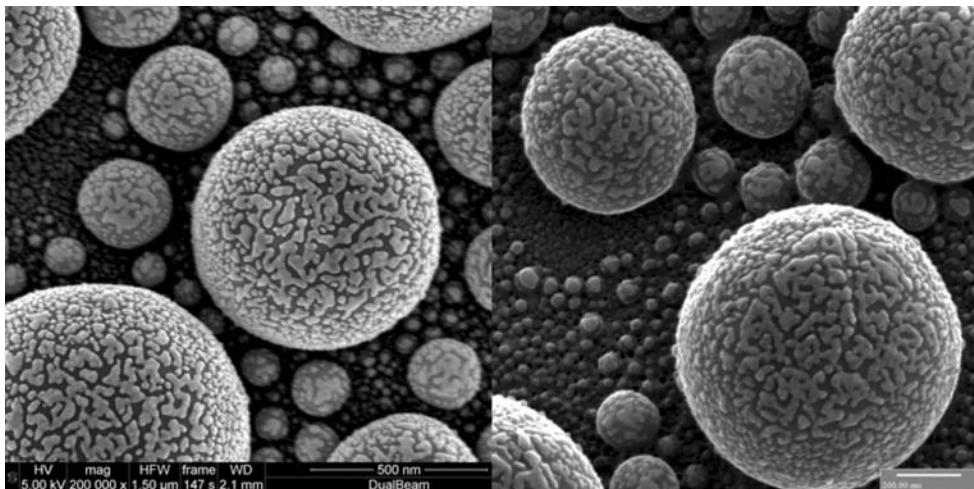
**FIGURE 9.57** The ion-optical column of the Orion helium ion microscope. (Courtesy Carl Zeiss, Inc.)

The HIM technology is not yet as optimized, developed, or mature as the SEM. As a new technique, HIM is just beginning to show promise, and the plethora of potentially advantageous applications for nanotechnology are yet to be exploited. Now that commercial instrumentation is available, further work can be done on the fundamental science of helium ion beam generation, helium ion beam-specimen interactions, and the signal generation and contrast mechanisms defining the image.

In addition to these areas of work, modeling needs to be done to correctly interpret the signal mechanisms and to understand the imaging mechanisms. These are indispensable for accurate nano-metrology. HIM and SEM have some overlapping territory, but they remain complementary techniques for nanoscience. Helium ion beam microscopy is forging into new scientific territory, and this new and innovative technology will develop new science and contribute to the progress in nanotechnology.



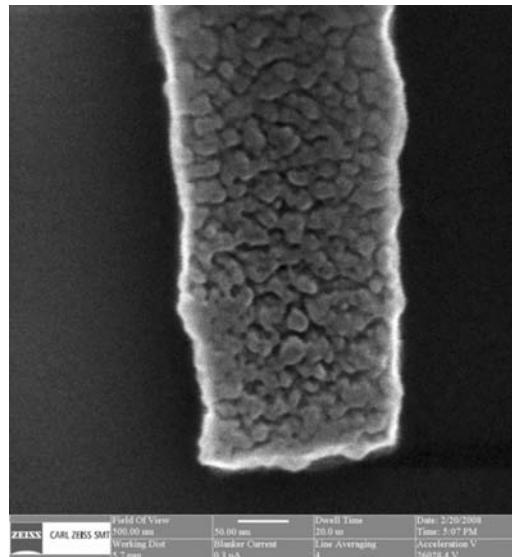
**FIGURE 9.58** Comparison of secondary electron images obtained by scanned helium ion (left) and electron beams (right) on a gold-on-carbon sample. 1  $\mu\text{m}$  field-of-view images, 1 pA beam current at 30 keV (helium ion microscope), and 20 pA beam current at 1 keV (scanning electron microscope) landing energies. The field-of-views are 400 nm. (Courtesy Carl Zeiss, Inc.)



**FIGURE 9.59** High vacuum, 5 keV landing energy SEM image of Au-decorated tin ball sample (left; the field of view is 1.5  $\mu\text{m}$ ). The same sample imaged in high vacuum with 31 keV landing energy HIM (right; the field of view is 1.5  $\mu\text{m}$ ). Micrographs by the authors.

Figure 9.59 shows a comparison of high vacuum, 5 keV landing energy SEM image of an Au-decorated tin ball sample (left) and a high vacuum, 31 keV landing energy HIM image of the same sample (right). Both images show excellent resolution and good contrast. Note the better depth-of-field (focus) of the HIM image. Both the small particles on the background surface and all regions of the large balls are sharp on the HIM image. The SEM image shows the typical restricted depth-of-field: only the top regions of the large particles are truly sharp.

With the HIM it is possible to apply an electron flood to the vicinity of the location where the primary ion beam hits the sample. This can compensate sample charging due to ions. Figure 9.60 proves that this can be done with good results even on samples that otherwise would charge up significantly. A Cr on quartz photolithography mask sample was imaged using the electron flood



**FIGURE 9.60** High vacuum, 26 keV landing energy HIM image of a Cr on quartz photolithography mask sample. (The field of view is 500 nm). An electron flood gun was used to compensate sample charging. Courtesy of Carl Zeiss Corporation.

gun and 26 keV landing energy ions. Fine details of the 70 nm thin Cr layer are clearly visible. The resolution and the image quality allow for mask dimensional measurements that are important for current and future semiconductor technology.

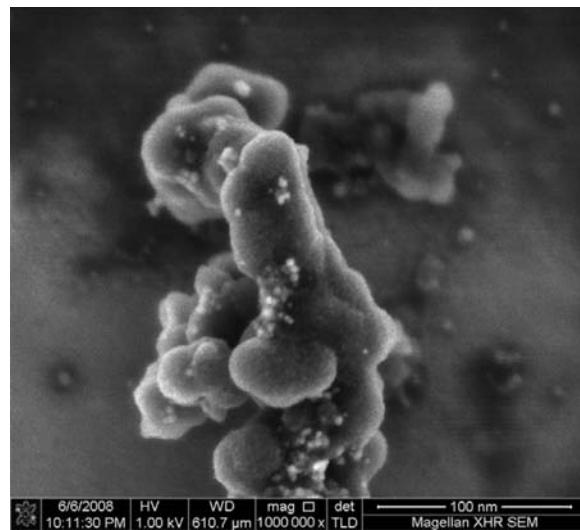
## 9.9 EXTRA HIGH-RESOLUTION SEMs

New designs that address various shortcomings of the older SEMs can bring unprecedented, sunanometer resolutions and capabilities that previously were not feasible. In-lens sample location (Hitachi S-5500), optimization of the electron optical column, reduction of energy spread of the primary electrons (FEI Magellan<sup>TM</sup> XHR SEM) and aberration corrected lenses (JEOL JEM-7700F) make it possible to achieve excellent, sub-nanometer resolution in both high and low landing energy operations. Low landing energy operation is especially advantageous, because it generally means less sample damage and more surface detail, equally essential in nano-scale imaging in scientific research and industrial development and production.

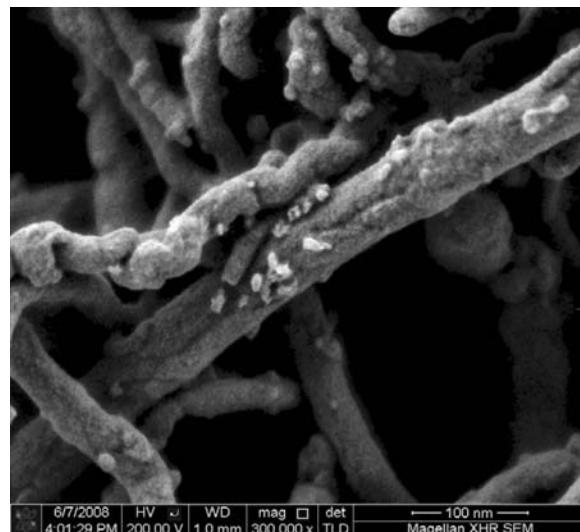
Figure 9.61 and 9.62 show that excellent resolution is possible at 1 keV or even at 200 eV landing energies and very high magnification images reveal fine details of various samples. Figure 9.61, a SE picture of a fuel cell sample clearly shows the Pt particles, which are only a few nm in size. Even surface details of the carbon globules are revealed.

Figure 9.62 is a 200 eV landing energy SEM image of bundled carbon nanotubes with bright catalytic particles. The very good resolution allows for high magnification imaging with exceptional surface details.

Sub-nanometer resolution and ultra-precise sample stages make the collection of meaningful images at close to 1 million times magnification and the reconstruction of real three-dimensional sample details at the 1 nanometer scale feasible. This is especially important for nano and semiconductor applications, where 1 nm is a “large” distance (The crystal lattice spacing of Si {111} is 0.32 nm.) The newest microscopes are getting close to the performance of the dedicated scanning transmission microscopes. While the current SEMs are not capable of resolving atoms, it is conceivable that this will be possible in the future on very thin samples. With these instruments it is not the focusing ability of the electron optical column that is the limiting factor, but the possible



**FIGURE 9.61** 1 keV landing energy SEM image of a fuel cell with bright Pt particles (the field of view is 300 nm). Courtesy of FEI Corporation.



**FIGURE 9.62** 200 eV landing energy SEM image of bundled carbon nano-tubes with bright catalytic particles (the field of view is 500 nm). Courtesy of FEI Corporation

shortcomings of the sample stage, adverse environmental effects (vibration, electromagnetic fields) and sample contamination, etc. All these must work favorably within very tight—essentially a few atoms worth of—region of performance, i.e., the adverse effects of these on the image quality must be kept at less than a single nanometer levels.

## 9.10 CONCLUSIONS

The SEM was developed in the late 1960s. This instrument has become a major research tool for many applications, providing a wealth of information not available by any other means. This instrument has also undergone a significant evolution in recent years into the production environment where it now holds a significant role in modern manufacturing.

## REFERENCES

1. Schick, J. D. 1981. Electron beam induced current analysis of integrated circuits. SEM Inc./1981/1, 295–304.
2. Qian, W., Scheinfein, M. R., and Spence, J. C. H. 1993. Brightness measurements of nanometer-sized field-emission-electron sources. *J. Appl. Phys.*, **73**(11), 7041–7045.
3. Crawford, C. K. 1979. Mounting methods and operating characteristics for LaB<sub>6</sub> cathodes. SEM Inc., pp. 19–30.
4. Postek, M. T., Lowney, J. R., Vladar, A. E., Keery, W. J., Marx, E., and Larrabee, R. D. 1993. X-ray lithography mask metrology: Use of transmitted electrons in an SEM for linewidth measurement. *NIST J. Res.*, **98**(4), 415–445.
5. Reimer, L. 1993. Image formation in low voltage scanning electron microscopy, *SPIE Tutorial*, **12**, 144, Text TT.
6. Reimer, L. and Volbert, B. 1979. Detector system for backscattered electrons by conversion to secondary electrons. *Scanning*, **2**, 238–248.
7. Wood, R. W. 1897. A new form of cathode discharge and the production of x-rays, together with some notes on diffraction. *Phys. Rev.*, **5**(1), 1–10.
8. Zworykin, V. A., Hillier, J., and Snyder, R. L. 1942. A scanning electron microscope. *ASTM Bull.*, **117**, 15–23.
9. Danilatos, G. D. 1988. Foundations of environmental scanning electron microscopy. *Adv. Electron. Electron Phys.*, **71**, 109–250.
10. von Ardenne, M. 1938. Das Elektronen-rastermikroskop. Praktische Ausführung. *Z. Tech. Phys.*, **19**(11), 407–416.
11. Orloff, J. 1984. A comparison of lanthanum hexaboride, cold field emission and thermal field emission electron guns for low voltage scanning electron microscopy. *Electron Optical Systems*. SEM Inc., 149–162.
12. Postek, M. T. 1994. Scanning electron microscope metrology. *SPIE Crit. Rev.*, **52**, 46–90.
13. Suzuki, S., Kawata, S. et al. 1995. U.S. Patent #5,396,067: Scan-type Electron Microscope, Nikon Corporation, Tokyo, Japan.
14. Volbert, B. and Reimer, L. 1979. Detector system for backscattered electrons by conversion to secondary electrons. *Scanning*, **2**, 238–248.
15. Phillips, M. R., Toth, M. et al. 1999. A novel technique for probe intensity profile characterization in the environmental scanning electron microscope. *Microsc. Microanal.*, **5**(2), 294.
16. Crewe, A. V., Isaacson, M., and Johnson, D. 1969. A simple scanning electron microscope. *Rev. Sci. Instrum.*, **40**, 241–246.
17. Danilatos, G. D. 1993. Introduction to the ESEM instrument. *Microsc. Res. Tech.*, **25**, 354–361.
18. Welter, L. M., 1975. Application of a field emission source to SEM. *Principles and Techniques of Scanning Electron Microscopy* (M. A. Hayat, ed.), Vol. 3, pp. 195–220.
19. Wells, O. C. 1986. Low loss electron images of uncoated photoresist in the scanning electron microscope. *Appl. Phys. Lett.*, **49**(13), 764–766.
20. Postek, M. T. and Vladar, A. E. 1995. Digital imaging for scanning electron microscopy. *Scanning*, **18**(1), 1–7.
21. Kruit, P. 1991. Magnetic through the lens detection in electron microscopy and spectroscopy, part 1. *Adv. Opt. Electron Microsc.*, **12**, 93–137.
22. Mulvey, T. 1974. Mini lenses and the SEM. SEM/1974/I, IITRI, 43–50.
23. Reimer, L. 1982. Electron signal and detector strategy. Electron beam interactions with solids. SEM Inc., 299–310.
24. Reimer, L. 1984. Electron signal and detector strategy. *Electron Beam Interactions with Solids for Microscopy, Microanalysis and Microlithography* (D. F. Kyser, H. Niedrig, D. E. Newbury, and R. Shimizu, eds), Scanning Microscopy International, Chicago, IL, pp. 299–310.
25. Postek, M. T., Vladar, A. E., Jones, S., and Keery, W. J. 1993. The NIST low accelerating voltage SEM magnification standard interlaboratory study. *NIST J. Res.*, **98**(4), 447–467.
26. Buckingham, J. D. 1965. Thermionic emission properties of LaB<sub>6</sub> rhenium cathode. *Br. J. Appl. Phys.*, **16**, 1812–1832.
27. Lowney, J. R., Postek, M. T., and Vladar, A. E. 1994. A Monte Carlo model for SEM linewidth metrology. *SPIE Proc.*, **2196**, 85–96.
28. Swanson, L. W. 1975. Comparative study of the zirconiated and built up W thermal field cathode. *J. Vac. Sci. Technol.*, **12**(6), 1228–1233.

29. Postek, M. T., Keery, W. J., and Larrabee, R. D. 1988. The relationship between accelerating voltage and electron detection modes to linewidth measurement in an SEM. *Scanning*, **10**, 10–18.
30. Welter, L. M. and Coates, V. J. 1971. A compact field emission scanning electron microscope. *Proc. EMSA* (C. J. Arceneaux, ed.), pp. 32–33.
31. Kyser, D. F. 1979. Monte Carlo simulation in analytical electron microscopy. *Introduction to Analytical Electron Microscopy* (J. J. Hren, J. I. Goldstein, and D. C. Joy, eds), Plenum Press, New York, pp. 199–221.
32. Kanaya, K., and Okayama, S. 1972. Penetration and energy-loss theory of electrons in solid targets. *J. Phys. D.*, **5**, 43–58.
33. Feuerbaum, H. P. 1983. Electron beam testing: Methods and applications. *Scanning*, **5**, 14–24.
34. Postek, M. T. 1994. Critical issues in scanning electron microscope metrology. *NIST J. Res.*, **99(5)**, 641–671.
35. Postek, M. T., Howard, K., Johnson, A., and McMichael, K. L. 1980. *Scanning Electron Microscopy — A Student's Handbook*. Ladd Research Industries, Burlington, VT.
36. Joens, S. W., Shah, K. K., and Shishido, H. T. 1982. Cross sectional characterization of semiconductor devices by a scanning electron microscopy technique. SEM Inc., 2/573–580.
37. Moll, S. H., Healey, F., Sullivan, B., and Johnson, W. 1978. A high efficiency non-directional backscattered electron detection mode for SEM. SEM/1978/1 SEM, Inc.
38. Moll, S. H., Healey, F., Sullivan, B., and Johnson, W. 1979. Further developments of the converted back-scattered electron detector. SEM/1979/II SEM Inc., 149–154.
39. Reisman, A., Merz, C., Maldonado, J., and Molzen, W. 1984. Low energy x ray and electron damage to IGFET gate insulators. *J. Electrochem. Soc.*, **131**, 1404–1409.
40. Russell, P. E. 1984. Microchannel plates as specialized scanning electron microscopy detectors. *Electron Optical Systems*, SEM Inc., Chicago, IL, pp. 197–200.
41. Russell, P. E. and Mancuso, J. F. 1985. Microchannel plate detector for low voltage scanning electron microscopes. *J. Microsc.*, **140**, 323–330.
42. Tuggle, D. W. and Watson, S. G. 1984. A low voltage field emission column with a Schottky emitter. *Proceedings EMSA* (G. W. Bailey, ed.) San Francisco Press, San Francisco, CA, 94101, 455–457.
43. Peters, K.-R., 1982. Conditions required for high quality high magnification images in secondary electron-I scanning electron microscopy. SEM/1982/IV, SEM Inc., pp. 1359–1372.
44. Peters, K.-R. 1985. Working at higher magnifications in scanning electron microscopy with secondary and backscattered electrons on metal coated biological specimens and imaging macromolecular cell membrane structures. Scanning Electron Microscopy/1985/IV SEM Inc., 1519–1544.
45. Everhart, T. E. and Thornley, R. F. M. 1960. Wideband detector for micro-microampere low energy electron currents. *J. Sci. Instrum.*, **37**, 246–248.
46. Postek, M. T. 1985. Low accelerating voltage inspection and linewidth measurement in the scanning electron microscope. SEM/1984/III SEM Inc., 1065–1074.
47. Qian, W., Scheinfein, M. R., and Spence, J. C. H. 1993. Electron optical properties of nanometer field emission electron sources. *Appl. Phys. Lett.*, **62(3)**, 315–317.
48. Reimer, J. 1977. Problem solving by SEM in IC wafer fabrication. *IITRI Symposium*. SEM Inc., **I**, 167–175.
49. Tuggle, D. W., Li, J. Z., and Swanson, L. W. 1985. Point cathodes for use in virtual source electron optics. *J. Microsc.*, **140(3)**, 293–301.
50. Tuggle, D. W. and Swanson, L. W. 1985. Emission characteristics of the ZrO/W thermal field electron source. *J. Vac. Sci. Technol.*, **B3(1)**, 220–223.
51. Ingrey, S. I. 1982. Applications of auger spectroscopy to semiconductor device analysis. SEM Inc., 1982/4/1475–1485.
52. Jensen, S. 1980. Planar micrometrology in the SEM. *Microbeam Analysis*, San Francisco Press, San Francisco, CA, pp. 77–84.
53. Jensen, S. and Swyt, D. 1980. Sub micrometer length metrology: Problems, techniques and solutions. *Scanning Electron Microsc.*, **I**, 393–406.
54. Mair, G. L. R. and Mulvey, T. 1984. High brightness sources for electron and ion beam microscopy and micro lithography. *Ultramicroscopy*, **15**, 255–260.
55. Swanson, L. W. 1984. Field emission source optics, *Electron Optical Systems*, SEM Inc., Chicago, IL, pp. 137–147.
56. <http://spectelresearch.com/semonitor.htm>
57. Swanson, L. W. and Crouser, L. C. 1969. Angular confinement of field electron and ion emission. *J. Appl. Phys.*, **40(12)**, 4741–4749.
58. Yacobi, B. G. and Holt, D. B. 1990. *Cathodoluminescence Microscopy of Inorganic Solids*, Plenum Press, New York, 302 pp.

59. Reimer, L. 1979. Electron-specimen interactions. SEM Inc., 1979/II. 111–124.
60. Joy, D. C. 1987. Low voltage scanning electron microscopy. *Inst. Phys Conf. Ser.*, No. 90, Chapter 7. EMAG, 175–180.
61. Griffith, E. and Newbury, D. E. 1996. Introduction to the environmental scanning electron microscopy issue. *Scanning*, **18**(7), 465.
62. Christman, L. A., Schwind, G. A., and Swanson, L. W. 1991. A comparison of CeB<sub>6</sub> and LaB<sub>6</sub> thermionic cathodes. *Proc. EMSA* (G. W. Bailey and E. Hall, eds), San Francisco Press, San Francisco, CA, pp. 46–347.
63. GSFC S 311 P12A. Specification for scanning microscope inspection of semiconductor device metallization. NASA/Goddard Space Flight Center, Greenbelt, MA.
64. Reimer, L. 1977. Scanning electron microscopy — Present state and trends. *Scanning*, **1**, 3–16.
65. Ahmed, T., Chen, S.-R., Naguib, H. M., Brunner, T. A., and Stuber, S. M. 1987. Low voltage metrology for pilot line applications. *Proc. SPIE*, **775**, 80–88.
66. Bhattacharya, P. K., Jones, S. K., and Reisman, A. 1989. A study of electron damage effects during low voltage SEM metrology. *Proc. SPIE*, **1087**, 9–16.
67. Postek, M. T., 1984. Low accelerating voltage inspection and linewidth measurement in the scanning electron microscope. SEM/1984/III, SEM Inc., 1065–1074.
68. Postek, M. T. 1987. Nondestructive submicron dimensional metrology using the scanning electron microscope. *Rev. Prog. NDE*, **6**(b), 1327–1338.
69. Bennett, M. H. and Fuller, G. E. 1986. In process inspection and metrology of semiconductor wafers with the use of an automated low voltage SEM. *Microbeam Anal.*, **21**, 649–652.
70. Orloff, J. 1985. Thermal field emission for low voltage scanning electron microscopy. *J. Microsc.*, **140**(3), 303–311.
71. Heinrich, K. F. J. 1964. Relationships of sample composition, backscatter coefficient, and target current measurement. *Adv. X-ray Anal.*, **7**, 325–339.
72. Heinrich, K. F. J. 1966. Electron probe microanalysis by specimen current measurement. X-ray optics and Microanalysis. *4th Int. Symp. on X-Ray Optics and X-Ray Microanalysis*, pp. 159–167.
73. Hembree, G. G., Jensen, S. W., and Marchiando, J. F. 1981. Monte Carlo simulation of submicrometer linewidth measurements in the scanning electron microscope. *Microbeam Analysis*, San Francisco Press, San Francisco, CA, pp. 123–126.
74. Tuggle, D. W., Swanson, L. W., and Orloff, J. 1979. Application of a thermal field emission source for high resolution, high current e beam micropipes. *J. Vac. Sci. Technol.*, **16**(6), 1699–1703.
75. MIL STD 883A Method 2018. *Scanning Electron Microscope (SEM) Inspection of Metallization*. U.S. Govt Printing Office, Washington, DC.
76. Gopinath, A., Gopinathan, K. G., and Thomas, P. R. 1978. Voltage contrast: A review. SEM Inc., 1978/1/375–380.
77. Baumgarten, N. 1989. Environmental SEM Premieres. *Nature*, **341**, 81–82.
78. Utterback, S. G. 1988. Semiconductor dimensional metrology using the scanning electron microscope. *Review of Progress in Quantitative non-Destructive Evaluation*, Plenum Press, New York, pp. 1141–1151.
79. Wight, S., Gillen, G., et al. 1997. Development of environmental scanning electron microscopy electron beam profile imaging with self-assembled monolayers and secondary ion mass spectroscopy. *Scanning*, **19**(2), 71–74.
80. Thiel, B. L., Bache, I. C. et al. 2000. Imaging the probe skirt in the environmental SEM. *Microsc. Microanal.*, **6**(2), 794.
81. Farley, A. N. and Shah, J. S. 1991. High-pressure scanning electron microscopy of insulating materials: A new approach. *J. microsc.*, **164**, 107–112.
82. Mancuso, J. F., Maxwell, W. B., and Danilatos, G. D. 1988. Secondary electron detector for use in a gaseous atmosphere, US Patent 4785182.
83. Postek, M. T., Keery, W. J., and Frederick, N. V. 1990. Low profile high efficiency microchannel plate detector system for scanning electron microscopy applications. *Rev. Sci. Instrum.*, **61**(6), 1648–1657.
84. Crewe, A. V., Eggenberger, D. N., Wall, J., and Welter, L. M. 1968. Electron gun using a field emission source. *Rev. Sci. Instrum.*, **39**, 576–583.
85. Rayleigh Resolution [http://en.wikipedia.org/wiki/Angular\\_resolution](http://en.wikipedia.org/wiki/Angular_resolution)
86. Spectel research <http://www.spectelresearch.com/>
87. RM 8091 Scanning Electron Microscope Sharpness Standard [https://srmors.nist.gov/view\\_detail.cfm?srm=8091](https://srmors.nist.gov/view_detail.cfm?srm=8091)
88. SMART Algorithm <http://web.utk.edu/~srcutk/smartzstuff/>
89. Evactron <http://www.evactron.com/>

---

# 10 The Scanning Transmission Electron Microscope

*Albert V. Crewe (updated by Peter D. Nellist)*

## CONTENTS

Editor's Preface .....	498
Preface to Updated Version .....	498
10.1 Introduction .....	498
10.2 General Principles.....	499
10.3 Imaging and Spectroscopy in the Scanning Transmission Electron Microscope .....	500
10.3.1 Bright-Field Imaging .....	500
10.3.2 Annular Dark-Field Imaging .....	501
10.3.3 Spectroscopy .....	502
10.3.4 Other Detection Modes and Signals .....	503
10.4 Formation of the Electron Probe.....	503
10.4.1 Geometric Size.....	504
10.4.2 Diffraction .....	504
10.4.3 Chromatic Aberration .....	505
10.4.4 Spherical Aberration .....	505
10.5 Brightness and the Effect of Source Size.....	506
10.6 Optimization of a Focused Probe .....	508
10.6.1 Spherical Aberration Plus Diffraction .....	509
10.6.2 Chromatic Aberration Plus Diffraction .....	510
10.7 Application to the Scanning Transmission Electron Microscope .....	510
10.8 Components of the Electron-Optical Column .....	511
10.8.1 The Electron Source .....	511
10.8.1.1 The Cold Field Emitter Source.....	512
10.8.1.2 The Thermally Assisted, Zirconium-Treated Field Emitter Source.....	513
10.8.2 The Electron Gun.....	513
10.8.3 The Magnetic Lenses .....	515
10.8.4 Column Designs and Performance .....	515
10.8.5 Other Features of the Optical Column .....	517
10.9 Scanning Transmission Electron Microscope Types .....	518
10.10 Correction of Spherical Aberration in Scanning Transmission Electron Microscope .....	518
Acknowledgments.....	519
References .....	520

## EDITOR'S PREFACE

I was saddened to learn of Professor Albert Crewe's ill health that prevented him from updating his chapter on "The Scanning Transmission Electron Microscope." Doctor Peter Nellist was kind enough to offer to assist with this effort by reading and updating the chapter where necessary. In this chapter where the word "author" appears, it is in reference to Professor Crewe.

## PREFACE TO UPDATED VERSION

For this new edition of the *Handbook of Charged Particle Optics* I was delighted to be asked to update Albert Crewe's original chapter on STEM, but it was also a project I approached with some trepidation. Albert Crewe is rightly acknowledged as the father of the modern STEM, and his original chapter elegantly described the particle-beam physics involved in its operation, the vast majority of which is still highly relevant to modern STEM instruments. I therefore decided to take a light touch and only make modifications where the modern technology and applications of STEM had moved on from the original chapter. Throughout most of this chapter I have limited my modifications to the addition of new references and relating the cited new work to the points Professor Crewe was making in his text. In two sections, however, I have made greater modifications. I have reworked the original sections on "Electron Interactions in the Specimen" and "Postspecimen Optics" into a single section, now near the start of the chapter, entitled "Imaging and Spectroscopy in the STEM," to enable the reader to understand how a modern STEM is typically used, before delving into the details of the beam physics in the rest of the chapter. I have also added a new section on the correction of spherical aberration, which may be regarded as the most important area of development since the original chapter. Indeed, the impact of spherical aberration correction on STEM is so great that I have also highlighted it at several other relevant places in the chapter. Like all those using modern STEM instruments I remain indebted to Professor Crewe, and if my modifications to his chapter seem presumptuous then I can only apologize—PDN.

### 10.1 INTRODUCTION

The seminal publication by Knoll and Ruska on the invention of the electron microscope appeared in 1932 (Knoll and Ruska, 1932) and initiated a flurry of activity, so much so that in the next few years several new inventions appeared, including the scanning electron microscope (SEM) (Zworykin et al., 1942) and the scanning transmission electron microscope (STEM) by von Ardenne (1938). In that article he described an instrument in which a focused probe of electrons was scanned across a specimen in an X-Y raster, and electrons that penetrated the specimen were recorded on photographic paper on a rotating drum. It was soon recognized that this ingenious device had some potential advantages over the conventional transmission electron microscope (CTEM) because the ultimate resolving power of the two instruments were identical but the images formed in the STEM would not be affected by the chromatic aberration of the electrons, which lose energy in the specimen. On the other hand, there were some obvious disadvantages. The formation of a focused probe with interestingly small dimensions entailed such a loss of probe current that the exposure time for a complete image was very long, and perhaps worst of all was the fact that the image could not be seen directly so that focusing the probe was difficult. It would seem now that the invention was too far ahead of the available technology of the times.

The modern version of the STEM was first proposed in an article by the author in 1963 (Crewe, 1963), a practical form of the instrument was described somewhat later (Crewe, 1966; Crewe et al., 1968b), and the first critical analysis of image formation was published in Zeitler and Thomson (1970). This STEM was a rediscovery or reinvention whose conception was based on techniques and processes in nuclear physics, and it was lamentably ignorant of any precedent. Very fortunately, there had been many significant advances in technology in the intervening years, which made the

concept viable, and we can mention in particular the invention of, and developments in, the television industry, the development of ultrahigh vacuum technology, and the development of understanding of the principles of field emission. At the time of this reinvention of the STEM some 40 years ago, these various technological advances had reached the point where, although in many respects the knowledge was inadequate, there was just enough available to allow further development of the microscope.

As a historical note we should perhaps point out that in the early years of the development of the STEM, it was merely a curiosity for the microscope community. For the first decade or so there was only one STEM and that was in the author's laboratory, and the developments were the result of the efforts of his students. Subsequent years saw the construction of a few similar instruments in other laboratories. But it was only with the introduction of a commercial instrument by VG Microscopes (no longer in existence) that the true power of the STEM began to be appreciated. Today, many research laboratories purchase field-emission instruments capable of both CTEM and STEM operation. The number of instruments being used as a STEM is only a fraction of the number of TEMs, but their number is rapidly increasing.

In Section 10.2 we will discuss the general principles of operation of this type of microscope, paying particular attention to the limits of performance. To provide details that are necessary for a complete understanding of this instrument, we must include some material that will overlap in content with other chapters in this book (see Chapters 2, 8, and 9). On these subjects we will provide only a brief introduction, and the reader is referred to those chapters for a more complete discussion.

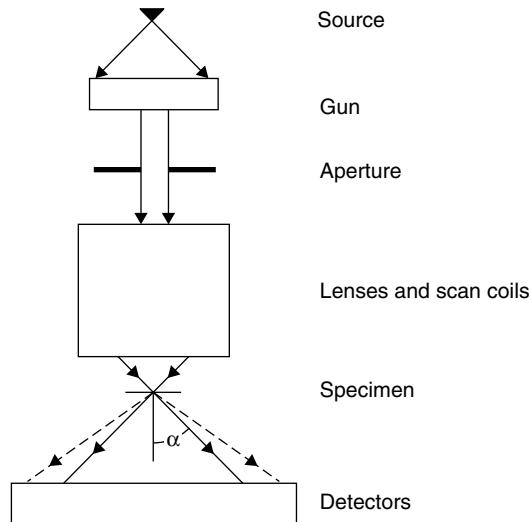
In this brief survey we cannot include all the developments that have occurred in the past 40 years, but we can refer the reader to more complete surveys that have appeared recently. In particular, Cowley (1995) has reviewed various applications of the STEM, and a detailed description of image formation in the STEM (particularly bright-field [BF] images) can be found in Hawkes and Kasper (1994). A detailed account of a commercial instrument is given by von Harrach (1994). More recently, Nellist (2007) has reviewed the imaging modes and range of applications available in STEM.

The discussion in this chapter stresses the instrumental aspects of the STEM and the electron optics that is involved because this is the area of interest for this book. Regrettably this means that there is little space to address the numerous applications of the STEM and the recent growth in the numbers of these applications. Fortunately, there are some excellent reviews and summaries of these applications, and this allows us to limit our discussion to the impact they have on the design of the instrument itself.

## 10.2 GENERAL PRINCIPLES

We suppose an electron source and a system of electrodes that can accelerate the electrons to some energy of interest, which could be anywhere from 100 V to several hundred thousand volts. This beam of electrons can then be brought to a focus in the plane of the specimen with the aid of one or more magnetic lenses. If the specimen is thin enough some electrons will penetrate and appear on the other side. These electrons can be detected directly or they can be sorted according to some preferred criteria and selectively detected. A set of deflector coils or electrodes can be used to scan the focused probe in an X-Y raster while the detected electrons can be collected and their intensity plotted as a function of the probe position to form an image. The magnification is simply the ratio of the two scan amplitudes whereas the resolution is determined largely by the size of the focused probe. Image contrast depends on the choice of electrons to be detected. A schematic representation is shown in Figure 10.1.

This arrangement has some possible advantages over the CTEM. (i) The image contrast is determined by the detection scheme, which can be tailored for specific uses. For example, use of an annular dark-field (ADF) detector leads to incoherent Z-contrast imaging, an imaging mode that has no analog in the CTEM. (ii) Perhaps more importantly, the resolution is largely determined by the properties of the electron beam prior to its interaction with the specimen. The critical electron optics involves only the primary electron beam and does not have to transmit either the image or the energy loss electrons (see, e.g., Crewe, 1970, 1971a). Recently, this feature of STEM has become



**FIGURE 10.1** Schematic representation of a scanning transmission electron microscope. Electrons emitted by the source are accelerated in an electron gun and passed through an aperture, which controls the quality of the beam. The electrons are focused onto the specimen with the aid of a lens, and the focused probe is scanned across the specimen in a two-dimensional pattern. A system of detectors responds to the electrons passing through the specimen. The detector output can be used to control the intensity in a synchronously scanned display device.

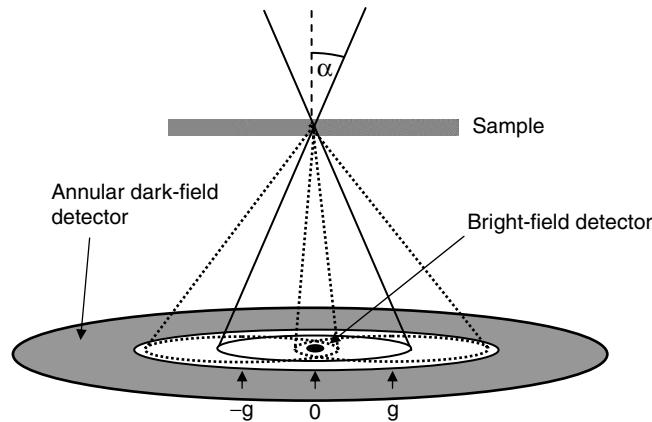
particularly important with the development of spherical aberration correctors, as will be discussed in Section 10.10.

### 10.3 IMAGING AND SPECTROSCOPY IN THE SCANNING TRANSMISSION ELECTRON MICROSCOPE

Although this chapter is primarily concerned with the principles of operation of the STEM, paying particular attention to the limits of performance, it is helpful first to understand the principles behind the formation of images and spectra in the STEM. The STEM has naturally been optimized for these modes of operation, and the performance of a modern STEM is judged using them. In this section we will restrict ourselves to describe the basic principles. Further details can be found in Nellist (2007).

#### 10.3.1 BRIGHT-FIELD IMAGING

Axial BF imaging is performed in the STEM by placing a small detector on the optic axis beyond the sample. To understand the type of image contrast we might expect it is the easiest way to first consider a crystalline sample. In a CTEM, the sample would be illuminated by a plane wave, and a set of small diffraction spots would be observed after the sample. In a STEM, the illumination is by a convergent beam of illumination, which means that rather than seeing diffraction spots, we see diffraction disks. If the illuminating beam is fully coherent, then interference features will be seen, often as a set of interference fringes in the overlap region. The form of the interference features will depend on the aberrations of the lens system forming the convergent beam, the phase relationship between the overlapping diffracted beams (which carries the structural information about the sample), and the position of the illuminating probe relative to the sample lattice. By moving the probe relative to the lattice, the interference in any given position will vary between bright (constructive) and dark (destructive). If the BF detector is in an overlap region (Figure 10.2), then the intensity variation as the probe is moved will show up as contrast in the image, in the form of lattice fringes. For this to happen, either the beam must have a high enough convergence angle or the diffraction



**FIGURE 10.2** A schematic diagram showing that for a crystalline sample, a small, axial bright-field scanning transmission electron microscope detector will record changes in intensity due to interference between three beams: the 0 unscattered beam and the  $+g$  and  $-g$  Bragg reflections. The annular dark-field detector collects higher-angle elastic and thermally scattered electrons.

beam close enough in scattering angle, for beam overlap to occur at the BF detector. In this case the microscope is able to resolve the crystal spacing. More detail on STEM lattice imaging is given by Spence and Cowley (1978).

We have concluded that BF imaging is a coherent interference process that involves the scattering by the sample and the aberrations of the lens system focusing the convergent beam. A similar description may apply to lattice imaging in the CTEM, indeed, the principle of reciprocity (as applied to STEM by Zeitler and Thomson (1970) and Cowley (1969)) shows that they are equivalent. The conventional transmission electron microscope (CTEM) and the STEM can be considered to be mirror images of one another, and imaging properties of one can be considered to be the same as the other except that the electrons travel in the opposite direction. The size of the BF detector in diffraction space plays the same role as the beam convergence angle in the CTEM. Indeed, the earliest STEM experiments used a small axial detector and Fresnel fringes identical to those in a TEM were observed (Crewe and Wall, 1970). An important point, however, is that because the BF detector is smaller than the convergence angle of the beam, many electrons are lost, and performing BF imaging is generally less efficient in STEM than in CTEM. It should also be noted that, like in CTEM, a minimum in contrast is usually seen when the microscope is precisely focused. Thin electron samples can often be regarded as phase objects, and finite lens aberrations, such as defocus, are required to give phase contrast.

### 10.3.2 ANNULAR DARK-FIELD IMAGING

The most commonly used imaging mode in STEM is that using an ADF detector. The ADF detector collects all the electrons scattering to angles between its inner and outer radii (Figure 10.2). The outer radius is usually so large that the electron scattering is negligible at that angle, and it is the size of the inner radius that is important. It is referred to as a dark-field detector because the image will only contain intensity when the probe is located over a sample that can scatter the electrons.

Naively we would expect that the electrons would be scattered by the atoms in the specimen and that, for a high enough inner radius, the scattering would be Rutherford scattering such that the scattering intensity is proportional to the square of the atomic number. However, Rutherford scattering normally applies only to a bare nucleus and for solid specimens the atomic nuclei are well screened by the electrons in the atom and by the free electrons in the solid. This leads to a modification of the usual expressions for Rutherford scattering, and the simplest method uses the Wentzel potential, which approximates the electron screening with an exponential function. The effect of this modification can be understood on a simple classic model. For high-angle scattering, the electrons must

penetrate close to the atomic nucleus, that is to say, *through* the electron cloud. Such a scattering event would then be expected to be closely approximated by Rutherford scattering, and indeed it is. A useful rule of thumb is that the intensity detected varies as approximately  $Z^{1.7}$ . The strong atomic number dependence of ADF imaging is a unique feature of this mode of STEM imaging and is sometimes referred to as Z-contrast. In modern instruments capable of lattice imaging, the Z-contrast nature of ADF imaging is extremely powerful and allows the direct identification of the presence of impurities (see, e.g., Voyles and Muller, 2002).

A further important feature of ADF results from the large extent of the detector. Any structure in the detector plane due to interference between diffracted beams is averaged over, resulting in an image that can be described as being incoherent. By reciprocity, the ADF detector is equivalent to a large incoherent source in a CTEM. Such a source is not available in CTEM, but if it were it would result in an incoherent image. The incoherence is further enhanced because the electrons collected at high angles are predominantly thermally scattered giving diffuse intensity. An analogy in light optics is to consider the difference between illuminating something with a laser beam and illuminating it with a fluorescent light tube. The incoherent image offered by the light tube is much simpler to interpret because of the lack of interference features. Because electrons interact so strongly with matter, they are multiply scattered in all but the very thinnest samples. For high-resolution CTEM, multiple scattering can complicate the image interpretation, and comparison of image simulations from trial structures to the experimental data is often the only way to proceed. The incoherent nature of STEM ADF reduces the complicating effects of the multiple scattering and is therefore particularly powerful at atomic resolution where the structure determination of defects can often be performed by direct observation (see, e.g., Chisholm and Pennycook, 2006).

The performance of the STEM is much simpler to assess in the incoherent ADF imaging mode. The image resolution is simply controlled by the extent of the intensity of the probe, and so the objective is to focus as much current into as small a probe as possible. Defocus is not necessary to observe image contrast, as it is for the coherent BF imaging mode.

### 10.3.3 SPECTROSCOPY

Spectroscopy is an important application in STEM, with the common ones being energy-dispersive x-ray (EDX) and electron energy loss spectroscopy (EELS). When high-energy electrons penetrate specimens, they can deposit energy in various ways. From a single atom point of view the electron can ionize an atom or excite it. For this reason one would expect many electron losses in the low-energy-loss region corresponding to these low-energy events. For deeper penetrations of the high-energy electrons into the atom the various electronic shells can be excited leading to the emission of x-rays in the secondary deexcitation processes, or characteristic core-loss edges in EELS spectra. Both of these can therefore give rise to compositional information.

This simple picture becomes much more complicated when the specimen is in the solid state, when it cannot be considered as a collection of independent and isolated atoms. All practical specimens must be considered to be in the solid state. Fine structure can then be observed in the EELS spectra that allow one to go further than just composition to measuring changes in the local electronic structure, or bonding, between the atoms.

A modern EELS spectrometer will detect the entire spectrum in parallel typically using a charge-coupled device (CCD) camera. A complete electron energy loss spectrum can be recorded for each point on the specimen (a technique known as spectrum imaging). The STEM then becomes an analytical tool capable of determining the chemical composition of specimens at high spatial resolution, or indeed for mapping local changes in electronic structure (see, e.g., Muller et al., 1993).

Inelastic scattering is predominantly an incoherent process. In a way similar to ADF imaging, maximizing the spectroscopy performance of the STEM is a matter of focusing as much current as possible into a probe as small as possible.

### 10.3.4 OTHER DETECTION MODES AND SIGNALS

The BF, ADF, EELS, and EDX signals described earlier are by far the most commonly detected signals for STEM work. However, it is clear that there are a very wide range of signals that could be detected as a function of the probe position. A few of these signals are discussed here.

We have already mentioned that a coherent convergent beam diffraction pattern is available in the diffraction plane and that this varies as a function of the probe position. Most modern instruments have cameras, based on electron scintillators coupled to CCD arrays, that allow this pattern to be directly observed. This has proved to be an extremely helpful diagnostic for microscope alignment because of the sensitivity of the interference features to the lens aberrations. Indeed, tuning of aberration correctors can be performed making use of this (Dellby et al., 2001). There has also been significant work (e.g., Nellist and Rodenburg, 1995) making use of the diffraction patterns recorded as a function of probe position to reconstruct high spatial resolution information regarding the sample.

There are many other events that occur when fast electrons strike a specimen, and most, if not all, of them have been used for imaging.

Slow secondary electrons are produced with high efficiency. When these electrons are used for imaging, the device is called the SEM and is outside the scope of this chapter (see Chapter 9). On the other hand, a STEM can be readily used for such purposes. In addition, the STEM has been used to make detailed studies of the production process, making use of the simultaneous appearances of inelastic scattering events (Voreades and Crewe, 1974; Voreades, 1976). There has been further interest in this type of study (Scheinfein et al., 1993; Mullejans et al., 1993). Auger electrons are also produced and can be used for imaging (e.g., Hembree and Venables, 1992).

Several authors have proposed using multiple detectors in the BF cone (Dekkers and Lang, 1974; Rose, 1975; Cowley, 1976; Hawkes, 1978), both subdivided into annular zones and various segments. Such detectors are particularly useful for the imaging of magnetic domains (Chapman et al., 1992).

One could think of many possible detector geometries, and detector geometry design is an area of research that can benefit greatly from the technology of silicon microfabrication (Hammel et al., 1990).

The task of extracting information from specimens at the high spatial resolution, which is afforded by the STEM, continues to be a challenge to investigators. This challenge is becoming even greater with the expanding use of aberration correctors and the subangstrom probes that are now available. In this respect it may be worthwhile to reiterate the special feature of the STEM, which was noted by the earliest proponents, namely, that the critical electron-optical elements do not have to deal with electrons that have lost energy in the specimen. When we add the obvious fact that these elements only need to image one point on the optical axis to another similar point, it becomes clear that the STEM is an ideal test bed for any system which can be used to correct aberrations.

## 10.4 FORMATION OF THE ELECTRON PROBE

The electron probe is formed by a lens that can be electrostatic, but today it is almost always magnetic. Inevitably, the electron beam will be defined by an aperture or restriction at some point along its length, which will have the effect of defining the semiangle of convergence  $\alpha$  at the focus. More specifically, this is the geometric angle that is obtained using only the first-order optics.

There are many types of lenses that could be used to form this probe, but we will restrict our attention to cylindrically symmetric ones. These lenses have the smallest aberration coefficients and consequently form the smallest probes.

There are several instrumental factors that determine the size (radius) of the final focused probe and in particular the size of the smallest probe which can be formed, and the ultimate resolution of the instrument is limited by the combinations of these factors. To understand the behavior of the STEM and its performance limits, we will describe the most significant of these factors and how they determine the probe size.

### 10.4.1 GEOMETRIC SIZE

All sources have some finite size and we can define a radius  $\Delta$ , and then the effect at the final probe is determined by the total first-order magnification  $M$ . We can write

$$\delta_{ss} = M\Delta \quad (10.1)$$

Values of  $\Delta$  range widely depending on the type of electron source. The largest values that may be of interest here are for the hot filament source where  $\Delta$  can be as high as 1 mm. The smallest values are for field-emission sources where we can have  $\Delta$  of the order of nanometers. In addition to this, we do not really expect the source to act as a disk with a uniform brightness so that we ought to consider  $\Delta$  to be a distribution function  $\Delta(r)$ . However, when we are interested only in the smallest probes, we show later that it is not necessary to enter into such details.

In the case of the hot filament source we can take  $\Delta$  to be the actual size of the emitting surface, but in other cases this may not be appropriate. For example, field-emission sources consist of a tungsten wire in the form of a narrow cone with a hemispherical emitting surface at the tip whose radius is in the range of  $\rho = 0.1 \mu\text{m}$ . If we take the energy of the emitted electrons to have an average value of  $\Delta V$  and consider all possible angles of emission, the trajectories will approach radial lines asymptotically. Conservation of angular momentum can be used to estimate the apparent source size when the electron energy reaches a value  $V$  (Gomer, 1993)

$$\Delta \sim \rho \sqrt{\frac{\Delta V}{V}} \quad (10.2)$$

However, the electrons are emitted with various energies so this is only an average value. In addition this is a classical calculation and any accurate description would take into account the wave nature of the electrons, because for the average energy of emission the wavelength is  $\sim 5 \text{ nm}$  and cannot be ignored.

For all these reasons we can only estimate the value  $\Delta$  for these small sources. From the various considerations that follow we can conclude that there are only two electron sources that are suitable for use with a STEM, the cold field emitter (CFE) and the thermally assisted, zirconium-treated field emitter (TFE). Various attempts have been made to measure or calculate the properties of these two sources, and here we will give the typically accepted values. For the CFE source we assume an energy distribution that is roughly Gaussian with a half width at half height of 0.3 eV and an effective source radius of 5 nm. For the TFE source we assume the values 0.8 eV and 15 nm (Orloff, 1988).

### 10.4.2 DIFFRACTION

The electron probe is defined by a physical aperture somewhere in the optical column, and this aperture has the effect of controlling the convergence angle of electrons at the focus. If we make the assumption that the aperture is uniformly illuminated by a parallel beam, then the diffraction pattern will be an Airy disk and the radius to the first zero is given by

$$\delta_d = \frac{0.61\lambda}{\alpha} \quad (10.3)$$

where  $\lambda$  is the wavelength.

$$\lambda \sim \frac{1.23}{\sqrt{V}} \text{ nm} \quad (10.4)$$

so that

$$\delta_d = \frac{0.75}{\alpha\sqrt{V}} \text{ nm} \quad (10.5)$$

where  $V$  is the accelerating voltage and this radius is normally taken to be the resolution of a diffraction-limited probe, as defined by Rayleigh.

The intensity distribution in a probe limited only by diffraction is identical to that of a similar probe in light optics, a problem which has been well discussed in various texts (Born and Wolf, 1980).

#### 10.4.3 CHROMATIC ABERRATION

All electron sources emit electrons with some small variation  $\Delta V$  in energy. Because the focusing action of a lens depends on the electron energy, the effect of this energy variation is to cause a variation in the axial position of the first-order focus such that an otherwise perfect focus is turned into a disk of confusion with an outer radius given by

$$\delta_c = C_c \alpha \frac{\Delta V}{V} \quad (10.6)$$

where  $C_c$  is the coefficient of chromatic aberration.

The effect of chromatic aberration can be important, and this equation indicates that there is an advantage to using sources with a low value of  $\Delta V$ . The definition of  $\delta_c$  is strictly a first-order geometric one such that if we construct the three rays corresponding to  $+\Delta V$ , 0, and  $-\Delta V$ , the maximum beam radius at the position of the true focus is that given earlier.

The source that has the smallest value is the CFE source where  $\Delta V \sim 0.3$  eV. The simple Equation 10.6 for the radius of the probe can be a little misleading because it indicates the probe radius in the paraxial imaging plane for a specific electron that has a small difference of energy from the one that is brought to a focus in the paraxial imaging plane. If we consider a distribution of  $\Delta V$ , let us say a Gaussian distribution, about the average energy  $V$ , then we would expect a smooth distribution of radii in the paraxial imaging plane, and we would also anticipate that a distribution along the optic axis would be symmetrical about that plane. However, this alone does not allow us to calculate the resolution that might be obtained with this probe. To do that it would be necessary to calculate the optical transfer function (OTF) of that probe.

This whole question has been investigated by several authors (Hanai and Hibino, 1984; Rempfer and Mauck, 1988; Sato and Orloff, 1991), and the results indicate that the effective resolution of a focused probe limited only by chromatic aberration can be significantly better than  $\delta_c$ , demonstrated experimentally for ADF imaging by Nellist and Pennycook (1998), who also considered in detail the effect of chromatic aberration on the microscope transfer function. However, it is conventional to use  $\delta_c$  as a measure, inadequate though it may be.

#### 10.4.4 SPHERICAL ABERRATION

Electron lenses are far from perfect and the main effect is that they suffer from substantial spherical aberration. While to first-order all trajectories intersect at the Gaussian focal plane, the effect of the third-order spherical aberration is that they intersect this plane at a point  $C_s \alpha^3$  away from the axis where  $\alpha$  is the first-order convergence angle. For all magnetic lenses  $C_s$  is positive and therefore the extreme focus is closer to the lens than the paraxial focus. Constructing a set of such rays for values from zero to a maximum value  $\alpha$ , we can observe a caustic that is the geometric outer limit of the beam envelope. The minimum radius of this envelope occurs at a point three quarters

of the way between the focal plane and the point of focus of the extreme rays. The minimum radius is given by

$$\delta_s = \frac{1}{4} C_s \alpha^3 \quad (10.7)$$

where  $C_s$  is the aberration coefficient, which is always positive for these lenses. We can note here that although this is a simple geometric construct, we cannot expect that the intensity distribution inside the envelope will be uniform at all, and in fact it varies from that of a strong central peak at the paraxial focus to a smaller peak with a bright peripheral ring at the position of the extreme focus. We also note that the minimum beam radius does not appear at the same axial position as the first-order focus where  $\delta_{ss}$ ,  $\delta_d$ , and  $\delta_c$  are defined.

It is conventional to use  $\delta_s$  as defined in Section 10.7 to describe the resolution of a probe that is limited only by spherical aberration, but this is very misleading. For example, at the paraxial focus the strong central peak has an effective width of the order of  $0.1C_s\alpha^3$  and one presumes that features of this size could be resolved (Hanai and Hibino, 1984; Rempfer and Mauck, 1988; Sato and Orloff, 1991).

The last few years have seen dramatic advances in the technology of spherical aberration correction for TEM, and commercial correctors are now available for both STEM and CTEM instruments. We shall discuss the impact of these correctors on STEM performance later in Section 10.10.

The next logical step is to combine all four effects to determine the probe radius for any particular system. Before doing so, however, it is important to discuss the beam current because it is quite possible to form a small focused probe with such a small total current as to be useless.

## 10.5 BRIGHTNESS AND THE EFFECT OF SOURCE SIZE

Of the four factors that can limit the resolution, the effect of the source size is the easiest one to deal with, and the most obvious way to proceed is by using the concept of brightness. The brightness of the source has been defined in several different ways in the literature, but there is only one definition that has a real physical meaning and has value here (Crewe, 1987), and that is

$$\beta = \frac{1}{\text{area} \times \text{solid angle} \times \text{voltage}} = \frac{1}{\pi^2 \delta^2 \alpha^2 V} \quad (10.8)$$

where the meaning of the various terms is quite clear. The electron energy is  $-eV$  and the electrons are presumed to be emitted into a cone of semiangle  $\alpha$  from a disk of radius  $\delta$ . There may be some residual conceptual difficulties stemming from the variations from these ideal conditions for any real source, but since we are only interested in the behavior near to the axis there is no practical problem. The usual units of brightness are amperes per square centimeter steradian volts.

The value of this definition lies in the fact that this parameter is conserved throughout the optical system. This means that at any focus of the electron beam we expect the same value of  $\beta$  whatever being the imaging conditions or the value of the electron energy, providing that we use the size of the first-order image rather than any observed probe size since this may have been affected by one of the other three contributions.

The fact that brightness is conserved is based on very firm foundations. It can be considered to be a consequence of Liouville's theorem or simply due to the fact that electron motions are described by a second-order differential equation whose Wronskian is a constant. However, one views this issue as the brightness is a conserved quantity.

As a consequence of this, if we assume that the electrons have been accelerated to some new energy and passed through lenses and apertures to arrive at a new focus, then the probe current will be given by

$$I_p = \beta\pi^2(M\delta\alpha_0)^2V_0 \quad (10.9)$$

where  $M$  is the magnification,  $\alpha_0$  the semiangle of convergence of the beam, and  $V_0$  the new voltage.

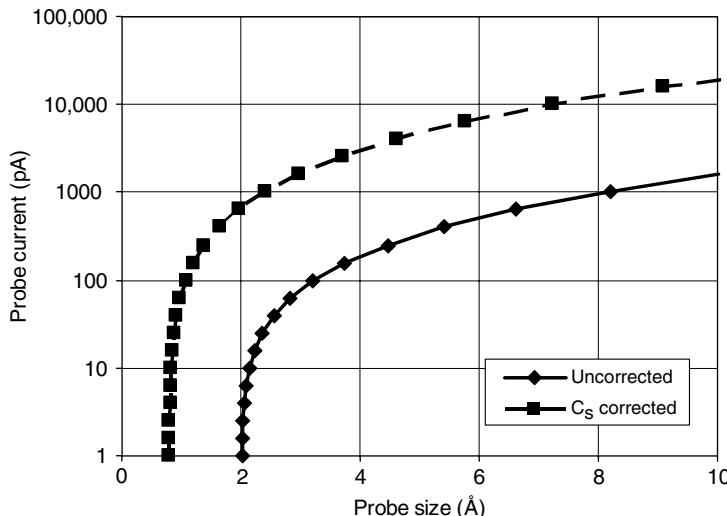
From Equation 10.9 it can be seen that we can trade off probe current for probe size. If we want a larger total current we must do with a larger probe. Also important in Equation 10.9 is  $\alpha_0$ , the semiangle of convergence of the beam, which is usually limited by the spherical aberration of the lens. Figure 10.3 shows how beam current and probe size can be traded off for a typical instrument. It is important to note that diffraction provides a fundamental limit to the probe size, and this lower limit to probe size can be seen clearly in Figure 10.3. Correction of spherical aberration (see Section 10.10) allows a significant enlargement of  $\alpha_0$  improving both the probe size and the available beam current.

If, however, we assume that we are interested in producing a focused probe that is at or near the diffraction limit, we are naturally interested in the ratio

$$\varepsilon = \frac{M\delta}{\delta_d} \quad (10.10)$$

which we would prefer to be a small number if we are interested in high resolution. As a rough guide, we could assume that the actual probe radius could be estimated by taking the two quantities in quadrature

$$\delta_{\text{eff}} = \delta_d \sqrt{(1 + \varepsilon^2)} \quad (10.11)$$



**FIGURE 10.3** A plot of probe size for incoherent imaging versus beam current for both a  $C_s$ -afflicted and  $C_s$ -corrected machine. The parameters used are 100 kV cold field-emission electron gun with  $C_s = 1.3$  mm. Note the diffraction-limited regime in which the probe size is independent of current, changing over to a source-size limited regime at large currents.

Then, when  $\epsilon$  is small, the resolution approaches the diffraction limit. Using this new parameter and Equations 10.3 and 10.6, we can then write the probe current as

$$I_p = \beta\pi^2\epsilon^2(0.61 \times 1.23 \times 10^{-7})^2 = 5.6 \times 10^{-15}\beta\epsilon^2 \quad (10.12)$$

This equation makes it very clear that we need large source brightness if we are to make  $\epsilon$  small, and in its simplicity it justifies our choice of  $\beta$  as the most appropriate definition of brightness. The brightness of a hot filament source in these units is close to unity, which makes it unsuitable for use in a STEM or SEM, and the *only* suitable source is field emission where the brightness can exceed  $10^4 \text{ A cm}^{-2} \text{ sr}^{-1} \text{ V}^{-1}$  for both types of source. Taking this value, we have

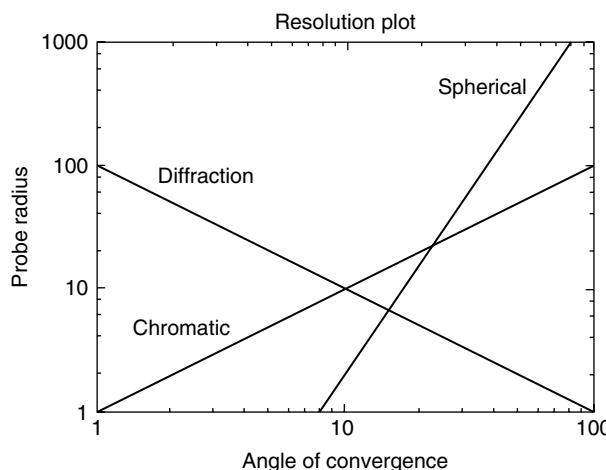
$$I_p \sim 3.4\epsilon^2 10^8 \text{ electrons/s} \quad (10.13)$$

so that even if  $\epsilon$  is 0.1 or less there is still an adequate number of electrons to form an image in a reasonable time (images at TV rates can be produced with an electron beam of a few nanoamperes).

Of more importance, however, is the significance of this argument in simplifying the design considerations of an optical column. This equation means that we can calculate the final probe current using only two pieces of information, the total linear magnification of the column and the condition that the final probe is diffraction limited. In practice, this generally means ensuring that the angle of convergence of the beam at the final focus is properly chosen.

## 10.6 OPTIMIZATION OF A FOCUSED PROBE

At the highest resolutions, the STEM is operated with a diffraction rather than source size-limited probe. In this case the remaining terms can be best appreciated by constructing what we might term a *resolution plot*. In this graph we plot the various contributions to the probe size as a function of the angle  $\alpha$ . A logarithmic plot is the most useful one in view of the various dependencies. A generic graph is shown in Figure 10.4. The diffraction line has a slope of  $-1$ , and we can never reach the region below this line since diffraction is synonymous with the uncertainty principle. Chromatic



**FIGURE 10.4** Each of the various factors affecting the probe radius has a power law dependence on the convergence angle of the electron beam at the final focus. This makes it convenient to use a logarithmic plot. The diffraction line has a slope of  $-1$  and we can never reach the region below this line. Chromatic aberration has a slope of  $+1$  whereas spherical aberration has a slope of  $+3$ . In regions where one or the other effects dominate, the resolution is determined by this plot. When two or more effects are involved we must make a careful evaluation. We cannot simply form the root mean square and assume that this number will give the resolution.

aberration is indicated with a line of slope +1 whereas the effect of spherical aberration is shown with a line of slope +3.

Although this plot allows us to gain some appreciation of the various contributions to the resolution of the microscope, it does not, in itself, allow us to calculate that resolution. The reason for this is given earlier, namely, that the various radii that are used to form the plot are of differing characters and may not even occur at the same axial position.

Several authors (von Ardenne, 1940; Colliex and Mory, 1984; Michael and Williams, 1987; Mikoyawa et al., 1994) have simply used the root mean square (RMS) values as a measure of the resolution, but this is clearly erroneous. The use of the RMS is quite appropriate for use with independent statistical errors, but even when the electron probe has a poor statistical quality the phenomena we are considering here cannot be treated as errors. In this respect it is probably a mistake to use the symbol  $\delta$  to represent both a radius and also the resolution, but this is quite common in the literature, which is why we have continued its use. However, it is quite obvious that the two are not the same. To illustrate the difference we only need to consider the case of spherical aberration. Here the symbol  $\delta$  represents the outer radius of a caustic, not the average radius or any other radius that has meaning. It is the radius of this outer disk that is described by the symbol  $\delta$ , not the much smaller radius of the bright spot, which is more nearly representative of the resolution. Furthermore, the minimum radius of the caustic does not even occur at the same position as the diffraction or chromatic aberration minima so that the simple algebraic sum has no meaning at all.

The correct way to proceed in the optimization of these effects is to describe the effect of the two aberrations in terms of a wave front defect and then calculate the radial intensity of the probe at many axial positions. Then one can conveniently take the Fourier transform of this intensity distribution to obtain the OTF. The function describes the fidelity by which the various spatial frequencies of the specimen are transmitted and from this one can define a resolution. This procedure has been used for the two primary aberrations (Rempfer and Mauck, 1988; Sato and Orloff, 1991), and the results are appropriate when diffraction is negligible. The inclusion of diffraction is described in Section 10.6.1. However, even when this process is followed it is still difficult to define a resolution because there is still some room for the interpretation of these OTF graphs, and it is not clear which one would be chosen by the operator of a microscope during the focusing process. For this reason there may be some disagreement with some of the numerical values that we present later. However, the range of disagreement is small and the reader could accommodate all of them within a 10% error range. We can now look at two extreme cases.

### 10.6.1 SPHERICAL ABERRATION PLUS DIFFRACTION

In the event that the effect of spherical aberration dominates, we need to find out how to combine it with diffraction. This has been studied by Black and Linfoot (1957) in light optics, and their work can be readily extended to electron optics (Crewe and Salzman, 1982). The result is that the best resolution can be written in several different ways

$$\delta = 0.61 \frac{\lambda}{\alpha} = 0.43 C_s^{1/4} \lambda^{3/4} = 0.15 C_s \alpha^3 = \frac{C_s^{1/4}}{V^{3/8}} \times 8.9 \times 10^{-8} \text{ m} \quad (10.14)$$

The first three expressions have self-consistent units, whereas in the last one we have written the resolution in meters and  $C_s$  is also specified in meters.

The optimum convergence angle to achieve this resolution is

$$\alpha = \left( \frac{4\lambda}{C_s} \right)^{1/4} \quad (10.15)$$

The numerical factor 0.43 in the second expression for  $\delta$  emerges from the calculation in a natural way although there is some leeway for other choices. In our case we chose 0.43 because it appeared

to be the best fit, and in addition it agrees with the value given by Haine and Mulvey (1954) several years earlier. The other two expressions are obtained by simple substitution.

At first sight this is a surprising result since the resolution is equal to  $\delta_d$  and is better than  $\delta_s$  and far better than would be given using the RMS method. However, we have already noted the different characters of the two definitions of  $\delta_d$  and  $\delta_s$ , the fact that they do not even occur in the same location and that  $\delta_s$  may not even respect the effective resolution when only  $C_s$  is involved. We can also note that identical relationships exist for the TEM with the sole difference being that the constant 0.43 is replaced by 0.78. A simple summary statement would be that in a STEM we can obtain a diffraction-limited resolution, but the maximum convergence angle is limited by the coefficient of spherical aberration.

The effects of spherical aberration are the dominant ones for all probe formation when the accelerating voltage is >5 kV or so. Where an instrument has been fitted with a spherical aberration corrector, the value of  $C_s$  can be set to zero, and the convergence semiangle  $\alpha$  can be dramatically increased. In the current commercially available correctors, it is found that the next limit is imposed by fifth-order aberrations. Even so,  $\alpha$  is usually increased by a factor of 2–3 times, reducing the diffraction-limited probe size (seen from Equation 10.14) and increasing the probe current available (seen from Equation 10.9).

### 10.6.2 CHROMATIC ABERRATION PLUS DIFFRACTION

This case has also been examined for light-optical systems (Hopkins, 1955), and the results can be carried over into electron optics. In this case the resolution is equal to the diffraction limit providing the chromatic aberration does not exceed that value. We can write

$$\delta = 0.61 \frac{\lambda}{\alpha} = C_c \alpha \frac{\Delta V}{V} = 6.3 C_c^{1/2} \lambda^{3/2} (\Delta V)^{1/2} \times 10^8 = 2.74 \times \frac{\Delta V^{1/2} C_c^{1/2}}{V^{3/4}} \times 10^{-5} \quad (10.16)$$

and we obtain the optimum convergence angle by equating the first two expressions:

$$\alpha = 2.74 \times 10^{-5} \frac{V^{1/4}}{C_c^{1/2} \Delta V^{1/2}} \quad (10.17)$$

All dimensions are in meters.

Once again, it may be surprising that a limited amount of chromatic aberration does not affect the numerical value of the resolution, as would be expected using the RMS method. However, the two radii  $\delta_d$  and  $\delta_c$  are totally different in character so that the main effect of introducing this aberration is to change the shape of the OTF slightly.

The effects of chromatic aberration are the predominant ones for all low-voltage instruments, especially in the range 0–2 kV, but they can generally be ignored in a STEM since the accelerating voltage is usually high because of the need for the electrons to penetrate that specimen.

## 10.7 APPLICATION TO THE SCANNING TRANSMISSION ELECTRON MICROSCOPE

To produce a small focused probe of electrons we need to use a low value of  $C_s$ , a high value of  $V$ , and a high value of brightness  $\beta$ . This situation is identical to that of the CTEM and we can simply use the same methods. Magnetic lenses can be designed with a  $C_s$  of 1 mm or even slightly less for electrons in the 100 kV range (e.g., Hawkes, 1982a). Using the expressions given in Section 10.6, we

calculate that this should allow the formation of probes capable of a resolution of 0.2 nm. Higher voltages allow higher resolution but the gains are slow. For example, the first STEM operated at 35 kV with a resolution of 0.24 nm (Crewe et al., 1968b) whereas recent ones operate at 300 kV with a resolution of 0.13 nm (e.g., von Harrach, 1994) or better if fitted with an aberration corrector. For a more complete discussion of the practical limits of the STEM, see, for example, (Crewe, 1995).

To design a high-performance STEM we need to produce a focused probe that is as small as possible and where the energy of the electrons is high enough to permit them to be focused into a small probe. This implies electrons with an energy of several tens of kilovolts. Taking this into account, the implications of Section 10.6 are the following. First of all, a high-resolution STEM must use one or other of the two types of field-emission source because these are the only two sources that have a brightness that is adequate for the formation of a very small electron probe. Next, the total magnification of the column must be such that the geometric image of the source can be made small compared with the diffraction-limited probe, but we will wish to control this to optimize our trade-off between probe size and current. In the case of the CFE source, which has an effective source size of  $\sim 5$  nm, this would imply a total column magnification of the order of 0.02, whereas if a TFE source is used the total magnification should be about one order of magnitude less. Next, we can assume that the effect of chromatic aberration is smaller than the effect of spherical aberration for the simple reason that we must use high-voltage electrons to produce a small focused probe, and given the small energy spread of the electrons from either of these two sources, the total effect of chromatic aberration is small. In that case, the factors controlling the size of the final focused probe are spherical aberration, diffraction, and the current required in the probe. The first two can be optimized by selecting the proper angle of convergence of the electron beam at the probe. Having found this value, it is then possible to plot the probe size versus the probe current by changing the magnification of the column (Figure 10.3). At low currents the probe becomes diffraction limited, whereas larger currents can be achieved by allowing a larger probe to be used.

The impact of spherical aberration correction can most powerfully be seen in Figure 10.3. The increased size of the convergence angle  $\alpha$  allows a smaller diffraction-limited probe, but can increase the current available in a STEM probe of given size by more than an order of magnitude.

For the highest spatial resolution work, we can say that the function of the optical column of the simplest form of a high-resolution STEM is to accelerate the electrons from a field-emission source to an energy of interest, to use a system of lenses that will produce a focus in the plane of the specimen such that the geometric image of the source is smaller than the diffraction limit of the probe and the angle of convergence is the one given in Equation 10.15. If we satisfy these conditions, then we will have a probe current given by Equation 10.12.

These are the elementary considerations for the column design, which have been discussed in several publications (Crewe, 1973, 1980, 1987). In the following sections we will examine these features in a little more detail.

## 10.8 COMPONENTS OF THE ELECTRON-OPTICAL COLUMN

### 10.8.1 THE ELECTRON SOURCE

We have already discussed the fact that we must use field-emission sources to obtain an adequate probe current and that there are two types of source, the CFE and the TFE. These two sources impose different requirements on the optical system, and, although this is not an appropriate place in which to enter into detailed discussion of this topic, it is nevertheless important that we establish some of the characteristics of these sources to appropriate the various design decisions.

### 10.8.1.1 The Cold Field Emitter Source\*

From some points of view the CFE source might be the obvious choice for a STEM since it provides us with the smallest effective source size and therefore simplifies the design of the lens system to be used in the column. With a total magnification requirement of 0.02, this can even be accomplished using only one magnetic lens. If for no other reason this makes the CFE source attractive. On the other hand, there are some characteristics that detract from its attractiveness. One major disadvantage of the CFE source is that it requires a very good vacuum system for successful operation. A vacuum of  $10^{-10}$  torr ( $10^{-8}$  Pa) is the minimum requirement.

The source itself consists of an etched tungsten point, which is formed by a heating process into a cone with a hemispherical end. When a voltage is applied to such a source, it is found that electrons are emitted when the field strength at the surface of the hemisphere is enough to allow significant tunneling of the electrons through the potential barrier (see Crewe, 1971a; see also Chapter 1). The emission current is a very strong function of the applied voltage, and it is very easy to destroy a tip by applying excessive voltage. Generally speaking, the sources that are most easily made will operate in the range of 3–5 kV with an emission current up to perhaps 50  $\mu$ A, though in the VG Microscopes they are usually operated at 5  $\mu$ A or below. This emission current is far from stable; if the voltage is held constant, it is found that the emission current steadily declines for a short period of time, remains constant for some other period of time and then increases steadily, becoming more and more erratic, and if left unattended, the source will destroy itself in one of these erratic increases. When emission begins to become erratic, the tip is *flashed* by passing a pulse of current through a filament attached to the source with enough current to heat the source to a white heat for a few milliseconds. If the pulse current is chosen correctly, the source will return to its original condition, and the process will begin over again. Initially, this process was a manual operation, which required experience and skill. More recently, programs have been written, which enable a PC to operate a CFE source automatically and indefinitely (Crewe et al., 1989; Ruan and Kapp, 1992).

The initial decline of the emission current appears to be due to the accumulation of gas molecules on the surface of the tip, which increases the potential barrier through which the tunneling takes place. This continues until a monomolecular layer has been formed at which point the increased tunneling distance is such as to cause the current to have declined by at least one order of magnitude from its initial value. The emission then remains reasonably stable until the effect of ionization in the gas surrounding the tip becomes significant. This ionization occurs because the vacuum is not perfect, and the electrons leaving the tip can ionize gas molecules, which are then attracted toward the tip because of its positive charge and deposit some energy in the tip, causing some small deformation in the surface. As time goes by, these accumulated deformations can themselves become emitters, thereby causing the erratic emission behavior. The effect of the flashing is simply to cause a local melting or annealing of the surface, which returns the tip to its original state.

In a vacuum of the order of  $10^{-10}$  torr, the initial decline can take 10–20 min, and the period of more or less constant emission can last for several hours. This can be improved by improving the vacuum conditions so that the residual pressure is better than  $10^{-10}$ . On the other hand, if the pressure is several times  $10^{-10}$  torr, then this whole process can occur so quickly that the available time for operation is very limited. For these reasons there is a great premium in any STEM using a CFE source on achieving the best possible vacuum in the vicinity of the tip.

---

\* The following brief account and indeed the very existence of the CFE source are entirely due to the efforts of the author's students Joe Wall, Mike Isaacson, and Dale Johnson. With great skill and patience and despite numerous disappointments and setbacks they never wavered in their conviction that it could be done. There are now thousands of CFE sources in use around the world and those owe a debt of gratitude to these three people.

### 10.8.1.2 The Thermally Assisted, Zirconium-Treated Field Emitter Source

The TFE or Zr/O/W Schottky emitter source has very different characteristics (see Chapter 1). The TFE source has a much narrower emission pattern than the CFE source, but the source size is greater. The brightness of the two sources is comparable. The TFE source is maintained at a high temperature, ~1800 K (a bright red heat), and this appears to serve two purposes. First of all, it allows the zirconium to migrate over the surface of the tip and should any be lost it is easily replenished from a reservoir that is generally placed at the junction between the heating filament and the tip. In addition to this, the tungsten tip itself is in a viscous state, so that the damage from returning ions is annealed quickly enough that there is no permanent damage to the surface (see Chapter 1). As a result, the operation can be continuous and can continue for several thousand hours. In addition to this, the vacuum requirements are less stringent than those of the CFE source by approximately one order of magnitude, so it is quite possible to operate a TFE source unattended in a vacuum of  $10^{-9}$  torr for several thousand hours. In addition to these, the fluctuations of the emission current are considerably smaller in the TFE source than in the CFE source and so from many points of view this source is preferable. Operationally, it is a much more attractive source than the CFE. Its major disadvantages are the fact that the source size is greater and therefore the demagnification required is also greater, and this can necessitate an additional magnetic lens in the column; in addition, the energy spread of the electrons is greater by about a factor of at least two. This can be important if chromatic aberration should become a factor, for example, when operating at low accelerating voltages, when the STEM is used for analytical purposes, and when energy loss data are required with a greater precision than 0.5 eV.

## 10.8.2 THE ELECTRON GUN

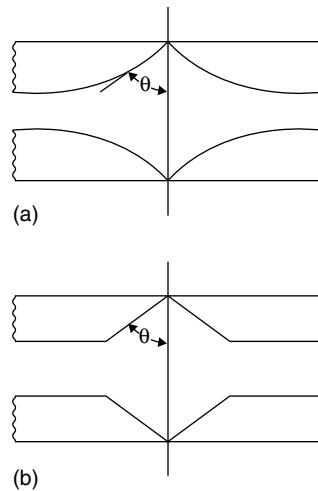
The characteristics of field-emission sources are significantly different from those of other sources, and, consequently, the conventional electron gun is inappropriate.

When using a CFE source the electrons will have energy in the range of 3–5 kV, and when using a TFE source it can range between 3 and 8 kV. In either case this is inadequate for a STEM where we normally require electrons in the energy range 50–300 kV. We therefore need an electron gun that will accept electrons in the 3–8 kV range and produce an electron beam in the 50–300 kV range with characteristics suitable for the rest of the electron-optical column.

The simplest way to satisfy these requirements is to use two electrodes. A potential  $V_1$  is applied between the tip and the first electrode, and by varying this extraction voltage we can control the emission current from the tip. A voltage  $V_0$  is applied between the tip and the second electrode and this serves to define the energy of the electrons at that point. Each of these electrodes must have an aperture to allow the electron beam to pass through.

It would be quite possible to use two flat parallel electrodes, but the difficulty here is that the focusing action of this arrangement would be controlled more by the apertures than by the spacing of the electrodes. This in itself is not an impossible situation, but it is difficult to make apertures very precisely. For this reason in the electron guns that are in common use an attempt is made to reduce the effect of the aperture. The first successful gun to accomplish this was the Butler gun. This electron gun has been adequately described elsewhere (Butler, 1966), but the main feature of the gun was that the potential distribution along the axis was used to determine the physical shape of the electrodes rather than the more usual approach in which the electrodes are chosen first and then the potential distribution is calculated.

In the Butler gun we consider an axial potential distribution between two points,  $z = 0$  and 1. The potential should have a value of  $V_1$  at  $z = 0$  and  $V_0$  at  $z = 1$ . Next we set the potential gradient to be zero at these two axial points. The simplest function that will satisfy all these conditions is a cubic expression in  $z$  from which one can calculate the shape of the electrodes. They are shown in Figure 10.5 where it can be seen that in the region near the aperture, the electrodes are tapered at an angle of approximately  $\arctan \sqrt{2}$ . This choice is made because it is the angle that reduces the field strength to zero in the aperture plane at the apex of the cone.



**FIGURE 10.5** The upper diagram (a) is a scale drawing of the electrode shape for a butler gun. The curved surfaces of the electrodes are the equipotentials for the cubic potential. The semiangle of the conical region  $\theta$  has a value of  $\arctan \sqrt{2}$ , which serves to reduce the field strength to zero at the apex. A small hole can be drilled to allow passage of the electron beam with minimal disruption of the potential distribution. In the lower diagram (b) we indicate how this primary shape can be modified to make the gun easier to fabricate. The potential along the axis in the two guns is virtually the same.

It is not clear whether or not this detailed shaping of the electrodes is necessary, and more recent versions of the gun simply use parallel plates but retain the tapered cone in the region of the aperture (Figure 10.5b).

This type of electron gun functions quite well (Crewe et al., 1968a), and a two-electrode gun has been used in all of the 100 kV VG Microscopes CFE guns. In some column designs it has one major disadvantage that for a fixed value of  $V_1$ , the position of the image of the tip will vary as  $V_0$  is varied, and this can be quite inconvenient if the image happens to fall inside one of the subsequent lenses.

This effect can be controlled by the addition of a third electrode placed between the other two.\* Such an approach was used in the 300 kV HB603 instruments of VG Microscopes. By applying a potential  $V_f$  to this electrode, we can control the position of the first-order focus of the electron gun, and in fact it is quite possible to keep the position of the focus fixed along the optical axis no matter what values of  $V_1$  and  $V_0$  are chosen. There are the two preferred operating conditions for such an electron gun, and there does not appear to be any great preference for one or the other. The first obvious choice is to place the image of the tip near to the exit of the gun in some fixed position so that it can be reimaged onto the specimen with a series of additional lenses. The other solution is to arrange the focusing electrodes so that the beam that emerges from the electron gun is approximately parallel to the axis, that is, to say the image of the tip is somewhere near infinity. This again allows the use of additional lenses to image the tip on the preferred final location. There is no electron-optical reason to prefer one condition over the other.

Many attempts have been made to reduce the aberration coefficients of the gun, but it is difficult to interpret the results. These coefficients are a strong function of the two focal lengths, the physical length of the gun and the internal field strength. The only reasonable comparison would be to compare the values of the coefficients at the same axial position for the same value of the gun magnification, but this is usually impossible to do. However, we can note that the best arrangement is to use small

---

\* The author is indebted to David A. Crewe who has calculated the optical properties of literally hundreds of electron guns and whose results make this summary possible.

entrance and exit apertures, small gaps between the electrodes, and a long central electrode with a wide bore (Crewe et al., 1989). Orloff (1989) has used a different geometry with some success.

When this gun is followed by a magnetic lens to form a small probe, the effect of the gun aberrations is generally small. A lens magnification  $M$  reduces the effective chromatic aberration of the gun by a factor of  $M^2$  and the spherical aberration by a factor of  $M^4$ , and since we have some freedom in choosing  $M$ , the effect of the gun aberrations can be made negligible.

### 10.8.3 THE MAGNETIC LENSES

The electron beam emerging from the electron gun must be brought to a focus on the specimen, and magnetic lenses are normally used for this purpose because the aberrations of magnetic lenses are lower than those of electrostatic lenses in the energy range that is suitable for a STEM. These lenses are no different from those in a TEM with the possible exception that they may be considered to be *upside down*. That is to say, if the magnetic lens is asymmetrical along the z-axis, the asymmetry would be in the opposite direction.

The reason for this is to be found in the reciprocity theorem, which is well known in light optics and can be applied directly to electron optics (Born and Wolf, 1980). In a TEM we see that a very small area on the specimen is imaged to a much larger area in the final image plane, and the smallest feature that can be resolved in the specimen is given by all the considerations described earlier. In a STEM, a macroscopic source is imaged onto the specimen to illuminate a very small area. In this sense, the entire optical column of a STEM can be thought of as being an upside-down version of a TEM column, and, indeed, some versions of these instruments have been built with the source at the bottom and the specimen at the top. This analogy also indicates that a STEM column can be simpler than that of a TEM because the required demagnification of the size of the source is only a few nanometers at the most, whereas the detector resolution in a TEM may be several microns. The reader is referred to any of the other standard publications on magnetic lenses if further information is needed, for example, (Hawkes (1982b) and accompanying Chapter 4 by Tsuno.

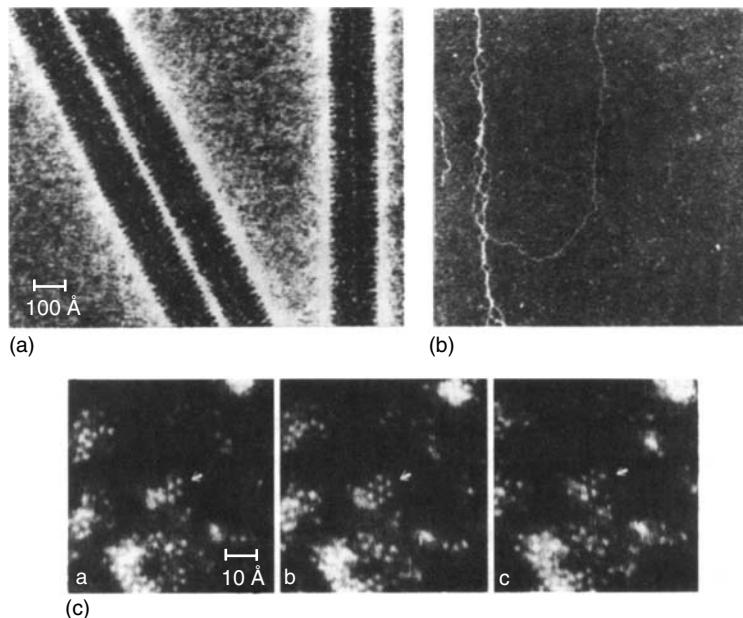
### 10.8.4 COLUMN DESIGNS AND PERFORMANCE

It is useful to examine the design of a complete STEM column from the point of view of the users and the constraints or the range of operating conditions that they would like to impose.

There are many such requirements depending on the type of investigation to be carried out. For example, the STEM could be viewed as just another microscope for obtaining images at high resolution, and, indeed, that was the initial goal. However, it can produce high-resolution images with attributes very different to those of the CTEM, and it is this alongside its analytical capability that attracts the user. Materials scientists have also become interested in its diffraction and interference capabilities (Nellist and Rodenburg, 1995) as well as in the direct images, and for the biological user the strong Z-contrast and elemental analysis is important. These interests vary from user to user and for this reason there are a range of possibilities for the design and operation.

A primitive STEM can be built using very few components. The first STEM used a simple electron gun with only two electrodes, the Butler gun referred to in Section 10.8.2, and there was a single fixed aperture in the plane of the first anode (Crewe et al., 1968a). This was followed by a single magnetic lens. In spite of the primitive nature of this instrument, it was competitive with the highest resolution TEM instruments of the time, and, for the reasons discussed in Section 10.3 the image contrast was very much higher. The instrument was operated at 30–35 kV with a resolution of 0.24 nm. The image-forming system will be described later, but the contrast was high enough that unstained biological objects could be easily recognized and even single-stranded DNA was imaged. Single heavy atoms were imaged for the first time (Crewe et al., 1970), and later the stability was found to be good enough to obtain sequences of the same atoms, which allowed studies of motion of atoms (Isaacson et al., 1977).

Some examples of these images are shown in Figure 10.6. These images are shown here to demonstrate that the probe-forming electron optics of the STEM can be calculated with some precision



**FIGURE 10.6** (a) Images of stained tobacco mosaic virus are similar to those of a transmission electron microscope, but the contrast is greater and there are no interference effects. (Unpublished micrograph by Ohtsuki, M.) (b) Unstained DNA. The U-shaped loop is single-stranded DNA (full-scale dimension 670 nm). (Unpublished micrograph by Wall, J.) (c) Uranium atoms on a thin carbon film. The images were taken at intervals of 17 s, and many individual atoms can be seen through the strong Z-contrast of the annular dark-field detector. The small crystal that forms appears to be that of  $\text{UO}_2$  and contains seven atoms with a spacing of 0.34 nm. (Unpublished micrograph by Ohtsuki, M.)

and that the performance of the instrument is in accordance with the calculations. The resolution and contrast in even a simple version of the STEM are of practical value. The images shown in Figure 10.6 were obtained using the annular detector. The image of stained tobacco mosaic virus shown in Figure 10.6a clearly demonstrates that the STEM can produce images of high quality and good contrast. The detail is as good as that in a TEM with the added advantage that there are no interference artifacts. The image of unstained DNA, shown in Figure 10.6b, shows that the STEM can produce images with much higher contrast than a TEM. The U-shaped strand was shown to be single-stranded DNA using the mass measurement techniques described earlier.

It was the DNA image that provided the encouragement for attempting to image single atoms. One of the best of these is the sequence of three shown in Figure 10.6c, taken 17 s apart. The specimen was a very dilute solution of uranyl acetate on a very thin carbon film. The acetate is converted into oxide by the intense irradiation, and one small crystal of just seven atoms exists as a perfect hexagon for a short time. The spacing is consistent with that of a small  $\text{UO}_2$  crystal.

All the images in Figure 10.6 were taken using signals obtained directly from the microscope. There was no subsequent image processing.

The flexibility of this particular instrument was severely limited. For any given source the value of  $V_1$  is well defined since the source can only be operated over a very small range of voltages, in this case  $\sim 3.5$  kV. The fact that the aperture in the gun was fixed meant that the only way to set the beam radius to the proper value in the magnetic lens was to adjust the accelerating voltage  $V_0$ . There are only two possibilities for achieving this condition, one with the gun focus below the specimen and one above, and since the former corresponded to an uncomfortably low voltage, there was only one practical solution and that was to operate at 30–35 kV.

This primitive STEM was built with only one purpose in mind and that was to demonstrate the principles of operation of this type of instrument in a way that would make it possible to make a

comparison with the most advanced forms of TEM. From a functional point of view this meant that there were only three essential variable parameters,  $V_1$ ,  $V_0$ , and the lens current, which allowed the operator to set only three electron-optical parameters, the probe current, the radius of the beam in the objective lens, and the final focus on the specimen.

For the more sophisticated uses of electron microscopes, there is a need to establish more electron-optical or experimental parameters; in that case we obviously need more independent adjustments, and this generally means providing more components.

The addition of another magnetic lens (the condenser lens) between the electron gun and the objective lens is another simple addition, and it allows the operator to control the diameter of the electron beam almost independently of the value of  $V_0$  because wherever the gun focus may be it can be refocused in the proper plane. The only exception to this is that the gun focus may occur near the center of this lens, making it inoperative. As a practical matter, this only affects the operation of medium-voltage SEM instruments.

The use of a fixed aperture in the gun limits the probe current that is available to the operator, and this may severely limit the analytical power of the instrument. Whenever a fixed aperture is used it would be expected to be the one appropriate for the highest resolution, and this implies the smallest aperture and, consequently, the smallest probe current. The maximum value of the probe current would then correspond to the value given earlier. To allow higher current probe formation, it is advisable to make the defining aperture in the electron gun large enough to permit the extraction of the largest value of the probe current that is desired and then to provide a variable aperture elsewhere in the column to limit that current and to define the beam size in the final lens. It is a common practice in the TEM to place that variable aperture in the objective lens itself, but this is not the ideal place for it since the optimum value of the beam radius may not correspond to the size of any available physical aperture, and its electron bombardment will generate x-rays thereby confusing any EDX analysis; therefore, some compromise may be necessary. In modern instruments, the beam limiting aperture is found in a variety of locations, including in the objective lens, in one of the condenser lenses, or between the gun and the condenser lenses.

The use of the STEM for analytical purposes presents another set of constraints. In this case the operator may wish to illuminate some small portion of the specimen with a well-defined and uniformly illuminated disk, perhaps with a parallel beam or at least a well-defined angle of convergence. This can be done by forming an image of the aperture (real or virtual) on the specimen, and to do so while still retaining the other capabilities requires one more lens. Venables and Cox (1987) have proposed a three-lens system for this purpose and have suggested that as many as four lenses might be required. The most recent STEM designs, for example, from Nion Co. or the FEI Titan instrument have made use of three-lens designs along with additional focusing optics (either a further anode or a magnetic gun lens) in the gun.

### 10.8.5 OTHER FEATURES OF THE OPTICAL COLUMN

To form a scanned image the focused probe of electrons is moved across the specimen in a rectangular raster scan, and to avoid introducing other aberrations it is necessary to pivot the electron beam about the coma-free point of the lens (see, e.g., Crewe and Parker, 1976). This is normally close to the position of the principal plane of the lens, quite close to the specimen, and this makes it necessary to use a double set of scan coils to produce the necessary deflection. The pivot point of the double-scan system can be determined by predetermining the number of turns in each scan coil and adjusting the current accordingly; but in this case it is only possible to fix the pivot point in space, which does not allow for any subsequent change of the position of the coma-free point such as might occur as the excitation of the lens is varied. A better scheme is to arrange for the currents in the two sets of scan coils to be independently adjustable so that the position of the pivot point can be set by the operator and changed if necessary. When the pivot point of the beam is in the coma-free plane,

then the probe size is independent of the size of the raster scan itself. In principle, it is possible to correct both isotropic coma by controlling the position of the pivot point (Crewe and Parker, 1976).

In instruments where the physical aperture is placed at or near to the coma-free point, it should be made adjustable in position and size. The position adjustment is normally done mechanically and the size adjustment by arranging a series of apertures mounted on a strip so that the proper one can be selected. However, it has been noted previously that this method suffers from the disadvantage that fine control of the beam diameter is not generally possible.

This may be a good point to discuss the vacuum requirements of the column. We have already seen that the vacuum level for a CFE source must be close to  $10^{-10}$  torr whereas that for a TFE source can be  $\sim 10^{-9}$  torr. However, this limitation extends only to the region near to the source. It is by no means necessary to reach these vacuum levels in the remainder of the column. For example, in the electron gun itself the vacuum level could be significantly different from that near the tip if we use differential pumping techniques, as is found currently on all STEM instruments. However, it is convenient to extend the ultrahigh vacuum throughout the electron gun. For the remainder of the electron-optical column, the vacuum requirements are by no means as stringent, with the possible exception of the vacuum level around the specimen itself, and here the vacuum level should be determined by the need to reduce or eliminate specimen contamination problems.

## 10.9 SCANNING TRANSMISSION ELECTRON MICROSCOPE TYPES

The preceding discussion has centered on the problem of focusing electrons from a source into the smallest possible focused probe or into a very well-defined one. An instrument of this type is called a dedicated STEM since the design is exclusively devoted to this problem. A dedicated STEM almost invariably uses a field-emission source, an electron gun of variable voltage, followed by one, two, or three magnetic lenses. The final focused probe depends on the choice of parameters, principally the accelerating voltage and the focal length of the objective lenses. All other parameters have minor influence on the final probe, although great care must be taken when striving for the ultimate resolution. In von Harrach (1994) there is a complete discussion of all these factors in conjunction with the design of the 300 kV HB603 made by VG Microscopes. For a discussion of the limits of resolution, see Crewe (1995).

Following the success of the dedicated STEM instruments several manufacturers of conventional TEMs have modified their instrument to allow operation in a STEM mode. The TEM has an electron-optical system between the source and the specimen, which is normally used to produce the illumination. In modern TEM/STEM instruments this usual TFE is followed by an electron gun and one or two condenser lenses. In normal operation these lenses are used to produce a parallel beam inside the objective lens, which then illuminates the specimen. However, it is quite possible to readjust the electron optics so that an image of the source is formed at the specimen. If one then adds a scanning system between the electron gun and the objective lens, one has the elements of the probe-forming system of a STEM.

## 10.10 CORRECTION OF SPHERICAL ABERRATION IN SCANNING TRANSMISSION ELECTRON MICROSCOPE

We have seen that a major limiting factor for STEM performance is the relatively small convergence angle  $\alpha_0$  allowed because of the spherical aberration of the lens system. Although Scherzer (1936) pointed out that spherical aberration was inherent to round electron optics, he also suggested that an optical system based on nonround lenses could be used to generate negative spherical aberration to compensate for that of the round lenses (Scherzer, 1949). In recent times, such spherical aberration correctors have become a commercial reality, and in STEM dramatic improvements in resolution have been seen (Nellist et al., 2004; Batson et al., 2002). Figure 10.3 also emphasizes

the improvements in probe size and beam current that result from a spherical aberration corrector. In this section we will explore some of the practicalities of installing aberration correctors in a STEM.

Compared with CTEM, developing a spherical aberration corrector for a STEM is simplified by two practical issues. (i) The double deflection scan coils can be placed between the corrector and the main imaging lens. This means that the beam is not scanned in the corrector, and is essentially always axial. Only the so-called axial aberrations need to be considered, unlike the use of a corrector in CTEM that needs to be able to cope with off-axial beams. (ii) The additional optics of the aberration corrector will add to the chromatic aberration of the lens system. As noted earlier, placing of the optics before the sample, along with the narrow energy spread of the field-emission guns used, means that the impact of the chromatic aberration is reduced in STEM compared with TEM. Nonetheless, the addition of a spherical aberration corrector to the STEM column adds some requirements. In particular, the aberration corrector will expect rays to enter it at a specific angle; that is to say with the beam focused on a specific crossover position. If one also wants to be able to control the convergence angle arising from a fixed aperture and the overall source demagnification while maintaining the crossover requirement of the corrector, three condenser lenses are required. Also, perhaps obviously, the resolution benefits of the aberration corrector will only be realized if the mechanical stability of the column and sample, and the electrical stability of the lens power supplies, allow it. In addition to the development of the corrector optics, much work has also gone into improving the microscope stability.

Shortly after the development of the high-resolution STEM, the author realized the potential of aberration correctors. Technology using both octopole–quadrupole and sextupole-based systems (see Chapter 12) was developed, and great strides forward were made. Despite all these activities, it was not possible to improve the spatial resolution over that achieved in an uncorrected instrument. Indeed, it was not until around the turn of the century that correction of aberration started to actually make improvements to the resolution of a microscope. One might ask why it took almost 60 years from the first proposals by Scherzer. The answer probably lies with the critical technological advances that were necessary. (i) The pole pieces of the correctors must be machined with an extremely high tolerance, and modern workshop instruments are required to achieve this. (ii) Despite our best efforts, there will still be inaccuracies in the mechanical alignment of the multipole element, leading to parasitic aberrations. It is necessary that the multipoles are designed with the flexibility and required extra windings to be able to null out the parasitic aberrations. (iii) Even to correct up to third-order aberrations (spherical aberration is one type of third-order aberration), it is necessary to keep track of 12 aberrations. Because of the need to null the parasitic aberrations, there may be around 30–40 separate power supplies. To measure the lens aberrations, and to feedback to necessary adjustment to the power supplies is a complex process, and it has required the development of computer control of the power supplies coupled with automated measurement of the aberrations to achieve accurate aberration correction. It is only recently that desktop computers with the required computational power have become available.

In recent years, STEM has become rapidly more popular, and many research laboratories now contain an instrument capable of atomic resolution STEM. The recent success of spherical aberration correction with STEM, has only served to further increase interest, and it now appears that STEM is well founded as an advanced microscopy technique.

## ACKNOWLEDGMENTS

The author is deeply indebted to all the students, graduate, undergraduate, and postdoctoral, who have worked in his laboratory. The knowledge and understanding that is contained in this chapter is largely due to their efforts. No advisor was ever blessed with a better group.

A deep debt of gratitude is also due to the Department of Energy and its predecessors, who have bravely supported this work for 30 years or so. Although there have been many individuals who

have been involved, we pay particular respects to Gerry Goldstein and Roland Hirsch. The author cannot refrain from saying with some degree of certainty that if peer review processes had been invoked in the early years, the STEM would not now exist.

Finally, the author is indebted to his friends and colleagues who provided information and reprints. It is most unfortunate that space did not permit the use and citation of this vast amount of information. The kind respondents included Philip Batson, John Cowley, Andreas Engel, Peter Hawkes, Mike Isaacson, Dale Johnson, David Joy, John Langmore, Jon Orloff, Steve Pennycook, Harald Rose, Andrew Somlyo, John Spence, J. Venables, Sebastian von Harrach, Joe Wall, and Elmar Zeitler.

## REFERENCES

- Batson, P. E., Dellby, N. and Krivanek, O. L. 2002. Sub-ångstrom resolution using aberration corrected electron optics, *Nature*, 418, 617–620.
- Black, G. and Linfoot, E. H. 1957. Spherical aberration and the information content of optical images, *Proc. R. Soc. Lond.*, A239, 522–540.
- Born, M. and Wolf, E. 1980. *Principles of Optics*, 6th ed., Pergamon Press, New York, p. 396.
- Butler, J. W. *Proc. 6th Int. Congr. for Electron Microsc.*, Kyoto (1960).
- Chapman, J. N., Ploessl, R. and Donnet, D. M. 1992. Differential phase contrast microscopy of magnetic materials, *Ultramicroscopy*, 47, 331–338.
- Chisholm, M. F. and Pennycook, S. J. 2006. Direct imaging of dislocation core structures by Z-contrast STEM, *Philos. Mag.*, 86, 4699–4725.
- Colliex, C. and Mory, C. 1984. Quantitative aspects of scanning transmission electron microscopy, in *Quantitative Electron Microscopy*, M. H. Jacobs, G. W. Lorimer and P. Doig, Eds, The Metals Society, London, p. 47.
- Cowley, J. M. 1969. Image contrast in a transmission electron microscope. *Appl. Phys. Lett.*, 15, 58–59.
- Cowley, J. M. 1976. Scanning transmission electron microscopy of thin specimens, *Ultramicroscopy*, 2, 3–16.
- Cowley, J. M. 1995. Scanning transmission electron microscopy, *Handbook of Microscopy*, 2, 563.
- Crewe, A. V. 1963. A new kind of scanning microscope, *J. Microsc.*, 2, 369.
- Crewe, A. V. 1966. A transmission scanning microscope, *Proc. 6th Int. Congr. for Elect. Microsc.*, Kyoto, (1960) 631.
- Crewe, A. V. 1970. The current state of high resolution scanning electron microscopy, *Q. Rev. Biophys.*, 3, 1–29.
- Crewe, A. V. 1971a. A high resolution scanning electron microscope, *Sci. Am.*, 224, 26–35.
- Crewe, A. V. 1971b. High intensity electron sources and scanning electron microscopy, *Electron Microscopy in Material Science*, Academic Press, New York, pp. 162–207.
- Crewe, A. V. 1973. Production of electron probes using a field emission source, in *Progress in Optics*, Vol. 11, E. Wolf, Ed., Braunschweig, North Holland, pp. 223–246.
- Crewe, A. V. 1980. The physics of the high resolution scanning microscope, *Rep. Prog. Phys.*, 43, 621–639.
- Crewe, A. V. 1987. Optimization of small electron probes, *Ultramicroscopy*, 23, 159–168.
- Crewe, A. V. 1995. Limits of electron probe formation, *J. Microsc.*, 178(1), 93–100.
- Crewe, A. V., Clayton, D. L., Crewe, D. A. and Moscicka, K. 1989. User-friendly field emission, *Proc. 47th Annu. EMSA Mtg.*, San Antonio.
- Crewe, A. V., Eggenberger, D. N., Wall, J. and Welter, L. M. 1968a. Electron gun using a field emission source, *Rev. Sci. Instrum.*, 39, 576–583.
- Crewe, A. V. and Parker, N. W. 1976. Correction of third order aberrations in the scanning electron microscope, *Optik*, 46(2), 183–194.
- Crewe, A. V. and Salzman, D. 1982. On the optimum resolution for a corrected STEM, *Ultramicroscopy*, 9, 373–378.
- Crewe, A. V. and Wall, J. 1970. Contrast methods in a high resolution scanning microscope—the use of multiple detectors, *7th Int. Congr. of Electron Microsc.*, Grenoble, 1, 35–36.
- Crewe, A. V., Wall, J. and Langmore, J. 1970. Visibility of single atoms, *Science*, 168, 1338–1340.
- Crewe, A. V., Wall, J. and Welter, L. M. 1968b. A high-resolution scanning transmission electron microscope, *J. Appl. Phys.*, 30, 5861–5868.
- Dekkers, N. H. and Lang, H. de. 1974. Differential phase contrast in a STEM, *Optik*, 41, 452–456.

- Dellby, N., Krivanek, O. L., Nellist, P. D., Batson, P. E. and Lupini, A. R., 2001. Progress in aberration-corrected scanning transmission electron microscopy, *J. Electron Microsc.*, 50, 177–185.
- Gomer, R. 1993. *Field Emission and Field Ionization*, American Institute of Physics (Am. Vacuum Soc. Classic), New York.
- Haine, M. E. and Mulvey, T. 1954. The application and limitations of the edge-diffraction test for astigmatism in the electron microscope, *J. Sci. Instrum.*, 31, 325.
- Hammel, M., Kohl, H. and Rose, H. 1990. *Int. Congr. on Electron Microsc.*, Vol. 1, Seattle, 120–121.
- Hanai, T. and Hibino, M. 1984. Effect of the spherical aberration on electron probe size, *J. Electron Microsc.*, 33(2), 116–122.
- Hawkes, P. W. 1978. Half-plane apertures in TEM, split detectors in STEM and ptychography, *J. Optics (Paris)*, 9, 235–241.
- Hawkes, P. W. 1982a. *Magnetic Electron Lenses*, Springer-Verlag, Berlin.
- Hawkes, P. W. 1982b. *Principles of Electron Optics*, Vol. 3, Springer-Verlag, Berlin.
- Hawkes, P. W. and Kasper, E. 1994. *Principles of Electron Optics*, Vol. 3, Academic Press, San Diego, CA.
- Hembree, G. G. and Venables, J. A. 1992. Nanometer resolution scanning Auger electron microscope, *Ultramicroscopy*, 49, 109.
- Hopkins, H. H. 1955. The frequency response of a defocused optical system, *Proc. R. Soc. Lond.*, A231, 91–103.
- Isaacson, M., Kopf, D., Utlaut, M., Parker, N. W. and Crewe, A. V. 1977. Direct observation of atomic diffusion by scanning transmission electron microscopy, *Proc. Natl Acad. Sci. USA*, 74, 1802–1806.
- Knoll, M. and Ruska, E. 1932. Electron microscope, *Z. Phys.*, 78, 318–339.
- Michael, J. R. and Williams, D. B. 1987. A consistent definition of probe size and spatial resolution in the analytical electron microscope, *J. Microsc.*, 147(3), 289–303.
- Mikoyawa, T., Kazumori, H., Nakagawa, S. and Nielsen, C. 1994. Ultra-high resolution semi-in-lens type FE-SEM, JSM-6320F with strong magnetic-field lens with built-in secondary-electron detector, *Proc. EMSA* (New Orleans), 52, 484–485.
- Mullejans, H., Bleloch, A. L. and Howie, A. 1993. Secondary-electron coincidence detection and time-of-flight spectroscopy, *Ultramicroscopy*, 52, 360–368.
- Muller, D. A., Tzou, Y., Raj, R. and Silcox, J. 1993. Mapping  $sp^2$  and  $sp^3$  states of carbon at sub-nanometre spatial resolution, *Nature*, 366, 725–727.
- Nellist, P. D. 2007. Scanning transmission electron microscopy, *Science of Microscopy*, Springer, New York.
- Nellist, P. D., Chisholm, M. F., Dellby, N., Krivanek, O. L., Murfitt, M. F., Szilagyi, Z., Lupini, A. R., Borisevich, A., Sides, W. H. Jr., Pennycook, S. J. 2004. Direct sub-angstrom imaging of a crystal lattice, *Science*, 305, 1741–1741.
- Nellist, P. D. and Pennycook, S. J. 1998. Subangstrom resolution by underfocussed incoherent transmission electron microscopy. *Phys. Rev. Lett.*, 81, 4156–4159.
- Nellist, P. D. and Rodenburg, J. M. 1995. Resolution beyond the information limit in transmission electron microscopy. *Nature*, 374, 630–632.
- Orloff, J. 1988. Survey of electron sources for high-resolution microscopy, *Ultramicroscopy*, 28, 88–97.
- Orloff, J. 1989. Analysis of an electrostatic gun with a Schottky cathode, *J. Vac. Sci.*, B7, 1870.
- Rempfer, G. and Mauck, M. 1988. A closer look at the effect of lens aberrations and object size on the intensity distribution and resolution in electron optics. *J. Appl. Phys.*, 63(7), 2187–2199.
- Rose, H. 1975. Theory of image formation in electron microscopy I, *Optik*, 42, 217–244.
- Ruan, S. and Kapp, O. H. 1992. An automatic field-emission tip conditioning system, *Rev. Sci. Instrum.*, 63, 4056–4060.
- Sato, M. and Orloff, J. 1991. A method for calculating the current density of charged particle beams and the effect of finite source size and spherical and chromatic aberrations on the focusing characteristics, *J. Vac. Sci. Technol.*, B9(5), 2602–2608.
- Scheinfein, M. R., Drucker, J. and Weiss, J. K. 1993. Secondary-electron production pathways determined by coincidence electron spectroscopy, *Phys. Rev.*, B47(7), 4068–4071.
- Scherzer, O. 1936. Über einige Fehler von Elektronenlinsen, *Z. Phys.*, 101, 593–603.
- Scherzer, O. 1949. Sphärische und Chromatische Korrektur von Elektronen-Linsen, *Optik*, 2, 114–132.
- Spence, J. C. H. and Cowley, J. M. 1978. Lattice imaging in STEM, *Optik*, 50, 129–142.
- Venables, J. A. and Cox, G. 1987. Computer modelling of field emission gun scanning electron microscope columns, *Ultramicroscopy*, 21, 33–46.
- von Ardenne, M. 1938. The electron scanning microscope: Practical construction, *Z. Tech. Phys.*, 19, 407–416.
- von Ardenne, M. 1940. Über das Auftreten von Schwarzungslinien bei der elektronenmikroskopischen Abbildung Kristalliner Lamellen, *Z. Phys.*, 116, 736–738.

- von Harrach, H. S. 1994. Medium voltage field-emission STEM—The ultimate AEM, *Microsc. Microstruct. Microanal.*, 5, 153–164.
- Voreades, D. 1976. Secondary electron emission from thin carbon films, *Surf. Sci.*, 60, 325–348.
- Voreades, D. and Crewe, A. V. 1974. Correlation between inelastic electron scattering and secondary electron production, *Proc. 32nd Annu. EMSA Mtg.*, 428–429.
- Voyles, P. M. and Muller, D. A., 2002. Atomic-scale imaging of individual dopant atoms and clusters in highly *n*-type bulk Si, *Nature*, 416, 826–829.
- Zeitler, E. and Thomson, M. G. R. 1970. Scanning transmission electron microscopy, I, *Optik*, 31, 258–280.
- Zworykin, V. K., Hillier, J. and Snyder, R. L. 1942. A scanning electron microscope, *ASTM Bull.*, 117, 15–23.

---

# 11 Focused Ion Beams

*M. Utlaut*

## CONTENTS

11.1	Introduction.....	523
11.2	A Short History .....	524
11.3	Uses of Focused Ion Beam: The Exploitation of Destruction.....	525
11.4	Ion-Sample Interactions .....	526
11.5	System Architecture Considerations.....	534
11.6	Focused Ion Beam Microscopy: Ions and Electrons.....	539
11.7	General Imaging Considerations .....	539
11.8	The Vexations of Charging Samples.....	544
11.9	Applications of Focused Ion Beam .....	549
11.9.1	Micromachining and Gas-Assisted Etching .....	549
11.9.2	Gas-Assisted Deposition of Materials .....	559
11.9.3	Scanning Ion Microscopy .....	567
11.9.4	Secondary-Ion Mass Spectrometry.....	571
11.9.5	Focused Ion Beam Implantation.....	581
11.9.6	Lithography .....	591
11.10	New Directions.....	593
11.11	The Future .....	594
	For Further Information.....	595
	Acknowledgments.....	595
A.1	Appendix: What is Boustrophedonic? .....	595
	References .....	596

## 11.1 INTRODUCTION

There is an old adage, perhaps from Confucius, that claims, “Theory without practice is futile. Practice without theory is fatal.” The latter half of this advice may appear a bit dramatic right up to the instant that one is in the middle of using an instrument that is not performing as expected, and only an intimate knowledge of how the instrument was designed and manufactured, and a knowledge of its interactions with its environment and whatever the instrument is working on can bring a solution to the encountered problem without spoiling the execution of the task at hand. Therefore, to alleviate both futility and fatality, we will cover both the theory of why focused ion beam (FIB) machines are the way they are and how they work, and also what is known of what to do with them. This is equivalent to knowledge being a luxury until it becomes necessary. Typically, in a treatment such as this, theoretical aspects of operation are first covered as a groundwork for describing the known applications. We could of course go the other way and first describe the uses of FIB, and then trace into what form these machines have evolved to satisfy those uses. The paradigm for an emerging technology usually is that *function follows form* (i.e., exploration), whereas for a reasonably

advanced evolving technology, the paradigm is that *form follows function* (i.e., refinement). As FIB based on liquid metal ion sources (LMISs) is now virtually all the way through the 30-year development phase that a new technology experiences, which determines what its predominate uses will be, it has now evolved along the lines of form following function. At present, the function is being defined and driven by requirements not only from the semiconductor industry, but also from industries as varied as mining, pharmaceuticals, and academic needs in biology. We will follow a mixed approach, in which we will briefly cover the constraints of history, nature, and demand, to derive the formation and evolution of FIB, and then lay a foundation of theory guided by what we now see FIB being used for, and then we will cover those applications in detail.

There have been eleven fairly comprehensive reviews of FIB technology and its applications in the analysis and modification of semiconductors. A chronological ordering of these reviews reveals the evolution of applications of the machines. Melngailis (1987, 1991) covered ion beam lithography and direct maskless implantation, whereas Harriott (1989, 1991) covered mask repair and secondary ion mass spectroscopy (SIMS), Namba (1989) and Nikawa (1991) reviewed the applications for semiconductors, involving sputtering and induced deposition. Orloff (1993) reviewed the entire history and evolution of the field, and the first book was devoted to FIB, by Prewett and Mair (1991), which covered in depth the theory and operation of the sources used in FIB and primarily the applications of mask repair, lithography, and implantation. In recent years, three books have appeared, which cover all aspects of the use of FIB. Orloff et al. (2003) and Giannuzzi and Stevie (2005) have substantially different approaches to the field, whereas the most recent contribution to the field by Yao (2007) is a comprehensive compendium of known FIB system applications. We have borrowed liberally from these works and the published and unpublished works of many researchers now in the field. Many bright people have spent precious parts of their lives, trying to figure out how parts of these systems work, or how to make or use the systems, and as it is usual in any field, nothing happens unless someone does it.

## 11.2 A SHORT HISTORY

Any short history is necessarily flawed, inadequate, and an adumbration as a result of lack of depth. We cover here only the most major milestones. FIB is a fairly old technology, which has among its roots scanning electron beam probes, which have been in use for nearly five decades, and ion thrusters, which is a relatively obscure space exploration technology. In the 1960s, systems were developed capable of producing  $\sim 0.1 \mu\text{m}$  ion beam probes of a few picoampères from low-brightness plasma sources (Hill, 1968). In the late 1960s, Crewe (Crewe et al., 1968; Crewe, 1973), at the University of Chicago, showed the importance of high-brightness electron field emission sources and low-aberration optical systems for obtaining very high-resolution atomic images with electrons. In the early 1970s, Levi-Setti (1974) at the University of Chicago and Orloff and Swanson (1975) at the Oregon Graduate Center (OGC) developed ion systems based on gas field ionization sources (GFIS), which have a small intrinsic source size, producing focused beam currents of about 10 pA, and an imaging resolution of about 50 nm. As the maximum beam current was so small, and because of the inherent temperamental nature and necessity of cryogenic cooling of these sources, this approach was abandoned for the most part in favor of the LMIS invented by Clampitt et al. (1975) in England. With the invention of a long-lived, stable LMIS by Jergenson (1982) in Santa Barbara, the first focusing columns were built by Krohn and Ringo (1975) at Argonne (who used a different style of LMIS) and by Seliger et al. (1979) at Hughes Research Laboratories (HRL), and many groups then followed these leads. In the early 1980s, Levi-Setti et al. (1984), in cooperation with Seliger and Ward, took the lead in developing a high-resolution system for microscopy and surface analysis, whereas Wagner (Atwood et al., 1984) at IBM began to use FIB for mask repair. Meanwhile, Gamo et al. (1980) in Japan, Wang et al. (1981) at HRL, and Ahmed's group (Cleaver and Ahmed, 1981) at Cambridge University built mass-separated columns whereas the development of LMIS with multiple components (based on the Jergenson design) (Clark et al., 1987) made it possible to perform direct

maskless implantation of devices, that were investigated by Reuss (Reuss, 1985; Reuss et al., 1986) at Motorola and Rensch et al. (1987) and Clark et al. (1988) at HRL and Ahmed's group (Evanson et al., 1988) at Cambridge in a variety of applications in Si and GaAs. In the mid-1980s, Ishitani et al. (1985) at Hitachi pioneered the use of FIB for integrated circuit (IC) cross sectioning, whereas Kubena et al. (1988a,b) studied ion lithography at HRL, and then based on an optics column design of Orloff and Swanson (1984), Doherty (1985), and Ward et al. (1985) of Micron Corp. and Tuggle and Swanson of FEI Co. with Boylan and Ward (Boylan et al., 1989) at Intel began to use FIB for commercial use in mask repair and IC production failure analysis and reconfiguration. Sudraud et al. (1988) at CNRS combined FIB with a scanning electron microscope, the first *dual beam* system used as a *microcircuit surgery* tool (Melngailis et al., 1986). Levi-Setti (Chabala et al., 1994) has continued to investigate the applications of FIB with SIMS and ion microscopy (taking the technology into biological investigations), Harriott (Harriot et al., 1986, 1987, 1990; Harriott, 1989) at Bell Labs has touched all aspects of instrumentation and application, whereas Crow (1992) at Intel has pursued the use of FIB in device modification, transmission electron microscope (TEM) sample preparation, failure analysis, and SIMS. At present, the field has become a mature and mainstay technology for the preparation of TEM samples, failure analysis for semiconductors and MEMS, and circuit edit procedures. The recent efforts of Ward (Notte and Ward, 2006) in producing a working GFIS have pushed the frontiers of ion microscopy to the limit. For applications, requiring very high currents or non-Gallium ions, Smith et al. (2005) have explored new plasma source configurations. Most of the topic covered in this chapter is based on work done with LMIS. Virtually every topic covered here is pertinent and relevant to other sources such as the GFIS and the plasma source.

### 11.3 USES OF FOCUSED ION BEAM: THE EXPLOITATION OF DESTRUCTION

FIB is a tool that, like many, does not have just one use. The multifaceted capability endears it to users and researchers, but causes heartburn and headache to designers and builders and debuggers. From outside, it appears like a scanning electron microscope (Figure 11.1), but instead of electrons, massive ions are made to do work on a variety of targets, and the interaction of ions with matter differs tremendously from the interactions of electrons with matter, and this is where a knowledge of the difference between scanning probes of electrons and ions should start.

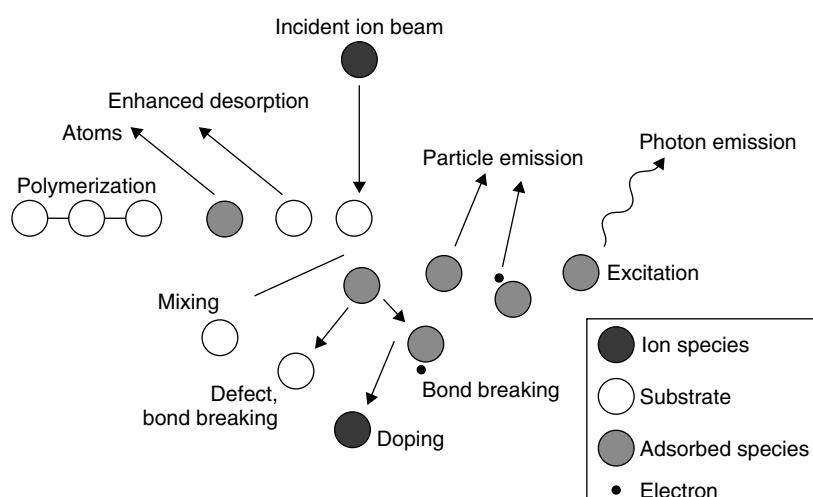


**FIGURE 11.1** (See color insert following page 340.) A state-of-the-art FIB system. (Courtesy of FEI Company.)

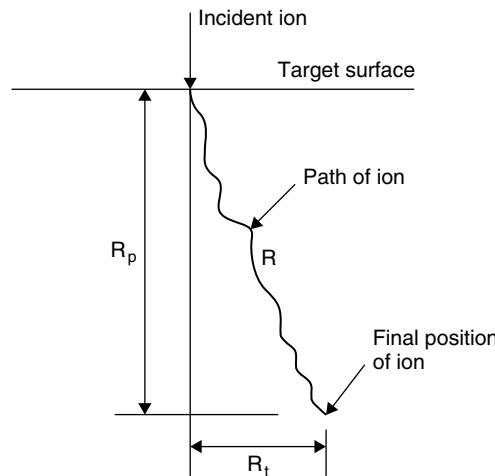
FIBs have been used for defect repair of optical and x-ray masks and direct maskless implantation of devices, and are presently in great demand for use in circuit and device modifications, failure analysis of circuits and devices, circuit and device cross sectioning, TEM sample preparation, microscopy, SIMS, MEMS, and general nanotechnology applications. With FIB, it is possible on a submicron scale reaching down to the several tens of nanometers scale, to remove selectively a predetermined amount and kind of material from a selected volume of a target, or to deposit at points of choice certain metals or insulators. Holes can be opened and refilled. By carefully removing volumes of material, one can use the FIB as a scalpel to carve out tiny, thin slices of material for examination in other instruments. Tiny wires can be fabricated to connect parts of ICs together that were left out of a design, or connections can be broken. By producing ions of high-enough energy, it is possible to implant directly the regions in devices, which cannot be done in a reasonable manner using normal masked implantation. Since some fraction of the sputtered atoms from the sample leave as ions, they can be collected to form an image of the sample with different information than secondary electrons, or they can be collected by a secondary optics system to give high spatial resolution chemical information of the sample through secondary ion mass spectroscopy (SIMS). Because of the recent advances in the theoretical understanding of source physics, ion optics, ion–gas chemistry, and image formation, both the technique of manufacturing systems and the daily operation of them have become very sophisticated. What were once academic curiosities have become necessary to various industries, and as the mass of common intellect of FIB continues to grow, more applications will be found.

#### 11.4 ION-SAMPLE INTERACTIONS

As FIBs produce focused beams (in diameter from a few nanometer to a few micrometer) with currents ranging from a fraction of a picoampere to 100 nA of positive ions of energies between 1 and 50 kV, it is instructive to have an overview of the main interactions that these ions have with the samples onto which they impinge. Figure 11.2 shows the generic interactions that incident ions produce. The most useful interactions are the removal of sample atoms through the mechanism of sputtering by which a small fraction of the sample atoms leave ionized, the production of secondary electrons from the sample and the ability of the primary ion beam to initiate surface reactions so that material may either be deposited from any surrounding gas or the gas may assist in selective and accelerated etching.



**FIGURE 11.2** Possible interactions of the incident ions in the beam with a solid sample.



**FIGURE 11.3** The geometric relationship between the range  $R$ , the projected range  $R_p$ , and the transverse straggling distance  $R_t$  for an ion impinging on a sample.

As the primary ions in the beam enter the sample, they lose energy through two main channels, and come to rest after all of their initial energy is lost. The path of a typical ion is shown in Figure 11.3 in which the range  $R$  is the total path length the ion takes as it losses energy, which is usually longer than the projected range,  $R_p$ , the final resting depth of the ion. The transverse straggle,  $R_t$ , is the sideways projection of the range. As the ions traverse into the sample, they loss energy through either *nuclear losses* or *electronic losses*. Nuclear losses are due to collisions between the screened nuclear charges of the incident ion and target atoms, whereas electronic losses are due to incident ions, interacting with lattice electrons of the sample, resulting in excitation and ionization.

The most useful interaction in the FIB, physical sputtering, used to modify the structural landscape of the sample by removing material, which occurs when the incident ions transfer sufficient momentum to surface or near-surface atoms to free one or more target atoms. The sputtering process can be chemically modulated with the addition of gasses into the system whereby chemical sputtering is initiated for increased or selective removal of material. Concomitant with material removal is the generation of sample information through the production of secondary ions and electrons from the sample. These secondary charged particles can be used to form images of the sample. Secondary ion emission occurs when surface atoms are ionized and sufficiently energized to be emitted from the surface. Secondary electrons are emitted mostly from ion bombardment from metallic surfaces, and their yield per incident ion is typically 10–1000 times greater than secondary ion yield.

Calculating ion–solid interactions can be difficult, but fortunately to know the range and the distribution of the extent to which ions have *straggled* both longitudinally and transversely from the mean range, there exists a good simulation with which to estimate these quantities. Transport of ions in matter (TRIM) (Ziegler, 1991) is a computer program using Monte Carlo simulations that calculate the three-dimensional penetration of ions into matter. Examples of TRIM results are shown in Figures 11.4 and 11.5. This simulation is also fairly accurate for estimates of the sputter yield (target atoms sputtered per incident ion) and damage incurred by the sample as ions impinge onto it. The program is useful for engineering solutions for specific applications, as both the sputter yield and damage are functions of incident ion species, energy, and angle. Such effects are shown in Figures 11.6 through 11.10, in which the incident ion type, energy, and angle are changed. Table 11.1 shows the results of TRIM calculations for  $\text{Ga}^+$  incident normally into Si (see also Table 11.2).

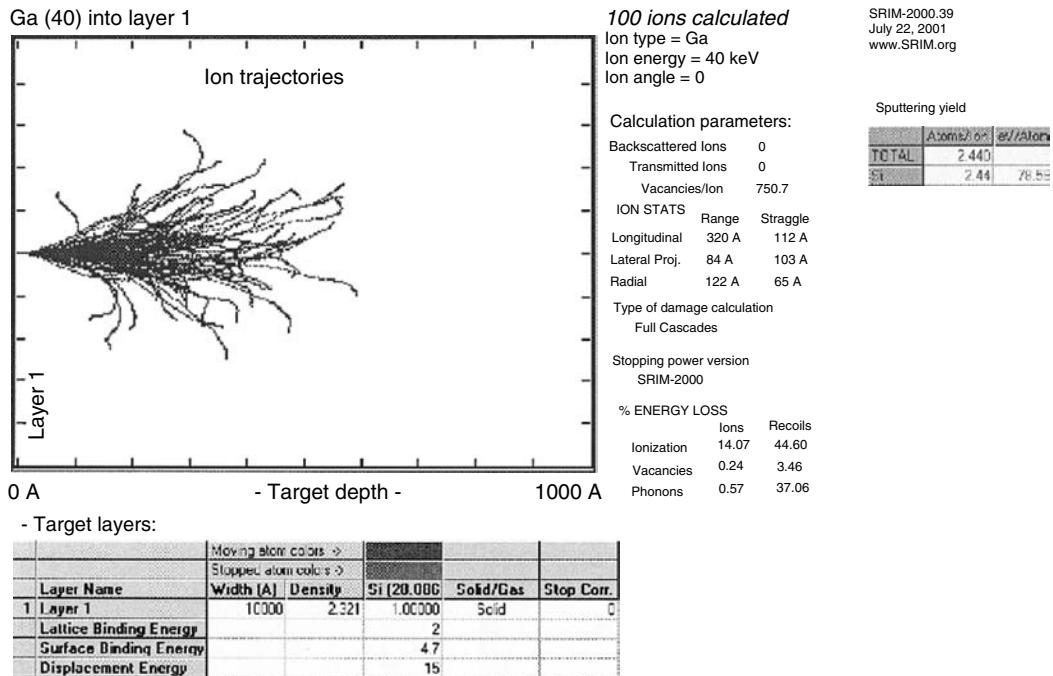


FIGURE 11.4 Example of a plot of the longitudinal distribution of ions from a TRIM calculation.

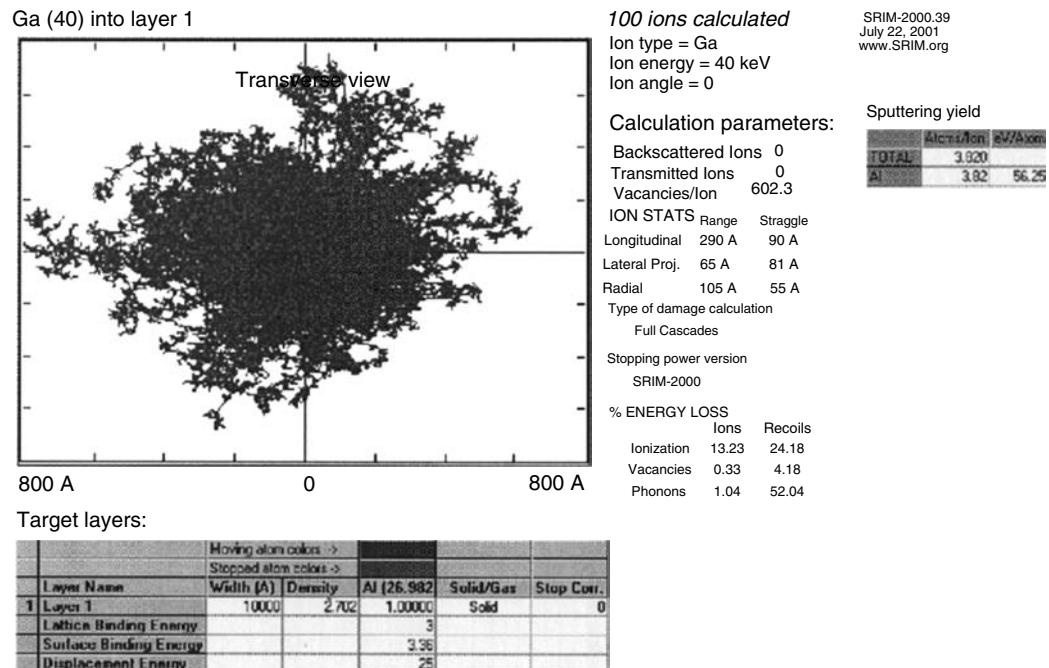
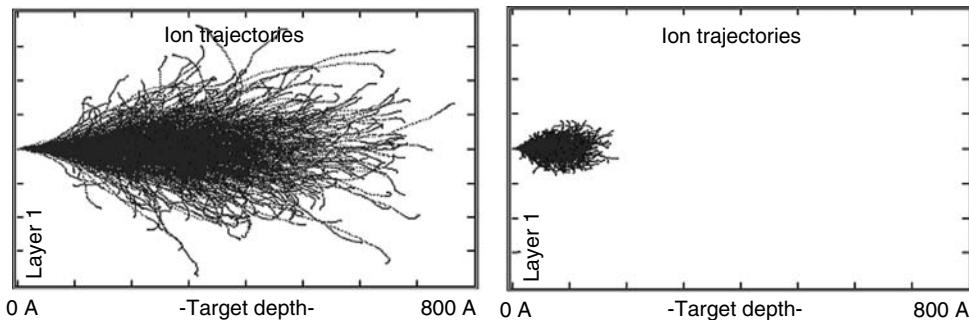
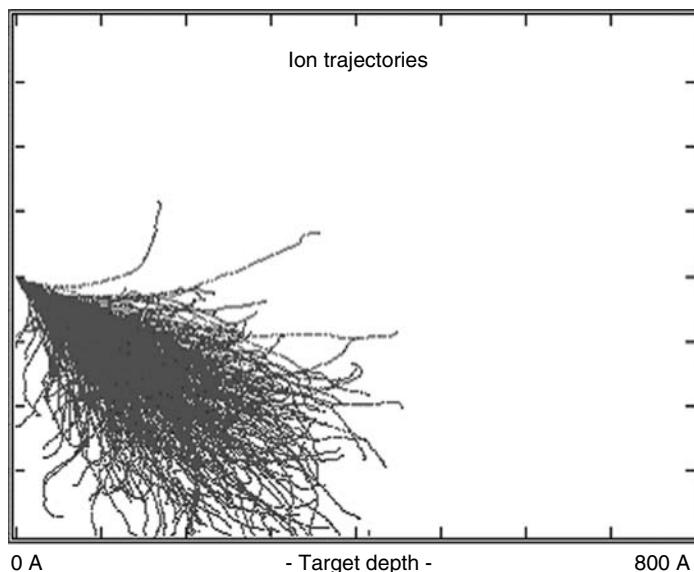


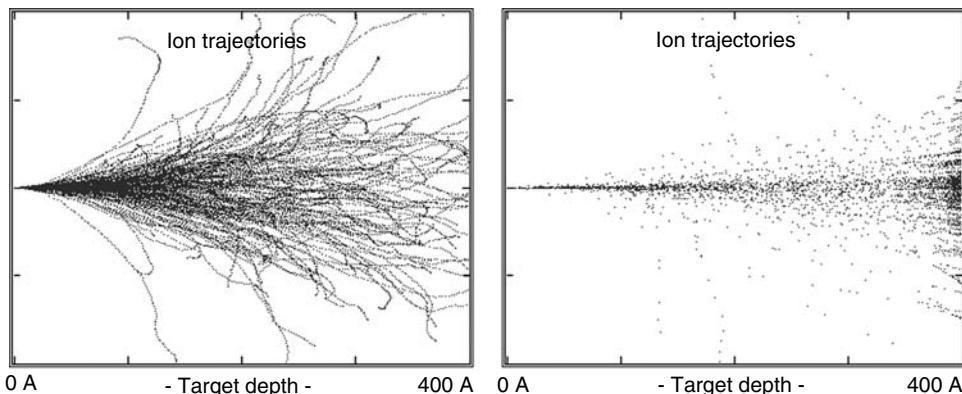
FIGURE 11.5 Example of a plot of the transverse distribution of ions from a TRIM calculation.



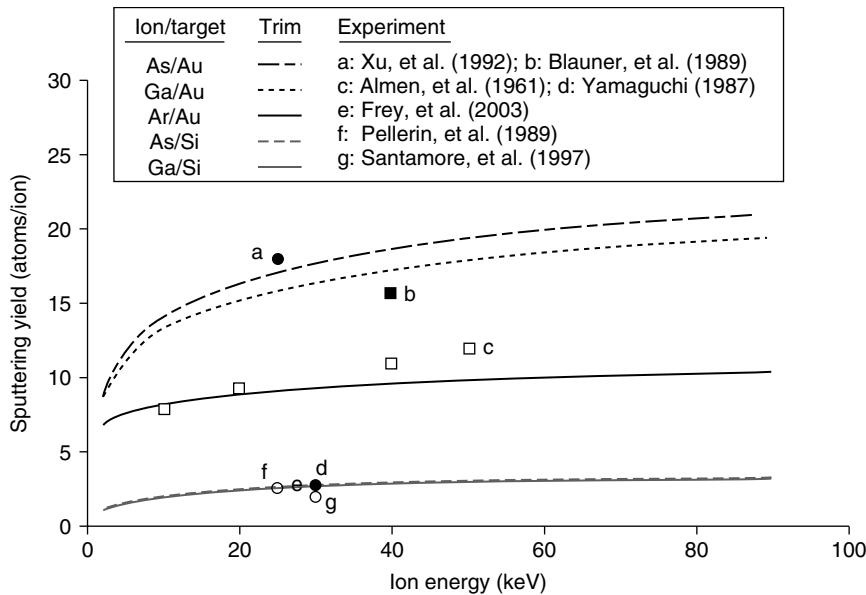
**FIGURE 11.6** TRIM calculations, showing the difference of energy. The left plot is of 40 keV  $\text{Ga}^+$  ions into Si. The right plot is for 5 keV  $\text{Ga}^+$  ions into Si.



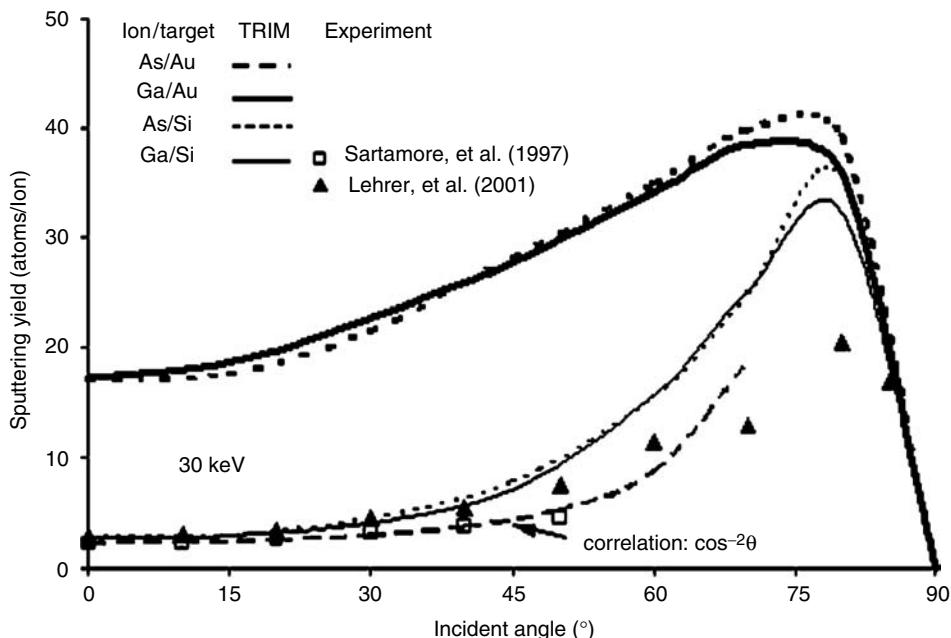
**FIGURE 11.7** TRIM simulation plot, showing the effect of angle for 40 keV  $\text{Ga}^+$  ions into Si at 45°. It is clear why the sputter yield increases as many more ions are near the surface.



**FIGURE 11.8** TRIM simulation plots, showing the effect of mass. The left plot is 30 keV  $\text{Ga}^+$  into Si, whereas the right plot is for 30 keV  $\text{Be}^+$ .



**FIGURE 11.9** Sputter yield as a function of incident ion energy for several incident ion-target combinations.



**FIGURE 11.10** Sputter yield as a function of incident angle.

Since the primary ion beam carries energy to the sample, some damage to the sample is incurred by the beam. The damage is in the form of rearrangement of the sample and is a strong function of incident ion type, energy, and angle. Such an effect is dramatically seen in Figure 11.11 in which a significant decrease in the damage layer in Si by impinging Ga is clearly evident with lower incident ion energy. The measured damage layers agree well with TRIM calculations.

**TABLE 11.1**  
**TRIM Calculated Quantities.  $\text{Ga}^+ \rightarrow \text{Si}$**

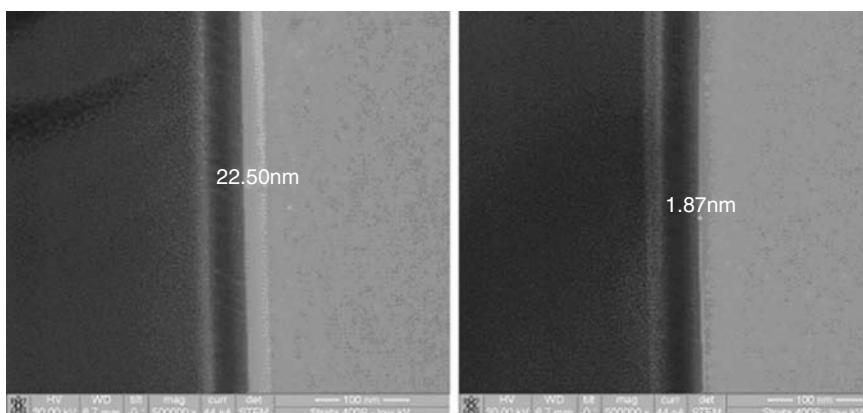
keV	Vacancy/Ion	$R_p$ (nm)	$\Delta R_p$ (nm)	$\Delta R_t$ (nm)	S (Atom/Ion)	eV/Atom
50	1177	41.1	14.8	11.6	2.3	80
40	967	34.9	12.6	9.6	2.4	64
30	748	28.0	9.9	6.3	2.5	52
20	516	20.3	7.4	4.6	2.4	41
10	274	12.9	4.6	3.8	1.9	35
5	144	8.3	3.0	2.6	1.6	23
2	61	4.9	1.7	1.6	1.1	17
1	32	3.5	1.2	1.2	0.6	13

Note: 0° (from normal) incidence.

**TABLE 11.2**  
**Comparison of Ranges, Straggles Sputter Yields, and Energy Transferred to Si Target Atoms for Several Different Incident Ions, 0° (from Normal) Incidence**

30 keV	Range	Longitudinal Straggle	Lateral Range	Lateral Straggle	Sputter	eV/Atom
He	280	79	82	102	0.03	112
Be	158	58	57	70	0.11	25
Al	53.1	23.1	15.8	20.2	1.3	144
Ga	28.1	10.1	6.4	8.3	2.6	39
In	24.8	7.3	5.1	6.4	3.6	39
Xe	24	6.5	4.2	5.3	3.4	30
Bi	22.8	4.5	3.5	4.4	3.8	37

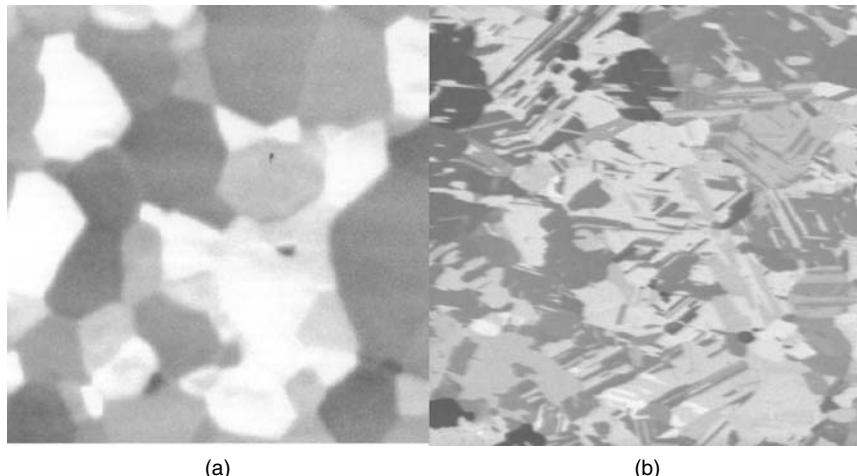
Note: All distances are in nanometer. Results are from TRIM.



**FIGURE 11.11** (See color insert following page 340.) STEM images of damage layers produced by 30 keV (left) and 2 keV (right).

**TABLE 11.3**  
**Beam Sizes (for a Typical System) and Ion Ranges for Several Beam Currents at Two Beam Energies for  $\text{Ga}^+$  into Si**

Beam Energy	Beam Current	Beam Size (nm)	Ion Range (nm)
30 keV	21 nA	274	28
	1 nA	44	28
	100 pA	24	28
2 keV	7 nA	1935	4.9
	1 nA	593	4.9
	10 pA	92	4.9



**FIGURE 11.12** Examples of ion channeling. In (a), 30 keV  $\text{Ga}^+$  impinges onto Al, whereas in (b) 40 keV  $\text{Ga}^+$  impinges into Cu. The field of view for each image is 8  $\mu\text{m}$ . The dark areas in each image are where the ions penetrate (channel) deep into the sample. In those cases, secondary electrons produced deep within the sample cannot reach the surface to be collected to form an image.

There is an inverse relationship between the size of the FIB and damage layer produced. One can produce a smaller beam with increasing energy, but the impinging ions inflict more damage. At a given energy, more beam current results in a larger radial beam size. As the data of  $\text{Ga}^+$  shown in Table 11.3, the beam size and damage depth become similar at about 100 pA for 30 keV Ga, and the beam size is always appreciatively larger ( $\sim 20\text{--}400\times$ ) than the damage layer for 2 keV Ga into Si. If one constructs a similar table for a heavier ion such as Bi, the beam size would nearly double (due to chromatic effects) and the ensuing damage would be reduced by  $\sim 40\text{--}50\%$ . For most applications (where absolute beam size matters), the use of such a source would be problematic. In almost all imaginable milling applications, Ga is the source of choice due to its ease of use.

One assumption that most calculations and simulations of ion–solid interactions use is that the sample is amorphous. In crystalline materials, a phenomenon known as channeling is possible, whereby incident ions have correlated collisions with target atoms if the ions enter in a direction of a low-indexed axis, and steering of the ions due to open channels can result in ranges several times the maximum range in nonsteering directions or amorphous materials. In this case, electronic losses determine the range, and there is little straggling (hence the term “channeling”), and near the end of their paths nuclear collisions with lattice atoms dominate the energy loss to stop the ions. This effect is shown in Figure 11.12.

The final ion–sample interaction, we will discuss, is the production of secondary electrons. These electrons are used to form images so that the FIB acts as a microscope much like a SEM. The interaction of ions with samples is very different than electron interactions with samples so that different information is available through the FIB-generated images. With impinging ions, secondary electrons are generated very near to the surface.

When ions impinge on the target, secondary electrons and ions are created. Secondary electron production is several orders of magnitude larger than the secondary ion yield. Typically for  $\text{Ga}^+$  ions the secondary-ion ionization efficiency is  $\sim 0.001\%$  which is five to six orders of magnitude lower than the secondary-electron production channel. Although secondary electrons are more useful for imaging, in some cases, they are concomitantly responsible for the majority of sample charging. Since the sample charges positively due to both the impinging positive ions and the expelled negative secondary electrons, some of the secondary electrons are attracted back to the sample reducing the magnitude of charging, but also reducing the number of information bearing secondary electrons available for imaging.

As the primary beam of ions interacts with target atom outer shell electrons, some of the electrons receive sufficient energy to be ejected from an atom. If they have sufficient energy and are within the secondary electron escape depth  $\lambda$  from the surface, they can overcome the surface barrier energy and be ejected from the sample. The secondary electron yield  $\delta$ , which is the ratio of the number of electrons produced per incident ion is given by

$$\delta = \int \frac{E}{\epsilon R} Be^{-x/\lambda} dx = \frac{BE}{\epsilon R} \lambda (1 - e^{-R/\lambda}) \quad (11.1)$$

where

$E$  and  $R$  = primary ion energy and range

$\epsilon$  = energy necessary for producing a secondary electron

$B$  = characteristic of the target material

To be useful, this formula requires knowledge of  $\lambda$ , which in practice is estimated from the relationship:

$$\delta = 1.11 \delta_m \left( \frac{E}{E_m} \right)^{-0.35} (1 - e^{-2.3(E/E_m)^{1.35}}) \quad (11.2)$$

where  $\delta_m$  and  $E_m$  are parameters tabulated for some materials of interest. As an example, one can calculate from tabulated data that for 40 keV  $\text{Ga}^+$  (Table 11.4).

Although these calculated secondary yields show great promise for imaging purposes, they are in fact considerably higher than the experimental values obtained in most FIB systems. The most likely reason for the discrepancy is due to surface sample charging. For example, the typical

---

**TABLE 11.4**  
**Parameters and Associated Values for Estimating  
Secondary-Electron Yields**

Material	$(B/\epsilon)$	$\lambda (\text{nm})$	$R (\text{nm})$	$\Delta$
Si	11.5	3.2	33	5
$\text{SiO}_2$	16	5.8	30	12
Interlayer dielectric (ILD)	9.3	6.2	30	8

---

numbers for many materials are around  $\delta = 2$ . A proper knowledge of the useful secondary yields is necessary if rational designs of system imaging parameters are to be expected.

The secondary yield is not the only parameter of interest for secondary electrons. The energy and angular distributions are also important. Typically, the energy distribution of the emitted electrons has a fairly broad maximum approximately 4–8 eV (the tail can extend to  $\sim 50$  eV), and a useful relationship for the number of electrons per unit energy is

$$\frac{dN}{dE} = k \frac{(E - E_F - \Phi)}{(E - E_F)^4} \quad (11.3)$$

where

$k$  = material dependent constant

$E_F$  = Fermi energy

$\Phi$  = work function

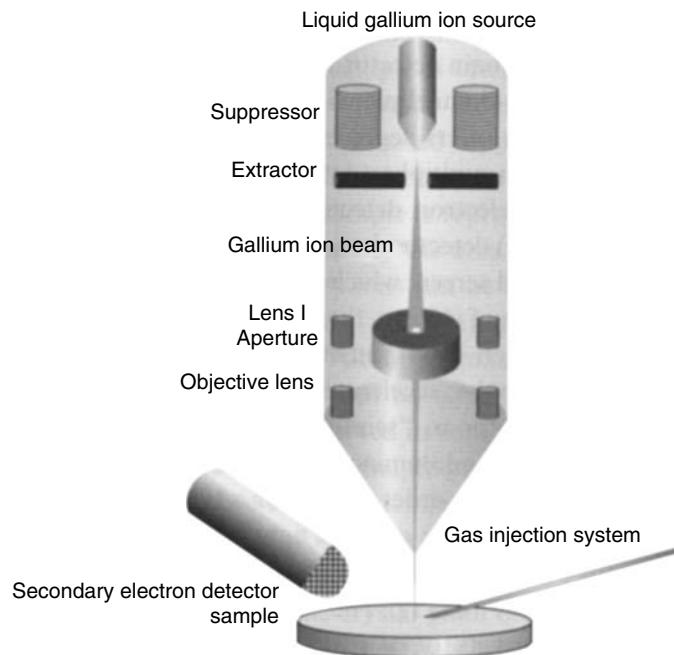
The angular distribution of emitted secondary electrons follows a cosine distribution measured normal to the target surface as

$$N(\theta) = N_0 \cos\theta \quad (11.4)$$

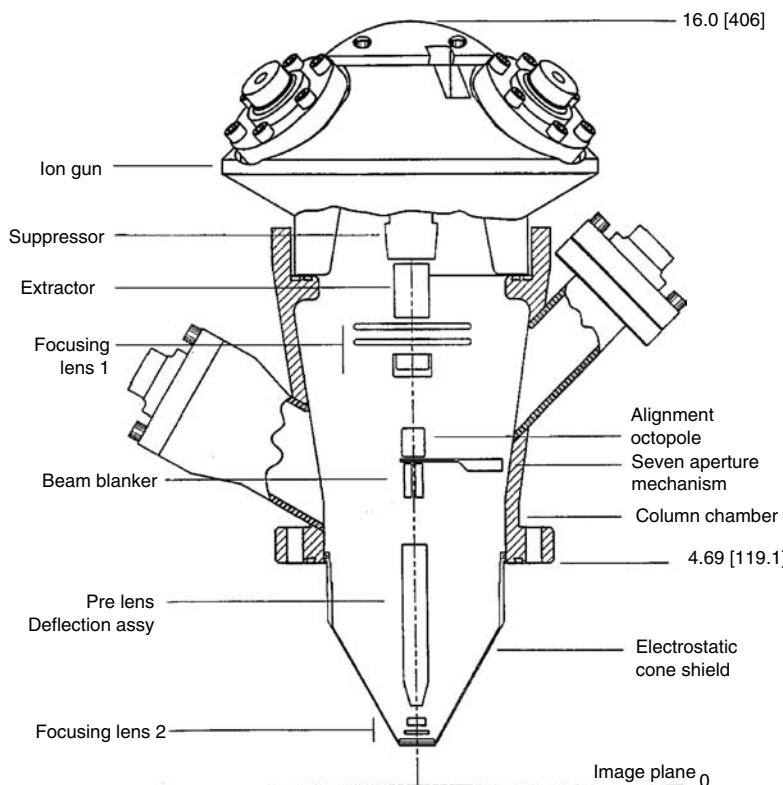
## 11.5 SYSTEM ARCHITECTURE CONSIDERATIONS

The first focusing columns built for LMIS consisted of a single electrostatic lens, a beam blanker, and a single octupole for scanning. These were attached to a vacuum chamber that had a movable stage and some kind of detector to form images. As the systems evolved for use at higher resolution, the column evolved to add a second and, in some cases a third lens. For LMIS that were multispecies, Wien ( $E \times B$ ) mass filters were added to provide a variety of different ions for the user. A schematic diagram of a typical two-lens system is shown in Figure 11.13. The elements of a typical two-lens focusing column are shown in Figure 11.14. For multiple species ion sources, some kind of mass selection is necessary, and the usual type of *filter* employed is the Wien velocity filter. One such version of a mass-separated column is shown in Figure 11.15. Details of the action of the mass filter can be found in Seliger (1972).

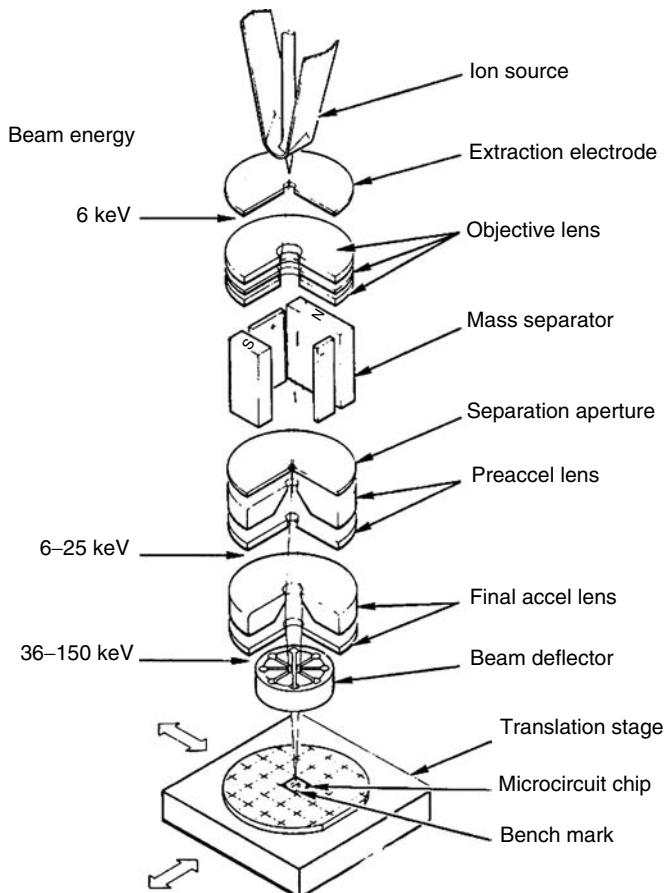
The theory and design of the LMIS and optical lenses for the system are covered in Chapters 2 and 5. There is one aspect of system design and performance that needs special attention. Coulomb interactions of ions near the source and in the beam as it traverses the length of the column play an important role in limiting the performance of the beam's size and shape (column effects are covered in Chapter 7). The Coulomb force is the second strongest force in the Universe, and the interion spacing can be fairly small at some positions in the column, causing repulsive effects to expand the beam in size as well as introducing chromatic effects that limit focusing capabilities. There are several guidelines for design of a column that were derived from simulations and verified by experiment. Owing to the statistical nature of these beam interactions Monte Carlo simulations are necessary. These effects were first noticed at the HRL where the effect of a *beam halo* was observed when resist-covered benchmarks were scanned by the ion beam to align the beam for subsequent exposures. Upon developing the resist, it was noticed that there appeared to be a *halo* around each benchmark that was scanned. Ward (1985) and Ward et al. (1987, 1988) carried out computer simulations, which indicated that Coulomb interactions were responsible for the effect, and the beam had very long spatial *tails* due to ion-ion interactions. The effect is shown in Figure 11.16, and subsequent measurement of the beam profile verified the theory (Figure 11.17). The beam profile is not Gaussian, but rather is described by the Holtsmark distribution which has very long tails. The distribution was first derived in connection with the Stark broadening of spectral lines and in electric field strength, which results from a random distribution of charged particles interacting



**FIGURE 11.13** The components of a modern FIB system. The system consists of a focusing column containing the ion source and lenses, a detector, gas injection needles, and a sample stage.

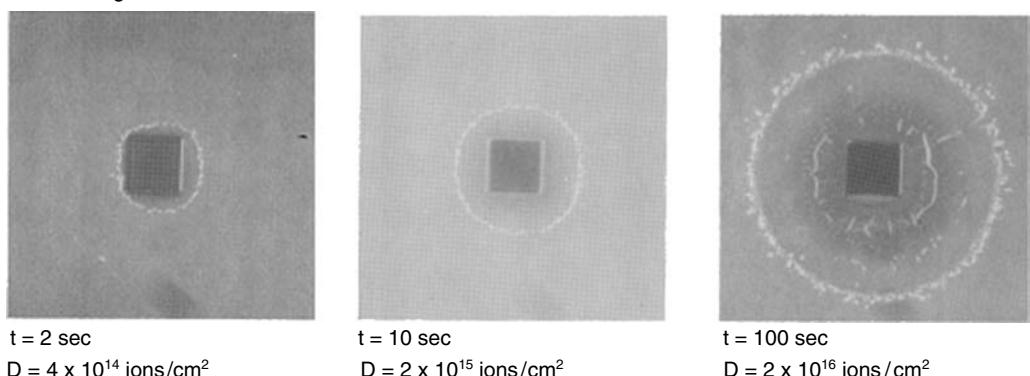


**FIGURE 11.14** Schematic diagram of the optical parts of a typical two-lens focusing column. The beam current is determined by selecting different sized apertures.

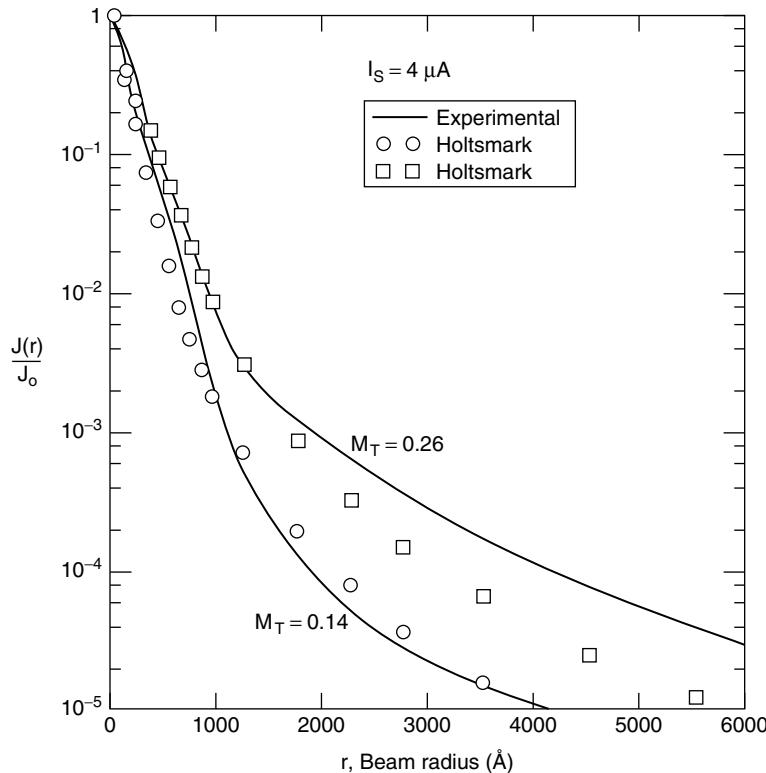


**FIGURE 11.15** Schematic diagram of a mass-separated column designed to perform maskless implantation built at the Hughes Research Laboratories. The column consisted of three electrostatic lenses, a mass separator, and operated at variable energies.

AuSi source  
0.2 nA beam current  
P4 TBS negative resist



**FIGURE 11.16** The *halo* effect due to the long tails of the radial distribution of charge in the beam. This illustration was obtained by milling boxes into a resist with an Au ion beam.



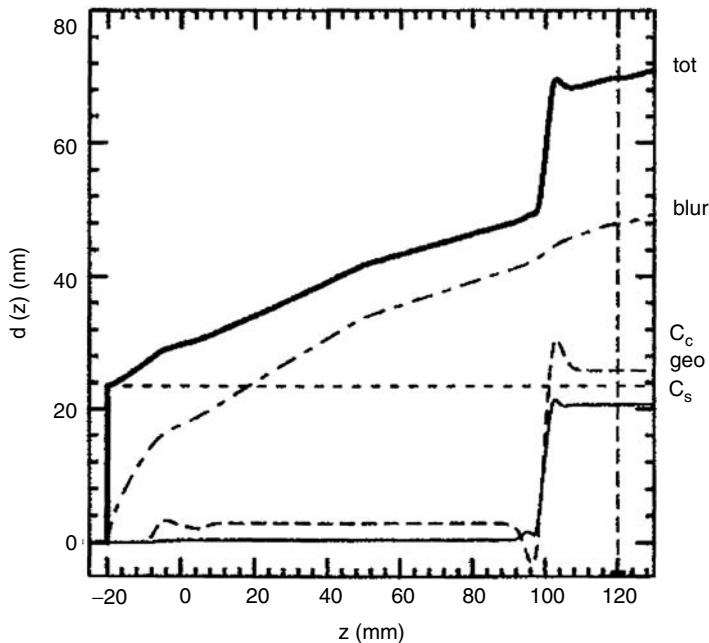
**FIGURE 11.17** Radial current density distributions from data obtained with 50 keV  $\text{Ga}^+$  ions in a two-lens system at two different optical magnifications. The solid lines are from a simulation for the system using the Holtsmark distribution.

with constant charge density. This effect occurs in the source region, and its imprint upon the beam distribution is carried down the length of the column.

In addition to the effects near the source, numerous studies have been made on the effect of ion-ion interactions along the beam path in the column, and they all indicate that careful attention must be given to the design and operating mode of the system to alleviate the growth of the beam size. In general, it has been found that both geometrical aberration effects and effects of Coulomb broadening must be considered simultaneously and the *danger areas* in designs and operating modes that should be avoided. Two of these *danger areas* result from either opening up aperture sizes to allow more current in the beam (many nanoamperes) so that geometrical aberrations dominate or at very low beam currents in which Coulomb effects drop in significance, leaving geometrical aberrations dominant again (typically chromatic effects). The *danger area* is where column performance will not be as anticipated if Coulomb interactions are ignored. Figure 11.18 shows the results of a simulation for a single lens column. In this case, ignoring Coulomb effects would disastrously underestimate the final beam size.

Monte Carlo simulations (Hoepfner, 1985; Swanson, 1994) near the source region yield quantitative dependencies on the physical parameters. The results predict that the virtual source size broadening,  $\Delta d$  (nm), and the longitudinal energy broadening  $\Delta E$  (eV) depend on the atomic mass  $m$  (amu), total current  $I$  ( $\mu\text{A}$ ), ion charge  $q$ , emission half-angle  $\alpha$  (rad), and source radius  $r$  ( $\text{\AA}$ ) and beam path length  $L$  as

$$\Delta d = \frac{6.5 m^{0.25} I^{0.42} L^{0.26} r^{0.5}}{V^{1.1} \alpha^{0.95}} \text{ nm} \quad (11.5)$$



**FIGURE 11.18** A computer simulation for a one-lens system. The total beam size at any point along the beam path is due to four parts. Two are the aberrations of the system, one is due to the source size (fixed) and the final contribution is due to the *blur* induced by Coulomb effects. In this example, the Coulomb blur is the major contributor to the final spot size.

---

**TABLE 11.5**  
**Application of the Results of Monte Carlo**  
**Calculations for Energy Broadening and Growth**  
**of Virtual Source Size for Selected LMIS**

---

Ion	$\Delta E$ (eV)	$\Delta d$ (nm)
Be	2.4	16
Al	4.8	24
Ga	7.1	30
In	9.4	41
Au	8.3	21

---

and

$$\Delta E = \frac{5 m^{0.41} I^{0.83} q^{0.44}}{r^{0.14} V^{0.47} \alpha^{1.1}} \text{ eV} \quad (11.6)$$

The results of these simulations for several different LMIS are shown in Table 11.5.

From these considerations, several general guidelines can be stated: (1) to reduce radial broadening, ions should be of low mass, (2) the ions should be accelerated to as high a voltage as feasible, and (3) as low a beam current as possible should be used. In the optical design, current should be stripped as soon as possible along the beam path, and the beam path length and number of beam crossover points should be minimized.

## 11.6 FOCUSED ION BEAM MICROSCOPY: IONS AND ELECTRONS

Since the incident ions produce both secondary electrons and ions, images can be obtained from either. Scanning ion microscopy (SIM) is the first use of FIB, because viewing in some form is necessary to focus the beam, and to see where it is on a sample. SIM is similar to scanning electron microscopy (SEM), but the contrast mechanisms are different, and as the beam is scanned over a sample it destroys it by sputtering, modifying it at a rate determined by the incident beam current and sputter yield of the sample, and limiting the time that it may be imaged. Because of this destructive interaction, and poorer beam current-probe size characteristics, SIM is not as widely used as SEM. In microscopy there are cases in which SIM is used for inspection and alignment, and it has been recently reported that SIM may be less destructive than SEM for use in the metrology of resists. In SIM, unlike SEM in which the electron probe is just a probe, the ion probe also acts like a tiny scalpel, which slices away layers of the sample, and it is possible to exploit this destruction to gain high-resolution three-dimensional structural, morphological, and chemical information (see Figures 11.47 through 11.50).

## 11.7 GENERAL IMAGING CONSIDERATIONS

To form a good image, there must be resolution, contrast, and adequate signal-to-noise (S/N) ratio. It is not only the instrument, which determines the quality of an image, but also the sample. Good images are a system issue, which we mean by the system, instrument, and sample. With ions a careful strategy must be employed to make a good image because of the destructive nature of the beam. The following analysis covers all the basics of FIB. The limitations in obtaining a good image can conveniently be broken into three categories: optics, sample, and environment. The limitations due to optics are source size and brightness, column aberrations, and beam noise. The limitations of the sample are the available contrast (either compositional or topographic), sputter rate, and charging. The environmental limitations are any outside influence that can degrade the performance of the system, such as mechanical vibrations, time-varying magnetic fields, electronic pickup noise, and grounded loops.

The effective (virtual) source size for the Ga LMIS is  $d_s \sim 50$  nm, and the brightness is  $B \sim 2 \times 10^6$  A/cm<sup>2</sup> · sr. The purpose of the optics column is to image the virtual source onto the image plane with a minimization of degradation of beam size due to the aberrations of the column. Spherical and chromatic aberration can be reduced by reducing the size of the beam-defining aperture of the system so that their contribution is negligible compared with the virtual source. The limiting beam aperture is an angle-limiting aperture, and the relationship between the object side angle  $\alpha_o$  and the image angle  $\alpha_i$  is  $\alpha_o = M\alpha_i\sqrt{V_i/V_o}$ , where  $M$  is the column magnification and  $V_i$  and  $V_o$  the extraction and beam voltages. The beam current  $I_b$  is related to the angular intensity of the source  $I' = 20$  μA/sr by

$$I_b = I'\pi\alpha_o^2 = I'\pi M^2\alpha_i^2\sqrt{V_i/V_o} \quad (11.7)$$

which is related to the brightness and source side  $d_s$  by

$$I_b = \frac{1}{4}B\pi^2d_s^2\alpha_i^2M^2\frac{V_i}{V_o} \quad (11.8)$$

If the contributions of the aberrations can be made small enough by reducing  $\alpha$ , so that the beam size  $d_b$  is effectively just the demagnified image of the virtual source, then the fundamental limitation to the size of the beam is  $d_b = Md_s$ . To make the beam size 5 nm, it is necessary for  $M \leq 0.1$ . This will mean that in a column with  $V_i = 30$  keV,  $V_o = 10$  keV, and  $\alpha_i = 1$  mrad,  $I_b \cong 2$  pA = 12 MHz of Ga<sup>+</sup> at the sample.

To obtain an image of a given S/N ratio, the beam must dwell at points on the sample for a time  $t_d$  to collect enough secondary quanta ( $N_e$  secondary electrons or  $N_i$  secondary ions). The number of detected quanta per pixel is

$$N_e = \frac{\gamma_e I_b t_d \eta_e}{e} \quad \text{or} \quad N_i = \frac{\gamma_i I_b t_d \eta_i}{e} \quad (11.9)$$

where

$\gamma_e$  and  $\gamma_i$  = yields for secondary electrons and ions  
 $\eta_e$  and  $\eta_i$  = corresponding collection efficiencies  
 $e$  = charge of the electron

The average human eye requires that the minimum S/N ratio in an image be  $>5$ . From Equation 11.20, we see that for a two-step process that  $(S/N)^2 \equiv N_{Ga} \gamma_e / (1 + \gamma_e)$ . For  $\gamma_e > 1$ , the noise bottleneck is due to the primary beam current, whereas for  $\gamma_e < 1$ , the bottleneck is due to the formation of secondary particles between the sample and detector. Because the secondary-electron yield is much greater than the secondary-ion yield, we assume that we detect electrons in the following discussions.

In practice, there is an overlap of the beam as it is moved along a line. Negative overlap means that there is a space between adjacent positions of the beam. Positive overlap means that the beam removes material from adjacent pixels. For a beam of diameter  $d_b$  and a step size of  $s$ , the overlap OL is related to  $s$  and  $d_b$  as  $s = d_b(1 - OL)$ . The volume of material  $V$  that is removed can be written as

$$V = \frac{YAN_{Ga}}{\rho N_A (1 - OL)^2} \quad (11.10)$$

where

$N_{Ga}$  = number of incident ions per pixel  
 $A$  = atomic mass

Folding in the S/N relationship gives

$$V = \frac{YAK^2(1 + \delta)}{\rho N_A (1 - OL)^2 \delta} \quad (11.11)$$

To achieve an S/N ratio = 20, it has been found from experience on a variety of samples, which are imaged flat that, assuming  $\gamma_e = 2$ , the number of incident ions must be  $N_{Ga} = 700 \pm 200$ . For imaging a cross section, this number might increase by a factor of 2 to 3 because of reduced collection efficiency by the detector. Assuming 1 pA of beam current, the pixel dwell time required to put 700 ions into a pixel is 0.11 ms, which for a  $1024 \times 1024$  pixel image implies 115 s of image acquisition time. This is not an unreasonable time to hold the electrical and mechanical drifts of the system in tight control.

If we assume that the characteristic size of a particle of volume  $V$  is  $V^{1/3} = d$ , then the minimum resolution can be calculated, assuming it is  $d$

$$d = \left( \frac{YAK^2(1 + \delta)}{\rho N_A (1 - OL)^2 \delta} \right)^{1/3} \quad (11.12)$$

**TABLE 11.6**  
**Fundamental Minimum Resolution Limits for an Image with S/N = 20, and S/N = 5**

Particle	C	Al	Si	Ti	Cr	Fe	Ni	Cu	Mo	Ta	W	Au
Yield	2.7	2.9	2.2	3.3	1.2	3.9	2.0	3.4	1.3	2.8	1.2	14
D <sub>S/N</sub> = 20 (nm)	4.7	5.5	5.3	5.9	3.8	5.5	4.3	5.2	4.1	5.7	4.1	9.5
D <sub>S/N</sub> = 5 (nm)	1.9	2.2	2.1	2.4	1.5	2.2	1.7	2.1	1.6	2.3	1.6	3.8

It is clear from Equation 11.12 that the limiting parameter for image formation that is not under the control of the user is the sputter yield Y for a given incident ion and specimen. We arrive at the same conclusion as Orloff (1995). The easiest way to increase resolution is to sacrifice S/N. If we let the S/N decrease from 20 to 5, then the minimum resolution decreases (i.e., improves) by 2.5. Table 11.6 gives values of the minimum resolution d for elemental particles, assuming S/N = 20 or S/N = 5, with  $\delta = 1.5$  and OL = 0.5.

It is also important to be able to display the information obtained. For a display (screen, photo, etc.) of length L and a scan dimension on the sample per scan where z depth is removed,  $V = 1^2z = L^2z/M$ .

It should be pointed out that the volume of material removed is independent of the scan field size. If higher magnification is used, less area is scanned, and hence the sputtered crater is deeper. The depth sputtered z per pixel is

$$z = \frac{4V}{\pi d_B^2} = \frac{4YAIt}{\pi \rho d^2 e N_A} \quad (11.13)$$

For positive overlap OL, the sputtered depth of a pixel not near the edge is  $z/(1 - OL)^2$ . The total depth sputtered per complete square frame of side 1 with  $N_p$  pixels is

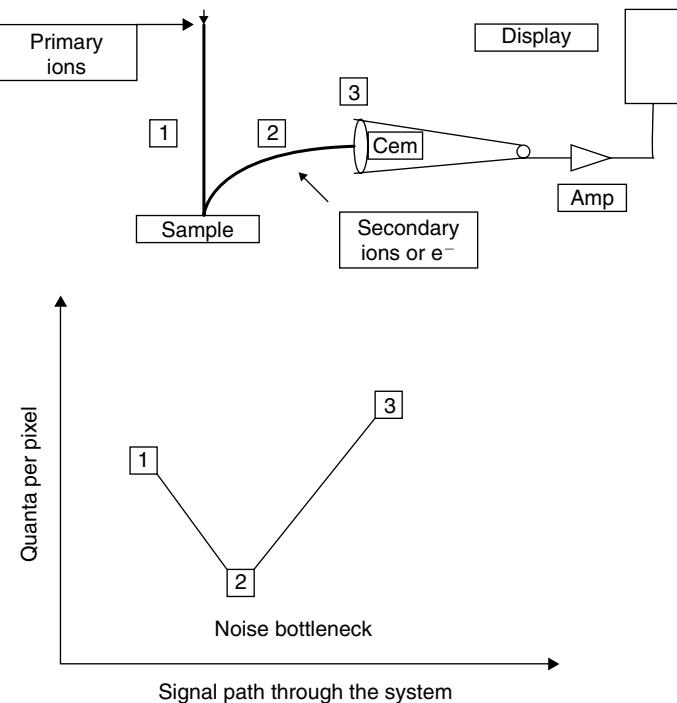
$$z = \frac{4YAItN_p^2}{\pi \rho e N_A I^2} = \frac{4YAItN_p^2 M^2}{\pi \rho e N_A} \quad (11.14)$$

where

M = magnification

The magnification is related to the overlap as  $M = L/(dN_p(1 - OL))$ . If d is the smallest object resolvable, its lateral size on the display will appear  $Md$  in extent. Thus the largest magnification that should be used to detect this minimum size depends on the size of the medium used for display, the beam overlap, and the number of pixels used. For an image size that is 6 in. with 1024 pixels and an overlap of 0.5,  $Md$  is  $1.1 \times 10^{-2}$  in., which is what the eye easily resolve from a foot. The high increase in magnification leads to the increasing overlap and oversampling (empty magnification). If the magnification is not high enough, then there is undersampling, and loss of information for the image.

Once a column has been designed to produce a suitable probe, attention must be given to the choice of what detector(s) are to be placed in the system to see what is happening. A minimum dynamic range is required, since the detector is used for low-current, high-resolution imaging, in which the beam may contain a picoampere or less, and high-current milling operations where quick glimpses are needed to check the progress of the operation, which might be using several tens of nanoamperes of current. Unlike electron beam instruments that produce no ions for detection, FIB



**FIGURE 11.19** Schematic diagram of the signal path in the FIB imaging system.

systems are more complicated because the ion–solid interactions produce both secondary electrons and secondary ions that contain different information about the sample, and in some applications of the instrument, prudentially reactive gases are introduced into the vacuum chamber which might degrade the sensitivity of the detector(s) with time.

The central issues of the choice of detectors are (1) sensitivity, (2) noise, (3) ion/electron detection capability, and (4) lifetime. Wells (1994) has made an analysis of the S/N ratio of the beam and detection system used in electron microscopes. We will use parts of that analysis in this chapter, where applicable.

The S/N in an image is often defined as the ratio of the RMS signal to the RMS fluctuations due to noise, and is not solely necessarily determined by the detector. Figure 11.19 shows the signal path in an FIB system, in which it can be seen that sources of noise occur in the beam, detector, amplifier, and recording system. During a pixel recording, at each point in the signal chain there are  $n_1, n_2, \dots, n_i$  signal quanta at each stage.  $n_1$  is the number of incident ions onto the sample during a pixel recording and the values of the other  $n_i$  depend on the conversion efficiencies at each point.

The RMS fluctuation in the incident ion beam is  $n_1$  so that at that point the total S/N =  $K_1$

$$K_1 = \frac{\bar{n}_1}{\sqrt{\bar{n}_1}} = \sqrt{\bar{n}_1} \quad (11.15)$$

At any following stage  $i$ , the S/N at that stage is  $K_i$  and is given in terms of  $n_1, n_2, \dots, n_i$ , and the conversion factor  $\epsilon_{i-1}$ , which gives  $n_i$  so that

$$\bar{n}_i = \bar{\epsilon}_{i-1} \bar{n}_{i-1} \quad (11.16)$$

If Poisson statistics are assumed, then the variance and mean are equal, and  $(\varepsilon - \bar{\varepsilon})^2 = \bar{\varepsilon}$ . The noise at any stage depends on fluctuations in the conversion factor and the quanta number from the preceding stage. It has been shown that the total fluctuation in  $n_i$  is then

$$\overline{(n_i - \bar{n}_i)^2} = \frac{1}{n_i^2} \frac{(n_{i-1} - \bar{n}_{i-1}^2)^2}{\bar{n}_{i-1}^2} + \frac{\bar{n}_i^2}{n_i} \quad (11.17)$$

which can be changed to

$$\frac{1}{K_i^2} = \frac{1}{K_{i-1}^2} + \frac{1}{\bar{n}_i} \quad (11.18)$$

which by substitution leads to

$$\frac{1}{K_i^2} = \frac{1}{\bar{n}_1} + \frac{1}{\bar{n}_2} + \dots + \frac{1}{\bar{n}_i} \quad (11.19)$$

For example, in a two-stage process such as in a primary beam striking a sample, generating secondary electrons or ions with coefficient  $\varepsilon$ :

$$K_2^2 = \frac{\varepsilon n_1}{1 + \varepsilon} \quad (11.20)$$

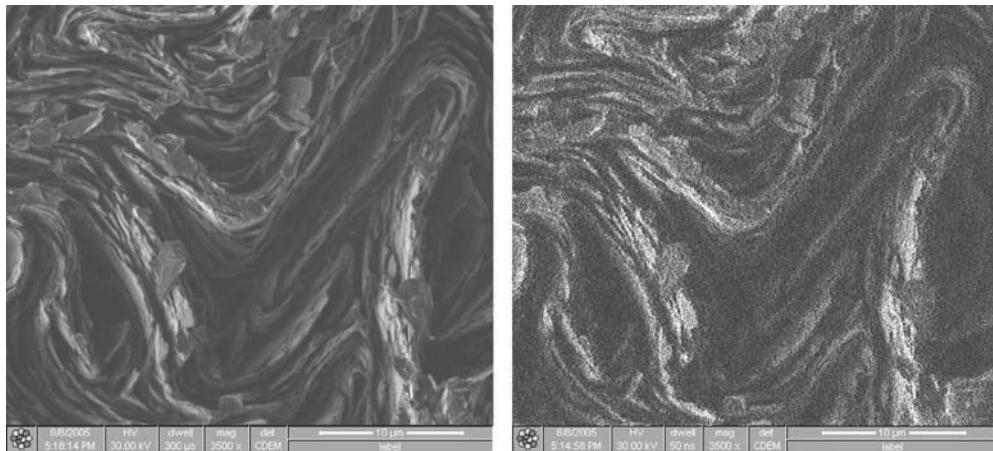
This analysis is valid for Poisson statistics, but it appears that the beam noise is more closely described by the Neyman (Neyman and Scott, 1958) distribution, in which the variance is greater than the mean. At any stage  $j$  where there is a different noise distribution than the Poisson, a factor can be inserted to account for this as  $b_j = (\varepsilon_j - \bar{\varepsilon}_j)^2 / \bar{\varepsilon}_j^2$ , so that

$$\frac{1}{K_i^2} = \frac{b_0}{\bar{n}_1} + \frac{b_1 \bar{\varepsilon}_1}{\bar{n}_2} + \dots + \frac{b_{i-1} \bar{\varepsilon}_{i-1}}{\bar{n}_i} \quad (11.21)$$

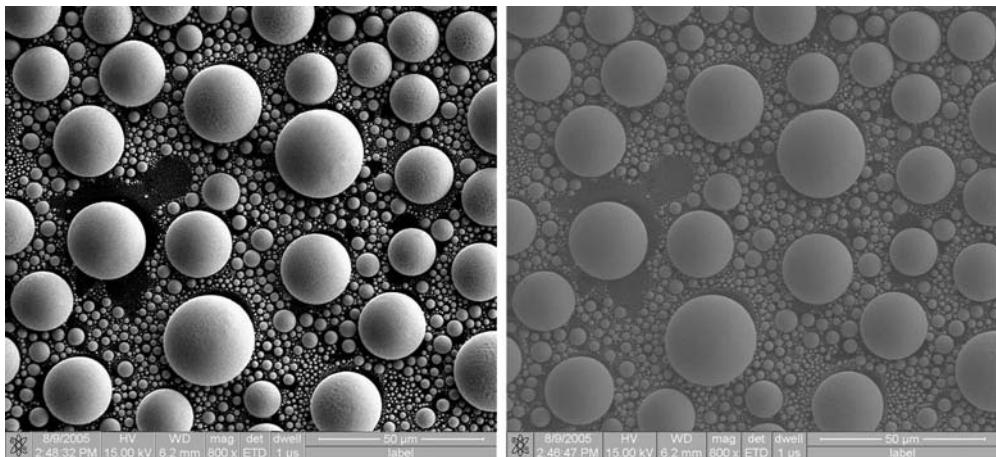
Wells (1974) has pointed out that there is a *noise bottleneck* where the smallest value of  $n$  exists. One can then calculate the noise fluctuation of the system as if it arose solely from that point. In a well-designed system, the bottleneck should be due to some natural limitation rather than an artificial design limitation. Figure 11.20 shows the effect of S/N on an image in which the resolution and contrast have been kept constant.

The responses to light of the human eye is logarithmic (Weber-Fechner Law), so that the contrast difference perceived in an image depends on the ratio of light flux into the eye from the two areas. It is also well-documented that the human eye can see  $\sim 20$  gray levels, so that images with many more gray levels may waste information that was obtained at the expense of eroding the sample with the beam. To obtain the maximum information from an image, the allocation of gray levels should be such that the S/N from the darkest to lightest pixels of the image is constant. This can be accomplished with a gamma control, where the lower levels in the image are expanded. This is easily accomplished with digital images, and optimum values are found empirically. Wells (1974) derives what he terms the *effective S/N from the darkest to the brightest parts of the image* to be

$$\left( \frac{S}{N} \right)_{\text{eff}} = \frac{2}{N_g} \left( \sqrt{\bar{n}_{\text{white}}} - \sqrt{\bar{n}_{\text{black}}} \right) \quad (11.22)$$



**FIGURE 11.20** Images of the same area of graphite with the only difference between the two is the dwell time per pixel. Clearly reduced S/N masks the information available from spatial resolution.



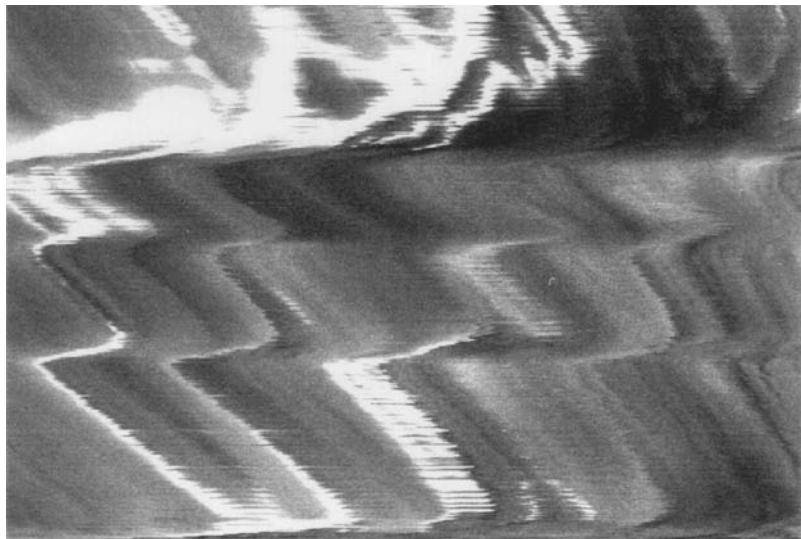
**FIGURE 11.21** The effect of contrast reduction in an image. The conditions of resolution and S/N are identical in each image, but the contrast is reduced in the right image.

where  $N_g$  is the number of gray levels. It is apparent that to make S/N as large as possible, that the minimum number of gray levels that can be perceived should be used.

Several workers have suggested that a minimum S/N of 5 will give an acceptable image. This is actually near the minimum value and is just at the point of *easy discernment*. Great images have a S/N ratio of 20–25, because that is all that the human eye can discern. S/N ratio greater than this is wasted on an observer because of lack of response of the eye, but requires excess sample destruction to form the image. Figure 11.21 shows the effect of loss of information for reduced contrast in an image.

## 11.8 THE VEXATIONS OF CHARGING SAMPLES

For samples that are not perfect conductors, and especially for samples that are insulators, charge will build-up on the surfaces whereas the primary beam is being scanned across them. As the surface acquires a charge (usually positive), a voltage is built up, and unwanted deflections of the primary beam as well as secondary electrons and ions will occur, with the resultant loss of faithful imaging or other controlled beam use. Charging shows up as a complete or partial loss of an image,



**FIGURE 11.22** Charging on a sample.

or of an image *moving around*, or a loss of pristine milling or deposition. In the latter cases, sharp features are unattainable, as would be evidenced by fuzzy sputtered areas and diffuse deposition. In the worst case, the image completely disappears and the milling and deposition process become amorphous. In moderate cases, features in the image are recognizable for periods of time, interlaced with *garbage* (Figure 11.22). In mild cases of charging, the system is still usable, but with decreased effectiveness. For example, when imaging an aluminum surface, the oxide layer is an insulator, and the image darkens with time because of charging. This same phenomenon is seen when imaging while moving the sample stage. The portions of the image just entering the rastered area of the beam are brighter than those portions that have received charge. There are four methods to combat this problem, and the choice of remedy will depend on the application. They are (1) coating the sample with a conductor (e.g., Au, Pt, or C), (2) imaging with secondary ions, (3) using a simultaneous beam of electrons, and (4) photonic methods of conductivity modulation.

Coating samples is straightforward, and always seems to work for dissipating charge buildup. The disadvantage is that an extra step is required, and in some cases, the thin coating is unwanted after the FIB is used. A thin film evaporator is used to make the films about 1000 Å thick, and if needed, the film can be removed by reactive ion etching (RIE). Carbon may be the easiest substance to use, although workers have reported that Au or Pt also works well. An interesting variation to external coating is the use of FIB to do selected coating, or by manufacturing small grounding wires with metal deposition. Basically, the technique here is to use the metal deposition capabilities of FIB to ground the parts of an IC, which are charging by making wire connections from the charging site to a known ground bus. If necessary, after the work is done, the wire can be milled away, and the device returned almost to its original configuration.

Making secondary-ion images either with or without a simultaneous electron beam works in many situations, but the former requires the addition of an electron source. To understand the physics of charge compensation, we analyze the problem here, assuming that a focused primary ion beam with current  $I_i$  is scanning an area  $A$  of the sample in a raster fashion and there is a detector, which is capable of collecting either secondary electrons ( $\alpha I_i$  and  $\beta I_i$ ) or secondary ions ( $\gamma I_i$ ) by means of a suitably biased grid, and secondary source of electrons ( $I_e$ ), which can impinge on the sample. The sample current  $I_s$  is then given by

$$I_s = I_i + \alpha I_i + \beta I_i - \gamma I_i - I_e \quad (11.23)$$

where

$\alpha$ ,  $\beta$ , and  $\gamma$  = secondary-electron yield due to primary ions, secondary-electron yield due to primary electrons, and secondary-ion yield due to primary ions, respectively.

A term representing the scattering of primary electrons by the surface has been neglected, as that elastic channel is very small for the situation considered here. Under the assumption that the surface can be simply modeled by a parallel R and C being charged by  $I_s$ , then the surface potential as a function of time is

$$V_s(t) = I_s R (1 - e^{-t/RC}) \quad (11.24)$$

Treating the insulator as a parallel plate capacitor of area A and thickness d, filled with a dielectric constant  $\epsilon$  and resistivity  $\rho$ , and then from the simple relationships,  $C = \epsilon_0 A/d$  and  $R = \rho d/A$ , we obtain

$$V(t) = \left( \frac{I_s \rho d}{A} \right) (1 - e^{-t/\rho \epsilon_0}) \quad (11.25)$$

The time it takes the sample to charge to one-half of its final value is

$$t_{1/2} = 6.1 \times 10^{-14} \epsilon \rho (s) \quad (11.26)$$

and the surface potential at that time is

$$V(t_{1/2}) = \frac{I_s \rho d}{2A} \quad (11.27)$$

For a 1-nA beam rastered over a  $10 \times 10 \mu\text{m}$  area on a sample with a 1- $\mu\text{m}$ -thick insulator with  $\epsilon = 2$  and  $\rho = 10^{10} \Omega\text{-cm}$

$$t_{1/2} = 1.2 \text{ ms} \quad (11.28)$$

and

$$V_{1/2} = 500 \text{ V} \quad (11.29)$$

In this case, the insulator will breakdown, and a discharge will occur. If instead a  $100 \times 100 \mu\text{m}$  area were rastered, the final surface potential would be 10 V, and that might be acceptable. It is clear that the process of charging can be fast and very high voltages can be encountered, so care must be taken in selecting the primary current, beam dwell time, and area to be rastered.

The effect of detector bias can also be analyzed. For imaging with secondary electrons, the front of the detector biased positively with respect to the sample, and if that bias is sufficient to collect all secondary electrons (a typical value is  $\alpha = 2$ ) and return all secondary ions to the sample ( $\gamma = 0$ ), then

$$I_s = I_i(1 + \alpha) + I_e(\beta - 1) \quad (11.30)$$

For the case  $I_e = 0$  with  $\alpha \approx 2$  and  $I_s \approx 3I_i$ , the sample will charge positively. When the sample reaches the potential of the detector grid bias, secondary electrons will be turned back to the sample, and  $I_s \approx I_i$ , and no secondary-electron image can be obtained. The image is black.

If an electron current  $I_e$  is used to compensate the charging,  $I_s = 0$  when

$$I_e = I_i \frac{(1 + \alpha)}{(1 - \beta)} \quad (11.31)$$

and for  $\alpha \sim 2$  and  $\beta < 1$ ,

$$I_e \approx \frac{3I_i}{(1 - \beta)} \quad (11.32)$$

So it should be possible to image insulators, but the detector would be saturated with electrons from the neutralizing source, unless discrimination is possible.

For the case of a negatively biased detector, positive ions are collected to form an image, and secondary electrons are returned to the sample. In this case,  $\alpha = 0$ , and

$$I_s = I_i(1 - \gamma) - I_e \quad (11.33)$$

For no compensating electron current,

$$I_s = I_i(1 - \gamma) \quad (11.34)$$

and since  $\gamma < 1$ ,

$$I_s \approx I_i \quad (11.35)$$

so that for ion imaging, there is some reduction of charging. Several users have reported that in many cases where charging is present, by switching to ion imaging the problem is alleviated. The disadvantage here is that there are not as many secondary ions as secondary electrons available to form an image ( $\alpha \ll \gamma$ ), and high S/N ratio may be a problem. If there are sufficient statistics available, then this method reduces charging substantially and it would be the preferred method since it is the simplest.

If an electron beam is also used with ion imaging,  $I_s = 0$  when

$$I_e = I_i(1 - \gamma) \approx I_i \quad (11.36)$$

In practical situations, when using a compensating electron beam, a low-energy *flood gun* is employed. For many situations of a few electron volts (10–100 eV) are made to drift over the sample, and areas which are positively charged force some of these electrons into the charged area. In addition, if a microchannel plate (MCP) detector is used, the extraction field of the grid aids in directing these electrons to the sample. The same phenomenon works, but at reduced efficiency with channel electron multiplier (CEM) detectors if the geometry is chosen carefully. The aiding field of the detectors is necessary, because the normal electric field produced at the sample surface falls off rapidly with distance from the surface. For a circularly charged area of radius  $r$  and charge density  $\alpha$ , the normal field to the surface is

$$E(z) = \sigma \epsilon \epsilon_0 (1 - 1/(1 + (r/z)^2)^{1/2}) \quad (11.37)$$

Photons may be used to generate carriers in metals and semiconductors. The phenomenon is wavelength dependent for either the photoelectric generation of electrons from metals, or the penetration of light into a semiconductor to induce conductivity modulation. The latter effect is

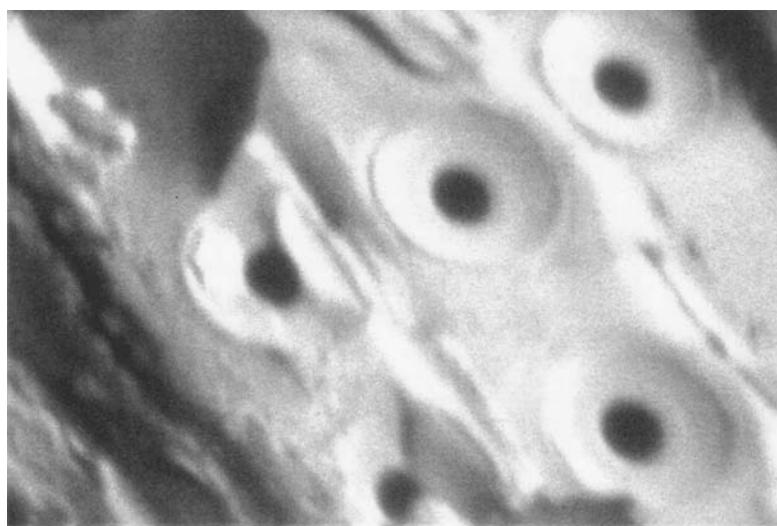
similar to optical beam-induced current (OBIC). The use of UV radiation for the elimination of charging has been demonstrated in SIMS for large-area of semiconductors (Hayashi and Hashiguichi, 1993), and has been demonstrated in FIB work on device modification.

If there are sufficient electrons to induce neutralization, the process is essentially self-regulating. When neutralization is achieved, there is no field with which to draw electrons, and the sample should begin to charge to negative voltages due to excess electrons, further charging will be limited by the sample as electrons are then rejected. The sample can only charge to the potential of the primary electron beam. In practice, the primary ion and electron beams are blanked and unblanked in unison, so that excess charge will not occur.

The use of an electron gun makes secondary-electron detection difficult. One means of recording secondary-electron images while using a flood gun is to *multiplex* the detector and gun. Basically, the flood gun is turned on to neutralize the sample, it is then blanked, the detector is enabled, and a line of data is taken. The primary beam and detector are then blanked (perhaps during a retrace), the electron gun unblanked, and the process continues repeatedly. Lines in the image are built up in this fashion, and the time overhead is negligible. For a  $1024 \times 1024$  image a 1 ms flood time per line adds 1 s to the image collection time. For typical beam dwell times, this represents about 1% overhead.

Another method of obtaining secondary-electron images is to set the grid on the front of the detector to reject electrons of energy equal to or less than those produced by the flood gun, so that only very energetic secondary electrons are collected. The image quality will suffer a little, but respectable images can be obtained this way. The images are not pure electron images, but have a contribution from the secondary-ion information. Figure 11.23 shows such an example.

In both cases, because the electrons are not focused and *flood* the surface with a current density less than that of the primary ion beam, not all of them land in the area that is charging and a rule of thumb for typical flood guns is that at least  $50\times$  as much electron current is needed as primary ion current. It should be noted that the impinging electrons reduce charging by two means. The first is direct neutralization by the primary electron beam, whereas the second is due to the liberation of secondary electrons by the primary electron beam at the surface near to where the charging is occurring. If the secondary yield is greater than unity ( $\beta > 1$ ), then some of the secondary electrons generated outside the region of charging can be pulled into the



**FIGURE 11.23** Image on a sample, which is flooded with electrons from a low-voltage gun. The sample is a piece of wood.

charging region to aid in neutralization. In a dual-beam system, the capability for neutralization is built in. The electron beam can be rastered over the same area as the ion beam, or it can be rastered off to the side or run in a spot mode. Again, there are probably two mechanisms of neutralization.

Because electric forces are so strong, it is very difficult without detailed modeling to understand precisely the exact nature of neutralization. By using simulation programs, like SIMION (Dahl et al., 1990), it is possible to infer trends which will aid in neutralization. Experiments have shown that it is possible to image such diverse insulators such as wood, PTE, Teflon<sup>TM</sup>, and passivation layers with the aid of an electron gun.

## 11.9 APPLICATIONS OF FOCUSED ION BEAM

In this section, we cover the major applications of FIB. The LMIS, which started as a laboratory curiosity, is now the basis of FIB systems that use a variety of heavy metal ions, which are focused into diameters smaller than  $0.1\text{ }\mu\text{m}$  with current densities of several ampere per centimeter square. The predominant use of these systems now is in various aspects of IC fabrication. The two features of this tool that are most important now are the capability to remove by sputtering (micromachining) and add by ion-induced reactions (deposition) material with submicron dimensions. In this role, the tool is used in failure analysis, defect characterization, design modification, and process control.

There are four other uses of FIB that will also be covered:

1. Scanning ion microscopy (SIM) is being used in the semiconductor industry because spatial resolution is now near 5 nm, and there are contrast mechanisms available other than those of electron microscopy.
2. SIMS when coupled with FIB (FIB/SIMS) gives very good high lateral and vertical spatial resolution chemical maps in addition to spectra. This technique is used in the semiconductor industry and in biological and materials science investigations.
3. For historical reasons and to cast the possible uses of FIB into perspective, we will also cover FIB implantation of semiconductors that was investigated in the past and has not yet found wide use in the fabrication of semiconductors, but may still find areas where it will be used.
4. FIB lithography which was also investigated and, like implantation, has not yet found a niche.

Mask repair will also be lightly covered in Sections 11.9.1 and 11.9.2 on micromachining and deposition. The book by Prewett and Mair (1991) covers mask repair in detail, and at the time of its publication mask repair was the major endeavor of FIB machines.

### 11.9.1 MICROMACHINING AND GAS-ASSISTED ETCHING

Micromachining is the programmed fabrication of features on a surface on a submicrometer scale by removal of material. The FIB system is used as a combined SIM and a milling machine, sometimes being referred to as a micromill. The term “milling” is used in analogy with larger machine tools for the sputtering of material. The small size of the ion beam allows structures to be produced that are smaller than conventional means that are capable of producing structures with sizes of  $\sim 25\text{ }\mu\text{m}$  (0.001 in.). Another feature of FIB is that the results of the work can then be imaged in the same machine. The majority of applications for micromachining are in the semiconductor industry. There are two main applications: (1) failure analysis (in which the beam is used to cut and image cross sections of a part of a circuit) and (2) device or IC modification (in which wires or tracks can

be either cut or joined in ICs that are faulty). There is also a range of other applications that include the preparation of TEM samples, materials analysis using SIMS, and the repair of both optical and x-ray masks.

The sputtering process is responsible for this effect, in which as previously mentioned, the incident beam transfers sufficient momentum to the surface and near-surface atoms for them to escape through a cascade of collisions. The sputtering yield  $Y$  is the number of sputtered atoms per incident ion and without the enhancements brought about by the introduction of reactive gases is typically  $0 < Y < 5$ . In the presence of reactive gasses such as Cl or I, the yield can be increased by a factor of  $20\times$  under suitable conditions in some materials due to chemical reactions initiated by the ion beam. Sometimes it is more convenient to express the sputter yield in terms of volume of material removed per quantity of charge. Owing to the scales of the beam size and current, typical numbers are quoted in cubic microns per nanocoulomb. The range of values of the sputter yield for 25 keV  $\text{Ga}^+$  is from 0.05 to 0.7  $\mu\text{m}^3/\text{nC}$ . An advantage of using these units is that the removal rate (volume per time) is the product of the incident current and sputter yield. For example, the sputter yield for Al is 0.3  $\mu\text{m}^3/\text{nC}$ , so that for an incident beam current of 100 pA, the sputter rate is 0.03  $\mu\text{m}^3/\text{s}$ . If it is desired to mill a 5- $\mu\text{m}$  square hole of 5  $\mu\text{m}$  deep, the total volume of 125  $\mu\text{m}^3$  will have material removed at about 7.2 nm/min and will take 69.4 min to complete. In this example, one might consider using 1–6 nA of beam current to decrease the removal time to a more reasonable time of from 1.2 to 7 min. Table 11.7 gives measured sputter rates and sputter yields for elements as measured by Leslie (1994). The sputter yields  $Y$  are calculated from the sputter rates  $Y_r$ , where the relationship is,

$$Y = 96.4 \frac{\rho Y_r}{m} \quad (11.38)$$

with  $m$  the mass in amu, and  $\rho$  the density in  $\text{g}/\text{cm}^3$ .

A serious negative complication to milling is redeposition. As material is sputtered away, some of it becomes redeposited in the volume that is being sputtered. In normal mechanical machining, buildup of machined material is avoided by the use of liquid or air streams, which carry the

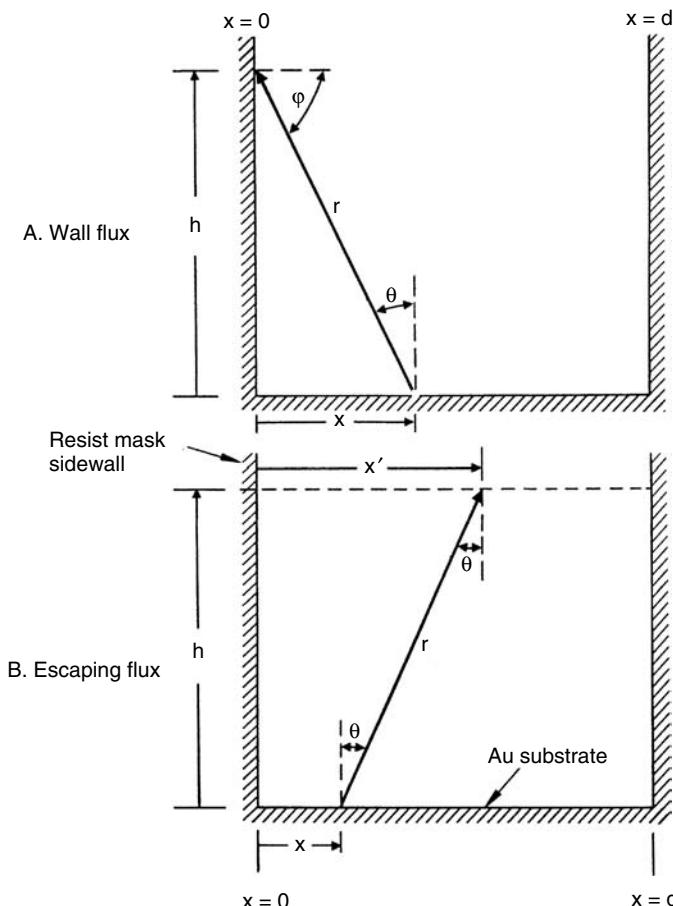
---

**TABLE 11.7**  
**Sputter Rates and Sputter Yields for Selected Elements**

Element	Density ( $\text{g}/\text{cm}^3$ )	Sputter Rates ( $\mu\text{m}^3/\text{nC}$ )	Sputter Yield (Atoms/Ion)
C (diamond)	3.57	0.18	2.73
Al	2.7	0.3	2.89
Si	2.33	0.27	2.08
Ti	4.5	0.37	3.35
Cr	7.19	0.09	1.20
Zn	7.13	0.34	3.57
Ge	5.32	0.22	1.55
Se	4.81	0.43	2.52
Mo	10.2	0.12	1.32
Ag	10.5	0.42	3.94
Sn	5.76	0.25	1.17
W	19.25	0.12	1.22
Pt	21.47	0.23	2.44

*Source:* Courtesy of A. Leslie.

---



**FIGURE 11.24** Geometry for the calculation of redeposition effects. (Courtesy of J. Ward.)

swarf away. Redeposition is critically dependent on how the milling is done. It has been shown that for the same total dose, scans that are repeated many times to mill a rectangular area cause less redeposition to occur than a single slow pass to mill the rectangle. For the slow scan, redeposited material is not removed, whereas for the fast repeated scans, some fraction of the primary beam is used to sputter away redeposited material, so that the sputter yield is less for the fast scans, but the total number of atoms sputtered is nearly the same. In an experiment that showed good agreement with results of Yamaguchi et al. (1985), Crow (1990) showed that the sputter yield as a function of scan speed in the range 0.05–1.0 cm/s remains constant. In addition, the sputter yield as a function of line pitch also remained constant.

A simple geometric one-dimensional analysis of redeposition illustrates the problem. For the case of milling a groove of width  $d$  and depth  $h$ , we can calculate the flux density of material sputtered from the bottom of the groove to the sidewalls of the groove. The geometry is shown in Figure 11.24. If we assume that the sputtered efflux has a cosine distribution, and falls off as  $1/r$ , and that the sticking coefficient is unity, then the flux density of sputtered material from the bottom redeposited onto a sidewall at a height  $h$  above the bottom is given by

$$F(h) = \frac{F_0}{2} \int_0^d \frac{\cos \phi \cos \theta dx}{r} = \frac{F_0 h}{2} \int_0^d \frac{x dx}{r^3} \quad (11.39)$$

where

$$F_0 = \text{total number of atoms emitted per unit length}$$

$$R = (x_2 + h^2)^{1/2}$$

$$\cos \phi = x/r$$

$$\cos \theta = h/r$$

The flux density of sputtered atoms which escape without striking the sidewalls is given by

$$F(x') = \frac{F_0}{2} \int_0^d \frac{\cos^2 \phi dx}{r} = \frac{F_0 h^2}{2} \int_0^d \frac{dx}{r^3} \quad (11.40)$$

where

$$x' = \text{distance from the left sidewall}$$

$$r = (h^2 + (x - x')^2)^{1/2}$$

The intensity distribution of the redeposited and escaping fluxes can be found by numerically integrating Equations 11.39 and 11.40. The total incident flux on a sidewall and the escaping flux are

$$Q_{\text{wall}} = \int_0^h F(h) dh \quad \text{and} \quad Q_{\text{lost}} = \int_0^h F(x') dx \quad (11.41)$$

For the case of two sidewalls, Figure 11.25 shows the normalized redeposited and escaping flux as a function of aspect ratio for the groove. At an aspect ratio  $< 1$ , more than half of the sputtered atoms from the bottom surface redeposit. For nonunity sticking coefficients,  $g \neq 1$ , then

$$Q_{\text{wall}} \Rightarrow gQ_{\text{wall}} \quad Q_{\text{lost}} \Rightarrow Q_{\text{lost}} + (1 - g)Q_{\text{wall}} \quad (11.42)$$

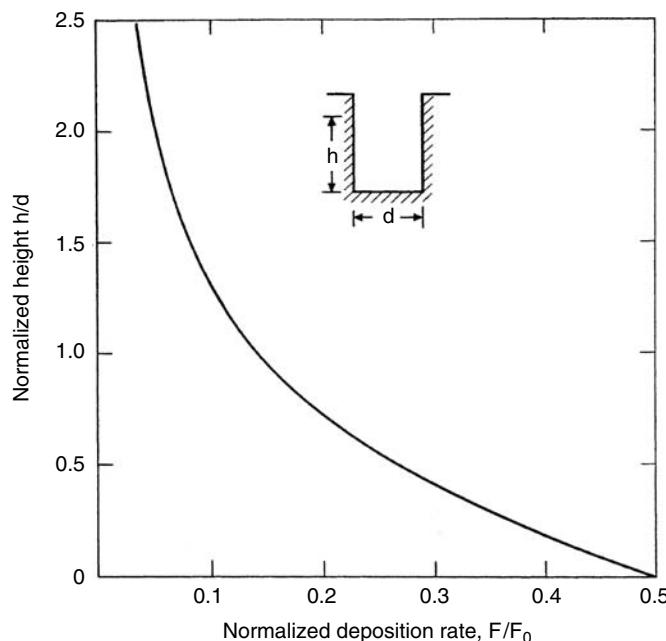


FIGURE 11.25 Normalized redeposition effect. (Courtesy of J. Ward.)

There is an effect, which can aid in the reduction of redeposition. Enhanced etching, sometimes called gas-assisted etching (GAE) was first investigated in FIB by Ochiai et al. (1985), and it greatly increases the rate at which material is removed from a surface by effectively reducing or eliminating redeposition without the need for special scan strategies. In this method, typically a reactive halogen gas is directed at the surface to be milled. Ochiai et al. described the effect of the presence of  $\text{Cl}_2$  as a function of pressure in a chamber on the etch rate of GaAs with a  $\text{Ga}^+$  beam. Takado et al. (1987) and Gandhi and Orloff (1990) used capillary needles to deliver  $\text{Cl}_2$  directly to the surface of the sample. Large enhancements were measured, and a model was developed which takes into account the relative fluxes of ions (in terms of scan rate and dwell time) and gas to the surface. In addition to an increased etch rate (up to 30× for the right conditions), there can be a selectivity of material removal. For the halogen gases  $\text{Cl}_2$  or  $\text{I}_2$ , materials such as Al, Si, and GaAs are selectively enhanced by 20–30× and  $\text{Si}_3\text{N}_4$  by 5–10×, whereas oxides like  $\text{SiO}_2$  and  $\text{Al}_2\text{O}_3$  have near unity enhancement. For the gas  $\text{XeF}_2$ , there is an enhancement for dielectric materials (Figure 11.26). The process of GAE is that the gas interacts with the sample, creating volatile compounds which are pumped out by the vacuum system. With the introduction of the proper gas, the effective sputter yield and rate are enhanced by some factor larger than unity. In many cases the yield enhancement is less than unity for some materials leading to selectivity. This is basically due to the existence or lack of chemical reactivity with the surface of the sample. Table 11.8 gives the etch rate enhancements for some materials of interest.

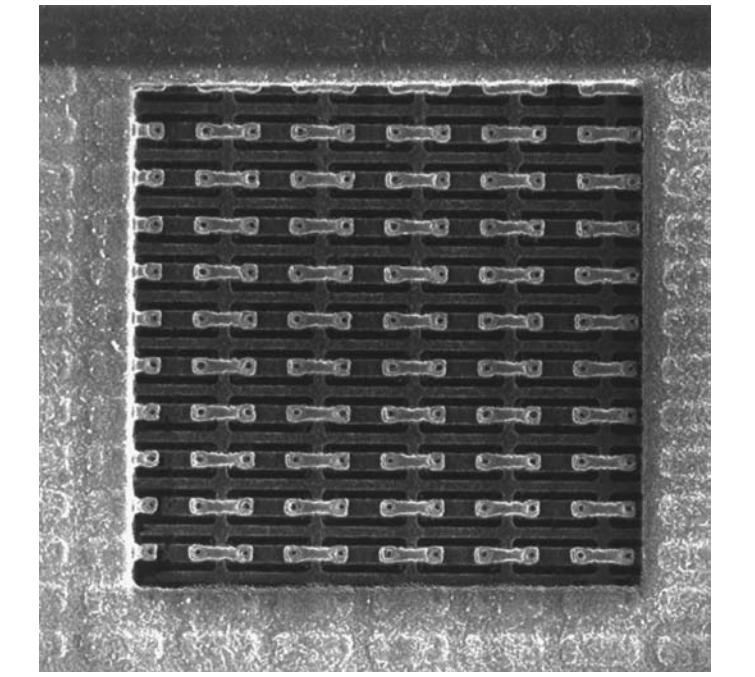
Figure 11.27 shows two vias in a semiconductor. The left one was done without an enhanced etch gas and the right with iodine introduced. It is clear that the sidewalls are better defined and cleaner, and the total time to mill a given volume is less, because milling of redeposition does not occur when iodine is used.

The scan parameters and method of gas delivery are critical. Typically, a capillary needle is used to deliver the gas close ( $\sim 50$ – $300 \mu\text{m}$ ) to the surface. Usually, either gaseous chlorine form a manifold or solid iodine heated in a small crucible is passed to the capillary. At larger distances, the flux from the end of the needle spreads, and delivery of the reactant atoms to the area of interest decreases. If too much gas is delivered to the surface, there can be a decrease in etch rate, because any gas absorbed onto the surface must be sputtered away by the ion beam. The loop time and dwell time must be adjusted so that about a monolayer of gas is present while the beam dwells at a pixel. Too long a dwell time uses up the gas, and normal sputtering then occurs. Too short a loop time also causes the same effect.

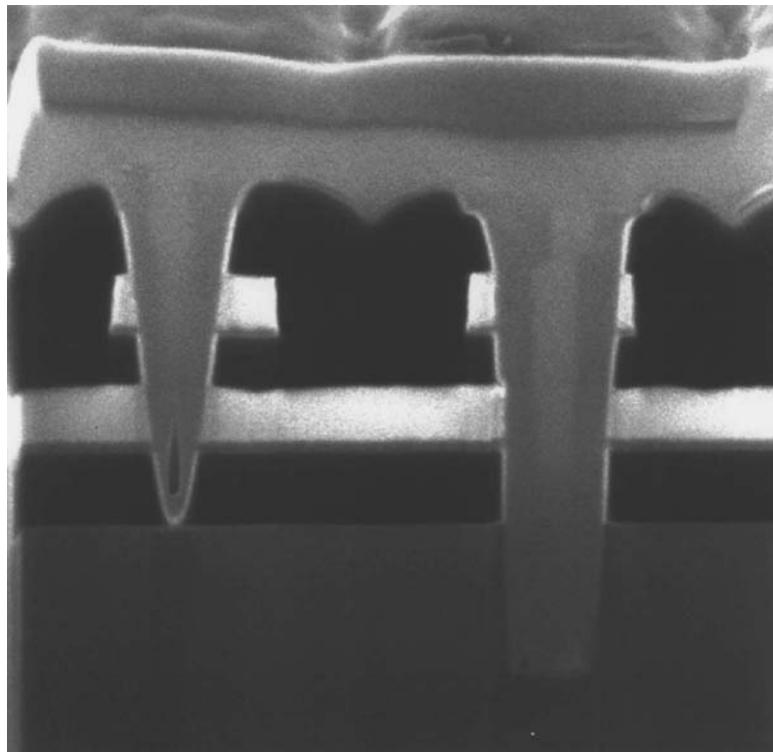
It is now well known that water enhances the sputtering of carbon-containing materials such as polymethyl methacrylate (PMMA), polyimide, and other resists by a factor of  $\sim 20$  relative to physical sputtering (Stark et al., 1995). In addition, there is a reduced removal rate from some other materials such as Si and Al by as much as a factor of 10, giving and increased selectivity of polymers over these other materials of a factor of 200. Removal rates of  $2000 \mu\text{m}^3/\text{min}$  have been reported, and the same trends for enhanced etch with Cl or I (increased gas pressure at the surface, short pixel dwell times, and long frame times) seem to hold. Table 11.9 (see Kang, H.H., Chandler, C., Weshcler, M. Gas assisted ion beam etching and deposition, in Yao (2007) pp. 67–86) gives yield enhancements for various gases on PMMA.

**TABLE 11.8**  
**Values for Rate Enhancements for Materials**  
**Sputtered by 30 keV  $\text{Ga}^+$  Ions**

	$\text{I}_2$	$\text{Cl}_2$	$\text{XeF}_2$
Si	7	11	10
Al	10	8	1
W		1	8
GaAs	12	50	
$\text{SiO}_2$	1.25	1	8



**FIGURE 11.26** Micrograph showing the etch selectivity of  $\text{XeF}_2$ . Metal 1, and metal 2 structures are left intact.



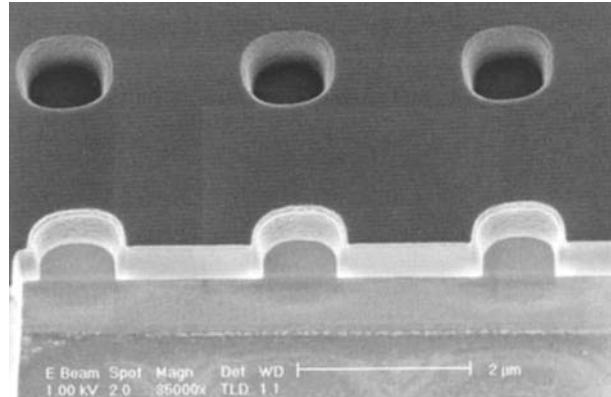
**FIGURE 11.27** Two vias showing the effect of enhanced etch. The via on the right was milled by the Ga beam with iodine gas present.

**TABLE 11.9**  
**Yield Enhancements for 30 keV Ga<sup>+</sup>**  
**on PMMA**

Gas	Yield ( $\mu\text{m}^3/\text{nC}$ )	Enhancement
None	0.4	—
I <sub>2</sub>	0.8	2
XeF <sub>2</sub>	1.7	4.2
O <sub>2</sub>	1.8	4.5
CH <sub>3</sub> OH	4.2	10
H <sub>2</sub> O	7.5	18

**TABLE 11.10**  
**Water Etch Enhancements**

Material	Enhancement
Polyimide	18
PMMA	18
Diamond	10
Teflon	6
Au	1
SiN	0.3
SiO <sub>2</sub>	0.3
Si	0.2
Al	0.15

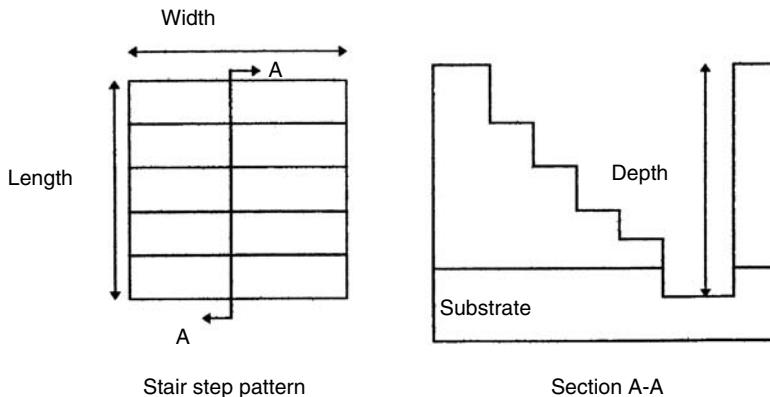


**FIGURE 11.28** Example of water enhanced etching to remove PMMA resist to expose the profile of the contact holes to the underlying silicon.

Water also inhibits the etch rate for several materials of interest. Table 11.10 shows the enhancement and inhibition for various materials.

Figure 11.28 shows the effect of water-enhanced etching on PMMA.

As an example of the usefulness of GAE, we describe the procedure for cutting a cross section. Cross sections are especially useful for investigations in failure analysis, low-yield analysis, and process control. Small staircase openings are made in a device, cleaned up, and polished, and then tilted up to view the face of the cross section in the FIB or an electron microscope. The steps in



**FIGURE 11.29** Staircase pattern used in FIB cross section for bulk removal. (Courtesy of G. Crow, Intel Corp.)

producing a good cross section are as follows (we assume that a generic wafer or packaged device is being used).

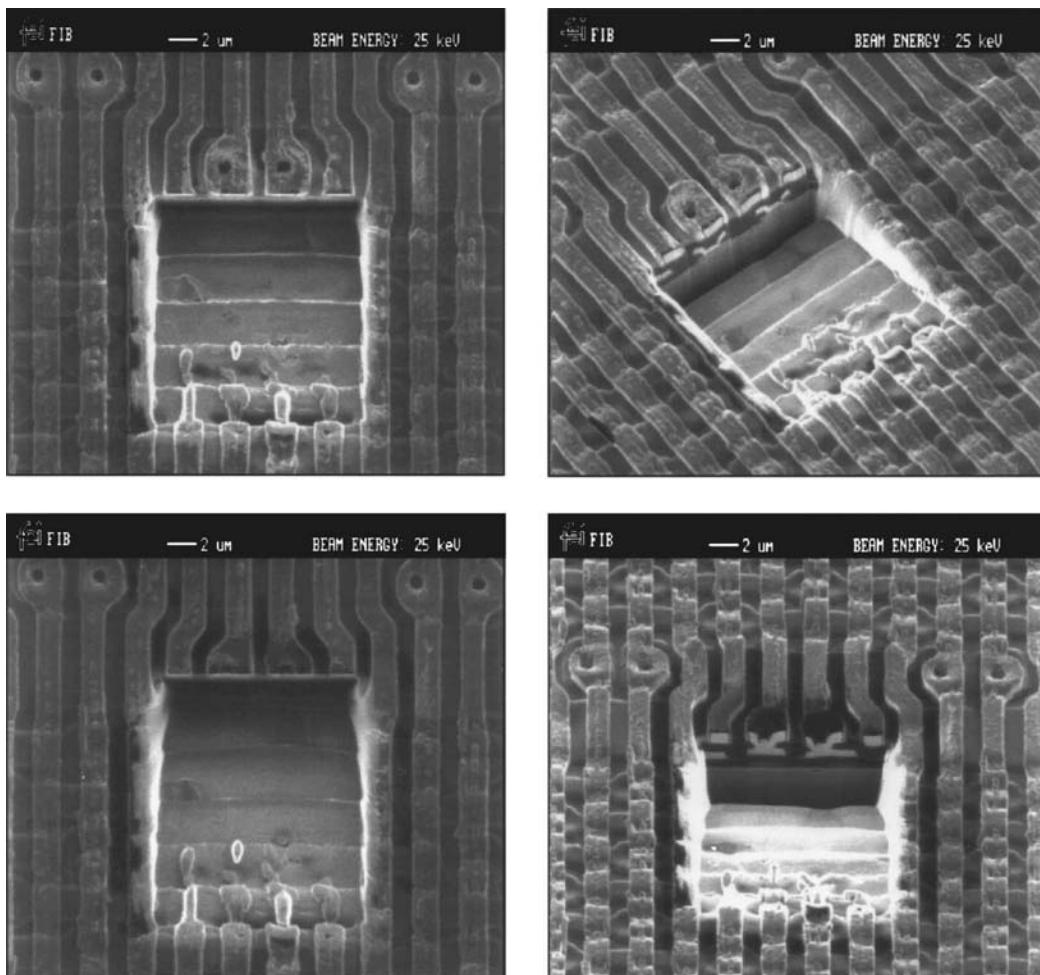
The first concern is to make sure that charging will not be a problem. It can make imaging difficult, or impossible and, in general, can cause the beam to drift relative to the sample so that clean milling will be impossible. To avoid charging, the sample should be grounded and, in many cases, it needs to be coated with a thin (several hundred angstroms) conductive film of a material like carbon. Carbon is slowly sputtered by the ion beam and, after the process is finished, it can be easily removed from oxygen plasma. In some cases, workers have used the FIB metal deposition system to run connections to ground points. There are many instances where an electron flood gun can be used with great success without the need for coating.

The determination of where to make the cross section requires some knowledge of the sample. If there are surface features that can be used for identification, then the FIB can be used as a microscope to find those features. For planarized surfaces with buried features, some combination of CAD navigation, coordinate transfer, and reasonably local landmarks is necessary. If any of these are combined with image overlay, a better estimate of where to perform the cross section can be obtained.

Once the area is identified where the cross section cut is to be made, a milling pattern needs to be defined in shape and position. A stair-step pattern is shown schematically in Figure 11.29, which is used for the removal of the bulk of the material. The stair-step shape is used so that when the cross section is tilted to  $\sim 45^\circ$ , the face can be imaged. The width of the pattern is defined by the size of the feature, whereas the length is determined by the depth of the feature.

A large beam current should be used in the first step of bulk removal to reduce the time. Typically, enough current should be used to remove the desired amount of material in 10–15 min. For typical cross sections, where several hundred cubic micrometers are to be removed, 1/2–2 nA of current is needed. For larger volumes, more current is needed. In this step, the use of enhanced etch is encouraged for saving time and keeping the shape of the stair step reasonably pristine.

After the bulk of the material has been removed, the face needs to be cleaned up before a final polishing step is performed. In the cleanup phase, the cleanup pattern width should match the width of the existing sputter crater, and the top edge of the pattern should fall just short of the final polish cut. The bottom edge should just overlap the existing sputter crater as shown in Figure 11.30. Enough beam current should be used so that each line in the cleanup will sputter for 15–20 s. Because of the geometry of the existing sputter crater, the sputter yield will be much higher (nonnormal angle of incidence). Typically, 250–500 pA of current is used in this step, with a 25% pixel overlap and a 1 ms dwell time. It is usually possible to monitor the progress of this step, and any drift should be compensated. Enhanced etch gas should not be used in this step because it generally causes too much sputtering to maintain control, and the selectivity can cause artifacts.

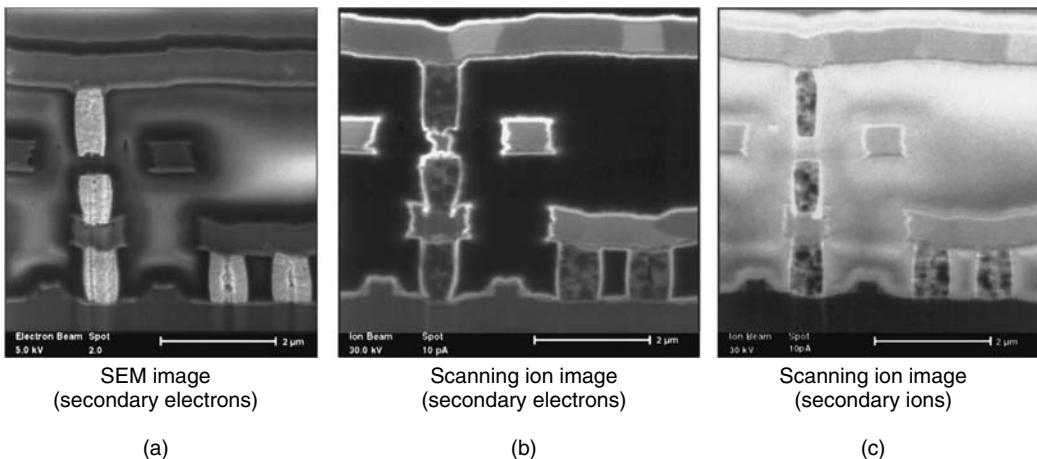


**FIGURE 11.30** Cross section. (Courtesy of G. Crow, Intel Corp.)

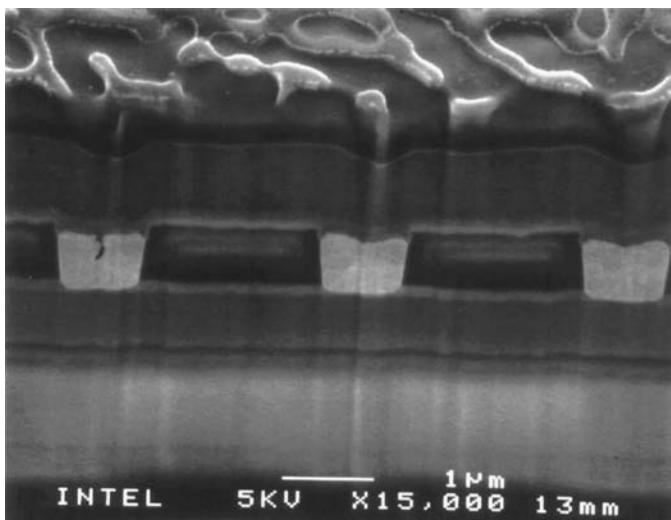
For the final polish, a pattern similar to the cleanup pattern is used, in which the width is the same as that of the crater. The top edge of the pattern should just pass through the center of the desired feature; whereas the bottom edge should just overlap the existing sputter crater. A beam current should be chosen that will sputter each line of the polish pattern for 30–45 s, and is typically between 50 and 100 pA, with the same overlap and dwell time as for the cleanup phase.

When the polish phase is finished, the sample can be tilted to 45°, and, with a minimum beam current, focusing and stigmation of the beam should be carried out on a region away from the area of interest to avoid damaging the cross section face. A single slow scan to enhance the S/N ratio of the image is usually superior to averaging frames. In secondary-electron images (Figure 11.31a), metals are bright and dielectrics are black, whereas in secondary-ion images (Figures 11.31b and 11.31c), both metals and dielectrics are bright. The S/N ratio of electron images are usually higher than ion images, because there are more electrons emitted for collection.

There are several artifacts that sometimes show up when cross sections are made. They fall into three categories: (1) waterfalls, (2) ridges, and (3) droplets. Waterfalls (Figure 11.32) are vertical stripes in the surface of the cross section, it arise from surface topography and has large differences in sputter rate. They can also arise from enhanced etch gas in the cleanup or final polish phase.



**FIGURE 11.31** (a) SEM image with secondary electrons, (b) SIM image with secondary electrons, and (c) SIM image with secondary ions. (Courtesy of P. Carleson, FEI Co.)



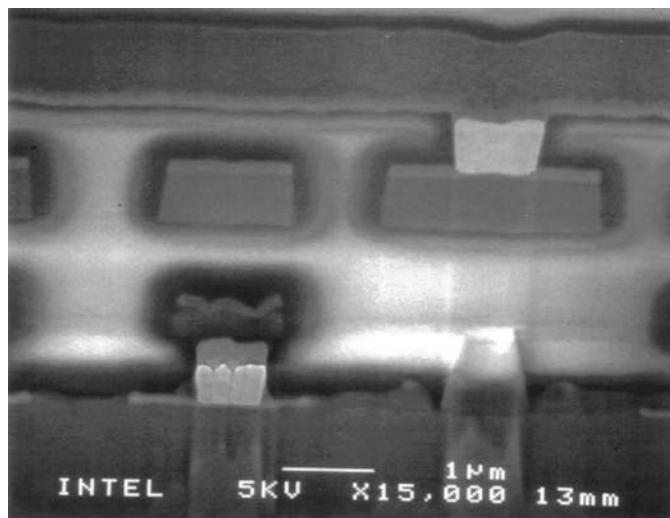
**FIGURE 11.32** Example of waterfall artifacts in an image.

Ridge (Figure 11.33) are horizontal stripes in the surface of the cross section. These are usually due to beam drift while sputtering or insufficient sputtering while in the cleanup or polish phase.

Small droplets sometimes appear on the cross section face (Figure 11.34). They are too small to analyze ( $\sim 20$  nm diameter with  $\sim 20$  nm diameter with  $\sim 10^5$  atoms) and almost certainly contain Ga from the primary beam. It is reported that they are generally easy to remove by sputtering another face or by chemical means.

Another application of FIB is in the manufacture of samples for TEM. These are thin sections that are carved out of a larger sample and then put into a TEM for further analysis (Basile et al., 1992). Figure 11.35 shows a TEM sample prepared in the FIB.

The steps for making a TEM sample are shown in Figure 11.36. This entire process up to lift-out has been completely automated. In the case of using a *dual-beam* machine, one can use a STEM mode to perform some analysis of the thin lamella before transferring the sample to a TEM for further analysis.



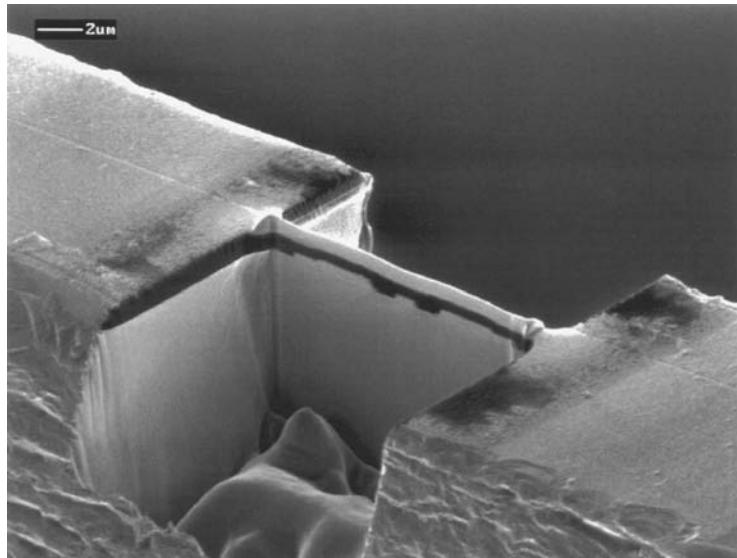
**FIGURE 11.33** Example of ridge artifacts in an image.



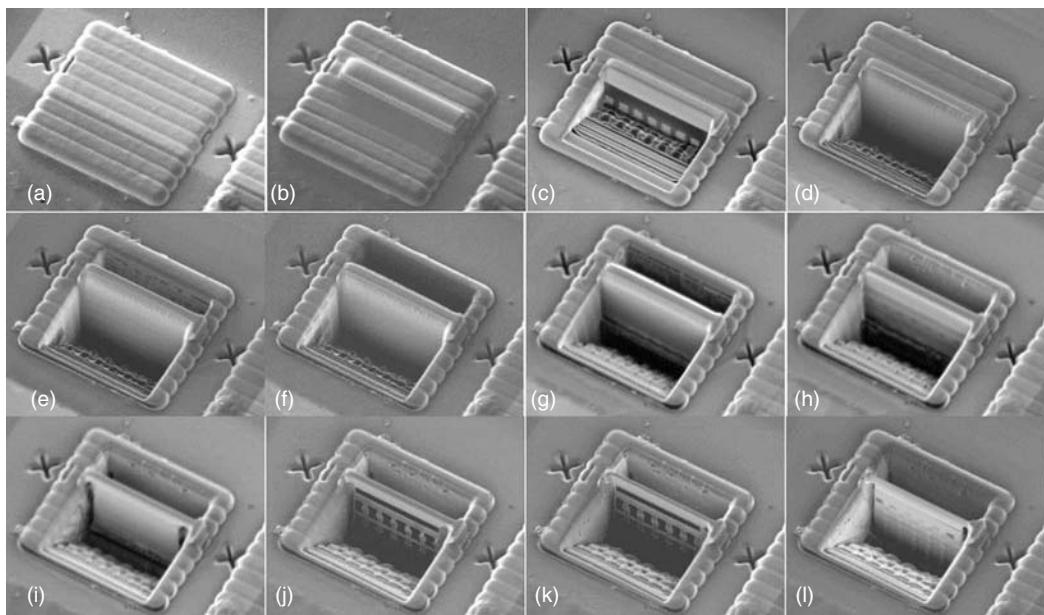
**FIGURE 11.34** Example of Ga droplets in a cross section.

### 11.9.2 GAS-ASSISTED DEPOSITION OF MATERIALS

In a process similar to the known phenomenon of deposition of materials from gases by electron beams, FIB-induced deposition of metals was first reported by Gamo et al. (1984) who studied the process for the deposition of Al from  $\text{Al}(\text{CH})_3$  and W from  $\text{WF}_6$ . The deposition process is important as an adjunct to milling and as a stand-alone process for the direct production of small high-resolution structures without the need of first forming a film that is etched by lithographic techniques. The primary uses of deposition have been for the repair of both optical and x-ray mask clear defects, where precisely defined structures made up of carbon for optical masks and W or Au for x-ray masks are necessary, and in the repair and modification of ICs, where fine low-resistance metal wires and pads are deposited for connections for probing or insulators are deposited for isolation purposes. At present, the capability exists to deposit a variety of metals and insulators.



**FIGURE 11.35** Example of a TEM specimen prepared using FIB.



**FIGURE 11.36** TEM sample preparation steps. (a) Once the site for the TEM sample is located it is identified by milling x's, (b) a protective metal pad is deposited by gas-assisted deposition, (c–d) a front-side staircase pattern is milled, (e–k) then backside staircase milling is performed, and (l) the sides are milled to enable lift-out.

Materials are usually deposited by directing a gaseous compound via capillary needle-sized nozzles over the surface where the deposit is to be made. The molecular gas deposited on the surface is then decomposed by the ion beam where it has been placed, leaving behind the metal, insulator, or carbon. In this process there is a problem of contamination of the residual film by the

**TABLE 11.11**  
**Precursor Gases Used for Metal and Insulator Deposition**

Material Deposited	Precursor Gas	References
Al	TMA (trimethyl aluminum)	Dubner and Wagner (1989)
	TMAA (trimethylamine alane)	Gross et al. (1990)
	TEAA (triethylamine alane)	Gross et al. (1990)
	TIBA (tri-isobutyl aluminum)	Kubena et al. (1988a,b)
W	WF <sub>6</sub>	Gamo et al. (1986)
	W(CO) <sub>6</sub>	Stewart et al. (1989)
	WF <sub>6</sub>	Xu et al. (1989)
Au	C <sub>7</sub> H <sub>7</sub> O <sub>2</sub> F <sub>6</sub> Au	Shedd et al. (1989) Blauner et al. (1989)
Pt	C <sub>9</sub> H <sub>16</sub> Pt	Tao et al. (1990) Puretz and Swanson (1992)
Cu	Cu(hfac) TMVS	Della Rata et al. (1993)
Ta	Ta(OC <sub>2</sub> H <sub>5</sub> ) <sub>5</sub>	Gamo et al. (1986)
Fe	Fe(CO) <sub>5</sub>	Kubena et al. (1988a,b)
Pa	[Pd(O <sub>2</sub> CCH <sub>3</sub> ) <sub>2</sub> ] <sub>3</sub>	Harriott et al. (1986)
TEOS	(C <sub>2</sub> H <sub>5</sub> ) <sub>4</sub> Si	Young and Puretz (1995)
SiO <sub>2</sub>	Si(OCH <sub>3</sub> ) <sub>4</sub>	Kamano et al. (1989)
C	C <sub>10</sub> H <sub>8</sub>	FEI (2003)

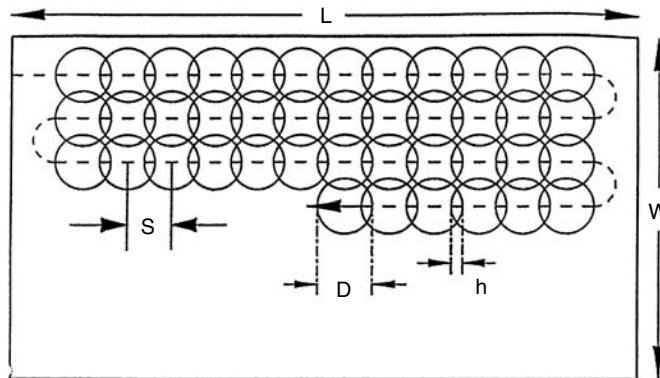
incident gallium or the organic part of the gas. Insulator deposition also has the additional complication of possible charging problems (an insulator is placed where the beam is scanning) and metal contamination by the primary beam (Ga is implanted into the deposition). Table 11.11 shows some of the compounds used for metal and insulator vapor deposition.

In an analysis of the processes that occur during deposition, we need to consider the effects of the rate of gas delivery to the surface and the interaction of the loop time for a pattern and the pixel dwell time of the beam. There are competing processes of layer growth, desorption, dissociation, and sputtering.

There have been two distinct models, which attempt to explain the mechanism of ion-induced deposition. In both models the gas precursor is assumed to adsorb to the surface in a monolayer, with no further adsorption during the ion–surface interaction. The surface onto which deposition occurs is the deposited material.

The thermal spike model assumes that there is a delta function increase in temperature when an ion impacts the surface and that there follows a temperature distribution with time from either a spatially expanding hemispherical flow of heat from a point on the surface or axially symmetric cylindrical flow of heat from an instantaneous line source. The temperature spike can be a few thousand degrees Kelvin with a lifetime of about a tenth of a nanosecond. A first-order rate equation is used to predict the fraction of molecules decomposed. The values predicted from this model are found to be smaller than experimentally measured values by about a factor of half.

The binary collision model is used in the TRIM Monte Carlo program (Zeigler, 1991) to simulate the collision cascade to estimate the number of sputtered atoms N(E) and distribution of excited surface atoms (from knock-on collisions) in the vicinity of the cascade as a function of energy. From this distribution, it is possible to calculate the probability of an excited surface atom colliding with the various atomic constituents of the adsorbed gas molecule. By summing the probabilities into a



**FIGURE 11.37** Schematic diagram of digital scan.

total fraction function  $f(E)$ , an estimation of the deposition yield  $Y_D$  is then

$$Y_D = \int f(E)N(E)dE \quad (11.43)$$

where the limits of integration can be taken from 0.01 to 100 eV.

The results of this model fit the experimental data better than the thermal spike model. It appears that the ion-induced deposition is mediated by the substrate, rather than being a result of the direct interaction of the ion beam with the adsorbed molecules on the surface or whereas the molecules are in the gas phase above the surface.

To understand what effect the controllable parameters in the system have upon deposition, we can make an analysis of the interaction of those parameters. The operator has control of the rate at which ions and gas molecules impinge on the surface, and in a reasonably sophisticated scan system, the operator has control of the pixel dwell and loop times and pixel overlap.

Figure 11.37 shows a schematic diagram of a digital scanning scheme which consists of a beam of diameter  $d$  and step size  $s$  which is boustrophedonically\* scanned over an area of length  $L$  and width  $W$ . For a digital scanning system, the step size is determined by the software by specifying the beam overlap  $OL$  and the magnification  $M$ . The step size is related to the overlap and beam diameter by

$$S = d(1 - OL) \quad (11.44)$$

and the total number of steps in the scanned area  $LW$  is  $N_p = LW/S^2$ . If we neglect beam blanking, the loop time  $t_l$ , which is the time it takes for the beam to cycle though all the points once, is related to  $N_p$  and the pixel dwell time  $t_d$  by

$$t_l = t_d N_p = \frac{t_d LW}{(d(1 - OL))^2} \quad (11.45)$$

We see that once an area has been determined in which it is desired to perform deposition, the loop time has two adjustable scan parameters, the beam overlap and pixel dwell time.

The parameters that determine the ion-induced deposition are listed in Table 11.12. The deposition yield  $Y_D$  (atoms deposited/ion) depends on the adsorbed coverage  $N$ , cross section for

\* Boustrophedonically—an ancient method of writing in which the lines run alternately from right to left and from left to right. From the Greek boustrophedon, like ox-turning (in plowing). See Figure A.1 in Appendix for an image of boustrophedonic writing.

**TABLE 11.12**  
**Parameters Used for Calculating**  
**Metal Deposition**

$g$	Sticking coefficient
$F$	Gas flux
$N_0$	Precursor monolayer density
$J$	Primary ion current density
$\sigma$	Decomposition cross section
$\theta$	Precursor monolayer coverage
$K_d$	Thermal desorption rate constant
$Y$	Film sputter yield

dissociation  $\sigma$ , and sputter yield for the deposited material  $Y_s$  as

$$Y_D = N\sigma - Y_s \quad (11.46)$$

and the mechanism of deposition can be summarized by the following steps:

1. The rate at which the precursor gas is absorbed onto the sample substrate is  $gF(1 - \theta) - k_d N_0 \theta$ . The first term is the rate at which gas molecules stick to the surface and the second term is the rate at which they are thermally desorbed.
2. The rate at which the precursor is disassociated by the incident ion beam is  $JN_0 \theta \sigma / e$ .
3. The rate at which the newly deposited film is sputtered away by the incident beam is  $JY(1 - \theta) / e$ , and the rate at which the adsorbed gas is sputtered is  $JYN_0 \theta / e$ .

The first step determines the amount of precursor adsorbed during the loop time  $t_l$ . The second step determines the amount of precursor converted to metal during the pixel dwell time  $t_d$ . If  $t_l$  is not sufficiently large for monolayer coverage to occur or if  $t_d$  is too long, then sputtering of the metal occurs. Conversely, if  $t_l$  is too long or if  $t_d$  is too small, then even though the overall yield will increase, it may happen that organic radicals from inner layers will become trapped in the film and the resistivity will increase.

Following Puretz and Swanson (1992), by combining these steps, we can write the rate of change in the density of adsorbed molecules  $n = N_0 \theta$  as

$$\frac{dn}{dt} = gF(1 - \theta) - k_d N_0 \theta - \frac{JN_0 \theta \sigma}{e} - \frac{JY(1 - \theta)}{e} \quad (11.47)$$

The rate of deposition  $D$  of metal is then

$$\frac{dD}{dt} = s\theta N_0 \sigma - (1 - \theta) \frac{JY}{e} \quad (11.48)$$

To optimize the deposition parameters that are controllable, namely,  $t_l$ ,  $t_d$ ,  $F$ , and  $J$ , we neglect the thermal desorption of the precursor, and find the relationship between the initial and final coverages  $\theta_i$  and  $\theta_f$  of the precursor during the loop time to be

$$(1 - \theta_f) = (1 - \theta_i) e^{gFt_l/N_0} \quad (11.49)$$

For  $t_d \ll t_l$ , which is almost always the case, and neglecting film sputtering, the relationship between the time  $t_d$  to convert the absorbed precursor and the initial and final coverages is

$$\theta_f = \theta_i e^{J\sigma t_d / N_0} \quad (11.50)$$

The necessary condition to convert all the absorbed precursor to metal can be approximated in the limit  $\theta_i \rightarrow 0$ ,  $\theta_f \rightarrow 1$ ,  $t_l \gg t_d$ , as

$$\frac{t_l Fe}{t_d J} = \frac{N_0 \sigma}{g} \quad (11.51)$$

The left side of Equation 11.51 is the ratio of the gas flux dose during the loop time to the ion dose during a pixel dwell time. If

$$\frac{t_l Fe}{t_d J} < \frac{N_0 \sigma}{g} \quad (11.52)$$

then excess precursor is left unconverted, and may be incorporated into the film. If

$$\frac{t_l Fe}{t_d J} > \frac{N_0 \sigma}{g} \quad (11.53)$$

then after all the precursor have been converted, sputtering of the film removes metal and decreases the amount of deposition.

If we take a monolayer to be  $N_0 = 1 \times 10^{15}$  molecules/cm<sup>2</sup>, the decomposition cross section  $\sigma \sim 1 \times 10^{-15}$  cm<sup>2</sup> and the sticking coefficient  $g \sim 0.1$ , then

$$\frac{t_l Fe}{t_d J} \approx 10 \text{ atoms/ion} \quad (11.54)$$

Puretz and Swanson (1992) have defined a figure of merit  $F_m$  for the deposition process of metals. Two useful measures of the quality of the deposition are the deposition yield  $Y_D$  and the film resistivity  $\rho$ . The resistance of the film in terms of the lateral resistance  $R_s$  ( $\Omega/\square$ ) is

$$R_s = \rho/z \Omega/\square \quad (11.55)$$

where  $z$  is the thickness of the film. The deposition yield (film volume/ion dose) is

$$Y_D = WLz/I_B t \text{ m}^3/\text{C} \quad (11.56)$$

where

$W$  and  $L$  = width and length of the film

$I_B$  and  $t$  = ion current and total deposition time

The figure of merit for the deposition can then be defined as

$$F_m \equiv R_s I_B t / LW \equiv \rho / Y_D \Omega C/\text{cm}^2 \quad (11.57)$$

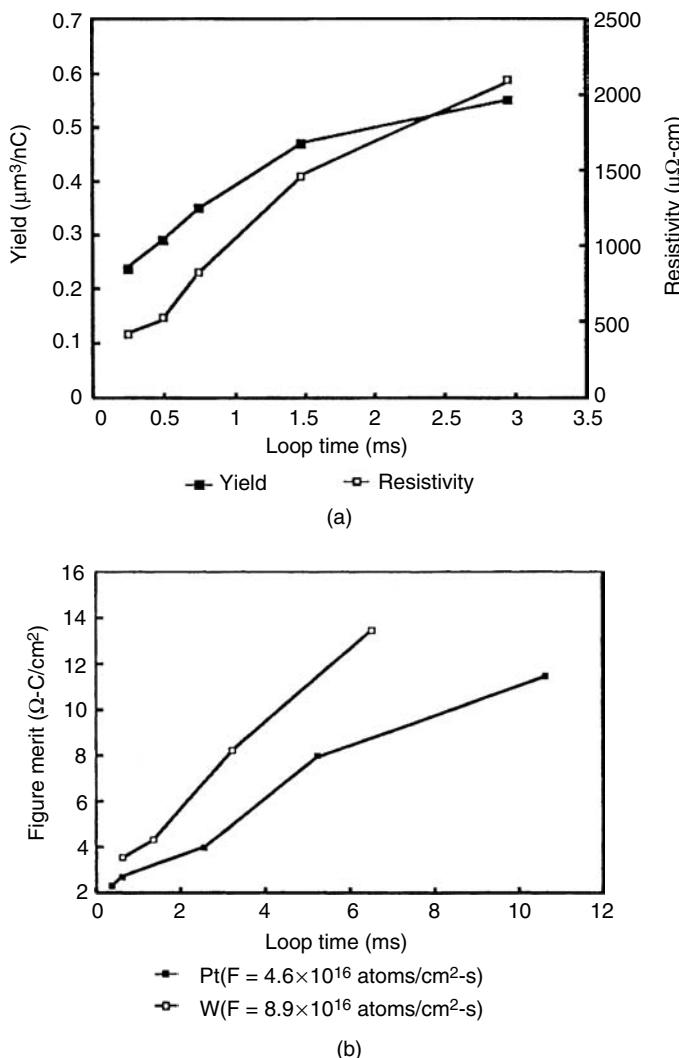
A small value of  $F_m$  is usually desired for low-resistance high-yield films. Once that is known for a given process, then to make a deposition with a particular desired  $R_s$ , the user can set the

quantity  $I_B t / WL$  equal to  $F_m / R_s$ . Since WL is usually already constrained by preexisting geometries on the sample, the user has to set the quantity  $I_B t$ .

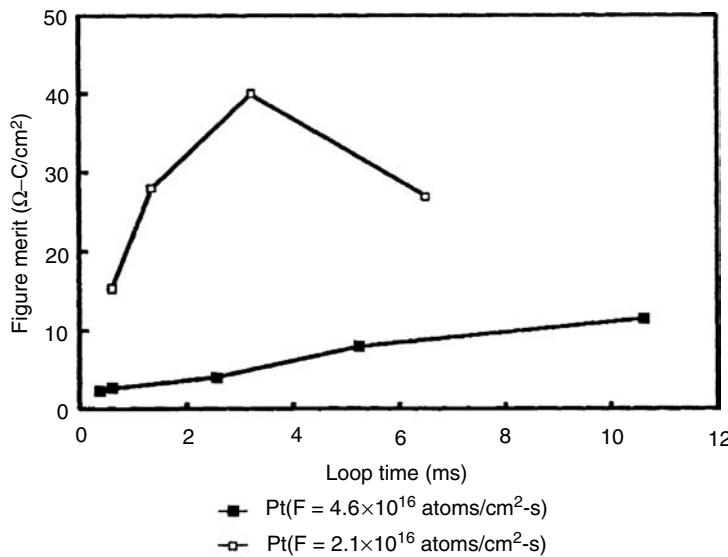
Figure 11.38a shows the yield and resistivity as a function of loop time for Pt deposition, whereas Figure 11.38b compares the figure of merit  $F_m$  for Pt and W deposition. A plot of the figure of merit  $F_m$  for Pt as a function of the ratio of the gas flux to ion flux is shown in Figure 11.39. The broad minimum suggests that 10–50 atoms/ion are deposited in that process.

An example of metal deposition is shown in Figure 11.40, in which a metal strap has been deposited between two tracks of an IC. Note that there is also a cut in the top track to the right of the deposited strap.

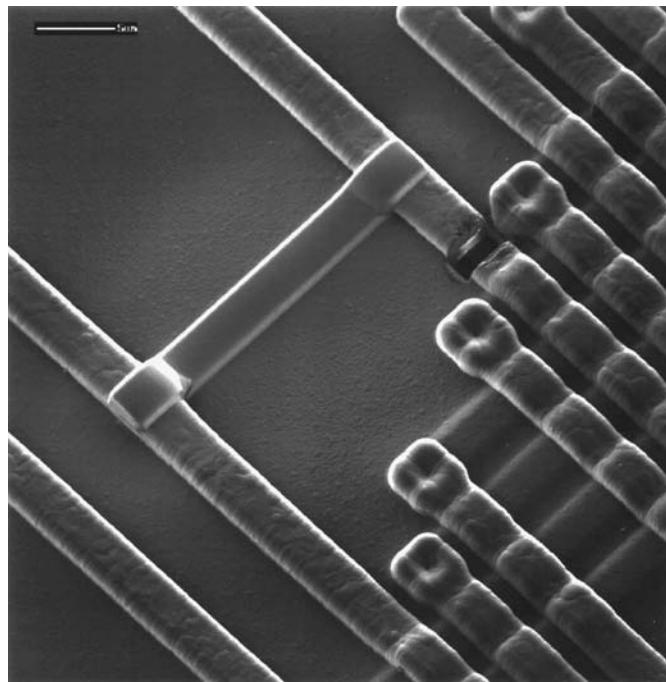
Figure 11.41 shows the use of both metal and insulator deposition. To gain access to metal 1 layer while maintaining electrical isolation from the metal 2 overlayer, an opening was first drilled to the metal 1 layer by sputtering with enhanced etch. Then an insulator was deposited into the hole,



**FIGURE 11.38** (a) Deposition yield and resistivity versus loop time. (From Puretz, J. and Swanson, L.W., *J. Vac. Sci. Technol.*, B10(6), 2695–2698, 1992. With permission.) (b) Comparison of the figure of merit for Pt and W. (From Puretz, J. and Swanson, L.W., *J. Vac. Sci. Technol.*, B10(6), 2695–2698, 1992. With permission.)



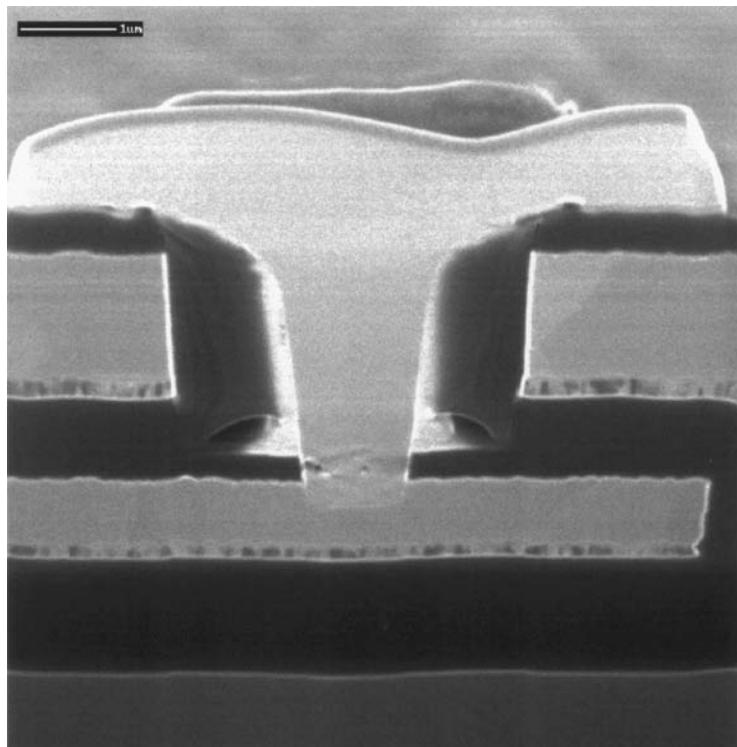
**FIGURE 11.39** Figure of merit for Pt versus gas flux/ion flux. (From Puretz, J. and Swanson, L.W., *J. Vac. Sci. Technol.*, B10(6), 2695–2698, 1992. With permission.)



**FIGURE 11.40** Milling and metal deposition used for circuit design edit. The metal strap to the left, angled 90° to the milled hole in FIB-deposited metal. (Courtesy of FEI Co.)

and a smaller hole opened into the insulator down to metal 1. The hole was then refilled with Pt, and a probe pad was manufactured with Pt also.

The deposition of hydrocarbons is important for mask repair. The analysis is similar to that found previously in this section, and the reader is referred to the book by Prewett and Mair (1991).



**FIGURE 11.41** Metal deposition and insulator deposition. Access to metal 1 through metal 2. The top probe pad is FIB-deposited metal surrounded by FIB-deposited insulator. (Courtesy of R. Young, FEI Co.)

### 11.9.3 SCANNING ION MICROSCOPY

SIM is similar to SEM in which charged particles are focused onto the surface of a sample, rastered across it, allowing image formation by the modulation of some sort of synchronously rastered display unit by interactions of the beam with the sample. At present, the highest resolution obtainable using FIB is slightly  $<5$  nm with an S/N ratio of about 20. The prospect of going much higher in resolution appears to be slim, as we discussed earlier. If the S/N ratio is sacrificed to the minimum level of about 5, then it is possible to image at a resolution near 1–3 nm.

The signals available in SIM are ion-induced secondary electrons (ISE), ion-induced secondary ions (ISI), ion-induced secondary atoms (ISA), and photons. The use of photons in SIM is not characterized or used because of the low yield for which most situations make it an impractical mode. It will not be discussed here. ISA is essentially sputtering and has not been used as a source for a contrast signal.

ISE emission coefficients ( $\gamma_e$ ) generally increase monotonically with incident energy in the range 10–100 keV, and typically have values of 1–5 for metals. There are two channels for ISE. At low ion energies, electrons are ejected when the ion is close to the surface by an Auger or resonance process (potential emission). In this process, when an ion approaches the surface of a metal within a few angstroms, the probability of neutralization by an Auger process becomes large. When a conduction band electron neutralizes the ion to its ground state an excess energy can be transferred (radiationless) to another conduction electron. With sufficient energy it can escape.

If the kinetic energy of the incoming ion is high enough, electrons can be ejected when the ion causes an inner shell ionization of an atom in the solid (kinetic emission). Electronic stopping is primarily responsible for kinetic ISE emission, and starts to become noticeable for ion velocities above  $10^5$  m/s (about 37 keV for  $\text{Ga}^+$ ). As a result of the inner shell vacancies created, electrons created

by an Auger process undergo collisions inside the target and may be eventually emitted from the surface. This type of emission strongly depends on the oxidation state of the surface. The energy distribution of secondary electrons typically has a maximum at around 5 eV and falls to 1% of the maximum at 15–20 eV.

For crystalline materials  $\gamma_e$  depends on the relative orientation of the incident beam to the lattice which strongly modulates  $\gamma_e$  with crystal orientation, leading to channeling contrast. The reason for this phenomenon is that ions impinging on an axial channeling direction of a crystal suffer energy loss dominated by electronic processes. By entering the crystal in a channel, nuclear collisions are effectively eliminated. In addition, electronic stopping is strongly reduced relative to amorphous materials because of the reduced electron density in the center of the channel. ISE is sharply reduced along these directions, and since well-channeled ions suffer few strong atomic collisions, both ISI and ISA are reduced along the channeling directions. Hence there is reduced sputtering.

Since there is a significantly reduced rate of energy loss for ions along a channel, the consequence is that the range increases, sometimes dramatically. Levi-Setti (1983) has pointed out that there seems to be a *remarkable coincidence* that for  $\text{Ga}^+$ , the electronic stopping power is near a minimum of  $Z_1$  oscillations, so that the increased range in channeling can be up to 50 times that of unchanneled ions, and that to reveal channeling effects in SIM, *the obvious choice would be  $\text{Ga}^+$* . The optics of the SIM also enhances the channeling effect because of the extreme collimation in the beam, which reduces the transverse component of the beam momentum.

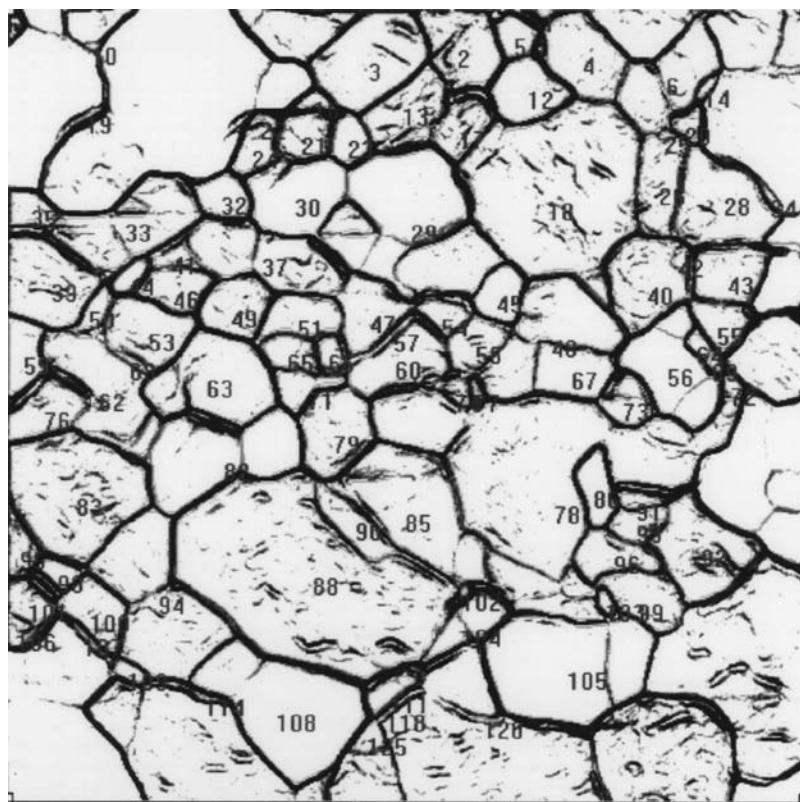
To produce channeling in an image it is usually necessary to remove the oxide layer, which scatters the incoming ions, and it is necessary to scan fairly slowly and to use sufficient beam current to continuously remove the oxide film that forms from the ambient gas. In general, the vacuum level should be below  $5 \times 10^{-7}$  torr. It should be noted that when channeling occurs, there will be reduced sputtering along the channeled axes, leading to a surface with topographic features.

An example of channeling is shown in Figure 11.42, which shows one analysis possible with channeling contrast, in which the grains are outlined, counted, and their sizes are measured. The resulting distribution of areas is also shown in the figure.

The ISA coefficients are essentially the sputter yields, and decrease slowly above a broad maximum around 30 keV for most LMIS species. Values in the range of 2–12 are typical, and because of shell effects large  $Z_2$  oscillations are seen. Since sputtering is used as a tool to remove material and the sputtered neutral atoms are not collected to form an image, to date there is no utilization of these atoms to form an image directly. The action of sputtering contributes to image formation by creating topographic structure from which emission of secondary electrons and ions can yield information.

ISI coefficients are less well known than ISE coefficients. The value of the coefficient is much less than the ISE coefficient, and in the general range of  $10^{-2}$  to  $10^{-4}$ . There is, of course, the added complication that the ions might be emitted in either positive or negative charge states and that molecules and fragments are formed. Useful data for  $\text{Ga}^+$  is sparse, and trends must be inferred from data obtained from other ions. The surface chemistry of the sample in addition to the energy and  $Z$  of the incident ions has a dramatic effect on the ISI coefficients. The presence of oxygen can greatly enhance the ISI yield. The enhancement depends on the amount of saturation and varies between factors of 10 and 1000. The uncertainty in the contribution of oxygen to enhanced yields makes quantification difficult not only in imaging but also in SIMS.

To understand the contrast mechanisms and the information content of an image generated by SIM, it is necessary to consider the escape depths of secondary electrons and ions. The main difference between SEM and SIM is the depth at which the primary energy is deposited in the sample (the range), which is generally shorter for ions than for electrons of the same energy. The mean escape depth for most of the ISE electrons is comparable with the ion penetration depths ( $\sim 10$  nm) for energies in the range of 20–40 keV. This means that the SIM ISE signal should be more sensitive to the surface topography for amorphous materials than the total secondary-electron signal in SEM, which has a significant contribution from long-range backscattered electrons. For crystalline materials,



**FIGURE 11.42** Channeling contrast in Al. Grain outlines and resulting distribution of sizes. (Courtesy of R. Young, FEI Co.)

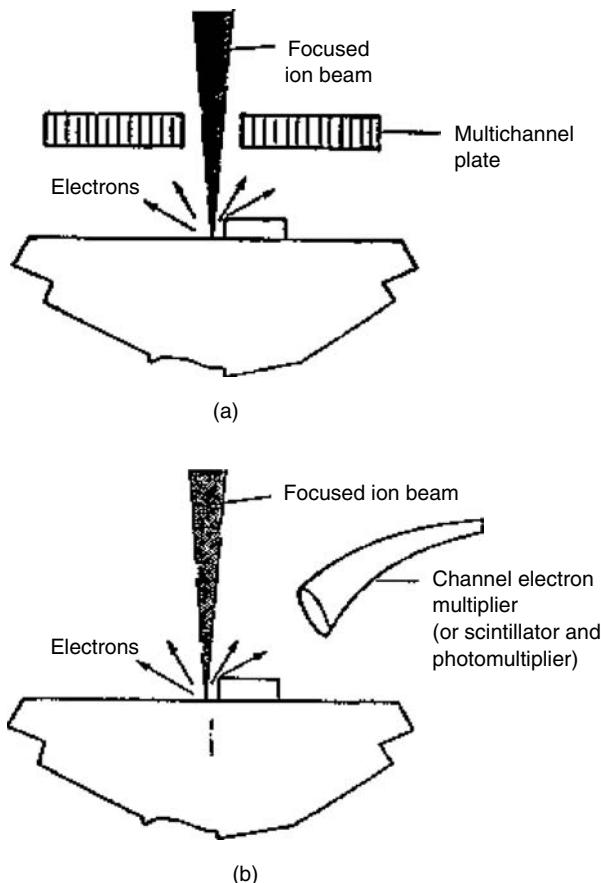
large differences in the dependence of ISE yields between channeled and unchanneled energy loss rates mean that the ISE signal can carry information about the bulk structure of the crystal.

The ISI signal arises from a two-step process in which sputtered atoms of a mean escape depth of 0.5–0.8 nm in amorphous materials become ions in a charge-exchange process that occurs a fraction of a nanometer above the surface of the material, and is strongly affected by the local surface chemistry, so that it is a true surface signal. For crystalline materials, sputtered atoms may be emitted from depths of 5–10 nm preferentially along lattice directions due to assisted focusing collision sequences.

For both the ISE and ISI signal, since the escape of either secondary electrons or ions is facilitated by incident ion energy being deposited along a shallow angle trajectory, the yields are proportional to  $1/\cos(\theta)$ , where  $\theta$  is the beam angle to the surface normal.

The detectors used in FIB are the CEM, the MCP, and the scintillator. The first two are capable of detecting either secondary electrons or ions, whereas the scintillator is an electron detector. It is possible to make the scintillator an ion detector by first converting the ions to electrons by using a conversion dynode and then detecting them. Key issues involved in the choice of detector are (1) gain, (2) collection efficiency, (3) noise, and (4) lifetime.

The CEM and MCP generate an electron signal by electron multiplication on a high-secondary-electron yield material. The gains are in the range of  $10^6$  to  $10^9$ . Figure 11.43 shows the schematic arrangement for these detectors. Typically, a grid at a potential of a few hundred volts can be placed in front of either to image with either electrons (grid positive) or ions (grid negative). The noise of each depends on the gain across the detector, and its life history. There seems to be much folklore concerning the use and rejuvenation of these detectors. In UHV systems where relatively



**FIGURE 11.43** Schematic diagram of detectors. (a) channeltron (CEM), (b) MCP.

small beam currents are used and no gases are introduced for enhanced sputtering or deposition, it is reported that CEM can operate satisfactorily for nearly a decade. A daily used system in which there is only high vacuum and for extended times high beam currents and corrosive gases are used, the lifetimes can be as short as 6 months. Rejuvenation seems to work sometimes, but the length of the rejuvenated period appears to be disappointingly short.

Scintillators are used in the Everhart–Thornley configuration, in which electrons strike a scintillating material, and the resultant light is reconverted to electrons and multiplied in a photomultiplier tube. These can be made with low noise figures, and with suitable care can last several years. There can be a concern of chemical reactions with the thin conductive coating on the scintillator.

There are differences in images obtained with CEM, MCP, and scintillator. The CEM and scintillator images are slightly different than the MCP image due to the asymmetry of collection of the former. The MCP can basically collect signal from both sides of an object, while the position to one side of the scanned area for the CEM and scintillator give the images a more topographic view. It should be noted that when forming ion images the image quality can be made excellent (near the same S/N as the electrons images), but at the expense of consuming more sample. This is because of the relatively low yield coefficients for ions as compared with electrons. Introducing oxygen onto the surface during imaging when channeling is not a concern should increase the ion yield to make better images for less loss of sample.

SIMION (Dahl, 1990) analysis of trajectories of ions and electrons indicates that the collection efficiency of the MCP is superior to the other two detectors. Experiments measuring the S/N ratio

of the MCP indicate that it is lower than for the CEM and scintillator, but most geometric arrangements used in FIB for secondary collection, all the three show similar S/N ratios. The slightly lower gain and higher noise of the MCP are offset by its greater collection efficiency.

We reiterate that the rule of thumb for obtaining high-quality micrographs pleasing to the eye is to have an S/N ratio of at least 20, and not much higher. The normal human eye is incapable of perceiving more than about 20 gray levels, and the use of more gray levels wastes sample. To make an S/N ratio of 20, about 400 quanta are necessary per pixel. Practical experience has shown that in fact about  $700 \pm 200$  ions are necessary incident per pixel for secondary-electron imaging and 2–3 times higher for secondary-ion imaging.

#### 11.9.4 SECONDARY-ION MASS SPECTROMETRY

Some fraction of the sputtered atoms at the target are ionized. If these ions are collected and mass analyzed, then the destructive properties of the FIB can be advantageously used for the chemical analysis of surfaces. Under suitable conditions it is possible to control the rate of erosion caused by the primary beam, so that either two- or three-dimensional chemical maps can be obtained. Due to the high spatial resolution inherent in FIB, the chemical maps obtained are also of high lateral spatial resolution, and under suitable conditions the depth resolution can also be very high. There are in relation to other surface analytical techniques, several trade-offs, which must be mentioned when considering which technique to use.

The advantages of FIB/SIMS are

1. The lateral and depth resolution are high. Laterally, 20 nm is possible, and in depth 1 nm is common.
2. The technique is sensitive to all elements and molecular fragments. The minimum detectable concentration can be as good as a few hundred parts per million, whereas for energy dispersive x-ray analysis the minimum detectable concentration is between a part per million and a part per hundred.
3. The S/N ratio can be  $>10^6$ .
4. The technique is sensitive to isotopes of the elements, so that the use of markers is possible.

The disadvantages of FIB/SIMS are

1. Quantification is usually difficult. There are large variations in sputtering and ionization yields.
2. The condition of the sample surface is critical. Surface oxides greatly enhance the ion yield. Ion yields are sensitive to the system base pressure.
3. The sample is consumed as it is being analyzed. No reexamination is possible, so that parallel collection techniques must be used in certain cases.
4. Primary ions are implanted into the sample during analysis and may react with it or confuse the interpretation of data.

There are three main types of FIB/SIMS systems: (1) RF quadrupole, (2) time of flight (TOF), and (3) magnetic sector. The RF quadrupole and magnetic sector use a standard FIB column to form a probe which is scanned across the sample, whereas secondary ions are collected and transported to the RF quadrupole or magnetic sector in which the ions are filtered and detected in a serial mode when the mass channel is swept through its range whereas the primary beam continuously strikes the sample. The TOF requires that a monoisotopic source material to be used, and that the primary beam to be pulsed (blanked off-on-off) for about 5 ns. The secondary ions produced at the sample are collected and directed into a flight tube after having fallen through an accelerating potential

drop. The ions are timed in their flight down the tube and detected in a parallel mode in which all mass channels are open for detection. In principle, almost all secondary ions can be detected, yielding very high sensitivity.

The evolution of the FIB/SIMS has followed a path starting with Castaing and Slodzian (1962) who used a stationary broad beam ( $\sim 0.5$  mm) of Ar ions directed to a target, from which the secondary ions were stigmatically imaged through a  $90^\circ$  sector magnet onto a fluorescent screen, whereby the entire area analyzed is viewed simultaneously. An ultimate spatial resolution of mass-resolved images is  $\sim 0.5$   $\mu\text{m}$  and is limited by the chromatic aberrations of the system due to the energy spread of the secondary ions.

Liebl (1967) was the first to use a scanning ion probe with SIMS. This instrument used a duoplasmatron source, producing a focused probe of  $1.5 \mu\text{m}$ , and used a double-focusing spectrometer for mass and energy filtering.

That instrument was improved upon most notably by Liebl (1972) by making short-focal-length lenses which simultaneously focused the primary beam onto the sample and secondary ions into a spectrometer. Slodzian et al. (1992) later developed a similar instrument. Rudenauer et al. (1979) used an RF quadrupole mass spectrometer and an electron bombardment ion source.

The LMIS which has about  $10^4$  times the brightness of other sources revolutionized scanning probe SIMS. Levi-Setti (1983) showed that it is possible to produce useful analytical information from probes as small as 20 nm. Currently, FIB/SIMS is used in obtaining high-resolution two-dimensional chemical maps, depth profiles, and spectra. The small beam size allows the analysis of tiny particles, the study of spatially segregated materials, and the use of isotope labeling in biological investigations, requiring high spatial resolution.

Figure 11.44 is a schematic diagram of the FIB/SIMS system which was jointly developed by the University of Chicago and HRL in the early 1980s (Levi-Setti et al., 1984). The system consists of a two-lens imaging column capable of 20 nm imaging resolution with 40 keV Ga ions. A prelens double deflection octopole scan system rasters the probe over a sample, from which secondary ions are collected by miniature electrostatic energy spectrometer, and transported to an RF quadrupole mass filter from which the results of mass filtering are detected by a CEM.

Levi-Setti (Chabala et al., 1992, 1994) has recently enlarged the Chicago FIB/SIMS to include a magnetic sector (Figure 11.45). The secondary transport system of this instrument consists of an extraction plate that produces a field ( $\pm 300$  V), forcing ions to enter a small  $90^\circ$  electrostatic spherical analyzer, which acts also to deflect the secondary ions out of the primary column. The ions are accelerated to 1 keV before the exit aperture of the spherical analyzer to increase transmission and mass resolution by reducing the chromatically aberrated width of the beam. The secondary-ion beam is then focused, shaped, and accelerated to 5 keV by a series of einzel lenses and a quadrupole. This optical arrangement can be used as a zoom lens to change the size of the beam at the entrance slit of the spectrometer.

There is a *switchyard* deflector to steer the beam either into the RF quadrupole or into the magnetic sector. The deflector can be used to descans the action of the primary beam at the sample, allowing a dynamic emittance matching into the sector magnet, so that large areas of the sample can be analyzed at increased mass resolution. Without emittance matching, only areas of about  $20 \times 20 \mu\text{m}$  at a mass resolution of  $M/\Delta M = 2000$  can be scanned although there is no practical area limitation when matching is used.

The magnetic sector is a modified Finnigan MAT uses a  $65^\circ$  magnet followed by a  $90^\circ$  electrostatic sector. The mass range is 0–3000. The ion detector is an active film pulse counting electron multiplier, which works with count rates up to 50 MHz.

A commercial FIB/SIMS system produced by FEI Co. is an RF quadrupole-based system that has a mass range of 0–200 amu. The extraction optics is a cylindrical extraction *snout* followed by a series of transport lenses, which carry secondary ions to the RF quadrupole. An off-axis detector is used to discriminate against neutrals.

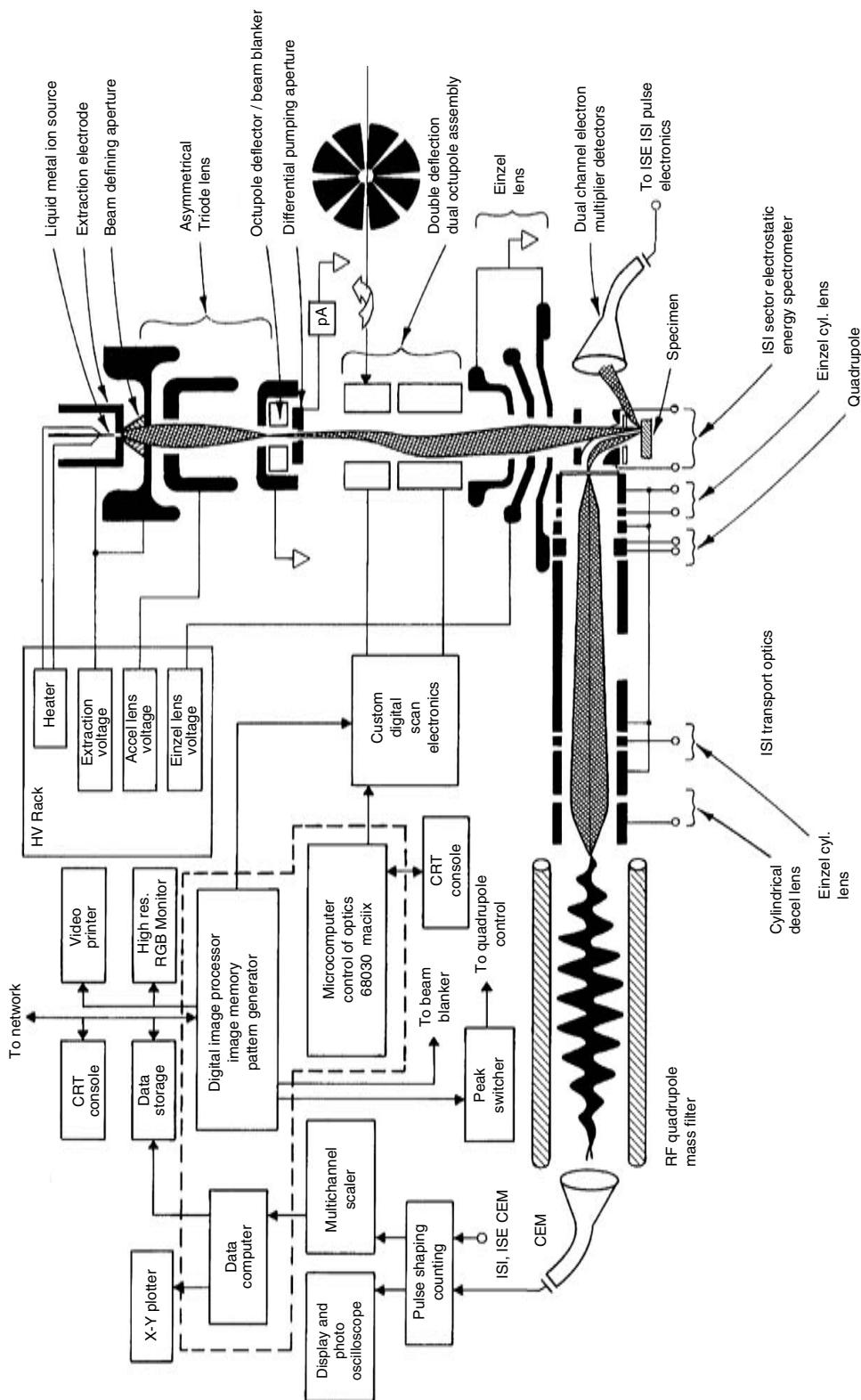
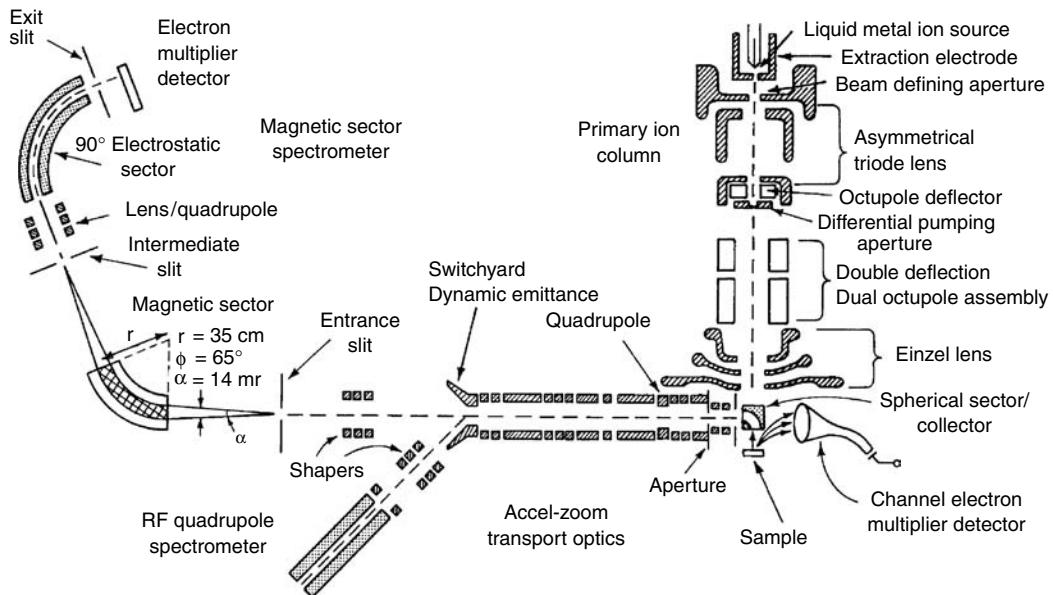


FIGURE 11.44 Schematic diagram of UC-HRL SIMS. (Courtesy of R. Levi-Setti, University of Chicago.)



**FIGURE 11.45** Schematic diagram of UC FIB/SIMS secondary-ion mass spectroscopy. (Courtesy of R. Levi-Setti, University of Chicago.)

FIB/SIMS is inherently a dynamic SIMS technique, because a large fraction of a monolayer of material is consumed for each map. Static SIMS measurements erode less than a tenth of a monolayer, and only recently have been able to demonstrate micrometer spatial resolution.

There are several quantities, which are important for judging how well a system is working. At present all FIB/SIMS use a  $\text{Ga}^+$  beam, so that the interaction of interest is between  $\text{Ga}^+$  and the surface. The sputter yield and ionization probability are the two most important quantities for the beam–surface interaction. Quantities that describe the performance of a given SIMS instrument are measured secondary-ion yield, the minimum detectable concentration the useful yield, and practical sensitivity.

The sputter yield  $Y$  is the number of sputtered atoms per incident primary ion. This is related to the rate at which the sample is consumed for a given beam current  $I_B$  and material. If the total sputter yield of all species is  $Y_{\text{TOT}}$ , then

$$Y = \frac{Y_{\text{TOT}}e}{I_B t} \quad (11.58)$$

The measured secondary-ion yield  $Y_A^\pm$  of an element or molecule A is an instrument-dependent quantity,

$$Y_A^\pm = I_B Y_A \gamma_A^\pm \eta C_A M_A \quad (11.59)$$

where

$I_B$  = the primary beam current

$Y_A$  = the sputter yield for A

$\gamma_A^\pm$  = the positive or negative ionization probability of A

$\eta$  = the overall efficiency of collection, transport, and detection of the instrument

$C_A$  = the fractional concentration of A at the surface

$M_A$  = the fraction of particles that are produced at the selected mass/charge of the mass spectrometer.

The only term in Equation 11.59 that is known straightaway is  $I_B$ ; every other term must be measured from calibration standards. Levi-Setti (Chabala et al., 1994) has estimated that  $\eta$  for the RF quadrupole-based system at Chicago is 0.2%, whereas for the magnetic sector system it is 4–20%. FEI Co. estimate that  $\eta$  for their commercial RF quadrupole system is 0.1–0.3% (Crow et al., 1995). Efficiencies for TOF systems can be as high as 70%.

Measuring the efficiency of transmission can be difficult, but one method that is fairly accurate utilizes the measurement of total stage current while collecting positive ion with the SIMS. In this procedure, positive ions are collected by biasing the transport optics extraction plate negative ( $-200$  V) which forces virtually all (except the few with energy  $>200$  eV) secondary electrons and negative ions to return to the sample, so that the stage current reading is the difference between the positive primary ion current and the positive secondary-ion current. The primary ion current can be measured by blanking the beam so that the difference between the primary ion current and the stage current with the beam unblanked is the total positive secondary-ion current. Measurements can be made for elements, and the isotopic distribution must be taken into account.

The minimum detectable concentration or sensitivity for a species A is the smallest concentration of A that can be measured with adequate statistics from a limited volume of material with a given instrument. In a typical SIMS map about  $2000\text{ nm}^3$  is eroded, which contains  $10^5$  atoms. For an instrument that has perfect efficiency and for total ionization, concentrations on the order of a part per thousand can be measured with statistical confidence. It is found that in practice a part per hundred can be achieved. Any instrument can obtain parts-per-billion resolution if there is virtually unlimited sample to be consumed.

The useful yield  $\tau_A(A)$  is the ratio of the detected SIMS signal for A ( $Y_A^\pm$ ) to the total number of sputtered particles of A, which is the product of the ionization probability and the instrumental sensitivity,

$$\tau_A(A) = \frac{Y_A^\pm}{I_B Y_A C_A M_A} = \gamma_A^\pm \eta \quad (11.60)$$

This dimensionless quantity is independent of the primary beam current and the sputtered volume, and it is useful for estimating the time required and the volume of material necessary to reach a given minimum concentration sensitivity. This also provides a means for comparing the efficiencies of instruments.

The practical sensitivity is the number of counts per second measured per beam current. Usually the unit of current used is the nA, so that the practical sensitivity is counts per second per nanoampere. This is useful for estimating the quantity of signal that can be expected for a given beam current. The assumption is that the signal rate scales with current, and experiments indicate that within the range of currents used in FIB/SIMS (10 pA to 2 nA) this scaling holds true.

Table 11.13 shows the useful yields and practical sensitivities measured in the FEI SIMS system. The increases in useful yield and practical sensitivity are also shown in the table, which result from the introduction of oxygen and iodine into the system (Crow et al., 1995).

There is one complication to measuring useful yields termed the *matrix effect*, which causes the yields to change by a factor of 10–1000×. Table 11.14 shows the effect as measured by Leslie (1994) for 25 keV  $\text{Ga}^+$ . The effect here arises from the presence of oxygen.

FIB/SIMS is usually operated in three modes: mass spectrum acquisition, elemental mapping, and depth profiling. In the acquisition of mass spectra, the primary beam is rastered over the sample either in a full frame mode or in a user-defined pattern. The primary beam generates ions from the sample that are collected and transported to the mass filter, which is being swept through a preset mass range. The output is a plot of signal intensity versus mass.

Table 11.15 shows the relative probability when using  $\text{Ga}^+$  for detecting an element as a positive or negative ion. Numbers greater than one indicate a likelihood of detecting a positive ion.

**TABLE 11.13**  
**Useful Yields and Practical Sensitivities of Elements for**  
**Background Gases (Bkgnd), Oxygen (O), and Iodine (I)**

Element	Gas	Useful Yield	Practical Sensitivity (Hz/nA)
Al	Bkgnd	$4.3 \times 10^{-5}$	$1.48 \times 10^5$
	O	$4.6 \times 10^{-4}$	$1.6 \times 10^6$
	I	$5.4 \times 10^{-4}$	$1.0 \times 10^7$
Si	Bkgnd	$4.4 \times 10^{-6}$	$1.25 \times 10^3$
	O	$2.6 \times 10^{-5}$	$7.0 \times 10^4$
	I	$1.2 \times 10^{-5}$	$1.31 \times 10^5$
Ti	Bkgnd	$8.3 \times 10^{-6}$	$4.5 \times 10^4$
	O	$4.5 \times 10^{-4}$	$2.5 \times 10^6$
	I	$6.9 \times 10^{-5}$	$1.55 \times 10^6$
Mo	Bkgnd	$9.3 \times 10^{-7}$	$2.1 \times 10^3$
	O	$8.7 \times 10^{-5}$	$2.0 \times 10^5$
	I	$8.3 \times 10^{-6}$	$7.5 \times 10^4$
W	Bkgnd	$3 \times 10^{-7}$	$5.35 \times 10^2$
	O	$3.8 \times 10^{-6}$	$7.45 \times 10^3$
	I	$3.0 \times 10^{-7}$	$1.67 \times 10^3$

Source: Courtesy of A. Leslie.

**TABLE 11.14**  
**Table of Useful Yield Showing**  
**the Matrix Effect**

Ion	Matrix	Useful Yield
$\text{Al}^+$	Al	$1.18 \times 10^{-6}$
	$\text{Al}_2\text{O}_3$	$2.95 \times 10^{-4}$
Ti	Ti	$1.93 \times 10^{-8}$
	$\text{TiO}$	$3.06 \times 10^{-7}$
	Tin	$4.44 \times 10^{-8}$
$\text{Cu}^+$	Cu	$2.61 \times 10^{-9}$
	$\text{Cu}_2\text{O}$	$3.63 \times 10^{-7}$
$\text{Fe}^+$	Fe	$1.04 \times 10^{-8}$
	$\text{FeO}$	$3.17 \times 10^{-7}$
	$\text{Fe}_1\text{O}_3$	$2.78 \times 10^{-7}$
$\text{O}^-$	$\text{Al}_2\text{O}_3$	$1.24 \times 10^{-7}$
	$\text{TiO}$	$1.5 \times 10^{-7}$

Source: Courtesy of A. Leslie.

In elemental mapping, the spectrometer is tuned to a specific mass peak, and whereas the primary beam scans over the sample, an image is generated in which the brightness level is derived from the output of the spectrometer signal. A two-dimensional chemical map is obtained in this manner. It is also possible by *peak switching* to collect several maps nearly simultaneously, by switching the selected mass peaks during the retrace period in the raster, cyclically collecting different masses on sequential raster lines. The images can later be deconvolved.

**TABLE 11.15**  
**Table of Relative Probabilities for Ga**

Ion	± Ion Yield Ratio
C	0.02
O	0.08
F	0.01
Mg	3.1
Al	$10^5$
Si	10
S	0.08
Ti	100
Cu	20
Sr	10

*Source:* Courtesy of A. Leslie.

Depth profiling (Benninghoven et al., 1987; Crow, 1992) is used to generate a one-dimensional map of a mass or masses as a function of depth. Typically, an area of the sample is milled by a user-defined pattern while ions are collected as a function of time, which is related to the depth drilled into the sample. For a given pattern size, if the primary current is known as well as the volume sputtered per charge for various materials, then an accurate vertical profile of a material can be measured.

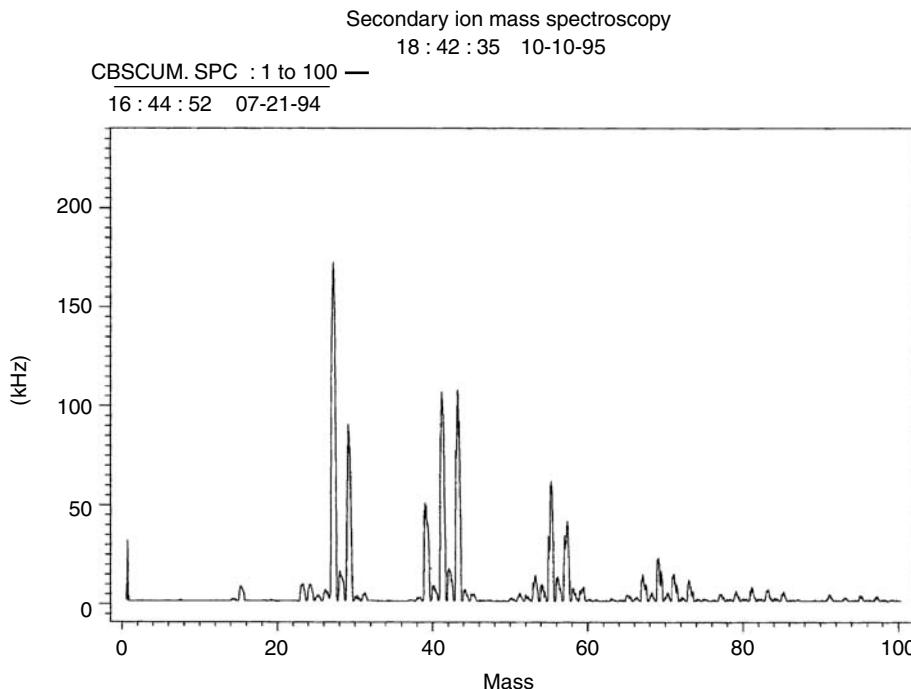
In practice there are considerations of available sample volume, primary beam current, and milling rate. Since the ion yield can vary about five orders of magnitude, the smallest volume that will give a usable amount of data varies over this range also. The proper beam current used for small available volumes must be sufficient to ensure proper spatial resolution, and then it is a trade-off between milling, which is quick enough to minimize the contribution from adsorbed species from the residual vacuum and slow enough to ensure that there are no aliasing effects (see the following paragraph). For example, with a background pressure of  $10^{-7}$  torr the arrival of background residuals is about 0.1 monolayers per second, so that it is probably reasonable to use a minimum milling rate of about 1 monolayers per second. A sense of the scale involved is that for an area of interest of a micron square and a tenth of a micron thick, a beam current of 40 pA will typically remove material at about 100 monolayers per second. Although the beam size would be much less than the feature size, the removal rate is too high because the total milling time is about 3 s, which is too short because of aliasing. Reducing the current to 5–10 pA that would stretch the acquisition time to overcome aliasing. Table 11.16 shows the monolayer removal rate in monolayers removed per second of beam time on the sample. This rate is nominally for 25–30 keV Ga, and is for materials commonly used in semiconductors in which the sputter yield is about 2–4.

Aliasing arises because of a nonsynchronicity between the beam, scanning the sample and mass spectrometer, scanning the mass range. It is possible that during a mass scan the primary beam may pass over the sample only a few times and that the portion of the sample that contains a specific element will not be visited at the same time that the mass spectrometer is tuned to the appropriate mass peak. This is part of the problem of serial data collection from sources with finite volumes.

For example, in certain survey spectra, where the operator is attempting to ascertain what materials are present, the mass scan range might be set to 1–70 amu. If the dwell time per channel is 5 ms and the number of bins/amu is 20 then the full mass scan will take 7 s. If the beam was scanned over the sample at a rate of 0.7 s per frame, then there would be 10 scans of the sample area during the mass scan, and if a specific material was localized within the scan area, it is highly unlikely that the beam would be on that area while the spectrometer was tuned to the appropriate mass.

**TABLE 11.16**  
**Monolayer Removal Rate (Monolayer/s)**

Current	Field of View				
	1 × 1 (μm)	3 × 3 (μm)	10 × 10 (μm)	30 × 30 (μm)	100 × 100 (μm)
10 pA	11	1.2	0.11	0.012	0.0011
30 pA	33	3.7	0.33	0.037	0.0033
100 pA	110	12	1.1	0.12	0.011
300 pA	330	37	3.3	0.37	0.033
1 nA	1100	120	11.0	1.2	0.11



**FIGURE 11.46** FIB/SIMS spectrum.

To minimize the effects of aliasing, it is preferable to use as long an acquisition time per mass bin and to scan the beam as fast as is reasonable over the sample, to maximize the number of times that each point on the sample is visited per mass bin. Figure 11.46 shows a typical survey spectrum from a wafer.

There are some practical considerations when interpreting spectra. The individual mass peaks can be obtained from an isotope of an element from the sample or primary beam, or from a multiply charged or fractionally charged isotope from the sample or primary beam, or from a multiply or fractionally charged complex.

The relative heights of the peaks in the spectra are not a good guide to the relative abundance of those elements because the yields may vary by orders of magnitude. For example, if a sample consisted mostly of gold with a few percent of aluminum, the mass spectrum would indicate almost exclusively the presence of the aluminum because of the great difference in ion yields. This is an extreme case, but it illustrates that it is possible to generate a spectrum with peak heights that are inversely related to the expected abundances.

Another major consideration is the effect that residual oxygen plays in the interpretation of spectra. Typically for the vacuum levels in most FIB systems ( $10^{-6}$  to  $10^{-7}$  torr) the presence of oxygen (Benninghoven et al., 1987) will increase the ion yield by a factor of 1–100. For metals like Al, the increase in yield can be quite dramatic, and is usually accompanied by the presence in the spectra of the oxide of the metal also. Since not all the ion yields are equally increased in the presence of oxygen, the effect can also skew the interpretation of peak height as representing abundance.

Mass interferences are a routine problem in any mass spectrometry. These occur when an elemental ion and some other complex ion have the same charge-to-mass ratio. The common examples are Al and  $C_2H_3$ . These two peaks are actually separated by about 0.04 amu. This is an unresolvable difference in a quadrupole system, but might be differentiated in a TOF or magnetic sector spectrometer. It is necessary to examine some other information to properly attribute an ion to a interference. For example, the mass spectrum from a piece of Al, which has not been properly cleaned or has come from a vacuum system that used oil will have large peaks at masses 29, 41, 43, 55, and 57 which are due to the hydrocarbon fragments  $C_2H_5$ ,  $C_3H_5$ ,  $C_3H_7$ ,  $C_4H_7$ , and  $C_4H_9$ . The peak at mass 27 is due to Al and  $C_2H_3$ .

There are also isotopic interferences. The SIMS yield for all isotopes of a given element is essentially the same, so that differences in peak heights for the isotopes are really due to their abundance.

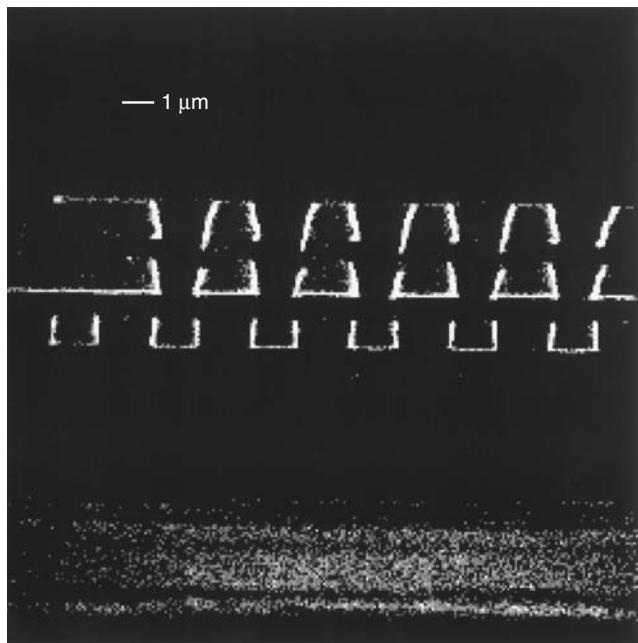
Tables of isotopic abundances are available so that the patterns can be predicted and discerned. An example of a interference would be finding mass peaks between 46 and 50, which are probably due to Ti, and the relative heights can be checked against tabulated data. It would be expected to see this same pattern repeated and displaced by 16, to the region 62–66. The indication of the change in the relative peak heights suggested that there might be the presence of Cu, which has isotopes at 63 and 65. This would be confirmed by the presence of the CuO peaks at 79 and 81.

There are some simple guidelines for the interpretation of mass spectra:

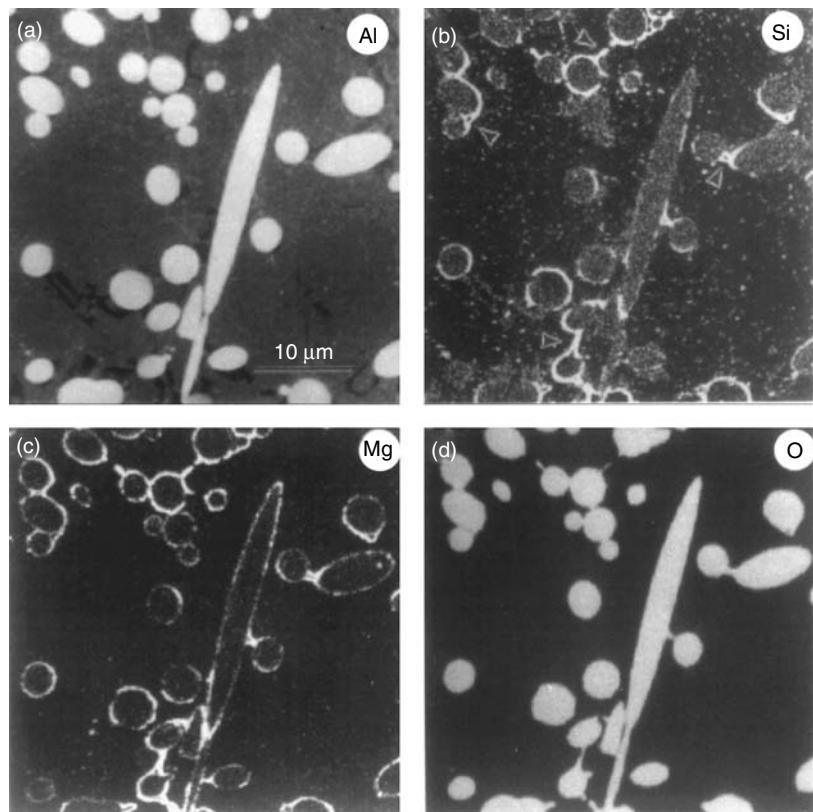
1. In a positive ion spectra, the peaks that occur most strongly are those of low work function.
2. The peaks that occur most strongly in negative ion spectra are those with high electron affinity.
3. Common complexes that occur in positive spectra are MO, MOH, MNa, MK, and hydrocarbon fragments, where M is the mass of some elemental species.
4. Common complexes in negative spectra are MO, MOH, MCl, and hydrocarbon fragments.

Elements maps (Chabala et al., 1994; Crow et al., 1995) either from the top surface of a sample or from the face of cross sections can be made. The general procedure is to find an area of interest, which has the primary beam focused and stigmatized and the spectrometer set to the mass peak of interest, and then acquire a map. Some consideration must be given as to the beam dwell time and primary current. Higher beam currents will produce more signal, but at the cost of reduced spatial resolution and more destruction of the sample. If it is desired to produce several maps of 1000 Å resolution that are within 1000 Å of the same plane in typical semiconductor materials, then for a 40 pA beam a dwell a time of about 1 ms would yield acceptable S/N images. For a magnification of about 10 k× with a 512 × 512 pixel map, the time for acquisition would be 250 s with a beam overlap of about 200 Å, and the depth milled on each map would be about 500 Å. This may consume more samples than desired, and one alternative is to use 256 pixels, which saves a factor of four on sample usage. Merely cutting the beam current with 512 pixels would reduce the sample consumption, but would also cause signal loss. It is found that in mapping cross sections it is sometimes necessary to do a light cleanup of the face after each map to remove the damage done by the beam. Figure 11.47 is a map from a semiconductor.

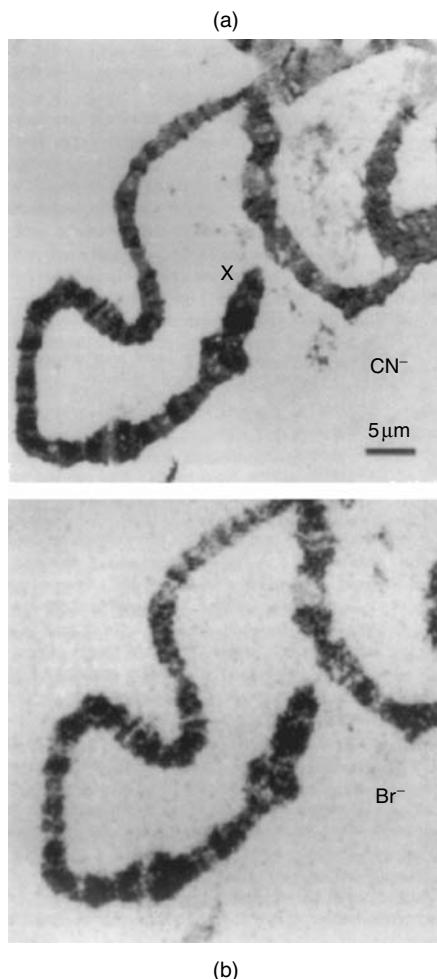
Levi-Setti's group (Chabala et al., 1994; Levi-Setti et al., 1994; Soni et al., 1995) have explored the mapping of a wide variety of biological structures and inorganic structures other than semiconductors, and have mapped with great success materials of interest such as ceramics and photoemulsion crystals. Figures 11.48 and 11.49 show examples of some of the recent work done by the Levi-Setti group. His group has taken the lead in exploring the use of FIB/SIMS in nonsemiconductor-related problems.



**FIGURE 11.47** Semiconductor map.



**FIGURE 11.48** SIMS map of the end region of the X chromosome of the fruit fly *Drosophila melanogaster*. CN<sup>-</sup> map. (Courtesy of R. Levi-Setti, University of Chicago.)

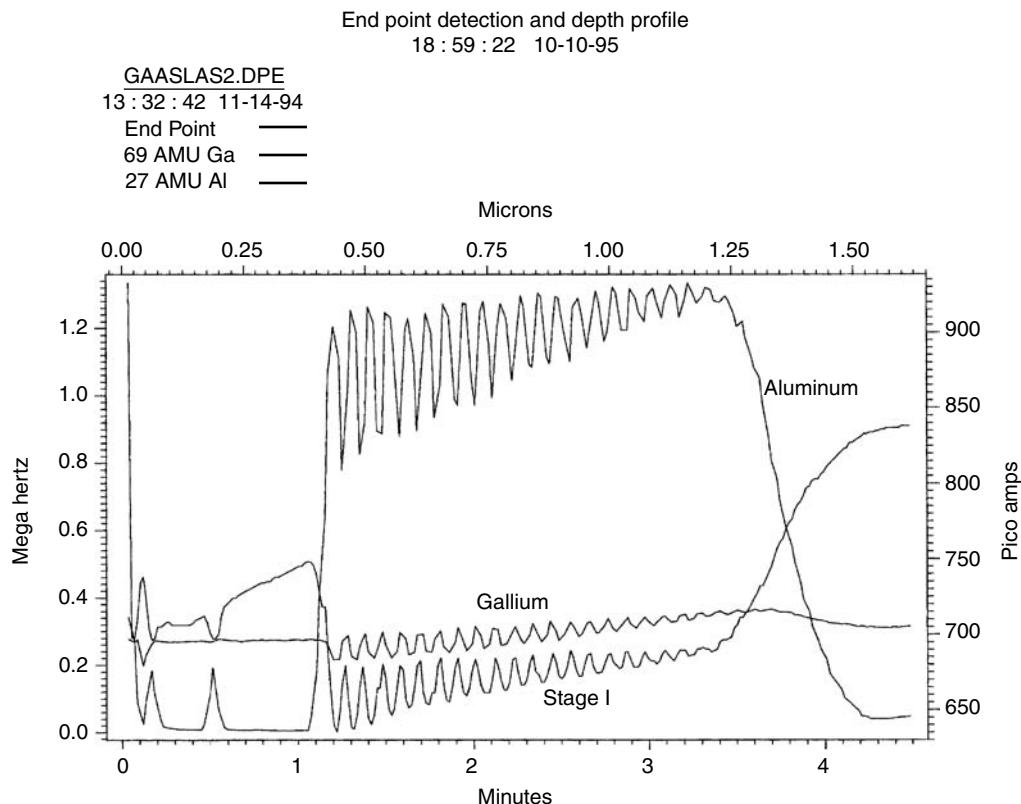


**FIGURE 11.49** SIMS map of the end region of the X chromosome of the fruit fly *D. melanogaster*.<sup>81</sup> Br<sup>-</sup> map. (Courtesy of Levi-Setti, R., University of Chicago.)

Depth profiling is usually accomplished by milling a pattern into the sample while the mass spectrometer monitors predetermined masses. Depth profiles are sometimes difficult to interpret, because there might be redeposition and, at oxide interfaces, increased yields. Larger pattern dimensions reduce the effects of sidewall redeposition and also redeposition on the crater bottom. It is possible that the use of an enhanced etch material may promote a cleaner crater and increase the yields. Figure 11.50 shows a typical depth profile from a modern semiconductor. Note that the interpretation is not so easy due to the matrix effect of oxygen at the interfaces, which enhance the yields.

### 11.9.5 FOCUSED ION BEAM IMPLANTATION

One of the first intended uses of FIB was in the production of semiconductor devices by direct, maskless implantation. Doping of semiconductors happens when ions are implanted into the material with a spatial distribution that has a maximum at some depth below the surface, and the material is then annealed so that the implanted atoms occupy lattice sites. If the implanted atoms have a different valence than the semiconductor lattice atoms, charge carriers are created. In conventional semiconductor device fabrication techniques, this process is accomplished by ion bombardment of the entire wafer, in which areas not to be implanted are covered by a lithographically patterned

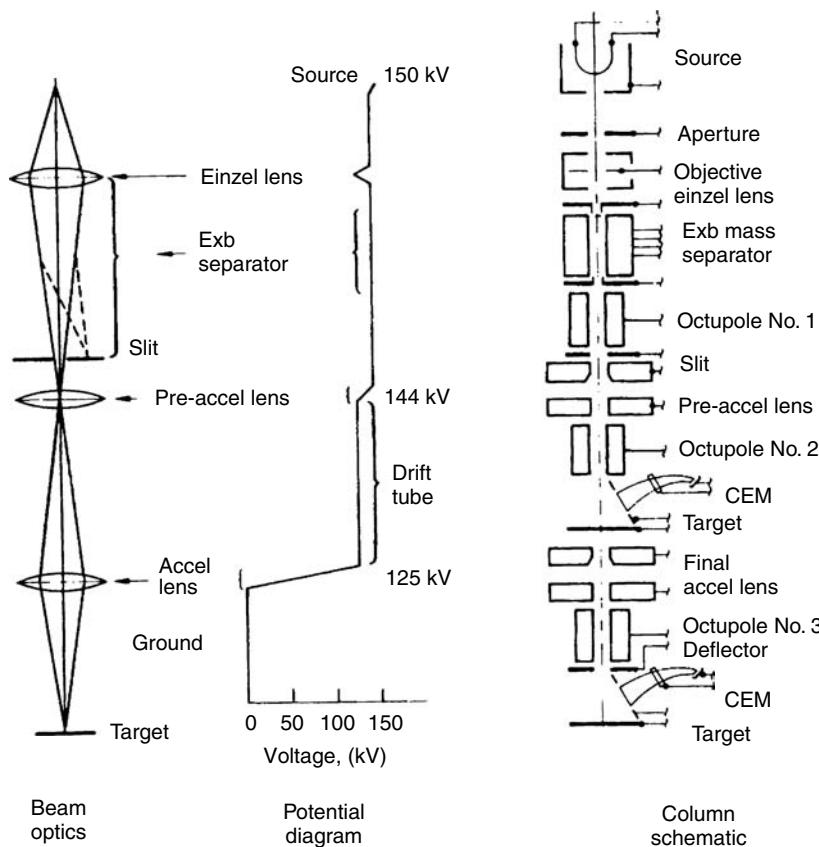


**FIGURE 11.50** Depth profile in IC.

film. FIBs provide a means to perform patterned implantation without the use of a mask or resist by directing the beam only to those areas, which require implantation. New degrees of freedom are allowed in this type of implantation. Laterally graded doping profiles are easily accomplished by varying the beam dwell time in adjacent pixels, and the dose and profile shape (and the depth of implantation also) can easily be varied from device to device on the same wafer for parameterization and test purposes. Structures fabricated this way include bipolar transistors, GaAs MESFETs, and tunable Gunn diodes. The removal of many lithographic steps makes experimentation and production much easier and less costly in both time and money. In a prototyping stage of development, devices that are identical except for implantation can be made side by side, so that wafer-to-wafer and process differences can be eliminated.

A disadvantage to the technique is the low throughput. However, there are applications in which throughput may not be an issue so that FIB may be of great value. In cases where doses are low and cover limited areas, direct implantation may simplify conventional processing, such as in channel implants in MOS transistors or GaAs MESFETs in which the implant area is small and the doses are in the range  $10^{11}$ – $10^{13}$  ions/cm<sup>2</sup>, or in analog-to-digital (A-D) converter threshold adjustments for devices that are critical, but not needed in large numbers.

Because the dopants used for Si are B and As, and for GaAs are Si and Be, sources which are not monatomic species are necessary, leading to ion optics, which require a means of separation of ion species. The most common approach has been used to use the  $\mathbf{E} \times \mathbf{B}$  velocity selector (Wien filter) in the focusing column as a means to achieve mass separation. Figure 11.51 shows the schematic of such a column. In this design, ions are extracted and brought to a focus by an einzel lens (there is no increase in beam energy) in the separation plane of the  $\mathbf{E} \times \mathbf{B}$  in which an aperture is placed to allow the transmission of ions with velocity  $v = E/B$  to pass through the aperture. This makes the velocity



**FIGURE 11.51** Schematic of the three-lens mass-separated variable energy FIB at HRI. (Courtesy of Hughes Research Labs, Malibu, CA.)

selector an effective mass selector, because all the ions from the source have been accelerated through the same potential difference  $V$ , and their velocities are then mass dependent as  $v = (2qV/m)^{1/2}$ .

Once the ions have been mass selected, those ions, which pass through the mass selection aperture (located in the second lens) are *preaccelerated*. This lens is used to vary the beam energy, entering the final lens. The image of the source by the objective einzel lens is placed near the center of the preaccelerator lens so that the beam appears to emanate from this lens with essentially no focusing effect and only a change in energy. Suitable simultaneous adjustment of the final lens voltage and the preaccelerator lens so that the final lens voltage ratio is nearly constant leaves the beam in focus as the final lens beam energy is varied. Also, imaging the objective lens near the center of the preaccelerator reduces the aberration coefficients of that lens, because the beam is always small in the lens. The beam can be blanked at the mass separation aperture, and octupole deflection above the  $\mathbf{E} \times \mathbf{B}$  filter allows column alignment in the upper section, and deflection below the final lens is used for placing the beam onto the sample.

The ion optical design of this column was based on specifications that the beam voltage be variable between 50 and 150 keV, for a source extraction voltage of 6 keV, with a final beam size of 0.1–0.5  $\mu\text{m}$  at a working distance of >5 cm.

Conventional ion implantation uses current densities that are orders of magnitude lower than those in FIB, and it seems reasonable that there might be differences in the properties of the materials implanted with the same dose rate differences. Differences have been observed, but the effect on the final carrier density seems to disappear after they have been annealed. The major differences noted seem to involve damage and amorphization.

The critical dose for forming a continuous amorphous layer is 3–10 times lower for FIB. TEM and electron diffraction also confirm the lower threshold for amorphization for FIB. Si implanted into GaAs by FIB is distributed deeper in the crystal than conventional implants. It has also been found that Si implanted with B to a dose of  $10^{15}$  ions/cm<sup>2</sup> by a slowly scanned beam is more electrically activated before annealing, as indicated by an order of magnitude lower sheet resistance than in the case of a rapidly scanned beam or conventional implantation. The difference disappears for annealing above 800°C. It is theorized that for higher dose rates, the substrate crystal cannot reach equilibrium in the area of an ion impact before the next ion arrives.

The rate of energy loss of ions in Si and GaAs is high enough that it is necessary to use high-voltage FIB. For Si the ranges for B and As are given in Table 11.17. The distribution of the implanted ions is described by the range  $R_p$ , that is, the average depth of penetration, the range straggle  $\Delta R_p$ , that is, the dispersion in depth, and by the transverse straggle  $\Delta R_t$  that is, the dispersion in the transverse delection to the direction of implantation.

The values in Table 11.17 for the straggling values are for the ions as implanted and do not include the effects of annealing. Annealing is necessary to remove some of the lattice damage and to help locating the dopant atoms on lattice sites where they can be electrically active. To see what this effect has on the ion distribution in the semiconductor, before annealing the ions have the spatial distribution  $N(z,r)$  in the semiconductor.

$$N(z,r) = \frac{N}{(2\pi)^{3/2} \Delta R_p \Delta R_t^2} \exp\left(\frac{(z - R_p)^2}{2\Delta R_p^2} - \frac{r^2}{\Delta R_t^2}\right) \quad (11.61)$$

where

$z = 0$  is at the surface

$z > 0$  is into the semiconductor

This assumes a Gaussian distribution for the straggles and that  $N$  ions enter the solid at  $z = r = 0$ . After annealing, the effective straggles grow because of diffusion, and if  $\Delta R'_p$  and  $\Delta R'_t$  are substituted into Equation 11.47, the distribution  $N'(z,r)$  of the ions after implantation and annealing can be estimated. The simplest way to account for the diffusion is to combine each straggle in quadrature with the diffusion length for the anneal time  $t$  as (Melingailis, 1987)

$$\Delta R'_p = \sqrt{(\Delta R_p^2 + 2Dt)} \quad (11.62)$$

$$\Delta R'_t = \sqrt{(\Delta R_t^2 + 2Dt)}$$

---

**TABLE 11.17**  
**Ranges ( $R_p$ ) and Straggles ( $\Delta R_p$ ,  $\Delta R_t$ ) in**  
**Angstroms for B and As into Si**

		Range ( $R_p$ )	$\Delta R_p$	$\Delta R_t$
B	50 keV	1876	534	529
	100 keV	3262	898	1017
	150 keV	4906	875	1228
	200 keV	6284	995	1407
As	50 keV	406	131	99
	1400 keV	696	241	172
	150 keV	1006	345	248
	200 keV	1403	451	301

---

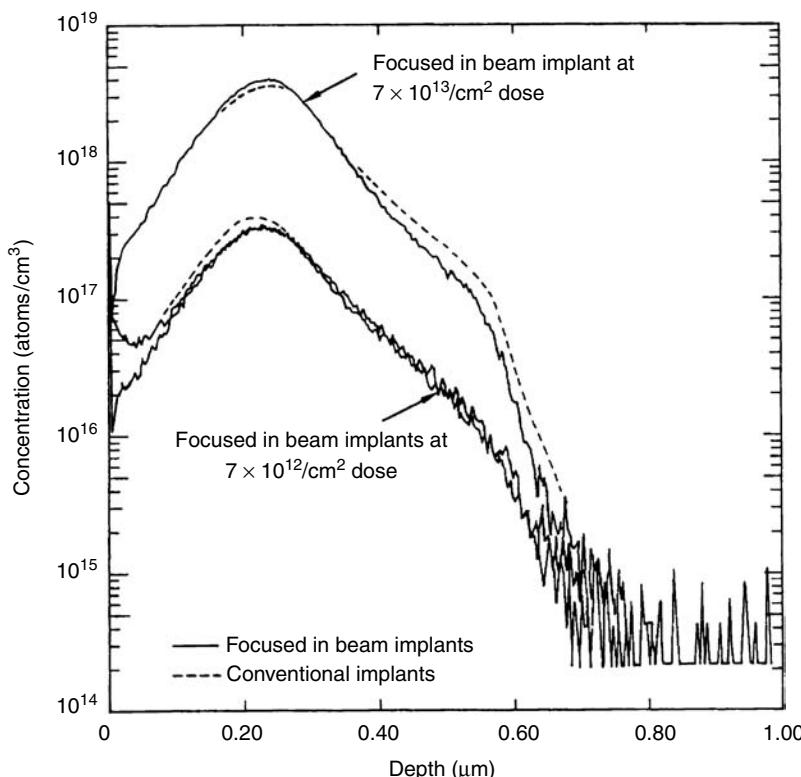
Self-annealing by the beam probably does not occur even though a 100 keV,  $1 \text{ A/cm}^2$  beam has a power density input of  $100 \text{ kW/cm}^2$ . Melngailis (1987) has shown that for a semi-infinite solid, the temperature  $T(\rho)$  at a radius  $\rho > r$ , the radius of the beam, for a beam of voltage  $V$ , current density  $J$ , and a solid with thermal conductivity  $\kappa$

$$T(\rho) - T(\infty) = \frac{V J r^2}{2 \kappa \rho} \quad (11.63)$$

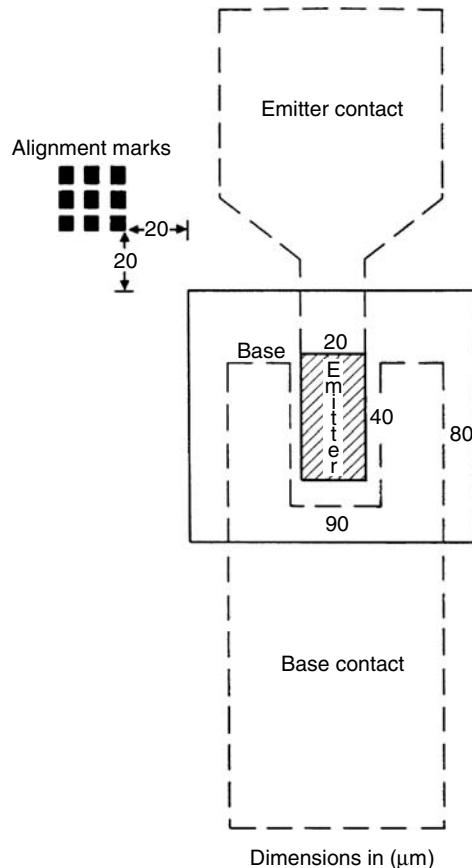
For the case of 100 keV ions with  $J = 1 \text{ A/cm}^2$ ,  $\rho = r = 0.1 \mu\text{m}$ , and  $\kappa = 2 \times 10^{-15} \text{ W/cm } ^\circ\text{C}$  (the value for  $\text{SiO}_2$ ),  $\Delta T = 267^\circ\text{C}$ . This is a near worst-case example, since the thermal conductivity of  $\text{SiO}_2$  is 100 times less than for Si and 25 times less than GaAs.

Reuss et al. (1985) investigated the differences between conventionally and FIB-produced NPN transistors and found essentially no differences. Analysis of the implant profiles by SIMS showed no significant differences, as shown in Figure 11.52, and a further study using Rutherford back scattering (RBS) and Auger microscopy also showed no significant differences.

In an effort to produce Si bipolar transistors by implanting both B and As by FIB, Rules et al. (1985) performed both conventional and FIB implants on the same wafer and found that the results for both techniques were about the same as measured by device characteristics. The bipolar transistor structure is shown in Figure 11.53. The implant procedure was first to find the alignment marks for a transistor, and then to perform the implants in the base and emitter regions. Unlike using the FIB for sputtering away material for a cross section or depositing a metal line, the operator can see no effect of the implantation process to monitor its progress. The beam is entirely computer controlled, with the operator supplying the computer the pattern dimensions and the desired implant



**FIGURE 11.52** SIMS measurement of conventional and FIB-implanted ions. (From Reuss, R.H. et al. *J. Vac. Technol.*, B3, 65, 1985. With permission.)

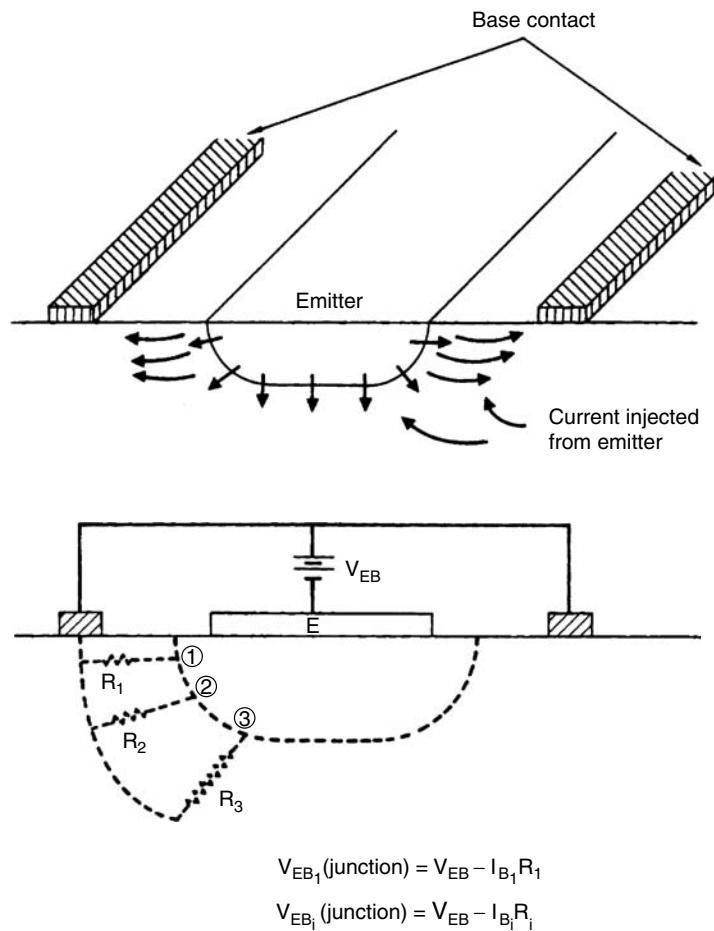


**FIGURE 11.53** FIB bipolar structure. (Courtesy of Hughes Research Labs, Malibu, CA.)

profile. In the mass-separated column used for these experiments, the beam pixel dwell time was 200 ns, and a pixel overlap of 80% was used to ensure uniform implantation. In this system, at the highest spatial resolution of 1500 Å, there was 10 pA of B and 25 pA of As at the target. The maximum source angular current intensities at a source extraction current are of 20 μA/sr for B and 9 μA/sr for As. The base pattern could be implanted in 15 s and the emitter in 60 s.

Several examples of device parameter and scaling studies and the production of novel devices made with FIB implantation include the laterally graded vertical bipolar transistors, FET scaling experiments, focused ion stripe transistor (FIST), the tunable Gunn diode, and threshold adjustments for a four-bit flash A-D converter.

One of the first parametric studies performed on devices with FIB was the lateral grading of the base implants on vertical bipolar transistors (Reuss et al., 1986; Evanson et al., 1988). This type of study is prohibitively difficult by conventional means and demonstrates one of the useful characteristics of FIB implants. Laterally profiled implants were performed to eliminate emitter current crowding by which the Kirk effect causes a decrease in the gain and gain bandwidth of these devices. Conventionally produced uniform base implants made through a mask with a uniform beam cause current to be crowded into the region in which the distance between the base contact and the junction is the smallest. Figure 11.54 is a diagram of the emitter and base active regions of a bipolar device with a two-fingered base contact. With this geometry the injected emitter current is concentrated at the periphery of the emitter because the emitter-base junction forward bias is greater there. For regions farther away from the base contact, the emitter-base forward bias and injected current is less because the resistive voltage drop  $I_{bi}R_i$  is larger because the path is longer. In the regions where there is higher



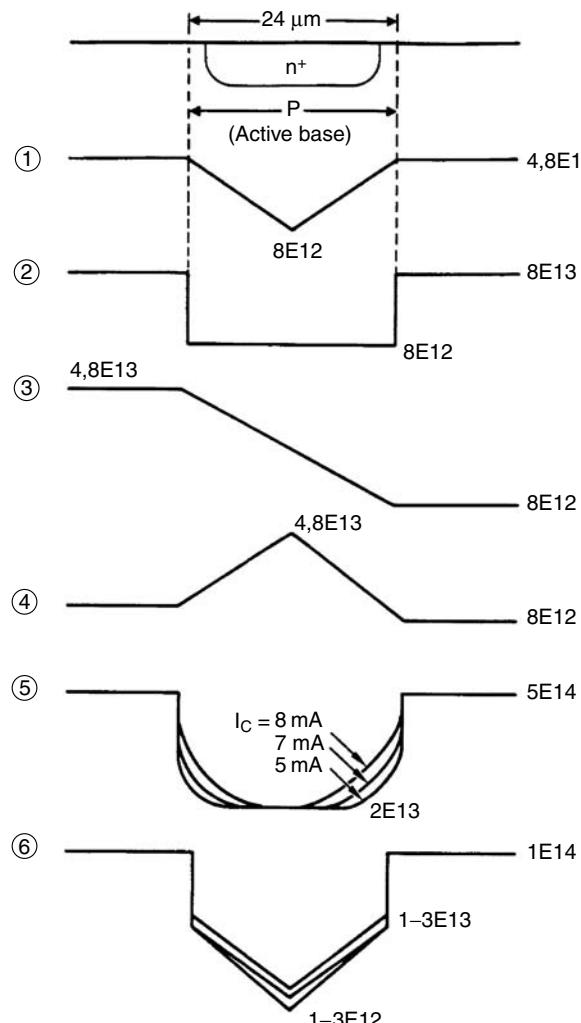
**FIGURE 11.54** Schematic diagram of Kirk effect. (Courtesy of Hughes Research Labs, Malibu, CA.)

injection current it is swept across the quasi-neutral base region to the collector–base junction and produces a charge density in the base that is comparable in magnitude to the base doping, modifying the base–collector background charge so that effective base width and Gummel number are increased. Increasing the Gummel number decreases the transistor current gain, which is called the *Kirk effect*.

It is therefore desirable to decrease the emitter current crowding to increase the transistor current gain. To accomplish this process it is necessary to dope selected regions near the emitter periphery more heavily than those near the center of the emitter. This decreases the resistive voltage drop for regions of the base–emitter junction that are farther away from the base contact and make the bias along the junction more uniform. A variety of laterally profiled base implants were performed to test this concept and are shown schematically in Figure 11.55. The technique used to produce the implants is shown in Figure 11.56.

The flexibility of maskless implantation makes it possible to create a wide variety of these implants on a single wafer along with uniform implants, so that effects can be determined. This type of comparison and profiled implants could not be obtained with conventional masked processes without increasing the number of mask levels to the degree that device yield would drop to nil. In addition, once a desirable implant profile is found, it becomes relatively easy to perform optimizing studies on a single wafer.

In a technique that might be more appropriately discussed in the lithography section, FIB can be used to delineate microstructures in semiconductor thin films by using the FIB to irradiate a

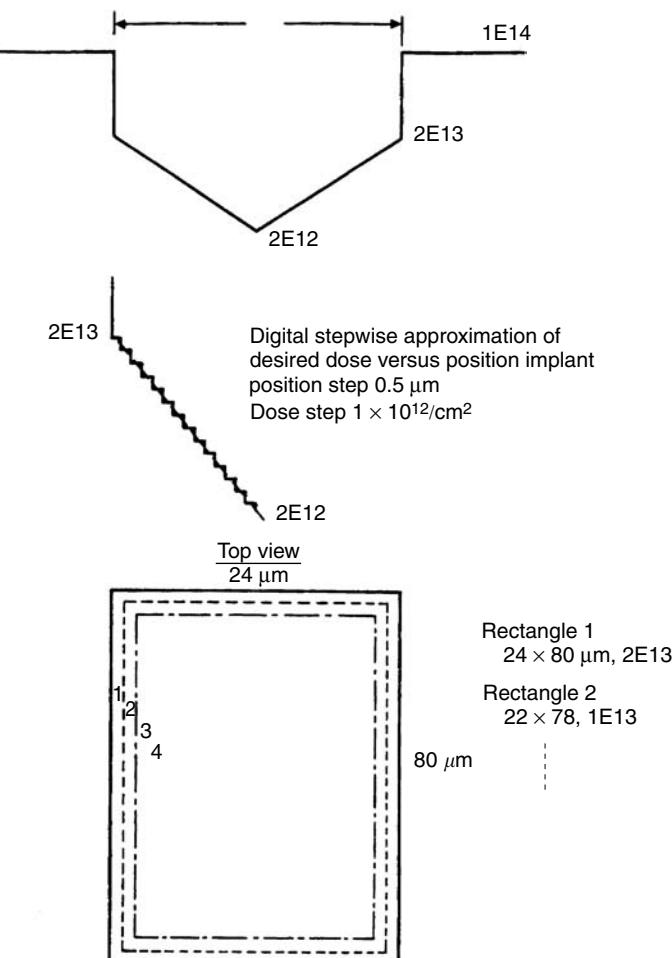


**FIGURE 11.55** Lateral profile shapes. (Courtesy of Hughes Research Labs, Malibu, CA.)

desired region thereby changing its chemical etch properties. The beam-irradiated region is made resistant to an etch step and is left behind when the rest of the thin film is removed chemically. This technique acts as a negative resist with the advantage that one wet fabrication step is eliminated in the processing of devices.

Three different implementations of this technique have been investigated:

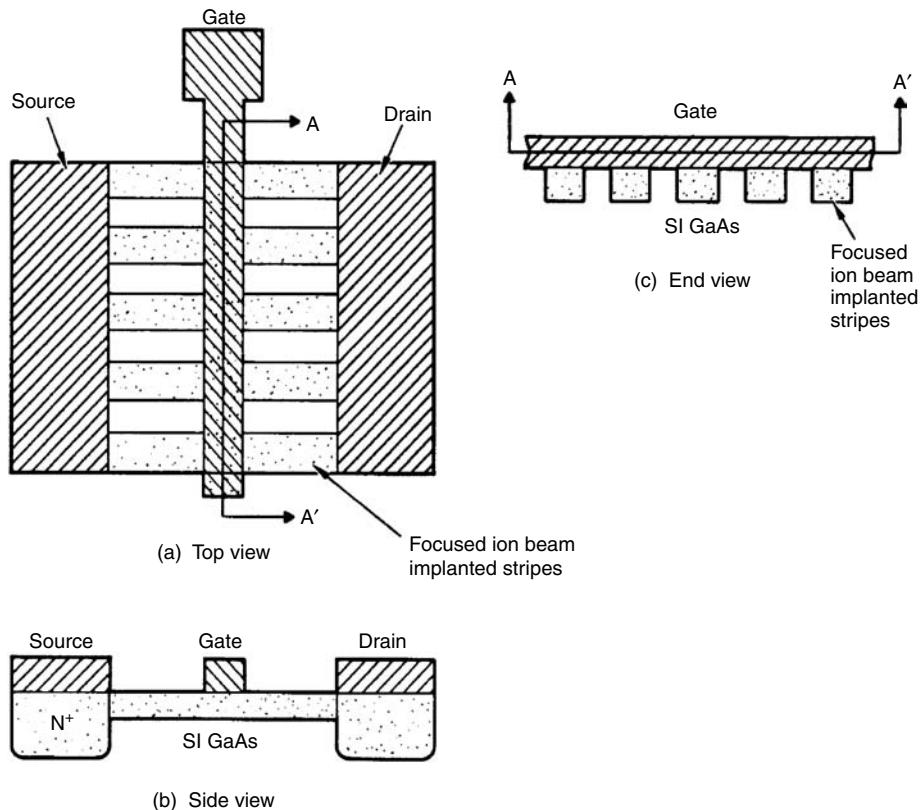
1. Polysilicon gates, where a polysilicon layer is exposed with  $B^+$ , which after etching leaves a boron-doped polysilicon gate.
2. Polysilicon–metal gates, where a polysilicon layer on top of a Ti/W film is exposed with a beam energy sufficient to make the ion range reach the interface (after etching, a gate contact of polysilicon on Ti/W is produced).
3. All-metal gates expose a metal film of chromium or aluminum on Ti/W. After FIB implantation this defines a gate contact, which after etching is transferred to the Ti/W layer by plasma or RIE etching. Because of the possible undercutting with plasma etching, this can produce a T-shaped gate for short-channel MESFET devices.



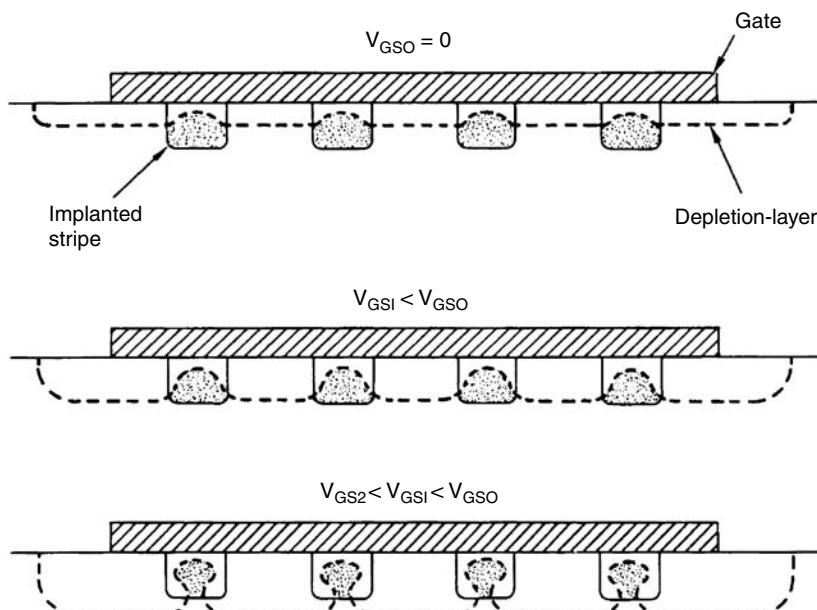
**FIGURE 11.56** Technique used to generate the lateral profile implants. (Courtesy of Hughes Research Labs, Malibu, CA.)

The FIB implanted stripes (focused ion stripe transistor or FIST) device structure (Rensch et al., 1987), which is shown in Figure 11.57, consists of parallel conducting channels formed by FIB implantation embedded in semi-insulating GaAs. The depletion region associated with the Schottky barrier gate extends into the channels from all sides as depicted in Figure 11.58. The modulation of the conducting channel from all sides increases the transconductance and improves the output conductance, both of which lead to a higher-gain device, which could enhance both digital and analog circuit performance in improved noise immunity and a higher-gain bandwidth product.

The principle of operation of the FIST device is that gate depletion layers can wrap around the channel stripes, making the gate region a two-dimensional structure in which the depletion region encroaches far into the undoped regions of the device, wrapping around the active layer as the gate voltage approaches pinch off and the conducting region becomes a conducting tube completely surrounded by the depletion layer. It should also be noted that as the stripe spacing is brought closer together, the pinch-off voltage increases, with the limiting condition that at overlap the device would be a standard MESFET. This feature allows the variation of pinch-off voltage from device to device in an IC without changing implant dose or energy.



**FIGURE 11.57** FIST structure showing parallel FIB implant stripes. (Courtesy of Hughes Research Labs, Malibu, CA.)



**FIGURE 11.58** FIST effect showing depletion region extending into the channel from all sides. (Courtesy of Hughes Research Labs, Malibu, CA.)

These devices were fabricated by implanting 50–70 keV Si<sup>+</sup> from a three-lens, mass-separated column using an Au–Si LMIS source. It is clear that the use of FIB for easy experimentation and optimization could be employed to evaluate theoretical notions and to optimize parameters and process steps on the same wafer.

Tunable Gunn diodes (Lezec et al., 1988) have been made using laterally profiled implantation to produce a doping gradient. This type of Gunn diode consisted of contacts on a GaAs surface with a conducting n-type channel between them. The channel is doped with Si where the dose varies linearly from one contact to the other. When a bias is applied, the electric field at the cathode is the highest due to the lowest doping, and a Gunn domain is launched from the cathode, which travels toward the anode until a threshold electric field is reached, quenching the domain, and launching a new one. The distance traveled (inverse of the oscillation frequency) by the domains can be tuned by adjusting the bias, yielding tunable oscillators in the range of 6–23 GHz.

The ability to adjust easily the threshold of transistors (Lee et al., 1988) has been used to fabricate four-bit 1 GHz flash A–D converters (Walden et al., 1988). For this IC 32 transistors had implants performed in the channel region with increasing doses of boron to make 16 comparators with threshold voltages spaced approximately 94 mV apart. This procedure also allowed the reduction in the total number of parts needed in the converter.

There are several other examples of direct implantation that we mention for completeness:

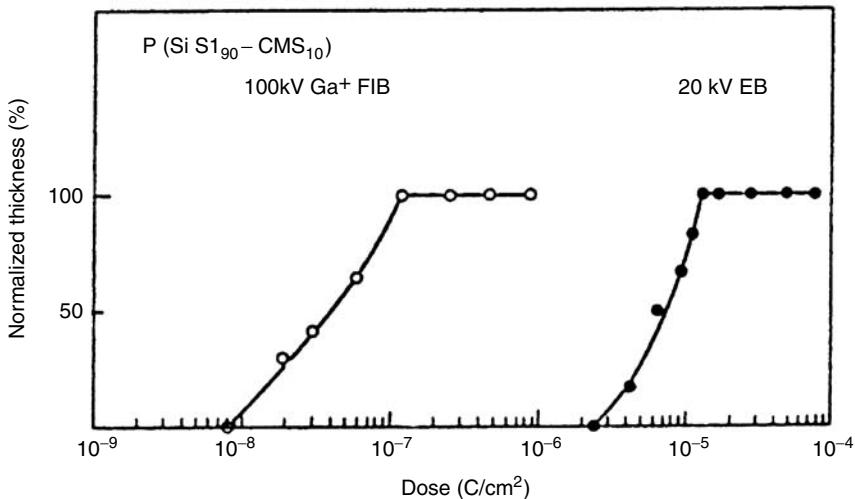
1. GaAs MESFETs with lateral step variations of dose under the gate were created and the position of the step was varied to determine an optimum (Rensch et al., 1987; Evanson et al., 1988).
2. AlGaAs/GaAs quantum wire structures were fabricated to study phase coherence lengths by creating two parallel insulating regions with a narrow (30 nm) conducting region between them by implanting 200 keV Si (Hiramoto et al., 1987).
3. In an optoelectronics application, FIB was used to implant a grating with 100 keV Si for GaAs/GaAlAs distributed Bragg reflector laser (Wu et al., 1988).
4. For multithreshold adjustment for transistors (Lee et al., 1988); channels were implanted in transistors with increasing doses of  $1 \times 10^{11}$  ions/cm<sup>2</sup> increments, the threshold voltage could be varied by 19 mV per increment.
5. Electron conduction channels were formed in a GaAlAs/GaAs heterostructure by unannealed FIB implantation of Be<sup>++</sup> at 260 keV. Electron focusing was observed as a function of magnetic field in the two-dimensional electron gas (Nakamura et al., 1990).

At present, the use of FIB in implantation has entered a relatively dormant state. The systems are more complex to manufacture and use than FIBs of the sort used for failure analysis, but they may still have a niche somewhere in the future production of some semiconductor devices.

### 11.9.6 LITHOGRAPHY

The use of FIB for lithography with resist exposure is the direct analog of electron beam lithography. The beam is scanned in a computer-generated pattern with the defined shapes needed in a layer of beam-sensitive material (“resist”), and is subsequently processed to form structures in the manufacture of semiconductors. This process is used to pattern a mask from which the semiconductor wafer is patterned by optical lithography, or the wafer is directly addressed in a *direct write* lithographic step. Figure 11.59 shows a comparison of the sensitivities of resist to ions and electrons.

Research in lithography using FIB was stimulated by the observations that many resists exhibited greater sensitivity to ions than electrons, sometimes by orders of magnitude, so that exposure rates could be increased, improving throughput and that there was an absence with ions of the



**FIGURE 11.59** Resist sensitivity comparison for electrons and ions for typical conditions in which they are used. (From Matsui, S. et al., *J. Vac. Sci. Technol.*, B4, 845, 1986. With permission.)

electron *proximity effect* due to backscattered electrons, which limited the minimum feature size obtainable with e-beam writing.

Even with increased resist sensitivity and the absence of proximity effects there are other problematic issues. Increased resist sensitivity is due to higher energy loss, resulting in shorter ranges, so resists have to be very thin, on the order of 0.1  $\mu\text{m}$  for even 75 keV Ga. Even if Si is used, only a fivefold increase in penetration is achieved. Si beams require a mass-separated column, so that instrumental complexity increases.

In addition to range problems associated with high resist sensitivity, the issue of statistical noise, which is a fundamental limitation, causes ion produced lithographic structures to become ragged due to statistical fluctuations of dose as a result of small numbers of ions. To impart a dose  $D$  (ions/ $\text{cm}^3$ ) from a beam of diameter  $d$  requires  $n$  ion such that

$$n = \frac{\pi d^2 D}{4} \quad (11.64)$$

The noise associated with  $n$  is  $\sqrt{n}$ , so that the S/N ratio is

$$\frac{S}{N} = \frac{d}{2} \sqrt{\pi D} \quad (11.65)$$

For beam diameters of 0.1  $\mu\text{m}$ , this translates to minimum dose of  $10^{12}$  for minimum well-defined structures. Low-dose statistical noise has been demonstrated by Matsui et al. (1986) and Kubena et al. (1988a,b). In a comparison of all the major lithographic techniques, Smith (1986) has compared them for 0.5  $\mu\text{m}$  linewidth control, which is shown in Table 11.18. It is obvious that the techniques involving parallel methods as opposed to serial scanning methods are superior for the most important quantity, the data transfer rate. It will be difficult to complete with technologies that offer 4–10 orders of magnitude in speed.

**TABLE 11.18**  
**Comparison of Lithographic Techniques**

	Nm	S	D	R (Hz)
Optical projection	1300	0.33 mJ/cm <sup>2</sup>	—	$1.2 \times 10^{14}$ to $1.2 \times 10^{17}$
X-ray (conventional)	24	23 mJ/cm <sup>2</sup>	—	$9.9 \times 10^{-8}$
X-ray (plasma)	24	15 mJ/cm <sup>2</sup>	—	$1.3 \times 10^{-9}$ to $1.1 \times 10^{12}$
X-ray (synchrotron)	24	15 mJ/cm <sup>2</sup>	—	$2.6 \times 10^{12}$
Masked ion beam	29	0.18 $\mu$ C/cm <sup>2</sup>	$1.1 \times 10^{12}$	$2.1 \times 10^{11}$
Scanning e-beam	900	4 $\mu$ C/cm <sup>2</sup>	$2.5 \times 10^{13}$	$2.0 \times 10^7$
Scanning ion beam	29	0.18 $\mu$ C/cm <sup>2</sup>	$1.1 \times 10^{12}$	$2.3 \times 10^7$

*Note:* Nm, minimum number of particles per pixel; S and D, minimum dose; and R, pixel transfer rate.

*Source:* Adapted from Smith, H.I., *J. Vac. Sci. Technol. B*, 4(1), 148–153, 1986. With permission.

## 11.10 NEW DIRECTIONS

There have been two significant developments that might expand FIB technology into new applications. Both are centered on new source technologies, which head in very different directions. The first advance of the technology is due to Ward (Notte and Ward, 2006) who has managed to produce a working stable gas field ionization source (GFIS) for use in a system that produces very high resolution, high contrast images with very little damage to the sample. The new GFIS differs from the field ion microscope (FIM) tip mostly in the shape of the tip (Tandare, 2005). The sharpened tip ( $R \sim 100$  nm) shape is manipulated so that there is an atomically precise pyramidal-shaped bump on the end much like the *super tip* of Kalbitzer (Tandare, 2005). The pyramid edges and apex are atomically sharp so that with this geometry the first few ionization discs (at the tip of the pyramid) begin emitting at a relatively low voltage whereas all the other atoms are not yet capable of emitting. This results in the arriving helium gas being shared by a few atoms instead of a few hundred atoms. By gimbaling the source, emission from a single atom is selected with an aperture, allowing beam to have  $\sim 100\times$  the beam current relative to the low current multiple *beamlets* from the FIM. This end pyramid can be readily removed by increasing the field to  $\sim 5$  V/Å until all pyramid atoms are removed, and then subsequently can be rebuilt, and removed, an unlimited number of times. The emission pattern from this ion source consists of a small number ( $\sim 3$ ) of beams each originating from an atom near the pyramid apex. The beam current can be modulated by changing the pressure of the imaging gas, and can be controlled over several orders of magnitude without any need to change the beam energy, aperture, focus, extraction field, or beam steering. Under typical conditions, the beam current from a single atom is 10 pA, but operation from 1 fA to 100 pA is practical. In addition, the energy spread of the ion source is of the order of 0.5 eV full width at half maximum (FWHM) arising from the finite thickness of the ionization disc ( $\sim 0.3$  Å) in conjunction with the very high electric field  $\sim 3$  V/Å throughout the disc. Estimates indicate that the virtual source size is  $\sim 3$  Å, making the brightness  $\sim 1.5 \times 10^9$  A/cm<sup>2</sup>-sr which is roughly 500× higher than an LMIS. Claims have been made that the ultimate imaging resolution of the instrument built around this source are  $\sim 0.25$  nm, which rivals the resolution achieved by Crewe with the STEM (see Chapter 10, this volume).

In an entirely different direction, Smith (Smith et al., 2006) has developed an inductively coupled plasma (ICP) ion source with high relatively brightness and low energy spread. This source when placed onto a conventional FIB system offers much more beam current than an LMIS can deliver into high current focused beams of the same size, and ions other than metals can be used

(e.g., Kr, Xe, O). The source has an automated gas delivery system and an external RF antenna operating at 13.56 MHz that inductively couples energy into the plasma. One of the major practical advantages of the ICP is that there is negligible cathode erosion due to sputtering.

In an experimental system using 20 keV Xe, sub-100 nm resolution was demonstrated at a beam current of 77 pA. The current density of the ICP system is lower than that for an LMIS based FIB up to about 50 nA, and in the region of 30–40 nA there is a *cross over* regime in which one would want to switch from a LMIS to an ICP-based system. With beam currents of 250 nA delivered into a 1  $\mu\text{m}$  spot size, for applications requiring  $\sim 10^6 \mu\text{m}^3$  of material removal is  $\sim 20$  times faster than using a 20 nA Ga LMIS-based FIB.

## 11.11 THE FUTURE

Prediction can be good unless it involves the future, and the future of FIB is fairly hard to predict. The use of FIB in failure analysis, circuit edit, and in device and IC restructuring and modification will continue into the future in the semiconductor industry. FIB now plays an essential role there. This can be seen from an analysis of the trends of the field by Mackenzie (1991), who cataloged publications. The trend then was clear that Ga is the source most used, and that micromachining was the top use. The trend since then is that micromachining and deposition are by far the main uses of FIB, and although SIMS has been slightly dormant, it is reemerging as a high spatial resolution chemical mapping tool, and lithography and implantation have seen diminished use.

Certainly there will be more work on GFIS sources as the first series of systems based on that idea have entered use. Whether or not this technology will have widespread value as a microscopy tool remains to be seen. One can imagine that conventional multispecies LMIS might become competitive with GFIS. Imaging resolution improvement with Ga LMIS is nearing fundamental physical limits, so that to achieve higher resolution, lower mass ions must be used to push this technology to its limit.

The newly developed low  $\Delta E$  plasma sources will certainly extend applications of FIB into higher volume removal realms. The high rate of bulk removal that they offer coupled with conventional FIB and SEM should prove useful in MEMS and nanotechnology arenas. In addition, as femtosecond lasers become more known, they offer some competition to FIB use because of their million times faster rate of material removal as compared to FIB without any damage due to heating. These lasers also seem to be able to operate at high-resolution levels near 50 nm, and can perform spectroscopy that cannot be achieved any other way. In addition, they offer two-photon polymerization deposition rates that are outstandingly fast.

New beam strategies will be found that enable gas selective and enhanced etching and deposition of materials to become more refined. Whether or not these capabilities and increased resolution will find their way into other fields like the biological sciences and materials science will be determined by whether useful structures can be fabricated and whether useful information from samples will be obtainable.

FIB implantation is still a back-burner topic, which a few solitary workers are exploring. It appears that FIB lithography at this time is a dead issue for semiconductors, but that there are a few researchers who still explore its use in forming MEMS-like and nano-scale structures.

Systems will continue to become more automated and reliable. They will find their way into most large industries, typically as multifunctional dual beam systems. One can imagine eventually that MEMS and nanotechnology will be processed or at least developed in *factories-in-a-can*.

The economic future of FIB looks very promising because of the continued fantastically high demand for semiconductors which requires the building of new FAB lines producing devices with ever shrinking geometries every year to meet consumer demand. The Information Age has arrived, and is fueled by semiconductors: the most numerous human-made structures are the transistor, and will likely be so in the foreseeable future. If suitable applications in areas other than semiconductors can be found, then new markets might sustain the research necessary for interesting development.

## FOR FURTHER INFORMATION

Seven excellent reviews of FIB are by Melngailis (1987, 1991), Harriott (1989, 1991), Namba (1989), Nikawa (1991), and Orloff (1993). There are four excellent books about this field all of which taken together show the historical development and use of FIB. Prewett and Mair (1991) were followed by Orloff et al. (2003), Giannuzzi and Stevie (2004), and Yao (2007). These reviews and books were written by the wizards of the field and contain more in-depth information and are abundant in references for the material covered here.

Many papers relevant to FIB are found in the proceedings of the Electron Ion Photon Beam and Nanotechnology Meeting held annually, which are usually published in the November/December or January/February issue of the *Journal of Vacuum Science Technology*. That journal also routinely has papers in the FIB field. There are often good sources to be found in the *Journal of Applied Physics* and sporadically in SPIE publications of conferences. In addition, there are several new journals dedicated to nanotechnology, which often feature FIB work. Examples are *Journal of Micromechanics and Microengineering* and *Small*.

Prewett and Mair (1991) as well as Orloff et al. (2003) are excellent references for LMIS. This handbook contains an excellent chapter by the late Mair (updated extensively by Richard Forbes) on LMIS (Chapter 2). Giannuzzi and Stevie (2005) and Yao (2007) contain excellent treatments of the theory and applications of FIB. For a thorough treatment of ion–solid interactions, Benninghoven et al. (1987) is an excellent source. That book is considered by many to be the bible of SIMS and covers many pertinent topics.

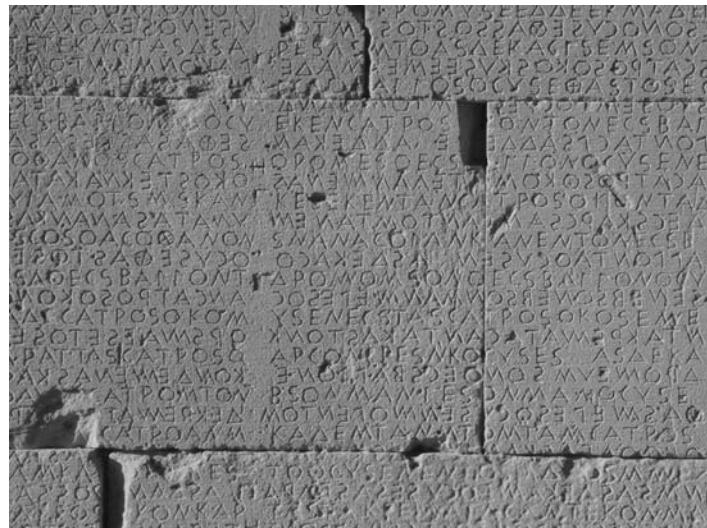
At present, there are five companies in the world that are making FIB systems. They are FEI Co. in Hillsboro, OR, USA; Hitachi Corp., JEOL Ltd., and Seiko in Tokyo, Japan; and Carl Zeiss in Oberkochen, Germany.

## ACKNOWLEDGMENTS

I have had the pleasure and good fortune to work with, know, and enjoy the friendship of a large number of the people whose work is represented in this chapter. I got to see firsthand (1) A. Crewe's physical intuition; (2) Levi-Setti's conversion to LMIS and SIMS; (3) founding of IBT and Micrion (the first and second U.S. FIB company, respectively) and the phenomenal growth of FEI, (4) exploration of FIB at HRL., OGC, and MIT; and (5) evolution and use of systems at FEI and Intel and beyond. Among those who have shared their ideas and showed me their secrets are I. Berry, R. Boylan, P. Carleson, C. Chandler, M. D. Courtney, A. V. Crewe, G. Crow, W. Clark, J. Doherty, R. Gerlach, L. Harriott, J. Jergenson, R. Levi-Setti, J. Melngailis, J. Orloff, N. W. Parker, R. Reuss, B. Samoyed, S. Samoyed, R. Seliger, N. Smith, M. Straw, L. Swanson, W. Thompson, P. Tesch, D. Tuggle, B. Ward, J. Ward, R. Young, and a government employee named Joe. I have also read a lot of other peoples' work and heard their speech. I would also like to thank my mother who let me read while eating breakfast.

## A.1 APPENDIX: WHAT IS BOUSTROPHEDONIC?

The term “boustrophedonically” is the adverbial form of boustrophedon (Greek: *βουστροφηδόν*—*turning like oxen in ploughing*), which is an ancient way of writing manuscripts and other inscriptions in which, rather than going from left to right as in modern English, or right to left as in Hebrew and Arabic, alternate lines must be read in opposite directions. The name is borrowed from the Greek language. Its etymology is from *βοῦς*, ox + *στρεφεῖν*, to turn because the hand of the writer goes back and forth like an ox drawing a plow across a field and turning at the end of each row to return in the opposite direction. This is different than *serpentine* which means: of, characteristic of, or resembling a serpent, as in form or movement. Only a snake on a treadmill comes close to exhibiting a boustrophedonic characteristics. Figure 11.A.1 shows an example of such ancient lithography. The photo was taken by the editor of this book.



**FIGURE 11.A.1** (See color insert following page 340.) This site was pointed out to the author by the editor of this book, and is on the island of Crete at Gortys, dating back ~2600 years–600 B.C. This shows one section of a wall containing the Laws of Gortyna, some 600 lines of law code that are the earliest recorded in the Greek world that are boustrophedonically written.

## REFERENCES

1. Atwood, D. K., Fisanick, G. J., Johnson, W. A., and Wagner, A. 1984. Defect repair techniques for X-ray masks, *SPIE*, 471, 127–137.
2. Basile, D., Boylan, R., Baker, B., Hayes, K., and Soza, D. 1992. FIBXTEM—focused ion beam milling for TEM sample preparation, *Mater. Res. Soc. Symp. Proc.*, 254, 23–41.
3. Benninghoven, A., Rudenauer, F. G., and Werner, H. W. 1987. *Secondary Ion Mass Spectrometry*, Wiley, New York, 218–224.
4. Blauner, P., Butt, Y., Ro, J., Thompson, C., and Melngailis, J. 1989. Focused ion beam induced deposition of low resistivity gold films, *J. Vac. Sci. Technol.*, B7, 1816–1818.
5. Boylan, R., Ward, M., and Tuggle, D. 1989. Failure analysis of micron technology VLSI using focused ion beams, in *Proc. Int. Sym. for Testing and Failure Analysis*, ASM International, Los Angeles, pp. 249–253.
6. Castaing, R. and Slodzian, G. 1962. Microanalyser par e' emission ionique secondaire, *J. Microsc.*, 1, 395–410.
7. Chabala, J. M., Levi-Setti, R., Li, L., Parker, N. W., and Utlaut, M. 1992. Development of a magnetic sector-based high lateral resolution scanning ion probe, in *Secondary Ion Mass Spectrometry, SIMS VIII*, A. Benninghoven, K. T. F. Janssen, J. Tumpner, and H. W. Werner, Eds, Wiley, Chichester, p. 179.
8. Chabala, J. M., Soni, K. K., Li, J., Gavrilov, K. L., and Levi-Setti, R. 1994. High-resolution chemical imaging with scanning ion probe, *SIMS. Int. J. Mass Spectr. Ion Proc.*, 143, 191–212.
9. Clampitt, R., Aitkin, K. L., and Jeffries, D. K. 1975. Intense field emission ion source of liquid metals, *J. Vac. Sci. Technol.*, 12 (Abstract), 1208.
10. Clark, W. M., Jr., Seliger, R. L., Utlaut, M., Bell, A. E., Swanson, L. W., Schwind, G. A., and Jergenson, J. B. 1987. Long-lifetime, reliable liquid metal ion sources for boron, arsenic, and phosphorus, *J. Vac. Sci. Technol.*, B5, 197–202.
11. Clark, W. M., Jr., Utlaut, M., Reuss, R. H., and Koury, D. 1988. High-gain lateral pnp bipolar transistors made using focused ion beam implantation, *J. Vac. Sci. Technol.*, B6(3), 1006–1009.
12. Cleaver, J. R. A. and Ahmed, H. 1981. A 100-kV ion probe microfabrication system with a tetrode gun, *J. Vac. Sci. Technol.*, 19(4), 1145–1148.
13. Crewe, A. V. 1973. Production of electron probes using a field emission source, in *Progress in Optics*, Vol. XI, E. Wolf, Ed., North-Holland, Amsterdam, Chap. 5.
14. Crewe, A. V. 1987. Optimization of small electron probes, *Ultramicroscopy*, 23, 159–168.

15. Crewe, A. V. and Salzman, D. 1982. On the optimum resolution for a corrected STEM, *Ultramicroscopy*, 9, 373–378.
16. Crewe, A. V., Eggenberger, D. N., Wall, J., and Welter, L. M. 1968. A high resolution scanning transmission electron microscope, *Rev. Sci. Instrum.*, 39, 576.
17. Crow, G. A. 1990. Thesis. Focused ion beam micromachining of Si and GaAs using Ga and Au liquid metal ion sources. Oregon Institute of Science and Technology; Beaverton, OR.
18. Crow, G. A. 1992. Endpoint detection and microanalysis with Ga FIB SIMS, in *Proc. ISTFA'91: The 17th Int. Symp. Testing and Failure Anal.*, ASM International, Materials Park, OH, pp. 401–407.
19. Crow, G. A., Christman, L., and Utlaut, M. 1995. A focused ion beam SIMS system, *J. Vac. Sci. Technol.*, B13(6), 2607–2612.
20. Dahl, D. A., Delaware, J. E., and Appelhans, A. D. 1990. SIMION PC/PS2 electrostatic lens program, *Rev. Sci. Instrum.*, 61, 607–609.
21. Della Ratta, A. D., Melngailis, J., and Thompson, C. V. 1993. Focused ion beam induced deposition of copper, *J. Vac. Sci. Technol.*, B11, 2195–2199.
22. Doherty, J. A. 1985. Mask repair with focused ion beams, *Proc. 9th Symp. on Ion Sources and Ion Assisted Technology (ISIAT)*, T. Takagi, Ed., Inst. of Electrical Engineers of Japan, Tokyo, Japan, pp. 65–69.
23. Dubner, A. D., Wagner, A. 1989. Mechanism of ion beam induced deposition of gold, *J. Vac. Sci. Technol.*, B7(6), 1950–1953.
24. Evanson, A. F., Cleaver, J. R. A., and Ahmed, H. 1988. Focused ion beam implantation of gallium arsenide metal-semiconductor field effect transistors with laterally graded doping profiles, *J. Vac. Sci. Technol.*, B6(6), 1832–1835.
25. FEI Technical Note. 2003. Carbon Deposition Technical Note. PN 4035 272 27241-A.
26. Gamo, K., Takakura, N., Hamaura, Y., Tomita, M., and Namba, S. 1986. *Microelectron. Eng.*, 5, 163–270.
27. Gamo, K., Takakura, N., Samoto, N., Shimizu, R., and Namba, S. 1984. Ion beam assisted deposition of metal organic films using focused ion beams, *Jpn J. Appl. Phys.*, 23, L293–295.
28. Gamo, K., Ukegawa, T., and Namba, S. 1980. Field ion sources using eutectic alloys, *Jpn J. Appl. Phys.*, 7, L379–382.
29. Gandhi, A. and Orloff, J. 1990. Parametric modeling of focused ion beam-induced etching, *J. Vac. Sci. Technol.*, B8, 1814–1819.
30. Giannuzzi, L. A. and Stevie, F. A., Eds. 2005. *Introduction to Focused Ion Beams: Instrumentation, Theory, Techniques and Practice*, Springer, Berlin.
31. Gross, M., Harriott, L., and Opila, R. 1990. Focused ion beam stimulated deposition of aluminum from trialkylamine alanes, *J. Appl. Phys.*, 68, 4820–4824.
32. Harriott, L. 1989. Microfocused ion beam applications in microelectronics, *Appl. Surf. Sci.*, 36, 432–442.
33. Harriott, L. 1991. Technology of finely focused ion beams, *Nucl. Instrum. Methods Phys. Res.*, B55, 802–814.
34. Harriott, L., Cummings, K. D., Gross, M. E., and Brown, W. E. 1986. Decomposition of palladium acetate films with a microfocused ion beam, *Appl. Phys. Lett.*, 49, 1661–1662.
35. Harriott, L., Cummings, K. D., Gross, M. E., Brown, W. E., Linnros, J., and Funsten, H. O. 1987. Fine line patterning by focused ion beam-induced decomposition of palladium acetate, *Mater. Res. Soc. Symp. Proc.*, 75, 99–105.
36. Harriott, L., Temkin, H., Wang, Y., Hamm, R., and Weiner, J. 1990. Vacuum lithography for 3-dimensional fabrication using finely focused ion beams, *J. Vac. Sci. Technol.*, B8(6), 1380–1384.
37. Hayashi, S. and Hashiguchi, Y. 1993. Photoirradiation-charge compensation for secondary ion mass spectrometers analysis of semiconductors, *J. Vac. Sci. Technol.*, A11(5), 2610–2613.
38. Hill, A. R. 1968. *Nature*, 218, 262.
39. Hiramoto, T., Hirakawa, K., Iye, Y., and Ikoma, T. 1987. One-dimensional GaAs wires fabricated by focused ion beam implantation, *Appl. Phys. Lett.*, 51, 1620–1622.
40. Hoepfner, P. J. 1985. A Monte Carlo calculation of virtual source size and energy spread for a liquid metal ion source, Thesis, Oregon Graduate Institute, Beaverton, OR.
41. Ishitani, T., Kawanami, Y., and Todokoro, H. 1985. Aluminum-line cutting end-monitor using scanning ion microscope voltage contrast, *Jpn J. Appl. Phys.*, 24, L133–134.
42. Jergenson, J. B. 1982. Liquid Metal Ion Source, U.S. Patent #4318029, March 2, 1982.
43. Komano, H., Ogawa, Y., and Takigawa, T. 1989. Film deposition with focused ion beams, *Int. Symp. Microprocess Conf. Proc.*, 303–306.

44. Krohn, V. E. and Ringo, G. R. 1975. Ion source of high-brightness using liquid metal, *Appl. Phys. Lett.*, 27, 479–481.
45. Kubena, R. L., Joyce, R. J., Ward, J. W., Garvin, H. L., Stratton, F. P., and Brault, R. G. 1988a. Dot lithography for zero-dimensional quantum well using focused ion beams, *J. Vac. Sci. Technol.*, B6, 353–356.
46. Kubena R. L., Stratton, F. P., and Mayer, T. M. 1988b. Selective area nucleation for metal chemical vapor deposition using focused ion beams, *J. Vac. Sci. Technol.*, B6, 1865–1868.
47. Lee, J. Y., Clark, W. M., Jr., and Utlaut, M. 1988. Multiple-threshold-voltage CMOS/SOS by focused ion beams, *Solid State Electron.*, 31(2), 155–158.
48. Leslie, A. 1994. Proc. Characterization and applications of FIB/SIMS for microelectronic materials and devices, *5th European Symposium on Reliability of Electron Devices, Failure Physics and Analysis*, p. 401.
49. Levi-Setti, R. L. 1974. Proton Scanning microscopy: feasibility and promise, *Scanning Electron Microsc.*, 1, 125.
50. Levi-Setti, R. 1983. Secondary electron and ion imaging in scanning ion microscopy, *Scanning Electron Microsc.*, 1–12.
51. Levi-Setti, R., Chabala, J., and Smolik, S. 1994. Nucleotide and protein distribution in BradU-labelled polytene chromosomes revealed by ion probe mass spectrometry, *J. Microsc.*, 175(1), 44–53.
52. Levi-Setti, R., Wang, Y. L., and Crow, G. 1984. High spatial resolution SIMS with the UC-HRL scanning ion microprobe, *J. Phys. C*, 9(45), 197–205.
53. Lezec, H. J., Ismail, K., Mahoney, L. J., Shepard, M. I., Antoniadis, D. A., and Melngailis, J. 1988. Tunable frequency Gunn diodes fabricated by focused ion beam implantation, *IEEE Electron Dev. Lett.*, 9, 476–478.
54. Liebl, H. 1967. Ion microprobe mass analyzer, *J. Appl. Phys.*, 38, 5277–5283.
55. Liebl, H. 1972. A coaxial combined electrostatic objective and anode lens for microprobe mass analyzers, *Vacuum*, 22, 619–621.
56. Mackenzie, R. 1991. Developments and trends in the technology of focused ion beams, *J. Vac. Sci. Technol.*, B9(5), 2561–2565.
57. Matsui, S., Mori, K., Saigo, K., Shiokawa, T., Toyoda, K., and Namba, S. 1986. *J. Vac. Sci. Technol.*, B4, 845.
58. Melngailis, J. 1987. Focused ion beam technology, *J. Vac. Sci. Technol.*, B5, 469.
59. Melngailis, J. 1991. Focused ion beam lithography and implantation, in *Handbook of VLSI Microlithography*, W. Glendinning and J. Helbert, Eds, Noyes Publications, Park Ridge, NJ.
60. Melngailis, J., Musil, C., Stevens, E. H., Utlaut, M., Kellogg, E., Post, R., Geis, M., and Mountain, R. 1986. The focused ion beams as an integrated circuit restructuring tool, *J. Vac. Sci. Technol.*, B4(1), 176–180.
61. Nakamura, K., Tsui, D. C., Nihey, F., Toyoshima, H., and Itoh, T. 1990. Electron focusing with multiparallel one dimensional channels made by focused ion beams, *Appl. Phys. Lett.*, 56, 385–387.
62. Namba, S. 1989. Focused ion beam processing, *Nucl. Inst. Methods Phys. Res.*, B39, 504–510.
63. Neyman, J. and Scott, E. L. 1958. Statistical approach to problems of cosmology, *J. R. Stat. Soc.*, B20, 1–29.
64. Nikawa, K. 1991. Applications of focused ion beam technique to failure analysis of very large scale integrations: a review, *J. Vac. Sci. Techno.*, B9(5), 2566–2577.
65. Notte J. and Ward B. 2006. Sample interaction and contrast mechanisms of the helium ion microscope. *Scanning*, 28, 13.
66. Ochiai, Y., Gamo, K., and Namba, S. 1985. Pressure and irradiation angle dependence of maskless ion beam assisted etching of GaAs and Si, *J. Vac. Sci. Technol.*, B3(1), 67–70.
67. Orloff, J. 1993. High-resolution focused ion beams, *Rev. Sci. Instrum.*, 64(5), 1105–1130.
68. Orloff, J. 1995. Limits on imaging resolution of focused ion beam systems, *SPIE Proc. Charged Particle Optics Conf.*, 3125, 92–95.
69. Orloff, J. and Swanson, L. W. 1975. Study of a field ionization source for microprobe applications, *J. Vac. Sci. Technol.*, 12, 1209–1213.
70. Orloff, J. and Swanson L. W. 1984. Charged Particle Beam Apparatus and Method Utilizing Liquid Metal Field Ionization Source and Asymmetric Three Element Lens System, U.S. Patent # 4426 582.
71. Orloff, J., Utlaut, M., Swanson, L. W. 2003. *High Resolution Focused Ion Beams*, Plenum, New York.
72. Prewett, P. D. and Mair, G. L. R. 1991. *Focused Ion Beams from Liquid Metal Ion Sources*, Wiley, New York.

73. Puretz, J. and Swanson, L. W. 1992. Focused ion beam deposition of Pt containing films, *J. Vac. Sci. Technol.*, B10(6), 2695–2698.
74. Rensch, D. B., Matthews, D. S., Utlaut, M., Courtesy, M. D., and Clark, W. M., Jr. 1987. Performance of the focused-ion-striped-transistor (FIST)—a new MESFET structure produced by focused-ion-beam implantation, *IEEE Trans. Electron Dev.*, ED-34, 2232–2237.
75. Reuss, R. H. 1985. Potential applications of focused ion beam technology for the semiconductor industry, *Nucl. Instrum. Methods*, B10/11, 515–521.
76. Reuss, R. H., Morgan, D., Goldenetz, A., Clark W. M., Jr., Rensch, D. B., and Utlaut, M. 1986. Fabrication of bipolar transistors by maskless implantation, *J. Vac. Sci. Technol.*, B4(1), 290–294.
77. Rudenauer, F. G., Steiger, W., and Kraus, U. 1979. Microanalysis with a quadrupole ion microprobe, *Mikrochim. Acta (Wein)*, (Suppl. 8), 51–58.
78. Sato, M. and Orloff, J. 1991. A method for calculating the current density of charged particle beams and the effect of finite source size and spherical and chromatic aberrations on the focusing characteristics, *J. Vac. Sci. Technol.*, B9, 2602–2608.
79. Seliger, R. L. 1972. E  $\times$  B Mass-separator design, *J. Appl. Phys.*, 43, 2352–2357.
80. Seliger, R. L., Ward, J. W., Wang, V., and Kubena, R. 1979. A high-intensity scanning ion probe with submicrometer spot size, *Appl. Phys. Lett.*, 34, 310.
81. Shedd, G. M., Dubner, A. D., Thompson, C. V., and Melngailis, J. 1986. Focused ion beam induced deposition of gold, *J. Appl. Phys. Lett.*, 49, 1584–1586.
82. Slodzian, G., Daigne, B., Girard, F., and Hillion, F. 1992. A high resolution scanning ion microscope with parallel detection of secondary ions, *Secondary Ion Mass Spectrometry SIMS VII*, A. Benninghoven, K. T. F. Janssen, J. Tumpner, and H. W. Werner, Eds, Wiley, Chichester, p. 169.
83. Smith, H. I. 1986. A statistical analysis of ultraviolet, x-ray, and charged-particle lithographies, *J. Vac. Sci. Technol.*, B4(1), 148–153.
84. Smith, N. S., Skocylas, W. P., Kellogg, S. M., Kinion, D. E., Tesch, P. P., Sutherland, O., Aanesland, A., and Boswell, R. W. 2006. High brightness inductively coupled plasma source for high current focused ion beam application, *J. Vac. Sci. Technol.*, B24(6), 2902–2906.
85. Soni, K., Tseng, M., Williams, D., Chabala, J., and Levi-Setti, R. 1995. An ion microprobe study of the microchemistry of Ni-bas superalloy-Al<sub>2</sub>O<sub>3</sub> metal-matrix composites, *J. Microsc.*, 178(2), 134–145.
86. Stark, T. J., Shedd, G. M., Vitarelli, J., Griffis, D. P., and Russell, P. E. 1995. H<sub>2</sub>O enhanced focused ion beam micromachining, *J. Vac. Sci. Technol.*, B13, 2565–2570.
87. Stewart, D., Stern, L., and Morgan, J. 1989. Electron beam x-ray and ion beam technologies: submicrometer lithographies VIII, *Proc. SPIE*, 1089, 18–25.
88. Sturrock, P. A. 1955. *Static and Dynamic Electron Optics*, Cambridge University Press, London, pp. 47–104.
89. Sudraud, P., Assayag, G., and Bon, M. 1988. Focused ion beam milling, scanning electron microscopy, and focused droplet deposition in a single microsurgery tool, *J. Vac. Sci. Technol. B*, 6(1), 234–238.
90. Swanson, L. W. 1994. Use of the liquid metal ion source for focused ion beam applications, *Appl. Surf. Sci.*, 76/77, 80–88.
91. Takado, N., Asakawa, K., Arimoto, H., Morita, T., Sugata, S., Miyauchi, E., and Hashimoto, H. 1987. Chemically enhanced GaAs maskless etching using a novel focused ion beam etching system with a fluorine molecular and radical beam, *Mater. Res. Symp. Proc.*, 75, 107–114.
92. Tandare, V. N. 2005. Quest for high brightness, monochromatic noble gas ion sources, *J. Vac. Sci. Technol.*, A23, 1498–1507.
93. Tao, T., Ro, J., and Melngailis, J. 1990. Focused ion beam deposition of platinum, *J. Vac. Sci. Technol.*, B8, 1826–1829.
94. Walden, R. H., Schmitz, A. E., Larson, L. E., Kramer, A. R., and Pasiecznik, J. 1988. A 4-bit, 1 Ghz sub-half micrometer CMOS/SOS flash analog-to-digital converter using focused ion beam implantation, *Proc. of the IEEE 1988 Custom Integrated Circuits Conf.*, pp. 18.7.1–18.7.4.
95. Wang, V., Ward, J. W., and Seliger, R. L. 1981. A mass separating focused ion beam system for maskless implantation, *J. Vac. Sci. Technol.*, 19, 1158–1163.
96. Ward, B., Shaver, D. C., and Ward, M. L. 1985. Repair of photomasks with focused ion beams, *SPIE*, 537, 110–114.
97. Ward, J. W. 1985. A Monte Carlo calculation of the virtual source size for a liquid metal ion source, *J. Vac. Sci. Technol.*, B3, 207.
98. Ward, J. W. and Seliger, R. L. 1981. Trajectory calculations of the extraction region of a liquid metal ion source, *J. Vac. Sci. Technol.*, 19(4), 1082–1086.

99. Ward, J. W., Kubena, R. L., and Joyce, R. J. 1991. An ion counting apparatus for studying the statistics of ion emission from liquid metal ion sources, *J. Vac. Sci. Technol.*, B9(6), 3090–3094.
100. Ward, J. W., Kubena, R. L., and Utlaut, M. 1988. Transverse thermal velocity broadening of focused beams from liquid metal ion sources, *J. Vac. Sci. Technol.*, B6(6), 2090–2094.
101. Ward, J. W., Utlaut, M., and Kubena, R. L. 1987. Computer simulation of current density profiles in focused ion beams, *J. Vac. Sci. Technol.*, B, 5(1), 169–174.
102. Wells, O. C. 1974. *Scanning Electron Microscopy*, McGraw-Hill, New York, pp. 29–36.
103. Wu, M. C., Boenki, M. M., Wang, S., Clark, W. M., Jr., Stevens, E. H., and Utlaut, M. 1988. GaAs/GaAlAs distributed Bragg reflector laser with focused ion beam, low dose dopant implanted grating, *Appl. Phys. Lett.*, 53(4), 265–267.
104. Xu, Z., Kosugi, T., Gamo, K., and Namba, S. 1989. *J. Vac. Sci. Tech.*, B7, 1959–1962.
105. Yamaguchi, H., Shimase, A., Haraichi, S., and Miyauchi, T. 1985. Characteristics of silicon removed by fine focused gallium ion beams, *J. Vac. Sci. Tech.*, B3(1), 71–75.
106. Yao, N. 2007. *Focused Ion Beam Systems*, Cambridge University Press, Cambridge, U.K.
107. Young, R. and Puretz, J. 1995. Focused ion beam insulator deposition, *J. Vac. Sci. Technol.*, B13(6), 2576–2579.
108. Zeigler, J. F. 1991. IBM research, Yorktown, New York.
109. Zworykin, V. 1945. *Electron Optics and the Electron Microscope*, Wiley, New York.

---

# 12 Aberration Correction in Electron Microscopy

*Ondrej L. Krivanek, Niklas Dellby, and Matthew F. Murfitt*

## CONTENTS

12.1	Introduction .....	601
12.2	Historical Background .....	602
12.2.1	Proof-of-Principle Correctors.....	603
12.2.2	Working Correctors.....	604
12.2.3	Third-Generation Correctors .....	605
12.3	Corrector Optics.....	605
12.3.1	Trajectory Calculation.....	605
12.3.2	The Aberration Function.....	606
12.3.3	The Effect of a Single Multipole .....	614
12.3.4	Combination Aberrations .....	617
12.3.5	Misalignment Aberrations .....	618
12.3.6	Corrector Types .....	619
12.3.7	Corrector Operation.....	625
12.3.8	Aberrations of the Total System.....	626
12.4	Aberration Diagnosis .....	627
12.4.1	Diagnostic Methods.....	627
12.4.2	Computer Control.....	631
12.5	Aberration-Corrected Optical Column.....	631
12.5.1	Description of the Column .....	631
12.5.2	Performance of the System .....	633
12.6	Conclusions .....	637
	Acknowledgments.....	638
	References .....	638
A.1	Appendix.....	640

## 12.1 INTRODUCTION

Like much else in real life, electron-optical systems are not perfect. Some imperfections arise because of fundamental physical reasons. Spherical and chromatic aberrations are well-known examples of this: in round lenses of the type normally used in the focusing of electrons and other charged particles, these aberrations are always positive and typically of the order of the focal length of the lens. Such aberrations are called *fundamental aberrations*. They can be characterized and additions to the optical systems can typically be built, which are able to correct them.

Other imperfections arise because of more practical reasons: lack of precision in machining, mistakes made in aligning the different components of the system, thermal drift, instabilities of power supplies, lack of homogeneity in the materials used to construct the system, etc. These lead to what are commonly called *parasitic aberrations*. The more complicated the optical system, the broader is the palette of parasitic aberrations that can arise and become important, and the more

serious their effects tend to become. This is particularly so because the performance of an improved system can no longer tolerate aberrations that would have been acceptable prior to the improvement. Just like the fundamental aberrations, the parasitic aberrations must also be characterized and made sufficiently small. The effort needed for this is often comparable to the effort expended on correcting the fundamental aberrations.

Similar to the increased sensitivity to parasitic aberrations, the improved system cannot tolerate instabilities that would have been acceptable prior to the improvement. This is best driven home by the dictum: aberration correctors correct aberrations, not instabilities. Another large part of the effort of producing a well-performing aberration-corrected instrument therefore needs to be expended on diagnosing and fixing instabilities.

Aberration correction is thus a four-part undertaking:

1. Aberrations that are important for each particular optical system must be understood.
2. Ways to correct them must be devised and devices for doing so built.
3. Parasitic aberrations that affect the corrected system's performance must be quantified and dealt with.
4. The overall performance of the system, particularly its stability, must be brought up to the new level made possible by aberration correction.

In this chapter, we present the subject of aberration correction from the point of view of *corrector constructors*, whose aim is to produce a corrected optical system that improves on the performance of a comparable uncorrected system. They need to decide what fundamental aberrations in which optical system to correct, identify and steer away from likely pitfalls inherent to various theoretical solutions, proceed to design and build an actual system, and make it work. We concentrate on correctors of geometric aberrations for transmission electron microscopes (TEMs), that is, on correctors of aberrations that depend on the angle or the position of an electron ray but not on the electron energy. Many of the concepts and practical solutions presented here will also be relevant to optical systems employing charged particles other than electrons, for example, ions. Practical examples in this chapter will be drawn from our own work, and will therefore focus on correctors for TEMs, and in particular scanning transmission electron microscopes (STEMs).

## 12.2 HISTORICAL BACKGROUND

In electron microscopy, a strong magnetic round lens is almost always used as the objective lens (OL), that is, the final lens before the sample in a probe-forming STEM, or the first lens after the sample in a conventional transmission electron microscope (CTEM). This is because the round lens has two very desirable properties: it images the same way in all azimuthal directions, and it provides a relatively high final demagnification (for STEM, initial magnification for CTEM), of  $\sim 100\times$ .

As is well known, all round magnetic lenses suffer from rather strong spherical and chromatic aberrations. The spherical aberration of the OL has been the main obstacle to attaining better STEM and CTEM resolution, since the 1960s. Correcting it has, therefore, become the chief aim of aberration correction in electron microscopy. The correction is performed by an aberration corrector, which makes sure that the beam arriving at the OL in a probe-forming STEM is preaberrated to just the right extent, or that the aberrations of the electron beam leaving the OL are compensated before the beam arrives at the detector of the CTEM.

Up to the present time, correctors that have managed to better the resolution of an uncorrected TEM have either used quadrupoles and octupoles, or sextupoles and round lenses as their principal optical elements. These are the kinds of correctors we will cover in this chapter. Correctors that use other optical elements such as mirrors, charge on axis, and phase plates have been covered by other accounts, for example, Hawkes and Kasper (1996) and Hawkes (2007) (see also Chapter 6 of this book).

### 12.2.1 PROOF-OF-PRINCIPLE CORRECTORS

The first two steps in the four-part approach outlined in the introduction were undertaken brilliantly by Scherzer. He showed that round lens systems of the type used in electron microscopy suffer from third-order spherical and first-degree chromatic aberration, and that these aberrations cannot be made to vanish no matter how careful the lens design (Scherzer, 1936). A decade later, he proposed several solutions to the problem (Scherzer, 1947), including the one that ultimately led to aberration correctors that have succeeded in improving the performance of today's electron microscopes: using nonround optical elements.

Many practical implementations of Scherzer's proposals were designed and built in the decades that followed. Correctors constructed up to the early nineties are best called *proof-of-principle* or *first-generation* correctors. They did not improve on the performance of the better electron microscopes of their day, but they allowed the field to progress by initiating many of the ideas that resulted in practical aberration correction later on.

Scherzer himself proposed a combined spherical and chromatic aberration corrector using electrostatic octupoles and cylindrical lenses (Scherzer, 1947). A corrector based on these principles was built by Seeliger (1953) and Möllenstedt (1956), and shown to work in principle, especially when the energy spread of the electron beam was artificially increased by wobbling the high voltage. But the resolution was not better than the resolution of the best electron microscopes of that time.

Archard in the United Kingdom proposed using quadrupoles instead of cylindrical lenses (Archard, 1954). Several research students in Cosslett's laboratory in Cambridge, United Kingdom explored the use of quadrupoles experimentally. Deltrap built an electromagnetic quadrupole–octupole corrector of spherical aberration for a probe-forming electron-optical column, and experimentally demonstrated the correction of the aberration (Deltrap, 1964a,b). Hardy built a combined electrostatic/electromagnetic quadrupole corrector, and demonstrated correction of both chromatic aberration and spherical aberration (Hardy, 1967). His design included many features of the corrector for a scanning electron microscope (SEM) that has recently reached practical success (Zach and Haider, 1995). Thomson built a quadrupole–octupole corrector using four quadrupoles and three separate octupoles (Thomson, 1968), which had many of the features of the Nion second-generation  $C_s$  corrector (Krivanek et al., 1999). Hawkes formulated the theory of nonround lenses and explored their aberration properties (Hawkes, 1965). The overall conclusion of the Cambridge group was that correctors were too complicated and that they would never produce a practical resolution improvement. Luckily for electron microscopy, they turned out to be as wrong in their pessimism as they were prescient in the approaches they took.

At about the same time as the correction work was being carried out in Cambridge, Crewe's group in Chicago developed the revolutionary cold field-emission scanning transmission electron microscope (CFE-STEM, Crewe, Wall and Welter, 1968). The small spherical aberration coefficient of the OL of this instrument ( $C_s = 0.3$  mm) allowed them to take the first electron microscope images of single atoms (Crewe, Wall and Langmore, 1970), despite the microscope's relatively low primary voltage of 30 kV. The group then devoted a significant effort to aberration correction. They started with a quadrupole–octupole spherical aberration corrector (Beck and Crewe, 1976), similar to Thomson's, realized that sextupole correctors could rival the performance of quadrupole–octupole correctors (Beck, 1979) and produced several designs (e.g., Crewe and Kopf, 1980; Crewe, 1982, 1984; Shao et al., 1988). They also worked out a theoretical design for a sextupole-round lens corrector able to correct all aberration up to sixfold astigmatism (Shao, 1988), whose theoretical properties were similar to the first practical TEM corrector (Rose, 1990; Haider et al., 1995). But they did not succeed in making any of their designs work. This had a chilling effect on funding for aberration correction research, particularly in the United States.

Correctors up to and including the Chicago ones did not manage to control adequately the parasitic aberrations. The earlier correctors typically comprised only the principal optical elements and few or no means of electrical adjustment such that say a parasitic quadrupole moment of an octupole could be nulled. Some of them did incorporate mechanical adjustments but these proved

impractical. This was a major reason that doomed their utility. An electron-optical corrector simply cannot be built perfectly enough, and incorporating enough adjustments for dealing with the parasitic aberrations is as important as incorporating elements that correct the primary aberrations.

The Chicago quadrupole-octupole corrector incorporated some 40 separate alignment controls, but no practical procedure was found for setting these controls so that the parasitic aberrations would be nulled (Beck, 1977). In retrospect it seems that the choice of permendur as the pole-piece material made hysteresis problems much worse than they needed to be, and that not enough attention was paid to developing diagnostic procedures able to quantify the parasitic aberrations.

A major step toward curing parasitic aberrations was made by the Scherzer/Rose group in Darmstadt. Their ambitious spherical/chromatic aberration corrector project incorporated electrical alignment and auxiliary multipole controls, and resulted in some 100 separate controls. The theoretical design was due to Rose (1970) and the practical implementation was worked on by several researchers (see Koops, 1978, for a review). Correction of the targeted aberrations was demonstrated. But there was no resolution improvement over the better microscopes of its day, and the instrument may have been too complicated to be operated routinely.

### 12.2.2 WORKING CORRECTORS

Much progress in nonround optics also occurred in the charged particle accelerators field. Accelerator designers introduced and started using quadrupole lenses (Courant et al., 1952) for *strong focusing* of charged particle beams, and later on also sextupoles and octupoles for correcting second- and third-order aberrations (e.g., Wilson, 2006). Meanwhile, in electron microscopy, non-round optical elements were used to disperse the electron beam in energy to form electron energy-loss spectra (Castaing and Henri, 1952; Wittry, 1969; Crewe et al., 1969; Senoussi et al., 1971). The first such spectrometers typically did not correct aberrations beyond first-order focus but by the end of the seventies, spectrometers and imaging filters that achieved complete correction of second-order aberrations began to appear (Rose and Pejas, 1979; Shuman, 1980). This technology reached a high level of sophistication in imaging filters, that is, instruments that disperse an electron beam into an energy spectrum, select a part of the spectrum, and transform it back into an energy-selected image or diffraction pattern (Krivanek et al., 1991; Rose and Krah, 1995). It is almost certainly not a coincidence that the two groups that made the first successful aberration correctors had previously worked on the optics of spectrometers and imaging filters.

One successful corrector effort originated in the Rose group at Darmstadt and in the European Molecular Biology Laboratory (EMBL) at Heidelberg. It led to two separate working correctors. One was a corrector of spherical and chromatic aberration for an SEM (Rose, 1971; Zach and Haider, 1995). The other was a sextupole-round lens-sextupole spherical aberration corrector for a TEM (Rose, 1990; Haider et al., 1995, 1998; see Rose, 2003 and Lentzen, 2006 for reviews). Both the correctors demonstrated a substantial improvement in the performance of the microscope they were built into. The TEM corrector also introduced new capabilities such as detecting oxygen atoms by phase-contrast imaging (Jia and Urban, 2004). More recently, the TEM sextupole corrector has been adapted for STEM applications (Sawada et al., 2005), and it has allowed the resolution at 300 kV to progress significantly beyond 1 Å (Freitag et al., 2007). Correctors of similar design have also been introduced in Japan (Mitsubishi et al., 2006; Hosokawa et al., 2006).

The other successful effort was our own. We first built a *proof-of-principle*  $C_s$  corrector for a STEM. It demonstrated complete third-order aberration correction and a resolution improvement over the uncorrected microscope it was built into (Krivanek et al., 1997). We then moved onto a mark II corrector design, which used four quadrupoles and three separate octupoles, and had better stability and lower chromatic aberration than the first corrector (Krivanek et al., 1999). It allowed the STEMs it was built into to reach directly interpretable sub-Å resolution for the first time in electron microscopy (Batson et al., 2002; Nellist et al., 2004). It also allowed a much larger beam current than previously possible to be focused into an atom-sized probe. This has led to the first

identification of the chemical type of a single atom by electron energy-loss spectroscopy (Varela et al., 2004) as well as to atomic-resolution elemental mapping, illustrated in this chapter.

Working correctors have recently also been developed for instruments other than TEMs, STEMs, and SEMs. A combined spherical aberration and chromatic aberration corrector for a photoemission electron microscope (PEEM) that uses a tetrode mirror was designed by Preikszas and Rose (1997) and built by Schmidt et al., (2003). A purely electrostatic quadrupole–octupole corrector that can correct chromatic aberration in addition to spherical aberration was designed for focused ion-beam systems by Weißbäcker and Rose (2001), and built by CEOS GmbH. Both these correctors are now working. They are, however, outside this chapter's focus on correctors for TEMs and will, therefore, not be reviewed here.

We may call the working correctors second-generation instruments. Their common characteristic is that they improve the resolution of the instrument they are built into by approximately a factor of two, and increase the probe current available in a given-size electron probe by 10 $\times$  or more. Automatic alignment is available for these systems but they can also be aligned manually by an experienced operator, especially in the STEM case (see Section 12.4.1). When built into a high-performance instrument, they introduce capabilities hitherto unavailable in electron microscopy.

### 12.2.3 THIRD-GENERATION CORRECTORS

Corrector progress will of course not stop with the development of the first practical correctors. New and more ambitious TEM and STEM correctors have been designed in the past few years. Some are now working and some are being built. They possess all the capabilities of second-generation instruments, and they also correct aberrations beyond the third order. Their operation has typically become so exacting that even the initial setup must be done completely under computer control. One such corrector is Nion's third- and fifth-order spherical aberration corrector (Krivanek et al., 2004; Dellby et al., 2006), which can correct aberrations up to and including all the fifth-order ones, and is now working. Another advanced corrector is Rose's quadrupole–octupole *superaplanator* (Rose, 2005), whose aim is to correct chromatic aberration in addition to geometric aberrations.

## 12.3 CORRECTOR OPTICS

### 12.3.1 TRAJECTORY CALCULATION

Quantifying the performance of a particular corrector system requires that the electron trajectories through the system be worked out in detail. Different methods exist for calculating the trajectories and quantifying the resultant aberrations. Several methods were described in Munro's chapter in the previous edition of this book (Munro, 1997). Here, we only mention points of specific interest for practical corrector design.

A convenient way to model the action of optical elements is to use matrix multiplication algebra (e.g., Grivet, 1972). In this approach, using the notation of the particle physics community, a paraxial ray (= ray passing near the optic axis) is described by a vector  $\mathbf{X} = (x, y, x', y', l, \partial)$ , where the optic axis points in the z direction, x and y are the transverse coordinates of the ray, x' and y' the angles of the ray to the optic axis ( $x' = dx/dz$ ,  $y' = dy/dz$ ), l describes the bunching of the charged particle beam and  $\partial$  is the fractional momentum deviation  $\Delta p/p$  of the particle.

Each element of the optical system is described by a matrix  $\mathbf{R}$ , and the passage of a charged particle through an element results in new ray vector  $\mathbf{X}(1)$  given by

$$\mathbf{X}(1) = \mathbf{R}(1)\mathbf{X}(0) \quad (12.1a)$$

Passage through several optical elements is then simply described by successive matrix multiplications, namely,

$$\mathbf{X}(h) = \mathbf{R}(h) \dots \mathbf{R}(3)\mathbf{R}(2)\mathbf{R}(1)\mathbf{X}(0) = \mathbf{Rt}(h)\mathbf{X}(0) \quad (12.1b)$$

where the new vector  $\mathbf{X}(h)$  describes, to first order, how the trajectory of the particle was modified by its passage through the different elements of the system, and  $\mathbf{Rt}(h)$  is a matrix describing the total action of all the optical elements up to  $h$ .

The matrix multiplication approach is readily extended to second and higher orders by the addition of extra terms. To second order, the components of the final coordinate vector, in terms of the original, is given by

$$\mathbf{X}_i(h) = \sum_j \mathbf{Rt}_{ij}(h) \mathbf{X}_j(0) + \sum_{jk} \mathbf{Tt}_{ijk}(h) \mathbf{X}_j(0) \mathbf{X}_k(0) \quad (12.2)$$

where  $\mathbf{Tt}_{ijk}(h)$  are the elements of the second-order transfer matrix of the optical system, and the matrix  $\mathbf{Tt}$  is obtained by multiplying the second-order  $\mathbf{T}$  matrices corresponding to all the elements of the beam line up to  $h$ .

The transfer matrices for bending magnets, multipoles, and round lenses have been worked out and incorporated into several computer programs by the particle accelerator community and are available for general use. One such program is Transport (Carey, 2006), which can calculate optical properties of beam lines up to third order.

A disadvantage of programs such as Transport is that their user interface is somewhat unwieldy. It makes them inconvenient during an initial search for an optimal configuration, which usually involves considering many different possibilities. This kind of a search typically starts by exploring just the first-order (Gaussian) optics of the system. A convenient tool for this is an Excel spreadsheet that uses the matrix formulation. It can be set up with one row per stage of the optical system, for example, a drift space, a quadrupole entrance or exit, a round lens, or a bending magnet entrance or exit, and with columns that track the distance traversed in the system ( $= z$ ), lens excitations, axial and field trajectories, plus additional parameters such as the locations and magnifications of intermediate images and diffraction patterns. The multidimensional Excel Solver is then available for homing in on a particular solution, making the initial exploration particularly convenient.

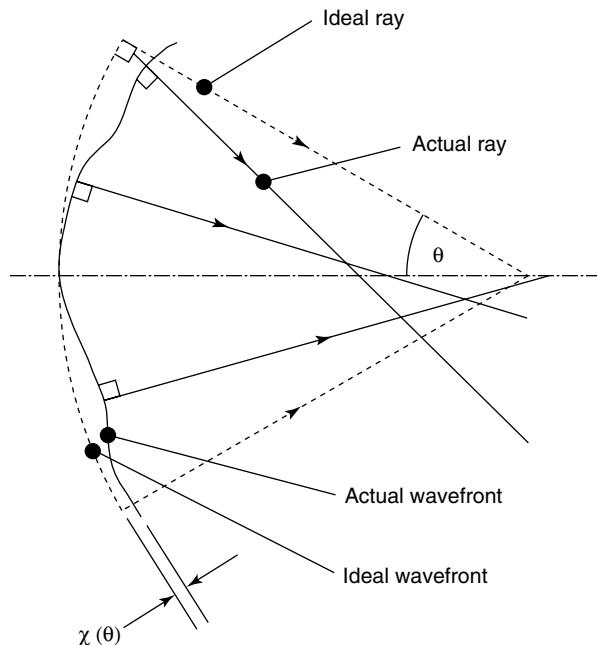
We have set up this kind of a spreadsheet in a form that allows us to introduce a quadrupole or a round lens at any desired stage, simply by entering the quadrupole properties (aperture, pole diameter, and the excitation) or the round lens focal length in an appropriate row and column. When no lens excitation is entered in a row, the row describes a drift space. Changing just one parameter immediately leads to new axial and field trajectories, and to new graphs showing what they are.

In this way, the search for the optimum configuration can be narrowed quickly and efficiently. The calculation then moves onto higher-order properties. Several tools are available for this. As mentioned earlier, the higher-order terms can be tracked in the matrix multiplication approach by the appropriate higher-order matrices. Another approach is trajectory tracing, in which the magnetic and electric fields of the elements of the system are calculated using a finite element approach, the exact trajectories are computed, and their deviations from the first-order trajectories analyzed in terms of aberration coefficients (Munro, 1997). Another powerful approach is the differential algebra one of Berz (1990) and Makino and Berz (1999), which can calculate aberrations up to any order for an arbitrary optical system consisting of multipoles, round lenses, and bending magnets.

### 12.3.2 THE ABERRATION FUNCTION

A limitation of the trajectory calculation approach is that it does not provide direct insight into the optics of the system. In other words, it is good at calculating the properties of any proposed configuration of optical elements, but it provides little insight into how to arrange the elements in a configuration that is likely to give optimum aberration characteristics.

The needed insight is reached more readily if we first calculate the first-order (Gaussian) optical properties of the system, for instance by the matrix approach, and then separately inquire about the system's deviations from the Gaussian trajectories, that is, about its higher-order optical properties.



**FIGURE 12.1** The aberration function concept.

This separation leads to the concept of the aberration function  $\chi(\theta)$  (Born and Wolf, 1980; Lenz, 1971; Hawkes and Kasper, 1996). For the STEM case, the aberration function is defined as the physical distance between the actual wavefront converging on the sample, and an ideal wavefront that would be produced by the Gaussian optical elements in the absence of all aberrations. The function is conventionally defined in the aperture plane, that is, in the front-focal plane of the STEM's OL. The concept is illustrated in Figure 12.1. Even though the aberration function cannot be *measured with a ruler*, the function does correspond to a simple distance, that is, a real physical property. The distance is typically of the order of the electron wavelength  $\lambda$ , which means that it is very small: a few picometers.

The aberration function concept is equally valid for the CTEM case, provided that we reverse the direction of the electron travel so that the electrons diverge from the sample instead of converging on it. The aberration function is then once more equal to the distance between the actual and ideal wavefronts, this time measured in the back-focal plane of the OL.

Reversing the direction of the electron travel, that is, interchanging the source and the detector, leaves the electron trajectories in an optical system unchanged, provided that the magnetic field polarities are reversed at the same time. In electron microscopy, this is known as the principle of reciprocity (Cowley, 1969). It means that the optics (and the image formation) of the TEM and the STEM are exactly equivalent. In the rest of this chapter we will concentrate on probe-forming in the STEM and point out the applicability to TEM when important. We will also use the ability to choose the direction of the electron travel to model the optical system as convenient. For instance, even though in the STEM the electrons first traverse the corrector and then the OL, when calculating the properties of corrector designs, we often proceed as though the electrons started at the sample, where the desirable exact shape of the beam is easy to define, propagated through the OL, and continued to the corrector.

As an interesting aside, the aberration function can be used to quantify the imaging performance by a figure of merit  $R$  familiar from optical astronomy: the diameter of the wavefront (e.g., the diameter of a telescope's primary mirror) divided by the precision with which the wavefront is shaped (the largest height deviations from the ideal shape of the mirror). The figure is roughly the

same for both aberration-corrected electron microscopes and the best optical telescopes. In aberration-corrected electron microscopes, the diameter of the wavefront entering the OL is  $\sim 100 \mu\text{m}$ , and the precision with which the wavefront needs to be shaped is  $\lambda/4 \sim 1 \text{ pm}$ , giving  $R \sim 0.1 \text{ mm}/1 \text{ pm} = 10^8$ . In the best optical telescopes,  $R \sim 10 \text{ m}/100 \text{ nm} = 10^8$  also. This demonstrates that these entirely separate fields of optics have presently reached a similar state of development.

As we shall see in Section 12.4, we can determine the aberration function experimentally for actual optical systems. If we also understand how it can be changed by the optical elements of the system, we will be in a strong position to control the aberrations as needed. A major part of this chapter, therefore, describes the connection between the actions of the various optical elements and the aberration function coefficients.

When defocus and spherical aberrations are the main contributing terms (as was the case in electron microscopy prior to aberration correction), the aberration function is given by

$$\chi(\boldsymbol{\theta}) = C_s \theta^4 / 4 + \Delta z \theta^2 / 2 \quad (12.3)$$

where  $\boldsymbol{\theta} = (\theta_x, \theta_y) = (x', y')$  is the angle with which the electron ray arrives at the sample,  $\theta = |\boldsymbol{\theta}|$ ,  $C_s$  is the coefficient of spherical aberration, and  $\Delta z$  the defocus. We are using a **bold** script for vectors and complex numbers, and a regular script for scalars.

The energy flow is always normal to the physical wavefront. This means that a ray arriving at the sample in direction  $\boldsymbol{\theta}$  is displaced a distance  $\mathbf{s}$  from the axis

$$\mathbf{s} = (-\partial \chi(\boldsymbol{\theta}) / \partial \theta_x, -\partial \chi(\boldsymbol{\theta}) / \partial \theta_y) \quad (12.4a)$$

which becomes

$$|\mathbf{s}| = C_s \theta^3 + \Delta z \theta \quad (12.4b)$$

This is the usual expression for the blurring of the probe (or image) due to the effects of spherical aberration and defocus.

When other aberrations are important, the aberration function can be expressed as a double sum over aberrations of different orders and multiplicities:

$$\chi(\theta, \phi) = \sum_n \sum_m \left\{ C_{n,m,a} \theta^{n+1} \cos(m\phi) + C_{n,m,b} \theta^{n+1} \sin(m\phi) \right\} / (n+1) \quad (12.5)$$

where  $\phi$  is the azimuthal angle of the ray arriving at the sample with angle  $\theta$ , that is,  $\phi = \text{atan}(\theta_x, \theta_y)$ .

Taking a hint from quantum mechanics in which  $n$  is used to denote the principal quantum number, we use  $n$  to denote the *order* of an aberration and  $m$  to denote its *multiplicity*, that is, the number of times  $\chi(\theta, \phi)$  reaches a maximum as  $\phi$  is swept from 0 to  $2\pi$ . The first summation is carried out for all  $n$ 's up to the highest aberration order of interest. The summation over  $m$ 's is carried out up to  $n+1$  over a series of even  $m$ 's ( $m = 0, 2, 4, \dots, n+1$ ), for  $n = \text{odd}$ , and odd  $m$ 's ( $m = 1, 3, \dots, n+1$ ) for  $n = \text{even}$ . (The aforementioned notation system was introduced by Krivanek, Dellby, and Lupini (1999), with one minor difference—it called the aberration coefficients  $C_{m,n}$  rather than  $C_{n,m}$  as is done here.)

The  $C_{n,m,a}$  coefficients describe aberrations whose effects are mirror-symmetric about the principal longitudinal plane of the system, while the  $C_{n,m,b}$  coefficients describe antisymmetric aberrations (called skew aberrations in accelerator design parlance). Note also that when  $m=0$ , there is no azimuthal variation in the effect of the aberration, and hence there is no  $C_{n,0,b}$  coefficient.

Equation 12.5 may be expressed more compactly writing  $\theta = \theta_x + i\theta_y = \theta e^{i\phi}$ , and  $C_{n,m} = C_{n,m,a} + iC_{n,m,b}$ . It then becomes

$$\chi(\theta) = \sum_n \sum_m \chi_{n,m}(\theta) = \sum_n \sum_m \operatorname{Re} \left\{ C_{n,m} \theta^{(n+1)} e^{-im\phi} \right\} / (n+1) \quad (12.6a)$$

where  $\operatorname{Re}$  denotes the real part of the expression in the curly brackets, and we have also shown how the aberration function  $\chi(\theta)$  itself may be expanded into aberration components. For computer manipulation, it is usually better to write Equation 12.6a out more explicitly as

$$\chi(\theta) = \sum_n \sum_m \operatorname{Re} \left[ C_{n,m,a} \left\{ (\theta_x - i\theta_y)^m \right\} + iC_{n,m,b} \left\{ (\theta_x - i\theta_y)^m \right\} \right] (\theta_x^2 + \theta_y^2)^{(n-m+1)/2} / (n+1) \quad (12.6b)$$

The terms of the expansion are readily generated up to any order with computer algebra. They are given for aberrations up to fifth order in Appendix A.1 of this chapter.

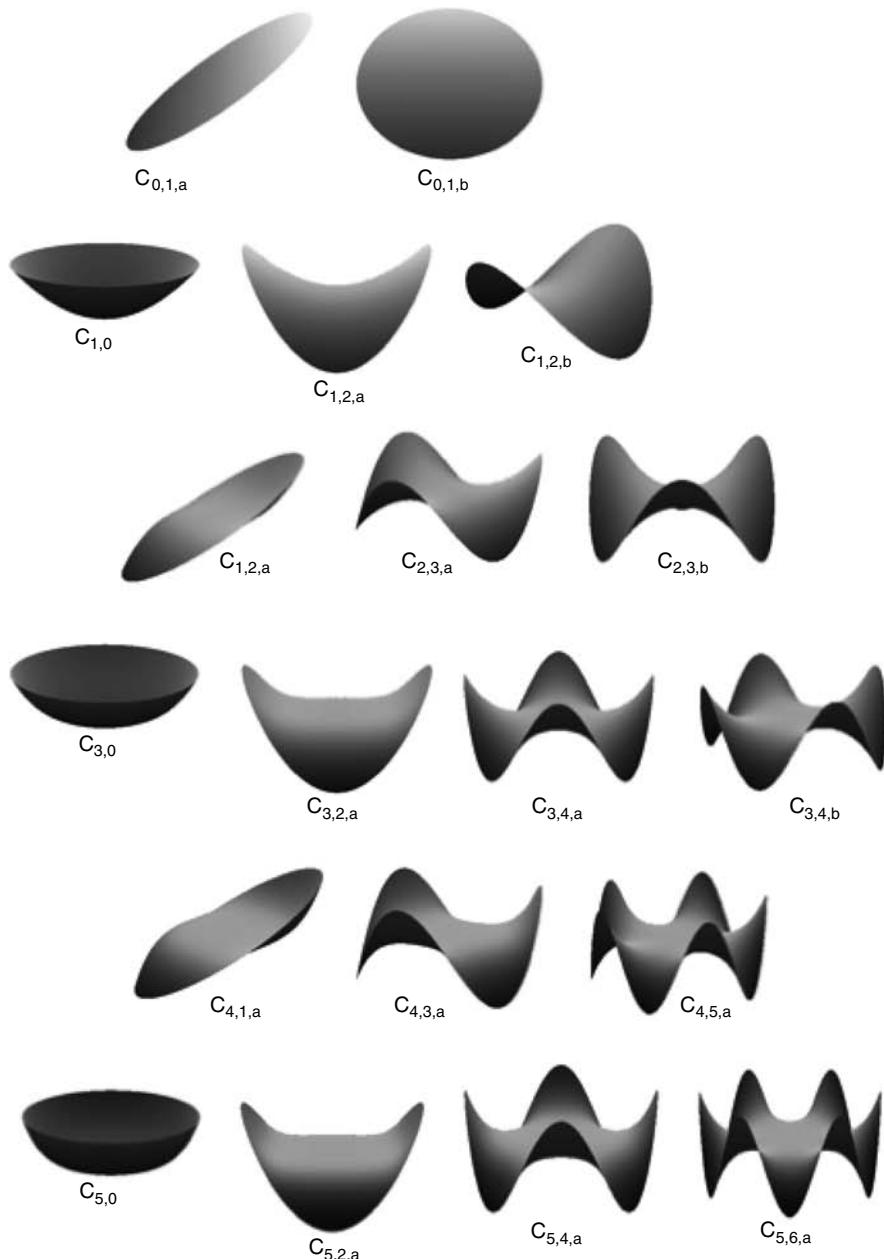
In Equations 12.5 and 12.6, spherical aberration  $C_s$  becomes  $C_{3,0}$ , which we often abbreviate as  $C_3$ . Defocus becomes  $C_{1,0}$  ( $C_1$  for short). Other important aberrations are first-order astigmatism  $C_{1,2}$ , axial coma  $C_{2,1}$ , threefold astigmatism  $C_{2,3}$ , and fourfold astigmatism  $C_{3,4}$ , that is, astigmatisms for which  $n = m - 1$ . Note that each such astigmatism needs to be called *m-fold astigmatism of n-th order*, as there are similar *m-fold astigmatisms* of higher orders for which  $n = m + 1, m + 3$ , etc. These all have the same azimuthal dependence, but different radial dependencies. Two other useful terms are  $C_{0,1,a}$  and  $C_{0,1,b}$ . They describe a simple displacement of the probe and hence of the whole image. Expressing displacement in this way makes it clear that image shift is related to all aberrations of multiplicity  $m = 1$ , such as axial coma  $C_{2,1}$  and fourth-order axial coma  $C_{4,1}$ .

In the large body of literature dealing with aberrations of optical systems, our notation for the aberrations is of course not unique. Notation systems for high-order aberrations can be found for instance in the work of Hawkes (1965), Born and Wolf (1980), Hawkes and Kasper (1996), Berz (1990), Saxton (2000), and Rose and coworkers (Uhlemann and Haider, 1998; Rose, 2003; Müller et al., 2006). However, we believe that our notation system has advantages that would make its wider adoption beneficial in the field of aberration correction. It is readily extensible to axial aberrations of all orders and multiplicities, and its indices make it immediately apparent what the nature of the aberration is. This is not true of aberrations notation systems presently in use in electron optics, which *mostly* assign different letters to aberrations with different multiplicities, keep the same letter for the highest multiplicity aberration of any order, and need to invent as yet unspecified letter assignments for aberrations of higher and higher orders (e.g., Müller et al., 2006).

Figure 12.2 shows surfaces corresponding to aberration function contributions from axial aberrations up to fifth order. For most lower-order aberrations, both the a- and b-type contributions are shown; for higher-order aberrations, only the a-terms that are symmetric about the  $\theta_x$  axis are shown. To generate the b-terms, the surfaces are simply rotated by  $\pi/m$  about the origin.

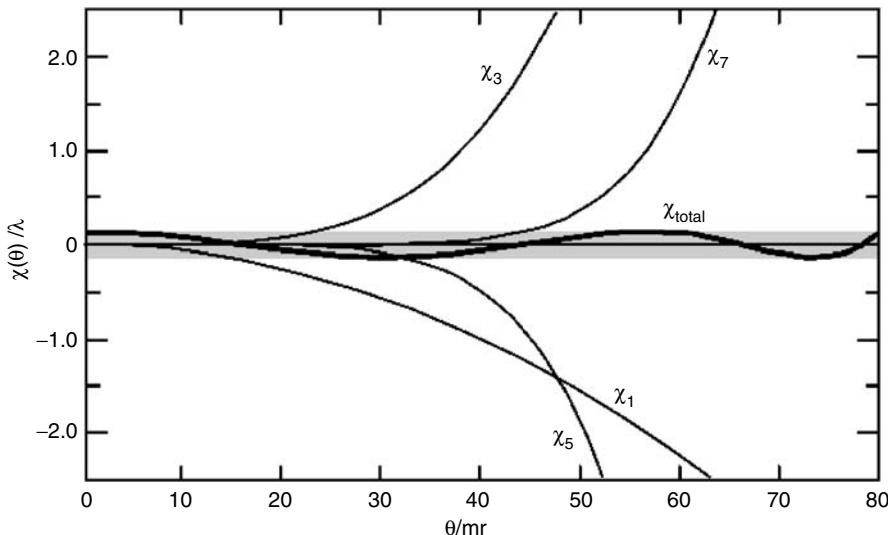
Note how each of the contributions has an instantly recognizable shape, and also that the contributions to the aberration function that have different azimuthal dependencies (= multiplicities) are orthogonal to each other. In other words, adjusting say threefold astigmatism  $C_{2,3}$  has no influence on twofold astigmatism  $C_{1,2}$ . Aberrations of different radial dependence but the same azimuthal dependence are not orthogonal to each other. This makes the shapes of the aberration functions due to, for instance defocus and spherical aberrations, quite similar, as can be seen in Figure 12.2. Other closely related contributions are simple shift ( $C_{0,1}$ ), which gives an inclined plane, and second- and fourth-order axial comas ( $C_{2,1}$  and  $C_{4,1}$ ), which can be viewed as distorted inclined planes.

These types of similarities mean that the effects of higher-order aberrations of a given multiplicity can be partially compensated by lower-order aberrations of the same multiplicity. Using defocus to partially compensate for spherical aberration is a stratagem familiar to most electron



**FIGURE 12.2** Contributions to the aberration function due to aberration coefficients from the zeroeth to the fifth order.

microscopists. The discussion just presented makes it clear that this is a universal concept, applicable to all aberrations of the same multiplicity. The effect of the compensation can be quite large. Figure 12.3 shows the practical case of  $C_{7,0}$  being compensated by properly selected values of  $C_{5,0}$ ,  $C_{3,0}$ , and  $C_{1,0}$ , as can be done in a corrector in which aberrations of up to fifth-order are freely adjustable, plus a constant offset. The compensation manages to keep the aberration function within a band spanning from  $-\lambda/8$  to  $\lambda/8$ , as is needed for proper imaging at maximum attainable resolution. Turning the compensation of the effect of  $C_{7,0}$  by lower-order aberrations off would result in the aberration function  $\chi_{7,0}$  reaching a value that is 128 $\times$  higher.



**FIGURE 12.3** Partial compensation of  $\chi_{7,0}(\theta)$  by lower-order aberration function components. The gray band near the x-axis covers the range  $-\lambda/8 < \chi < \lambda/8$ .

Note that although the average value of a high-order aberration can be compensated to a large degree, it is not readily possible to compensate for fluctuations in the aberration (which would involve tracking the fluctuations and making the lower-order aberrations fluctuate correspondingly). This means that once the compensation is optimally adjusted, the high-order aberration has to stay close to the value the compensation was set up for.

Table 12.1 shows how precisely different aberration coefficients of up to seventh order need to be controlled for 1 Å resolution in high-angle annular dark-field (HAADF) imaging in a 100 kV STEM and for 0.5 Å HAADF resolution in a 200 kV STEM. There are two columns for each case, one showing the maximum value an aberration can have for the desired resolution if not compensated, and one showing the maximum value permissible when lower-order aberrations are compensating for it optimally. The smaller value also represents the maximum allowed deviation from the value that is being compensated.

The table makes it clear that large aberration coefficients of higher order can be tolerated much more readily than lower-order ones. This arises because the angles  $\theta$  used for image formation in electron microscopy are  $\ll 1$  rad, and the effect of an aberration of order  $n$  is weighed by  $\theta^{(n+1)}$ . It also shows that the tolerance for high-order aberrations depends on their multiplicity. Low-multiplicity high-order aberrations can be partially compensated by lower-order aberrations of the same multiplicity. Highest multiplicity aberrations ( $m = n + 1$ ) have no equivalent lower-order aberrations and their effect, therefore, cannot be diminished.

The table also shows that for aberrations with no azimuthal variation such as  $C_{1,0}$ , and  $C_{3,0}$ , a value twice higher than for aberrations with  $m > 0$  can be tolerated. This is because the  $m = 0$  aberrations produce only positive values of  $\chi$ . Imposing the condition that the aberration function has to stay within a band that is  $\lambda/4$  wide then means that twice as much of the aberration can be tolerated than for aberrations that produce both positive and negative  $\chi$  values.

When considering the combined effect of several aberrations of the same order, it is convenient to use a weighting factor  $F_{n,m}$  that accounts for the different value of each aberration that produces the same maximum deviations of the aberration function. In the presence of optimal compensation, the values of the weighing factor are the reciprocals of the values given in the *compensation factor* column in Table 12.1, except that they are further reduced by 0.5× for all  $m = 0$  aberrations. The weighing factors for all aberrations up to the seventh order are also shown in Table 12.1.

**TABLE 12.1**  
**Maximum Possible Compensation Factors, Weighing Factors, and Maximum Allowed Aberration Coefficients up to Seventh Order for 1 Å STEM Probe Size at 100 keV Primary Energy, and 0.5 Å Probe Size at 200 keV**

Coefficient	Maximum Compensation Factor	Weighing Factor $F_{n,m}$	Maximum Size for 1 Å Probe at 100 keV (nm) (No Compensation)		Maximum Size for 1 Å Probe at 100 keV (nm) (With Compensation)		Maximum Size for 0.5 Å Probe at 200 keV (nm) (No Compensation)		Maximum Size for 0.5 Å Probe at 200 keV (nm) (With Compensation)	
			at 100 keV (nm)	(No Compensation)	at 100 keV (nm)	(With Compensation)	at 200 keV (nm)	(No Compensation)	at 200 keV (nm)	(With Compensation)
$C_{0,1}$	1	1.0	$2.00 \times 10^{-2}$		$2.00 \times 10^{-2}$		$1.00 \times 10^{-2}$		$1.00 \times 10^{-2}$	
$C_{1,0}$	1	$5.0 \times 10^{-1}$	3.60		3.60		1.30		1.30	
$C_{1,2}$	1	1.0	1.80		1.80		$6.70 \times 10^{-1}$		$6.70 \times 10^{-1}$	
$C_{2,1}$	4.1	$2.4 \times 10^{-1}$	$1.20 \times 10^2$		$4.90 \times 10^2$		$3.30 \times 10^1$		$1.30 \times 10^2$	
$C_{2,3}$	1	1.0	$1.20 \times 10^2$		$1.20 \times 10^2$		$3.30 \times 10^1$		$3.30 \times 10^1$	
$C_{3,0}$	4	$1.3 \times 10^{-1}$	$1.40 \times 10^4$		$5.70 \times 10^4$		$2.80 \times 10^3$		$1.10 \times 10^4$	
$C_{3,2}$	6	$1.7 \times 10^{-1}$	$7.10 \times 10^3$		$4.30 \times 10^4$		$1.40 \times 10^3$		$8.60 \times 10^3$	
$C_{3,4}$	1	1.0	$7.10 \times 10^3$		$7.10 \times 10^3$		$1.40 \times 10^3$		$1.40 \times 10^3$	
$C_{4,1}$	17	$5.9 \times 10^{-2}$	$3.90 \times 10^5$		$6.70 \times 10^6$		$5.80 \times 10^4$		$9.90 \times 10^5$	
$C_{4,3}$	7.6	$1.3 \times 10^{-1}$	$3.90 \times 10^5$		$3.90 \times 10^6$		$5.80 \times 10^4$		$4.40 \times 10^5$	
$C_{4,5}$	1	1.0	$3.90 \times 10^5$		$3.90 \times 10^5$		$5.80 \times 10^4$		$5.80 \times 10^4$	
$C_{5,0}$	16	$3.1 \times 10^{-2}$	$4.20 \times 10^7$		$6.70 \times 10^8$		$4.60 \times 10^6$		$7.30 \times 10^7$	
$C_{5,2}$	25	$4.0 \times 10^{-2}$	$2.10 \times 10^7$		$5.20 \times 10^8$		$2.30 \times 10^6$		$5.70 \times 10^7$	
$C_{5,4}$	9.7	$1.0 \times 10^{-1}$	$2.10 \times 10^7$		$2.00 \times 10^8$		$2.30 \times 10^6$		$2.20 \times 10^7$	
$C_{5,6}$	1	1.0	$2.10 \times 10^7$		$2.10 \times 10^7$		$2.30 \times 10^6$		$2.30 \times 10^6$	
$C_{6,1}$	58	$1.7 \times 10^{-2}$	$1.10 \times 10^9$		$6.30 \times 10^{10}$		$8.70 \times 10^7$		$5.00 \times 10^9$	
$C_{6,3}$	38	$2.6 \times 10^{-2}$	$1.10 \times 10^9$		$4.10 \times 10^{10}$		$8.70 \times 10^7$		$3.30 \times 10^9$	
$C_{6,5}$	11	$9.1 \times 10^{-2}$	$1.10 \times 10^9$		$1.20 \times 10^{10}$		$8.70 \times 10^7$		$9.70 \times 10^8$	
$C_{6,7}$	1	1.0	$1.10 \times 10^9$		$1.10 \times 10^9$		$8.70 \times 10^7$		$8.70 \times 10^7$	
$C_{7,0}$	64	$7.8 \times 10^{-3}$	$1.10 \times 10^{11}$		$7.00 \times 10^{12}$		$6.40 \times 10^9$		$4.20 \times 10^{11}$	
$C_{7,2}$	110	$9.1 \times 10^{-3}$	$5.50 \times 10^{10}$		$6.00 \times 10^{12}$		$3.20 \times 10^9$		$3.60 \times 10^{11}$	
$C_{7,4}$	53	$1.9 \times 10^{-2}$	$5.50 \times 10^{10}$		$2.90 \times 10^{12}$		$3.20 \times 10^9$		$1.70 \times 10^{11}$	
$C_{7,6}$	13	$7.7 \times 10^{-2}$	$5.50 \times 10^{10}$		$7.10 \times 10^{11}$		$3.20 \times 10^9$		$4.20 \times 10^{10}$	
$C_{7,8}$	1	1.0	$5.50 \times 10^{10}$		$5.50 \times 10^{10}$		$3.20 \times 10^9$		$3.20 \times 10^9$	

*Note:* When higher-order aberrations are being compensated by lower-order ones, the lower-order aberrations need to be set to non-zero values, with precision corresponding to the their maximum values allowed in the non-compensated case.

In a real system, the influence of all the different aberrations needs to be weighed at the same time. Considering the a and b components of nonround aberrations separately and leaving out the simple shifts, there are 33 aberrations up to and including seventh-order ones that are relevant ( $C_{1,0}$ – $C_{7,8,b}$ ). Some combinations of aberrations lead to a broadened probe, others to a probe whose width is essentially unchanged but whose intensity is lowered because its *tail* has grown more intense. It is, therefore, difficult to provide a general description for which combination of aberrations is permissible and which is not. A better approach is to determine all the relevant aberrations experimentally, and to calculate the expected probe shape (or point spread function in TEM) directly. This is now becoming a common feature in autotuning software. The user can then judge directly whether the combination of aberrations that the instrument is set up for is acceptable, or whether the instrument needs further adjustment.

The requirements for phase-contrast imaging (both in CTEM and STEM modes) are similar but stricter for a given resolution. The 1 Å and 0.5 Å resolution columns shown in the table correspond to the requirements of phase-contrast imaging with 1.64 Å resolution at 100 kV, and 0.82 Å resolution at 200 kV, respectively.

Returning now to the overall considerations of the aberration function, the expressions given in Equations 12.5 and 12.6 assume that the optical properties of the system are the same over the whole image (= field) area of interest. This is known as the isoplanatic approximation. It holds reasonably well for small fields of view used by traditional high-resolution electron microscopy, but not for corrected system with high field aberrations or for wide fields of view used at lower magnifications. The aberration coefficients can be generalized to describe their field dependence by adding superscripts p and q to describe the power and the multiplicity of the field dependence, respectively. The aberration function expression then becomes

$$\chi(\theta, r) = \sum_p \sum_q \sum_n \sum_m r^{p\theta(n+1)} \left\{ C_{n,m,a}^{p,q,a} \cos(q\omega) \cos(m\phi) + C_{n,m,a}^{p,q,b} \sin(q\omega) \cos(m\phi) \right. \\ \left. + C_{n,m,b}^{p,q,a} \cos(q\omega) \sin(m\phi) + C_{n,m,b}^{p,q,b} \sin(q\omega) \sin(m\phi) \right\} / (n+1) \quad (12.7)$$

where  $\omega$  is the azimuthal angle in the field. Equation 12.7 is a rather general expression that allows for many types of nonround aberrations not normally considered in light optics but which may arise with nonround optics. The field multiplicity index q can take on any even value from 0 to p for p = even, and any odd value from 1 to p for p = odd. Typically only the lowest orders in p and q are of practical interest, such as field curvature  $C_{1,0}^{2,0}$ , field astigmatisms  $C_{1,2}^{2,0}$  and  $C_{1,2}^{2,2}$ , and coma  $C_{2,1}^{1,1}$ . For regular coma, which rotates in the sample (field) plane so that the aberrated rays make a *comet tail* that points away from the center of the field,  $C_{2,1,a}^{1,1,a} = C_{2,1,b}^{1,1,b}$  and  $C_{2,1,a}^{1,1,b} = C_{2,1,b}^{1,1,a} = 0$ .

Fully expanded versions of Equation 12.7 are rather lengthy, but they can be readily programmed into a computer. Our notation system has the advantage of making the order of the dependencies of each particular aberration on the image and the angular coordinates clearly apparent. For instance, axial coma, which does not depend on the image coordinates, is denoted by  $C_{2,1}^{0,0} = C_{2,1}$ , while regular coma familiar from light optics, which depends linearly on the image coordinates, is denoted by  $C_{2,1}^{1,1}$ .

The aberration function represents the effect of the entire optical system up to the plane in which it is defined. When relating it to the actions of the individual elements in the system, it can be useful to define the contribution to the total aberration function that arises from each separate element:

$$\chi(\theta) = \sum_h \chi_h(\theta) \quad (12.8)$$

where h denotes the different elements of the system. Note, however, that unlike the transfer matrices  $\mathbf{R}(h)$  (Equation 12.1) that are independent of each other, the contribution of each particular

optical element  $X_h$  is influenced by the action of all the elements that precede it, which means that the separate contributions to  $X$  are not independent. Combined actions of different elements produce combination aberrations, which are discussed in Section 12.3.4. Nevertheless, when the optical system is near its desired state and the adjustments remaining to be made are small, the combination aberrations are often negligible. The contributions from the different elements can then be treated independently, and this provides useful insights into the overall properties of the system.

The action of an element that is extended in the  $z$  direction can similarly be expressed as

$$\chi_h(\theta) = \int_0^{Z_s} \chi'_h(\theta, z) dz \quad (12.9)$$

where  $\chi'_h(\theta, z) = d\chi_h(\theta, z)/dz$  and the integration limits are taken over the optical length of the element, which may be greater than its physical length.

We next turn to chromatic aberrations, which describe the dependence of the aberration function on the electron's energy. The aberration coefficients given earlier then become functions of the electron energy and gain appropriate expansions to account for chromatic aberrations of first order, second order, and so on. For the needs of electron microscopy, which uses rather monochromatic beams, considering only the linear terms is typically sufficient. Moreover, only the energy dependence of the lowest order aberrations is typically of interest, that is, of defocus and astigmatism, and at times also the image displacement. These are given by

$$C_1(E) = C_1 + C_{1,c}(E - E_0)/E_0 = C_1 + C_c(E - E_0)/E_0 \quad (12.10a)$$

$$C_{1,2}(E) = C_{1,2} + C_{1,2,c}(E - E_0)/E_0 \quad (12.10b)$$

$$C_{0,1}(E) = C_{0,1} + C_{0,1,c}(E - E_0)/E_0 \quad (12.10c)$$

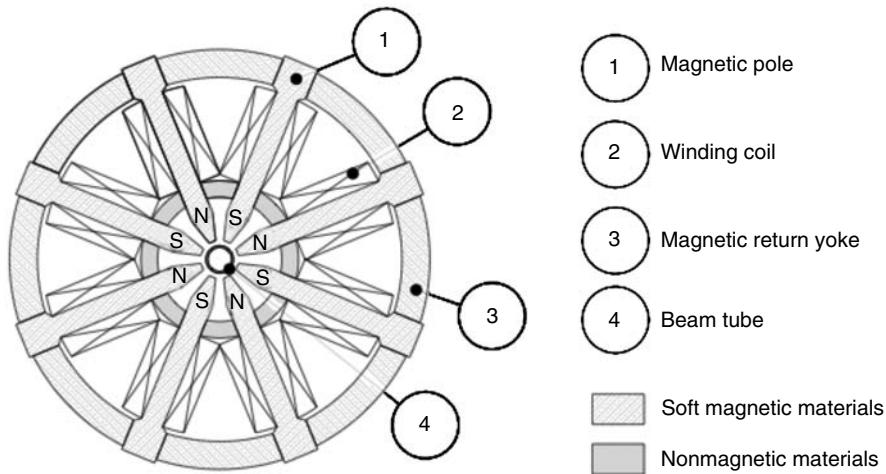
Here  $E$  is the actual electron energy, and  $E_0$  the peak value of the electron energy distribution. We abbreviate  $C_{1,c}$  as  $C_c$ , in accordance with the usual convention.

### 12.3.3 THE EFFECT OF A SINGLE MULTIPOLE

The aberration function is changed by all the optical elements of the system, which typically include accelerators, round lenses, and multipoles. The aberrations of round lenses have been treated extensively in electron-optical literature (e.g., Hawkes and Kasper, 1996). They depend on parameters such as the shape of the lens pole-pieces and the perfection of its machining. But once a round lens is made, its aberrations typically cannot be changed independently of its first-order focusing properties, that is, without changing the Gaussian optics of the system. The same limitation applies to accelerators.

Multipoles of multiplicity  $m > 2$ , on the other hand, can be used to adjust a variety of aberrations, without changing the Gaussian trajectories. This gives them a central role in aberration correction. We focus our description on the change in the aberration function that is produced by the excitation of a single multipole. A more comprehensive discussion of the electron-optical effects of the elements used in corrector construction may be found in Grivet (1972) or Carey (2006).

In correctors employed in transmission electron microscopy, magnetic multipoles are typically used. A schematic cross section of a magnetic octupole is shown in Figure 12.4. The same current is run through the eight windings of the octupole, but with opposite polarities for neighboring poles. This produces alternating north–south pole pattern at the magnetic poletips. A nonmagnetic beam tube in the center of the octupole provides the vacuum needed to pass a beam of electrons through the octupole in a direction generally normal to the plane of the figure.



**FIGURE 12.4** A schematic cross section of a magnetic octupole.

The principal effect of a single multipole of  $2m$  appropriately shaped poles acting on a round beam is to produce an aberration of the type  $C_{m-1,m}$ . The octupole depicted in Figure 12.4 thus produces a fourfold astigmatism  $C_{3,4}$ , a sextupole produces a threefold astigmatism  $C_{2,3}$ , and so on.

To relate the change in  $C_{m-1,m}$  to the current supplied to the windings, we use the sharp cut-off of fringing fields (SCOFF) approximation, under which the magnetic scalar potential near the axis of a multipole with  $2m$  appropriately shaped poles is given by

$$\Phi(\mathbf{r}) = \Phi_a(r/a)^m \cos(m\phi) \quad (12.11)$$

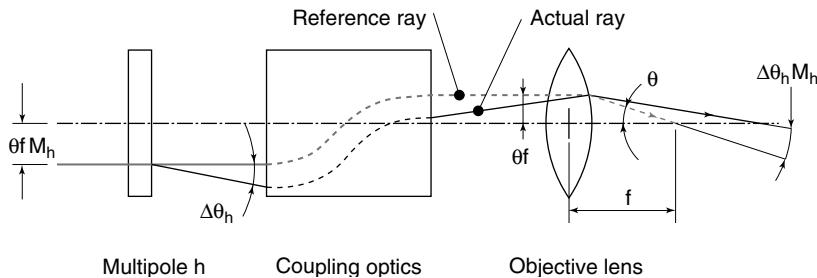
in a region spanning the *effective length*  $L$  of a multipole, and zero everywhere else.  $\mathbf{r} = x + iy$ ,  $\Phi_a$  is the magnetic potential at the poles, and  $a$  is the distance between each pole tip and the axis (= multipole aperture radius). The potential at the poles is given by

$$\Phi_a = \mu_0 v NI \quad (12.12)$$

where  $\mu_0$  is the permeability of free space,  $v$  a factor characterizing the efficiency of the magnetic circuit ( $v \sim 1$  in a loss-less, nonsaturating circuit with  $\mu_r \gg 1$ ), and  $\pm NI$  are the ampere-turns applied to each pole of the  $2m$  multipole. The proper shape of the poles is one which follows the same iso-intensity contour in the distribution described by Equation 12.11 for the plus poles, and the corresponding negative intensity contour for the minus poles.

The potential distribution naturally satisfies the Laplace equation (Grivet, 1972). Its form is the same as the form of the aberration function due to aberration coefficient  $C_{m-1,m,a}$  (Equation 12.5). This is why there is a direct correspondence between a multipole  $2m$  and the aberration  $C_{m-1,m}$ .

To describe the effect of the multipole on the aberration function quantitatively, we need to define the optical arrangement linking the multipole to the OL of the system. A suitable arrangement is shown in Figure 12.5. A multipole of  $2m$  poles is coupled by a system of first-order focusing lenses (which could be round lenses, quadrupoles, and drift spaces, or a combination thereof) to an OL of focal length  $f$ . For the moment we assume that the lens arrangement projects the mid-plane of the multipole into the plane in which the aberrations of the OL appear to act on the beam (also known as the coma-free plane of the OL), with the same magnification in all directions, that



**FIGURE 12.5** Schematic drawing of the coupling scheme of a multipole to the objective lens.

is, that no combination aberrations arise. We define magnification  $M_h$  at multipole  $h$  as the ratio (beam diameter in multipole  $h$ )/(beam diameter in the OL), that is, we relate  $r$  to  $\theta$  by  $r = \theta f M_h$ . We also assume that if the coupling lens arrangement causes the beam to rotate, the coordinate system has been rotated correspondingly.

The magnetic field  $\mathbf{B}$  in the  $2m$  multipole is then

$$\mathbf{B}_x(\mathbf{r}) = -\delta\Phi(\mathbf{r})/\delta x \quad (12.13a)$$

$$\mathbf{B}_y(\mathbf{r}) = -\delta\Phi(\mathbf{r})/\delta y \quad (12.13b)$$

$$\mathbf{B}(\mathbf{r}) = -\frac{m}{a}\Phi_a\left(\frac{r}{a}\right)^{m-1} e^{-i(m-1)\phi} \quad (12.13c)$$

where we are writing  $\mathbf{B}(\mathbf{r}) = B_x(\mathbf{r}) + iB_y(\mathbf{r})$ . The effect of a magnetic field  $\mathbf{B}$  is to cause the electron trajectory to deviate by

$$r'' = \partial^2 r / \partial z^2 = 1/\rho = e\mathbf{B}/p \quad (12.14)$$

where  $\rho$  is the radius of curvature of the deviated ray,  $e$  the electron charge and  $p$  the relativistic momentum of the electron. To relate the deviation to the aberration function, we note that the change in the angle of the ray on traversing the multipole  $k$  whose effective length is  $L$  is given by

$$\Delta\theta_k = L/\rho \quad (12.15)$$

and that the corresponding change in the angle in the OL is  $\Delta\theta = \Delta\theta_h M_h$  (which follows from Liouville's theorem). We also note that  $\Delta\theta$  is related to the displacement of the ray in the image plane  $s$  and hence to the change in the aberration function by

$$\Delta\theta = s/f = -(d\chi(\theta)/d\theta)/f = -(\Delta C_{m-1,m}/f)\theta^{m-1} \quad (12.16)$$

This leads to

$$\Delta\theta/M_h = \Delta\theta_h = LeB/p = -(\Delta C_{m-1,m}/(fM_h))\theta^{m-1} \quad (12.17)$$

and finally

$$\Delta C_{m-1,m} = -LeBfM_h/(p\theta^{m-1}) = mLe\Phi_a(fM_h/a)^m/p \quad (12.18a)$$

This expression may be written in terms of the number of ampere-turns NI per pole of the multipole and the relativistically corrected primary energy of the electrons  $E^*$  (where  $p = \sqrt{2m_0 E^*}$ )

$$\Delta C_{m-1,m} = mLe(f_0 M_h/a)^m \mu_0 v NI / \sqrt{2m_0 E^*} \quad (12.18b)$$

where  $m_0$  is the mass of the electron at rest.  $E^* = E_0 (1 + E_0/2 m_0 c^2)$ , where  $E_0$  is the actual primary energy and  $c$  the speed of light. For clarity we may also neglect all unit-dependent scaling factors and simply note the proportionality as

$$\Delta C_{m-1,m} \propto m L N I (M_h/a)^m / \sqrt{E^*} \quad (12.18c)$$

which shows that the change in an aberration coefficient of the  $(m - 1)$ th order depends linearly on the magnetic excitation and the effective length of a  $2m$  multipole, on the magnification of the beam diameter in the multipole to power  $m$ , on the multipole aperture to power  $-m$ , and on the inverse square root of the *relativistically corrected* electron energy. Several insights follow from the expression, such as that to attain a stronger effect from a high- $m$  multipole, it is much more effective to increase the magnification  $M_h$ , or to decrease the multipole aperture  $a$ , than to increase the multipole length  $L$  or its excitation NI.

When the proper shape of the multipole poles is not followed in practice, parasitic aberrations with multiplicities equal to odd multiples of  $m$  arise ( $m' = 3m, 5m, \text{etc.}$ ). An imperfect dipole will be accompanied by a parasitic sextupole and a weaker parasitic decapole, an imperfect quadrupole by a dodecapole, etc. The extra deviation of the ray caused by these parasitic aberrations is of comparable magnitude to the primary  $m$ -fold aberration right at the multipole's polefaces, but it decays as  $r^{3m-1}$  or faster toward the axis of the multipole. The principal ray deviation due to an  $m$ -fold multipole decays only as  $r^{m-1}$ . This means that for multipoles whose polefaces are situated a few millimeters from the axis, and an electron beam extending only over tens of microns from the axis, as is typical in electron microscopy, the effect of the higher-multiplicity parasitic multipoles will be millions of times or more weaker than the primary effect of the multipole. The parasitic multipole effects can, therefore, almost always be neglected and the poles shaped as convenient for the chosen mechanical construction method.

Dipoles ( $m = 1$ ) are an exception to this rule. They typically need to be designed so that their parasitic sextupole moment is zero. Another exception is quadrupoles acting on wide beams, as in postsample projection optics, where the parasitic effects of the inexact pole shape can become important too.

The effect of a small change in the excitation of a round lens can be modeled in the aberration function picture too. Disregarding the rotation of the electron beam, the small change causes a change of defocus  $C_{1,0}$ , and there are typically also spherical aberration contributions of third and higher orders ( $C_{3,0}, C_{5,0}, \text{etc.}$ ). Because of the round symmetry of the lens, aberrations other than ones with zero multiplicity are not changed by it. The symmetry is of course not perfect and parasitic aberrations such as twofold and threefold astigmatisms do arise in a real round lens, but their effects are typically three orders of magnitude weaker than the principal effects of the lens. They can be modeled as though the aperture plane of the round lens contained weak additional multipoles.

### 12.3.4 COMBINATION ABERRATIONS

If round lenses and multipoles acting on round beams were the only components available to corrector constructors, the number of aberrations that could be controlled would be very limited. In reality, however, combining just a few types of optical elements can produce a rich variety of optical effects due to two phenomena: combination and misalignment aberrations.

Combination aberrations arise when a beam shaped by element 1 (with an aberration function  $\chi_1$ ) arrives at element 2, separated by a distance D from element 1, and imparting an aberration function  $\chi_2$ . They also arise in extended (not infinitely thin) optical elements. The basis of the effect is that the passage of the beam through an element changes the shape of the beam and thus the location where a particular ray traverses element 2, or where it traverses an element that is extended in the z direction. This modifies the optical effect of element 2, and thereby gives rise to combination aberrations. A general rule valid for small distortions of the beam is that combining  $C_{n,m}$  aberration produced by the first element with  $C_{u,v}$  aberration produced by the second element gives rise to aberrations  $C_{n+u-1,m-v}$  and  $C_{n+u-1,m+v}$ . The higher-multiplicity combination aberration does not arise when combining two aberrations of highest multiplicity for a given order, for example,  $C_{2,3}$  and  $C_{3,4}$  only give rise to  $C_{4,1}$ .

When the beam distortion produced by element 1 is no longer slight, all aberrations of the  $n + u - 1$  order are excited. As an example, projecting a beam made slightly elliptical by a quadrupole onto an octupole, that is, combining  $C_{1,2}$  with  $C_{3,4}$  results in a combination aberration  $C_{3,2}$ , in addition to the octupole's primary effect ( $C_{3,4}$ ). Projecting a line or a very elliptical beam produced by one or more quadrupoles onto an octupole excites  $C_{3,0}$ ,  $C_{3,2}$ , and  $C_{3,4}$ , in proportions that depend on the aspect ratio of the beam ellipse. This is the basis of quadrupole–octupole correctors, which use beam shapes of differing ellipticities projected into three or more octupoles to produce adjustable  $C_{3,0}$ , while at the same time canceling  $C_{3,2}$  and  $C_{3,4}$  produced in intermediate stages of the system. An example of such a system is described in Section 12.3.6.

As another example, combining two sextupoles produces an adjustable  $C_{3,0}$  as a combination aberration. The same applies to an extended sextupole, which can be regarded as a series of thin sextupoles. In other words, sextupole combinations and extended sextupoles produce third-order spherical aberration that is the same at all azimuthal angles. The sign of the aberration in an extended sextupole is negative, that is, opposite to the spherical aberration of round lenses. This forms the basis of sextupole correctors, also described using a concrete example in Section 12.3.6.

Combining two octupoles or an extended octupole similarly produces an adjustable  $C_{5,0}$ . So does combining two optical elements producing  $C_{3,0}$ , such as a corrector and a round lens. In the case of a corrector producing negative  $C_3$  so as to compensate for positive  $C_3$  of the OL, if the *effective* plane of the corrector, from which the corrective action appears to emanate, and the coma-free plane of the OL are separated by a distance D, the resultant combination  $C_5$  is given by (e.g., Krivanek et al., 2003)

$$C_5 = 3D(C_3/f)^2 \quad (12.19)$$

where f is once more the focal length of the OL. Projecting an image of the corrector into the OL by a system of lenses and adjusting D as needed to null the total  $C_5$  of the optical system thus allows simultaneous correction of  $C_{5,0}$  and  $C_{3,0}$ .

Combining multiple elements can also produce adjustable chromatic aberration. This is based on the fact that the chromatic aberration coefficient  $C_{1,2,c-e}$  of an electrostatic quadrupole is larger than  $C_{1,2,c-m}$  of an electromagnetic quadrupole. A combined quadrupole that has both electrostatic and electromagnetic poles can, therefore, produce negative  $C_{1,2,c}$  (Yavor et al., 1964). When acting on a line or elliptical beam, such a combined quadrupole produces both an adjustable  $C_{1,2c}$  and  $C_{1,0c}$ . Combining two such elements in which the ellipse directions are different gives the compound system an ability to set both  $C_{1,0,c}$  and  $C_{1,2,c}$  of the total apparatus to zero.

### 12.3.5 MISALIGNMENT ABERRATIONS

Misalignment aberrations arise when the beam is shifted or tilted relative to the center of an optical element or its axis. A shifted round beam passing through an ideal multipole picks up aberrations of lower orders than the multipole's primary effect. For example, if a round beam traversing an ideal

octupole that produces  $C_{3,4a}$  as its primary effect is shifted in the octupole such that it arrives at the sample tilted by an angle  $\tau = (\alpha, \beta)$ , the threefold astigmatism, the twofold astigmatism, and the probe shift caused by the octupole change by

$$\begin{aligned} C_{2,3a} &= 3\alpha C_{3,4a} & C_{2,3b} &= -3\beta C_{3,4a} \\ C_{1,2a} &= 3(\alpha^2 - \beta^2)C_{3,4a} & C_{1,2b} &= -6\alpha\beta C_{3,4a} \\ C_{0,1a} &= (\alpha^3 - 3\alpha\beta^2)C_{3,4a} & C_{0,1b} &= (\beta^3 - 3\alpha^2\beta)C_{3,4a} \end{aligned} \quad (12.20)$$

For small misalignments, the second-order aberrations dominate.

More generally, shifting the center of a round beam incident on a  $2m$  multipole by a small amount has the principal effect of changing the aberration coefficient  $C_{m-2,m-1}$ . In this way, shifting the beam center in an octupole ( $m = 4$ ) can be used to adjust  $C_{2,3}$ , as is often done in quadrupole–octupole correctors. Shifting the beam in a sextupole can be used to adjust twofold astigmatism, and this is done in sextupole correctors.

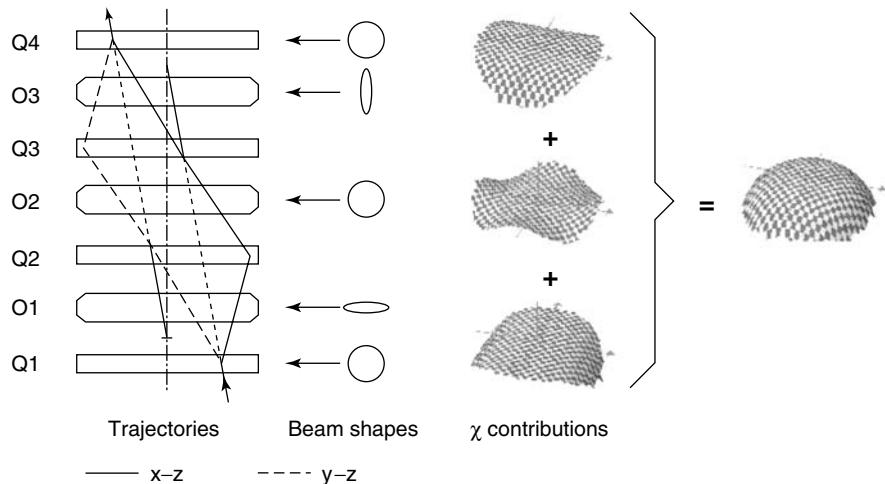
The analysis just presented assumes that the centers of the lower-order multipoles that are implicitly present in a high-order multipole all coincide. In practice, this is not the case, due to imperfections of the multipole that lack the prescribed symmetry, for example, differences in the sizes and magnetic potentials of the different poles, uneven spacing between the poles, etc. To align a multipole in practice so that only pure  $C_{m-1,m}$  is obtained from it, the beam typically needs to be centered so that the highest order miscentering aberration  $C_{m-2,m-1}$  is nulled. If the remaining lower-order miscentering aberrations need to be nulled too, auxiliary coils of  $2(m - 2)$ ,  $2(m - 3)$  etc. multiplicities must be incorporated in the multipole. For instance, to obtain a pure octupole effect, a round beam must be centered on the octupole's *no sextupole* center, and the parasitic quadrupole and dipole moments need to be canceled by appropriately exciting auxiliary coils cowound with the principal ones.

### 12.3.6 CORRECTOR TYPES

Correction of geometric aberrations amounts to identifying and incorporating optical elements able to adjust the different terms in the expansions of the aberration function given in Equations 12.5 through 12.7. The type of the element depends on the aberration to be corrected. Regular astigmatisms (twofold, threefold, and so on) are among the simplest: all that is needed is an adjustable multipole with the appropriate number of poles acting on a round beam. A simple quadrupole corrects regular astigmatism ( $C_{1,2}$ ), a sextupole corrects threefold astigmatism ( $C_{2,3}$ ), as is well known. These elements can, therefore, be regarded as very simple aberration correctors.

Other aberrations need more complicated optical arrangements in which the individual elements act on more than one aberration at a time, but their combined effect amounts to adjusting only the aberration(s) of interest. In transmission geometries (no mirrors) in which there is no charge on or near the optic axis, two broad classes of such arrangements exist.

1. Direct-action correctors, in which the principal aberration is corrected by a correcting element whose main effect is producing aberrations of the needed order. In these systems, undesirable by-product aberrations tend to arise that are of the same order and of similar magnitude as the aberrations being corrected. The by-product aberrations are typically canceled by creating different beam shapes in different stages of the apparatus, which permits the same type of optical element to produce different mixtures of aberration terms.
2. Indirect-action correctors, in which the aberration to be corrected is acted on by a combination aberration arising as a by-product of a lower-order aberration intentionally created in the corrector. The undesirable lower-order aberration coefficients are typically of the same order of magnitude as the coefficients of the aberrations to be corrected. For the



**FIGURE 12.6** Trajectories and aberration function contributions in a quadrupole–octupole corrector.

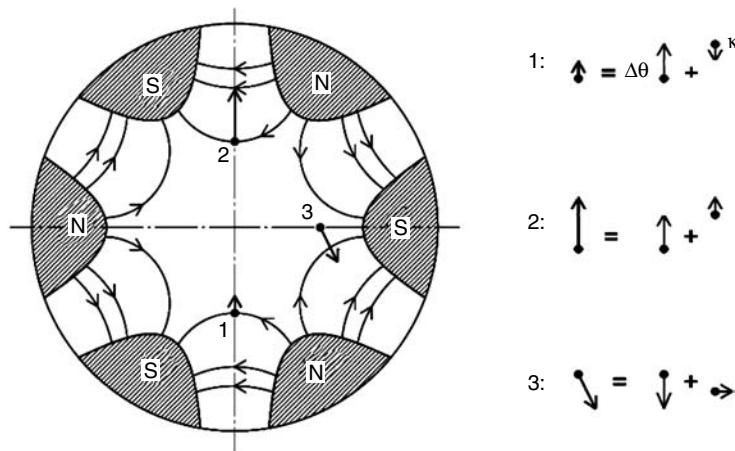
small angles used in electron microscopy, this means that the effect of the lower-order aberrations on the beam are very strong. These aberrations, therefore, need to be canceled very precisely, typically by using an equivalent optical element that cancels the undesirable low-order effect while at the same time increasing the desirable higher-order effect.

For correction of spherical aberration, the direct-action correctors rely on octupoles, whose potential has the right radial dependence to make  $C_3$  adjustable. The azimuthal dependence is of course wrong: the potential switches its sign  $8\times$  around the circumference. But this can be accommodated by sending a line (or elliptical) beam into the octupole, so that the octupole potential is sampled in only (or mostly) one direction. The elliptical beam shapes are produced by quadrupole lenses.

A practical illustration of a quadrupole–octupole corrector is provided by our second-generation corrector. The left side of Figure 12.6 shows the schematic first-order trajectories through the corrector and the beam shape at each important stage of the corrector. A round beam converging on a point inside the corrector is sent into the corrector by the final condenser of the microscope. Quadrupole 1 transforms the round beam into an elliptical beam in octupole 1, quadrupole 2 changes it back into a round beam in octupole 2, quadrupole 3 changes it into an elliptical beam in octupole 3 (with the ellipse orientation at  $90^\circ$  to the ellipse in octupole 1), and quadrupole 4 changes the beam back into a round beam that appears to emanate from a virtual crossover inside the corrector.

The contributions to the aberration function from the octupoles are shown on the right side of Figure 12.6. The contributions from the octupoles with the elliptical beams (octupoles 1 and 3) have the right radial dependence in the long direction of the ellipse in each octupole. Combining the contributions from the two octupoles gives an aberration function that corresponds to the needed adjustable-strength fourth-order paraboloid, plus fourfold astigmatism. The astigmatism is canceled by octupole 2, which acts on the round beam, leaving an adjustable fourth-order paraboloid as the principal effect of the corrector.

Other quadrupole–octupole arrangements that employ the aforementioned principles are readily possible. Examples of different solutions are the Archard–Deltrap corrector (Deltrap, 1964 a,b), Thomson's (Thomson, 1968), and Beck's (Beck and Crewe, 1976) separate quadrupole–octupole correctors, whose optical solutions were similar to the one illustrated in Figure 12.6, the quadrupole–octupole corrector proposed by Rose (1970) and built in Darmstadt, our third-generation quadrupole–octupole corrector (Krivanek et al., 2004; Dellby et al., 2006), and Rose's *ultracorrector* (Rose, 2004) and *superaplanator* (Rose, 2005).



**FIGURE 12.7** The origin of third-order deflections and hence  $C_{3,0}$  in an extended sextupole.

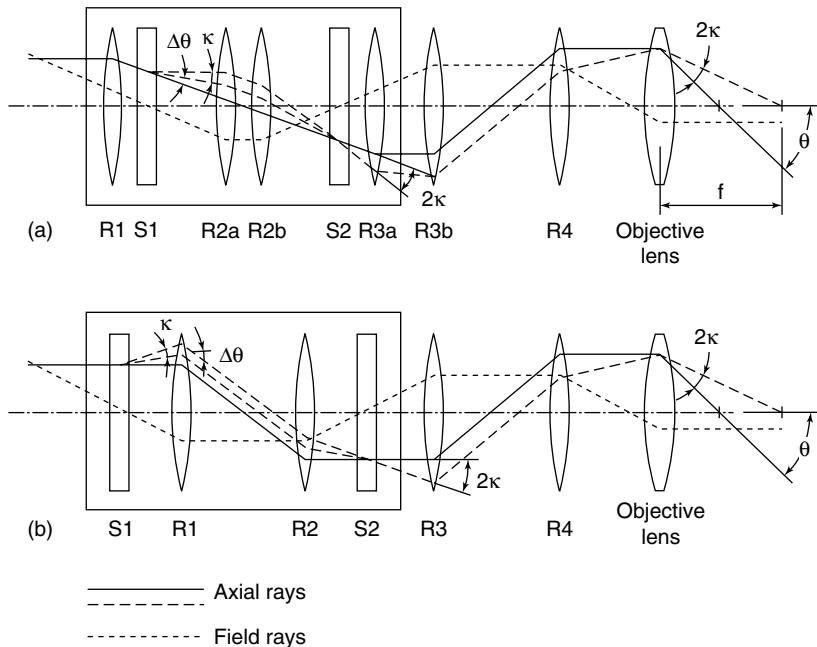
Indirect-action correctors rely on combination aberrations to produce the desired effect. Figure 12.7 shows schematically how an extra deflection that amounts to negative spherical aberration arises in an extended sextupole. The sextupole pole-tip region is shown in cross section. An electron entering the sextupole at point 1 and initially traveling parallel to the optic axis (i.e., perpendicular to the plane of the paper) will be deflected toward the center of the sextupole, as the primary effect. This will bring it into a region where the sextupole field is weaker, thus producing a weaker deflection toward the axis. The decreased deflection due to the weaker field amounts to a small extra deflection  $\kappa$  away from the axis compared to the deflection the electron would have experienced in a thin sextupole, in which the sextupole field would have terminated before the electron trajectory was substantially changed.

An electron entering the sextupole at point 2 will be displaced outward, that is, toward stronger fields. This will also produce an additional outward deflection  $\kappa$  compared to a thin sextupole. Finally, an electron entering the sextupole at point 3 is initially displaced azimuthally. This brings it into a region in which the field is pointing in a different direction, thus also producing an additional outward deflection  $\kappa$  compared to a thin sextupole.

One can also see that the extra deflection will always point outward by looking at the  $C_{2,3}$  surface in Figure 12.2. The direction of the primary deflection  $\Delta\theta$  points in the direction of the steepest downward slope at any point on the surface, and the extra deflection  $\kappa$  results from extra phase shift due to the rate of change of the aberration function being different in the sextupole regions the beam moves over to. The extra phase shift is negative and independent of the azimuthal direction of the deflection. It varies as  $\theta^4$ , exactly as needed for  $C_{3,0}$  correction.

Figure 12.8 shows how the negative  $C_3$  arising in an extended sextupole can be used for correcting the positive  $C_3$  of the OL of a microscope, while compensating for the  $C_{2,3}$  due to the sextupole, and also avoiding exciting a large combination  $C_{5,0}$ . The figure shows the axial and field trajectories of the schemes proposed by Shao (1988) (Figure 12.8a) and by Rose (1990) (Figure 12.8b).

In both designs, sextupole S1 is projected by a pair of round lens into sextupole S2, and by further round lenses into the front-focal plane of a probe-forming OL. (In Shao's article, only one round lens is shown between S1 and S2. But the lens was actually an antirotation round lens doublet.) After traversing the first sextupole, the beam that entered it at radius  $f\theta$  (where  $f$  is the focal length of the OL), is deflected by an angle  $\Delta\theta + \kappa$ , where  $\Delta\theta$  is the principal deflection due to the threefold field of the sextupole (that gives rise to threefold astigmatism  $C_{2,3}$ ) and  $\kappa$  is an additional deflection that amounts to  $C_{3,0}$  arising from the extended nature of the sextupole. Next, the



**FIGURE 12.8** Sextupole—round lens—sextupole corrector designs. (Shao, Z., *Rev. Sci. Instrum.*, **59**, 2429, 1988; Rose, H., *Optik*, **85**, 19, 1990.)

round lenses image S1 into S2. Sextupole S2 affects the beam with opposite polarity to S1, and  $\Delta\theta$  (and thus also the  $C_{2,3}$  due to the first sextupole) is canceled. The combination aberration  $C_{3,0}$ , which is independent of the orientation of the sextupole, is however doubled in size. The beam thus emerges from S2 deflected by a small additional angle equal to  $2\kappa$ . This deviation is propagated all the way to the OL, where it cancels the equal but opposite deviation due to the spherical aberration of the lens. At the same time, S2 is imaged into the so-called coma-free plane of the OL. Imaging S2 (and thus also S1) into this plane avoids introducing a large combination fifth-order spherical aberration  $C_{5,0}$ .

To determine the third-order effect of an extended sextupole quantitatively, we model the sextupole with SCOFF, an effective length of L, and threefold astigmatism that the sextupole contributes at the sample of  $C_{2,3}$ . The paraxial ray equation then describes that the ray deviation as

$$r'' = Sr^2 \quad (12.21)$$

where  $r$  is the  $(x, y)$  position in the sextupole and  $S$  is the sextupole strength. Assuming that the electron entered the sextupole parallel to the axis, we can expand  $r$  in even powers of  $z$  as

$$r(z) = R + az^2/2 + bz^4/4 + \dots \quad (12.22)$$

where  $z = 0$  at the start of the sextupole. We then get

$$a + 3bz^2 + \dots = SR^2 + SRaz^2 + \dots \quad (12.23)$$

Identifying powers leads to  $a = SR^2$  and  $b = SRA/3 = S^2R^3/3$ , that is, to

$$r(z) = R + SR^2z^2/2 + S^2R^3z^4/12 + \dots \quad (12.24)$$

We are interested in the angle with which the ray leaves the exit face of the sextupole of length  $L$ :

$$r'(L) = (dr/dz)_L = SR^2L + S^2R^3L^3/3 - \dots \quad (12.25)$$

The first term describes the primary effect of the sextupole that leads to  $C_{2,3}$ , the second term describes the combination aberration that leads to  $C_{3,0}$ . At the OL, the changes  $\Delta\theta_2$  and  $\Delta\theta_3$  of the slope of the ray due to these two terms are given by  $-C_{2,3}\theta^2/f$  and  $-C_{3,0}\theta^3/f$ , respectively, where  $\theta$  is defined at the OL and  $f$  is again the lens's focal length. Calling the ratio of the beam diameters in the sextupole and in the OL once more  $M$ , we have  $R = -\theta f M$ ,  $\kappa = \Delta\theta_3/M$  and further

$$\Delta\theta_2 = -C_{2,3}\theta^2/f = -SR^2LM \quad (12.26a)$$

$$S = -C_{2,3}/(f^3M^3L) \quad (12.26b)$$

$$\Delta\theta_3 = -C_{3,0}\theta^3/f = S^2R^3L^3M/3 \quad (12.26c)$$

and hence

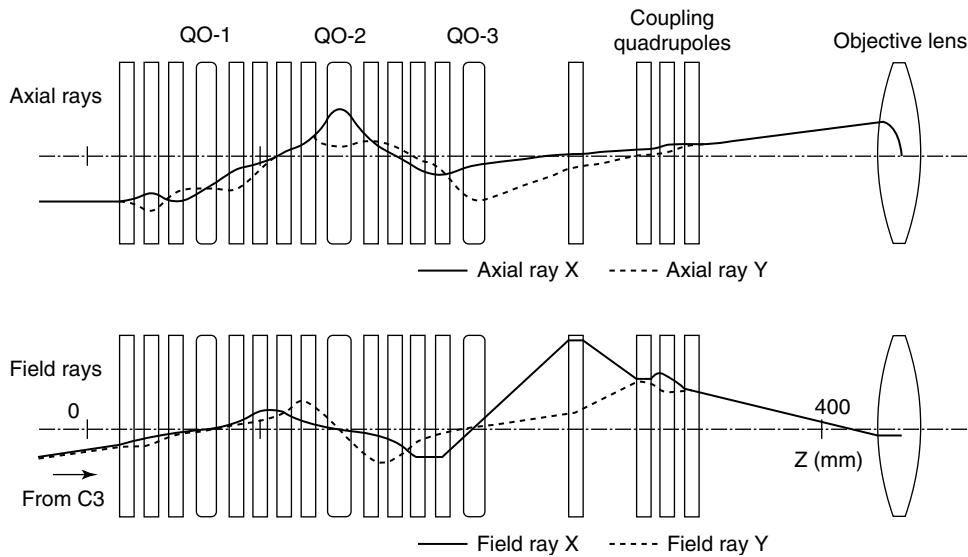
$$C_{3,0} = -(C_{2,3})^2L/(3f^2M^2) \quad (12.27)$$

In other words, the contribution to the spherical aberration from an extended sextupole scales as the second power of the threefold astigmatism that the sextupole would cause at the OL if its primary effect was not compensated elsewhere, and linearly with the length of the sextupole. In practice the  $C_{2,3}$  contribution due to the first sextupole is of course compensated by the second sextupole, which also doubles the  $C_{3,0}$  contribution. Nevertheless, it's useful to examine the primary effect of each sextupole separately.

For typical values of  $f = 2$  mm,  $L = 30$  mm,  $M = 1$ , and total corrector  $C_3$  of  $-1$  mm, the coefficient  $C_{2,3}$  that each sextupole has to produce is equal to 0.45 mm. This is  $\sim 4,000\times$  larger than the largest  $C_{2,3}$  that can be tolerated for 1 Å HAADF resolution imaging, and  $14,000\times$  larger than can be tolerated for 0.5 Å imaging (c.f. Table 12.1). The cancellation of the threefold astigmatism of sextupole 1 by sextupole 2, therefore, has to be very accurate. We will return to this point in the next section.

Both the direct and indirect types of correctors need to be coupled into the OL whose positive spherical aberration they are counterbalancing. The coupling may be done simply without any lenses, but the fact that a positive  $C_3$  of the OL is being compensated some distance from the OL then causes combination  $C_5$ . A better arrangement is to image the corrector into the OL's coma-free plane, that is, the plane in which the  $C_3$  effect of the OL arises. This can be done by a combination of round lenses (as shown in Figure 12.8), or by quadrupole lenses.

Figure 12.9 shows the axial and field trajectories of a complete corrector system based on a combination of quadrupoles and octupoles (Dellby et al., 2006), which can correct aberrations up to and including fifth-order aberrations, without introducing any fundamental aberrations lower than seventh order. The elements labeled Q–O denote combined quadrupole–octupole elements, the unlabeled rectangles simple quadrupoles. This corrector slightly exceeds the performance of current



**FIGURE 12.9** Axial and field trajectories through a  $C_3/C_5$  quadrupole–octupole corrector.

sextupole-round lens correctors, which would need to correct  $C_{5,6}$  by a dodecapole stigmator to match its degree of correction (Müller et al., 2006).

The same type of corrector could be placed after the OL in a mirror-symmetric arrangement about the midplane of the OL. It would then serve as a CTEM corrector. A by-product of  $C_5$  correction is that the field aberrations are considerably reduced for a  $C_3/C_5$  corrector compared to a  $C_3$ -only corrector. Indeed, they are quite comparable (Murfitt et al., 2005) to the sextupole corrector that has become established as the corrector of choice for TEM applications (Haider et al., 1995, 1998).

There are subtle differences between the direct (quadrupole–octupole) and indirect (sextupole-round lens) corrector types. The quadrupole–octupole system needs a larger number of power supplies (~70), whereas the sextupole-round lens system only needs ~30 supplies. But the quadrupole–octupole supplies are all low power and can be packed with high density on printed circuit boards. At 100 keV primary energy, the whole quadrupole–octupole corrector depicted in Figure 12.9 only consumes ~0.4 W of power. The round lenses in the sextupole corrector require hundreds of watts of power, making their power supplies much bulkier. Further, the quadrupole–octupole system does not need any water cooling, which simplifies its construction and makes it more stable, whereas the round lenses in the sextupole corrector would overheat without water cooling.

As was explained in Section 12.3.3, correction of chromatic aberration is achieved by incorporating combined electrostatic/electromagnetic quadrupoles in the optical system. In quadrupole–octupole correctors, replacing two quadrupoles in which the beam is not round by the combined quadrupoles allows correcting  $C_{1,0c}$  and  $C_{1,2c}$ , in addition to all the geometric aberrations. No such extension is available for sextupole correctors.

It needs to be noted that the task of chromatic aberration correction in transmission electron microscopy is rendered harder by the fact that the effects being corrected are quite small. At 200 keV primary energy  $E_0$ , an energy spread of 0.4 eV (readily attainable with a CFE gun) represents a deviation of just  $\pm 1$  ppm relative to  $E_0$ . The power supply stabilities needed for successful chromatic aberration correction are about an order of magnitude higher, and this taxes the capabilities of modern-day electronics. The precision of the alignment of the corrector needs to be improved too. Although the principles are clear, no successful combined geometric and chromatic TEM (or STEM) aberration corrector has been completed to date. This may change in the next year or two

with the  $C_s/C_c$  corrector being built for the TEAM project (Haider et al., 2007). But as its success has not yet been demonstrated, we will leave a more detailed discussion of such a system for the future.

### 12.3.7 CORRECTOR OPERATION

Getting any type of a corrector to work satisfactorily requires that both the principal and the parasitic aberrations be adjusted with sufficient precision, and that the whole system be stable long enough to acquire useful results. The precise correction of the aberrations relies on diagnostic methods described in Section 12.4. In this section, we briefly review the main challenges of day-to-day corrector operation.

Aberrations of different orders need to be nulled with very different precision (c.f. Table 12.1). Defocus  $C_1$  and twofold astigmatism  $C_{1,2}$  need to be controlled rather exactly, coma  $C_{2,1}$  and threefold astigmatism  $C_{2,3}$  typically only need touching up from time to time. Aberrations of third and higher orders can typically be set and not changed again, provided that the height of the sample in the OL is not changed. This requires a sample stage with height ( $z$ ) control that is precise to approximately  $\pm 100$  nm. Such stages are now fairly standard in transmission electron microscopy.

The major effects that can upset the corrector setup are as follows. Misaligning the  $C_{3,0}$  phase-shifted wavefront produced by the corrector sideways in the OL leads to coma. When the wavefront that is preaberrated with the right amount of  $-C_3$  arrives at the OL miscentered by a small distance  $\tau f$  (measured in the coma-free plane of the OL), the main effect is to introduce axial coma of magnitude

$$C_{2,1(\text{miscenter})} = 3C_3\tau \quad (12.28)$$

where the coma points in the same direction as  $\tau$ . This leads to the requirement

$$\tau_{\max} = C_{2,1,\max} / (3C_3) \quad (12.29)$$

The effect occurs in all  $C_3$  corrector types. For maximum tolerable  $C_{2,1}$  of 490 nm (as required for 1 Å HAADF imaging at 100 keV; c.f. Table 12.1) in a system with an OL with  $C_3 = 1$  mm,  $\tau_{\max} = 1.6 \times 10^{-4}$  rad. For an OL focal length of 2 mm, this amounts to a maximum physical displacement of the beam entering the OL by  $\sim 330$  nm. This is a reasonably large distance by electron microscopy standards, but even so coma drift can seriously affect STEMs with condenser lenses prone to sideways drift, such as those in 100 kV VG STEMs.

The sensitivity to misalignment of one octupole with respect to another one in a quadrupole–octupole system is similar to the sensitivity to misalignment of the whole corrector to the OL. This is because the magnitude of the third-order aberrations produced within the corrector is of the same order as the  $C_3$  of the OL. The aberrations introduced when an octupole is misaligned with respect to another octupole are both threefold astigmatism  $C_{2,3}$  and coma  $C_{2,1}$ . Their ratio depends on the beam shapes in the two octupoles. In actual correctors, the octupoles tend to be physically close together, and there are only a few optical elements between them. This makes a large change in their relative position unlikely. As a result, drift of coma or threefold astigmatism due to shifting octupoles is almost never seen in practice.

Within the sextupole corrector system, threefold stigmatism due to the first sextupole needs to be canceled by the second sextupole. When the beam arriving into the second sextupole is miscentered by a small distance  $\tau f$ , the principal resultant effect is twofold astigmatism of magnitude

$$C_{1,2(\text{miscenter})} = 2C_{2,3}\tau \quad (12.30)$$

which leads to

$$\tau_{\max} = C_{1,2,\max} / (2C_{2,3}) \quad (12.31)$$

Twofold astigmatism  $C_{1,2}$  is a first-order aberration, which has to be strictly controlled. With  $C_{2,3}$  of 0.45 mm produced by each sextupole as is typical of sextupole correctors and a maximum allowed  $C_{2,1}$  of 1.8 nm (Table 12.1), the misalignment  $\tau$  has to be smaller than  $4 \times 10^{-6}$  rad. For  $f = 2$  mm, it amounts to a physical displacement of the beam aberrated by sextupole 1 arriving at sextupole 2 by  $<8$  nm. This is about 40× stricter than the misalignment tolerance for projecting the beam from the whole corrector into the OL. It makes sextupole correctors susceptible to astigmatism drift.

Control over misalignment aberrations is used in working correctors to bring parasitic aberrations to within acceptable limits. Sometimes a single control couples into a single parasitic aberration. A dipole between the corrector and the OL is typically used as a coma adjustment; a dipole between the two sextupoles in a sextupole corrector is used as an adjustment of twofold astigmatism. Other adjustments are more complicated. In a quadrupole–octupole corrector, introducing  $C_{2,3}$  by deliberate misalignment of an octupole that contains a round beam, propagating the distorted beam through a  $C_{3,0}$ -producing element such as an octupole with a line beam in it, and then removing the  $C_{2,3}$  with a sextupole leaves nonzero  $C_{4,3}$ . Using this kind of a combination of adjustments,  $C_{4,3}$  can be set to zero.

Finding workable combinations of controls for adjusting important parasitic aberrations is a major part of getting a corrector to work properly. Without it, it is simply not possible to make the actual performance of a working corrector match the theoretical performance that was predicted with no regard to parasitic aberrations. We call these combinations of controls *knobs*. For our quadrupole–octupole correctors, we have developed knobs for up to and including fifth-order aberrations. This is typically done by a combination of theoretical insight and practical experience. It is the one area of corrector design that continues to evolve even after a corrector has been built and successfully brought up.

### 12.3.8 ABERRATIONS OF THE TOTAL SYSTEM

In complete optical systems, optical elements other than the OL or the corrector also contribute to the total sum of aberrations. The contributions from these additional elements depend on the extent of the axial trajectories in them. To weigh the contributions properly, we need to consider the coupling lens arrangement linking optical element  $h$  to the OL, similar to the discussion leading up to Equation 12.18. When the beam radius in the element  $h$  is  $M_h \theta f$ , where  $M_h$  is the magnification factor for element  $h$ , the  $C(h)_{n,m}$  contribution of the element to the total aberration is scaled as  $M_h^{n+1}$ . The aberration  $C_{n,m}$  for the whole optical system is then given by a sum over all the elements contributing to the aberration

$$C_{n,m} = \sum_h C(h)_{n,m} M_h^{(n+1)} \quad (12.32)$$

(Note that we are assuming that the individual aberration contributions are small enough so that combination aberrations can be neglected.)

For round lenses, whose aberrations are conventionally expressed in terms of the angle a ray makes with the optic axis in the back-focal plane of the lens, expressing the phase shift due to the lenses in terms of the angles at the sample involves an additional scaling by  $(f/f_h)^{n+1}$ , where  $f$  is the focal length of the OL and  $f_h$  the focal length of round lens  $h$ . The sum of the aberrations then becomes

$$C_{n,m} = \sum_h C(h)_{n,m} ((f/f_h) M_h)^{(n+1)} \quad (12.33)$$

In the case of spherical aberration  $C_3$ , the contributions from the different round lenses in the column are, therefore, proportional to  $((f/f_h) M_h)^4$ . Because condenser lenses normally have focal lengths that are  $>10\times$  larger than the focal length of the OL, the contributions to the total spherical

aberration of the illuminating column is typically unimportant even when the beam reaches a width in the condensers that is similar to its width in the OL (i.e.,  $M_h \sim 1$ ). Similarly, in the case of the CTEM, only the spherical aberration of the OL typically makes a significant contribution to the spherical aberration of the final image.

An important exception to this rule occurs when a small virtual source such a cold field-emission electron gun (CFEG) tip is imaged onto the sample with hardly any demagnification, to obtain a large beam current. In this case, the beam diameter in the condensers may reach several times the diameter in the OL, and the total  $C_3$  may become dominated by aberrations arising in the source and the condensers, instead of in the OL. A total  $C_3$  tens or hundreds of times larger than that due to OL alone can then easily result. A  $C_3$  corrector in the illumination system can of course correct this kind of aberration too, provided that the first-order trajectories through the system are changed to reflect the new situation.

For chromatic aberration  $C_{1,c}$ , the contributions from the different parts of the imaging system scale as  $((f/f_h)M_h)^2$ . The weaker power of the dependence means that contributions from other stages of the total optical apparatus are more common for  $C_{1,c}$  than for  $C_3$ . Contributions from the corrector and from any condenser lenses in which the beam is wide, therefore, need to be minimized, by keeping the beam width outside of the OL as small as possible. In the corrector, this requires increasing the strength of the higher-order multipole correcting elements, which then allows the beam diameter in the corrector to be smaller for the same correcting effect. In the condenser lenses, it requires avoiding regimes in which a large diameter beam is focused by a strong lens.

## 12.4 ABERRATION DIAGNOSIS

A key development that helped make aberration correction practical was the advent of computer-based autotuning methods that can analyze aberrations quantitatively. Without them, the user of an aberration-corrected system would most likely face an array of many controls that all change a fuzzy image subtly, and no guidance for how to adjust them.

### 12.4.1 DIAGNOSTIC METHODS

The aim of aberration diagnosis is to determine the actual optical state of the apparatus. This amounts to an experimental determination of the aberration coefficients. The usual way to do this is to determine the shape of the aberration function  $\chi(\theta)$  experimentally, and then work out the aberration coefficients that gave rise to it. Several different methods of achieving this are available, each one with advantages and disadvantages.

For TEM applications, diffractogram-based autotuning (Fan and Krivanek, 1990; Krivanek and Fan, 1992a,b; Krivanek, 1994a; Typke and Dierksen, 1995; Uhlemann and Haider, 1998) has become dominant, owing to its good precision at high magnification and immunity to artifacts arising from image drift. The essence of the method is that a power spectrum (= diffractogram) of a phase-contrast TEM image of an amorphous sample can be quantified to determine the apparent defocus and astigmatism of the image. These give the second derivatives of the aberration function  $\partial^2\chi/\partial\theta_x^2$  and  $\partial^2\chi/\partial\theta_y^2$  at the location in the back-focal plane of the OL that the unscattered beam passed through. Recording and analyzing diffractograms for different angles of illumination and analyzing them automatically, therefore, maps the second derivatives of the aberration function  $\chi(\theta)$  over the angular range of the illumination tilts used. Once the derivatives are known, it is a straightforward matter to determine the exact form of  $\chi(\theta)$ , and to work out the strength of the aberration coefficients that gave rise to it.

Another form of autotuning that is applicable to TEM is based on measuring image displacements that result when the illumination is tilted (Koster and de Ruijter, 1992; Krivanek and Fan, 1992b). This determines the first derivatives of the aberration function at the locations of the tilts employed, and thus also the aberration function and the aberration coefficients. However, this

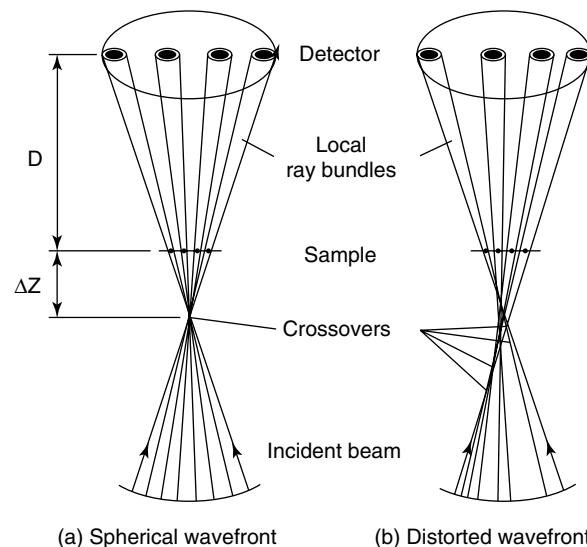
method has not become widely used in aberration correction, probably because sample drift produces spurious image displacements and decreases the accuracy of the method, and because image features do not stay unchanged in an image recorded at close to 1 Å resolution when the beam is tilted. The second problem makes an accurate determination of the image shifts more difficult than analyzing the apparent defocus and astigmatism in a diffractogram.

For STEM, there are even more choices. Exact equivalents of the TEM autotuning methods are obtained simply by recording bright-field (BF) images for different detection angles, which by reciprocity correspond to the TEM illumination angles. A major advantage of the STEM is that with a parallel recording device such as a  $4 \times 4$  PMT array, the needed images can be recorded in parallel. This makes sure that spurious shifts of the sample do not invalidate the measurement. Even with a single-channel recording device, the spurious shifts can be compensated by simultaneously recording a dark-field image, which is not affected by changing the BF detection angle, and using it as a reference that quantifies the shifts.

We implemented both the shift and the diffractogram-based methods for our first STEM corrector, and successfully used them to quantify the aberrations (Krivanek et al., 1997, 1999). Even though both the methods determined the aberration coefficients with sufficient accuracy, they had one major disadvantage: they did not make good use of the signals available for recording in the STEM. Placing a small aperture in the detection plane, as needs to be done to preserve sharp features in a tuning image that is likely to be misadjusted, meant that much of the signal that was potentially available was wasted. A consequence of this was that to achieve adequate signal-to-noise ratios in experimental images used for the tuning, total acquisition times of the order of a minute or more were necessary.

An STEM method utilizing nearly all the transmitted electrons would clearly be preferable, as the data it would produce would be far less noisy. Such a method needs to be based on a Ronchigram (Lin and Cowley, 1986), that is, the shadow image of the sample that is formed in the far field and can be efficiently recorded with a charge coupled device (CCD) camera.

Figure 12.10 shows how Ronchigrams are formed. Figure 12.10a depicts Ronchigram formation on a detector screen a distance  $D$  away from the sample in an optical system free of aberrations other than overfocus. In such a system, the wavefront converging on the sample is spherical, and all the local ray bundles come to a perfect focus at a single height in the column. The bundles



**FIGURE 12.10** The formation of a Ronchigram (= shadow image) in a scanning transmission electron microscope.

illuminate the sample, and propagate to the detector plane. This plane contains an image of the sample recorded with uniform magnification

$$M = (D + \Delta z) / \Delta z \sim D / \Delta z = D / (\partial^2 \chi(\theta) / \partial \theta^2) \quad (12.34)$$

where  $\Delta z$  is the defocus and we have used the facts that  $\Delta z$  is equal to the curvature (second derivative) of the aberration function, and that  $D \gg \Delta z$ . If the defocus is set to zero,  $M$  becomes infinite. All the electrons contributing to the Ronchigram then pass through the same location on the sample, and the Ronchigram becomes devoid of all features.

Figure 12.10b depicts Ronchigram formation in the presence of aberrations. The wavefront is not spherical, and local ray bundles go through crossovers at different heights in the column. The magnification of the shadow image, therefore, becomes position-dependent, and also different in different directions. Determining the magnification experimentally then leads to a map of the local second derivative of the aberration function, and from there to the aberration function itself.

The expression linking the local magnification to the aberration function is given by (Dellby et al., 2001; Krivanek et al., 2003)

$$M_i = D \begin{pmatrix} \partial^2 \chi(\theta) / \partial \theta_x^2 & \partial^2 \chi(\theta) / \partial \theta_x \partial \theta_y \\ \partial^2 \chi(\theta) / \partial \theta_y \partial \theta_x & \partial^2 \chi(\theta) / \partial \theta_y^2 \end{pmatrix}^{-1} \quad (12.35)$$

where  $M_i$  is an element of a discrete matrix  $\mathbf{M}$  that describes how the local magnification varies with position in the Ronchigram, and  $\theta = (\theta_x, \theta_y)$  is the angle that the electron trajectories make to the optic axis at the sample. The derivatives are worked out at positions  $\theta_i = (\theta_{xi}, \theta_{yi})$  in the Ronchigram that correspond to the elements  $M_i$ .

To determine the local magnification experimentally on a general sample whose features are not known a-priori, we use the following stratagem: we move the probe by a small amount in the  $x$  and  $y$  directions, and record a Ronchigram for each probe displacement (Krivanek et al., 2003). Cross-correlating small Ronchigram subareas then shows how much that part of the Ronchigram had shifted because of the probe shift, and hence to the Ronchigram local magnification.

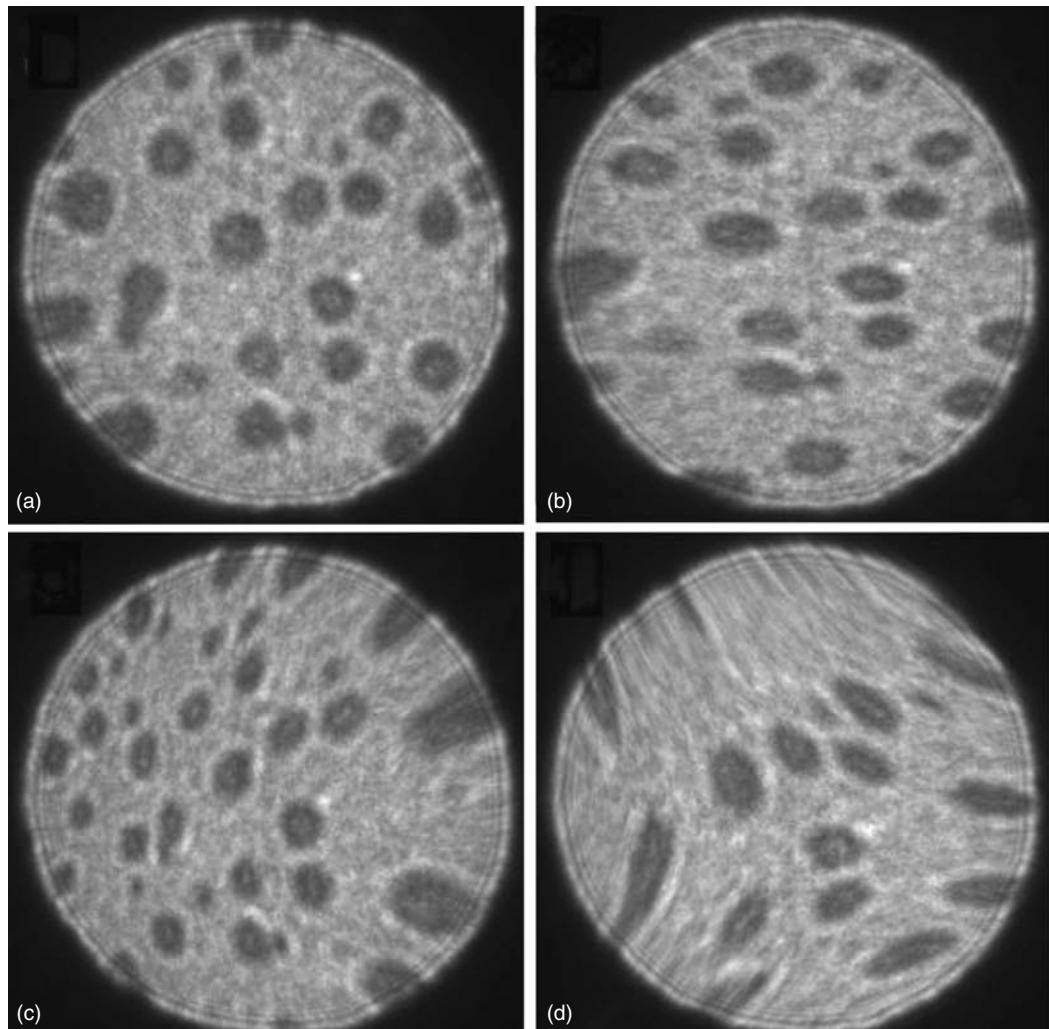
The highest order of the aberrations that can be determined using the Ronchigram-based method is given by the number of regions in the Ronchigram for which the local magnification is determined. For a full characterization of all axial aberrations up to fifth order, we determine  $M_i$  on an array of 57 points in each Ronchigram, and the characterization takes  $\sim 10$  s. For determining aberrations only up to third order or second order, we typically use arrays of 16 and 7 points, respectively, with a corresponding increase in the speed of the characterization.

At the present time, we do not measure aberrations beyond fifth order. This is because their admissible values are quite large (Table 12.1) and they do not greatly affect the performance of our correctors. As the resolution improves further in the future, they will, however, become more important. It will then be a straightforward matter to extend the aforementioned technique to measure them too.

Once the aberration coefficients have been determined, the aberrations can be retuned as required by a computer that instructs the hardware to change the corrector excitations by the appropriate amounts.

Ronchigrams also contain much information that can be assessed visually. A Ronchigram in fact contains a similar amount of information about the aberration function as a tableau of diffractograms recorded for many different angles of detection (illumination for TEM). Whereas the recording and computing of such a tableau would take several seconds even with reasonably powerful cameras and computers, a Ronchigram can be viewed at TV rate even with a probe current of only a few picoamperes.

Experimental Ronchigrams corresponding to particular aberrations are shown in Figure 12.11. A defocus of  $-600$  nm had been applied to produce Ronchigrams in which there was an average



**FIGURE 12.11** Experimental Ronchigrams of the same area of a sample of Au particles on an amorphous carbon film, corresponding to (a) defocus  $C_{1,0} = -600$  nm, (b)  $C_{1,0} = -600$  nm plus twofold astigmatism  $C_{1,2} = 200$  nm, (c)  $C_{1,0} = -600$  nm plus coma  $C_{2,1} = 10 \mu\text{m}$ , and (d)  $C_{1,0} = -600$  nm plus threefold astigmatism  $C_{2,3} = 6 \mu\text{m}$ . 100 keV, Ronchigram field of view =  $\pm 30$  mrad.

noninfinite magnification whose variation can be gauged visually. Areas in which the local Ronchigram magnification is higher correspond to parts of the aberration function where the curvature is less. Unequal magnification in different directions means that the curvature of the aberration function is different in different directions. Twofold astigmatism  $C_{1,2}$ , therefore, gives rise to elliptical Au particles (see Figure 12.11b), most of which were originally fairly round (Figure 12.11a). Coma  $C_{2,1}$  gives rise to a left-right magnification asymmetry as seen in (Figure 12.11c), in which we applied just  $C_{2,1a} \cdot C_{2,1b}$  would have given a top-bottom asymmetry, a mixture of  $C_{2,1a}$  and  $C_{2,1b}$  would have given a general asymmetry between one side of the Ronchigram and the opposite side. Threefold astigmatism  $C_{2,3}$  produces a threefold Ronchigram that can be reminiscent of the *star of Mercedes* when there are no other significant aberrations present except  $C_{2,3}$  and  $C_1$ , and the ratio between them is roughly 1:2.

Note that in a single diffractogram of a high-resolution image, even-power aberrations such as axial coma and threefold astigmatism do not leave a readily identifiable imprint. That is why

several diffractograms recorded for different beam tilts or whole diffractogram tableaus must be used to identify and measure these aberrations (e.g., Krivanek, 1994b).

When recording a high-resolution STEM image, the defocus is set close to zero and the center of the Ronchigram then contains a large featureless patch. Rays traversing the featureless patch are all brought to focus in the same place on the sample, and thus contribute to the formation of a sharp probe. Rays corresponding to regions showing varying contrast in the Ronchigram traverse the sample away from the probe maximum and, therefore, only contribute to the probe tail. The featureless patch denotes the part of the OL front-focal plane that should be used for imaging, that is, that should be selected by a probe-defining aperture placed either in the front-focal plane of the OL or upstream from it. The larger the extent of the patch, the better the resolution will be, and the higher the probe current. By analogy with terminology established for tennis rackets, in which the sub-area of a racket that will steer the ball impinging on it in the selected direction is called the sweet spot, we like to refer to the featureless patch as the Ronchigram sweet spot. Making the sweet spot as large as possible is the essence of STEM aberration correction.

For noncrystalline samples thicker than a few nanometers, different depths of the sample are imaged at different defocus values, and  $C_1 = 0$  is not valid throughout the sample. The diffractogram is then never truly featureless. Instead, the sweet spot contains large features of similar magnification, which shift across it at constant speeds when the probe is slowly moved over the sample.

#### 12.4.2 COMPUTER CONTROL

A major problem with optical systems that provide complete control over parasitic aberrations is that the effects of their various controls are almost never *pure*: each optical element changes several aberrations at the same time, and most aberrations are affected by more than one element. In modern aberration correctors, this problem is solved by computer control: the various elements are linked together in the proportions required to produce a pure(ish) effect for each control, which makes the adjustment of the corrector tractable.

The mixture of element excitations that correspond to a pure effect is typically worked out experimentally. It involves setting various controls that affect a particular aberration (plus a mixture of other aberrations) to different excitations, and determining the resultant aberrations for each separate optical state. Provided that there are more controls than the number of aberrations affected and that the different controls produce different effects, the experimentally determined matrix linking the excitations of the optical elements to the aberrations can be inverted, resulting in a matrix linking each aberration coefficient to a particular set of changes in the excitations of the optical elements. The resultant controls are then used by the computer to adjust the optical system each time the aberration coefficients have been measured experimentally. The controls are also made available to experienced users, who use them to adjust the optical system as needed based on their assessment of freshly obtained Ronchigrams or even recorded images.

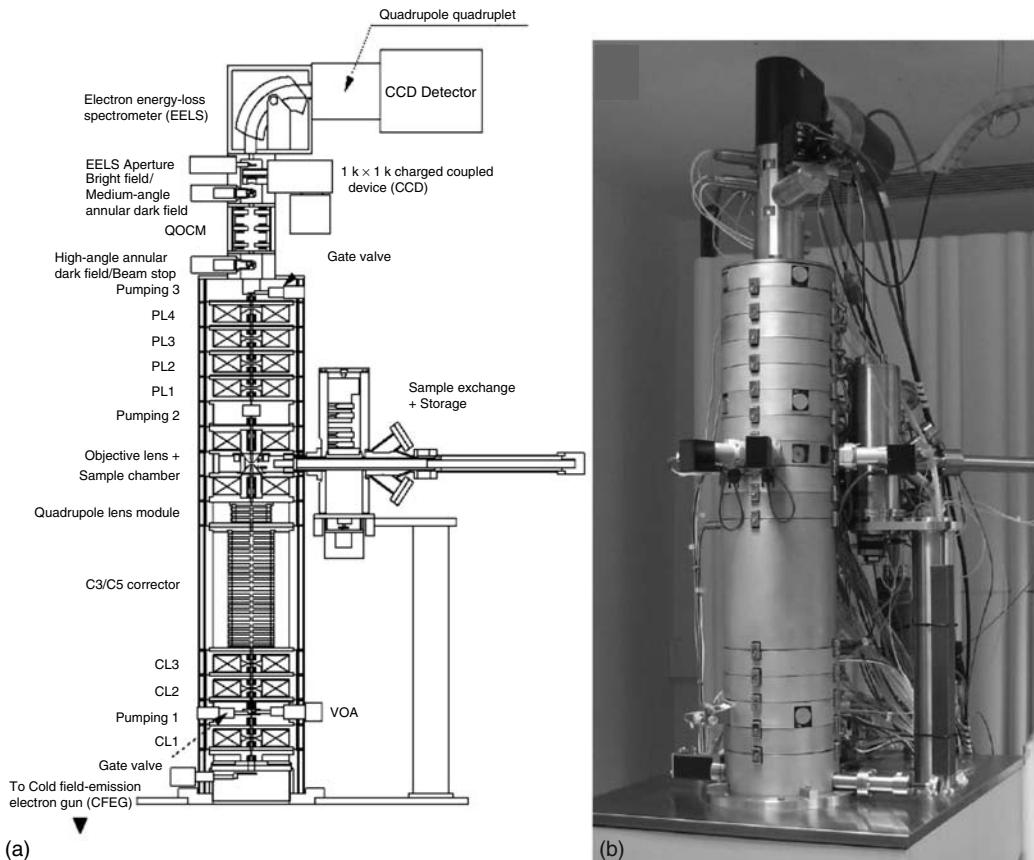
### 12.5 ABERRATION-CORRECTED OPTICAL COLUMN

In this section, we describe an optical system whose design was based on the principles described in Sections 12.1 through 12.4. The example is drawn from our own work.

#### 12.5.1 DESCRIPTION OF THE COLUMN

An aberration corrector does not work in isolation: its function is to correct the optical defects of other optical elements. Figure 12.12a shows a cross section through the optical column of a STEM designed specifically to work with the  $C_3/C_5$  corrector described in Section 12.3. Figure 12.12b shows an overall picture of the system.

The electron source of the microscope is a CFEG, which provides excellent brightness ( $B > 10^9$  A/(cm<sup>2</sup> sr)) and good energy spread ( $\Delta E \sim 0.3$  eV). The four round condenser lenses of the column



**FIGURE 12.12** (See color insert following page 340.) Scanning transmission electron microscope column that includes the corrector of Figure 12.9. (a) Schematic cross section and (b) the actual column. Column diameter = 280 mm.

(one condenser is mounted in the electron gun and is not shown in the schematic) allow the source demagnification to be adjusted as needed, allow the beam's angular convergence to be changed, and allow the height of the crossover that the beam entering the corrector appears to emanate from to be set to the value needed by the corrector. This section of the column also includes an electrostatic beam blanker, which can turn the beam on the sample off (and on) in a few microseconds, which is useful for preventing radiation damage when no data is being acquired. It also includes a beam-defining aperture and a precorrector set of scan coils. Placing the scan coils before the corrector means that the beam is scanned in the entire corrector-OL assembly, which makes sure that the scanning is not affected by uncorrected aberrations of the OL. This is particularly important for beam rocking, as needed for precession electron diffraction (Own et al., 2007).

The condensers are optimized for low aberrations but they nevertheless have nonzero spherical and chromatic aberrations. Under conditions of high source demagnification (= best possible resolution) these make a negligible contribution to the total aberration budget. But when the source demagnification is less, as for instance when beam currents of the order of 0.5 nA and higher are needed for rapid spectroscopic analysis, the aberrations of the condensers can no longer be neglected. Their spherical aberration contribution is then canceled by the corrector, along with the contribution from the OL. This is because the aberration diagnosing software makes no distinction as to where the aberrations came from, and simply monitors the coefficients of the total system.

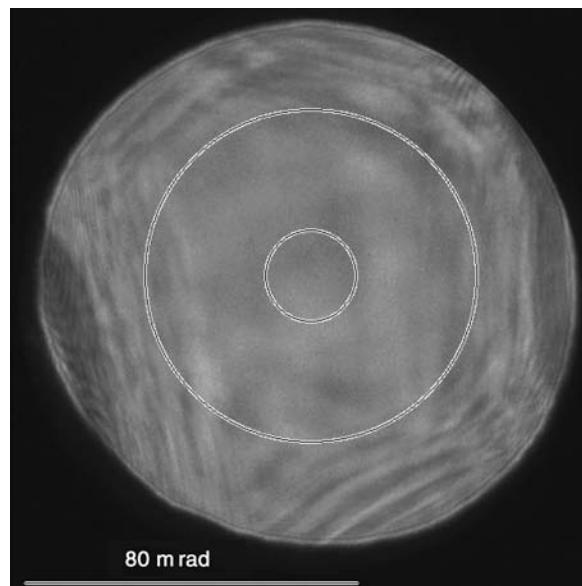
The corrector is placed between the condenser section of the microscope and a coupling lens arrangement whose task is to image the corrector into the coma-free plane of the OL. The coupling arrangement consists of four quadrupole lenses, three of which are located in a separate quadrupole lens module (QLM) between the corrector and the OL. The fourth one is built into the corrector.

The OL produces the small electron probe incident on the sample. It is of the condenser-objective type, with the final pole-piece shape optimized for low chromatic aberration (which is not corrected in this system) rather than low spherical aberration (which is corrected). The sample is held by a five-axis ( $x$ ,  $y$ ,  $z$ ,  $\alpha$ ,  $\beta$ ) goniometer near the center of the lens's pole-piece gap. The goniometer is stable, precise, and responsive: it allows the sample to be moved mechanically by as little as 1 nm in  $x$  and  $y$ , and 5 nm in  $z$ . Focusing and sample area selection are therefore mostly done by moving the sample mechanically, with electrical adjustments of the beam reserved for the last few nanometer of each displacement. This arrangement allows the electron-optical setup to remain essentially unchanged no matter which area of the sample is being looked at.

The OL is followed by several coupling (projector) lenses, a quadrupole–octupole triplet that acts as a pre-electron energy-loss spectrometer (pre-EELS) aberration corrector, and a variety of detectors that include HAADF, medium-angle annular dark field (MAADF) and BF STEM detectors, a 1 k  $\times$  1 k fast read-out CCD camera for Ronchigram recording, and a parallel-detection EELS. There is also a clean and bakeable vacuum system that attains pressures in the  $10^{-11}$  torr region in the gun and in the  $10^{-9}$  torr region in the rest of the column, an extensive set of electronics, and custom software for running the microscope. The whole system is described in detail in Krivanek et al. (2007).

### 12.5.2 PERFORMANCE OF THE SYSTEM

The electron-optical success of STEM aberration correction is best demonstrated by the size of the sweet spot the corrected system can attain in an experimental Ronchigram, or by the size of the experimentally measured aberration coefficients. Figure 12.13 shows a Ronchigram obtained



**FIGURE 12.13** Experimental Ronchigram obtained with the Nion C<sub>3</sub>/C<sub>5</sub> corrector. The optimum aperture corresponding to the size of the sweet spot (40 mrad half-angle) as well as the optimum aperture in an equivalent uncorrected microscope (the small circle corresponding to 11.6 mrad half-angle) are indicated by the large and small circles, respectively.

**TABLE 12.2**  
**Measured Aberration Coefficients**  
**for the Ronchigram of Figure 12.13**

Coefficient	Value (nm)
$C_{2,1,a}$	$1.2 \times 10^2$
$C_{2,1,b}$	$5.0 \times 10^1$
$C_{2,3,a}$	-2.9
$C_{2,3,b}$	$-1.0 \times 10^2$
$C_{3,0}$	$5.7 \times 10^3$
$C_{3,2,a}$	$-4.0 \times 10^3$
$C_{3,2,b}$	$-8.5 \times 10^3$
$C_{3,4,a}$	$-4.5 \times 10^2$
$C_{3,4,b}$	$1.7 \times 10^3$
$C_{4,1,a}$	$-1.7 \times 10^5$
$C_{4,1,b}$	$-7.2 \times 10^4$
$C_{4,3,a}$	$6.0 \times 10^3$
$C_{4,3,b}$	$2.1 \times 10^5$
$C_{4,5,a}$	$2.2 \times 10^4$
$C_{4,5,b}$	$6.4 \times 10^4$
$C_{5,0}$	$-4.1 \times 10^6$
$C_{5,2,a}$	$2.8 \times 10^6$
$C_{5,2,b}$	$8.2 \times 10^6$
$C_{5,4,a}$	$-2.1 \times 10^6$
$C_{5,4,b}$	$9.9 \times 10^4$
$C_{5,6,a}$	$4.4 \times 10^5$
$C_{5,6,b}$	$2.4 \times 10^5$

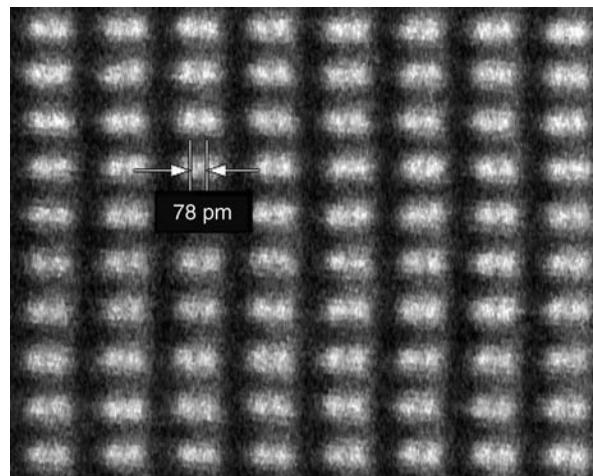
with the 100 kV STEM described in the previous section. The radius of the sweet spot is 40 mrad. Experimental aberration coefficients measured a short time before the Ronchigram was recorded are shown in Table 12.2. The table omits the first-order coefficients ( $C_{1,0}$  and  $C_{1,2}$ ), which were changed by the operator just prior to the recording of the Ronchigram. Comparing the measurements of Table 12.2 with the maximum permissible values of Table 12.1 shows that the coefficients were more than adequate for 1 Å probe size at 100 kV.  $C_{2,3}$  appears to be the aberration that was the most limiting, but its value was actually chosen to oppose non-zero  $C_{4,3}$ .

The spherical aberration coefficient  $C_3$  of the microscope's OL was 1.0 mm. Without an aberration corrector, the largest sweet spot the microscope could attain would correspond to a half-angle  $\theta_{\max}$  of

$$\theta_{\max} = 1.5(\lambda/C_3)^{1/4} = 11.6 \text{ mrad} \quad (12.36)$$

In this particular instance, the action of the corrector therefore increased  $\theta_{\max}$  3.5×.

When setting the microscope up for the best possible resolution, the full extent of the sweet spot may not be usable due to the effects of chromatic aberration. This limitation is discussed in detail by Krivanek et al. (2008). At 100 kV with a CFEG, the extent of the usable sweet spot in an optical system with total  $C_c$  of 1.2 mm (as is the case with the Nion column) is ~35 mrad half-angle when aiming for the best resolution, and slightly more when aiming for a large current in a probe. At 200 kV, the maximum usable half-angle for the same  $C_c$  is >40 mrad.



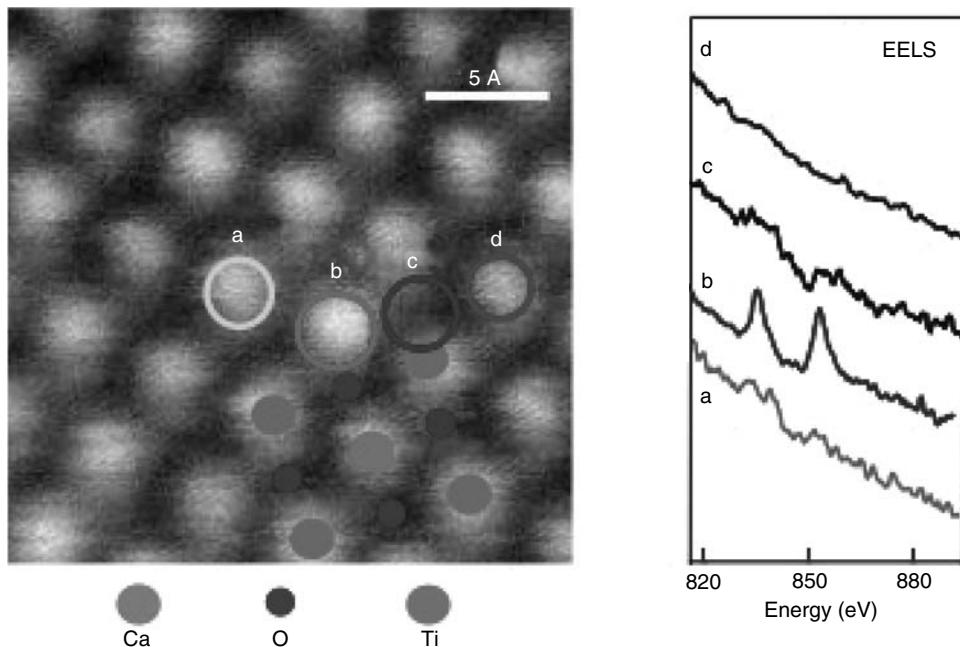
**FIGURE 12.14** High-angle annular dark-field image of silicon recorded in the  $\langle 112 \rangle$  direction. Atoms 0.78 Å apart in the projection are resolved. VG HB603 STEM, Nion aberration corrector, 300 keV. (Courtesy Drs P.D. Nellist, M.F. Chisholm, and S.J. Pennycook, permission Science.)

Increasing the illumination angle in the STEM results in two very important practical improvements.

1. When no other changes are made to the optical system, the probe current increases as  $\theta_{\max}^2$ . The resolution improves slightly at the same time, because the diffraction limit contribution to the probe size becomes less.
2. When the magnification of the virtual source imaged by the OL is decreased in proportion with the increase of the illumination angle, the probe size decreases as  $\theta_{\max}^{-1}$ . This presupposes that the illumination angle does not exceed the size of the sweet spot, and that instabilities of the system are much smaller than the resolution. The probe current remains the same as it was before.

Figures 12.14 through 12.16 show practical results from different materials using three different aberration-corrected systems. Figure 12.14 shows an HAADF image of a single crystal of silicon projected in the  $\langle 112 \rangle$  direction (Nellist et al., 2004). The image was recorded in a 300 kV VG HB603 STEM equipped with a second-generation Nion C<sub>s</sub> corrector. Si atoms just 0.78 Å apart in the projection are clearly resolved in the image. A power spectrum of the image showed that spacings down to 0.6 Å have been recorded, that is, that the resolution of this instrument was substantially better than 1 Å.

Figure 12.15 shows that with the electron probe size and current made possible by aberration correction, EELS can now be recorded from single atoms (Varela et al., 2004). The left side of Figure 12.15 shows a HAADF image of calcium titanate (CaTiO<sub>3</sub>) doped with La. The atoms are arranged in columns as marked by the filled disks. The empty circles show locations from which EEL spectra were recorded, and the spectra themselves are shown on the right side of the figure. It is expected that La (Z = 57) dopant atoms will substitute for the lighter Ca (Z = 20). The empty circle marked b is positioned over a Ca column whose image is brighter than the images of other similar columns. The corresponding spectrum (b) shows very strong La M<sub>4,5</sub> edge threshold lines. Spectra recorded just one column (~4 Å) away (a and c) show a small hint of these lines; a spectrum recorded two columns away shows none. The concentration of La was so low that it is nearly certain that this was a single La atom. The images and spectra were obtained in a 100 kV VG equipped with a Nion second-generation C<sub>s</sub> corrector plus a parallel-detection EELS (Varela et al., 2004).



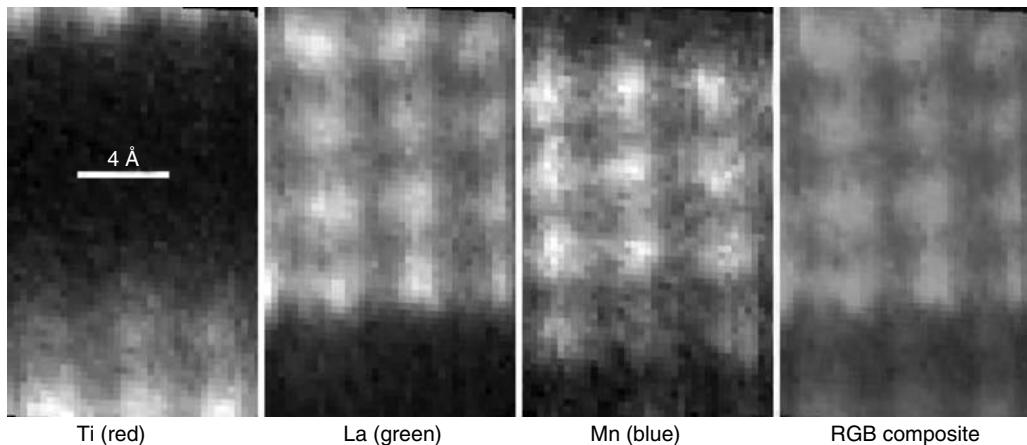
**FIGURE 12.15** (See color insert following page 340.) (a) High-angle annular dark-field image and electron energy-loss spectra from  $\text{CaTiO}_3$  doped with La. The spectrum (b) originated from a single atom of La. VG HB501 STEM, Nion aberration corrector, 100 keV. (Courtesy Drs M. Varela and S.J. Pennycook, permission Physical Review Letters.)

The coupling of the EELS to the OL of the microscope was far from perfect in this work: only  $\sim 8\%$  of the signal available within the BF disk was admitted into the EEL spectrometer. It can be expected that with better-optimized instruments that have since become available, single atoms of all atomic species that have nondelayed edges and preferably also strong white threshold lines at energy losses between  $\sim 100$  and  $1000$  eV (i.e., the first row transition metals, lanthanides, actinides, and possibly Be to F and Na to P), will be detectable by EELS, provided of course that radiation damage is not too severe. As the instruments improve further, single atom detection by EELS is likely to become applicable to almost all the elements in the periodic table.

Figure 12.16 shows an example of a two-dimensional chemical map obtained with the  $\text{C}_3/\text{C}_5$  corrected system of Figure 12.12. The sample was a  $\text{SrTiO}_3\text{-La}_{0.7}\text{Sr}_{0.3}\text{MnO}_3$  multilayer structure grown by pulsed laser deposition. The HAADF intensity and a 650 eV-wide EEL spectrum were recorded at each pixel in a  $64 \times 64$  pixel spectrum-image (Jeanguillaume and Colliex, 1989), 60% of which is shown in the figure. EELS L-edges of Ti and Mn, and the M-edge of La were background-subtracted and quantified in terms of the atomic concentrations at every pixel using principal component analysis.

The resultant chemical maps of the Ti, La, and Mn concentrations are shown in the figure, in black-and-white. The right side of the figure displays an RGB color composite obtained by combining the partial images (red = Ti, green = La, and blue = Mn). The green and blue dots that are clearly visible in the color image show that La and Mn occupy different columns in the  $\text{La}_{0.7}\text{Sr}_{0.3}\text{MnO}_3$  layers (top 3/4 of the image). The red dots visible at the bottom of the image show that Ti in the  $\text{SrTiO}_3$  layers occupies columns that are equivalent to the columns occupied by Mn in the  $\text{La}_{0.7}\text{Sr}_{0.3}\text{MnO}_3$  layers. Very interestingly, three purple dots are visible at the interface between the two layers. They show that at this particular interface, there was a mixing of Ti and Mn within individual atomic columns.

The data was acquired with the aberration-corrected Nion STEM column described in this chapter, equipped with a parallel-detection EELS (Gatan Enfina). The EELS collection angle was



**FIGURE 12.16** (See color insert following page 340.) Electron energy-loss spectroscopy spectrum-images of Ti, La, and Mn in a  $\text{SrTiO}_3\text{-La}_{0.7}\text{Sr}_{0.3}\text{MnO}_3$  multilayer, plus a combined false color image showing the locations of various atomic columns in the structure. Nion column with  $C_3/C_5$  aberration corrector, 100 keV. (Courtesy Prof D.A. Muller, L.F. Koukoutis, and M.F. Murfitt, multilayer structure courtesy Drs H.Y. Hwang and J.H. Song.)

~50 mrad, which meant that >70% of the available EELS signal was coupled into the spectrometer. The per-pixel acquisition time was 14 ms, and the live collection time for the data set was therefore 57 s (~2 min with per-pixel processing). There was a sample drift of ~1 nm during the acquisition. The resultant image distortion was removed by an unwarping algorithm. Radiation damage was present, but it was not severe enough to prevent the experiment. Similar data that was acquired with 7 ms per pixel in a subsequent run during the same experimental session and processed differently has been shown by Muller et al. (2008).

Because of the advanced corrector and the small energy spread of the CFEG (~0.3 eV), an illumination half-angle of 40 mrad could be used without a marked loss of resolution due to chromatic effects. In other words, the illumination aperture was 3.5× larger (in angle, 12× in area) than the largest aperture useable in an equivalent uncorrected instrument. It allowed the current in the probe to be 0.7 nA, even though the probe still contained spatial frequencies of ~1 Å (see Muller et al., 2008). The large current in a probe roughly 1.3 Å in size plus efficient EELS coupling are the chief factors that made the experiment possible in a relatively short acquisition time.

Atomic-resolution elemental maps have recently also been obtained by others, but with considerably longer acquisition times. For instance, Bosman et al. (2007) have used a VG 100 kV STEM with a Nion second-generation ( $C_3$ -only) aberration corrector to obtain atomic-resolution EELS chemical maps of O and Mn in several projections of  $\text{Bi}_{0.5}\text{Sr}_{0.5}\text{MnO}_3$ , with an acquisition time of 0.2 s per pixel.

Results such as those presented in Figures 12.15 and 12.16 show that EEL spectroscopy and spectrum-imaging in an aberration-corrected STEM have now joined atom-probe microscopy as techniques able to detect single atoms of particular chemical types, and to map the chemical composition of solid samples at atomic resolution. Unlike atom-probe microscopy, EELS-STEM is also able to analyze the local bonding and electronic states. It is likely to become preeminent among the techniques available for exploring the chemical composition and other properties of the sample at the nanoscale and beyond.

## 12.6 CONCLUSIONS

Aberration correction in electron microscopy stands apart from many other scientific developments by how long it took: more than 50 years from the first ideas to truly useful implementations. But the wait was a worthwhile one, as the results from aberration-corrected instruments show.

Further exciting developments may be expected especially in the domains that have not yet seen as much effort devoted to aberration correction as transmission electron microscopy: SEMs, and ion microscopes and microanalyzers. In both these cases, correction of chromatic aberration will be as important as correction of geometric aberrations.

The developments that have already taken place and those that are currently being implemented are advancing electron microscope optics to a new state of perfection. The ultimate goal is that the amount of information that can be extracted from each particular sample should be limited only by the nature of the beam-sample interaction, and not by the optics of the microscope. Operating parameters such as the electron exposure, illumination and collection angles, and so on, will then be determined by the properties of the sample rather than the properties of the microscope. The scientists using such a microscope will be secure in knowing that whatever the investigation ahead, the instrument will be optimally suited for it.

## ACKNOWLEDGMENTS

We are grateful to our many colleagues for encouragement, financial support, and numerous discussions, and to George Corbin and Dr. Chris Own for help with the preparation of the figures.

## REFERENCES

- Archard G.D. 1954. *Br. J. Appl. Phys.*, **5**, 294.  
 Batson P.E., Dellby N. and Krivanek O.L. 2002. *Nature*, **418**, 617.  
 Beck V. 1977. *Proc. 33rd EMSA Meeting*, 90.  
 Beck V. 1979. *Optik*, **53**, 241.  
 Beck V. and Crewe A.V. 1976. *Proc. 32nd EMSA Meeting*, 578.  
 Berz M. 1990. *Nucl. Instrum. Meth.*, **A298**, 473.  
 Born M. and Wolf E. 1980. *Principles of Optics* (Pergamon Press, Oxford).  
 Bosman M., Keast V.J., Garcia-Munoz J.L., D'Alfonso A.J., Findlay S.D. and Allen L.J. 2007. *Phys. Rev. Lett.*, **99**, 086102.  
 Carey D.C. 2006. Third order TRANSPORT with MAD input, <http://www.slac.stanford.edu/pubs/slacreports/slacr-530.html>.  
 Castaing R. and Henry L. 1962. *C.R. Acad. Sci. (Paris)*, **B255**, 76.  
 Courant E.D., Livingston M.S. and Snyder H.S. 1952. *Phys. Rev.*, **88**, 1168.  
 Cowley J.M. 1969. *Appl. Phys. Lett.*, **15**, 58.  
 Crewe A.V. 1982. *Optik*, **60**, 271.  
 Crewe A.V. 1984. *Optik*, **69**, 24.  
 Crewe A.V., Isaacson M. and Johnson D. 1969. *Rev. Sci. Instrum.*, **40**, 241.  
 Crewe A.V. and Kopf D. 1980. *Optik*, **55**, 1.  
 Crewe A.V., Wall J. and Langmore J. 1970. *Science*, **168**, 1338.  
 Crewe A.V., Wall J. and Welter L.M. 1968. *J. Appl. Phys.*, **39**, 5861.  
 Dellby N., Krivanek O.L. and Murfitt M.F. 2006. *CPO7 Proc.*, p. 97, available at: <http://www.mebs.co.uk/CPO7.htm> and also Physics Procedia in print.  
 Dellby N., Krivanek O.L., Nellist P.D., Batson P.E. and Lupini A.R. 2001. *J. Electron Microsc.*, **50**, 177.  
 Deltrap J.H.M. 1964a. Ph. D. dissertation, University of Cambridge.  
 Deltrap J.H.M. 1964b. *Proc. 3rd Eur. EM Congr.* (Prague), 45.  
 Fan G.Y. and Krivanek O.L. 1990. *Proc. 12th Int. EM Congr.* (Seattle), **1**, 532.  
 Freitag B., Stekelenburg M., Ringnalda J. and Hubert D. 2007. *Proc. Microsc. Microanal. Meeting*, Ft Lauderdale, *Microsc. Microanal.*, **13**(Suppl. 2), 1162CD.  
 Grivet P. 1972. *Electron Optics, Part I* (Pergamon Press, Oxford and New York).  
 Haider H., Loebau U., Hoeschen R., Müller H., Uhlemann S. and Zach J. 2007. *Proc. Microsc. Microanal. Meeting*, Ft. Lauderdale, *Microsc. Microanal.*, **13**(Suppl. 2), 1156CD.  
 Haider M., Braunshausen G. and Schwan E. 1995. *Optik*, **99**, 167.  
 Haider M., Rose H., Uhlemann S., Kabius B. and Urban K. 1998. *J. Electron Microsc.*, **47**, 395.  
 Haider M., Uhlemann S., Schwann E., Kabius B. and Urban K. 1998. *Nature*, **392**, 768.

- Hardy D.F. 1967. Ph. D. dissertation, University of Cambridge.
- Hawkes P.W. 1965. *Phil. Trans. R. Soc. (London)*, **A257**, 479.
- Hawkes P.W. 2007. In: *Science of Microscopy* (Hawkes P.W. and Spence J.C.H., eds, Springer, New York), p. 696.
- Hawkes P.W. and Kasper E. 1996. *Principles of Electron Optics* (Academic Press, New York), Vol. 2, pp. 857–863.
- Hosokawa F., Sawada H., Sannomiya T., Kaneyama T., Kondo Y., Hori M., Yuasa S., Kawazoe M., Nakamichi T., Tanishiro Y., Yamamoto N. and Takayanagi K. 2006. *Proc. 16th Int. Microsc. Congr.* (Ichinose H. and Sasaki T., eds, Sapporo, Japan), 582.
- Jeanguillaume C. and Colliex C. 1989. *Ultramicroscopy*, **78**, 252.
- Jia C.L. and Urban K. 2004. *Science*, **303**, 2001.
- Koops H. 1978. *Proc. Int. EM Congr.*, Toronto, **3**, 185.
- Koster A.J. and de Ruijter W.J. 1992. *Ultramicroscopy*, **40**, 89.
- Krivanek O.L. 1994a. *Ultramicroscopy*, **55**, 419.
- Krivanek O.L. 1994b. U.S. patent #5,300,776.
- Krivanek O.L., Corbin G.J., Dellby N., Elston B.F., Keyse R.J., Murfitt M.F., Own C.S., Szilagyi Z.S. and Woodruff J.W. 2008. *Ultramicroscopy*, **108**, 179.
- Krivanek O.L., Dellby N. and Lupini A.R. 1999. *Ultramicroscopy*, **7**, 1.
- Krivanek O.L., Dellby N. and Lupini A.R. 2003. U.S. patent #6,552, 340.
- Krivanek O.L., Dellby N. and Nellist P.D. 2004. U.S. patent #6,770, 887.
- Krivanek O.L., Dellby N., Spence A.J., Camps R.A. and Brown L.M. 1997. In: *Inst. Phys. Conf. Ser.* 153 (*Proc. 1997 EMAG Meeting*, Rodenburg J.M., ed.), 35.
- Krivanek O.L. and Fan G.Y. 1992a. *Proc. 50th EMSA Meeting*, 96.
- Krivanek O.L. and Fan G.Y. 1992b. *Scanning Microsc.*, Suppl. 6, 105.
- Krivanek O.L., Gubbens A.J. and Dellby N. 1991 *Microsc. Microanal. Microst.*, **2**, 315.
- Krivanek O.L., Nellist P.D., Dellby N., Murfitt M.F. and Szilagyi Z. 2003. *Ultramicroscopy*, **96**, 229.
- Lentzen M. 2006. *Microsc. Microanal.*, **12**, 191.
- Lenz F.A. 1971. In: *Electron Microscopy in Materials Science* (Valdre U., ed., Academic Press, New York and London), p. 541.
- Lin J.A. and Cowley J.M. 1986. *Ultramicroscopy*, **19**, 31.
- Makino K. and Berz M. 1999. *Nucl. Instrum. Meth.*, **A427**, 338.
- Mitsuishi K., Takeguchi M., Kondo Y., Hosokawa F., Okamoto K., Sannomiya T., Hori M., Iwama T., Kawazoe M. and Furuya K. 2006. *Microsc. Microanal.*, **12**, 456.
- Möllenstedt G. 1956. *Optik*, **13**, 2009.
- Muller D.A., Fitting-Kourkoutis L., Murfitt M.F., Song J.H., Hwang H.Y., Silcox J., Dellby N. and Krivanek O.L. 2008. *Science*, **319**, 1073.
- Müller H., Uhlemann S., Hartel P. and Haider M. 2006. *Microsc. Microanal.*, **12**, 442.
- Munro E. 1997. In: *Handbook of Charged Particle Optics* (1st edition) (Orloff J., Ed., CRC Press, Boca Raton and New York), p. 1.
- Murfitt M.F., Dellby N. and Krivanek O.L. 2005. Unpublished results.
- Nellist P.D., Chisholm M.F., Dellby N., Krivanek O.L., Murfitt M.F., Szilagyi Z.S., Borisevich A., Sides W.H. Jr. and Pennycook S.J. 2004. *Science*, **305**, 1741.
- Own C.S., Dellby N., Krivanek O.L., Marks L.D. and Murfitt M.F. 2007. *Proc. Microsc. Microanal. Meeting*, Ft. Lauderdale, *Microsc. Microanal.*, **13**(Suppl. 2), 96.
- Preikszas D. and Rose H. 1997. *J. Electron Microsc.*, **1**, 1.
- Rose H. 1970. *Optik*, **32**, 144.
- Rose H. 1971. *Optik*, **33**, 1.
- Rose H. 1990. *Optik*, **85**, 19.
- Rose H. 2003. In: *High-Resolution Imaging and Spectrometry of Materials* (Ernst E. and Rühle M., eds, Springer, Berlin), p. 189.
- Rose H. 2004. *Nucl. Instrum. Meth.*, **519**, 12.
- Rose H. 2005. *Ultramicroscopy*, **103**, 1.
- Rose H. and Krahl D. 1995. In: *Energy-Filtering Transmission Electron Microscopy* (Reimer L., ed., Springer, Berlin), 43.
- Rose H. and Pejas W. 1979. *Optik*, **54**, 235.
- Sawada H., Tomita T., Naruse M., Honda T., Hambridge P., Hartel P., Haider M., Hetherington C., Doole R., Angus K.A., Hutchison J., Titchmarsh J. and Cockayne D. 2005. *J. Electron Microsc.*, **54**, 119.

- Saxton W.O. 2000. *Ultramicroscopy*, **81**, 41.
- Scherzer O. 1936. *Z. Phys.*, **101**, 593.
- Scherzer O. 1947. *Optik*, **2**, 114.
- Schmidt T. et al. 2003. *Microsc. Microanal.*, **9**, 122.
- Seeliger R. 1953. *Optik*, **10**, 29.
- Senoussi S., Henry L. and Castaing R. 1971. *J. Microsc.*, **11**, 19.
- Shao Z. 1988. *Rev. Sci. Instrum.*, **59**, 2429.
- Shao Z., Beck V. and Crewe A.V. 1988. *J. Appl. Phys.*, **64**, 1646.
- Shuman H. 1980. *Ultramicroscopy*, **5**, 45.
- Thomson M.G.R. 1968. Ph.D. dissertation, University of Cambridge.
- Typke D. and Dierksen K. 1995. *Optik*, **99**, 155.
- Uhlemann S. and Haider M. 1998. *Ultramicroscopy*, **72**, 109.
- Varela M. et al. 2004. *Phys. Rev. Lett.*, **92**, 095502.
- Weißbäcker Ch. and Rose H. 2001. *J. Electron Microsc.*, **50**, 383.
- Wilson E.J.N. 2006. *An Introduction to Particle Accelerators* (Oxford University Press, Oxford and New York).
- Wittry D.B. 1969. *Br. J. Appl. Phys. (J. Phys. D)*, **2**, 1757.
- Yavor S. Ya., Dymnikov A.D. and Ovsyannikova L.P. 1964. *Nucl. Instrum. Meth.*, **26**, 13.
- Zach J. and Haider M. 1995. *Optik*, **99**, 112.

## A.1 APPENDIX

Aberration function expansion up to fifth order.

$$\begin{aligned}
 \chi(\theta) = & C_{0,1,a} \theta_x + C_{0,1,b} \theta_y \\
 & + C_{1,0} (\theta_x^2 + \theta_y^2)/2 \\
 & + C_{1,2,a} (\theta_x^2 - \theta_y^2)/2 + C_{1,2,b} \theta_x \theta_y \\
 & + C_{2,1,a} (\theta_x^3 + \theta_x \theta_y^2)/3 + C_{2,1,b} (\theta_x^2 \theta_y + \theta_y^3)/3 \\
 & + C_{2,3,a} (\theta_x^3 - 3\theta_x \theta_y^2)/3 + C_{2,3,b} (3\theta_x^2 \theta_y - \theta_y^3)/3 \\
 & + C_{3,0} (\theta_x^4 + 2\theta_x^2 \theta_y^2 + \theta_y^4)/4 + C_{3,2,a} (\theta_x^4 - \theta_y^4)/4 + C_{3,2,b} (\theta_x^3 \theta_y + \theta_x \theta_y^3)/2 \\
 & + C_{3,4,a} (\theta_x^4 - 6\theta_x^2 \theta_y^2 + \theta_y^4)/4 + C_{3,4,b} (\theta_x^3 \theta_y - \theta_x \theta_y^3) \\
 & + C_{4,1,a} (\theta_x^5 + 2\theta_x^3 \theta_y^2 + \theta_x \theta_y^4)/5 + C_{4,1,b} (\theta_x^4 \theta_y + 2\theta_x^2 \theta_y^3 + \theta_y^5)/5 \\
 & + C_{4,3,a} (\theta_x^5 - 2\theta_x^3 \theta_y^2 - 3\theta_x \theta_y^4)/5 + C_{4,3,b} (3\theta_x^4 \theta_y + 2\theta_x^2 \theta_y^3 - \theta_y^5)/5 \\
 & + C_{4,5,a} (\theta_x^5 - 10\theta_x^3 \theta_y^2 + 5\theta_x \theta_y^4)/5 + C_{4,5,b} (5\theta_x^4 \theta_y - 10\theta_x^2 \theta_y^3 + \theta_y^5)/5 \\
 & + C_{5,0} (\theta_x^6 + 3\theta_x^4 \theta_y^2 + 3\theta_x^2 \theta_y^4 + \theta_y^6)/6 \\
 & + C_{5,2,a} (\theta_x^6 + \theta_x^4 \theta_y^2 - \theta_x^2 \theta_y^4 - \theta_y^6)/6 + C_{5,2,b} (2\theta_x^5 \theta_y + 4\theta_x^3 \theta_y^3 + 2\theta_x \theta_y^5)/6 \\
 & + C_{5,4,a} (\theta_x^6 - 5\theta_x^4 \theta_y^2 - 5\theta_x^2 \theta_y^4 + \theta_y^6)/6 + C_{5,4,b} (2\theta_x^5 \theta_y - 2\theta_x \theta_y^5)/3 \\
 & + C_{5,6,a} (\theta_x^6 - 15\theta_x^4 \theta_y^2 + 15\theta_x^2 \theta_y^4 \theta_y^4 - \theta_y^6)/6 + C_{5,6,b} (3\theta_x^5 \theta_y - 10\theta_x^3 \theta_y^3 + 3\theta_x \theta_y^5)/3
 \end{aligned}$$

---

# Appendix: Computational Resources for Electron Microscopy

*J. Orloff (with valuable information from  
Peter W. Hawkes and Bohumila Lencová)*

## CONTENTS

Further Reading .....	643
Books .....	643
Papers and Conference Proceedings .....	643
Additional References .....	644

To design a charged particle instrument it is necessary to be able to make an accurate prediction of its properties. This generally means predicting the properties of the lenses and other optical elements (such as deflectors or stigmators) and determining how the elements behave when added together to make up a whole system. To do this it is necessary to calculate the trajectories of the electrons or ions in the optical elements. The first edition of this book (in 1997) contained an extensive chapter which covered many aspects of numerical calculations relevant to this problem. In recent years software for the design and analysis of electron optics for microscopy has become a mature subject and a variety of tools are commercially available. For this reason and in the interest of space (two new chapters have been added) we have decided to provide only this appendix, which briefly outlines the nature of the problem. For further information the reader should probably begin with the first edition of this book.

To calculate the properties of an electron microprobe, for example, one could, in principle, find the distribution of electric or magnetic fields (or both) of a given element (i.e., lens) and then, solve Newton's equations of motion for the charged particles that start at a source and end up on a target. The distribution of the particles would give the current density distribution on the target from which one could estimate the system resolution or other properties. The other approach is to characterize the lenses, for example, in terms of their Gaussian focal properties by solving the paraxial ray equation to find the cardinal elements. Once the trajectories are known, the aberrations can be found in terms of aberration integrals. Classical optical methods are then used to find the current density distributions or other characteristics in terms of the lens focal properties and aberrations. The latter approach is perhaps preferable because it is not only easier and faster to do, but it gives one a better insight into for what the optical system is doing: "optics is a distinct advance over ballistics" (P.W. Hawkes). For example, if a system is being designed for a particular application it is easier to decide how to build the lenses and how to combine them if one knows their focal and aberration properties—the lenses can be treated as building blocks and tailored for a particular use.

The problem then consists of two parts: (1) the calculation of the electric or magnetic fields inside a lens given the geometrical shape of the electrodes (electrostatic lenses) or pole pieces (magnetic lenses) and the potentials applied to the electrodes or the excitation of the pole pieces (in the case of magnetic lenses it is also necessary to take into account the permeability of the metal of the pole pieces); (2) the calculation of the particle trajectories as they move through the lens field(s).

When the trajectories are known the Gaussian lens properties (e.g., focal planes, principal planes) can be calculated, and given the trajectories and the fields, the lens aberrations can be found.

Consider the case of an electrostatic lens. The potential  $\Phi(r, z)$  is everywhere determined by solving Laplace's equation subject to the boundary conditions of the voltages placed on the various electrodes and at infinity (see Section 5.2.1). Because the lens electrodes may have arbitrary shapes it is almost always not possible to do this analytically, and a number of numerical ways of solving the problem have evolved. Three families of methods have been found useful for this purpose: the finite-difference method (FDM), the finite-element method (FEM), and the boundary-element (or charge density) method (BEM). In the FDM the space in which the potential of field is to be calculated is covered with a mesh and Laplace's equation is used to relate the values at the nodes. A systematic procedure (relaxation) then enables us to find a set of values that are compatible with the boundary values and  $i$ th iteration of Laplace's equation. This is described elaborately in earlier texts such as Klemperer (and references therein) and also in Hawkes and Kasper (1989). In the early 1970s, the use of finite elements was introduced into electron optics by Munro (*Image Processing and Computer-Aided Design in Electron Optics*). In the FEM the mesh is no longer necessarily rectangular and meshes of very different sizes can be used in different areas, which means that the mesh size can be matched to the rate at which the potential varies—slowly far from the electrodes, for example, and rapidly close to areas where the shape changes, such as corners. Once again, the values at the nodes are related, this time with the aid of a variational approach, and a solution is found. Unlike the FDM, the FEM is in widespread use and, as mentioned later, is the principal tool in many software suites. The present trend is to use the first-order FEM, simply increasing the number of points until no further improvement is found (Lencová, 2007) or using second-order versions of the methods (Munro et al., 2006), the latter providing higher accuracy with fewer points. Finally, the BEM has the attraction of providing a solution when the geometry of the system is such that the other methods are difficult to use. Because high accuracy is necessary in these calculations, a lot of effort has gone into developing ways to ensure that good accuracy can be had at a reasonable cost in computing time.

Once the potential  $\Phi(r, z)$  is known, the usual approach for a trajectory calculation is to assume the electrons stay close to the optical axis and to make what is called the paraxial approximation, in which powers of the radial coordinate higher than the third (or sometimes, the fifth) can be ignored (we assume circular symmetry in this example, see Section 5.2.2). The potential is then expanded in powers of  $r$  and the result is used to solve the equations of motion, with the  $z$ -axis being the axis of symmetry of the lens and the electrons traveling generally in the  $z$ -direction close to the axis. The potential close to the axis is found in terms of the axial potential  $\phi(z)$  and its derivatives  $\phi'(z)$ ,  $\phi''(z)$  etc., where the ' denotes a derivative with respect to  $z$ . Accurate calculation of the derivatives of  $\Phi(r, z)$  requires an accurate calculation of the potential itself, of course. The potential near the axis given by  $\Phi(r, z) = \phi(z) - 1/4\phi''(z)r^2 + 1/64\phi'''(z)r^4 - \dots$  and the solution  $r(z)$  to the paraxial ray equation

$$r''(z) + \frac{\phi'(z)}{2\phi(z)} r'(z) + \frac{\phi''(z)}{4\phi(z)} r(z) = 0$$

found by solving the Lagrange equations of motion, is obtained by numerical integration with appropriate boundary conditions. Aberration coefficients can be calculated in terms of  $r$  and  $\phi$  and their derivatives (see Chapter 6); for example, the chromatic aberration is given by

$$C_{ci} = \frac{\sqrt{\phi(z_i)}}{r'^2(z_i)} \int_{z_0}^{z_i} \left( \frac{\phi'(z)}{2\phi(z)} r'(z) + \frac{\phi''(z)}{4\phi(z)} r(z) \right) \frac{r}{\sqrt{\phi(z)}} dz$$

(the subscript  $i$  means the aberration coefficient is referred to the image plane).

This is meant only to give an idea of how the subject is approached and how to find a solution, and is by no means a complete outline. A very great deal of effort has gone into finding fast and accurate ways of solving these problems and there is a considerable body of literature on the subject. Software is also available: we refer the reader in the direction of the extensive literature on the subject, as well as several web sites where computational tools are available.

*Software developed specifically for electron microscopy/lithography:*

- CPO Ltd.: [www.electronoptics.com](http://www.electronoptics.com)
- Munro Electron Beam Software (MEBS): [www.mebs.co.uk](http://www.mebs.co.uk)
- Software for Particle Optics Computations (SPOC): [www.lencova.cz](http://www.lencova.cz)

*Other useful software for general charged particle optics work:*

- Field Precision: <http://www.fieldp.com/>
- Lorentz 2D: [www.integratedsoft.com/products/lorentz/](http://www.integratedsoft.com/products/lorentz/)
- SIMION: [www.simion.com](http://www.simion.com)
- Vector Fields: [www.vectorfields.com/](http://www.vectorfields.com/)

## FURTHER READING

### BOOKS

- J. Orloff (Ed.), *Handbook of Charged Particle Optics*, Chapter 1, First Edition, CRC Press, Boca Raton, FL, 1997.  
 P.W. Hawkes and E. Kasper, *Principles of Electron Optics*, Vols 1 and 2, Academic Press, London, 1989.  
 P.W. Hawkes (Ed.), *Image Processing and Computer-Aided Design in Electron Optics*, Academic Press, London and New York, 1973.  
 S. Humphries, *Charged Particle Beams*, John Wiley & Sons, Hoboken, NJ, 1990.  
 E. Harting and F.H. Read, *Electrostatic Lenses*, Elsevier Publishing Company, Amsterdam, 1976.  
 P.W. Hawkes (Ed.), *Magnetic Electron Lenses*, Springer Verlag, New York, 1982.  
 O. Klempner and M.E. Barnett, *Electron Optics*, Cambridge University Press, London and New York, 1971.

### PAPERS AND CONFERENCE PROCEEDINGS

- D. Cubric, B. Lencová and F.H. Read, Comparison of finite difference, finite element and boundary element methods for electrostatic charged particle optics, *Electron Microscopy and Analysis*, Institute of Physics Conference Series, **153**, 1997, pp. 91–94  
 D. Cubric, B. Lencová, F.H. Read and J. Zlámal, Comparison of FDM, FEM and BEM for electrostatic charged particle optics, *Nucl. Instrum. Meth. Phys. Res.*, **427(1)**, 1999, pp. 357–362.  
 D. Dahl, SIMION for the personal computer in reflection, *Int. J. Mass Spectrom.*, **200**, 2000, pp. 3–25.  
 M. Lenc and B. Lencová, Analytical and numerical computation of multipole components of magnetic deflectors, *Rev. Sci. Instrum.*, **68**, 1997, pp. 4409–4414.  
 B. Lencová, Computation of electrostatic lenses and multipoles by the first order finite element method, *Nucl. Instrum. Meth. Phys. Res.*, **A363**, 1995, pp. 190–197.  
 B. Lencová, Accurate computation of magnetic lenses with FOFEM, *Nucl. Instrum. Meth. Phys. Res.*, **A427**, 1999, pp. 329–337.  
 B. Lencová, Accuracy estimate for magnetic electron lenses by first-order FEM, *Nucl. Instrum. Meth. Phys. Res.*, **A519**, 2004, pp. 149–153.  
 B. Lencová and J. Zlámal, The development of EOD program for the design of electron optical devices, *Microsc. Microanal.*, **13**(Suppl. 3), 2007, pp. 2–3 (Proc. Micro. Conf., 2007, Sauerbruecken).  
 E. Munro and H.C. Chu, Numerical analysis of electron beam lithography systems. Part I: Computation of fields in magnetic deflectors. Part II: Computation of fields in electrostatic deflectors, *Optik*, **60**(4), 1982, pp. 371–390 and **61**(1), 1982, pp. 1–16.

- E. Munro, J. Rouse, H. Liu, L. Wang and X. Zhu, Simulation software for designing electron and ion beam equipment, *Microelec. Eng.*, **83**(4–9), 2006b, pp. 994–1002.
- F.H. Read and N.J. Bowring, Ultimate numerical accuracy of the surface charge method for electrostatics, in *Computation in Electromagnetics*, Conference Publication, 420, 1996, pp. 57–61 (Institution for Electrical Engineers, London).

## ADDITIONAL REFERENCES

For further information, a considerable amount of research on computer methods in charged particle optics may be found in the Proceedings of the International Conferences on Charged Particle Optics (CPO), which are published in *Nuclear Instruments and Methods A*. These Conferences have taken place every 4 years, beginning in 1980. The Proceedings of CPO-7 (2006) have not yet (June 2008) been published, but we understand that the proceedings are to appear online in *Physics Procedia* (Elsevier). Also, beginning in 1993 there has been a regular adjunct to the SPIE meetings, having to do with electron optics, in which useful references can be found.

- CPO-1**, H. Wollnik (Ed.), *Nucl. Instrum. Meth. Phys. Res.*, **187**, 1981.
- CPO-2**, S.O. Schriber and L.S. Taylor (Eds), *Nucl. Instrum. Meth. Phys. Res.*, **A258**, 1986.
- CPO-3**, P.W. Hawkes (Ed.), *Nucl. Instrum. Meth. Phys. Res.*, **A298**, 1990.
- CPO-4**, K. Ura, M. Hibino, M. Komura, M. Kurashige, S. Kurokawa, T. Matsuo, S. Okayama, H. Shimoyama and K. Tsuno (Eds), *Nucl. Instrum. Meth. Phys. Res.*, **A363**, 1995.
- CPO-5**, P. Kruit and P.W. van Amersfoort (Eds), *Nucl. Instrum. Meth. Phys. Res.*, **A427**, 1998.
- CPO-6**, A. Dragt and J. Orloff (Eds), *Nucl. Instrum. Meth. Phys. Res.*, **A519**, 2004.
- W.B. Thompson, M. Sato and A.V. Crewe (Eds), *Charged-Particle Optics*, San Diego CA, July 15, 1993, *Proc. SPIE*, **2014**, 1993.
- E. Munro and H.P. Freund (Eds), *Electron-Beam Sources and Charged-particle Optics*, San Diego, CA, July 10–14, 1995 *Proc. SPIE*, **2522**, 1995.
- E. Munro (Ed.), *Charged-Particle Optics II*, Denver CO, August 5, 1996, *Proc. SPIE*, **2858**, 1996.
- E. Munro (Ed.), *Charged-Particle Optics III*, San Diego CA, *Proc. SPIE*, **3155**, 1997.
- E. Munro (Ed.), *Charged-Particle Optics IV*, Denver CO, *Proc. SPIE*, **3777**, 1999.
- O. Delage, E. Munro and J.A. Rouse (Eds), *Charged Particle Beam Optics Imaging*, San Diego CA, July 30, 2001. In *Charged Particle Detection, Diagnostics and Imaging*, *Proc. SPIE*, **4510**, 2001, 71–236.
- For completeness, we mention also the series on Problems of theoretical and applied electron optics [Problemyi Teoreticheskoi i Prikladnoi Elektronnoi Optiki]
- Proceedings of the First All-Russia Seminar*, Scientific Research Institute for Electron and Ion Optics, Moscow, 1996. *Prikladnaya Fizika*, 1996, No. 3.
- Proceedings of the Second All-Russia Seminar*, Scientific Research Institute for Electron and Ion Optics, Moscow, April 25, 1997. *Prikladnaya Fizika*, 1997, No. 2–3.
- Proceedings of the Third All-Russia Seminar*, Scientific Research Institute for Electron and Ion Optics, Moscow, March 31–April 2, 1998. *Prikladnaya Fizika*, 1998, Nos 2 and 3/4.
- A.M. Filachev and I.S. Gaidukova (Eds), *Proceedings of the Fourth All-Russia Seminar*, Scientific Research Institute for Electron and Ion Optics, Moscow, October 21–22, 1999. *Prikladnaya Fizika*, 2000, Nos 2 and 3 and *Proc. SPIE*, **4187**, 2000.
- A.M. Filachev (Ed.), *Proceedings of the Fifth All-Russia Seminar*, Scientific Research Institute for Electron and Ion Optics, Moscow, November 14–15, 2001. *Prikladnaya Fizika*, 2002, No. 3 and *Proc. SPIE*, **5025**.
- A.M. Filachev and I.S. Gaidukova (Eds), *Proceedings of the Sixth All-Russia Seminar*, Scientific Research Institute for Electron and Ion Optics, Moscow, May 28–30, 2003. *Prikladnaya Fizika*, 2004 and *Proc. SPIE*, **5398**, 2004.
- A.M. Filachev and I.S. Gaidukova (Eds), *Proceedings of the Seventh All-Russia Seminar*, Scientific Research Institute for Electron and Ion Optics, Moscow, May 25–27, 2005. *Prikladnaya Fizika*, 2006 and *Proc SPIE*, **6278**, 2006.

# Index

## A

- Aberration coefficient, evolution of, 608–614  
Aberration correction  
    aberration function limitations, 606–614  
    all-electrostatic correctors, 300–301  
    combination aberrations, 617–618  
    computer control, 631  
    corrector classification, 619–625  
    corrector operation, 625–626  
    diagnostic methods, 627–631  
    fifth order aberration function, 640  
    future research, 637–638  
    general multipole correctors, 301–304  
    historical background, 602–605  
    magnetic lens design, 145–150  
        dodecapole multipole field component generation, 145  
        electrostatic/magnetic quadrupole lens, 148–150  
        hexapole spherical corrector, transfer doublet, 146–148  
    minimization, 296–297  
    mirrors, 304–305  
    misalignment aberrations, 618–619  
    miscellaneous techniques, 305–310  
    mixed quadrupole correctors, 301  
    monochromators and, 210  
    optical columns, 631–637  
    overview, 294–295, 601–602  
    proof-of-principle correctors, 603–604  
    quadrupoles and octupoles, 297–298  
    scanning transmission electron microscope, 518–519  
    sextupoles, 298–300  
    single multipole effect, 614–617  
    third-generation correctors, 605  
    total system aberrations, 626–627  
    trajectory calculation, 605–606  
    working correctors, 604–605  
Aberration function concept, aberration corrector optics, 606–614  
Aberrations  
    basic properties, 210–211  
    calculation methods, 211–221  
    chromatic aberrations  
        curved optic axes, 265–271  
        mirrors and cathode lenses, 259–265  
        quadrupole lenses, 249–252  
        round lenses, 247–249, 254–259  
        sextupole lenses, 252–254  
        superimposed deflection fields, 254–259  
        Wien filter, 271–286  
    electrostatic lenses, 176  
    fundamental aberrations, 601–602  
    gas field ionization source, 112–114  
        beam acceptance aperture, 113  
        column aberrations, 114  
    local angular distortion, 113  
    source radius, 113–114  
    space charge, 112  
    spherical and chromatic, 113  
geometric aberrations  
    curved optic axes, 265–271  
    mirrors and cathode lenses, 259–265  
    quadrupole lenses, 249–252  
    round lenses, 221–249, 254–259  
    sextupole lenses, 252–254  
    superimposed deflection fields, 254–259  
    Wien filter, 271–286  
information sources on, 310  
optical transfer function, monoenergetic beam, 407–409  
parasitic aberrations, 292–294  
representation, 286–290  
space charge effects, 347–348  
symmetry, 290–292  
Absorbed electrons, scanning electron microscope, 463–464  
Acceleration voltages, space charge effects, 348–349  
Acceptance factor, resolution, image quality, 404–405  
Achromatic aberrations, quadrupole lenses, 252  
Acryl glass, electrostatic lens, insulator materials and design, 189–190  
Activation energy, liquid metal ion sources, field evaporation, 37–38, 40–42  
Adsorbed molecule density, focused ion beam technology, gas-assisted deposition, 562–566  
Airy pattern, resolution  
    charged particle image formation, 397  
    contrast performance and, 401–402  
    optimum condition, 429–430  
    Rayleigh's criterion, 395–396  
Aliasing, focused ion beam/secondary-ion mass spectrometry, 577–581  
Alignment accuracy, electrostatic lens, 190–191  
All-electrostatic aberration correctors, 300–301  
Alloy ion sources, liquid metal ions, 66–67  
Ampere turns (AT), pole-piece lens design, 132, 134–136  
Analytical theory  
    Boersch effect  
        equation summary, 372–373  
        extended two-particle approximation, 362–364  
        microbeam column design, 381  
        numerical examples, 374–375  
        thermodynamic limits, 373–374  
    Monte Carlo simulation vs., 385–386  
    space charge defocus and aberration, 305–306, 347–348  
        defocus conditions, 346–347  
        laminar flow, 344–346  
        numerical examples, 348–349  
        ray equation, 344–345

- Analytical theory (*contd.*)
- statistical coulomb effect
    - beam segment displacement, 364–365
    - closest encounter approximation, 357
    - extended two-particle approximation, 358–359
      - displacement distribution, 359–364
    - first-order perturbation models, 355–357
    - mean square field fluctuation approximation, 357–358
  - microbeam column design, 377–382
    - Boersch effect, 381
    - interaction calculations, 381–382
    - statistical interaction minimization, 381
    - trajectory displacement and probe size, 377–380
  - nonmicrobeam instruments, 382–385
  - N*-particle reduction, 352–353
  - plasma physics models, 354–355
  - problem formulation, 349–352
  - two-particle dynamics, 353–354
- trajectory displacement, 350–352
- equation summary, 368
  - first-order perturbation models, 355–357
  - mean square field approximation, 357–358
  - nonmicrobeam instruments, 383–385
  - numerical examples, 369–370
- Angular change
- statistical coulomb effect
    - closest encounter approximation, 357
    - first-order perturbation models, 356–357
  - trajectory displacement, 366–368
- Angular intensity
- liquid metal ion sources, ion optical effects, 63–64
  - ZrO/W Schottky emission, 10–12
- Angular magnification
- Müller emitter design, 109
  - spherical charged particle emitter, 105–106
- Anisotropic aberration coefficients
- chromatic and geometric aberrations, 223–246
  - spherical aberration, 235–240
- Anisotropy constant, magnetic lens design, 142–145
- Annular dark-field (ADF) imaging
- aberration correction, 611–614
  - optical columns, 633–637
  - scanning transmission electron microscope, 501–502
- Aperture lenses
- aberrations, quadrupole lenses, 250–252
  - action, 172–173
- Apex parameters
- liquid metal ion sources, boundary conditions, 50–51
  - steady high-electrical-conductivity Gilbert–Gray cone-jet, 47–48
- Applied field, gas field ionization emitters, 93
- Approximation method, density-of-information passing capacity, 416–423
- fitting functions, 422–423
  - wave/geometric optics, 418–422
- Archard–Deltrap aberration corrector, 620–625
- Astigmatic lenses
- chromatic and geometric aberrations, 222, 228–246
    - spherical aberration, 242
  - optical properties, 183
  - parasitic aberrations, 292–294
- Astigmatism
- aberration corrector operation, 625–626
  - magnetic lens design, 143–145
- Asymptotic aberration
- quadrupole lenses, 250–252
  - spherical aberration, 240–241
  - superimposed round lens deflection fields, 258–259
- Atomic ions, liquid metal ion sources, 65–67
- Atomic-level ion source (ALIST<sup>TM</sup>)
- gas field ionization emitters, 93
    - source properties, 114–115
    - liquid metal ion sources, 31
- Attainable resolution, optimum condition and, 425–430
- Attempt frequency, gas field ionization emitters, 97
- Attractors, liquid metal ion source
- chaotic attractors, 58–59
  - Taylor cones, 46–47
- Auger electrons
- focused ion beam implantation, 585–591
  - scanning electron microscope, 464–465
- Axial-gap lens, low-voltage scanning electron microscope, 155–156
- Axial magnetic field distribution, pole-piece lens design, 134–136
- Azimuthal variation, aberration correction, 608–614
- B**
- Background gas temperature, gas field ionization emitters, 95
- Backscattered electrons (BSEs)
- scanning electron microscope
    - absorbed electrons, 464
    - collection, 460–461
    - electron beam-specimen interaction, 452–454
    - signal properties, 456, 459–461
    - variable vacuum scanning electron microscope, gas signal detection, 475–477
- Backward aberration coefficients, chromatic and geometric aberrations, 224–246
- Beam acceptance aperture diffraction, gas field ionization source, 113
- Beam axis trajectory, statistical coulomb effect, 350–352
- Beam current instability
- focused ion beam technology, gas-assisted etching, 556–559
  - source-limited system, 428–430
  - space charge effect, 342–343
  - ZrO/W cathode
    - shape stability, 18–19
    - total energy distribution, 20–22
- Beam halos, focused ion beam technology, 534–538
- Beam intensity, resolution research, 393–394
- Beam size, resolution and, 398–399
- Beam temperatures, statistical coulomb effect, 354–355
- Bell-shaped lens design
- chromatic and geometric aberrations, 225–246
    - minimization, 296–297
    - historical background, 130
- Bernoulli equation, liquid metal ion sources, viscous-loss terms, 49

- Bessel functions, resolution  
contrast performance and, 401–402  
Rayleigh's criterion, 395–396
- Best focus position, basic properties, 415–416
- Best image field (BIF), gas field ionization emitters, 93, 96
- Best image voltage (BIV), gas field ionization emitters, 93, 96
- Best source field (BSF), gas field ionization emitters, 96–97
- Binary collision model, focused ion beam technology, gas-assisted deposition, 561–566
- Bipolar transistors, focused ion beam implantation, 582–591
- Blunt-needle liquid metal ion source, 31
- Blurred beam design, spherical charged particle emitter, 105
- Blurring cone, spherical charged particle emitter, 104–105
- Blurring ratio/magnification, spherical charged particle emitter, 105–106
- Boersch effect  
equation summary, 372–373  
extended two-particle approximation, 362–364  
microbeam column design, 381  
Monte Carlo simulation, 377  
numerical examples, 374–375  
overview, 342  
parameter dependencies, weak collisions, 370–372  
plasma physics, 354–355  
statistical coulomb effect  
closest encounter approximation, 357  
first-order perturbation models, 356–357  
mean square field approximation, 357–358  
thermodynamic limits, 373–374  
two-particle approximation, 343
- ZrO/W cathode, total energy distribution, 20–22
- Boustrophedonic writing, focused ion beam technology, gas-assisted deposition, 562–566
- Brandon's criterion, liquid metal ion sources, field evaporation, 39–40
- Bremsstrahlung radiation, scanning electron microscope, x-ray collection, 462–463
- Bright-field imaging  
aberration correction and diagnosis, 628–631  
scanning transmission electron microscope, 500–501
- Brightness properties  
scanning transmission electron microscope, 506–508  
ZrO/W cathode, 22–23
- C**
- Calculated resolution, experimental results *vs.*, 423–425
- Canonical coordinates, chromatic aberration calculation, 221
- Captured-flux enhancement factor, gas field ionization emitters, Southon gas-supply theory, 118
- Cartesian coordinates  
aberration calculation, 211–221  
chromatic and geometric aberrations, mirrors and cathode lenses, 260–265
- Cathode lenses  
action model, 174–175  
chromatic and geometric aberration, electron mirror, 259–265  
classification, 171  
objective lenses, 179–181
- Cathode ray tube (CRT)  
resolution  
image formation, 396–397  
image quality, 404–405  
scanning electron microscope, 446–448
- Cathodoluminescence (CL), scanning electron microscope, 464
- Center of curvature, chromatic and geometric aberrations, mirrors and cathode lenses, 262–265
- Central emission current, ZrO/W cathode emitter stability, 13–19
- Cerium hexaboride cathode, scanning electron microscope, 440–443
- Channel electron multiplier (CEM) detector, focused ion beam technology, 547–548  
scanning ion microscopy, 569–571
- Channeling contrast, focused ion beam technology, scanning ion microscopy, 568–571
- Charge control  
focused ion beam technology, 544–549  
variable vacuum scanning electron microscope, 478
- Charged particle (CP) optics  
electrostatic lens evaluation, 169–170  
gas field ionization emitters, 90–91  
resolution  
image formation and, 396–397  
image quality, 404–405
- Charge-draining mechanism, liquid metal ion sources, field evaporation, 40–42
- Charged-surface models, liquid metal ion sources, 36–37
- Charge-exchange effects, gas field ionization emitters, gas-pressure trade-off, 91–92
- Charge-hopping model, liquid metal ion sources, field evaporation, 40–42
- Child–Langmuir law, steady-state current-related liquid metal ion sources, 55–56
- Chromatic aberrations  
aberration correctors, 614  
combination aberrations, 617–618  
transmission electron microscopy, 624–625  
calculation methods, 212–221  
coefficient of, 247–249  
correctors, 294–295  
all-electrostatic correctors, 300–301  
mixed-quadrupole correctors, 301  
curved optic axes, 265–271  
defined, 210  
density-of-information passing capacity, wave and geometric optic equivalence, 418–420  
gas field ionization emitters, 113  
liquid metal ion sources, 59–64  
microbeam column design, trajectory displacement, 377–378  
mirrors and cathode lenses, 259–265  
monoenergetic beam current density, rotational symmetry, 415

- Chromatic aberrations (*contd.*)
  - optical transfer function, electron beam, 409–410
  - quadrupole electrostatic/magnetic chromatic aberration corrector, 148–150
  - quadrupole lenses, 249–252
  - resolution, beam size and, 398
  - round lenses, 247–249, 254–259
  - scanning transmission electron microscope, 505
    - diffraction and, 510
  - sextupole lenses, 252–254
  - superimposed deflection fields, 254–259
  - Wien filter, 271–286
  - working correctors, 604–605
- Chromatic angular intensity, liquid metal ion sources, 60–64
- Closest encounter approximation, statistical coulomb effect, 357
  - extended two-particle approximation, 359
- Clouded-glass effect, aberration correction, 305–306
- Cluster ions, liquid metal ion sources, 65–67
- Coil design
  - magnetic lenses, 136–140
  - scanning electron microscope architecture, 439–440
- Cold field emission (CFE) cathodes
  - scanning electron microscope, 440–443
  - ZrO/W cathode, 3–5
    - applications, 27
    - total energy distribution, 20–22
- Cold field emission electron gun (CFEG), aberration correction, 627
  - optical columns, 631–637
- Cold field emission scanning transmission electron microscope (CFE-STEM)
  - chromatic aberration, 505
  - electron sources, 512
  - proof-of-principle aberration correctors, 606–604
- Collapse voltages, steady-state current-related liquid metal ion sources, 54
- Collision dynamics
  - statistical coulomb effect, extended two-particle approximation, 358–359, 361–364
  - trajectory displacement, weak/incomplete collisions, 365–368
- Collision free interactions, liquid metal ion sources, energy spreads, 61–62
- Column aberrations, gas field ionization source, 114
- Column design and performance, scanning transmission electron microscope, 515–517
- Coma, chromatic and geometric aberrations, 227–246
- Combination aberrations, aberration correctors, 617–618
- Compensation factor, aberration correction, 611–614
- Complete collision
  - space charge effect, extended two-particle approximation, 359
  - statistical coulomb effect, two-particle dynamics, 353
- Compound lenses, scanning electron microscopes, 196–199
- Compression factor, Müller emitter design, projection magnification, 111
- Computer controls, aberration correction and diagnosis, 631
- Conceptual object generation, spherical charged particle emitter, 105–106
- Condenser-objective lens, historical background, 130
- Cone-jets
  - gas field ionization emitters, Southon gas-supply theory, 119–120
  - liquid metal ion source, classification of, 34
- Conical lens technology, scanning electron microscope architecture, 443
- Contamination reduction, variable vacuum scanning electron microscope, 478
- Continuum condition
  - laminar flow, 346
  - scanning electron microscope, x-ray collection, 462–463
- Contrast performance, resolution based on, 399–402
- Converted backscattered secondary electron (CBSE), scanning electron microscope, 461
- Corrector constructors, aberration corrections, 602
- Correlation, statistical coulomb effect, beam segment displacement, 364–365
- Correspondences, aberration calculation, 216–221
- Coulomb interactions, ZrO/W cathode
  - emitter brightness, 22–23
  - total energy distribution, 20–22
- Coupled differential equations, aberration calculation, 217–221
- Critical dimension scanning electron microscope, linewidth measurements, 466–467
- Critical surface, gas field ionization emitters, gas-atom behavior, 94
- Crossing surface, gas field ionization emitters, 93–94
- Crossover location parameter
  - Boersch effect, 372–373
    - numerical examples, 374–375
    - statistical coulomb effect, 349–352
    - plasma physics, 355
    - trajectory displacement, 367–368
  - Crossover mode thermionic cathodes, historical background, 1–3
- Cross section analysis, focused ion beam technology, gas-assisted etching, 555–559
- Current angular intensity, Schottky emission, 6
- Current conditioning, electrostatic lens, 187–188
- Current density ratio
  - monoenergetic beam
    - rotational symmetry, 412–415
    - spatial frequency response, 411–412
    - space charge defocus, 346–347
    - steady-state liquid metal ion sources, apex boundary conditions, 51
    - total energy distribution, ZrO/W cathodes, 19–22
- Current stability, ZrO/W cathodes
  - emitter stability, 14–19
  - short-term frequency vs. long-term current drift, 23–24
- Current voltages
  - space charge effects, 348–349
  - steady-state current-related liquid metal ion sources, 54–57
    - flow impedance, 56
    - Mahoney's equation, 57–58
    - Mair's equation, 54–55
    - temperature effects, 57
- Curved foils, aberration correction, 305–308

- Curved optic axes, chromatic and geometric aberrations, 265–271
- Curve fit method, ZrO/W Schottky emission, 8–10
- Cusp length, steady-state current-related liquid metal ion sources, emission current and, 57
- Cylinder supply, gas field ionization emitters, Southon gas-supply theory, 118–119
- Cylindrical lenses, optical properties, 183
- Cylindrical polar coordinates, aberration calculation, 213–221
- ## D
- Damage layers, focused ion beam technology, 532–534
- Danger areas, focused ion beam technology, 537–538
- Debye screening, statistical coulomb effect, *N*-particle problem, 352–353
- Deflection fields
- chromatic and geometric aberrations, 254–259
  - scanning electron microscope architecture, 439–440
- Defocus distance
- aberration correction, 608–614
  - space charge defocus, 346–347
- Degree of aberration, defined, 211
- Delta-Dirac function, statistical coulomb effect, 351–352
- Density-of-information passing capacity (density-of-IPC) approximation methods, 416–423
  - fitting functions, 422–423
  - wave/geometric optics, 418–422
- best focus position, 415–416
- calculated resolution *vs.* experimental results, 423–425
- calculation *vs.* experimental results, 423–425
- current density, monoenergetic beam, 411–412
- image quality, 402, 405–406
- ion beam spatial frequency response, 411
- optical transfer function
  - chromatic aberration and, 409–410
  - monoenergetic beam, 407–409
- optimum condition, source-limited system, 429–430
- resolution
  - image quality, 404–405
  - Rayleigh's criterion and, 405–406
- rotationally symmetric systems, monoenergetic beam current density, 412–415
- spatial frequency response
  - electron beam, 406–410
  - ion beam, 411–415
- two-dimensional Fourier transform, source intensity distribution, 406–407
- Depth-of-field, scanning electron microscope architecture, 439
- Depth profiling, focused ion beam/secondary-ion mass spectrometry, 575–581
- Detection modes and signals, scanning transmission electron microscope, 503
- Detector bias, focused ion beam technology, 546–549
- Diagnostic methods, aberration correction, 627–631
- Differential algebra
- basic principles, 339–340
  - chromatic aberration calculation, 219–221
- Differential part, defined, 340
- Differential unit, defined, 340
- Diffraction-limited optical systems, optimum condition, 426–428
- Diffraction plane
- chromatic and geometric aberrations, curved optic axes, 268–271
  - scanning transmission electron microscope, 504–505
- Diffractogram-based autotuning, aberration diagnosis, 627–631
- Digital image storage and analysis, scanning electron microscope, 446–449
  - transmitted electrons, 465
- Dimensionless quantity, aberration calculation, 211–221
- Direct-action aberration correctors, 619–625
- Direct ray-tracing, electrostatic lens evaluation, 167–168
- Disk of least confusion, chromatic and geometric aberrations, 224–246
- Displacement distribution, statistical coulomb effect, 349–352
  - beam segments, 364–365
  - closest encounter approximation, 357
  - extended two-particle approximation, 358–359, 359–364
  - first-order perturbation models, 355–357
  - N*-particle problem, 352–353
  - regime characteristics and, 361–364
- Distortion, chromatic and geometric aberrations, 228–246
- Distribution shape, liquid metal ion sources, ion optical characteristics, 62
- Dodecapole, magnetic lens multipole field component generation, 145
- Doubly symmetric electrostatic corrector (DECO), chromatic and geometric aberrations, 301–302
- Droplet emission, liquid metal ion sources, upper unsteady regime, 68–69
- Dual-beam machines, focused ion beam technology, transmission electron microscopy, 558–559
- Dual-channel deflection system, third-order geometric aberrations, round lenses, 257–259
- Dulling rate, ZrO/W cathode shape stability, 17–19
- Dynamic low-field equilibrium (DLFE), gas field ionization emitters, 99–100
- ## E
- Effective capture area, gas field ionization emitters, 99
- Eikonal methods, chromatic aberration calculation, 220–221
- Einzel lenses, 171
  - all-electrostatic aberration correctors, 301
  - chromatic and geometric aberrations, 226–227
- Electrical insulation, electrostatic lens, 187–188
- Electrical surface properties, liquid metal ion sources, charge-surface models and Maxwell stress, 37
- Electric field stresses, ZrO/W cathode shape stability, 18–19
- Electroforming, gas field ionization emitters, 93

- Electrohydrodynamic (EHD) effect, liquid metal ion sources, 31–33
- apex boundary conditions, 50–51
  - equilibrium and stability conditions, 45–46
  - spraying, 35
  - steady high-electrical-conductivity Gilbert–Gray cone-jet, 47–50
  - surface pressure jump, 44–45
  - Taylor's mathematical cone, 46–47
- Electrohydrostatic equilibrium, liquid metal ion source, 45–46
- Electron beam
- best focus position, 415–416
  - scanning electron microscope
    - architecture, 438–439
    - landing position, 483–485
    - resolution measurements, 478–480
    - specimen interactions, 452–456
  - spatial frequency response, 406–410
    - source intensity distribution, two-dimensional Fourier transform calculation, 406–407
- Electron beam induced current (EBIC) imaging, scanning electron microscope, 464
- applications, 470
  - contamination, 485–487
- Electronegative gases, emitter environmental requirements, 24–25
- Electron energy distribution, aberration correction, 308–310
- Electron energy loss spectroscopy (EELS)
- aberration correction and diagnosis, optical columns, 634–637
  - scanning transmission electron microscope, 502
- Electron flooding, helium ion microscopy, 491–492
- Electron gun
- focused ion beam technology, 548–549
  - scanning transmission electron microscope, 513–515
- Electronic losses, focused ion beam technology, ion-sample interactions, 527–534
- Electron mirror
- aberration correctors, 304–305
  - chromatic and geometric aberration, 259–265
  - Wien filter, 271–286
  - electrostatic lens design, 171
  - optical properties, 181
- Electron-optical design
- chromatic and geometric aberrations, Wien filter, 273–286
  - column components, scanning transmission electron microscopy, 511–518
    - cold field emitter source, 512
    - design and performance, 515–517
    - electron gun, 513–515
    - electron source, 511
    - magnetic lenses, 515
    - thermally assisted, zirconium-treated field emitter source, 513
  - Electron probe formation, scanning transmission electron microscope, 503–506
    - chromatic aberration, 505
    - diffraction, 504–505
- geometric size, 504
- spherical aberration, 505–506
- Electron range, scanning electron microscope, electron beam-specimen interaction, 452–454
- Electron source characteristics, scanning electron microscope architecture, 440–443
- Electron-stimulated desorption (ESD), emitter environmental requirements, 24–25
- Electron trajectories
- cathode design, 1–2
  - electrostatic/magnetic chromatic aberration corrector, 148–150
- Electrostatic detector objective lens (EDOL), scanning electron microscopes, 198–199
- Electrostatic field, aberration calculation, 212–221
- Electrostatic lenses
- aberration correction, 308–309
  - applications, 199
  - chromatic and geometric aberrations, 226–246
    - anisotropic components, 243–246
    - minimization, 296–297
    - mirrors and cathode lenses, 264–265
  - cylindrical/astigmatic lenses, 183
  - design problems, 185–186
  - environmental and system considerations, 191–192
  - future research issues, 200–201
  - insulator materials, 188–190
  - ion microscopy and lithography, 199–200
  - low-energy and photoemission electron microscopes, 194–196
  - low-voltage transmission electron microscopy, 194
  - manufacturing and alignment accuracy, 190–191
  - optical properties, 166–185
    - aberrations, 176
    - action models, 174–175
    - aperture lens, 172–173
    - electron mirror, 181
    - equation of motion and trajectory equation, 168–170
    - grid and foil lenses, 181–183
    - homogeneous fold action, 173–174
    - immersion objective lenses, 179–181
    - multielement and zoom lenses, 179
    - round lenses, 177–183
    - thick lenses, 170–171
    - three-electrode unipotential lenses, 178–179
    - trajectory matrix, 172
    - two-electrode immersion lenses, 177–178
  - optimization, 183–185
  - overview, 162–166
  - scanning electron microscopes, compound and retarding lenses, 196–199
  - transmission electron microscopes, 192–194
  - vacuum gap electrical insulation and breakdown, 187–188
- Electrostatic/magnetic lens design
- low-voltage scanning electron microscope, 156–157
  - quadrupole chromatic aberration corrector, 148–150
- Electrostatic potential, aberration calculation, 212–221
- Elemental ion sources, liquid metal ions, 64–66
- Elemental mapping, focused ion beam/secondary-ion mass spectrometry, 575–581
- Ellipticity, electrostatic lens manufacturing, 191

- EM8 electrostatic lens, historical background, 162–166  
Emission current  
gas field ionization emitters, theoretical background, 97–98  
steady-state current-related liquid metal ion sources, cusp length and, 57  
Emission regimes, liquid metal ion sources, 75  
Emission-site radius, gas field ionization emitters, 101–102  
Emittance, conservation principle of, statistical coulomb effect, 354–355  
Emitter-base junction forward bias, focused ion beam implantation, 586–591  
Emitter brightness, ZrO/W cathode, 22–23  
Emitter environmental requirements, ZrO/W cathodes, 24–25  
Emitter formation, gas field ionization emitters, 93  
Emitter geometry  
liquid metal ion source, 34  
ZrO/W cathode, 8–10  
emitter life mechanisms and, 26–27  
shape stability, 12–19  
Emitter life mechanisms, ZrO/W cathodes, 26–27  
Emitter temperature, ZrO/W emitter life and, 26–27  
Energy dispersive x-ray spectrometer (EDS)  
scanning electron microscope, 461–463  
scanning transmission electron microscope, 502  
Energy distribution  
gas field ionization emitters, 97, 108  
liquid metal ion sources, 59–64  
energy spreads, 61–62  
ion energy distributions, 60–63  
ion optical effects, 63–64  
monoenergetic beam current density, rotational symmetry, 415  
Energy filters, chromatic and geometric aberrations, curved optic axes, 268–271  
Energy scanning, electrostatic lenses, 171  
Environmental factors, electrostatic lens manufacturing, 191–192  
Environmental scanning electron microscope (ESEM), evolution of, 472  
Epoxy resin, electrostatic lens, insulator materials and design, 189–190  
Equation of motion  
aberration calculation, 211–221  
electrostatic lens evaluation, 168–170  
Equivalency functions  
fitting functions for, 420–422  
wave and geometric optics, density-of-information passing capacity, 418–420  
Equivalent emission current, liquid metal ion source, 33–34  
Error analysis, Monte Carlo simulation, 376–377  
Escape charge states, liquid metal ion sources, field evaporation, 39–40  
Escape mechanism, liquid metal ion sources, field evaporation, 40  
Euler equations, aberration calculation, 218–221  
Evaporation rate, ZrO/W emitter life and, 26–27  
Everhart-Thornley (ET) detector  
focused ion beam technology, scanning ion microscopy, 570–571  
scanning electron microscope  
backscattered electron collection, 459–461  
secondary electron collection, 457–459  
Excitation NI  
low-energy electron microscope/photoelectron emission microscope, 150–152  
magnetic lens coil design, 137–140  
pole-piece lens design, 134–136  
Excitation parameters, quadrupole lens aberrations, 251–252  
Extended beam condition, statistical coulomb effect, two-particle dynamics, 353  
Extended field lens technology, scanning electron microscope architecture, 444–446  
Extended Schottky emission (SE)  
basic equations, 6–8  
field factor  $\beta$ , emitter radius, and work function, 8–10  
total energy distribution, 19–22  
ZrO/W cathode, 3–5  
total energy distribution, 20–22  
Extended two-particle approximation  
limitations of, 385–386  
space charge effect, 358–359  
Extinction current, steady-state current-related liquid metal ion sources, 51–52, 57–58  
Extinction voltage, steady-state current-related liquid metal ion sources, 51–59  
Extraction voltage  
Schottky emission, 6  
ZrO/W Schottky emission, 10–12  
Extra high-resolution scanning electron microscopy, 492–493
- ## F
- Fabrication process, liquid metal ion sources, 71–73  
Facet lenses, optical properties, 182  
Failure analysis, scanning electron microscope, 470–472  
Fall-off half-angle (FOHA)  
gas field ionization emitter  
field ion microscope resolution, 107  
spherical charged particle emitter, 104–105  
gas field ionization emitters, 101–102  
Müller emitter design, 106–107  
Fall-off radius (FORD)  
gas field ionization emitters, 101–102  
beam acceptance aperture diffraction, 113  
minimum values, 108  
gas field ionization source, virtual sources, 107  
Müller emitter source size, 111  
Faraday constant, liquid metal ion source, 33–34  
Fermi level, ZrO/W cathode emissions, 3–4  
Field curvature, chromatic and geometric aberrations, 228–246  
anisotropic components, 242–246  
Field desorption, liquid metal ion sources, 31  
Field emission electron propulsion (FEEP), liquid metal ion source, 32  
Field emission gun (FEG), objective lens design, low-voltage scanning electron microscope, 153–156

- Field-emitted vacuum space charge (FEVSC), liquid metal ion sources, field evaporation, 42–44
- Field evaporation (FEV)  
gas field ionization emitters, 93  
liquid metal ion sources, 37–42  
activation energy dependence, 40–42  
equilibrium and supply limits, 38–39  
escape mechanism, 40  
evaporation-field values prediction, 39–40  
rate and time constants, 37–38
- Field expansions  
chromatic and geometric aberrations, curved optic axes, 265–271
- ZrO/W cathode emitter stability, 12–19
- Field factor  $\beta$ , ZrO/W cathode, 8–10
- Field ion emission sources, liquid metal ion sources, 31
- Field ionization  
gas field ionization emitters, 88–89  
liquid metal ion sources, 31  
field evaporation and, 37
- Field ion microscopy (FIM)  
electrostatic lens, 165–166  
gas field ionization emitters, 89–91  
resolution criterion, 103, 107  
helium ion microscopy, 489–492  
image-spot size, 111–112  
liquid metal ion sources, 65–67  
Müller emitter design, projection magnification, 111
- Field particles, statistical coulomb effect,  $N$ -particle problem, 352–353
- Field-stabilized forms, ZrO/W cathode, 4–5
- Field strength, electrostatic lens, insulator materials and design, 189–190
- Fifth-order aberration coefficient, calculation of, 220–221
- Fifth-order aberrations, 210
- Figure of merit  
focused ion beam technology, gas-assisted deposition, 564–566  
gas field ionization emitters, 115  
resolution, image quality and, 403–405
- Finite difference method (FDM), electrostatic lens evaluation, 166–168
- Finite element method (FEM)  
electrostatic lens evaluation, 166–168  
lens design, 131  
pole-piece design, 132–136
- Finite size effects, Monte Carlo simulation, 376–377
- Firmly field-adsorbed layer, gas field ionization emitters, gas-atom behavior, 94
- First-generation aberration correctors, 603–604
- First-order perturbation models  
space charge effect, extended two-particle approximation, 359  
statistical coulomb effect, 355–357
- Fitting functions  
equivalency functions, wave and geometric optics, 420–422  
information passing capacity, 422–423  
numerical computation vs., 423
- Flashing process, scanning electron microscope cathodes, 441–443
- Flood gun, focused ion beam technology, 547–548
- Flow impedance, steady-state current-related liquid metal ion sources, 56
- Flux density, liquid metal ion source, 33–34
- Focused ion beam (FIB)  
applications, 525–526, 549–593  
gas-assisted materials deposition, 559–567  
implantation, 581–591  
lithography, 591–593  
micromachining and gas-assisted etching, 549–559  
scanning ion microscopy, 567–571  
secondary-ion mass spectrometry, 571–581
- architecture, 534–537
- boustrrophedonic principle, 595–596
- charging sample vexation, 544–549
- electrostatic lenses, 199–200
- future research issues, 593–594
- gas field ionization emitters, 90–91
- historical background, 131–132
- imaging properties, 538–544
- ion-sample interactions, 526–534
- ions and electrons, 538
- liquid metal ion sources, 31–33  
emitter shape formation, 34
- overview, 523–524
- research history, 524–525
- resolution research, 392–393
- spherical charged particle emitter, Müller emitter design, 106–107
- Focused ion stripe transistor (FIST) device, focused ion beam implantation, 588–591
- Focused probe optimization, scanning transmission electron microscope, 508–510
- Focus size condition, laminar flow, 346
- Foil lenses, optical properties, 181–182
- Fokker–Planck equation, statistical coulomb effect, 354–355
- Forbes 1982 formula, liquid metal ion sources, field evaporation, 40–42
- Force-current ratios (FCRs), steady-state current-related liquid metal ion sources, 52–53
- Forward aberration coefficients, chromatic and geometric aberrations, 224–246
- Fourier transform  
resolution  
contrast performance and, 400–402  
scanning electron microscope, 482–483  
statistical coulomb effect, extended two-particle approximation, 360–364
- Fowler–Nordheim constant, gas field ionization emitters, 97
- Free-space field ionization (FSFI)  
gas field ionization emitters, 99–100  
liquid metal ion sources, 37
- Fresnel functions, monoenergetic beam, optical transfer function, 408–409
- Fringing fields, chromatic and geometric aberrations, curved optic axes, 265–271
- Full width at half maximum (FWHM) values  
Boersch effect, kinetic energy relaxation, 373–374  
gas field ionization emitters, 97  
limitations of, 385–386

- liquid metal ion sources, chromatic aberration, 59–64  
resolution research, 392–393  
  basic properties, 393–394  
space charge effect, 343  
statistical coulomb effect, 351–352  
  extended two-particle approximation, 363–364  
total energy distribution, ZrO/W cathodes, 19–20  
trajectory displacement, 367–368  
upper unsteady regime, most-steady regime, 35
- Fundamental aberrations, defined, 601–602  
FW50 values, ZrO/W cathode, total energy distribution, 21–22
- ## G
- Gallium arsenide MESFETs, focused ion beam implantation, 582–591
- Gallium liquid metal ions  
  emission regimes, 75  
  gas field ionization emitters, 114–115
- $\gamma$ -rule for beam segments, statistical coulomb effect, 365
- Gap resolution  
  measurement using, 431–433  
  scanning electron microscope, 480–483
- Gas-assisted etching (GAE), focused ion beam technology, 553–559
- Gas-atom behavior, gas field ionization emitters, 94  
  history, 95–96
- Gas cascade cell, variable vacuum scanning electron microscope, 476–477
- Gas concentration, gas field ionization emitters, 94
- Gas conditioning, electrostatic lens, 188
- Gas effects, variable vacuum scanning electron microscope, primary electron beam, 474–475
- Gas field ionization source (GFIS)  
  aberrations, 112–114  
    beam acceptance aperture, 113  
    column aberrations, 114  
    local angular distortion, 113  
    source radius, 113–114  
    space charge, 112  
    spherical and chromatic, 113  
  advantages of, 91  
  corrected Sounthon gas-supply theory, 117–122  
  emission current theory, 97–98  
  emission-site radius, 101–102  
  emitter formation, 93  
  energy spreads, 97  
  field definitions, 93  
    illustrative values, 103  
    ion microscope resolution criterion, 103  
  focused ion beam technology, 524–525  
    future applications, 593–594  
  gas-atom behavior, 94–96  
  gas-pressure trade-off, 91–92  
  glossary of terms, 122–124  
  helium ion microscopy, 488–492  
  ion generation theory, 96–97  
  ionization regimes, 99–100
- ionization temperature, 100  
ionoptical surface, 93–94  
liquid metal ion sources, 31  
nanotechnology applications, 116–117  
overview, 88–91  
real-source current-density distribution, 94  
real-source data, 101–103  
secondary electrons, 92  
source properties, 114–115  
spherical charged particle emitter, 103–108  
  blurred beam, 105  
  field ion microscope resolution, 107  
  FORD minimum value, 108  
  ion energy spread, 108  
  Müller emitter-based machines, 106–107  
  optical model, 103  
  optical object generation, 105–106  
  radial trajectory ions, 103  
  total energy distribution, 108  
  transverse velocity, 103–105  
  transverse zero-point energy, 108  
  virtual source size, 107
- supply-and-capture regime, 100–101  
supply current and effective capture area, 99  
technological development, 92–93  
weak lens effects, 108–112  
  angular magnification, 109–110  
  image-spot size, 111–112  
  Müller emitter source size, 111  
  numerical trajectory analyses, 112  
  projection magnification, 111  
  transverse magnification, 110
- Gas pressure, variable vacuum scanning electron microscope, path optimization, 477–478
- Gas temperature, gas field ionization emitters, 100
- Gaussian emission distribution  
  beam size and resolution, 398–399  
  Boersch effect, 372–373  
    kinetic energy relaxation, 373–374  
  emitter brightness, ZrO/W cathode, 22–23  
  monoenergetic beam current density, rotational symmetry, 415  
  statistical coulomb effect  
    extended two-particle approximation, 359, 363–364  
    nonmicrobeam instruments, 383–385  
    trajectory displacement, 367–368
- Gauss's theorem  
  laminar flow equation, 346  
  space charge effects, ray equation, 344–345
- Geometric-aberration-free condition, Wien filter, 281–286
- Geometric aberrations  
  calculation methods, 212–221  
  correctors, 294–295, 619–625  
  curved optic axes, 265–271  
  defined, 210  
  mirrors and cathode lenses, 259–265  
  quadrupole lenses, 249–252  
  round lenses, 221–249, 254–259  
  sextupole lenses, 252–254  
  superimposed deflection fields, 254–259  
  Wien filter, 271–286

Geometric optics, density-of-information passing capacity, 418–420  
 Geometric size, scanning transmission electron microscope, 504  
 GG cone-jets, liquid metal ion source, 34  
 Ghost charges, Monte Carlo simulation, 376–377  
 Gibbs function, liquid metal ion source  
     electrohydrodynamics, surface pressure jump formula, 44–45  
 Gilbert–Gray (GG) cone-jet  
     research background, 33  
     steady high-electrical-conductivity, 47–50  
         nonturbulent flow, 48–49  
         parameter limits, 48  
         pressure drop, 50  
         pressure relationships, 47–48  
         quasi-ellipsoidal model, liquid cap, 49  
         viscous-loss terms, 49  
     zero-base pressure approximation, 50  
 Globule emission, liquid metal ion sources, 69–70  
 Glow discharge conditioning, electrostatic lens, 188  
 Grid lenses, optical properties, 181–182  
 Gummel number, focused ion beam implantation, 587–591

## H

Hairpin tungsten design, liquid metal ion sources, 73  
 Halo effect, focused ion beam technology, 534–538  
 Heating properties, liquid metal ion sources  
     fabrication, 72  
 Heat transfer, magnetic lens coil design, 138–140  
 Helium gas field ionization source  
     current models, 117  
     gas field ionization emitters, 114–115  
 Helium ion microscopy (HIM), current research on, 488–492  
 Helix detector, variable vacuum scanning electron microscope, 476–477  
 Helmholtz free energy, liquid metal ion source  
     electrohydrodynamics, surface pressure jump formula, 44–45  
 Hexapole planator, multiple aberration correctors, 301–304  
 Hexapole spherical aberration correction, magnetic lens design, 146–148  
 High-angle annular dark-field imaging (HAADF),  
     aberration correction, 611–614  
     optical columns, 633–637  
 Higher-order aberrations, aberration correction, 609–614  
 High-frequency excitation, aberration correction, 308  
 High-voltage focused ion beam implantation, energy loss and, 584–591  
 Holmium, magnetic lens design using, 143–145  
 Holography, aberration correction, 309–310  
 Holtsmark regime  
     Boersch effect, 372–373  
         potential energy relaxation, 373–374  
     focused ion beam technology, 534–538  
     microbeam column design, trajectory displacement, 378–380  
     statistical coulomb effect, nonmicrobeam instruments, 383–385

trajectory displacement, 367–368  
     equations, 368  
 Homogeneous field, lens action, 173–174  
 Hydrocarbon deposition, focused ion beam technology, 566

**I**

Image formation  
     focused ion beam technology, 539–544  
     resolution, charged particle beams, 396–397  
     scanning electron microscope, architecture, 438–439  
 Image-hump mechanism, liquid metal ion sources, field evaporation, 39–40  
 Image plane aberration  
     calculation, 222–246  
     mirrors and cathode lenses, 262–265  
 Image quality, resolution and, 402–406  
     figure of merit, 403–405  
 Image spot size  
     aberration calculation, deflection fields, round lenses, 255–259  
     microbeam column design, trajectory displacement, 377–378  
 Immersion lens  
     action model, 174–175  
     classification, 171  
     low-voltage scanning electron microscope, 154–155  
     objective lenses, 179–181  
     scanning electron microscope architecture, 443–444  
     three-electrode lenses, 179  
     two-electrode immersion lenses, 177–178  
 Implantation techniques, focused ion beam technology, 581–591  
 Incident flux enhancement factor, gas field ionization emitters, Sounth gas-supply theory, 118–122  
 Incomplete Cholesky conjugate gradient (ICCG) method, pole-piece lens design, 133–136  
 Incomplete collisions, trajectory displacement, 365–368  
 Indirect-action aberration correctors, 619–625  
 Inductively coupled plasma (ICP), focused ion beam technology, 593–594  
 Infinitesimal method, differential algebra, 339–340  
 Information passing capacity (IPC). *See also* Density-of-information passing capacity (density-of-IPC)  
     approximation of, 417–423  
     calculated resolution *vs.* experimental results, 423–425  
     fitting functions, 422–423  
     optimum condition  
         diffraction-limited system, 426–428  
         source-limited system, 429–430  
     resolution  
         development of, 394–395  
         research history, 393  
         specimen information size, 433–434  
 Insulator materials and design  
     electrostatic lens, 188–190  
     focused ion beam technology, gas-assisted deposition, 559–566

Integrated circuits, scanning electron microscope inspection, 466–467  
Interaction effects calculation, microbeam column design, 381–382  
Internal energy relaxation, statistical coulomb effect, 354–355  
Ion beams  
    best focus position, 416  
    spatial frequency response, 411–415  
Ion emission theory, liquid metal ion sources (LMIS), 36–44  
    charged-surface models and Maxwell stress, 36–37  
    field evaporation, 37–44  
Ion generation theory, gas field ionization emitters, 96–97  
Ionic liquid ion source (ILIS), liquid metal ion sources, 31  
Ion-induced secondary atoms, focused ion beam technology, scanning ion microscopy, 567–571  
Ion-induced secondary electrons, focused ion beam technology, scanning ion microscopy, 567–571  
Ion-induced secondary ions, focused ion beam technology, scanning ion microscopy, 567–571  
Ionization regimes, gas field ionization emitters, 99–100  
Ion lithography, electrostatic lenses, 199–200  
Ion microscopy, electrostatic lenses, 199–200  
Ion-optical characteristics  
    helium ion microscopy, 489–492  
    liquid metal ion sources, 59–64  
        ion energy distributions, 60–63  
        ion optical effects, 63–64  
Ionoptical surface, gas field ionization emitters, 93–94  
Ion-sample interactions, focused ion beam technology, 525–534  
Iron and iron–cobalt alloys, magnetic lens design, anisotropy constant, 142–145  
Isoplanatic approximation, aberration correction, 613–614  
Isoplanatism patch, image formation, resolution, 396–397  
Isotopic abundance, focused ion beam/secondary-ion mass spectrometry, 579–581  
Isotropic aberration coefficients, spherical aberration, 231–234

## K

Kanaya–Okayama range, scanning electron microscope, electron beam–specimen interaction, 452–454  
Kepler problem, statistical coulomb effect, two-particle dynamics, 353  
Kinetic energy relaxation, Boersch effect, 373–374  
Kinetic-theory equilibrium gas flux density, gas field ionization emitters, theoretical background, 97–98  
Kingham and Swanson model, liquid metal ion sources, apex boundary conditions, 50–51  
Kirk effect, focused ion beam implantation, 586–591  
Köhler illumination, space charge effect, 343  
Krivanek (Nion) aberration corrector, 298–299

## L

Laminar flow  
    current and acceleration voltages, 348–349  
    equation, 345–346  
    space charge defocus and, 344  
Lanthanum hexaboride cathode, scanning electron microscope, 440–443  
Laplace equation  
    aberration correctors, magnetic multipoles, 615–617  
    electrostatic lens properties, 166–168  
    force parameters, steady-state current-related liquid metal ion sources, 52–53  
    liquid metal ion sources, field-emitted vacuum space charge, 43–44  
Laser interferometry, scanning electron microscope, electron beam landing position, 484–485  
Lateral broadening  
    gas field ionization source, 112  
    liquid metal ion sources, ion optical effects, 64  
Lens properties. *See also* Electrostatic lenses; Magnetic lens design, electron microscopy  
    gas field ionization emitters, 108–109  
    scanning electron microscope architecture, 443–446  
Lie algebra, chromatic and geometric aberrations, curved optic axes, 265–271  
Light milling penumbra, liquid metal ion sources, ion optical characteristics, 62  
Linear effect, optical elements, 210  
Linear projection condition, Müller-type angular magnification, 109–110  
Linewidth measurements, scanning electron microscope, 466–467  
Linfoot theory, image resolution quality, 404–405  
Liouville's theorem  
    aberration correctors, magnetic multipoles, 616–617  
    scanning transmission electron microscope, source brightness, 506–508  
Liquid alloy ion source (LAIS), defined, 32  
Liquid cap model, liquid metal ion sources, quasi-ellipsoidal model, 49  
Liquid cone formation, liquid metal ion source, 34  
Liquid metal alloy ion source (LMAIS), 32, 66–67  
Liquid metal ion sources (LMIS)  
    alternative geometries, 74  
    electrohydrodynamics theory, 44–51  
        apex boundary conditions, 50–51  
        electrospraying, 35  
        equilibrium and stability conditions, 45–46  
        steady high-electrical-conductivity Gilbert–Gray cone-jet, 47–50  
        surface pressure jump, 44–45  
        Taylor's mathematical cone, 46–47  
    emitter shape, 34  
    energy distribution and ion-optical characteristics, 59–64  
        ion energy distributions, 60–63  
        ion optical effects, 63–64  
    fabrication process, 71–73  
    field ion emission sources, 31  
    focused ion beam/secondary-ion mass spectrometry, 572–581  
    focused ion beam technology, 524–525

- Liquid metal ion sources (LMIS) (*contd.*)  
     applications, 549–593  
     architectural characteristics, 534–538  
     future applications, 593–594  
     image formation, 539–544  
     future research issues, 74–756  
     ion emission theory, 36–44
- Liquid metal ion sources  
     charged-surface models and Maxwell stress, 36–37  
     field-emitted vacuum space charge, 42–44  
     field evaporation, 37–42  
     ionization mechanisms and emitted species, 65–67  
         alloy ion sources, 66–67  
         elemental ion sources, 65–66  
     overview, 31  
     research background, 32–33  
     source operation, 73–74  
     steady-state current-related characteristics, 51–59  
         current-voltage characteristic above extinction, 54–57  
         cusp length emission current dependence, 57  
         extinction and collapse voltages, 54  
         extinction current, 57–58  
         physical chaotic attractor applications, 58–59  
         theoretical formulation, 52–54  
     terminology and conventions, 33–34  
     time-dependent behavior, 35–36, 67–71  
         droplet emission, upper unsteady regime, 68–69  
         globule emission, needle and cone, 69–70  
         liquid-shape numerical modeling, 71  
         pulsation, 67–68  
         source turn-on, 70–71
- Liquid-shape development, liquid metal ion sources, numerical modeling, 71
- Lithography  
     focused ion beam implantation, 591–593  
     space charge effect, 343
- Load characteristics, chromatic and geometric aberration, 291–292
- Local angular distortion, gas field ionization source, 113
- Longitudinal ion distribution, focused ion beam technology, 528–534
- Long-range polarization potential-energy well, gas field ionization emitters, 95
- Long-term current drift, ZrO/W cathode current fluctuations, 23–24
- Lorentz distribution  
     Boersch effect, 372–373  
     statistical coulomb effect  
         beam segments, 365  
         extended two-particle approximation, 363–364
- Lotus root lens, multibeam electron lithography system, 152–153
- Low accelerating voltage inspection, scanning electron microscope, 466–468
- Low-energy electron microscope (LEEM)  
     electrostatic lens, 164–166, 194–196  
     design properties, 185–186  
     hexapole spherical aberration correction, 147–148  
     historical background, 131–132  
     objective lens design, 150–152
- Low-energy tail, liquid metal ion sources, ion optical characteristics, 62
- Low-vacuum microscopy, evolution of, 472
- Low-voltage scanning electron microscope  
     electrostatic/magnetic lens design, 156–157  
     objective lens design, 153–156
- Low-voltage transmission electron microscopy, electrostatic lenses, 194
- Low-work-function plane, emitter life mechanisms, 26–27
- M**
- Macroscopic surface tension, ZrO/W cathode shape stability, 18–19
- Magnetic circuit design, microscopic lenses, 136–141
- Magnetic field  
     aberration calculation, 212–221  
     resolution limits and stray fields, 434
- Magnetic lenses  
     chromatic and geometric aberrations, 221–249  
         anisotropic components, 244–246  
         minimization, 296–297  
     electron microscopy  
         aberration corrections, 145–150  
         coil design, 136–140  
         dodecapole magnetic field component generation, 145  
         electrostatic/magnetic quadrupole lens, chromatic aberration corrector, 148–150  
         hexapole spherical aberration corrector, transfer doublet, 145–148  
         lens materials, magnetic homogeneity, 143–145  
         lotus root lens, multibeam electron lithography, 152–153  
         low-energy electron microscope/photoelectron emission microscope lens design, 150–152  
         low-voltage scanning electron microscope lens, 153–157  
         magnetic circuit, 136–141  
         pole pieces, 132–136, 140–141  
         research background, 129–132  
         saturation magnetic flux densities, 141–143  
         scanning transmission electron microscope, 515
- Magnetic materials properties, magnetic lens design, 141–145  
     homogenous magnetic properties, 143–145
- Magnetic multipoles, aberration correctors, 614–617
- Magnetic saturation, pole-piece lens design, 134–136
- Magnetic sector, focused ion beam/secondary-ion mass spectrometry, 571–581
- Mahoney's equation, steady-state current-related liquid metal ion sources, 55–56
- Mair's equation, steady-state current-related liquid metal ion sources, 54–55
- Manufacturing process, electrostatic lens, 190–191
- Mask repair, focused ion beam technology, 549  
     gas-assisted deposition, 559–566
- Mass interference, focused ion beam/secondary-ion mass spectrometry, 579–581
- Mass spectrum acquisition, focused ion beam/secondary-ion mass spectrometry, 575–581
- Material-specific parameters

- focused ion beam technology, gas-assisted deposition, 559–566  
steady high-electrical-conductivity Gilbert–Gray cone-jet, 48
- Matrix desorption, electrostatic lenses, 172
- Matrix effect, focused ion beam/secondary-ion mass spectrometry, 575–581
- Matrix multiplication, aberration corrector optics, 605–606
- Matrizant, chromatic aberration calculation, 220–221
- Maxwell–Boltzman statistics, Boersch effect, potential energy relaxation, 373–374
- Maxwell stress  
liquid metal ion sources, 36–37  
field-emitted vacuum space charge, 43–44  
normalized apex, 42, 49  
surface pressure jump formula, 44–45
- steady-state current-related liquid metal ion sources, 53–54
- MCP detector, focused ion beam technology, scanning ion microscopy, 570–571
- Mean information content, resolution, image quality and, 403–405
- Mean probability of capture, gas field ionization emitters, Southon gas-supply theory, 118
- Mean square field fluctuation approximation, statistical coulomb effect, 357–358
- Measurement techniques, resolution, 431–433
- Medium-angle annular dark field (MAADF), aberration correction, optical columns, 633–637
- Message state, image resolution quality, 403–405
- Metals deposition, focused ion beam technology, gas-assisted deposition, 559–566
- Method of limits, differential algebra, 339–340
- m-fold astigmatism of n-th order, aberration correction, 609–614
- Microbeam column design, statistical coulomb effect, 377–382  
Boersch effect, 381  
interaction calculations, 381–382  
statistical interaction minimization, 381  
trajectory displacement and probe size, 377–380
- Microchannel plate (MCP) detector, focused ion beam technology, 547–548
- Microdischarges, electrostatic lens, insulator materials and design, 189–190
- Micromachining, focused ion beam technology, 549–559
- Microscopic column, variable vacuum scanning electron microscope, 473–474
- Minimization  
aberration correctors, 296–297  
microbeam column design, statistical interactions, 381
- Minimum steady emission current, steady-state current-related liquid metal ion sources, 51–52
- Misalignment aberrations  
aberration corrector operation, 626–626  
aberration correctors, 618–619
- Mixed quadrupole correctors, chromatic and geometric aberrations, 301
- Monochromatic beams, statistical coulomb effect, 350–352
- Monoenergetic beam  
current density, spatial frequency response, 411–412  
optical transfer function, 407–409
- Monte Carlo simulation  
analytical theory vs., 385–386  
focused ion beam technology, 537–538  
limitations of, 385–386  
microbeam column design, Boersch effect, 381  
particle-particle interaction, 375–377  
scanning electron microscope, signal modeling, 455–456
- space charge effect, 343  
statistical coulomb effect  
first-order perturbation models, 356–357  
nonmicrobeam instruments, 384–385
- Most-steady regime, liquid metal ion sources, time-dependent behavior, 35
- Müller emitter design  
angular magnification, 109–110  
definitions and illustrative values, 116–117  
gas field ionization emitters, 89–91  
magnification properties, 116–117  
numerical trajectory analyses, 112  
projection magnification, 111  
source properties, 114–115  
source sizes, 111  
spherical charged particle emitter, 103–107  
total captured flux, Southon gas-supply theory, 120–122  
weak lens, 108–109
- Müller escape field, liquid metal ion sources, field evaporation, 39–40
- Multibeam electron lithography, lotus root lens, 152–153
- Multielectrode lenses, computation of, 178
- Multielement lenses, 179
- Multiplicity effects, aberration correction, 608–614
- Multipole field components  
aberration correctors, 301–304  
single multipole effects, 614–617  
dodecapole generation, 145–146
- N**
- Nanotechnology  
extra high-resolution scanning electron microscopy, 492–493  
gas field ionization emitters, 117  
liquid metal ion source applications, 32  
surface pressure jump formula, 45
- Nearest-neighbor approximation, statistical coulomb effect, nonmicrobeam instruments, 383–385
- Nebulosity, gas field ionization emitters, gas-pressure trade-off, 91–92
- Needle shapes, gas field ionization emitters, 93
- Needle wetting, liquid metal ion sources fabrication, 71–72
- Net plane collapse, emitter environmental requirements, 24–25
- Newton–Raphson method, pole-piece lens design, 134–136
- Noise limits  
focused ion beam technology, 543–544  
resolution, scanning electron microscope, 481–483
- Noncrossover mode thermionic cathodes, 1–3  
extended Schottky regime, 3–4

- Nondestructive inspection, scanning electron microscope, 450–451
- Nonmicrobeam instruments, statistical coulomb effects, 382–385
- Normalized apex Maxwell stress, liquid metal ion sources, field evaporation, 42
- N*-particle problem
- Monte Carlo simulation, 376–377
  - space charge effect, extended two-particle approximation, 358–359
  - statistical coulomb effect, 352–353
  - closest encounter approximation, 357
  - extended two-particle approximation, 359–364
- Nth-order aberration, defined, 211
- Nuclear losses, focused ion beam technology, ion-sample interactions, 527–534
- Numerical data
- field ion microscopy, trajectory analyses, 112
  - fitting functions vs., 423
  - liquid metal ion sources, 76–80
  - Southon gas-supply theory, 120–122
  - trajectory displacement, 368
- Numerical modeling, liquid metal ion sources, liquid-shape development, 71
- ## O
- Objective lens design
- aberration correctors, magnetic multipoles, 615–617
  - electrostatic lenses, 171
  - historical background, 130–131
  - immersion lenses, 179–181
  - low-energy electron microscope/photoelectron emission microscope, 150–152
  - low-voltage scanning electron microscope, 153–156
- Object plane, chromatic and geometric aberrations, curved optic axes, 268–271
- Octopoles
- aberration correction, 619–625, 620–625
  - aberration correctors, 294–295, 297–298
  - aberration data
    - excitation parameters, 251–252
    - Wien filter, 285–286
- Onset energy deficits, liquid metal ion sources, ion optical characteristics, 60–61
- Onset voltage, steady-state current-related liquid metal ion sources, 51–59
- Optical beam-induced current (SBIC), focused ion beam technology, 548–549
- Optical column properties
- aberration correction, 631–637
  - scanning transmission electron microscope, 517–518
- Optical model, gas field ionization source, spherical charged particle emitter, 103–108
- Optical source size, liquid metal ion sources, ion optical effects, 64
- Optical transfer function (OTF)
- monoenergetic beam current density, rotational symmetry, 414–415
  - resolution, contrast performance and, 401–402

- scanning transmission electron microscope, focused probe optimization, 509–510
- spatial frequency response
- chromatic aberration, 409–410
  - electron beam, 406–410
  - monoenergetic beam, 407–409
- Optimization
- chromatic and geometric aberration minimization, 296–297
  - electrostatic lenses, 183–185
  - microbeam column design, trajectory displacement, 378–380
  - scanning electron microscope operations, 449
- Optimum condition, image resolution, 425–430
- diffraction-limited system, 426–428
  - source-limited system, 428–430
- Overshoot/decay-back effect, liquid metal ion sources, 70–71
- ## P
- Parameter dependencies, trajectory displacement, 365–368
- Parasitic aberrations
- basic properties and notations, 292–294
  - defined, 210, 601–602
  - electrostatic lens manufacturing, 190–191
  - proof-of-principle correctors, 604
- Paraxial approximation
- aberration calculation, 211–221
  - chromatic and geometric aberrations, 222–246
  - chromatic and geometric aberrations
    - mirrors and cathode lenses, 259–265
    - quadrupole lenses, 249–252
    - sextupole lenses, 252–254
    - Wien filter, 271–286
- Particle metrology, scanning electron microscope, 467–468
- Path equations, chromatic and geometric aberrations, curved optic axes, 265–271
- Path optimization, variable vacuum scanning electron microscope, 477–478
- P/D curves, two-electrode immersion lenses, 177–178
- Peak switching, focused ion beam/secondary-ion mass spectrometry, 576–581
- Pencil beam condition
- Boersch effect, 364
  - equations, 372–373
  - weak collisions, 370–372
  - statistical coulomb effect
    - nonmicrobeam instruments, 383–385
    - two-particle dynamics, 353
    - trajectory displacement, 367–368
    - equations, 368
- Permalloy, magnetic lens design using, 143–145
- Permendur, magnetic lens design
- homogeneous properties, 143–145
  - saturation magnetic flux density, 141–142
- Perturbation theory, statistical coulomb effect, first-order perturbation models, 355–357
- Phase condition, aberration correction, 308–309

- Photoelectron emission microscope (PEEM)  
aberration correction, 305  
electrostatic lens, 164–166, 194–196  
historical background, 131–132  
objective lens design, 150–152  
working correctors for, 605
- Picht transformation  
all-electrostatic aberration correctors, 300–301  
electrostatic lens evaluation, 169–170  
aberrations, 176
- Picoprober, nanotechnology and development of, 117
- Pinhole lens technology, scanning electron microscope architecture, 443–446
- Plasma physics, statistical coulomb effect, 354–355
- Plies-Typke theory, chromatic and geometric aberrations, Wien filter, 275–286
- Point-by-point representation, scanning electron microscope architecture, 439–440
- Point cathode electron sources, scanning electron microscope, 440–443
- Point spread function (PSF)  
aberration correction, 613–614  
resolution research, 392–393
- Poisson statistics, focused ion beam technology, 543–544
- Poisson surface field, liquid metal ion sources, field-emitted vacuum space charge, 42–44
- Polarization potential-energy (PPE) well, gas field ionization emitters, 88–89
- Pole and yoke design, magnetic lenses, 140–141
- Pole-piece lens  
design procedures, 132–136  
historical background, 130–131  
homogeneous magnetic properties of lens materials, 143–145
- Post-field ionization, liquid metal ion sources, 37
- Potential energy relaxation, Boersch effect, 373–374
- Power supply, magnetic lens coil design, 139–140
- Preaccelerated ions, focused ion beam implantation, 583–591
- Pre-electron energy loss spectrometer (pre-EELS), aberration correction, 633–637
- Pressure drop, steady high-electrical-conductivity Gilbert-Gray cone-jet, 50
- Pressure relationships  
gas field ionization emitters, gas-pressure trade-off, 91–92  
steady high-electrical-conductivity Gilbert-Gray cone-jet, 47–48
- Primary electrons  
image formation, resolution, 396–397  
scanning electron microscope, architecture, 438–439  
variable vacuum scanning electron microscope, gas effects on, 474–475
- Probe current-probe size interaction, microbeam column design, 382
- Probe size  
microbeam column design  
interaction effects, 382  
trajectory displacement, 377–380  
resolution research, 392–393
- Projection ion lithography  
electrostatic lenses, 200  
statistical coulomb effects, 382–385
- Projection magnification, Müller emitter design, 111
- Projective lens design, electrostatic lenses, 171
- Proof-of-principle aberration correctors, basic properties, 603–604
- Proportionality, aberration correctors, magnetic multipoles, 617
- Proximity effect, focused ion beam implantation, 592–593
- Pseudorandom number generation, Monte Carlo simulation, 375–377
- Pulsation regime  
liquid metal ion sources, time-dependent behavior, 35, 67–68  
upper unsteady regime, 36
- Pupil function, optical transfer function, monoenergetic beam, 407–409
- Q**
- Quadrupole lenses  
aberration correction, 620–625  
optical columns, 633–637  
chromatic and geometric aberration, 249–252  
all-electrostatic correctors, 300–301  
correctors, 294–295, 297–298  
Russian quadruplet representation, 291–292  
chromatic and geometric aberrations, mixed aberration correctors, 301  
electrostatic/magnetic chromatic aberration corrector, 148–150  
proof-of-principle aberration correctors, 603–604  
working correctors, 604–605
- Qualitative voltage contrast, scanning electron microscope, 468–470
- Quantitative voltage contrast, scanning electron microscope, 470
- Quasi-ellipsoidal model, liquid metal ion sources, 49
- R**
- Radial-gap lens, low-voltage scanning electron microscope, 155–156
- Radial trajectory ions, gas field ionization source, spherical charged particle emitter, 103
- Radiofrequency (RF) quadrupole systems, focused ion beam/secondary-ion mass spectrometry, 571–581
- Radius aberration  
gas field ionization source, 113–114  
mirrors and cathode lenses, 262–265
- Rate constant, liquid metal ion sources, field evaporation, 37–38
- Rate of decrease equation, ZrO/W cathode shape stability, 16–19
- Ray equation  
aberration correctors, sextupole lenses, 622–623  
Monte Carlo simulation errors, 376–377  
space charge effects, 344–345

- Rayleigh's criterion, resolution  
 charged particle image formation, 397  
 contrast performance and, 399–402  
 density-of-information passing capacity, 405–406  
 origins of concept, 395–396  
 point-to-point measurement, 431–433  
 scanning electron microscope, 481–483
- Reactive ion etching, focused ion beam technology, 545–549
- Real part, defined, 340
- Real-source current-density distribution, gas field ionization emitters, 94
- Real-source data, gas field ionization emitters, 101–103
- Real-time image analysis and processing, scanning electron microscope, 449
- Redeposition mechanisms, focused ion beam technology, micromachining and gas-assisted etching, 551–559
- Refractive index, chromatic and geometric aberrations, Wien filter, 277–286
- Relativistic correction, electrostatic lens evaluation, 169–170
- Representation, aberration data, 286–290
- Repulsion distance, liquid metal ion sources, charge-surface models and Maxwell stress, 37
- Reservoir properties, liquid metal ion sources fabrication, 72
- Resolution  
 basic principles, 393–394  
 beam size definition, 398–399  
 charged particle image formation, 396–397  
 contrast performance, optical systems, 399–402  
 density-of-information passing capacity  
     approximation methods, 416–423  
     fitting functions, 422–423  
     wave/geometric optics, 418–422  
 best focus position, 415–416  
 calculation vs. experimental results, 423–425  
 current density, monoenergetic beam, 411–412  
 image quality, 402, 405–406  
 ion beam spatial frequency response, 411  
 optical transfer function  
     chromatic aberration and, 409–410  
     monoenergetic beam, 407–409  
 rotationally symmetric systems, monoenergetic beam current density, 412–415  
 spatial frequency response  
     electron beam, 406–410  
     ion beam, 411–415  
 two-dimensional Fourier transform, source intensity distribution, 406–407
- focused ion beam technology, limitations, 540–544
- image quality, 402–406  
 figure of merit, 403–405  
 limitations, 406  
 optical image, 402–403  
 Rayleigh's criterion and density-of-information passing capacity, 405–406
- limitations, 433–434  
 measurement, 431–433  
 optimization and attainable resolution, 425–430
- diffraction-limited system, 426–428  
 source-limited system, 428–430
- Rayleigh's criterion, 395–396
- research background, 392–393
- scanning electron microscope  
 architecture, 438–439  
 image resolution, 480–483  
 measurements, 478–480
- scanning transmission electron microscope, focused probe optimization, 508–510
- Retarding lenses, scanning electron microscopes, 196–199
- Richardson–Dushman equation, Schottky emission, 5–6
- Ronchigram formation, aberration correction and diagnosis, 628–631
- optical columns, 633–637
- Root-mean-square (RMS) value  
 Boersch effect, kinetic energy relaxation, 373–374  
 focused ion beam technology, 542–544  
 scanning transmission electron microscope, focused probe optimization, 509–510
- statistical coulomb effect, 351–352  
 closest encounter approximation, 357  
 plasma physics, 354–355
- Rose gauge  
 aberration calculation, 216–221  
 chromatic and geometric aberrations, Wien filter, 275–286
- Rose ultracorrector, 298
- Rotating coordinates  
 chromatic aberration calculation, 217–221  
 deflection fields, round lenses, 255–259
- Rotational symmetry  
 aberration calculation, 217–221  
 monoenergetic beam current density, 412–415  
 resolution, contrast performance and, 400–402
- Round lenses, aberrations, 210  
 geometric/chromatic coefficients, 221–249, 254–259  
 minimization, 296–297  
 representation, 286–290  
 superimposed deflection fields, 254–259
- Runge–Kutta–Fehlberg numerical integration, extended Schottky emission, 6–8
- Russian quadruplet, aberration data representation, 291–292
- Rutherford backscattering (RBS), focused ion beam implantation, 585–591

## S

- Sample preparation, scanning electron microscope, 449–452
- Saturation magnetic flux densities, magnetic lens design, 141–143
- Scaled linear particle density, statistical coulomb effect, two-particle dynamics, 353
- Scanning electron microscopy (SEM)  
 applications, 466–478  
     electron beam induced current, 470  
     failure analysis, 470–472  
     linewidth measurements, 466–467

- low accelerating voltage inspection, 466–467  
particle metrology, 467–468  
variable vacuum SEM, 472–478  
voltage contrast, 468–470
- architecture, 438–449  
cold field emission cathodes, 441–442  
digital image storage and analysis, 446–449  
electron source types, 440–443  
extended field lens technology, 444–446  
immersion lens technology, 443–444  
lens design, 443–446  
operating conditions optimization, 449  
point cathode electron source types, 440–443  
real-time image analysis, 449  
thermally assisted field emission, 442–443  
TV rate scanning, 448–449
- chromatic and geometric aberrations, 224–246  
electron beam-specimen interactions, 452–456  
electron range, 452–455  
signal modeling, 455–456
- electrostatic lens, 164–166  
compound and retarding lenses, 196–199  
extra high-resolution SEM, 492–493  
future research issues, 493
- historical background, 131–132, 438
- imaging and measurement accuracy, 478–487  
electron beam induced contamination, 485–487  
electron beam landing position, 483–485  
image resolution, 480–483  
resolution measurements, 478–480
- objective lens design, low-voltage scanning electron microscope, 153–156
- resolution research, 392–393  
gap resolution measurement, 431–433  
imaging and measurement accuracy, 478–483  
specimen information size, 433–434
- sample preparation and size, 449–452  
nondestructive inspection, 450–451  
total electron emission, 451–452
- signal properties, 456–465  
absorbed electrons, 463–464  
Auger electrons, 464–465  
backscattered electrons, 459–461  
cathodoluminescence, 464  
secondary electrons, 456–459  
transmitted electrons, 465  
x-rays, 461–463
- variable vacuum SEM, 472–478  
charge control, 478  
column properties, 473–474  
contamination reduction, 478  
gas effects, primary electron beam, 474–475  
gas pressure and path optimization, 477–478  
secondary electron signal detection in gas, 475–477
- ZrO/W Schottky cathode, 27
- Scanning ion microscopy (SIM), focused ion beam technology, 539  
applications, 549  
procedures and properties, 567–571
- Scanning transmission electron microscopy (STEM)  
aberration correction, 294–295  
aberration function concept, 607–614
- diagnostic methods, 628–631  
optical columns, 633–637  
overview, 602–604  
quadrupoles and octopoles, 297–298
- annular dark-field imaging, 501–502  
applications, 510–511  
basic principles, 499–500  
bright-field imaging, 500–501  
chromatic and geometric aberrations, 224–246, 505  
focused probe optimization, 508–510
- design principles, 518  
detection modes and signals, 503  
diffraction, 504–505  
diffraction plus chromatic aberration, 510  
diffraction plus spherical aberration, 509–510  
electron-optical column components, 511–518  
cold field emitter source, 512  
design and performance, 515–517  
electron gun, 513–515  
electron source, 511  
magnetic lenses, 515  
thermally assisted, zirconium-treated field emitter source, 513
- electron probe formation, 503–506  
focused probe optimization, 508–510  
geometric size, 504
- historical background, 131–132
- immersion lens technology, 443–444
- research background, 498–499
- source size, brightness and, 506–508
- spectroscopy applications, 502
- spherical aberration, 505–506  
correction, 518–519  
transmitted electrons, 465  
working correctors for, 605
- Scherzer's theorem  
chromatic and geometric aberrations, 225–246  
proof-of-principle correctors, 603–604
- Schottky emitter (SE)  
basic equations, 5–6  
field factor  $\beta$ , emitter radius, and work function, 8–10  
microbeam column design, trajectory displacement, 377–380
- scanning electron microscope, 441–443  
scanning transmission electron microscope, 513
- ZrO/W cathode, 3–5  
total energy distribution, 20–22
- Schottky point cathodes, historical background, 1–2
- Scintillator detector, focused ion beam technology, scanning ion microscopy, 569–571
- Secondary electrons  
focused ion beam technology, 533–534  
gas field ionization emitters, 92  
helium ion microscopy, 488–492  
liquid metal ion sources, 74  
scanning electron microscope  
absorbed electrons, 464  
collection of, 457–461  
electron beam-specimen interaction, 452–454  
immersion lens technology, 443–444  
signal properties, 456–459
- variable vacuum scanning electron microscope, gas signal detection, 475–477

- Secondary-ion images, focused ion beam technology, 545–549
- Secondary-ion mass spectrometry, focused ion beam technology, 571–581
- Self-annealing, focused ion beam implantation, 585–591
- Semiconductor thin films, focused ion beam implantation, 587–591
- Sextupole lenses
- aberration correctors, 298–300, 621–625
  - aberration data, 252–254
  - Wien filter, 285–286
- Shadow method, electrostatic lens evaluation, 167–168
- Shaped beam lithography instruments, statistical coulomb effects, 382–385
- Shape stability, ZrO/W cathode emitter geometry, 12–19
- Sharp-cutoff-(SCOFF) approximation
- aberration correctors
  - magnetic multipoles, 615–617
  - sextupole lenses, 622–623
  - chromatic and geometric aberrations
  - curved optic axes, 267–271
  - Wien filter, 271–286
- Short-term frequency, ZrO/W cathode current fluctuations, 23–24
- Side pole-gap lens, low-voltage scanning electron microscope, 154–155
- Signal power/noise (P/N) ratio
- density-of-information passing capacity, wave and geometric optic equivalence, 418–420
  - resolution, image quality, 404–405
- Signal properties
- helium ion microscopy, 490–492
  - scanning electron microscope, 456–465
  - absorbed electrons, 463–464
  - Auger electrons, 464–465
  - backscattered electrons, 459–461
  - cathodoluminescence, 464
  - modeling, 455–456
  - secondary electrons, 456–459
  - transmitted electrons, 465
  - x-rays, 461–463
  - scanning transmission electron microscope, 503
- Signal-to-noise ratio (S/N)
- density-of-information passing capacity
  - calculated resolution *vs.* experimental results, 423–425
  - wave and geometric optic equivalence, 418–420
- focused ion beam technology
- gas-assisted etching, 557–559
  - image formation, 539–544
- optimum condition, source-limited system, 429–430
- resolution
- image quality, 404–405
  - optimum condition, diffraction-limited system, 426–428
- resolution research, 392–393
- contrast performance and, 400–402
- Silicon drift detector (SDD), scanning electron microscope, x-ray collection, 462–463
- Single multipole effects, aberration correctors, 614–617
- Slater–Pauling curve, saturation magnetic flux density, magnetic lens design, 142–145
- Slender-body approximation, steady-state current-related liquid metal ion sources, 53–54
- Slit plane, chromatic and geometric aberrations, curved optic axes, 268–271
- Slow-scan operations, scanning electron microscope, 446–448
- SMART project
- aberration correctors, 305
  - scanning electron microscope resolution, 482–483
- Smith–Helmholtz formula, transverse magnification, 110
- Snorkel lens, scanning electron microscope architecture, 444–445
- Source brightness
- optimum condition, 428–430
  - scanning transmission electron microscope, 506–508
- Source intensity distribution, two-dimensional Fourier transform of, electron beam, 406–407
- Source-limited system, optimum condition, 428–430
- Source properties
- gas field ionization source, 114–115
  - monoenergetic beam current density, rotational symmetry, 414–415
- Source turn-on, liquid metal ion sources, 70–74
- Southon gas-supply theory, gas field ionization emitters, 117–122
- Space charge effect
- aberration, 305–306, 347–348
  - defocus conditions, 346–347
  - laminar flow, 344–346
  - numerical examples, 348–349
  - ray equation, 344–345
  - gas field ionization source, 112
  - overview, 342
- Space intensity distribution, electron beam, two-dimensional Fourier transform, 406–407
- Spark conditioning, electrostatic lens, 188
- Spatial frequency
- density-of-information passing capacity
  - electron beam, 406–410
  - ion beams, 411–415
  - resolution, contrast performance and, 400–402
  - statistical coulomb effect, extended two-particle approximation, 360–364
- Spatial resolution, scanning electron microscope, 480–483
- Specimen current imaging, scanning electron microscope, 464
- Specimen-electron beam interaction, scanning electron microscope, 452–456
- Specimen size, scanning electron microscope, 449–452
- Spectral density function, ZrO/W cathode current fluctuations, current stability, 23–24
- Spectroscopic applications, scanning transmission electron microscope, 502
- Sphere-on-orthogonal-cone (SOC) model, numerical trajectory analyses, 112
- Spherical aberration
- aberration correction
  - aberration function concept, 608–614
  - direct-action correctors, 619–620
  - best focus position, 415–416
  - calculation, 223–246

- density-of-information passing capacity, wave and geometric optic equivalence, 418–420  
gas field ionization emitters, 113  
microbeam column design, trajectory displacement, 377–378  
monoenergetic beam current density, rotational symmetry, 413–415  
resolution, beam size and, 398  
scanning transmission electron microscope, 505–506  
correction, 518–519  
diffraction and, 509–510  
working correctors, 604–605
- Spherical charged particle emitter (SCPE)  
definitions and illustrative values, 116–117  
gas field ionization source, 103–108  
blurred beam, 105  
field ion microscope resolution, 107  
FORD minimum value, 108  
ion energy spread, 108  
Müller emitter-based machines, 106–107  
optical model, 103  
optical object generation, 105–106  
radial trajectory ions, 103  
total energy distribution, 108  
transverse velocity, 103–105  
transverse zero-point energy, 108  
virtual source size, 107  
magnification properties, 116–117
- Spherical coordinates with increasing mesh (SCWIM)  
model, ZrO/W cathodes, 8–10  
angular intensity/extraction voltage relationships, 10–12
- Spread function, resolution, Rayleigh's criterion, 395–396
- Sputter yield, focused ion beam technology, 530–534  
focused ion beam/secondary-ion mass spectrometry, 574–581  
micromachining and gas-assisted etching, 549–559
- Square root dependency, statistical coulomb effect, mean square field approximation, 358
- Stability conditions, liquid metal ion source, 45–46
- Stage drift, scanning electron microscope, 485
- Star of Mercedes, aberration correction and diagnosis, 630–631
- Statistical coulomb effect  
beam segment displacement, 364–365  
closest encounter approximation, 357  
extended two-particle approximation, 358–359  
displacement distribution, 359–364  
first-order perturbation models, 355–357  
focused ion beam technology, 534–538  
mean square field fluctuation approximation, 357–358  
microbeam column design, 377–382  
Boersch effect, 381  
interaction calculations, 381–382  
statistical interaction minimization, 381  
trajectory displacement and probe size, 377–380  
nonmicrobeam instruments, 382–385  
*N*-particle reduction, 352–353  
plasma physics models, 354–355  
problem formulation, 349–352  
two-particle dynamics, 353–354
- Statistical interaction minimization, microbeam column design, 381
- Steady flow properties, liquid metal ion source, 33–34  
Steady high-electrical-conductivity Gilbert–Gray cone-jet, liquid metal ion source, 47–50  
nonturbulent flow, 48–49  
parameter limits, 48  
pressure drop, 50  
pressure relationships, 47–48  
quasi-ellipsoidal model, liquid cap, 49  
viscous-loss terms, 49  
zero-base pressure approximation, 50
- Steady-state current-related characteristics, liquid metal ion sources, 51–59  
current-voltage characteristic above extinction, 54–57  
cusp length emission current dependence, 57  
extinction and collapse voltages, 54  
extinction current, 57–58  
physical chaotic attractor applications, 58–59  
theoretical formulation, 52–54
- Steady-state liquid ion molecule sources, material-specific parameters, 48
- Steady-state liquid metal ion sources, apex boundary conditions, 51
- Stigmatic focusing, chromatic and geometric aberrations, Wien filter, 271–286
- Superaplanator  
aberration corrections, 620–625  
multiple aberration correctors, 303–304  
third-generation correctors, 605
- Supply-and-capture regime, gas field ionization emitters, 100–101
- Supply current density, gas field ionization emitters, 99
- Supply limit, liquid metal ion sources, field evaporation, 38–39
- Surface analysis, electrostatic lenses, 1989
- Surface diffusion-based flicker noise, ZrO/W cathode current fluctuations, 23–24
- Surface pressure jump formula, liquid metal ion source electrohydrodynamics, 44–45
- Surface-tension stress, liquid metal ion sources, material parameters, 49
- Switchyard deflector, focused ion beam/secondary-ion mass spectrometry, 572–581
- Symmetry, aberration data, 290–292
- Symplectic condition, chromatic and geometric aberrations, curved optic axes, 265–271

## T

- $\tau$ -variation method, chromatic aberration calculation, 220–221
- Taylor cones  
liquid metal ion source  
electrohydrodynamics theory, 46–47  
formation of, 34  
steady-state current-related liquid metal ion sources, 53–54
- Taylor expansion, statistical coulomb effect, extended two-particle approximation, 361–364
- Temperature effects  
liquid metal ion sources, ion optical characteristics, 62–63  
steady-state current-related liquid metal ion sources, 57

- Thermal annealing, ZrO/W cathode, 4–5
- Thermally assisted field emission (TFE) cathodes
- scanning electron microscope, 440–443
  - scanning transmission electron microscope, zirconium-treated field emitter source, 513
- Thermal spike model, focused ion beam technology, gas-assisted deposition, 561–566
- Thermionic emission cathodes, scanning electron microscope, 440–443
- Thermodynamic equilibrium limit
- Boersch effect, 373–374, 385–386
  - liquid metal ion sources, field evaporation, 38–39
- Thick lens design, electrostatic lens, 170–171
- Third-generation correctors, 605
- Third-order geometric aberrations
- deflection fields, round lenses, 257–259
  - Wien filter, 282–286
- Thomson's quadrupole–octupole corrector, 620–625
- Three-dimensional beam
- Boersch effect, potential energy relaxation, 373–374
  - scanning electron microscope, resolution measurements, 478–480
- Three-electrode unipotential lenses, 178–179
- Through-the-lens detection, immersion lens technology, 443–444
- Tilted image displacement, aberration diagnosis, 627–631
- Time constant, liquid metal ion sources, field evaporation, 37–38
- Time-dependent behavior, liquid metal ion sources, 35–36, 67–71
- droplet emission, upper unsteady regime, 68–69
  - globule emission, needle and cone, 69–70
  - liquid-shape numerical modeling, 71
  - pulsation, 67–68
  - source turn-on, 70–71
- Time intervals, statistical coulomb effect, mean square field approximation, 358
- Time of flight (TOF), focused ion beam/secondary-ion mass spectrometry, 571–581
- Total captured flux, Müller emitter design, Sauthon gas-supply theory, 120–122
- Total electron emission, scanning electron microscope, 451–452
- Total energy distribution (TED)
- extended Schottky emission, 8
  - gas field ionization emitters, 108
  - ZrO/W cathode, 19–22
- Total system aberrations, correction of, 626–627
- Trajectory calculations, aberration corrector optics, 605–606
- Trajectory displacement
- equation summary, 368
  - limitations of, 385–386
  - liquid metal ion sources, ion optical effects, 64
  - microbeam column design, 377–380
  - interaction effects, 382
  - Monte Carlo simulation, 375–377
  - numerical examples, 369–370
  - overview, 342–343
  - parameter dependencies, weak/incomplete collisions, 365–368
  - statistical coulomb effect, 350–352
  - first-order perturbation models, 355–357
- mean square field approximation, 357–358
- nonmicrobeam instruments, 383–385
- two-particle approximation, 343
- Trajectory equation, electrostatic lens evaluation, 168–170
- Trajectory method, aberration calculation, 218–221
- Transfer doublet, hexapole spherical aberration correction, 146–148
- Transmission electron microscopy (TEM)
- aberration correction, 294–295
  - aberration function concept, 607–614
  - diagnostic methods, 627–631
  - overview, 602–604
  - chromatic and geometric aberrations, 224–246
  - all-electrostatic aberration correctors, 301
  - electrostatic lenses, 192–194
  - overview, 162–166
  - focused ion beam technology, 558–559
  - historical background, 130–132
  - objective lens design, low-voltage scanning electron microscope, 153–156
  - resolution research, 392–393
  - scanning transmission electron microscope modifications, 518
  - ZrO/W Schottky cathode applications, 27
- Transmission of digital images, scanning electron microscope, 449
- Transmitted electrons, scanning electron microscope, 465
- Transmitted intensity, resolution measurement, 432–433
- Transport of ions in matter (TRIM), focused ion beam technology, 527–534
- gas-assisted deposition, 561–566
- Transverse magnification
- Müller emitter design, 110
  - spherical charged particle emitter, 105–106
- Transverse velocity effects
- gas field ionization source, spherical charged particle emitter, 103–105
  - liquid metal ion sources, thermal velocity broadening, 62
- Transverse zero-point energy, gas field ionization emitters, 108
- Tunable Gunn diodes, focused ion beam implantation, 582, 590–591
- Tungsten cathodes, scanning electron microscope, 440–443
- Tunneling current, extended Schottky emission, 6–8
- TV rate scanning, scanning electron microscope, 448–449
- Two-dimensional Fourier transform
- resolution
  - contrast performance and, 400–402
  - image quality, 404–405
  - source intensity distribution, electron beam, 406–407
- Two-electrode immersion lenses, 177–178
- Two-particle approximation
- space charge effect, 343
  - extended approximation, 358–359
  - statistical coulomb effect
  - dynamics analysis, 353
  - extended two-particle approximation, 360–364
  - N*-particle problem, 352–353
  - plasma physics, 354–355

**U**

- Ultracorrector
  - basic properties, 298
  - multiple aberration correctors, 303–304
  - quadrupole–octupole arrangements, 620–625
- Unblurred distributions, gas field ionization source, spherical charged particle emitter, 103
- Uniform current distributions, space charge effect, 343
- Unipotential lenses
  - action model, 174–175
  - classification, 171
  - three-electrode unipotential lenses, 178–179
- Upper unsteady regime, liquid metal ion sources, droplet emission, 68–69

**V**

- Vacuum gap breakdown, electrostatic lens, 187–188
- Vacuum space charge, liquid metal ion sources, field-emitted, 42–44
- Variable vacuum scanning electron microscopy (VPSEM), 472–478
  - charge control, 478
  - column properties, 473–474
  - contamination reduction, 478
  - gas effects, primary electron beam, 474–475
  - gas pressure and path optimization, 477–478
  - secondary electron signal detection in gas, 475–477
- Vector potential, aberration calculation, 214–221
- Ventura contracta, liquid metal ion sources, emitter shape formation, 34
- Vertex, mirrors and cathode lenses, 262–265
- Vertical bipolar transistors, focused ion beam implantation, 586–591
- Vibration
  - resolution limits and, 433–434
  - scanning electron microscope resolution, 485
- Virtual source point cathodes
  - emitter brightness, 22–23
  - gas field ionization source, 107–108
  - historical background, 1–2
- Viscous-loss terms, liquid metal ion sources, 49
- Voltage contrast, scanning electron microscope, 468–470
- Volume flow rate/impedance, steady-state current-related liquid metal ion sources, 56

**W**

- Wafer metrology, scanning electron microscope, 446
- Wavelength dispersive x-ray spectrometer (WDS), scanning electron microscope, 461–463
- Wave optics, density-of-information passing capacity, 418–420
- Weak collisions
  - statistical coulomb effect
    - extended two-particle approximation, 359, 362–364
    - nonmicrobeam instruments, 383–385
    - two-particle dynamics, 353
    - trajectory displacement, 365–368
- Wien filter
  - chromatic and geometric aberrations, 271–286
  - focused ion beam implantation, 582–591
- Wire selection criteria, magnetic lens coil design, 139–140
- Work function, 8–10
  - current fluctuations and, 23–24
  - emitter life mechanisms, 26–27
  - Schottky reduction, total energy distribution, 19–22
- Working correctors, aberration correction, 604–605

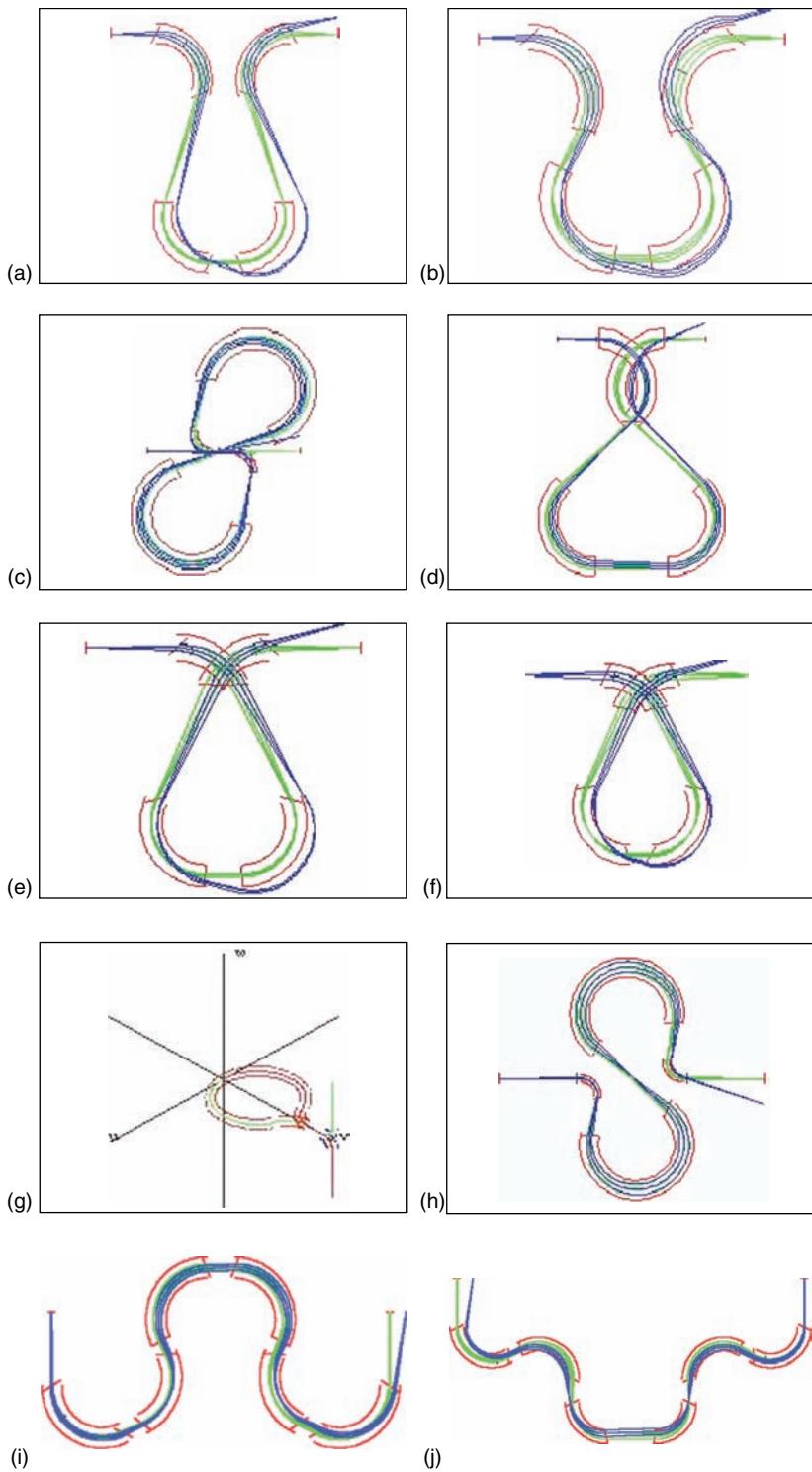
**X**

- X-ray, scanning electron microscope, signal properties, 461–463

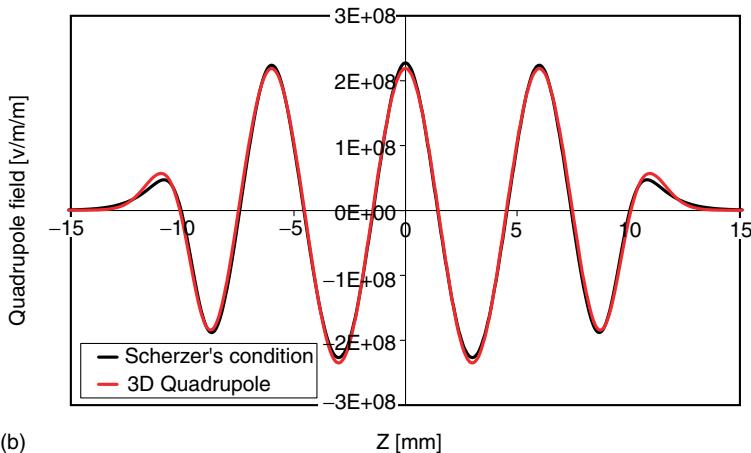
**Z**

- Zero-base-pressure approximation, liquid metal ion sources, 50
- Zero-Q evaporation field (ZQEF), liquid metal ion sources, 39–40
- Zoom lenses, 171
  - multielement movable lenses, 179
- ZrO/W cathodes
  - angular intensity/extraction voltage relationships, 10–12
  - crossover vs. noncrossover mode, 2
  - current fluctuations, 23–24
  - emitter environmental requirements, 24–25
  - emitter life mechanisms, 26–27
  - field factor  $\beta$ , emitter radius, and work function, 8–10
  - research background, 2–5
  - shape stability, 12–19
  - total energy distribution, 19–22

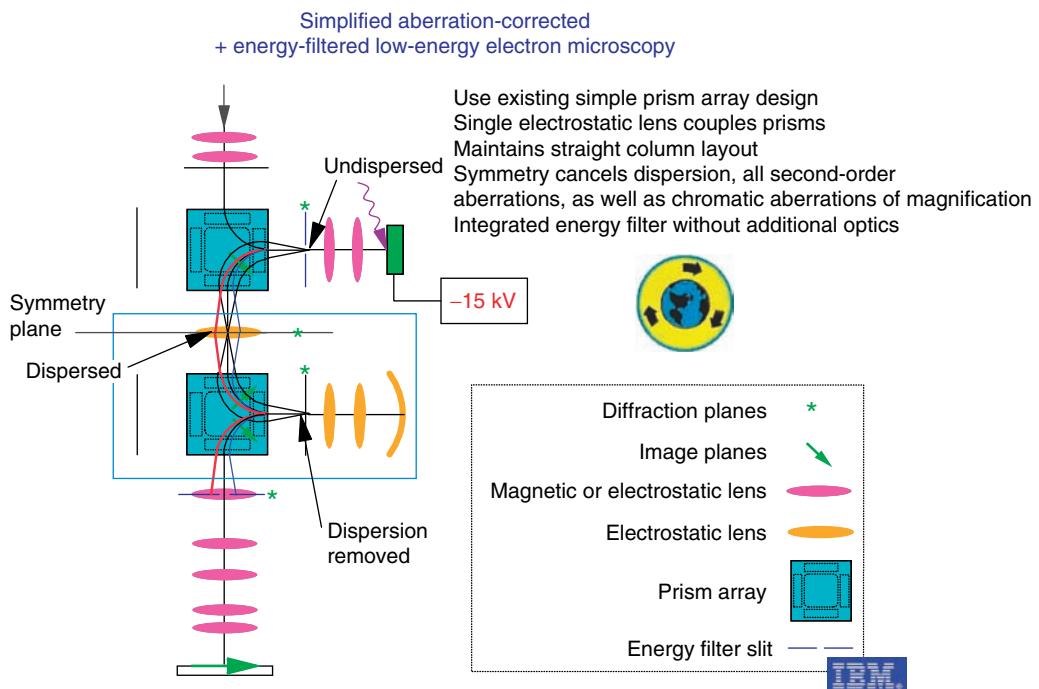




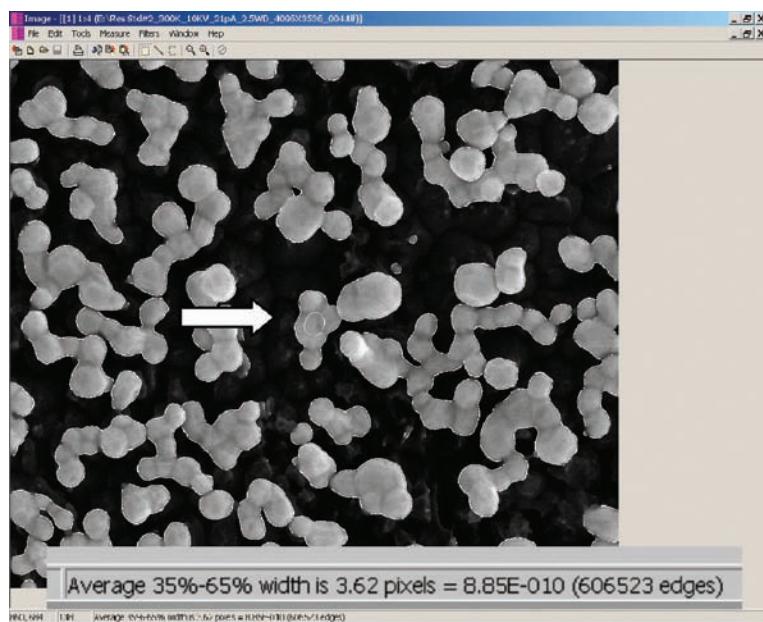
**COLOR FIGURE 6.4** Tableau showing the optic axes and typical rays in the various imaging filters: (a)  $\Omega$ -filter, A-type; (b)  $\Omega$ -filter, B-type; (c) infinity filter; (d) mandoline filter; (e)  $\alpha$ -filter, A-type; (f)  $\alpha$ -filter, B-type; (g)  $\phi$ -filter; (h) S-filter; (i) twin-column W-filter; and (j) variant twin-column geometry. (Courtesy of Dr. K. Tsuno.)



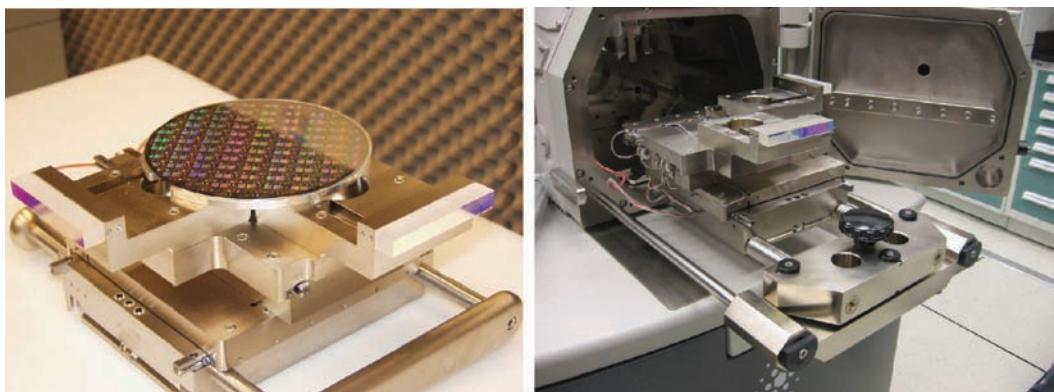
**COLOR FIGURE 6.12** Scherzer's proposal for electrostatic correction of chromatic aberration. (b) match between the potentials needed to satisfy condition 6.269 and those in the corrector. (After Maas D.J. et al., *Proc. SPIE*, 4510, 205–217, 2001; Mass D. et al., *Microsc. Microanal.*, 9(Suppl. 3), 24–25, 2003. Courtesy of the authors, SPIE, and the Microscopy Society of America.)



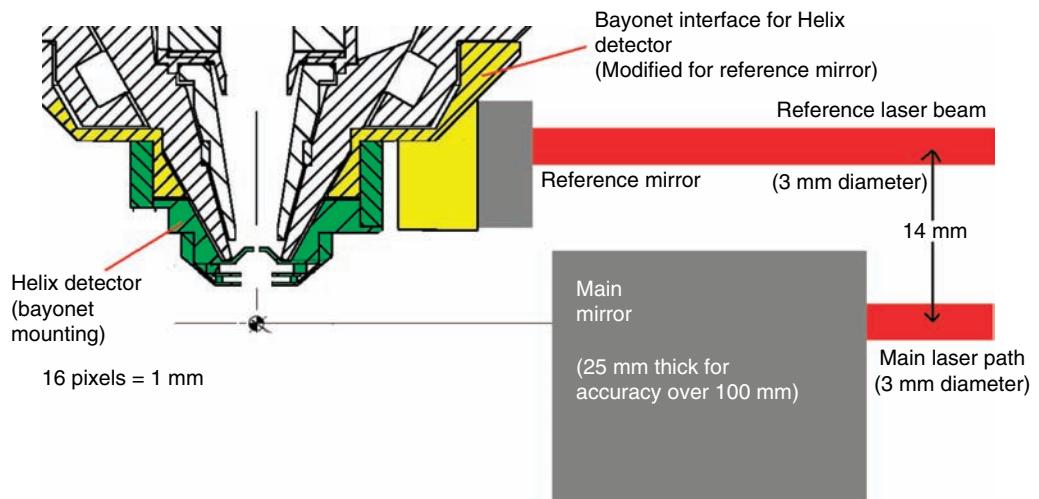
**COLOR FIGURE 6.16** The disposition of the components of the IBM photoemission electron microscope showing the two prism units, the electron mirror, and the connecting electrostatic lens. The two sets of conjugate planes—image planes and diffraction planes—are identified by arrows and stars, respectively. (Courtesy of R. Tromp.)



**COLOR FIGURE 9.47** The 0.9 nm spatial resolution performance of the environmental scanning electron microscope (ESEM). Note the good, circular beam shape in the center of the image. The field-of-view of the gold-on-carbon sample is 200 nm.



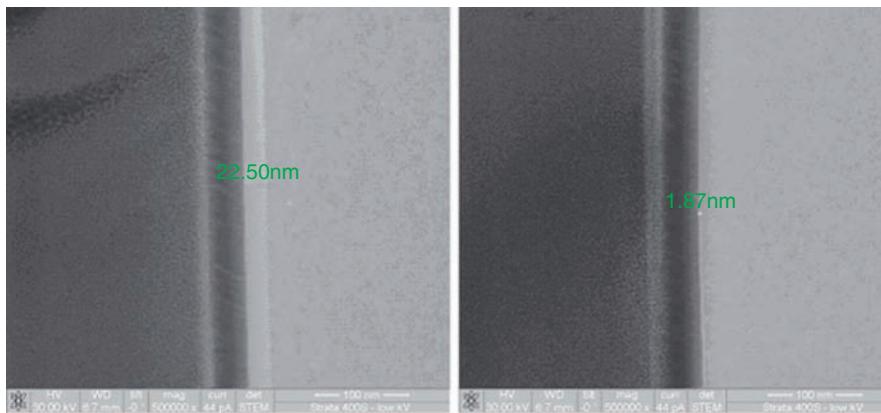
**COLOR FIGURE 9.48** Laser interferometer sample stage with a 200 mm wafer (left). Laser interferometer sample stage before insertion into the environmental scanning electron microscope (ESEM) (right).



**COLOR FIGURE 9.49** Schematic drawing of the beam path of the laser interferometer sample stage. The system is optimized at 4 mm working distance.



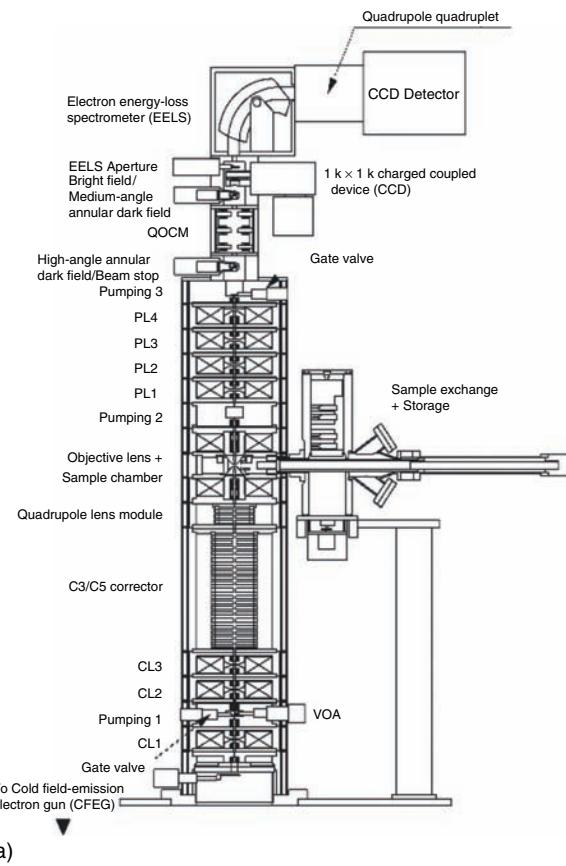
**COLOR FIGURE 11.1** A state-of-the-art FIB system. (Courtesy of FEI Company.)



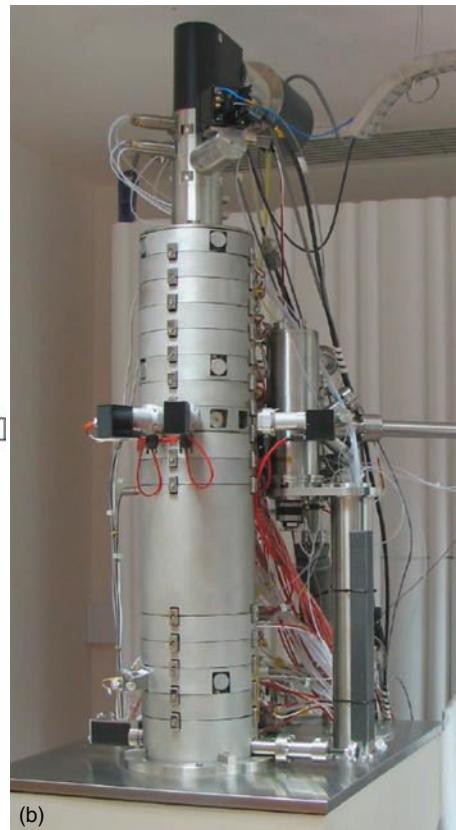
**COLOR FIGURE 11.11** STEM images of damage layers produced by 30 keV (left) and 2 keV (right).



**COLOR FIGURE 11.A.1** This site was pointed out to the author by the editor of this book, and is on the island of Crete at Gortys, dating back ~2600 years–600 B.C. This shows one section of a wall containing the Laws of Gortyna, some 600 lines of law code that are the earliest recorded in the Greek world that are boustrophedonically written.

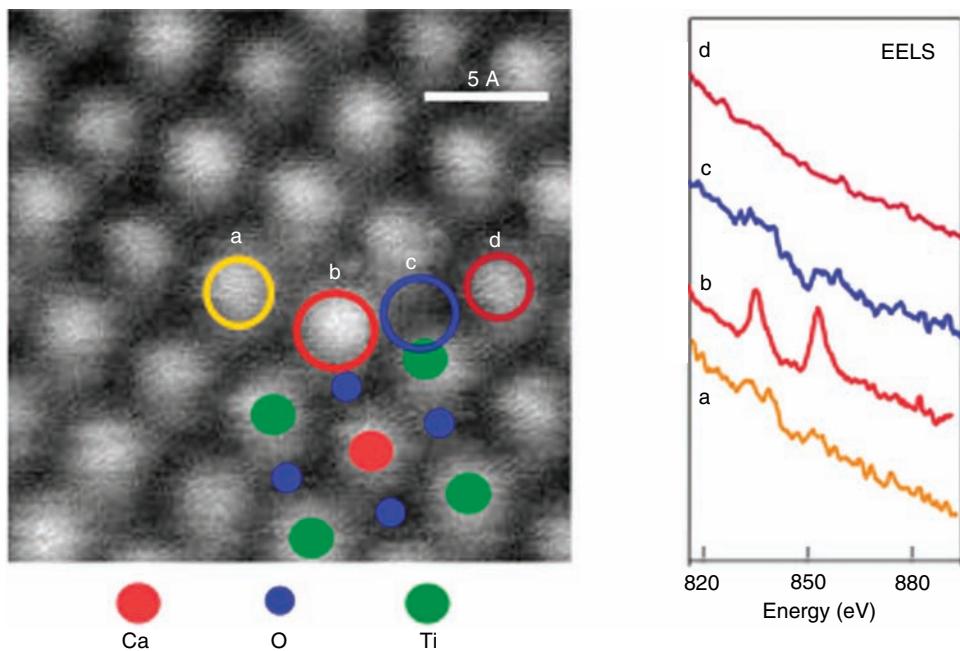


(a)

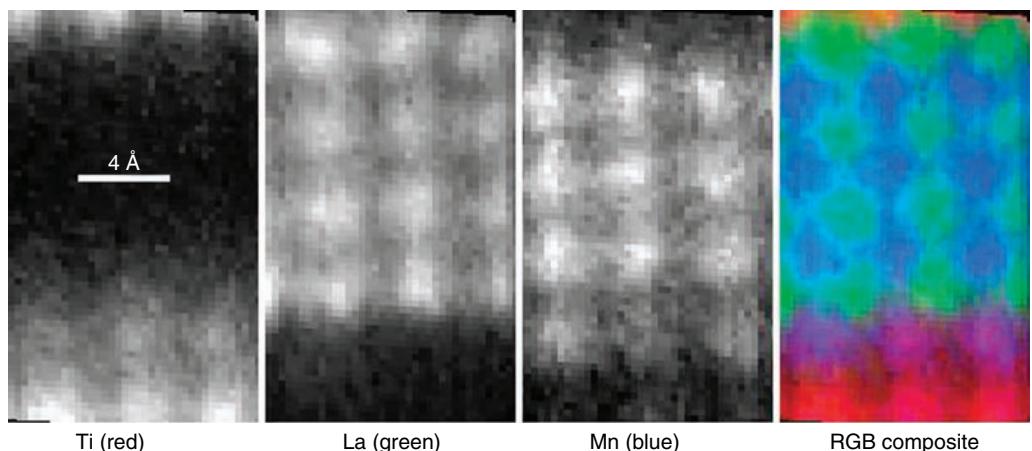


(b)

**COLOR FIGURE 12.12** Scanning transmission electron microscope column that includes the corrector of Figure 12.9. (a) Schematic cross section and (b) the actual column. Column diameter = 280 mm.



**COLOR FIGURE 12.15** (a) High-angle annular dark-field image and electron energy-loss spectra from  $\text{CaTiO}_3$  doped with La. The spectrum (b) originated from a single atom of La. VG HB501 STEM, Nion aberration corrector, 100 keV. (Courtesy Drs M. Varela and S.J. Pennycook, permission Physical Review Letters.)



**COLOR FIGURE 12.16** Electron energy-loss spectroscopy spectrum-images of Ti, La, and Mn in a  $\text{SrTiO}_3$ - $\text{La}_{0.7}\text{Sr}_{0.3}\text{MnO}_3$  multilayer, plus a combined false color image showing the locations of various atomic columns in the structure. Nion column with  $C_3/C_5$  aberration corrector, 100 keV. (Courtesy Prof D.A. Muller, L.F. Koukoutis, and M.F. Murfitt, multilayer structure courtesy Drs H.Y. Hwang and J.H. Song.)



# Handboook of Charged Particle Optics

## Second Edition

With the growing proliferation of nanotechnologies, powerful imaging technologies are being developed to operate at the sub-nanometer scale. A new edition of a bestseller, the **Handbook of Charged Particle Optics, Second Edition** provides essential background information for the design and operation of high resolution-focused probe instruments.

The book's unique approach covers both the theoretical and practical knowledge of high-resolution probe-forming instruments. With the inclusion of additional references to past and present work in the field, this second edition offers perfectly calibrated coverage of the field's cutting-edge technologies with added insight into how they work.

- Provides up-to-date information on high brightness ion and electron sources
- Covers SEM, STEM and FIB operation
- Offers a detailed analysis of aberration theory and aberration correction
- Includes information on space charge analysis of probe-forming instruments

Written by the leading research scientists, the second edition of the **Handbook of Charged Particle Optics** is a complete guide to understanding, designing, and using high-resolution probe instrumentation.

45547

ISBN : 978-1-4200-4554-3

90000



9 781420 045543



CRC Press  
Taylor & Francis Group  
an Informa business  
[www.crcpress.com](http://www.crcpress.com)

6000 Broken Sound Parkway, NW  
Suite 300, Boca Raton, FL 33487  
270 Madison Avenue  
New York, NY 10016  
2 Park Square, Milton Park  
Abingdon, Oxon OX14 4RN, UK