Transmission Electron Microscopy

A Textbook for Materials Science

Transmission Electron Microscopy

A Textbook for Materials Science

David B. Williams
C. Barry Carter

Springer

David B. Williams
The University of Alabama in Huntsville
Huntsville AL, USA
david.williams@uah.edu

C. Barry Carter
University of Connecticut
Storrs, CT, USA
cbcarter@engr.uconn.edu

ISBN 978-0-387-76500-6 hardcover
ISBN 978-0-387-76502-0 softcover (This is a four-volume set. The volumes are not sold individually.)
e-ISBN 978-0-387-76501-3

Library of Congress Control Number: 2008941103

Printed on acid-free paper

springer.com

To our parents

Walter Dennis and Mary Isabel Carter

and

Joseph Edward and Catherine Williams,
who made everything possible.

About the Authors

David B. Williams

David B. Williams became the fifth President of the University of Alabama in Huntsville in July 2007. Before that he spent more than 30 years at Lehigh University where he was the Harold Chambers Senior Professor Emeritus of Materials Science and Engineering (MS&E). He obtained his BA (1970), MA (1974), PhD (1974) and ScD (2001) from Cambridge University, where he also earned four Blues in rugby and athletics. In 1976 he moved to Lehigh as Assistant Professor, becoming Associate Professor (1979) and Professor (1983). He directed the Electron Optical Laboratory (1980–1998) and led Lehigh's Microscopy School for over 20 years. He was Chair of the MS&E Department from 1992 to 2000 and Vice Provost for Research from 2000 to 2006, and has held visiting-scientist positions at the University of New South Wales, the University of Sydney, Chalmers University (Gothenburg), Los Alamos National

Laboratory, the Max Planck Institut für Metallforschung (Stuttgart), the Office National d'Etudes et Recherches Aérospatiales (Paris) and Harbin Institute of Technology.

He has co-authored and edited 11 textbooks and conference proceedings, published more than 220 refereed journal papers and 200 abstracts/conference proceedings, and given 275 invited presentations at universities, conferences and research laboratories in 28 countries.

Among numerous awards, he has received the Burton Medal of the Electron Microscopy Society of America (1984), the Heinrich Medal of the US Microbeam Analysis Society (MAS) (1988), the MAS Presidential Science Award (1997) and was the first recipient of the Duncumb award for excellence in microanalysis (2007). From Lehigh, he received the Robinson Award (1979), the Libsch Award (1993) and was the Founders Day commencement speaker (1995). He has organized many national and international microscopy and analysis meetings including the 2nd International MAS conference (2000), and was co-chair of the scientific program for the 12th International Conference on Electron Microscopy (1990). He was an Editor of *Acta Materialia* (2001–2007) and the *Journal of Microscopy* (1989–1995) and was President of MAS (1991–1992) and the International Union of Microbeam Analysis Societies (1994–2000). He is a Fellow of The Minerals Metals and Materials Society (TMS), the American Society for Materials (ASM) International, The Institute of Materials (UK) (1985–1996) and the Royal Microscopical Society (UK).

C. Barry Carter

C. Barry Carter became the Head of the Department of Chemical, Materials & Biomolecular Engineering at the University of Connecticut in Storrs in July 2007. Before that he spent 12 years (1979–1991) on the Faculty at Cornell University in the Department of Materials Science and Engineering (MS&E) and 16 years as the 3 M

Heltzer Multidisciplinary Chair in the Department of Chemical Engineering and Materials Science (CEMS) at the University of Minnesota. He obtained his BA (1970), MA (1974) and ScD (2001) from Cambridge University, his MSc (1971) and DIC from Imperial College, London and his DPhil (1976) from Oxford University. After a postdoc in Oxford with his thesis advisor, Peter Hirsch, in 1977 he moved to Cornell initially as a postdoctoral fellow, becoming an Assistant Professor (1979), Associate Professor (1983) and Professor (1988) and directing the Electron Microscopy Facility (1987–1991). At Minnesota, he was the Founding Director of the High-Resolution Microscopy Center and then the Associate Director of the Center for Interfacial Engineering; he created the Characterization Facility as a unified facility including many forms of microscopy and diffraction in one physical location. He has held numerous visiting scientist positions: in the United States at the Sandia National Laboratories, Los Alamos National Laboratory and Xerox PARC; in Sweden at Chalmers University (Gothenburg); in Germany at the Max Planck Institut für Metallforschung (Stuttgart), the Forschungszentrum Jülich, Hannover University and IFW (Dresden); in France at ONERA (Chatillon); in the UK at Bristol University and at Cambridge University (Peterhouse); and in Japan at the ICYS at NIMS (Tsukuba).

He is the co-author of two textbooks (the other is *Ceramic Materials; Science & Engineering* with Grant Norton) and co-editor of six conference proceedings, and has published more than 275 refereed journal papers and more than 400 extended abstracts/conference proceedings. Since 1990 he has given more than 120 invited presentations at universities, conferences and research laboratories. Among numerous awards, he has received the Simon Guggenheim Award (1985–1986), the Berndt Matthias Scholar Award (1997/1998) and the Alexander von Humboldt Senior Award (1997). He organized the 16th International Symposium on the Reactivity of Solids (ISRS-16 in 2007). He was an Editor of the *Journal of Microscopy* (1995–1999) and of *Microscopy and Microanalysis* (2000–2004), and became (co-)Editor-in-Chief of the *Journal of Materials Science* in 2004. He was the 1997 President of MSA, and served on the Executive Board of the International Federation of Societies for Electron Microscopy (IFSEM; 1999–2002). He is now the General Secretary of the International Federation of Societies for Microscopy (IFSM; 2003–2010). He is a Fellow of the American Ceramics Society (1996) the Royal Microscopical Society (UK), the Materials Research Society (2009) and the Microscopy Society of America (2009).

Preface

How is this book different from the many other TEM books? It has several unique features but what we think distinguishes it from all other such books is that it is truly a *textbook*. We wrote it to be read by, and taught to, senior undergraduates and starting graduate students, rather than studied in a research laboratory. We wrote it using the same style and sentence construction that we have used in countless classroom lectures, rather than how we have written our countless (and much-less read) formal scientific papers. In this respect particularly, we have been deliberate in *not* referencing the sources of every experimental fact or theoretical concept (although we do include some hints and clues in the chapters). However, at the end of each chapter we have included groups of references that should lead you to the best sources in the literature and help you go into more depth as you become more confident about what you are looking for. We are great believers in the value of history as the basis for understanding the present and so the history of the techniques and key historical references are threaded throughout the book. Just because a reference is dated in the previous century (or even the antepenultimate century) doesn't mean it isn't useful! Likewise, with the numerous figures drawn from across the fields of materials science and engineering and nanotechnology, we do not reference the source in each caption. But at the very end of the book each of our many generous colleagues whose work we have used is clearly acknowledged.

The book consists of 40 relatively small chapters (with a few notable Carter exceptions!). The contents of most of the chapters can be covered in a typical lecture of 50-75 minutes (especially if you talk as fast as Williams). Furthermore, each of the four softbound volumes is flexible enough to be usable at the TEM console so you can check what you are seeing against what you should be seeing. Most importantly perhaps, the softbound version is cheap enough for all serious students to buy. So we hope you won't have to try and work out the meaning of the many complex color diagrams from secondhand B&W copies that you acquired from a former student. We have deliberately used color where it is useful rather than simply for its own sake (since all electron signals are colorless anyhow). There are numerous boxes throughout the text, drawing your attention to key information (green), warnings about mistakes you might easily make (amber), and dangerous practices or common errors (red).

Our approach throughout this text is to answer two fundamental questions:

Why should we use a particular TEM technique?
How do we put the technique into practice?

In answering the first question we attempt to establish a sound theoretical basis where necessary although not always giving all the details. We use this knowledge to answer the second question by explaining operational details in a generic sense and showing many illustrative figures. In contrast, other TEM books tend to be either strongly theoretical or predominantly descriptive (often covering more than just TEM). We view our approach as a compromise between the two extremes, covering enough theory to be reasonably rigorous without incurring the wrath of electron physicists yet containing sufficient hands-on instructions and practical examples to be useful to the materials engineer/nanotechnologist who wants an answer to a

materials problem rather than just a set of glorious images, spectra, and diffraction patterns. We acknowledge that, in attempting to seek this compromise, we often gloss over the details of much of the physics and math behind the many techniques but contend that the content is usually approximately right (even if on occasions, it might be precisely incorrect!).

Since this text covers the whole field of TEM we incorporate, to varying degrees, *all* the capabilities of the various kinds of current TEMs and we attempt to create a coherent view of the many aspects of these instruments. For instance, rather than separating out the broad-beam techniques of a traditional TEM from the focused-beam techniques of an analytical TEM, we treat these two approaches as different sides of the same coin. There is no reason to regard 'conventional' bright-field imaging in a parallel-beam TEM as being more fundamental (although it is certainly a more-established technique) than annular dark-field imaging in a focused-beam STEM. Convergent beam, scanning beam, and selected-area diffraction are likewise integral parts of the whole of TEM diffraction.

However, in the decade and more since the first edition was published, there has been a significant increase in the number of TEM and related techniques, greater sophistication in the microscope's experimental capabilities, astonishing improvements in computer control of the instrument, and new hardware designs and amazing developments in software to model the gigabytes of data generated by these almost-completely digital instruments. Much of this explosion of information has coincided with the worldwide drive to explore the nanoworld, and the still-ongoing effects of Moore's law. It is not possible to include all of this new knowledge in the second edition without transforming the already doorstop sized text into something capable of halting a large projectile in its tracks. It is still essential that this second edition teaches you to understand the essence of the TEM before you attempt to master the latest advances. But we personally cannot hope to understand fully all the new techniques, especially as we both descend into more administrative positions in our professional lives. Therefore, we have prevailed on almost 20 of our close friends and colleagues to put together with us a companion text (TEM; a companion text, Williams and Carter (Eds.) Springer 2010) to which we will refer throughout this second edition. The companion text is just as it says—it's a friend whose advice you should seek when the main text isn't enough. The companion is not necessarily more advanced but is certainly more detailed in dealing with key recent developments as well as some more traditional aspects of TEM that have seen a resurgence of interest. We have taken our colleagues' contributions and rewritten them in a similar conversational vein to this main text and we hope that this approach, combined with the in-depth cross-referencing between the two texts will guide you as you start down the rewarding path to becoming a transmission microscopist.

We each bring more than 35 years of teaching and research in all aspects of TEM. Our research into different materials includes metals, alloys, ceramics, semiconductors, glasses, composites, nano and other particles, atomic-level planar interfaces, and other crystal defects. (The lack of polymeric and biological materials in our own research is evident in their relative absence in this book.) We have contributed to the training of a generation of (we hope) skilled microscopists, several of whom have followed us as professors and researchers in the EM field. These students represent our legacy to our beloved research field and we are overtly proud of their accomplishments. But we also expect some combination of these (still relatively young) men and women to write the third edition. We know that they, like us, will find that writing such a text broadens their knowledge considerably and will also be the source of much joy, frustration, and enduring friendship. We hope you have as much fun reading this book as we had writing it, but we hope also that it takes you much less time. Lastly, we encourage you to send us any comments, both positive and negative. We can both be reached by e-mail: david.williams@uah.edu and cbcarter@engr.uconn.edu.

Foreword to First Edition

Electron microscopy has revolutionized our understanding of materials by completing the *processing-structure-properties* links down to atomistic levels. It is now even possible to tailor the microstructure (and mesostructure) of materials to achieve specific sets of properties; the extraordinary abilities of modern transmission electron microscopy—TEM—instruments to provide almost all the structural, phase, and crystallographic data allow us to accomplish this feat. Therefore, it is obvious that any curriculum in modern materials education must include suitable courses in electron microscopy. It is also essential that suitable texts be available for the preparation of the students and researchers who must carry out electron microscopy properly and quantitatively.

The 40 chapters of this new text by Barry Carter and David Williams (like many of us, well schooled in microscopy at Cambridge and Oxford) do just that. If you want to learn about electron microscopy from specimen preparation (the ultimate limitation); or via the instrument; or how to use the TEM correctly to perform imaging, diffraction, and spectroscopy—it's all there! This, to my knowledge, is the only complete text now available that includes all the remarkable advances made in the field of TEM in the past 30 to 40 years. The timing for this book is just right and, personally, it is exciting to have been part of the development it covers—developments that have impacted so heavily on materials science.

In case there are people out there who still think TEM is just taking pretty pictures to fill up one's bibliography, please stop, pause, take a look at this book, and digest the extraordinary intellectual demands required of the microscopist in order to do the job properly: crystallography, diffraction, image contrast, inelastic scattering events, and spectroscopy. Remember, these used to be fields in themselves. Today, one has to understand the fundamentals of *all* these areas before one can hope to tackle significant problems in materials science. TEM is a technique of characterizing materials down to the atomic limits. It must be used with care and attention, in many cases involving teams of experts from different venues. The fundamentals are, of course, based in physics, so aspiring materials scientists would be well advised to have prior exposure to, for example, solid-state physics, crystallography, and crystal defects, as well as a basic understanding of materials science, for without the latter, how can a person see where TEM can (or may) be put to best use?

So much for the philosophy. This fine new book definitely fills a gap. It provides a sound basis for research workers and graduate students interested in exploring those aspects of structure, especially defects, that control properties. Even undergraduates are now expected (and rightly) to know the basis for electron microscopy, and this book, or appropriate parts of it, can also be utilized for undergraduate curricula in science and engineering.

The authors can be proud of an enormous task, very well done.

G. Thomas
Berkeley, California

Foreword to Second Edition

This book is an exciting entry into the world of atomic structure and characterization in materials science, with very practical instruction on how you can see it and measure it, using an electron microscope. You will learn an immense amount from it, and probably want to keep it for the rest of your life (particularly if the problems cost you some effort!).

Is nanoscience "the next industrial revolution"? Perhaps that will be some combination of energy, environmental and nanoscience. Whatever it is, the new methods which now allow control of materials synthesis at the atomic level will be a large part of it, from the manufacture of jet engine turbine-blades to that of catalysts, polymers, ceramics and semiconductors. As an exercise, work out how much reduction would result in the transatlantic airfare if aircraft turbine blade temperatures could be increased by 200°C. Now calculate the reduction in CO_2 emission, and increased efficiency (reduced coal use for the same amount of electricity) resulting from this temperature increase for a coal-fired electrical generating turbine. Perhaps you will be the person to invent these urgently needed things! The US Department of Energy's Grand Challenge report on the web lists the remarkable advances in exotic nanomaterials useful for energy research, from separation media in fuel cells, to photovoltaics and nano-catalysts which might someday electrolyze water under sunlight alone. Beyond these functional and structural materials, we are now also starting to see for the first time the intentional fabrication of atomic structures in which atoms can be addressed individually, for example, as quantum computers based perhaps on quantum dots. 'Quantum control' has been demonstrated, and we have seen fluorescent nanodots which can be used to label proteins.

Increasingly, in order to find out exactly what new material we have made, and how perfect it is (and so to improve the synthesis), these new synthesis methods must be accompanied by atomic scale compositional and structural analysis. The transmission electron microscope (TEM) has emerged as the perfect tool for this purpose. It can now give us atomic-resolution images of materials and their defects, together with spectroscopic data and diffraction patterns from sub-nanometer regions. The field-emission electron gun it uses is still the brightest particle source in all of physics, so that electron microdiffraction produces the most intense signal from the smallest volume of matter in all of science. For the TEM electron beam probe, we have magnetic lenses (now aberration corrected) which are extremely difficult for our X-ray and neutron competitors to produce (even with much more limited performance) and, perhaps most important of all, our energy-loss spectroscopy provides unrivalled spatial resolution combined with parallel detection (not possible with X-ray absorption spectroscopy, where absorbed X-rays disappear, rather than losing some energy and continuing to the detector).

Much of the advance in synthesis is the legacy of half a century of research in the semiconductor industry, as we attempt to synthesize and fabricate with other materials what is now so easily done with silicon. Exotic oxides, for example, can now be laid down layer by layer to form artificial crystal structures with new, useful properties. But it is also a result of the spectacular advances in materials characterization, and our ability to see structures at the atomic level. Perhaps the best example of this is the discovery of the carbon nanotube, which was first identified by using an electron

microscope. Any curious and observant electron microscopist can now discover new nanostructures just because they look interesting at the atomic scale. The important point is that if this is done in an environmental microscope, he or she will know how to make them, since the thermodynamic conditions will be recorded when using such a 'lab in a microscope'. There are efforts at materials discovery by just such combinatorial trial-and-error methods, which could perhaps be incorporated into our electron microscopes. This is needed because there are often just 'too many possibilities' in nature to explore in the computer — the number of possible structures rises very rapidly with the number of distinct types of atoms.

It was Richard Feynman who said that, "if, in some catastrophe, all scientific knowledge was lost, and only one sentence could be preserved, then the statement to be passed on, which contained the most information in the fewest words, would be that matter consists of atoms." But confidence that matter consists of atoms developed surprisingly recently and as late as 1900 many (including Kelvin) were unconvinced, despite Avagadro's work and Faraday's on electrodeposition. Einstein's Brownian motion paper of 1905 finally persuaded most, as did Rutherford's experiments. Muller was first to see atoms (in his field-ion microscope in the early 1950s), and Albert Crewe two decades later in Chicago, with his invention of the field-emission gun for his scanning transmission electron microscope (STEM). The Greek Atomists first suggested that a stone, cut repeatedly, would eventually lead to an indivisible smallest fragment, and indeed Democritus believed that "nothing exists except vacuum and atoms. All else is opinion." Marco Polo remarks on the use of spectacles by the Chinese, but it was van Leeuwenhoek (1632-1723) whose series of papers in *Phil. Trans.* brought the microworld to the general scientific community for the first time using his much improved optical microscope. Robert Hooke's 1665 Micrographica sketches what he saw through his new compound microscope, including fascinating images of facetted crystallites, whose facet angles he explained with drawings of piles of cannon balls. Perhaps this was the first resurrection of the atomistic theory of matter since the Greeks. Zernike's phase-plate in the 1930s brought phase contrast to previously invisible ultra-thin biological 'phase objects', and so is the forerunner for the corresponding theory in high-resolution electron microscopy.

The past fifty years has been a wonderfully exciting time for electron microscopists in materials science, with continuous rapid advances in all of its many modes and detectors. From the development of the theory of Bragg diffraction contrast and the column approximation, which enables us to understand TEM images of crystals and their defects, to the theory of high-resolution microscopy useful for atomic-scale imaging, and on into the theory of all the powerful analytic modes and associated detectors, such as X-rays, cathodoluminescence and energy-loss spectroscopy, we have seen steady advances. And we have always known that defect structure in most cases controls properties — the most common (first-order) phase transitions are initiated at special sites, and in the electronic oxides a whole zoo of charge-density excitations and defects waits to be fully understood by electron microscopy. The theory of phase-transformation toughening of ceramics, for example, is a wonderful story which combines TEM observations with theory, as does that of precipitate hardening in alloys, or the early stages of semiconductor-crystal growth. The study of diffuse scattering from defects as a function of temperature at phase transitions is in its infancy, yet we have a far stronger signal there than in competing X-ray methods. The mapping of strain-fields at the nanoscale in devices, by quantitative convergent-beam electron diffraction, was developed just in time to solve a problem listed on the Semiconductor Roadmap (the speed of your laptop depends on strain-induced mobility enhancement). In biology, where the quantification of TEM data is taken more seriously, we have seen three-dimensional image reconstructions of many large proteins, including the ribosome (the factory which makes proteins according to DNA instructions). Their work should be a model to the materials science community in the constant effort toward better quantification of data.

Like all the best textbooks, this one was distilled from lecture notes, debugged over many years and generations of students. The authors have extracted the heart from

many difficult theory papers and a huge literature, to explain to you in the simplest, clearest manner (with many examples) the most important concepts and practices of modern transmission electron microscopy. This is a great service to the field and to its teaching worldwide. Your love affair with atoms begins!

J.C.H. Spence
Regent's Professor of Physics
Arizona State University and Lawrence
Berkeley National Laboratory

Acknowledgments

We have spent over 20 years conceiving and writing this text and the preceding first edition and such an endeavor can't be accomplished in isolation. Our first acknowledgment must be to our respective wives and children: Margie, Matthew, Bryn, and Stephen and Bryony, Ben, Adam, and Emily. Our families have borne the brunt of our absences from home (and occasionally the brunt of our presence). Neither edition would have been possible without the encouragement, advice, and persistence of (and the fine wines served by) Amelia McNamara, our first editor at Plenum Press, then Kluwer, and Springer.

We have both been fortunate to work in our respective universities with many more talented colleagues, post-doctoral associates, and graduate students, all of whom have taught us much and contributed significantly to the examples in both editions. We would like to thank a few of these colleagues directly: Dave Ackland, Faisal Alamgir, Arzu Altay, Ian Anderson, Ilke Arslan, Joysurya Basu, Steve Baumann, Charlie Betz, John Bruley, Derrick Carpenter, Helen Chan, Steve Claves, Dov Cohen, Ray Coles, Vinayak Dravid, Alwyn Eades, Shelley Gillis, Jeff Farrer, Joe Goldstein, Pradyumna Gupta, Brian Hebert, Jason Hefflefinger, John Hunt, Yasuo Ito, Matt Johnson, Vicki Keast, Chris Kiely, Paul Kotula, Chunfei Li, Ron Liu, Charlie Lyman, Mike Mallamaci, Stuart McKernan, Joe Michael, Julia Nowak, Grant Norton, Adam Papworth, Chris Perrey, Sundar Ramamurthy, René Rasmussen, Ravi Ravishankar, Kathy Repa, Kathy Reuter, Al Romig, Jag Sankar, David A. Smith, Kamal Soni, Changmo Sung, Caroline Swanson, Ken Vecchio, Masashi Watanabe, Jonathan Winterstein, Janet Wood, and Mike Zemyan.

In addition, many other colleagues and friends in the field of microscopy and analysis have helped with the book (even if they weren't aware of it). These include Ron Anderson, Raghavan Ayer, Jim Bentley, Gracie Burke, Jeff Campbell, Graham Cliff, David Cockayne, Peter Doig, the late Chuck Fiori, Peter Goodhew, Brendan Griffin, Ron Gronsky, Peter Hawkes, Tom Huber, Gilles Hug, David Joy, Mike Kersker, Roar Kilaas, Sasha Krajnikov, the late Riccardo Levi-Setti, Gordon Lorimer, Harald Müllejans, Dale Newbury, Mike O'Keefe, Peter Rez, Manfred Rühle, John-Henry Scott, John Steeds, Peter Swann, Gareth Thomas, Patrick Veyssière, Peter Williams, Nestor Zaluzec, and Elmar Zeitler. Many of these (and other) colleagues provided the figures that we acknowledge individually at the end of the book.

We have received financial support for our microscopy studies through several different federal agencies; without this support none of the research that underpins the contents of this book would have been accomplished. In particular, DBW wishes to thank the National Science Foundation, Division of Materials Research for over 30 years of continuous funding, NASA, Division of Planetary Science (with Joe Goldstein) and The Department of Energy, Basic Energy Sciences (with Mike Notis and Himanshu Jain), Bettis Laboratories, Pittsburgh, and Sandia National Laboratories, Albuquerque. While this edition was finalized at the University of Alabama in Huntsville, both editions were written while DBW was in the Center for Advanced Materials and Nanotechnology at Lehigh University, which supports that outstanding electron microscopy laboratory. Portions of both editions were written while DBW was on sabbatical or during extended visits to various microscopy labs: Chalmers University, Göteborg, with Gordon Dunlop and Hans Nordén; The Max Planck

Institut für Metallforschung, Stuttgart, with Manfred Rühle; Los Alamos National Laboratory with Terry Mitchell; Dartmouth College, Thayer School of Engineering, with Erland Schulson; and the Electron Microscope Unit at Sydney University with Simon Ringer. CBC wishes to acknowledge the Department of Energy, Basic Energy Sciences, the National Science Foundation, Division of Materials Research, the Center for Interfacial Engineering at the University of Minnesota, The Materials Science Center at Cornell University, and the SHaRE program at Oak Ridge National Laboratories. The first edition was started while CBC was with the Department of Materials Science and Engineering at Cornell University. This edition was started at the Department of Chemical Engineering and Materials Science at the University of Minnesota where the first edition was finished and was finalized while CBC was at the University of Connecticut. The second edition was partly written while CBC was on Sabbatical Leave at Chalmers University with Eva Olssen (thanks also to Anders Tholen at Chalmers), at NIMS in Tsukuba with Yoshio Bando (thanks also to Dmitri Golberg and Kazuo Furuya at NIMS at Yuichi Ikuhara at the University of Tokyo) and at Cambridge University with Paul Midgley. CBC also thanks the Master and Fellows of Peterhouse for their hospitality during the latter period.

CBC would also like to thank the team at the Ernst Ruska Center for their repeated generous hospitality (special thanks to Knut Urban, Markus Lenzen, Andreas Thust, Martina Luysberg, Karsten Tillmann, Chunlin Jia and Lothar Houben)

Despite our common scientific beginnings as undergraduates in Christ's College Cambridge, we learned our trade under different microscopists: DBW with Jeff Edington in Cambridge and CBC with Sir Peter Hirsch and Mike Whelan in Oxford. Not surprisingly, the classic texts by these renowned microscopists are referred to throughout this book. They influenced our own views of TEM tremendously, contributing to the undoubted bias in our opinions, notation, and approach to the whole subject.

Contents

Part 1

Basics

1

The Transmission Electron Microscope

CHAPTER PREVIEW

A typical commercial transmission electron microscope (TEM) costs about $5 for each electron volt (eV) of energy in the beam and, if you add on all available options, it can easily cost up to $10 per eV. As you'll see, we use beam energies in the range from 100,000 to 400,000 eV, so a TEM is an extremely expensive piece of equipment. Consequently, there have to be very sound scientific reasons for investing such a large amount of money in one microscope. In this chapter (which is just a brief overview of many of the concepts that we'll talk about in detail throughout the book) we start by introducing you to some of the historical development of the TEM because the history is intertwined with some of the reasons why you need to use a TEM to characterize materials. Other reasons for using a TEM have appeared as the instrument continues to develop, to the point where it can seriously be claimed that no other scientific instrument exists which can offer such a broad range of characterization techniques with such high spatial and analytical resolution, coupled with a *completely* quantitative understanding of the various techniques. Indeed as nanotechnology and related areas seize both the public and the technological community's imaginations, it is increasingly obvious that the TEM is the central tool for complete characterization of nanoscale materials and devices. Unfortunately, coupled with the TEM's advantages are some serious drawbacks and you must be just as aware of the instrument's limitations as you are of its advantages, so we summarize these also.

A TEM can appear in several different forms, all of which are described by different acronyms such as HRTEM, STEM, and AEM, and we'll introduce you to these different instruments. We'll also use the same acronyms or initials (go back and read p. xxi) to denote both the technique (microscopy) and the instrument (microscope). We regard all of the different types of TEM as simply variations on a basic theme and that is why only 'TEM' is in the book title. We will also describe some of the basic physical characteristics of the electron. Throughout the book you'll have to confront some physics and mathematics every now and again because understanding what we can do with a TEM and why we operate it in certain ways is governed by the fundamental physics of electrons, how electrons are controlled by magnetic fields in the microscope, how electrons interact with materials, and how we detect the many signals emitted from a specimen in the TEM.

Finally we will summarize some of the most popular computer software packages for TEM. We will refer to many of these throughout the text. We are including them in the first chapter to emphasize the central role of the computer in today's TEM operation and analysis. A basic lesson to take away from this chapter is not just the versatility of the TEM but the fact that it is fundamentally a signal-generating and detecting instrument rather than simply a microscope for high-resolution images and diffraction patterns (we'll call them DPs), which is how it operated for many decades.

1.1 WHAT MATERIALS SHOULD WE STUDY IN THE TEM?

The materials scientist has traditionally examined metals, alloys, ceramics, glasses, polymers, semiconductors, and composite mixtures of these materials, with sporadic adventures into wood, textiles, and concrete. In addition to thinning them from the bulk state, particles and fibers of some of these materials are also commonly studied and, in such shapes, they are sometimes thin enough for direct TEM examination. Nanotechnology, which will feature as a common theme throughout

this book, is defined as "the ability to understand and control matter at dimensions of roughly 1 to 100 nanometers, where unique phenomena enable novel applications. Encompassing nanoscale science, engineering and technology, nanotechnology involves imaging, measuring, modeling, and manipulating matter at this length scale" (URL #1).

THE CRUCIAL WORDS

"Imaging, measuring, modeling, and manipulating matter" can be accomplished with the help of the TEM and are often thrown together as part of the emerging field of 'nanocharacterization,' a term which we will try not to use too often.

When we create nanoscale materials, they come with specific dimensional limits in 1D, 2D, or 3D and the TEM is well suited to observing them, precisely because of these limits. We will include examples of archetypal dimensionally limited structures throughout the book. For example, single layers (such as graphene sheets or quantum wells), nanotubes and nanowires, quantum dots, nanoparticles, and most catalyst particles can be viewed as 1D structures. We can put all of these types of specimen into the TEM without modification, since 1D is always thin enough for direct observation; 2D nanomaterials include interfaces, and complex 3D nanomaterials are typified by multilayer, semiconductor devices, functional materials, or nanoporous structures such as substrates for catalyst-particle dispersions. Lastly, we should note the rapidly growing interface between the nano- and the bio-worlds. While much of biological electron microscopy has been superceded in the last decade or more by less-damaging techniques such as confocal, two-photon, multi-photon, and near-field light microscopies, there is still a major role for TEM in biomaterials, bio/inorganic interfaces, and nano-bio/biomaterials.

1.2 WHY USE ELECTRONS?

Why should we use an electron microscope? Historically TEMs were developed because of the limited image resolution in light microscopes, which is imposed by the wavelength of visible light. Only after electron microscopes were developed was it realized that there are many other equally sound reasons for using electrons, most of which are utilized to some extent in a modern TEM. By way of introduction to the topic, let's look at how the TEM developed and the pros and cons of using such an instrument.

1.2.A An Extremely Brief History

Louis de Broglie (1925) first theorized that the electron had wave-like characteristics, with a wavelength substantially less than visible light. Then in 1927 two

FIGURE 1.1. The electron microscope built by Ruska (in the lab coat) and Knoll, in Berlin in the early 1930s.

research groups, Davisson and Germer and Thomson and Reid, independently carried out their classic electron-diffraction experiments, which demonstrated the wave nature of electrons. It didn't take long for the idea of an electron microscope to be proposed, and the term was first used in the paper of Knoll and Ruska (1932). In this paper they developed the idea of electron lenses into a practical reality and demonstrated electron images taken on the instrument shown in Figure 1.1. This was a most crucial step, for which Ruska received the Nobel Prize ("somewhat late" he was quoted as saying), in 1986, shortly before his death in 1988. Within a year of Knoll and Ruska's publication, the resolution limit of the light microscope was surpassed. Ruska, surprisingly, revealed that he hadn't heard of de Broglie's ideas about electron waves and thought that the wavelength limit didn't apply to electrons. Some idea of the power of Ruska's breakthrough is the fact that commercial TEMs were first developed only 4 years later. The Metropolitan-Vickers EM1 was the first such instrument and was built in the UK in 1936. Apparently it didn't work very well and regular production of commercial TEMs was really started by Siemens and Halske in Germany in 1939. TEMs became widely available from several other sources (Hitachi, JEOL, Philips, and RCA, inter alia) after the conclusion of World War II.

For materials scientists a most important development took place in the mid-1950s when Bollman in Switzerland and Hirsch and co-workers in Cambridge, in the UK, perfected techniques to thin metal foils to electron transparency. (In fact, because so much of the early TEM work examined metal specimens, the word 'foil' came to be synonymous with 'specimen' and we'll often use it this way.) In addition, the Cambridge group also developed the theory of electron-diffraction contrast with which we can now identify, often in a quantitative manner, *all* known line and planar crystal defects

in TEM images. This work is summarized in a formidable but essential text often referred to as the 'Bible' of TEM (Hirsch et al. 1977). For the materials scientist, practical applications of the TEM for the solution of materials problems were pioneered in the United States by Thomas and first clearly expounded in his text. Other materials-oriented texts followed, notably the first student-friendly 'hands-on' text by Edington.

Today TEMs constitute arguably the most efficient and versatile tools for the characterization of materials over spatial ranges from the atomic scale, through the ever-growing 'nano' regime (from < 1 nm to ~ 100 nm) up to the micrometer level and beyond. If you want to read a history of the TEM, the book by Marton (1968) is a compact, personal monograph and the text edited by Hawkes in 1985 contains a series of individual reminiscences. Fujita's (1986) paper emphasizes the substantial contribution of Japanese scientists to the development of the instrument. The field is now at the point where many of the pioneers have put their memoirs down on paper, or Festschrifts have been organized in their honor (e.g., Cosslett 1979, Ruska 1980, Hashimoto 1986, Howie 2000, Thomas 2002, Zeitler 2003) which detail their contributions over the decades, and compile some useful overview papers of the field. If you enjoy reading about the history of science, we strongly recommend the review of *Fifty Years of Electron Diffraction* edited by Goodman (1981) and *Fifty Years of X-ray Diffraction* edited by Ewald (1962) (the spelling of X-ray is discussed in the *CBE Manual*, 1994). More recently, Haguenau et al. (2003) compiled an extensive list of references describing key events in the history of electron microscopy. As always, there is a wealth of information, some of it accurate, available on the Web.

1.2.B Microscopy and the Concept of Resolution

When asked "what is a microscope?," most people would answer that it is an instrument for magnifying things too small to see with the naked eye, and most likely they would be referring to the visible-light microscope (VLM). Because of the general familiarity with the concept of the VLM, we will draw analogies between electron and light microscopes wherever it's instructive.

The smallest distance between two points that we can resolve with our eyes is about 0.1–0.2 mm, depending on how good our eyes are, and assuming that there's sufficient illumination by which to see. This distance is the *resolution* or (more accurately) the *resolving power* of our eyes. So any instrument that can show us pictures (or *images* as we'll often refer to them) revealing detail finer than 0.1 mm could be described as a microscope, and its highest useful magnification is governed by its resolution. A major attraction to the early developers of the TEM was that, since electrons are smaller than atoms, it should be possible, at least theoretically, to build a microscope that could 'see' detail well below the atomic level. The idea of being able to 'see' with electrons may be confusing to you. Our eyes are not sensitive to electrons. If a beam of high-energy electrons was aimed into your eye, you would most likely be blinded as the electrons killed your retinal cells, but you wouldn't see anything (ever again!). So an integral part of any electron microscope is a viewing screen of some form (now usually a flat-panel computer display), which displays electron intensity as light intensity, which we first observe and then record photographically or store digitally. (We'll discuss these screens and other ways of recording electron images in Chapter 7.)

VLM

We'll try to avoid the phrases 'optical microscope' (they all are) and 'light microscope' (some are very heavy).
'Visible-light microscope/y' is simple and appropriate use of the hyphen.

The resolution of a TEM means different things for different functions of the instrument, and we'll discuss them in the appropriate chapters. It's easiest to think of the image resolution in TEM in terms of the classic Rayleigh criterion for VLM, which states that the smallest distance that can be resolved, δ, is given approximately by

$$\delta = \frac{0.61\lambda}{\mu \sin \beta} \tag{1.1}$$

In equation 1.1, λ is the wavelength of the radiation, μ the refractive index of the viewing medium, and β the semi-angle of collection of the magnifying lens. For the sake of simplicity we can approximate $\mu \sin \beta$ (which is sometimes called the numerical aperture) to unity and so the resolution is equal to about half the wavelength of light. For green light in the middle of the visible spectrum, λ is about 550 nm, and so the resolution of a good VLM is about 300 nm. In TEMs we can approximate the best resolution using an expression similar to equation 1.1 (actually $\sim 1.22\lambda/\beta$) which, as we'll see later, is very small.

Now although 300 nm is a small dimension to us, it corresponds to about 1000 atom diameters, and, therefore, many of the features that control the properties of materials are on a scale well below the resolution of the VLM. Also, 300 nm is well above the upper limit of the nano regime which we defined earlier. So there's a real need in nano/materials science and engineering to image details, all the way down to the atomic level, if we want to understand and ultimately control the properties of materials, and that's a major reason why TEMs are so useful.

This limit of light microscopy was well understood at the turn of the last century and prompted Ernst Abbe, one of the giants in the field, to complain that "it is poor comfort to hope that human ingenuity will find ways

and means of overcoming this limit." (He was right to be so depressed because he died in 1905, some 20 years before de Broglie's ingenuity solved the problem.) Louis de Broglie's famous equation shows that the wavelength of electrons is related to their energy, E, and, if we ignore relativistic effects, we can show approximately (and exactly in Section 1.4 below) that (ignoring the inconsistency in units)

$$\lambda = \frac{1.22}{E^{1/2}} \qquad (1.2)$$

In this equation E is in electron volts (eV) and λ in nm.

So from equation 1.2 you can work out that for a 100 keV electron, $\lambda \sim 4$ pm (0.004 nm), which is much smaller than the diameter of an atom.

V AND eV

Remember that we should be precise in our use of these units: V represents the accelerating voltage of the microscope while eV refers to the energy of the electrons in the microscope (look ahead to equation 1.4 to see the relation between the two).

We'll see later that we cannot yet build a 'perfect' TEM that approaches this wavelength-limited limit of resolution, because we can't make perfect electron lenses (see Chapter 6). Until recently, a top of the line lens could rightly be compared to using the bottom of a Coca-Cola™ bottle as a lens for light microscopy. Progress was rapid after Ruska's early work on lenses and since the mid-1970s many commercial TEMs have been capable of resolving individual columns of atoms in crystals, creating the field of high-resolution transmission electron microscopy or HRTEM, which we'll discuss in Chapter 28. A typical HRTEM image is shown in Figure 1.2A.

The advantages of shorter wavelengths led in the 1960s to the development of high-voltage electron microscopes (HVEMs), with accelerating potentials between 1 and 3 MV. In fact, rather than push the resolution limits, most of these instruments were used to introduce controlled amounts of radiation damage into specimens, in an attempt to simulate nuclear-reactor environments. Three-Mile Island and Chernobyl contributed to changes in the emphasis of energy research; recently there has not been much call for HVEMs. Today, climate change is forcing a reconsideration of nuclear power. Only one HVEM (1 MV) for HRTEM imaging was constructed in the 1980s and three 1.25 MV machines in the 1990s. Intermediate voltage electron microscopes (IVEMs) were introduced in the 1980s. These TEMs operate at 200–400 kV, but still offer very high resolution, close to that achieved at 1 MV. In fact, progress is such that most IVEMs purchased these days are, effectively, HRTEMs with atomic resolution.

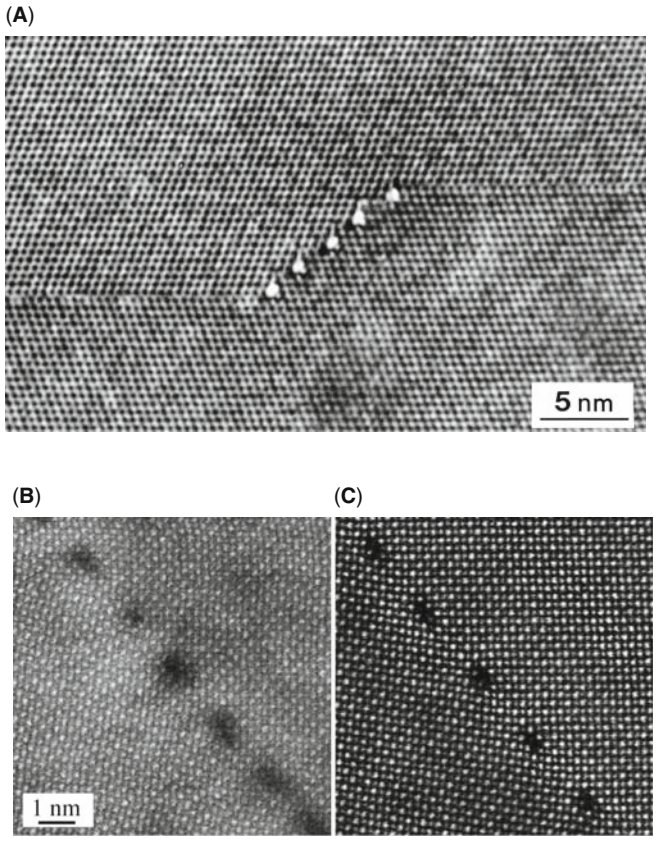


FIGURE 1.2. (A) A twin boundary in spinel stepping from one {111} plane to another parallel plane. The white dots are columns of atoms. The change in atomic orientation across the twin boundary can be readily seen even if we do not know what causes the white dots or why indeed they are white. (B) A grain boundary in $SrTiO_3$ imaged without C_s correction and (C) with C_s correction. As you can see, the effect is just as dramatic as putting on your reading glasses (if you need them).

We are still improving the resolution, and recent breakthroughs in spherical- and chromatic-aberration corrections (see Chapters 6 and 37, respectively) are revolutionizing the TEM field. Among many advantages, corrections of spherical aberration (which, for reasons we'll explain in Chapter 6, we abbreviate to C_s) and chromatic aberration (C_c) allow us to produce sharper atomic-resolution images. By filtering out electrons of different wavelengths we can also better image thicker specimens.

The combination of IVEMs and C_s correction has pushed TEM image resolution to well below the 0.1 nm (1 Å) barrier. Today the point has perhaps been reached where the drive for much better resolution is now no longer paramount and the TEM will develop more constructively in other ways. As we'll illustrate many times throughout the book and elaborate in the companion text, C_s correction is perhaps the most exciting advance in TEM technology in several decades and Figure 1.2B and C shows beautifully the difference in a typical atomic-resolution image with and without C_s correction. The advantages of C_s and C_c aberration correction in

TEM are explored in depth in chapters on C_s correction and energy-filtered TEM (EFTEM) in the companion text.

C_S,C_C AND MAGNIFICATION

Having extolled the virtues of C_s correction it is worth pointing out that most TEM images are recorded at magnifications where such correction makes *no* discernible difference. Most TEM specimens are not thin enough to produce images with resolution that benefits from C_s correction. For thicker specimens C_c correction via energy filtering is much more useful.

1.2.C Interaction of Electrons with Matter

Electrons are one type of ionizing radiation, which is the general term given to radiation that is capable of removing the tightly bound, inner-shell electrons from the attractive field of the nucleus by transferring some of its energy to individual atoms in the specimen.

One of the advantages of using ionizing radiation is that it produces a wide range of secondary signals from the specimen and some of these are summarized in Figure 1.3. Many of these signals are used in analytical electron microscopy (AEM), giving us chemical information and a lot of other details about our specimens. AEM uses X-ray energy-dispersive spectrometry (XEDS) and electron energy-loss spectrometry (EELS). For example, Figure 1.4A shows X-ray spectra from very small regions of the TEM specimen imaged in Figure 1.4B. The spectra exhibit characteristic peaks, which identify the different elements present in different regions. We can transform such spectra into quantitative

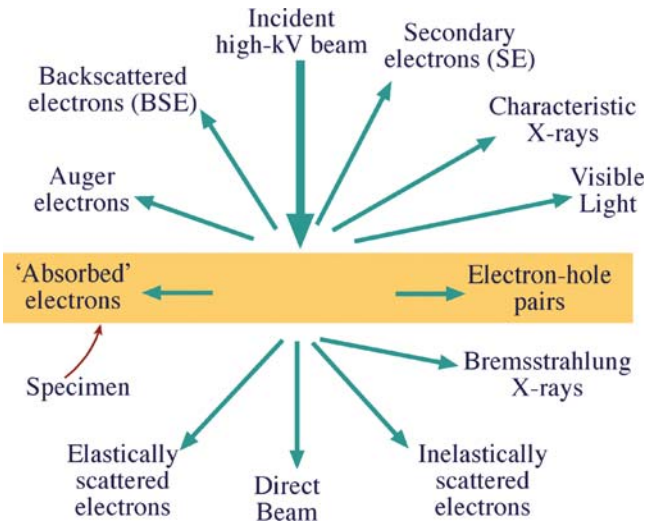


FIGURE 1.3. Signals generated when a high-energy beam of electrons interacts with a thin specimen. Most of these signals can be detected in different types of TEM. The directions shown for each signal do not always represent the physical direction of the signal, but indicate, in a relative manner, where the signal is strongest or where it is detected.

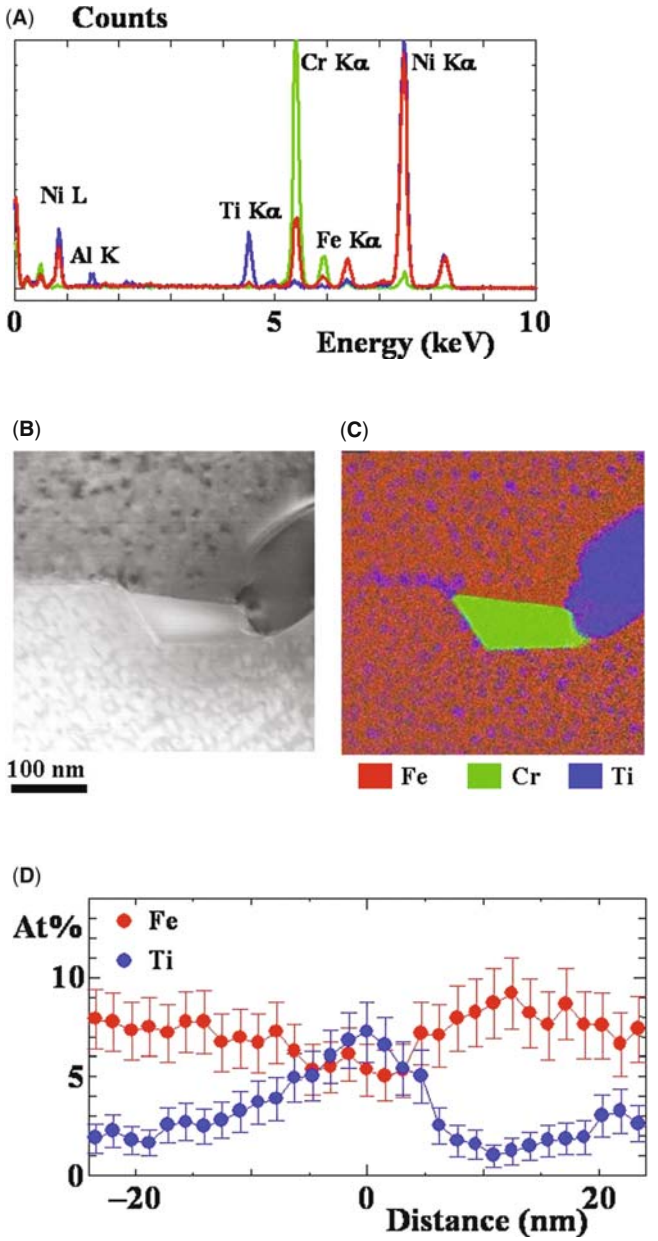


FIGURE 1.4. (A) X-ray spectra from three different regions of a specimen of Ni-base superalloy imaged in (B). The spectra are color-coded to match the different regions of the specimen highlighted in (C) which is a quantitative map showing the distribution of the elements detected in the spectra in (A) (e.g., green areas are rich in Cr, blue areas contain predominantly Ti, etc.). Quantitative composition profiles showing the localized changes in composition across one of the small matrix precipitates in (C) are shown in (D).

images of the distributions of all the elements present in the specimen (Figure 1.4C) and from such images extract quantitative data describing elemental changes associated with inhomogeneous microstructures as shown in Figure 1.4D. This and similar analyses with EELS comprise Part 4 of the book. In contrast, microscopes using non-ionizing radiation, such as visible light, usually only generate light (but not much heat, which is good).

In order to get the best signal out of our specimens we have to put the best signal in, and so the electron source is critical. We are now very accomplished in this respect, as you'll see in Chapter 5; modern TEMs are very good signal-generating instruments. To localize these signals we need our TEM to produce a very small electron beam (or *probe* as it is often called), typically <5 nm and at best < 0.1 nm in diameter. We combine TEM and scanning electron microscope (SEM) technology to create a scanning transmission electron microscope (STEM). The STEM is both the basis for AEMs and a unique scanning-imaging (or scanned-probe) microscope in its own right. In fact there are instruments that are only capable of operating in scanning mode and these are sometimes referred to as dedicated STEMs or DSTEMs. AEMs offer improved analytical performance at intermediate voltages, similar to the improvement in image resolution gained in standard TEMs.

Most importantly, C_s correction permits the generation of smaller electron probes with higher currents, thus significantly improving both analytical spatial resolution and sensitivity. Chromatic-aberration correction (i.e., energy filtering) also offers the opportunity to form images of electrons with a whole range of specific energies, permitting such breakthroughs as band-gap imaging and chemical-bond imaging.

1.2.D Depth of Field and Depth of Focus

The depth of field of a microscope is a measure of how much of the *object* that we are looking at remains in focus at the same time; the term depth of focus refers to the distance over which the *image* can move relative to the object and still remain in focus. If you are confused, it may help to recall that *depth of field* and *field of view* both refer to the *object* in everyday photography. The lenses in the TEM govern these properties just as they determine the resolution. Electron lenses are not very good, as we've already mentioned, and one way to improve their performance is to insert very small limiting apertures, narrowing the beam down to a thin 'pencil' of electrons which at most is a few micrometers across. These apertures obviously cut down the intensity of the electron beam, but they also act to increase the depth of field of the specimen and depth of focus of the images that we produce, as we explain in detail in Chapter 6.

While this large depth of field is chiefly used in the SEM to produce 3D-like images of the surfaces of specimens with large changes in topography, it is also critical in the TEM. It turns out that in the TEM, your specimen is usually in focus from the top to the bottom surfaces at the same time, independent of its topography, so long as it's electron transparent! Figure 1.5 shows a TEM image of some dislocations in a crystal.

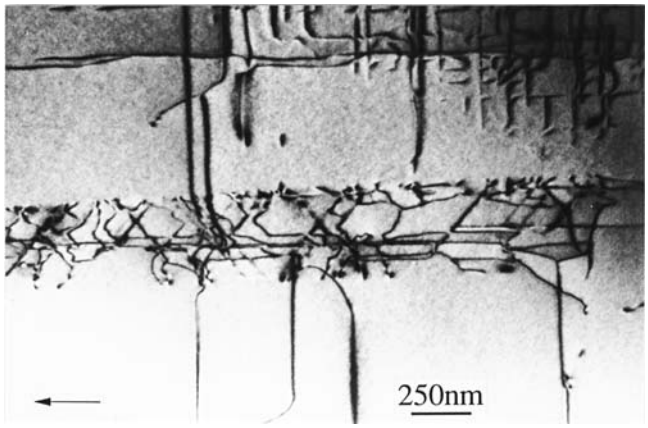


FIGURE 1.5. TEM image of dislocations (dark lines) in GaAs. The dislocations in the band across the middle of the image are on slip planes close to 90° to one another and thread through the thin specimen from the top to the bottom but remain in focus through the foil thickness.

The dislocations appear to start and finish in the specimen, but in fact they are threading their way through the specimen from the top to the bottom surfaces, and they remain in sharp focus at all times. (By the time you finish reading this book, you should be able to work out which is the top and which is the bottom surface of the specimen.) Furthermore, you can record the final image at different positions below the final lens of the instrument and it will still be in focus (although the magnification will change). Compare these properties with the VLM where, as you probably know, unless the surface of your specimen is flat within the wavelength of light, it is not all in focus at the same time. This aspect of TEM gives us both advantages and disadvantages in comparison to the VLM. You should note that, in this rare situation, C_s correction is not an advantage since it permits the use of larger apertures without degrading the resolution of the lens. But *smaller* apertures are the ones that give better depth of focus and depth of field (see Section 6.7). However, if you are using a C_s corrector, your specimen has to be so thin that it will still remain in focus except under extreme conditions. We'll see more on this topic in the companion text and also mention using TEM in a 'confocal' mode.

1.2.E Diffraction

As we've noted, Thompson, Reid, Davisson, and Germer independently showed that electrons could be diffracted when passing through thin crystals of nickel. Performing electron diffraction in TEMs was first realized by Kossel and Möllenstedt (1939). Today, electron diffraction is an indispensable part of TEM and is arguably the most useful aspect for materials scientists and nanotechnologists for whom crystal structure (and particularly crystal defects) is an essential characteristic when it comes to controlling properties. Figure 1.6

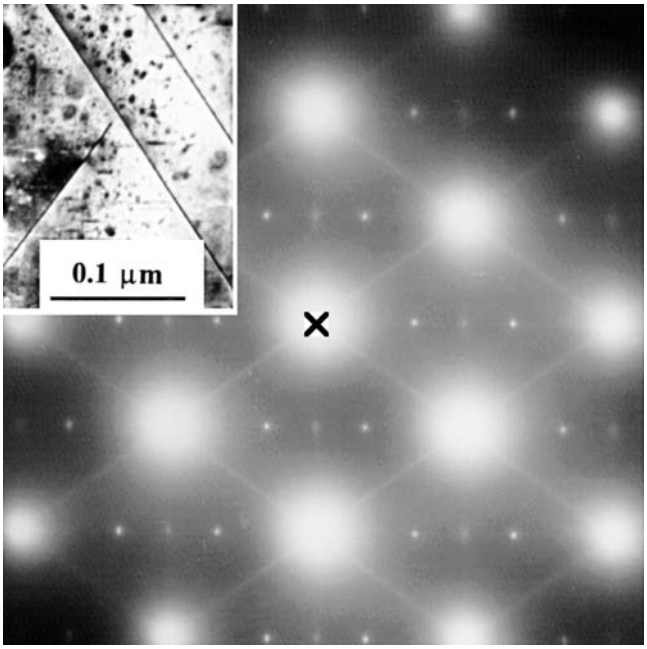


FIGURE 1.6. TEM DP from a thin foil of Al-Li-Cu containing various precipitate phases, shown in the inset image. The central spot (X) contains electrons that come directly through the foil and the other spots and lines are diffracted electrons which are scattered from the different crystal planes.

shows a TEM DP that contains information on the crystal structure, lattice repeat distance, and specimen shape (as well as being a most striking pattern). We'll see that the pattern can always be related to the image of the area of the specimen from which it came, in this case shown in the inset. You will also see in Part 2 that, in addition to the things we just listed, if you converge the usually parallel TEM beam to a focused probe, then you can produce even more striking convergent-beam patterns (see Figure 2.13D) from which you can conduct a complete crystal-symmetry analysis of minuscule crystals, including such esoteric aspects as point-group and space-group determination. You shouldn't be surprised by now if we tell you that aberration correction can produce even better DPs, which are both sharper (by reducing chromatic aberration) and come from smaller regions of the specimen (by reducing C_s). The crystal structure produces no diffraction information in a VLM because of the relatively large wavelength of visible light.

KEY POINT TO REMEMBER

At all times the crystallographic information in the DP (and all the analytical information) can be related to the image of your specimen.

So a TEM can produce atomic-resolution images, it can generate a variety of signals telling you about your specimen chemistry and crystallography, and you can always produce images that are in focus. There are many other good reasons why you should use electron microscopes. We hope they will become evident as you read through this book. At the same time there are many reasons why you should *not* always seek to solve your problems with the TEM, and it is most important that you realize what the instrument *cannot* do, as well as knowing its capabilities.

1.3 LIMITATIONS OF THE TEM

1.3.A Sampling

All the above advantages of the TEM bring accompanying drawbacks. First of all, the price to pay for any high-resolution imaging technique is that you only look at a small part of your specimen at any one time. The higher the resolution therefore, the worse the sampling abilities of the instrument. Von Heimendahl (1980) reported a calculation by Swann around 1970 estimating that all TEMs, since they first became available commercially (~15 years), had only examined 0.3 mm^3 of material! Extending that calculation to the present time probably increases this volume to no more than 10^3 mm^3. So we have an instrument that is not a good sampling tool! This sampling problem only serves to emphasize that, if you're just starting your research, before you put your specimen in the TEM you must have examined it with techniques that offer poorer resolution but better sampling, such as your eyes, the VLM, and the SEM. In other words, know the forest before you start looking at the veins in the leaves on the trees.

1.3.B Interpreting Transmission Images

Another problem is that the TEM presents us with 2D images of 3D specimens, viewed in transmission. Our eyes and brain routinely understand reflected light images but are ill-equipped to interpret TEM images and so we must be cautious. Hayes illustrates this problem well by showing a picture of two rhinoceros side by side such that the head of one appears attached to the rear of the other (see Figure 1.7). As Hayes puts it "when we see this image we laugh" (because we understand its true nature in 3D) "but when we see equivalent (but more misleading) images in the TEM, we publish!" So beware of artifacts which abound in TEM images.

One aspect of this particular drawback (sometimes called the projection-limitation) is that generally all the TEM information that we talk about in this book (images, DPs, spectra) is averaged through the thickness of the specimen. In other words, a single TEM image has no depth sensitivity. As we noted in Figure 1.5 there often is information about the top and bottom surfaces of the thin foil, but this is not immediately apparent. So

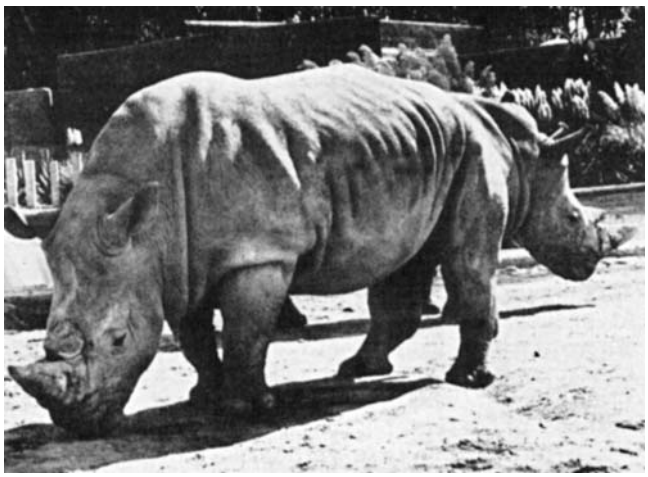


FIGURE 1.7. Photograph of two rhinos taken so that, in projection, they appear as one two-headed beast. Such projection artifacts in reflected-light images are easily discernible to the human eye but similar artifacts in TEM images are easily mistaken for 'real' features.

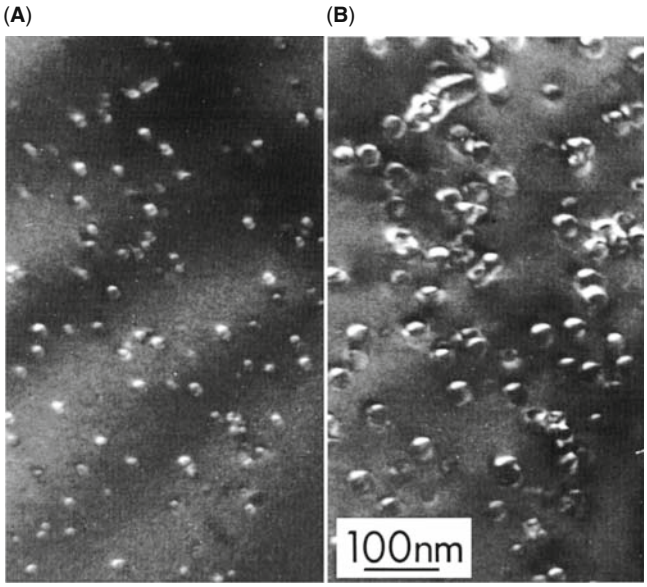


FIGURE 1.8. Beam damage (bright bubble-like regions) in quartz after bombardment with 125 keV electrons. With increasing time from (A) to (B) the damaged regions increase in size.

other techniques which are more clearly surface sensitive or depth sensitive, such as field-ion microscopy, scanning-probe microscopy, Auger spectroscopy, and Rutherford backscattering, are necessary complementary techniques if you want a full characterization of your specimen.

However, there has been progress in overcoming this limitation, which was much more of a problem for biologists interested in the shape of complex molecules, cells, and other natural structures. So they invented the technique of electron tomography, which uses a sequence of images taken at different tilts to create a 3D image, identical in principle to the more familiar medical CAT (computerized-axial tomography) scans using X-rays. Recently, there has been rapid improvement in specimen-holder design to permit full 360° rotation and, in combination with easy data storage and manipulation, nanotechnologists have begun to use this technique to look at complex 3D inorganic structures such as porous materials containing catalyst particles. This relatively new aspect of TEM for materials scientists is explored in depth in the companion text.

1.3.C Electron Beam Damage and Safety

A detrimental effect of ionizing radiation is that it can damage your specimen, particularly polymers (and most organics) or certain minerals and ceramics. Some aspects of beam damage are exacerbated at higher voltages, and with commercial instruments offering up to 400 kV, beam damage can now limit much of what we do in the TEM, even with refractory metals. The situation is even worse with more intense beams made possible because of advances in C_s correction. Figure 1.8 shows an area of a specimen damaged by high-energy electrons.

However, all is not lost and we can combine more intense electron sources with more sensitive electron detectors and use computer enhancement of noisy images to minimize the total dose received by the specimen to levels below the damage threshold. Minimum-dose microscopy techniques, often combined with specimen cooling (cryo-microscopy) and low-noise, charge-coupled device (CCD) cameras (see Chapters 7 and 31, respectively), are standard approaches in biological TEM and permit images to be obtained even when only a few hundred electrons/nm^2 are hitting the specimen. These approaches are finding increasing usage in TEM of materials where digital control of the beam in STEMs is another way to minimize radiation damage.

The combination of high kV beams with the intense electron sources that are available means that you can destroy almost any specimen, if you are not careful. At the same time comes the danger that should *never* be forgotten, that of exposing yourself to ionizing radiation. Modern TEMs are remarkably well engineered and designed with safety as a primary concern, but *never* forget that you are dealing with a potentially dangerous instrument that generates radiation levels that will kill tissue (and managed to damage some operators in the early days of the technique). So *never* modify your microscope in any way without consulting the manufacturer and without carrying out routine radiation-leak tests. If in doubt, don't do it!

1.3.D Specimen Preparation

Your specimens have to be thin if you're going to get any information using transmitted electrons in the TEM. 'Thin' is a relative term, but in this context it means electron transparent. For a specimen to be transparent to electrons, it must be thin enough to transmit sufficient electrons such that enough intensity falls on the screen, CCD, or photographic plate to give an interpretable image in a reasonable time. Generally this requirement is a function of the electron energy and the average atomic number (Z) of your specimen. Typically for 100-keV electrons, specimens of aluminum alloys up to $\sim$1 μm would be thin, while steel would be thin up to about several hundred nanometers. However, it is an axiom in TEM that, almost invariably, *thinner is better* and specimens $<$100 nm should be used wherever possible. In extreme cases such as doing HRTEM or electron spectrometry, specimen thicknesses $<$50 nm (even $<$10 nm) are essential. These demands become less strict as the beam voltage increases, but this is offset by the production of beam damage.

Again these words of caution are balanced by the development of a specialized specimen-preparation tool called a focused ion beam (FIB) which is essential to the semiconductor-device fabricators who use them (by the dozen) to prepare, in a matter of a few tens of minutes, thin foils of specific, individual gates or junctions from one of the many millions of such on a 12-in. VLSI wafer. The only drawback is that, to buy a FIB, you have to pay as much as it costs to buy a TEM. We'll talk about this and other methods of specimen preparation in Chapter 10 and in the companion text.

THIN SPECIMENS

A *major* limitation of the TEM is we need thin specimens. Methods to prepare thin specimens exist for almost all materials, and we talk about them in Chapter 10. But as a general rule, the thinning processes that we use do affect the specimens, changing both their structure and chemistry. So you need to be aware of the drawbacks of specimen preparation and learn to recognize the artifacts introduced by standard preparation methods.

A terminological distinction is worth noting here. The words sample and specimen are often used interchangeably but, in this text, we'll distinguish the specimen specifically as the thin piece of material that you insert into the TEM in a specimen holder and we'll assume that the specimen was thinned from a much larger, bulk sample of the material you are interested in studying. Sometimes we'll mix the terms to test your understanding.

So it should be obvious to you by now that while TEM and associated techniques are tremendously powerful characterization tools when used properly, they should *never* be used in isolation to solve a materials problem. You must understand your material at low magnification with your eyes and with VLM and SEM before venturing into TEM studies. Otherwise you may fall-foul of some of the limitations we have just listed. But you should also have got the message that we are constantly improving the technique and overcoming these limitations while, at the same time, making the positive aspects even better.

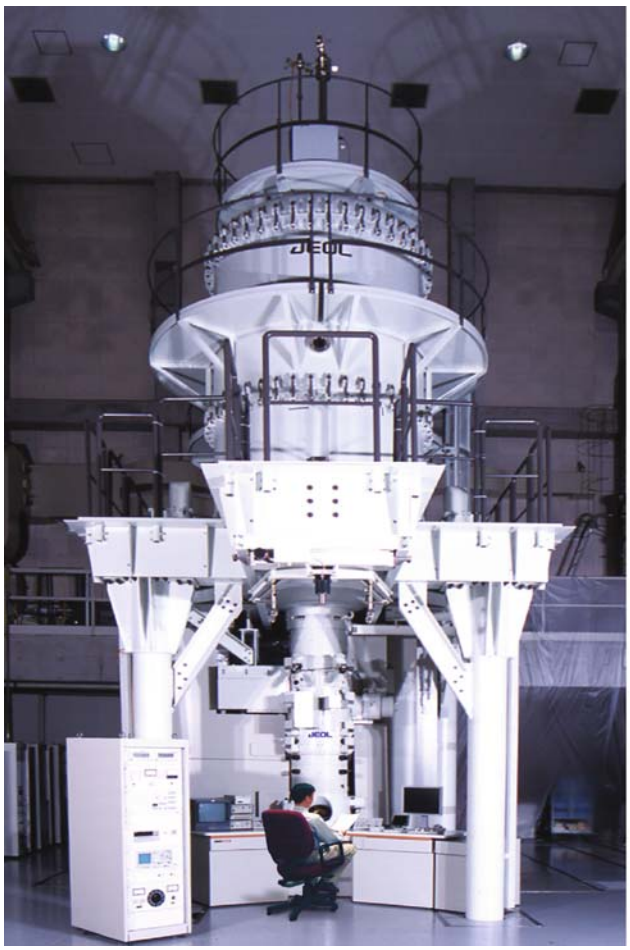
1.4 DIFFERENT KINDS OF TEMs

As you read through the previous sections you will have seen that TEMs come in a wide variety of types: HRTEMs, HVEMs, IVEMs, STEMs, and AEMs. Complete books have been written on each of these instruments, but it is our philosophy that all these are simply different forms of the basic TEM. So in this book we intend to treat them as such. Indeed a current 200 or 300 keV TEM can combine aspects of *all* the above microscope types. Figure 1.9 shows several of the different kinds of TEMs we have mentioned. It is instructive to consider some of the features of the instruments shown here. An HVEM usually requires a two- or three-story room; the operator shows the scale of this instrument. A modern machine essentially is an electron-optical column in which we can maintain a good vacuum but the lenses and most other functions can be controlled by one or more computers. Note that the DSTEM only has a flat-panel display: there is no (internal) viewing screen. This aspect is becoming a more popular design feature in TEMs because, if the screen is not in the microscope, then the operator doesn't have to be in the room or in the building or even in the same country. Removing the operator from proximity to the TEM overcomes many of the factors that limit the performance of the best instruments. Remote operation (or telepresence) is an increasingly attractive feature, which will give many more researchers access to the most sophisticated TEMs, as already happens in the world of astronomy with major telescopes.

1.5 SOME FUNDAMENTAL PROPERTIES OF ELECTRONS

Many times in the book we'll have to refer to some of the basic properties of electrons. You know that electrons show both particle and wave characteristics, illustrating one of the great puzzles of quantum physics that we all seem to accept without too much trouble. In fact the TEM routinely demonstrates both the particle and wave characteristics of the electron, repeating the electron analog of GI Taylor's famous experiment in which he

(A)



(B)

(C)

(D)

FIGURE 1.9. (Continued)

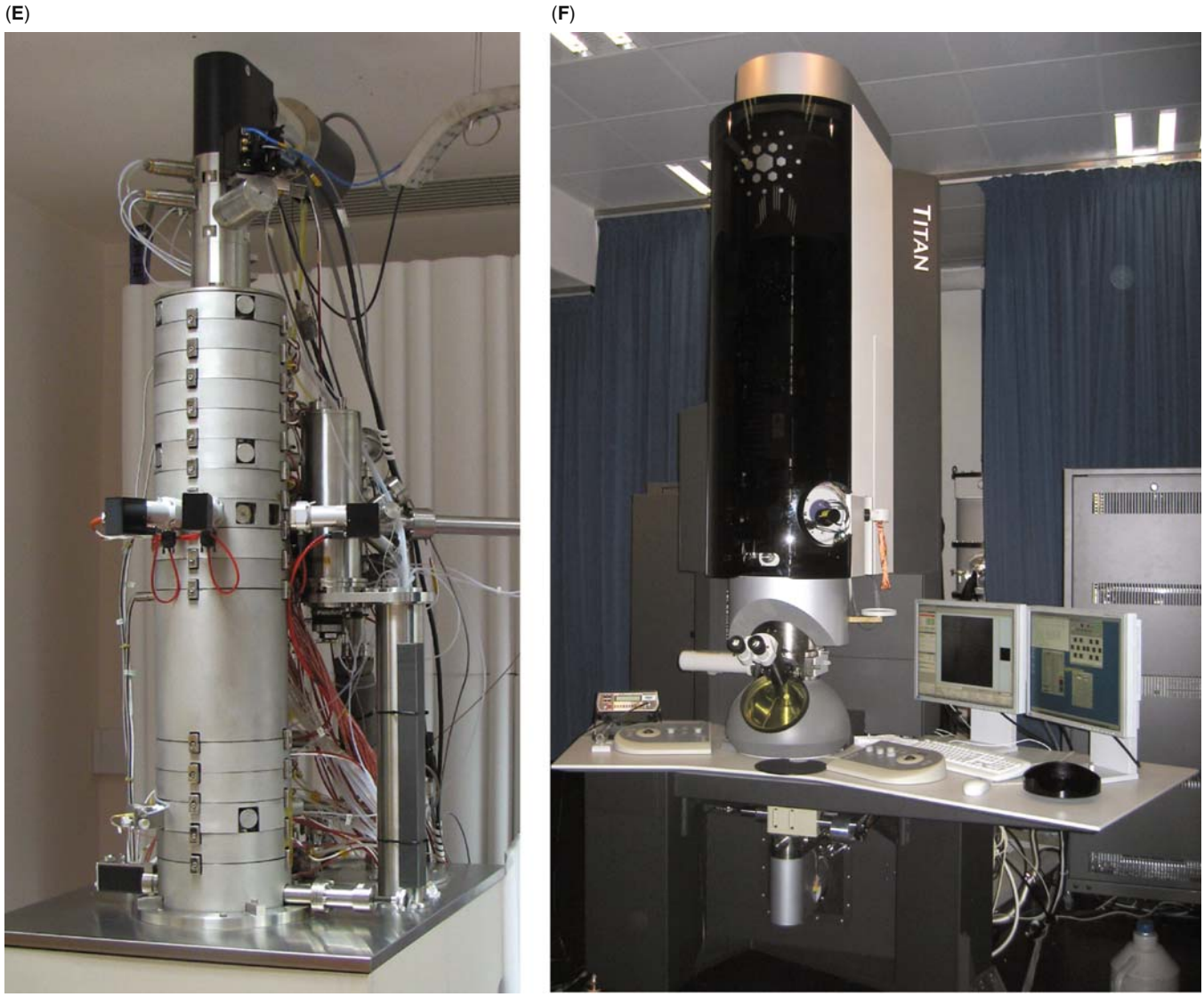


FIGURE 1.9. A selection of different commercial TEMs: (A) JEM 1.25 MeV HVEM. Note the size of the instrument; often the high-voltage tank is in another room above the column. (B) Zeiss HRTEM with a C_s corrector and an in-column energy filter. Note the large frame to provide high mechanical stability for the highest-resolution performance. (C) Hitachi 200 keV dedicated STEM; note the absence of a viewing chamber. Such instruments are often designed to aid failure analysis for the semiconductor device manufacturers. Specimens thinned from wafers on the production line can be easily transferred and examined. (D) JEOL 200 keV TEM/STEM; note also the absence of a viewing chamber. (E) Nion 200 keV ultrahigh vacuum SuperSTEM; the only US-manufactured (S)TEM and current holder of the world record image resolution (F) FEI Titan. Comparison with Ruska's instrument (Figure 1.1), which is 70–80 years older than these instruments, is instructive.

demonstrated Young's slits interference patterns despite using such a weak light source that only one photon passed through either slit at any one time. The electron beam current in a TEM can be as high as ~ 0.1–$1\ \mu A$, which corresponds to about 10^{12} electrons passing through the specimen plane. But, as we'll see below, with 100 keV energy, these electrons travel at about 0.5c (actually $\sim 1.6 \times 10^8$ m/s), so they are separated by ~ 1.6 mm and this means that there is never more than one electron in the specimen at any one time. Nevertheless, electron diffraction and interference occur, both of which are wave phenomena, and require interaction between different electron waves. Despite this dilemma, we know a lot about the electron and its behavior, and some of the basic characteristics are summarized in Table 1.1, along with some relevant physical constants.

There are a few important equations that you should know. First of all, based on de Broglie's ideas of the wave-particle duality, we can relate the particle momentum p to its wavelength λ through Planck's constant, thus

$$\lambda = \frac{h}{p} \tag{1.3}$$

TABLE 1.1 Fundamental Constants and Definitions

Charge (e)	(–) 1.602×10^{-19} C
1 eV	1.602×10^{-19} J
Rest mass (m_0)	9.109×10^{-31} kg
Rest energy (m_0c^2)	511 keV
Kinetic energy (charge × voltage)	1.602×10^{-19} N m (for 1 volt potential) = J
Planck's constant (h)	6.626×10^{-34} N m s
1 A	1 C/s
Speed of light in vacuum (c)	2.998×10^{8} m/s

In the TEM we impart momentum to the electron by accelerating it through a potential drop, V, giving it a kinetic energy eV. This potential energy must equal the kinetic energy, thus

$$eV = \frac{m_0 v^2}{2} \tag{1.4}$$

Now we can equate the momentum p to the electron mass, m_0, times the velocity, v, and substituting for v from equation 1.4

$$p = m_0 v = (2m_0 eV)^{1/2} \tag{1.5}$$

These three simple equations define the relationship between the electron wavelength λ and the accelerating voltage of the electron microscope, V

$$\lambda = \frac{h}{(2m_0 eV)^{1/2}} \tag{1.6}$$

If you look back, this equation is equivalent to equation 1.2. The inverse relationship between λ and V introduces a very important concept: by increasing the accelerating voltage we decrease the wavelength of the electrons. So you, the operator, can do this whenever you wish!

Equations 1.2 and 1.6 are useful expressions for deducing ballpark estimates, but be careful to note the differences. We can use equation 1.6 to calculate the non-relativistic electron wavelength for typical commercial TEM operating voltages as listed in Table 1.2.

The simple treatment we just went through neglects relativistic effects and, unfortunately for electron microscopists, relativistic effects cannot be ignored at energies $> \sim$100 keV because the velocity of the electrons (as particles) becomes greater than half the speed of light! So to be exact we must modify equation 1.6 to give

$$\lambda = \frac{h}{\left[2m_0 eV\left(1 + \frac{eV}{2m_0 c^2}\right)\right]^{1/2}} \tag{1.7}$$

A full listing for many more voltages can easily be generated by putting equations 1.6 and 1.7 into a spreadsheet. The effect of relativity is greater for higher accelerating voltages as shown in Table 1.2 which comprises all the commercial TEM accelerating voltages.

There will be many times when it's useful to refer back to these numbers, especially when we consider the resolution of the microscope and when we need to make calculations about the way electrons interact with matter.

A word about units: as we noted above, we should all be using SI units. We don't for two reasons: first, some special units are ideal for the purpose at hand; second we forget to include special conversion factors in some formulas. The difference between, e.g., the Gaussian system of units and SI units is summarized in the invaluable reference by Fischbeck and Fischbeck (1987) or in the electronic version of the almost 100-year-old standard source, Kaye and Laby (1986) (URL #2), or on the NIST database from which you can quickly find any number or constant that you need (URL #3).

TABLE 1.2 Electron Properties as a Function of Accelerating Voltage

Accelerating voltage (kV)	Non-relativistic wavelength (nm)	Relativistic wavelength (nm)	Mass ($\times m_0$)	Velocity ($\times 10^8$ m/s)
100	0.00386	0.00370	1.196	1.644
120	0.00352	0.00335	1.235	1.759
200	0.00273	0.00251	1.391	2.086
300	0.00223	0.00197	1.587	2.330
400	0.00193	0.00164	1.783	2.484
1000	0.00122	0.00087	2.957	2.823

1.6 MICROSCOPY ON THE INTERNET/ WORLD WIDE WEB

TEM users are well integrated into the Internet and the World Wide Web (WWW) and this is a source of useful information (and also some useful knowledge!) about what's going on in the field. You can view research TEMs in real time over the Internet and, as we've seen, you may not only see other instruments but you may be able to operate them remotely (a quick search on telemicroscopy or telepresence microscopy will get you some useful hits). So rather than spend a week in sunny Storrs operating an advanced TEM you can (regretfully) do it from the comfort of your own office or lab (or even from the beach in Aruba if there's a good broadband connection). As Internet2 and the National Lambda Rail further penetrate research labs in the United States and equivalent systems spread throughout the world, the rapid exchange of experimental information and the parallel simulation of data via access to high-end computational resources will continue to expand the options available to microscopists and analysts.

In addition, specialized software packages are also available on the Web which allow you to carry out many of the advanced analyses that we will introduce in this text (e.g., DP analysis and image/diffraction/spectral simulation). In many cases access to this software is limited (i.e., you have to pay for it) but any serious microscopy operator should have access to such software on site. Sometimes it is useful to explore the possibilities before you purchase.

We have already referenced a few important URLs in this chapter and we will continue to do so throughout the book and give lists at the end of each chapter. We have tried to keep the referenced URLs to sites associated with longstanding, reputable organizations such as national labs, professional societies, and major publishers. A list of broadly useful sites is included below and, not surprisingly, much of this list is very different from the one that we gave more than a decade ago when the first edition of this text was published. Nothing much has changed in this respect and so we cannot guarantee that these same URLs will still be active when you get round to searching for them. Such is life in the age of the Web.

1.6.A Microscopy and Analysis-Related Web Sites

http://www.amc.anl.gov This is the best source for TEM information on the Web in the United States and it is run by NJ Zaluzec at Argonne National Laboratory (ANL). Through it you can get access to the Microscopy ListServer and a Software Library. There is a connection to the Microscopy & Microanalysis FTP Site and to access Software/Image Libraries. Other useful connections through this site include:

http://microscopy.com/MicroscopyListserver/ This is an email-based discussion forum giving members of the EM community a centralized Internet address to which questions/comments/answers in the various fields of microscopy or analysis can be rapidly distributed to a list of (subscribed) individuals by electronic mail.

http://www.microscopy.com/MMMeetingCalendar.pl Listing of currently planned meetings, symposia, and courses.

http://zaluzec.com/cgi-bin/ANLWWWListingSQL.pl?SearchOrg=society This is a listing of the Web sites of various national and international microscopy and analysis societies. There are also connections to university, government, and individual microscopy-related sites.

http://cimewww.epfl.ch/EMYP/emyp.html This is the home of the Electron Microscopy Yellow Pages which is a similar operation to the ANL site, but based at the Ecole Polytechnique Fédérale de Lausanne in Switzerland, run by P Stadelmann. The Yellow Pages contain electron microscopy laboratories, software for electron microscopy, learned societies, instruments, equipment and consulting education in electron microscopy, data and databases, news and publications-related sources of information, conferences, workshops and schools, and getting somewhere else on the Web.

http://cimesg1.epfl.ch/CIOL/ and **http://cimesg1.epfl.ch/CIOL/summary.html** Stadelmann also offers access to very sophisticated EMS software for high-resolution image analysis and diffraction simulation and much more detail about this is covered in Chapter 30 and in the companion text.

http://iucr.org/resources/commissions A great resource operated for the International Union of Crystallography by JCH Spence at Arizona State University giving listings of electron diffraction and microscopy-related software, etc.

http://www.numis.northwestern.edu/IUCR_CED/ L Marks resource site for electron microscopy and diffraction

http://tem.msae.wisc.edu/emdb/index.html NSF supported TEM data base maintained by P. Voyles.

1.6.B Microscopy and Analysis Software

There is a lot of software available on the WWW and this is an aspect of TEM which is changing on a rapid basis, but you can now buy excellent software packages for all the fundamental aspects of microscopy: diffraction, imaging, and analysis. Many of these programs will be referenced throughout the text, but here is a brief summary of

the best that are currently used (with an indication of the source of the software) some of which are still free! There are many more packages than we have listed here but these are the ones with which we are familiar.

- **Cross sections for electron scattering:** The NIST version 3.0 of this database provides values of differential elastic-scattering cross sections, total elastic-scattering cross sections, phase shifts, and transport cross sections for elements with atomic numbers from 1 to 96 and for electron energies between 50 eV and 300 keV (in steps of 1 eV). Free downloads at http://www.nist.gov/srd/nist64.htm.
- **CRISP**: It is a commercial package running under Windows on a PC. It is designed for imaging process of HRTEM images. It can be combined with ELD (see below) and is available from Calidris, Manhemsvägen 4, S-191 46 Solltuna, Sweden (46 8 625 00 41). www.calidris-em.com/crisp.htm—the site does not change very often.
- **DigitalMicrograph™ (DM)**: It is a complete system for the acquisition, control, and processing of digital images from any electron microscope and is the central software for the Gatan Microscopy Suite™ (GMS) (see below). Gatan is to TEM as Microsoft is to PC.
- **DTSA** (Desk-Top Spectrum Analyzer): The NIST/NIH Desktop Spectrum Analyzer generates, interprets, and analyzes X-ray spectra from specimens under electron bombardment. This remarkable software/database package simulates the experimental environment and emulates specimen properties to generate spectra reflecting the relevant physics, chemistry, and statistics of a real-world application. DTSA incorporates many widely accepted X-ray data analysis procedures. Technical Contact: johnhenry.scott@nist, (301) 975-4981; http://www.cstl.nist.gov/div837/Division/outputs/DTSA/DTSA.htm.
  Really essential for the X-ray analyst but currently only available from NIST for Macintosh users. A PC version has been developed by Masashi Watanabe at Lehigh University.
- **ELD**: It is a commercial package from the producers of CRISP running under Windows on a PC. It is intended for quantitative analysis of DPs and is available from Calidris, Manhemsvägen 4, S-191 46 Solltuna, Sweden (46 8 625 00 41). www.calidris-em.com/eld.htm—still not changing often.
- **ELP**: The original Gatan energy-loss acquisition, processing, and analysis program now incorporated as part of the Gatan Microscopy Suite™ (GMS) (see below).
- **EMS** and **jEMS**: Image simulation program; see the listing of its capabilities in Section 1.6.A.
- **Gatan Microscopy Suite™ (GMS)**: It permits data acquisition, processing, and analysis for a wide range of TEM applications including EELS analysis, and energy-filtered compositional mapping, DP analysis, and 3D tomography acquisition. Open to individual scripting for user-specific needs. Widely used by many microscopists and analysts for acquiring, analyzing, and processing any kind of image data. Use with a CCD camera that provides digital images from the TEM or interface to any STEM system. From Gatan Inc., 5933 Coronado Lane, Pleasanton, CA 94588, (925) 463-0200; info@gatan.com.
- **ImageJ**: It is a versatile and extremely powerful image-processing and analysis open-source freeware available from NIH (has largely replaced the popular NIH Image). Developed by Wayne Rasband and aided by input from more than 1400 users round the world, ImageJ runs on Linux, Mac OS 9, Mac OS X, Windows. It is the world's fastest pure Java image-processing program. It can filter a 2048 × 2048 image in 0.1 s (40 million pixels per second!). Open and save all supported data types as TIFF (uncompressed) or as raw data. Open and save GIF, JPEG, BMP, PNG, PGM, FITS, and ASCII. Open DICOM. Open TIFFs, GIFs, JPEGs, DICOMs, and raw data using a URL. Open and save many other formats using plugins. Supports smoothing, sharpening, edge detection, median filtering, and thresholding on both 8-bit grayscale and RGB color images. Measure area, mean, standard deviation, min, and max of selection or entire image. Measure lengths and angles. Use real-world measurement units such as millimeters. Calibrate using density standards. Generate histograms and profile plots and much more.
  Available from http://rsb.info.nih.gov/ij/.
  Also look for Image SXM, a version of NIH Image extended to handle the loading, display, and analysis of scanning images: http://www.liv.ac.uk/~sdb/ImageSXM/.
- **Maclispix**: A Macintosh-based image-processing program, which works in conjunction with NIH Image (see below) or ImageJ (see above); permits analysis of
  Stacks: Movies, Depth profiles, Cropping and saving of large data sets.
  Groups: Coordinated measurements, Color overlays, Scatter diagrams.
  Pixels: bit, byte, integer, RGB, real, complex.
  Statistical measurements. Signal/noise determinations.
  Also diffraction analysis, segmentation (blobbing and measurement), registration, principal component analysis.
  Developed by David Bright at NIST (micro@nist.gov) and can be downloaded (free) from http://www.nist.gov/lispix/MLxDoc/mlx.html.
- **MacTempas and CrystalKit**: MacTempas is a Macintosh-based image analysis program for the simulation of high-resolution images, DPs, and crystal structures. Features include

Full multislice calculation of HRTEM Images and dynamical DPs.
Automatic calculation of the correct unit cell for any given beam orientation.
Automatic selection of correct array size and shape for the multislice calculation.
Display of atomistic models from any direction in both color and shades of gray.
Display of projected potentials, Fourier coefficients, diffractogram of images.
Plotting of amplitudes/phases of diffracted beams vs. thickness, contrast transfer functions.
Database on the 230 space groups; calculates the associated symmetry operators and atomic coordinates.
http://www.totalresolution.com/MacTempas.html.

CrystalKit: It works within MacTempas or EMS and builds models of crystalline defects, from point defects to grain boundaries and precipitates. A geometric grain boundary involving several thousand atoms can be generated in a matter of minutes by specifying the orientation relation between the grains, the interface plane, and the zone axis. CrystalKit allows the user to freely rotate the crystal, identify planes, measure angles and distances between atoms, visually move atoms, delete atoms, and add new atoms or create an arbitrary path interface.
http://www.totalresolution.com/CrystalKit.html.

- **Monte Carlo Simulations**: Software to simulate electron-beam trajectories through materials for estimating the spatial resolution of X-ray analysis or the backscattered electron yield. The best source of information is NIST's Web site http://www.cstl.nist.gov/div837/837.02/epq/index.html.
  For thin specimens you should use David Joy's Monte Carlo program described in his book (Joy 1995) and available at the University of Tennessee Web site http://web.utk.edu/~srcutk/htm/simulati.htm.
- **NIH-Image** (and **ImageJ**): It is a public domain software from NIH, developed by Wayne Rasband for general image manipulation with a limited set of image-processing tools. Details can be found at http://rsb.info.nih.gov/nih-image/about.html but it has been largely replaced by ImageJ (see above).
- **Adobe Photoshop**: Professional page layout programs for presentations and labeling your figures. Still the standard and available for a nominal fee at http://www.adobe.com/products/photoshop/.
- **Diffraction-Pattern Indexing**: Start with http://emaps.mrl.uiuc.edu JM Zuo's excellent free site at the University of Illinois. After this you should check the EM Yellow Pages and the EMS software and also the Web site for the International Union of Crystallography (both already referenced in Section 1.6.A). Also try SingleCrystal™; part of the extensive CrystalMaker® suite of DP and crystal structure software. Free demo download at: http://www.crystalmaker.co.uk/singlecrystal/index.html.

CHAPTER SUMMARY

TEMs comprise a range of different instruments that make use of the properties of electrons, both as particles and as waves. The TEM generates a tremendous range of signals so we can obtain images, DPs, and several different kinds of spectra from the same small region of the specimen. In the rest of this book we'll take you through the fundamental aspects of TEM, trying to explain at all times *why* we do certain things in certain ways. We'll also explain to some degree *how* we carry out certain operations. Since many different commercial TEMs exist, there's no point in being specific in how to operate a particular TEM, but we can explain in a generic sense, in many cases, what you have to do to get your particular one to deliver the enormous amounts of information that it can generate. Not least of course, we also describe what you need to know to *interpret* the images, DPs, and spectra that you obtain.

If you count up the different imaging, diffraction, and spectroscopic operations that are available in a TEM there are almost 40 different modes of forming an image, DP, or spectrum, each of which produces different information about your specimen. (Your last homework, when you've finished the book, is to validate this claim!) Each of these information planes can be understood in a *quantitative* manner, to the extent that we can simulate all TEM images, DPs, and spectra in a computer. *No* other characterization technique comes close to the combination of versatility and quantification that is produced by this remarkable instrument, particularly when you consider the enormous range of magnifications over which the information is obtainable.

There is a wealth of other sources of information about TEM and in the general reference list below we give a selection of appropriate books that emphasize materials science and

nanotechnology (most of which remain in print) as well as some standard journals and regular conference proceedings. We also encourage you to get on the WWW and see what's out there, but be careful about the validity of the content on any particular site. It is probably reasonable to trust the Web sites from reputable institutions such as government labs, commercial TEM and related equipment manufacturers organizations where, once you get past any sales pitch, there's often great educational material, and of course the Web pages of professional electron microscopists such as that associated with this text (URL #4).

REFERENCES PHILOSOPHY

In the reference sections throughout the book, we will annotate the references but we won't necessarily spell them out in the text (we'll give clues); we don't want the chapters to look like journal papers themselves and we want to encourage you to look at these reference sections. (This chapter is the exception.) In addition, don't forget the references for the figures right at the end of the textbook where you'll find much more than is in the figure captions. We do encourage you to explore the literature—you really have no choice if you want to understand TEM.

GENERAL TEM BOOKS

Amelinckx, S, van Dyck, D, van Landuyt, J and van Tendeloo, G (Eds.) 1997 *Electron Microscopy: Principles and Fundamentals* VCH Weinheim Germany. An expensive text containing review articles by leading microscopists covering TEM and much more. A good overview to put TEM in context, but make sure your library buys this along with the even more expensive *Handbook of Microscopy: Applications in Materials Science, Solid State Physics and Chemistry* edited by the same authors and published by VCH in 1997.

DeGraef, M 2003 *Introduction to Conventional Transmission Microscopy* Cambridge University Press New York. A complementary text to this one, in many ways. It uses different materials specimens as a thread to introduce different techniques. Supported by an excellent Web site (http://ctem.web.cmu.edu/) but not really an introduction!

Edington, JW 1976 *Practical Electron Microscopy in Materials Science* Van Nostrand-Reinhold New York. The original out-of-print 1976 edition has been reprinted by TechBooks, 2600 Seskey Glen Court, Herndon, VA 22071. A very helpful, if outdated, text full of examples and hands-on operations; no AEM or HREM, just diffraction-based imaging.

Egerton, RF 2006 *Physical Principles of Electron Microscopy; An Introduction to TEM, SEM, and AEM* Springer New York. If you need a general introduction to EM, this is a good choice.

Ernst, F and Rühle, M (Eds.) 2003 *High-Resolution Imaging and Spectrometry of Materials* Springer Series in Materials Science 50 Springer Berlin. A collection of review articles covering some aspects of TEM and other high-resolution techniques.

Fultz, B and Howe, JM 2002 *Transmission Electron Microscopy and Diffractometry of Materials* 2nd Ed. Springer New York. A broad-based text emphasizing diffraction-based imaging and crystallography via studies with X-ray and electrons.

Goodhew, PJ, Humphreys, FJ and Beanland, R 2001 *Electron Microscopy and Analysis* 3rd Ed. Taylor & Francis New York. A succinct summary of SEM, TEM, and AEM.

Hall, CE 1953 *Introduction to Electron Microscopy* McGraw-Hill New York. A wonderful but nowadays neglected book. The level is very close to this text. Historically minded students will enjoy the preface.

Hawkes, PW and Spence, JCH (Eds.) 2007 *Science of Microscopy* Springer New York. Comprehensive, up-to-date, multi-author review of many forms of microscopy. Get your library to buy it.

Heidenreich, RD 1964 *Fundamentals of Transmission Electron Microscopy* Interscience Publisher New York NY. Another wonderful, but sometimes forgotten, classic.

Hirsch, PB, Howie, A, Nicholson, RB, Pashley, DW and Whelan, MJ 1977 *Electron Microscopy of Thin Crystals* 2nd Ed. Krieger Huntington NY. For many years, the 'Bible' for TEM users and still required reading for true TEM converts!

Marton, L 1968 *Early History of the Electron Microscope* San Francisco Press San Francisco.

McLaren, AC 1991 *Transmission Electron Microscopy of Minerals and Rocks* Cambridge University Press New York. Invaluable for the geologist or ceramist.

Reimer, L 1997 *Transmission Electron Microscopy; Physics of Image Formation and Microanalysis* 4th Ed. Springer New York. Essential reference text. Strong physics background required; never uses a few words where a triple integral will do.

Ruska, E 1980 *The Early Development of Electron Lenses and Electron Microscopy* (translated by T Mulvey) S Hirzel Verlag Stuttgart.

Sawyer, LC, Grubb, DT and Meyers, GF 2008 *Polymer Microscopy* 3rd Ed. Springer New York. An expensive but useful qualitative introduction to TEM and SEM of polymers.
Thomas, G and Goringe, MJ 1979 *Transmission Electron Microscopy of Metals* Wiley New York. Invaluable for classical imaging and diffraction topics. The original out-of-print 1979 edition has been reprinted by TechBooks, 2600 Seskey Glen Court, Herndon, VA 22071.
Watt, IM 1997 *The Principles and Practice of Electron Microscopy* 2nd Ed. CUP New York NY. A basic, practical introduction to SEM and TEM.
Wenk, H-R 1976 *Electron Microscopy in Mineralogy* Springer New York NY. Required reading for microscopy of geological or ceramic materials. From the library.
Yao, N and Wang, ZL 2005 *Microscopy for Nanotechnology* Kluwer New York. In-depth review articles on techniques for nano-characterization. Half of the 22 chapters are devoted to EM methods and more than 80% of those are TEM.

SPECIALIZED TEM BOOKS

Ahn, CC (Ed.) 2004 *Transmission Electron Energy Loss Spectrometry in Materials Science and the EELS Atlas* 2nd Ed. Wiley-VCH Berlin. An excellent, in-depth review of EELS and the best database for ionization-edge energies.
Brydson, R 2001 *Electron Energy Loss Spectroscopy* Bios (Royal Microsc. Soc.) Oxford UK. You must read this before going on to Egerton's classic.
Champness, PE 2001 *Electron Diffraction in the TEM* Bios (Royal Microsc. Soc.) Oxford UK. Outstanding, essential introductory text by a crystallographer who also knows TEM.
Cowley, JM (Ed.) 1992 *Electron Diffraction Techniques* Vols. 1 and 2 Oxford University Press New York. Another collection of excellent individual review articles.
Egerton, RF 1996 *Electron Energy Loss Spectroscopy in the Electron Microscope* 2nd Ed. Plenum Press New York. The quintessential text for this aspect of AEM.
Frank, J 1992 *Electron Tomography* Plenum Press New York. A biological EM text but the basis for much of what's going to happen in TEM of materials.
Garratt-Reed, AJ and Bell, DC 2002 *Energy-Dispersive X-ray Analysis in the Electron Microscope* Bios (Royal Microsc. Soc.) Oxford, UK. Basic introduction pitched at a similar level to this text.
Hawkes, PW and Kasper, E 1989, 1994 *Principles of Electron Optics* Vols. 1–3 Academic Press New York, 1900 pp. Comprehensive but advanced. The third volume deals with many aspects of imaging in the TEM, simulation and processing with ~118 pages of TEM references; an exceptional resource.
Head, AK, Humble, P, Clarebrough, LM, Morton, AJ and Forwood, CT 1973 *Computed Electron Micrographs and Defect Identification* North-Holland New York NY. Long out of print but often referenced.
Horiuchi, S 1994 *Fundamentals of High-Resolution Transmission Electron Microscopy* North-Holland Amsterdam. As it says; a one topic book.
Kirkland, EJ 1998 *Advanced Computing in Electron Microscopy* Plenum Press New York. Required reading for the mathematically inclined microscopist.
Jones, IP 1992 *Chemical Microanalysis Using Electron Beams* Institute of Materials London. Great introduction to AEM with lots of calculations to introduce the principles of quantitative analysis.
Loretto, MH 1994 *Electron Beam Analysis of Materials* 2nd Ed. Chapman and Hall New York. A concise overview of the subject.
Royal Microscopical Society Handbook series: a broad range of introductory texts covering many aspects of TEM but also SEM and visible-light microscopy. Easy reading and not expensive (URL #5).
Shindo, D and Oikawa, T 2002 *Analytical Electron Microscopy for Materials Science* Springer New York. Brief summary of X-ray and electron spectrometry.
Spence, JCH 2003 *High Resolution Electron Microscopy* 3rd Ed. Oxford University Press New York. Practical HREM combined with lots of sound theory.
Spence, JCH and Zuo, JM 1992 *Electron Microdiffraction* Plenum Press New York. Quantitative convergent-beam diffraction in great depth.
Tonomura, A 1999 *Electron Holography* Springer New York. A good introduction.
von Heimendahl, M 1980 *Electron Microscopy of Materials* Academic Press New York NY. An introductory-level text, no AEM or HRTEM component. From the library.
Wang, ZL 1995 *Elastic and Inelastic Scattering in Electron Diffraction and Imaging* Plenum Press New York. Everything you ever need to know about scattering of electrons in the TEM.

THE COMPANION TEXT

Throughout this chapter and the rest of the text we will refer to the companion text. This new textbook is not a required reading for all students but does contain chapters on special topics that are only covered briefly in the present text.

JOURNALS

Advances in Imaging and Electron Physics Ed. PW Hawkes. Peter Hawkes is an editor, author, and historian without peer in the field of electron microscopy. Any text he edits or book he writes is worth reading. This journal merges two long-running serial journals *Advances in Electronics and Electron Physics* and *Advances in Optical and Electron Microscopy*. Reviews on particle optics at high and low energies, image science and digital image processing, electromagnetic-wave propagation, electron microscopy, and computing methods associated with all these topics. http//www.elsevier.com/wps/find/bookdescription.cws_home/711044/description#description.

Journal of Electron Microscopy Official Journal of the Japanese Society for Microscopy, becoming a more widely appreciated resource now that the papers are all in English. Oxford University Press Oxford. http//jmicro.oxfordjournals.org/archive/.

Journal of Microscopy Official Journal of the Royal Microscopical Society, the International Society for Stereology, the Microscopical Society of Ireland, the Polish Society for Microscopy, and the Austrian Society for Electron Microscopy. Many groundbreaking papers in the TEM of materials have appeared in this journal despite the strong biological theme of many of the papers. Blackwell Publishing Ltd Oxford, UK. http//www.rms.org.uk/journal.shtml.

Micron The International Research and Review Journal for Microscopy. Elsevier Amsterdam The Netherlands. http//www.elsevier.com/wps/find/journaldescription.cws_home/475/description#description.

Microscopy and Microanalysis Official Journal of the Microscopy Society of America, Microbeam Analysis Society (USA), Microscopical Society of Canada/Société de Microscopie du Canada, Mexican Microscopy Society, Brazilian Society for Microscopy and Microanalysis, Venezuelan Society for Electron Microscopy, European Microbeam Analysis Society, Australian Microscopy and Microanalysis Society. Has the largest circulation of any EM journal. Cambridge University Press New York NY. http//www.msa.microscopy.org/MSAUnits/Journal/MscopyManalysis.html.

Microscopy, Microanalysis, Microstructure Official Journal of the Société Française des Microscopies. Publication continued by The European Physical Journal (Applied Physics) in 1998 since when it has lost much visibility, but there are key papers in here before 1998.

Microscopy Research and Technique A place to publish your new techniques and methods for microscopy, specimen preparation, or any related aspect of TEM. John Wiley & Sons Hoboken NJ. http//www3.interscience.wiley.com/cgi-bin/jhome/38527.

Ultramicroscopy An international journal affiliated with multiple national societies, committed to the advancement of new methods, tools, and theories in microscopy, where much cutting-edge TEM research is published. Look out for Peter Hawkes' occasional reviews of the state of the microscopy literature. Elsevier Amsterdam The Netherlands http//www.elsevier.com/wps/find/journaldescription.cws_home/505679/description#description.

SELECTED CONFERENCE PROCEEDINGS

Asia-Pacific Electron Microscopy Conference organized by the Committee of Asia-Pacific Societies for Electron Microscopy (CAPSEM) every 4 years (2012).

Australian Microscopy and Microanalysis Society every 2 years (2010).

Electron Microscopy and Analysis Group (EMAG), Institute of Physics, London, UK every 2 years (2011).

European Microbeam Analysis Society every 2 years (2011).

European Microscopy Congress organized by the European Microscopy Society, every 4 years (2012).

Inter-American Congress for Electron Microscopy organized by Comité Interamericano De Sociedades De Microscopía Electrónica (CIASEM), every 2 years (2011).

International Congress for Microscopy organized by the International Federation of Societies for Microscopy (IFSM), every 4 years (2010 in Rio de Janeiro). The world gathers here.

International Union of Microbeam Analysis Societies every 4 years (2012).

Japanese Society for Microscopy annually.

Microscopy & Microanalysis combined meeting of the Microscopy Society of America, the Microbeam Analysis Society and occasionally others (including the Canadians); annually.

USEFUL SOURCES OF NUMERICAL DATA AND CONSTANTS

Fischbeck, HJ and Fischbeck, KH 1987 *Formulas, Facts and Constants* 2nd Ed. Springer New York. An invaluable reference. SI units are described in Chapter 2. Relevant equations in Gaussian units are related to SI units on page 127.

Jackson, AG 1991 *Handbook for Crystallography for Electron Microscopists and Others* Springer New York. Ideal for the microscopist but see the review by JA Eades Microsc. Res. Tech. **21** 368.

Kaye, GWC and Laby, TH 1986 *Tables of Physical and Chemical Constants* 15th Ed. This invaluable resource was first published in 1911 and is now online courtesy of the National Physical Laboratory (UK) (URL #2).

SPECIFIC REFERENCES IN CHAPTER 1

Usually we will group these into topics but this time the topic is 'The Introduction.'

CBE (Council of Biology Editors) 1994 *Scientific Style and Format* 6th Ed. Cambridge University Press New York.

Cosslett, VE 1979 *The Cosslett Festschrift* J. Microsc. **117** 1–184.

Davisson, CJ and Germer, LH 1927 *Diffraction of Electrons by a Crystal of Nickel* Phys. Rev. **30** 705–740. Early work by Clinton J. Germer and Lester H. German; read the introduction for encouragement.

de Broglie, L 1925 *Recherches sur la Theorie des Quanta (Researches on the Quantum Theory)* Ann. Phys. **3** 22–128.

Ewald, PP 1962 *Fifty Years of X-ray Diffraction* International Union of Crystallography D. Reidel Dordrecht.

Fujita, H 1986 *History of Electron Microscopes* Business Center for Academic Societies Japan.

Goodman, P 1981 *Fifty Years of Electron Diffraction* International Union of Crystallography Utrecht.

Haguenau, F, Hawkes, PW, Hutchison JL, Satiat-Jeunemaitre, B, Simon, G and Williams, DB 2003 *Key Events in the History of Electron Microscopy* Microsc. Microanal. **9** 96–138.

Hashimoto, H 1986 *High Resolution and High Voltage Electron Microscopy* J. Elec. Microsc. Tech. **3** 1.

Hawkes, PW (Ed.) 1985 *The Beginnings of Electron Microscopy, Advances in Electronics and Electron Physics* Academic Press New York NY.

Hayes, TL 1980 *Biophysical Aspects of Scanning Electron Microscopy* SEM-1980 **1** 1–10 Ed. O Johari SEM Inc. AMF O'Hare IL.

Howie, A 2000 *A Symposium in Honor of Professor Archie Howie's 65th Birthday* Eds. PL Gai, ED Boyes, CB Carter, DJH Cockayne, LD Marks and SJ Pennycook. Microsc. Microanal. **6** 281–284.

Joy, DC 1995 *Monte Carlo Modeling for Electron Microscopy and Microanalysis* Oxford University Press New York.

Knoll, M and Ruska, E 1932 *Das Elektronenmikroskop (Electron Microscope)* Z. Phys. **78** 318–339.

Kossel, W and Möllenstedt, G 1939 *Electroneninterferenzen im Konvergenten* Ann. Phys. **36** 113–140.

Thomas, G 2002 *A Symposium in Honor of Gareth Thomas' 70th Birthday* Eds. DG Howitt, CB Carter, U Dahmen, R Gronsky, DB Williams and R Sinclair Microsc. Microanal. **8** 237–364.

Thomson, GP 1928 *Experiments on the Diffraction of Cathode Rays* Proc. Roy. Soc. Lond. **117** 600–609. George Paget Thomson was the son of J. J. Thomson; he shared the 1937 Nobel Prize for physics with Davisson. A good read.

Zeitler, E 2003 *Zeitler Festschrift* Eds. GA Botton, K Moore and D Su Micron **34** 119–260.

SPECIFIC URLs

1. http//www.nano.gov/html/facts/whatIsNano.html
2. http//www.kayelaby.npl.co.uk/
3. http//www.physics.nist.gov/cuu/Constants/archive1998.html
4. http//www.TEMbook.com
5. http//www.rms.org.uk/other-publications.shtml

SELF-ASSESSMENT QUESTIONS

Q1.1 Define ionizing radiation and explain why it's useful.
Q1.2 List four signals generated by electrons interacting with the atoms in your specimen.
Q1.3 What makes electrons interact strongly with matter?
Q1.4 What is the resolution or the resolving power of a microscope?
Q1.5 What limits the resolution of VLMs?
Q1.6 What limits the resolution of TEMs?
Q1.7 How can we get high resolution in a TEM?
Q1.8 Define depth of field for a specimen in the TEM.
Q1.9 Explain why sampling is a problem with TEMs and how to combat this limitation.

Q1.10 List three benefits of using a TEM rather than any other instrument to characterize nanoparticles.
Q1.11 What is the major difficulty with interpreting all TEM images?
Q1.12 Why is TEM so powerful a characterization tool?
Q1.13 What signals are most commonly used in the conventional TEM?
Q1.14 Define the acronyms TEM, AEM, STEM, HREM, HVEM. Are there any other kinds of electron microscopes?
Q1.15 Which two signals produced when electrons interact with matter are used in AEM?
Q1.16 What happens to the wavelength of the electron beam when the accelerating voltage of the TEM is increased?
Q1.17 What does 'thin' mean when referring to a TEM specimen?
Q1.18 Give a ballpark figure for a suitably 'thin' TEM specimen.
Q1.19 Why is TEM such an important experimental tool for materials scientists and nanotechnologists?
Q1.20 What is a typical electron beam current in a TEM?
Q1.21 What are the typical causes of specimen damage in the TEM?
Q1.22 What effect should we consider in any calculations when the energy of the electrons is 100 keV or higher and why?

TEXT-SPECIFIC QUESTIONS

T1.1 If you were a physicist in the 1920s why would you think that the fact that electrons can be diffracted by crystals is a good reason to develop an electron microscope?
T1.2 Calculate an approximate wavelength for 50 kV electrons via equations 1.1 and 1.2 and then estimate the Rayleigh-defined resolving power of a microscope using such electrons. We tell you to assume $\mu \sin \beta \sim \beta$ so what does this tell you about what numbers might be reasonable for the refractive index and numerical aperture in a TEM?
T1.3 Examine Figure 1.2A; draw a schematic diagram to summarize all that you can deduce about the crystallography of the specimen with the information given in the text.
T1.4 Look at the arrows in Figure 1.3. From what little you know at this stage about how electrons interact with matter, can you work out which signals are most likely to be used in a TEM and which are less likely to be used? (Hint: relate the diagram to the actual construction of the TEM.)
T1.5 Examine Figure 1.4; as a materials scientist, do you expect there to be more of the Ti (blue) alloying element present in the smaller precipitates in the matrix, the larger precipitates on the boundary or the large particle on the boundary, the big dark lumps up the middle of the image or in the matrix? Are there reasons to be concerned about trying to estimate compositions from the color in the image only?
T1.6 In Figure 1.5 some dislocations appear to begin or end in the material but you will recall that dislocation lines cannot begin or end in a crystal. (Refer, e.g., to Hull, D and Bacon, DJ Introduction to Dislocations, Butterworth-Heinemann, 2001 ISBN 0750646810 or Weertman, J and Weertman, J, Elementary Dislocation Theory, Oxford University Press; 1992 ISBN 0195069005.) So why do the lines appear to do just this?
T1.7 Why do you think it is useful to be able to relate the directions in the DP in Figure 1.6 to the directions of the plate-like precipitates in the inset in the same figure?
T1.8 Can you think of an analogy to Figure 1.7 involving images viewed in transmission using visible light? (Hint: think about looking at images of objects at varying distance apart viewed through a transparent medium, e.g., air or water).
T1.9 Examine Figure 1.8; what information in the images would lead you to infer that electron-beam damage was occurring and what could you do to try and avoid further damage? (Hint: what is the difference between the two images and what is causing it?)
T1.10 Would you expect a higher-voltage TEM to be capable of more damage to your specimen than a lower-voltage instrument?
T1.11 Use equations 1.6 and 1.7 in a spreadsheet to reproduce Table 1.2 and add in lines for electrons accelerated to 50 and 250 kV.
T1.12 Compare the usefulness of Google searches for 'TEM' and 'transmission electron microscopy,' 'AEM', and 'analytical electron microscopy.' What does this tell you about acronyms (TMBA?)
T1.13 See how many of the URLs listed in Section 1.6 are still active. Make your own list of useful URLs related to your own interests in electron microscopy.
T1.14 Confirm for yourself the calculation referred to in Section 1.4 that there is never more than one electron inside a typical thin specimen in a 100 keV TEM.
T1.15 Looking at Figure 1.9, you'll notice that high-voltage microscope is bigger than other ones. Can you think of any reasons why this is so?
T1.16 Why do some TEMs have viewing screens and some only have computer displays?
T1.17 Why might you want to build an ultrahigh vacuum TEM and why are there so few available?

2

Scattering and Diffraction

CHAPTER PREVIEW

The electron is a low-mass, negatively charged particle. As such, it can easily be deflected by passing close to other electrons or the positive nucleus of an atom. These Coulomb (electrostatic) interactions cause electron scattering, which is the process that makes TEM feasible. We will also discuss how the wave nature of the electron gives rise to diffraction effects. What we can already say is that if the electrons weren't scattered, there would be no mechanism to create TEM images or DPs and no source of spectroscopic data. So it is essential to understand both the particle approach and the wave approach to electron scattering in order to be able to interpret all the information that comes from a TEM. Electron scattering from materials is a reasonably complex area of physics, but it isn't necessary to develop a detailed comprehension of scattering theory to be a competent microscopist.

We start by defining some terminology that recurs throughout the book and then we introduce a few fundamental ideas that have to be grasped. These fundamental ideas can be summarized in the answers to four questions.

- What is the probability that an electron will be scattered when it passes near an atom?
- If the electron is scattered, what is the angle through which it is deviated?
- What is the average distance an electron travels between scattering events?
- Does the scattering event cause the electron to lose energy or not?

The answer to the first question concerning the probability of scattering is embodied in the idea of a cross section. The angle of scattering (usually determined through the differential cross section) is also important because it allows you as the TEM operator to control which electrons form the image and therefore what information is contained in the image. We will develop this point much further when we talk about image contrast in Part 3 of the book. The third question requires defining the mean-free path, an important concept given that we use thin specimens. The answer to the fourth question requires distinguishing elastic and inelastic scattering. The former constitutes most of the useful information in DPs obtained in the TEM, discussed in Part 2, while the latter is the source of X-rays and other spectroscopic signals discussed in Part 4. The distinction between electrons that lose energy and those that don't is important enough that we devote the next two chapters to each kind of electron and expand on the basic ideas introduced here.

2.1 WHY ARE WE INTERESTED IN ELECTRON SCATTERING?

We need to know about electron scattering because it is fundamental to all electron microscopy (not just TEM). You know well that your eyes cannot see any object unless it interacts with visible light in some way, for example through reflection or refraction, which are two forms of scattering (e.g., we can't see a light beam unless it is scattered by dust within it or it hits a surface). Similarly, we cannot see anything in EM images unless the specimen interacts with and scatters the electrons in some way. Thus, any non-scattering object is invisible and we will come across situations where 'invisibility' is an important criterion in TEM images. In the TEM we are usually most interested in those electrons that do not deviate far from the incident-electron direction. This is because the TEM is constructed to gather these electrons primarily and they also give us the information we seek about the internal structure and chemistry of the specimen. Other forms of scattering, such as electrons which are scattered through large angles (e.g., backscattered

electrons) and electrons ejected from the specimen (such as secondary electrons) are also of interest and we will not totally neglect them (although they are of much greater interest in the SEM where they provide atomic number contrast and surface-sensitive, topographical images, respectively).

WAVE AND PARTICLE

The electron is treated in two different ways throughout this book: in electron scattering it is a succession of particles, while in electron diffraction it is treated by wave theory. The analogy to X-rays or visible light would be to compare a beam of photons and an electromagnetic wave. However, you must always remember that electrons are charged particles and that Coulomb forces are very strong.

In this chapter we introduce the fundamental ideas of electron scattering; then, in the next two chapters, we discuss the two principal forms of scattering, namely, elastic and inelastic. Both forms are useful to us, but you'll see that the latter has the unfortunate side effect of being responsible for specimen damage and ultimately limits what we can do with a TEM.

To give you some feel for the importance of electron scattering, it is worth illustrating at this stage the basic principles of the TEM. You will see in due course that in a TEM we illuminate a thin specimen with a broad beam of electrons in which the intensity is uniform over the illuminated area.

We will often refer to incident and scattered electrons as beams of electrons, because we are dealing with many electrons, not an individual electron; these electrons are usually confined to well-defined paths in the microscope. So the electrons that hit the specimen are often called the incident beam and those scattered by the specimen are called scattered (or sometimes specifically, diffracted) beams. Electrons coming through a thin specimen are separated into those that suffer no angular deviation and those scattered though measurable angles. We call the undeviated electrons the 'direct beam' (in contrast to most texts that describe this as the 'transmitted beam' despite the fact that all electrons coming through the specimen have been 'transmitted'). As the electrons travel through the specimen they are either scattered by a variety of processes or they may remain unaffected. The end result, however, is that a non-uniform distribution of electrons emerges from the exit surface of the specimen, as shown schematically in Figure 2.1. It is this non-uniform distribution that contains all the structural, chemical, and other information about our specimen. So everything we learn about our specimen using TEM can be attributed to some form of electron scattering.

DIRECT BEAM

The beam that comes through the specimen, but remains parallel to the direction of the incident electrons is a very important beam, which we will term the **direct beam**.

We'll see in Chapter 9 that the TEM is constructed to display this non-uniform distribution of electrons in two different ways. First the *spatial distribution* (Figure 2.1A) of scattering can be observed as contrast in *images* of the specimen, and the *angular* distribution of scattering (Figure 2.1B) can be viewed in the form of scattering patterns, usually called *diffraction* patterns. A simple (and fundamental) operational step in the TEM is to use a restricting aperture, or an electron detector, of a size such that it only selects electrons that have suffered more or less than a certain angular deviation. So you as the operator have the ability to choose which electrons you want to use and thus you control what information

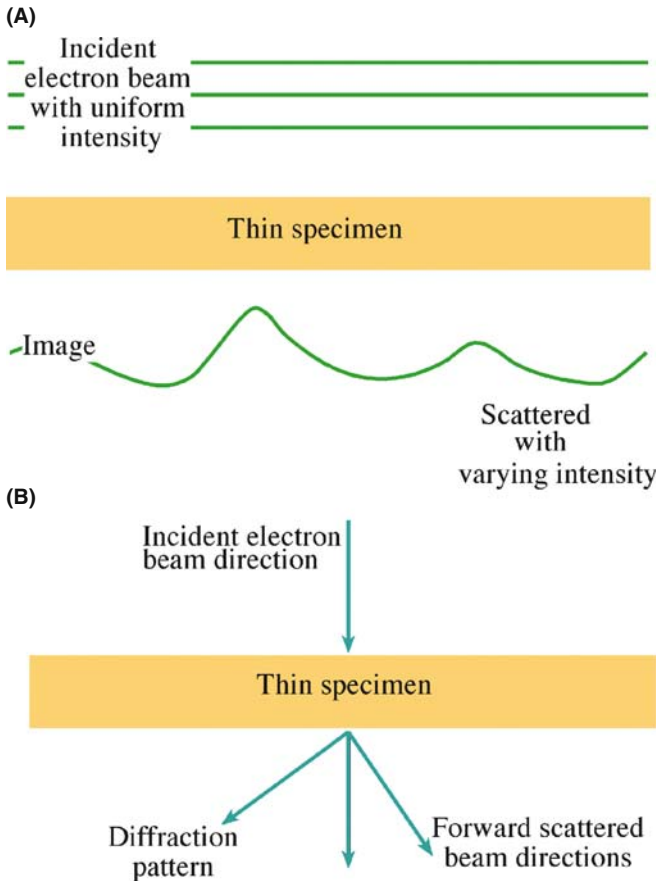


FIGURE 2.1. (A) A uniform intensity of electrons, represented by the horizontal lines, falls on a thin specimen. Scattering within the specimen changes both the spatial and angular distributions of the emerging electrons. The spatial distribution (intensity) is indicated by the wavy line. (B) The change in angular distribution is shown by an incident beam of electrons being transformed into several forward-scattered beams.

will be present in the image. Therefore, to comprehend these images, you have to understand what causes electrons to scatter in the first place. The same is true for DPs since you can also control (to a lesser extent) the angular-scattering distribution, e.g., by tilting your specimen.

We devote the whole of Part 2 to diffraction phenomena and Part 3 to images. Lastly, Part 4 deals with ways in which we use inelastic scattering for analytical electron microscopy (AEM) to study, e.g., the chemistry and the bonding of the atoms in our specimen.

2.2 TERMINOLOGY OF SCATTERING AND DIFFRACTION

Electron-scattering phenomena can be grouped in different ways. We've already used the most important terms: *elastic* and *inelastic* scattering. These terms, respectively, describe scattering that results in no loss of energy or in some measurable loss of energy (usually very small with respect to the beam energy). In either case, we can consider the beam electrons and specimen atoms as particles, and scattering of the incident electrons by the atoms in the specimen can often be approximated to something like billiard balls colliding. The billiard-ball analogy will be good through Section 2.7 after which we'll be talking about waves.

ELECTRON SCATTERING
This theme permeates the whole text and connects ALL aspects of TEM.

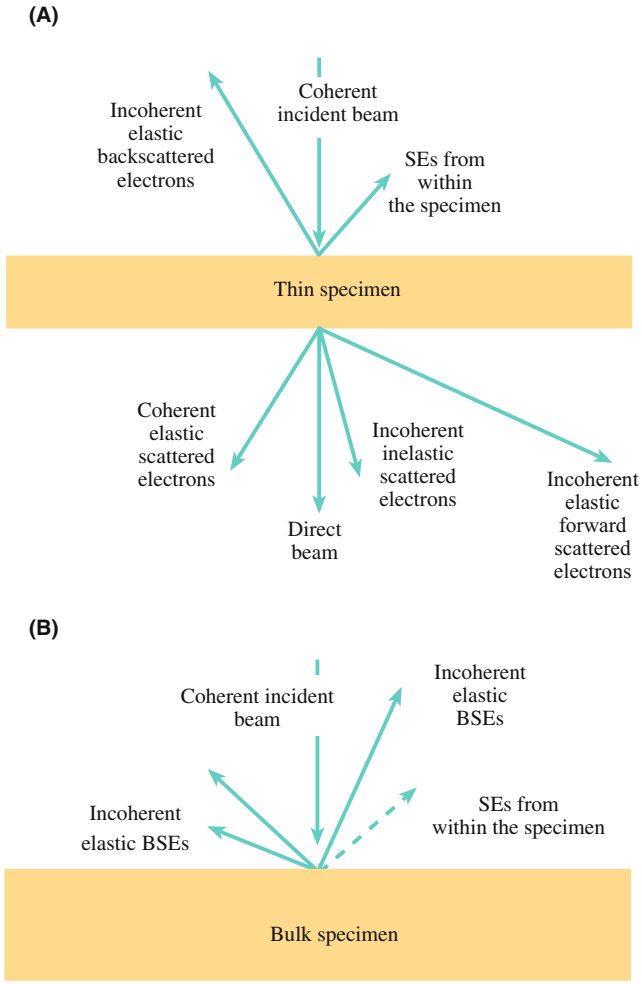


FIGURE 2.2. Different kinds of electron scattering from (A) a thin specimen and (B) a bulk specimen: a thin specimen permits electrons to be scattered in both the forward and back directions while a bulk specimen only backscatters the incident-beam electrons.

However, we can also separate scattered electrons into *coherent* and *incoherent*, which refers, of course, to their wave nature. These distinctions are related since elastically scattered electrons are usually coherent and inelastic electrons are usually incoherent (note the modifier 'usually'). Let's assume that the incident electron waves are coherent, that is, they are essentially in step (in phase) with one another and of a fixed wavelength, governed by the accelerating voltage. You'll see that this isn't a bad assumption in most circumstances. Then, coherently scattered electrons are those that remain in step and incoherently scattered electrons have no phase relationship, after interacting with the specimen.

The nature of the scattering can result in different angular distributions. Scattering can be either *forward scattering* or *back scattering* (usually written as one word) wherein the terms refer to the angle of scattering with respect to the incident beam and a specimen that is normal to that beam. (Note: you will sometimes see the term 'forward scattering' used in another sense.) If an electron is scattered through $< 90^\circ$, then it is forward scattered and $> 90^\circ$ it is backscattered. These various terms are related by the following general principles, summarized in Figure 2.2.

- Elastic scattering is usually coherent, if the specimen is thin and crystalline (think in terms of waves).
- Elastic scattering usually occurs at relatively low angles (1–10°), i.e., it is strongly peaked in the forward direction (waves).
- At higher angles ($>\sim 10^\circ$) elastic scattering becomes more incoherent (now think of particles).
- Inelastic scattering is almost always incoherent and is very low angle ($<1^\circ$) scattering (think particles).
- As the specimen gets thicker, fewer electrons are forward scattered and more are backscattered. Incoherent, backscattered electrons are the only remnants of the incident beam that emerge from bulk, non-transparent specimens (think particles).

The notion that electrons can be scattered through different angles is related to the fact that an electron can also be scattered more than once. Generally, the more scattering events, the greater the angle of scattering (although sometimes a second scattering event can redirect the electron back into the direct beam, so it appears to have undergone no scattering).

The simplest scattering process is *single scattering* and we often approximate all scattering within a TEM specimen to this process (i.e., an electron either undergoes a single-scattering event or it suffers no scattering). We'll see that this can be a very reasonable assumption if the specimen is very thin (something you can control). If the electron is scattered more than once, we use the term *plural scattering* and if it is scattered >20 times, we say *multiple scattering*. It is generally safe to assume that, unless you have a particularly thick specimen (through which you probably can't see anything anyhow), multiple scattering will not occur in the TEM. The greater the number of scattering events, the more difficult it is to predict what will happen to the electron and the more difficult it is to interpret the images, DPs, and spectra that we gather. So, once again, we emphasize the importance of the 'thinner is better' criterion, i.e., if you create thin enough specimens so that the single-scattering assumption is plausible, your TEM research will be easier.

Diffraction is a very special form of elastic scattering and the terminology used can be confusing. Collins' Dictionary defines *diffraction* as 'a deviation in the direction of a wave at the edge of an obstacle in its path' while *scattering* is defined as 'the process in which particles, atoms, etc., are deflected as a result of collision.' The word scatter can also be a noun denoting the act of scattering. So scattering might best apply to particles and diffraction to waves; both terms thus apply to electrons! You should also note that the term diffraction is not limited to Bragg diffraction which we'll emphasize in TEM; it refers to *any* interaction involving a wave, but many texts are not consistent in this respect.

DEFINE DIFFRACTION
An interaction between a wave of any kind and an object of any kind (Taylor 1987).

In the TEM we utilize the electrons that go through a specimen; it is important to note that such electrons are not simply 'transmitted' in the sense of visible light through window glass. Electrons are scattered mainly in the forward direction, i.e., parallel to the incident beam direction (and we've already noted the confusion between 'direct' and 'transmitted'). We'll tell you in a short while what fraction of the electrons are forward scattered and how this varies with the thickness of the specimen and atomic number of the 'target' atom. This scattering is a direct consequence of the fact that there is such a strong interaction between electrons and matter.

Forward scattering includes the direct beam, most elastic scattering, diffraction, particularly Bragg diffraction (see Chapter 3), refraction, and inelastic scattering (see Chapter 4). Because of forward scattering through our thin specimen, we see a DP or an image on the viewing screen, and detect an X-ray spectrum or an electron energy-loss spectrum outside the TEM column. But don't neglect backscattering; it is an important imaging mode in the SEM.

FORWARD SCATTERING
The cause of most of the signals used in the TEM.

When physicists consider the theory of electron interactions within a solid, they usually consider scattering of electrons by a single, isolated atom, then progress to agglomerations of atoms, first in amorphous solids and then in crystalline solids and we'll follow a similar path.

2.3 THE ANGLE OF SCATTERING

When an electron encounters a single, isolated atom it can be scattered in several ways which we will cover in the next two chapters. For the time being, let's imagine simply that, as shown in Figure 2.3, the electron is scattered through an angle θ (radians) into some solid angle ω, measured in steradians (sr). We have to define this angle first because you'll see that it plays an important role in the subsequent discussion of cross sections.

SEMI-ANGLE
Note that the scattering angle θ is in fact a semi-angle, not a total angle of scattering. Henceforth, whenever we say "scattering angle" we mean "scattering semi-angle."

Often we assume that θ is small enough such that $\sin\theta \approx \tan\theta \approx \theta$. When θ is this small, it is often convenient to use milliradians or mrads; 1 mrad is 0.0573°, 10 mrads is ~0.5°.

SMALL ANGLE
A convenient upper limit is <10 mrads.

The characteristics of the scattering event are controlled by factors such as the incident-electron energy and the atomic number/weight of the scattering atom. When we consider a specimen rather than a single

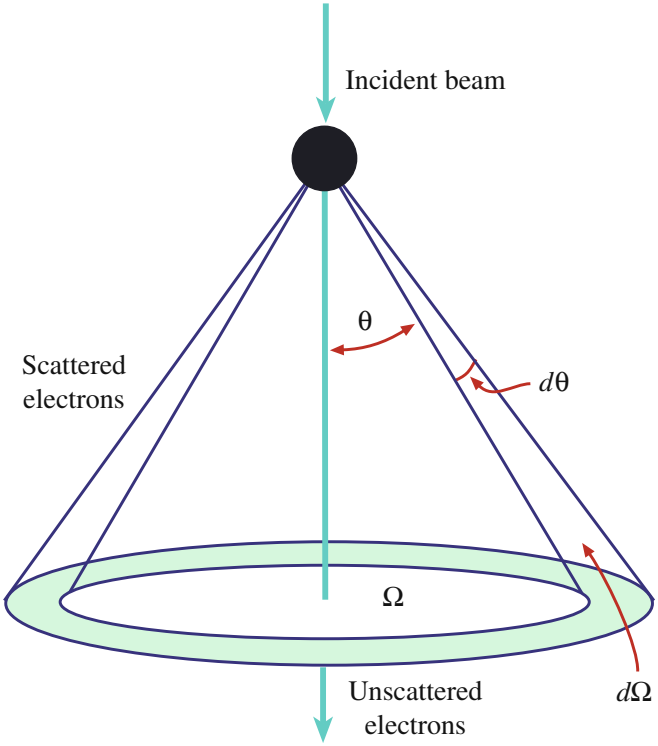


FIGURE 2.3. Electron scattering by a single isolated atom. The electrons are scattered through an angle θ and the total solid angle of scattering is Ω. An incremental increase in scattering angle $d\theta$ gives an incremental increase in a solid angle $d\Omega$, which is the basis for determining the differential scattering cross section.

atom, factors such as the thickness, density, crystallinity, and angle of the specimen to the incident beam also become important. To understand these variables, we need to examine the physics of scattering in more detail. Of necessity, we'll be rather brief and often imprecise since we're trying to condense much of Mott and Massey's substantial and classic textbook into just a few pages.

2.4 THE INTERACTION CROSS SECTION AND ITS DIFFERENTIAL

The chance of a particular electron undergoing any kind of interaction with an atom is determined by an interaction *cross section*. The concept of a cross section is well described by the following analogy given by Rudolf Peierls (Rhodes 1986)

> "If I throw a ball at a glass window one square foot in area, there may be one chance in ten that the window will break and nine chances in ten that the ball will just bounce. In the physicist's language this particular window, for a ball thrown in this particular way, has a disintegration (inelastic!) cross section of 0.1 square feet and an elastic cross section of 0.9 square feet."

So each possible interaction has a different cross section which depends on the energy of the particle, in our case the beam energy. The cross section (for which we'll use the Greek letter σ) has units of area (not square feet as used in Peierls' analogy, but a tiny fraction of the area of an atom termed a 'barn'). One barn is 10^{-28} m^2 (that's $(10^{-5}$ nm$)^2$) and the name arises because of the perverse sense of humor of some of the early atomic physicists who considered that this unimaginably small area is 'as big as a barn door.' The cross section does *not* represent a physical area but, when divided by the actual area of the atom, it represents a *probability* that a scattering event will occur.

2.4.A Scattering from an Isolated Atom

First of all we'll consider the scattering cross section for a single isolated atom, then extend the concept to a specimen with many atoms. We'll use a generalized form to start with in this chapter and then break down the concept of a total cross section into cross sections for individual processes such as elastic scattering and the various inelastic processes in the next two chapters.

SCATTERING PROBABILITY
The larger the cross section, the better the chances of scattering.

Following Heidenreich (1964), we can define the cross section (an area) in terms of the *effective radius* of a single, isolated atom, r

$$\sigma_{\text{atom}} = \pi r^2 \tag{2.1}$$

where r has a different value for each scattering process as we'll see in the next chapter. What interests us in the TEM is whether or not the scattering process deviates the incident-beam electrons outside a particular scattering angle θ such that, e.g., they do not go through the aperture in the lens or they miss the electron detector. So we have to know the *differential cross section* $(d\sigma/d\Omega)$ which describes the angular distribution of scattering from an atom. As shown in Figure 2.3 electrons are scattered through an angle θ into a solid angle Ω and there is a simple geometrical relationship between θ and Ω

$$\Omega = 2\pi(1 - \cos\theta) \tag{2.2}$$

and therefore

$$d\Omega - 2\pi \sin\theta d\theta \tag{2.3}$$

So the differential scattering cross section for a single, isolated atom can be written as

$$\frac{d\sigma}{d\Omega} = \frac{1}{2\pi \sin\theta}\frac{d\sigma}{d\theta} \tag{2.4}$$

Now, we can calculate σ_{atom} for scattering into all angles *greater than* θ by integrating equation 2.4 from θ to π

$$\sigma_{atom} = \int_{\theta}^{\pi} d\sigma = 2\pi \int_{\theta}^{\pi} \frac{d\sigma}{d\Omega} \sin\theta \, d\theta \tag{2.5}$$

The limits of the integration are governed by the fact that the values of θ can vary from 0 to π, depending on the specific type of scattering. If we work out the integral we find that σ decreases as θ increases (which makes physical sense). Since $d\sigma/d\Omega$ is often what is measured experimentally (but not in the TEM), equation 2.5 gives us an easy way to determine the cross section for a given atom (σ_{atom}) for all values of θ, by working out the integral of $(d\sigma/d\Omega)\sin\theta$ from 0 to π.

INTEGRATE

If we integrate from 0 to θ then we determine the cross section for scattering into all angles less than θ which is in fact more relevant to the TEM situation.

2.4.B Scattering from the Specimen

So let's move on from the cross section for a single isolated atom (with units of area) and consider that the specimen contains N atoms/unit volume. Therefore, we can define the total cross section for scattering from the specimen (in units of m^{-1}) as

$$\sigma_{total} = N\sigma_{atom} \tag{2.6}$$

Since $N = N_0\rho/A$ where N_0 is Avogadro's number (in units of atoms mol^{-1}), A is the atomic weight of the scattering atoms in the specimen (kg mol^{-1}) which has density ρ (kg m^{-3}), we can write

$$\sigma_{total} = N\sigma_{atom} = \frac{N_0\sigma_{atom}\rho}{A} \tag{2.7}$$

Thus, σ_{total} is the number of scattering events per unit distance that the electron travels through the specimen. If the specimen has thickness t, then the probability of scattering from the specimen is given by

$$\sigma_{total}t = \frac{N_0\sigma_{atom}(\rho t)}{A} \tag{2.8}$$

Here we've gathered together the product of ρ and t because this is called the mass thickness of the specimen (e.g., doubling ρ produces the same effect as doubling t) and we'll come across this term again when we discuss image contrast and also X-ray absorption. Equation 2.8 is an important expression, since it contains all the variables that affect the scattering probability from a real specimen. We'll use it again when we consider how certain kinds of image contrast arise in the TEM.

So we can now appreciate, through a few (rather simplified) equations, the relationship between the physics of electron scattering and the information we collect in the TEM.

Expressions for the cross section become more complicated as they are modified to give better approximations for the scattering in a real specimen, as we'll see in the next couple of chapters. However, the more complex equations don't alter the basic scattering behavior predicted by the simple equations we've just given.

2.4.C Some Numbers

Because of all the variables that affect σ_{atom} and σ_{total}, it is only possible to give a ballpark (barnyard?) value for the cross section. For TEM electron energies (100–400 keV), the elastic cross section is almost always the dominant component of the total scattering. If you look ahead to Figure 3.3, typical small-angle, elastic cross sections for transition metals bombarded by 100-keV electrons are $\sim 10^{-22}$ m^2. This is a good number to remember for typical elastic scattering. Inelastic cross sections are generally smaller and range from $\sim 10^{-22}$ m^2 down to 10^{-26} m^2 depending on the specific type of scattering and the material. Going back to equation 2.1, a typical scattering radius (r) is $\sim 10^{-11}$ m or ~ 0.01 nm, which might seem a bit small (about a tenth of an atomic radius) but, since the scattering is localized to the inner or core shells which are closer to the nucleus or to a particular electron-electron interaction, perhaps this isn't such a bad estimate, given all the caveats.

2.5 THE MEAN FREE PATH

Instead of using an area to describe the interaction we can use a length since the distance an electron travels between interactions with atoms is clearly going to be important when we are using thin specimens. This new parameter is then the average distance that the electron travels between scattering events. This distance is important because, if we know what it is, we can work out how thin we have to make our specimen, so plural scattering is not significant, thus making it easier to interpret our images and spectroscopic data in terms of single-scattering theory. The term σ_{total} can be expressed as the inverse

of the mean free path, λ. Because the dimensions of σ_{total} are m^{-1} (you can check this using equations 2.1 and 2.7) there is a simple expression for the mean free path λ which has units of length

$$\lambda = \frac{1}{\sigma_{total}} = \frac{A}{N_0 \sigma_{atom} \rho} \qquad (2.9)$$

Typical values of λ for scattering at TEM voltages are of the order of tens of nm, so single-scattering approximations imply specimen thicknesses of this order. It is, unfortunately, conventional to use λ to denote the mean free path; it is *not* the wavelength of the electron. From equation 2.9 we can define a probability of scattering p as the electron travels through a specimen thickness t

$$p = \frac{1}{\lambda} = \frac{N_0 \sigma_{atom} (\rho t)}{A} \qquad (2.10)$$

which is $(\sigma_{total} t)$ from equation 2.8.

Although computational resources are constantly improving, our knowledge of the values of σ, λ, and θ is imprecise at best, particularly at the 100–400 keV beam energies used in TEMs. Cross sections and mean free paths for particular scattering events may only be known within a factor of two, but we can often measure θ very precisely in the TEM. We can combine all our knowledge of scattering in Monte Carlo simulations to predict the electron paths as a beam is scattered through a thin foil.

MONTE CARLO SIMULATION

So called because random numbers are used in the computer programs; the outcome is always predicted by statistics!

The first Monte Carlo calculation was developed by two of the United States' foremost mathematicians, John von Neumann and Stanley Ulam, at Los Alamos in the late 1940s. Ulam actually rolled dice and made hand (!) calculations to determine the paths of neutrons through deuterium and tritium which proved that Teller's design for the 'Super' (H-bomb) was not feasible (Rhodes 1995). Monte Carlo methods are used more often in SEM imaging and X-ray calculations (see, e.g., NIST's Web site (URL #1), Joy 1995, Goldstein et al. 2003) but they have a role in TEM in determining the expected spatial resolution of X-ray analysis as we'll discuss in Chapter 36. Figure 2.4 shows Monte Carlo simulations of electron paths through thin foils of Cu and Au.

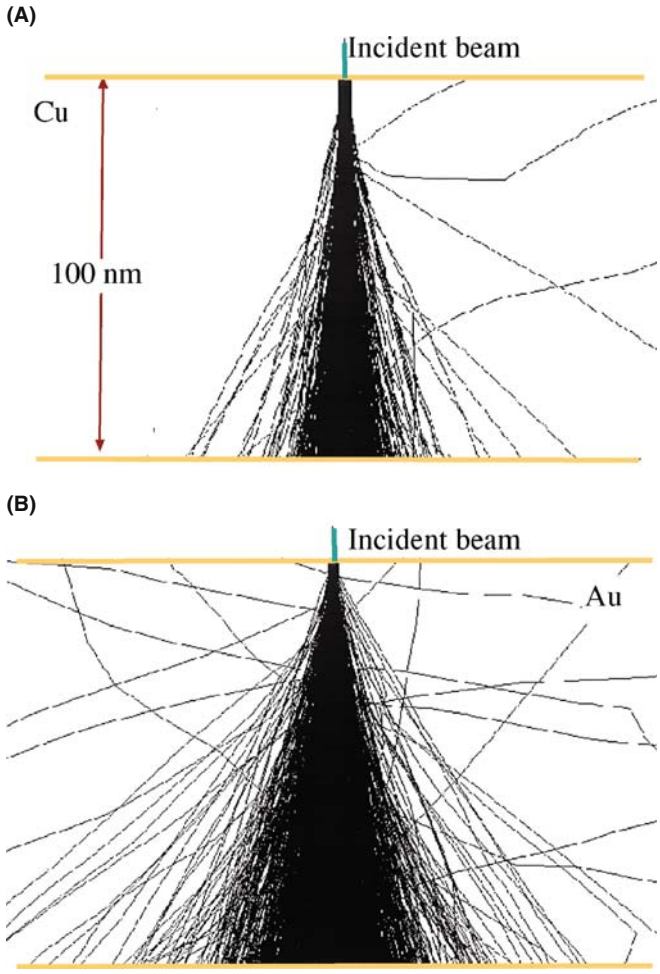


FIGURE 2.4. Monte Carlo simulation of the paths followed by 10^3 100-keV electrons as they pass through thin foils of (A) Cu and (B) Au. Notice the increase in scattering angle with atomic number and the small number of electrons that are scattered through >90°.

2.6 HOW WE USE SCATTERING IN THE TEM

So why have we made you go through all this math? Because when you select electrons that have been scattered through a certain angle (choosing a θ), you are changing the effective scattering cross section (σ_θ), because the scattering strength generally decreases as the angle of scattering increases. Therefore, there will generally be less scattering at higher angles, which explains why we said at the start of the chapter that we are mainly interested in forward scattering in the TEM. Most scattered electrons are contained within ±5° of the direct beam.

300 VERSUS 100 kV

Total σ decreases as E_0 increases; electron scattering at 300 kV will be less than at 100 kV. Higher-density regions of your specimen scatter more than lower-density regions. The target becomes smaller as the bullets become faster!

You also have control of the scattering cross section in other ways. First, the accelerating voltage, which determines the electron energy E_0 (eV), will affect the cross section as implied in equation 2.3 (specifically for elastic scattering). In fact, for all forms of scattering, the total cross section decreases as E_0 increases. Therefore, intermediate- and higher-voltage TEMs will result in *less* electron scattering than typical 100-kV instruments and, as we'll see in Chapter 4, this has important implications for electron-beam damage in delicate specimens, such as polymers. Second, and more intuitively, you can choose specimens with different densities. Denser materials scatter more strongly, so you have to make them thinner to keep the single-scattering approximation valid.

We shall see in the next two chapters that the effect of the atomic number of the atom is more important in elastic than inelastic scattering and, as Z increases, elastic scattering dominates. This behavior helps when we consider ways to enhance scattering (and therefore contrast) in low-Z materials such as polymers and biological tissue.

2.7 COMPARISON TO X-RAY DIFFRACTION

There is a very good reason why electrons are used in microscopy: they have a 'suitable interaction' with matter. Most descriptions of the interaction of electrons with matter are based on scattering. You will come across such topics as kinematical scattering and dynamical scattering in addition to elastic and inelastic scattering, and we will use the formalism of a scattering factor to describe the process mathematically. It is this scattering process that varies with the structure or composition of the specimen, permitting us ultimately to image a microstructure, record a DP, or collect a spectrum. We'll see in the next chapter that scattering factors are used when we consider electrons as waves and their diffraction as a specific form of scattering.

So now it's time to move from billiard balls to waves. Historically, it was diffraction that provided most of the crystallographic information we have about materials, and the majority of those studies used X-rays. This is partly why X-ray diffraction is so well documented in the scientific literature. A good understanding of X-ray diffraction helps considerably in understanding electron diffraction; however, the primary processes by which electrons are scattered are very different to the processes by which X-rays are scattered. Electron scattering is much more complex.

X-rays are scattered by the *electrons* in a material through an interaction between the negatively charged electrons and the electromagnetic field of the incoming X-rays. The electrons in the specimen respond to the applied field of the X-ray flux, oscillating with the period of the X-ray beam. These accelerated charged particles then emit their own electromagnetic field, identical in wavelength and phase to the incident X-rays. The resultant field, which propagates radially from every scattering source, is called the scattered wave.

ELECTRONS VERSUS X-RAYS

Electrons are scattered much more strongly than X-rays.

Electrons are scattered by *both* the electrons and the nuclei in a material; the incoming negatively charged electrons interact with the local electromagnetic fields of the specimen. The incoming electrons are therefore directly scattered by the specimen; it is not a field-to-field exchange as occurs for the case of X-rays.

2.8 FRAUNHOFER AND FRESNEL DIFFRACTION

Diffraction of visible light is well understood, so we should carry over as much of the analysis as possible. Optics is a venerable discipline with a history of several hundred years and what we're trying to do here is condense the principal messages from classic texts such as Hecht (2003) into a few pages. So, as with electron scattering, we'll be making a few simplifications. If you have any experience with diffraction of visible light you will have encountered Fraunhofer and Fresnel diffraction.

- Fraunhofer diffraction occurs when a flat wavefront interacts with an object. Since a wave emitted by a point becomes planar at large distances, this is known as far-field diffraction.
- Fresnel diffraction occurs when it's not Fraunhofer. This case is also known as near-field diffraction.

We will see later that electron-diffraction patterns correspond closely to the Fraunhofer case while we 'see' the effects of Fresnel diffraction in our images.

In TEM we will find both forms of diffraction. We will briefly go through the Huygens' explanation of how a wave propagates, then consider Fraunhofer diffraction from two slits (Young's slits) and then extend this process to many slits. So why discuss these topics now? There are two reasons for reviewing this analysis

- It reminds us that coherent interference is purely a matter of physical optics.
- We can introduce the concept of phasor diagrams which we'll use in later chapters.

Huygens explained the propagation of any wavefront by imagining that each point on the wavefront itself acts as a new source for a spherical wavelet. The wavelets interfere with one another to give the new wavefront and the process is repeated.

2.9 DIFFRACTION OF LIGHT FROM SLITS AND HOLES

In this section we will very briefly review the topic known as physical or geometric optics as it relates to diffraction. Much of what we know about diffraction of electron waves has been carried over from the understanding of the diffraction of visible light and X-rays. There are textbooks on this topic if you don't vaguely remember it from high school.

Two slits (the Young's slits experiment)

We start with diffraction caused by a wavefront incident on a pair of very narrow slits. We then select just two of the Huygens wavelets; these wavelets then must have the same phase at the slits. As they propagate past the slits, their phases differ depending on the position of the detector. The important term is the path difference $L = d \sin\theta$ as shown in Figure 2.5. The two wavelets propagating in direction $\boldsymbol{r}$ have a path difference of L and a phase difference of $2\pi L/\lambda$, i.e., $2\pi d \sin\theta/\lambda$. If d and λ are such that this phase difference is actually a multiple of 2π (so $d \sin\theta/\lambda =$ an integer, n) then the rays are again in phase and their amplitudes add. The condition for this additive interference is thus that $d \sin\theta = n\lambda$. Therefore, there is an inverse relationship between d and θ for a given d; as d decreases, $\sin\theta$ increases. If we think of each wavelet as having an amplitude and a phase we can represent each by a vector—a phasor. When the phasors are parallel to one another (in phase) they add; when they are antiparallel, they cancel (since they have equal lengths). A phasor diagram is a way to plot the amplitude and phase of the total scattered wave; in other words, when we add the amplitudes of beams we must take into account their phase.

THE INVERSE RELATIONSHIP

$\theta \propto d^{-1}$ solely due to the positions of the slits. We'll come across an identical relationship when we talk about electron diffraction.

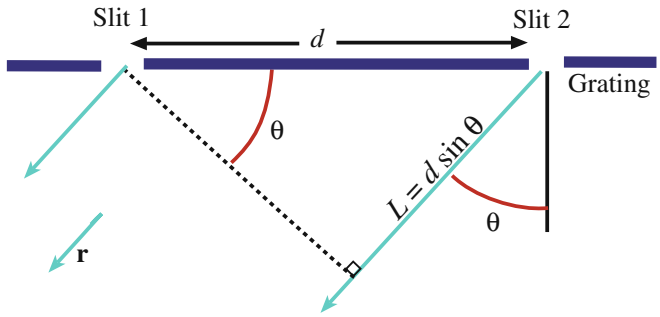


FIGURE 2.5. An incident plane wave is scattered by two slits, distance d apart. The scattered waves are in phase when the path difference $d \sin\theta$ is $n\lambda$.

Many slits (the diffraction grating)

When we extend this analysis to more than two slits we see a similar result but with added subsidiary peaks. The origin of the subsidiary peaks can best be illustrated by considering a series of phasor diagrams. (We'll find similar diagrams useful when we discuss TEM images in Chapter 27.) We'll examine the case of five slits. Each of the polyhedra in Figure 2.6 represents a different value of θ. When θ is zero, the five rays are all in phase and we simply add all of the amplitudes (the phasor vectors are aligned); as θ increases the rays become out of phase but the phasors can still add to give a large resultant vector but can also add to give zero. For example, when θ is exactly 72° (360°/5 for 5 slits) the phasor diagram is a closed pentagon (shown in the figure) and the resultant amplitude is zero. This process repeats at 144° (2×720°/5) and 216° (3×360°/5). In between these values at 108°

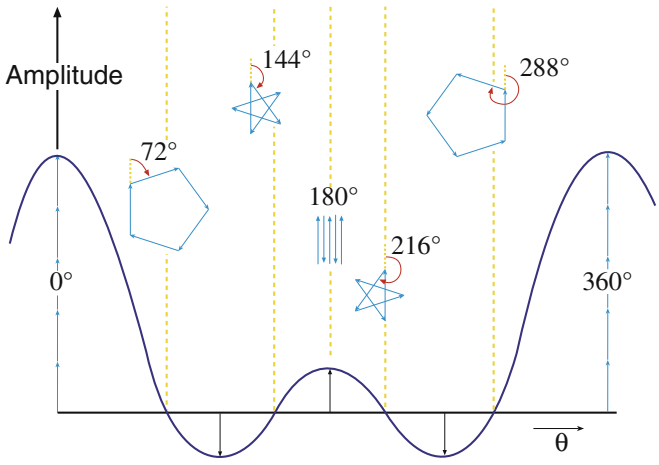


FIGURE 2.6. A phasor diagram showing how the total amplitude produced by summing five waves produced by five slits varies with the phase angle θ between the different waves. The individual phasors from each of the five slits sum to create a total amplitude of zero at $\theta = 72°, 144°, 216°$, and 288°, large positive amplitudes at $\theta = 0°$ and 360°, a single-phasor negative amplitude at $\theta = 108°$ and 252°, and a single-phasor positive amplitude at $\theta = 180°$. Remember the intensity is governed by the square of the amplitude, so positive and negative values both contribute to diffracted intensity.

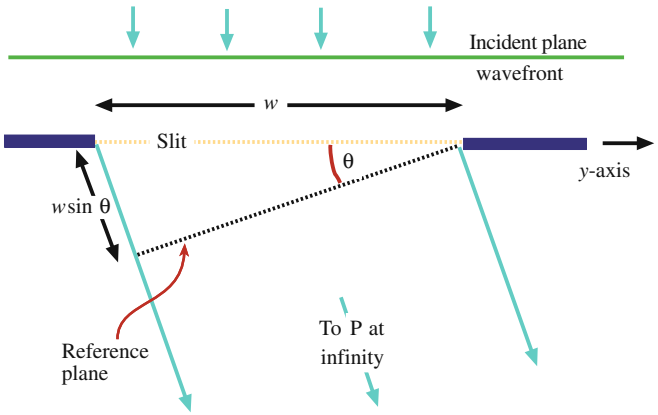


FIGURE 2.7. Geometry for the scattering from an individual slit.

$(1.5\times360^\circ/5)$ we produce a local maximum in amplitude which is repeated at 180° $(2.5\times360^\circ/5)$. If we plot the amplitude as a function of θ, we produce the curve with a series of subsidiary maxima shown in Figure 2.6. From this figure you see that the amplitude is a strong function of θ and you'll learn in the next chapter that the electron intensity (which is what we see in images and DPs) is proportional to the square of the amplitude (so negative amplitudes don't mean anything) and the scattered electron intensity is, therefore, a similarly strong function of θ.

A single wide slit

What happens if we allow the slit to have some width as shown in Figure 2.7? Now the rays from *within* a single slit will interfere with each other. We can think of the single slit as being many adjacent slits of width δw. Imagine dividing the one slit into 11 slits of width $\delta w/11$. This one slit would then produce a phasor diagram as shown in Figure 2.8; if we make δw increasingly small, the phasor diagram becomes a curve: instead of having Figure 2.8, we have Figure 2.9 (for several different values of θ). If you do the full analysis you'd find that the amplitude from a single slit varies as $A = A_0\phi^{-1}\sin\phi$ where ϕ is the phase $\pi w \sin\theta/\lambda$ for a slit of width w (which reminds you of the analysis from Figure 2.5). For one slit, we would see a zero in the phasor diagram when $\phi = \pm n\pi$. If we plot the intensity (rather than A) we obtain the Airy plot shown in Figure 2.10.

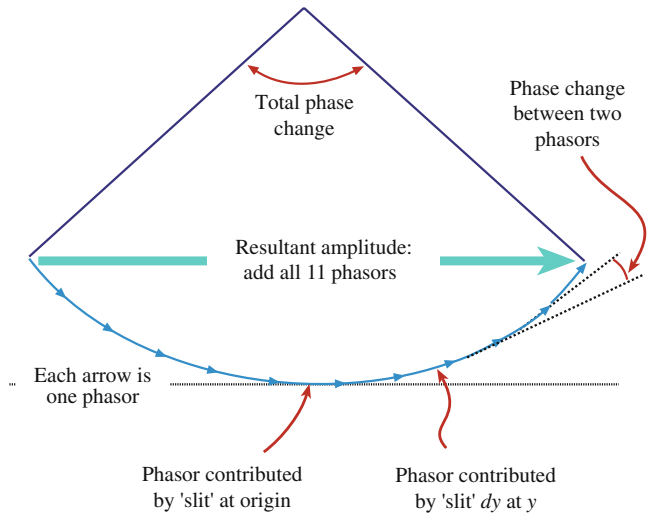


FIGURE 2.8. How the phasors from within an individual slit can be added to give the total phasor for the slit shown in Figure 2.7.

THE AIRY DISK

The disk of radius $r = 1.22\lambda/D$ is named after Airy and is one of the fundamental limits on the achievable resolution in TEM, as we will discuss in Chapter 6. If we introduce *any* aperture into *any* microscope we will limit the ultimate resolution of the instrument.

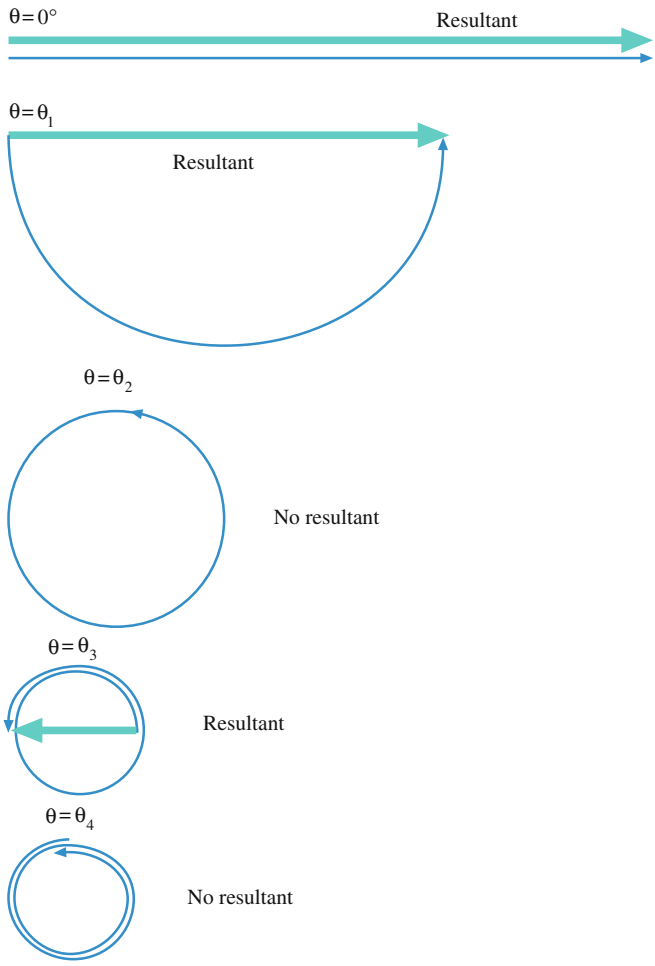


FIGURE 2.9. How a single slit can produce a beam which has zero amplitude for certain values of θ in Figure 2.7. The circles are directly comparable to the polyhedra in Figure 2.6. The total length of the phasor increments (from each dy) is the same in each figure.

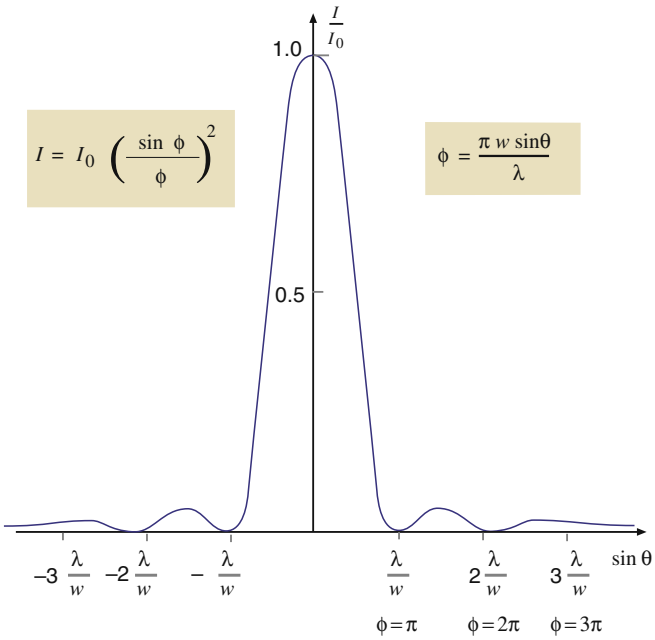


FIGURE 2.10. The plot of the resulting intensity for scattering from the slit shown in Figure 2.7; this is known as the Fraunhofer DP from a single slit; w is the slit width defined in Figure 2.7.

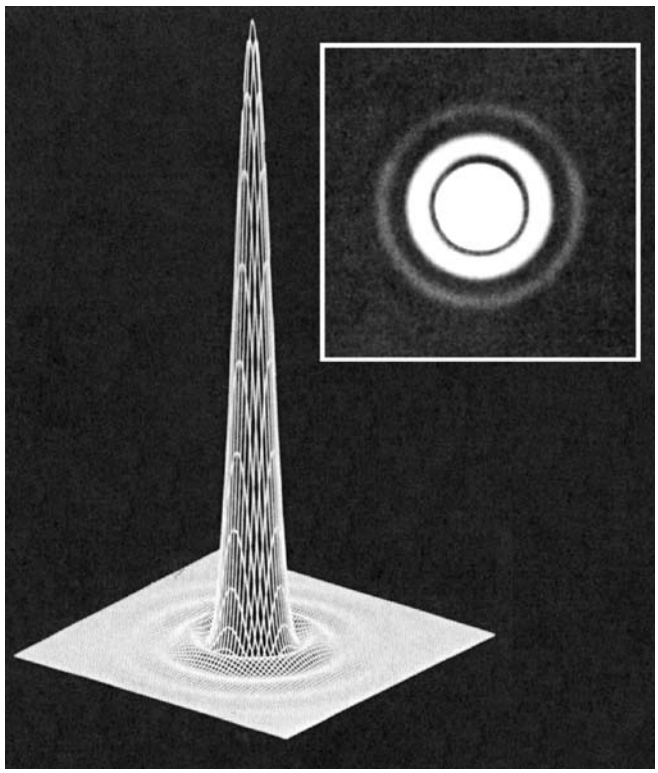


FIGURE 2.11. The visible-light intensity produced by a 0.5-mm-diameter circular aperture and the observed Airy rings (inset). The width of the central intense region is $1.22\lambda/D$.

Scattering from a circular hole

Now the real purpose of the exercise: without going into the detailed math we can replace the slit of width w, by a circular hole or aperture of diameter D. The resulting peak width in the plot of amplitude versus θ then has a maximum at $1.22\lambda/D$ as shown in Figure 2.11, which is a 3D representation of Figure 2.10 (but the third dimension is I, not I/I_0).

Because of the circular symmetry of the aperture, the calculation needed to obtain the number 1.22 involves the use of Bessel functions which you can find in texts on physical optics, a few of which we reference at the end of the chapter.

As the diameter of the aperture, D, decreases, the minimum resolvable spacing, r, increases (i.e., the resolving power gets worse). This expression for the Airy disk diameter also shows that as λ decreases, r decreases (so decreasing λ by increasing the accelerating voltage of the TEM will improve resolution).

Why is this relevant to TEM?

The important point about this analysis for TEM is that we'll see the same relationship in several later chapters. In those chapters, we will replace the slits by an aperture or we'll replace the hole by an atom or by your specimen. In other words, this analysis of diffraction from slits and holes is just geometry applied to optics—it's geometric optics.

2.10 CONSTRUCTIVE INTERFERENCE

To expand on this point, consider an infinite plane wave described by the usual characteristics of amplitude and phase. We can describe the wave function ψ by the standard expression

$$\psi = \psi_0 \exp[i\phi] \tag{2.11}$$

where ψ_0 is the amplitude and ϕ the phase of the wave. The phase depends on position x, such that if the path length changes by one wavelength λ, the phase difference is 2π. Stated another way, the phase difference $\Delta\phi$ between any two monochromatic (same wavelength) waves is related to the path difference Δx they must travel in going from the source to the detector. The relationship is

$$\Delta\phi = \frac{2\pi}{\lambda}\Delta x \tag{2.12}$$

This phenomenon of constructive interference is precisely what we discussed in Figure 2.6. Constructive interference between waves relies on the fact that the amplitudes of the waves add when you take account of the phase. If all waves scattered by all of the atoms in the specimen are to interfere constructively, they must all differ in phase by integral multiples of 2π. Clearly this condition requires that the path differences

traveled by all of the waves be integer multiples of the wavelength of the incident wave. We can ensure this by requiring that the scattering centers be periodically spaced; fortunately this can happen for all crystals. So the mathematical description of constructive interference is simplified (as we'll see in Part 2 of this text). The point here is that this analysis was carried out for X-rays and was not changed for electrons since it does not depend on the scattering mechanism, only on the geometry.

2.11 A WORD ABOUT ANGLES

Since angles (remember we mean semi-angles) are so important in the TEM (you can control some of them and the specimen controls others) we want to try to be consistent in our terminology.

- We can control the angle of incidence of electrons on the specimen and we will define the angle of incidence as α, as shown in Figure 2.12.
- In the TEM we use apertures or detectors to collect a certain fraction of the scattered electrons and we will define any angle of collection as β.
- We will define all scattering angles controlled by the specimen as θ. This may be a specific angle, such as twice the Bragg angle (where $\theta = 2\theta_B$) (see Section 11.4) or a general scattering angle θ. So θ is the scattering semi-angle for diffraction even though it is $2\theta_B$!

In fact the only angle of interest in TEM which is not given as a semi-angle is the solid angle of collection of X-rays by the XEDS detector (see Chapter 32) which is such a miserably small fraction of the total solid angle of X-ray generation (4π sr) that it is traditionally given in terms of the full collection angle!

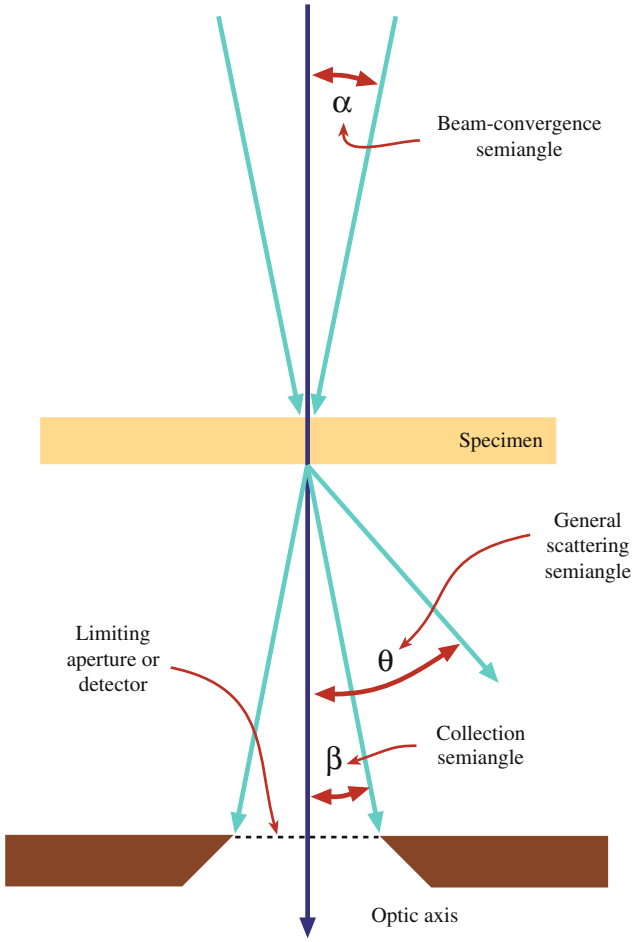


FIGURE 2.12. Definition of the major angles (i.e., semi-angles) in TEM. Any incidence/convergence angle of the beam is α; any collection angle is β and general scattering angles are θ. All the angles are measured from the optic axis, an imaginary line along the length of the TEM column.

2.12 ELECTRON-DIFFRACTION PATTERNS

We've mentioned a couple of times that the TEM is uniquely suited to take advantage of electron scattering because it can form a picture (DP) of the distribution of scattered electrons, which we'll discuss in Part 2 in much more detail. To understand fully how a DP is formed in the TEM, you need to go to Chapter 6 to see how electron lenses work and then to Chapter 9 to find how we combine several lenses to create the TEM imaging system. But before we take you through these concepts it is worth just showing a few of the many kinds of DPs that can be formed in the TEM. At this stage, all you have to do is imagine that a photographic film is placed directly after the thin specimen and that electrons scattered by the specimen as in Figure 2.1B impinge directly on the film. Under these circumstances, the greater the angle of scatter, the further off center the electron hits the film.

ON THE 'FILM'
Thus in a DP, distances on the film correspond to angles of scatter at the specimen.

Even using this simple description, however, you can comprehend some of the basic features of DPs. Figure 2.13 is a montage of several kinds of DPs, all of which are routinely obtainable in a TEM. You can see that several points we've already made about scattering are intuitively obvious in the patterns. First, most of the intensity is in the direct beam, in the center of the pattern, which means that most electrons appear to travel straight through the specimen. Second, the scattered intensity decreases with increasing θ (increasing distance from the direct beam), which reflects the

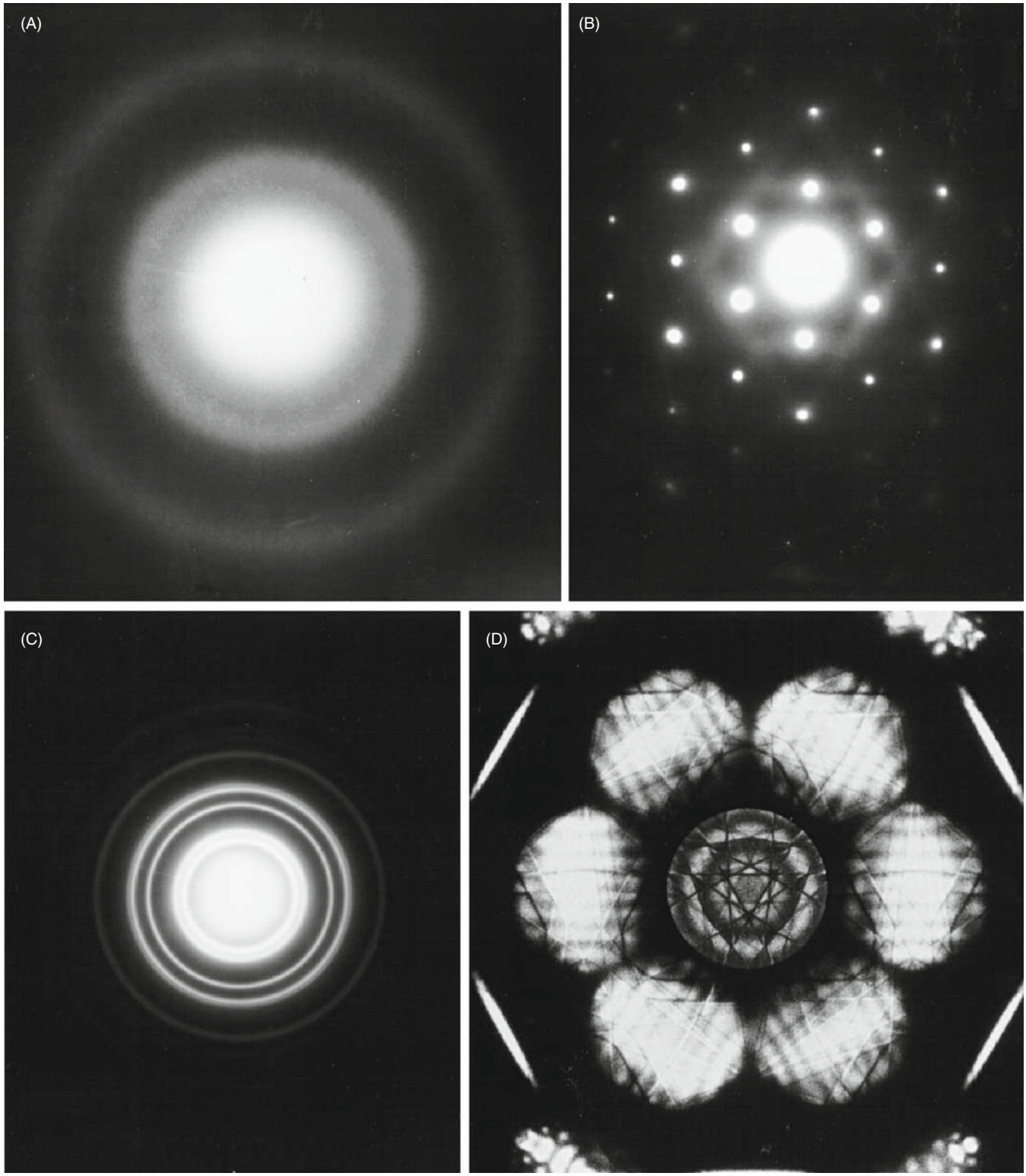


FIGURE 2.13. Several kinds of DPs obtained from a range of materials in a conventional 100-kV TEM: (A) amorphous carbon, (B) an Al single crystal, (C) polycrystalline Au, (D) Si illuminated with a convergent beam of electrons. In all cases the direct beam of electrons is responsible for the bright intensity at the center of the pattern and the scattered beams account for the spots or rings that appear around the direct beam.

decrease in the scattering cross section with increasing θ. Third, the scattering intensity varies strongly with the structure of the specimen. You'll see much more of this in Part II.

ANGLE OF SCATTER AND DISTANCES IN DPs
This relationship is different to the usual interpretation of images in which distances correspond to distances in the specimen, but it is critical to our understanding of diffraction patterns.

So far, we've only considered the amplitude/intensity of the electron wave and neglected the phase. When a wave is scattered, it will change its phase with respect to the incident wave. This is because a wave cannot change direction and remain in step with a wave that is not scattered. The phase of the scattered wave is most important in the specific topic of phase-contrast images, which have until recently been the principal form of high-resolution, atomic-level images such as shown back in Figure 1.2. We'll also come across the importance of the phase of the scattered wave when we consider the intensity of diffracted electron beams and the intensity in diffraction-contrast images. But at this stage all you need to know is that the electrons in the beam are in phase when they hit the specimen and the process of scattering, in any form, results in a loss of phase between the scattered and direct beams.

CHAPTER SUMMARY

Remember that electrons are strongly scattered because they are charged particles. This is the big difference compared to X-rays. Electrons are scattered by the electron cloud and by the nucleus of an atom. Remember X-rays are only scattered by the electron cloud. (In case you are physics oriented, a quantum-mechanical calculation does give the same distribution as the classic calculation for the Coulomb force.)

We have defined four important parameters in this chapter:

σ_{atom}	the scattering cross section of one atom
σ_{total}	the number of scattering events per unit distance traveled in the specimen
$d\sigma/d\Omega$	the differential scattering cross section of one atom
λ	the mean free path of (average distance traveled by) an electron between scattering events

Finally, a note on grammar! Should we discuss electron *scatter* or electron *scattering*? Electrons are scattered and we observe the results of this scattering (a gerund) but in fact we see the scatter (noun) of the electrons, which can be measured. However, if you've been observant you'll have noticed that we have always used *scattering* to denote the effect. Our practice is also consistent with the popular usage, which goes back to the early work of Bragg and others.

SCATTERING AND CROSS SECTIONS

Born, M and Wolf, E 1999 *Principle of Optics* 7th (yes, 7th!) Ed. Cambridge University Press New York. Perhaps *the* optics textbook in terms of classical treatments and number of editions.

Heidenreich, RD 1964 *Fundamentals of Transmission Electron Microscope* Interscience Publisher New York NY.

Jones 1992 gives a succinct introduction to scattering and Newbury (1986) gives a clear exposition on the units of cross sections. If you want to see a fuller description, read Wang (1995). If you're a glutton for punishment, the classic text is by Mott and Massey (1965) as we've already mentioned. You should realize that we've introduced you to some of the giants of electron optics, e.g., Airy, Fresnel, and Fraunhofer, who never knew about electron waves.

Jones, IP 1992 *Chemical Microanalysis Using Electron Beams* The Institute of Materials London.

Mott, NF and Massey, HSW 1965 *The Theory of Atomic Collisions* Oxford University Press Oxford.

Newbury, DE 1986 in *Principles of Analytical Electron Microscopy* p 1 Eds. DC Joy, AD Romig Jr and JI Goldstein Plenum Press New York.

Wang, ZL 1995 *Elastic and Inelastic Scattering in Electron Diffraction and Imaging* Plenum Press New York. An in-depth treatment of scattering using a much more rigorous mathematical approach than in this chapter.

OPTICS

We should have references to some of the founders of optics here, especially Abbe, Airy, Fraunhofer, and Fresnel, but we'll leave you to chase those up in the optics texts.

Fishbane, PM, Gasiorowicz, S and Thornton, ST 2004 *Physics for Scientists and Engineers* 3rd Ed. Prentice Hall Englewood Cliffs NJ.

Goodman, JW 2004 *Introduction to Fourier Optics* 3rd Ed. Roberts & Company Greenwood Village CO. An excellent source for the advanced student.

Hecht, E 2003 *Optics* 4th Ed. Addison-Wesley Reading MA. A favorite.

Klein, MV and Furtak, TE 1985 *Optics* 2nd Ed. Wiley & Sons New York NY. Not for the faint-hearted.

Smith, FG and Thomson, JH 1988 *Optics* 2nd Ed. Wiley & Sons New York.

Taylor, C 1987 *Diffraction* Adam Hilger Bristol UK.

SOME MICROANALYSIS AND MORE

Goldstein, JI, Newbury, DE, Joy, DC, Lyman, CE, Echlin, P, Lifshin, E, Sawyer, LC and Michael, JR 2003 *Scanning Electron Microscopy and X-ray Microanalysis* 3rd Ed. Kluwer New York.

Joy, DC 1995 *Monte Carlo Modeling for Electron Microscopy and Microanalysis* Oxford University Press New York.

Rhodes, R 1986 *The Making of the Atomic Bomb* Simon and Schuster New York. See p 282.

Rhodes, R 1995 *Dark Sun: The Making of the Hydrogen Bomb* Simon and Schuster New York. See p 423.

These books are well worth reading because of both the historical and the scientific content.

URLs

1) http://www.cstl.nist.gov/div837/837.02/epq/index.html

SELF-ASSESSMENT QUESTIONS

Q2.1 What is a cross section and in what units is it measured?
Q2.2 Distinguish between total, atomic, and differential cross sections.
Q2.3 Why are we interested in variations in the scattering intensity and the angular distribution of electron scattering?
Q2.4 What is the mean free path of an electron?
Q2.5 What do we mean by the term *electron beams* and why do we ask this question?
Q2.6 How is the direct beam different from or similar to the scattered beams?
Q2.7 Distinguish scatter and scattering.
Q2.8 What's the difference between forward scattering and backscattering?
Q2.9 What distinguishes elastic and inelastic scatterings?
Q2.10 Distinguish between coherent and incoherent scattering.
Q2.11 Describe what distinguishes diffraction from other kinds of scattering.
Q2.12 Distinguish between Fraunhofer and Fresnel diffractions.
Q2.13 Distinguish the angles α, β, θ, and Ω.
Q2.14 List the different ways a specimen can scatter electrons.
Q2.15 How many different ways can you control the scattering processes in the TEM?
Q2.16 How can you select electrons that have suffered a specific kind of scattering?
Q2.17 What's the fundamental difference between electron scattering and X-ray scattering?
Q2.18 What is a phasor diagram?
Q2.19 Why would you want to draw a phasor diagram in TEM?
Q2.20 How small is a small angle in the TEM and why are scattering angles in the TEM usually this small?

TEXT-SPECIFIC QUESTIONS

T2.1 Write down concise definitions of coherent, incoherent, elastic, and inelastic as we use them and link these definitions to the information in Figure 2.2.
T2.2 Explain in a paragraph the relationship between scattering cross section and atomic scattering factor, mentioning the important factors that influence them.
T2.3 Explain the link between the information in Figures 1.3 and 2.1.
T2.4 Distinguish the scattering angles θ and Ω in Figure 2.3 and the information that can be gathered within them. Relate these angles to the relevant angles in a TEM described in Figure 2.12.
T2.5 Sketch the intensity projected onto a photographic plate or viewing screen from the scattering produced by the Cu and Au specimens in Figure 2.4. The result does not look like the intensity in either a typical TEM image or DP shown in many figures throughout the book. Explain why this is so.

T2.6 Why are the backscattered electrons so few in number in Figure 2.4A and B and why do they all scatter to the one side in Figure 2.4A?

T2.7 Draw equivalent diagrams to Figure 2.5 for (a) 2 slits $d/2$ apart; (b) 2 slits $2d$ apart; (c) 5 slits d apart. What does this tell you about the effect on the scattering distribution of both the number and the spacing of the scattering centers?

T2.8 Draw a phasor diagram like Figure 2.6, but for three slits only.

T2.9 What is the relationship between Figure 2.10 and Figure 2.11?

T2.10 Make a copy of Figure 2.13. Cut out two circular holes with diameters ~ 5 and ~ 40 mm in another sheet of paper corresponding to different collection angles (β) in Figure 2.12. Superimpose the smaller circular hole on the different patterns in different positions to simulate the selection of electrons for forming images in a TEM. Note how easy it is to select electrons scattered in specific directions, but also note how many electrons are excluded when you do this. (a) What does this tell you about the advantages and disadvantages of a small selection aperture (or small detector)? Now superimpose the larger hole and note how many more electrons can be selected. (b) What does this tell you about the advantages and disadvantages of a large selection aperture (or large detector)?

3

Elastic Scattering

CHAPTER PREVIEW

Elastically scattered electrons are the major source of contrast in TEM images. They also create much of the intensity in DPs, so we need to understand what controls this process. First we'll consider elastic scattering from single, isolated atoms and then from many atoms together in the specimen. To comprehend elastic scattering we need to invoke both particle and wave characteristics of electrons.

Scattering from isolated atoms can occur either as a result of electrons interacting with the negatively charged electron cloud, which results in angular deviations of only a few degrees, or by attraction toward the positive nucleus which scatters the electrons through much larger angles, up to 180°. Such scattering can often be interpreted in terms of billiard-ball type, particle-particle collisions, cross sections, and mean free paths that we introduced in the previous chapter. We'll introduce the Rutherford differential cross section, which explains the strong dependence of high-angle elastic scattering on the atomic number (Z) of the atom. Later in the book, we'll show how to use this Z dependence to form images that reflect the chemistry of the specimen. When we treat the electrons as waves, their *coherency* becomes important. The coherency of the scattered electrons is related to their *angle* of scattering (θ). As this angle becomes larger, the degree of coherency becomes less and electrons that are Rutherford-scattered out to high angles are incoherent.

In contrast to Rutherford scattering, electrons that are scattered elastically through small angles (which we'll define as $<3°$) are coherent. The intensity of this low-angle scattering is strongly affected by the arrangement of atoms within the specimen. As we introduced in the previous chapter, such collective scattering by the atoms is referred to as *diffraction* and can only be understood if we treat the electron as a wave. Diffraction is controlled mainly by the angle of incidence of the electron beam to the atomic planes in the specimen, the spacing between these planes, and interatomic distances within the planes. So this small-angle, coherent scattering is invaluable for characterizing the crystallography of the specimen and is undoubtedly *the* most significant scattering phenomenon in the TEM.

So as we discuss elastic scattering, you'll see that we often use the wave-particle duality simultaneously, because both lines of thought are necessary for a full understanding.

3.1 PARTICLES AND WAVES

We have two different ways of looking at how an electron beam interacts with our TEM specimens. We can consider the beam as a succession of particles or as a number of waves. What we want to do is understand the relationship between the two approaches and we can summarize the two viewpoints thus:

Electrons are *particles* so they have the following properties, which we introduced in Chapter 2.

- They have a scattering cross section and a differential scattering cross section.
- They can be scattered through particular angles (remember our angles are semi-angles).
- The electrons interact with the nucleus and the electron cloud through Coulomb forces.
- We can relate this process to scattering of other particles such as α particles, so lots of analysis can carry over from other systems.

PARTICLES AND WAVES

When we discuss X-ray and electron spectrometry you'll see that we have to use a particle description. When we discuss imaging, HRTEM, and DPs you'll see that we use a wave description.

Electrons have a *wave* nature and the electron beam is almost a *plane wave*, hence:

- Waves are diffracted by atoms or 'scattering centers.'
- How strongly a wave is scattered by an atom is determined by the atomic-scattering amplitude.
- When we gather atoms together into a solid, the diffraction process gets much more complicated but it is central to TEM.
- We can relate the process to the diffraction of X-rays, so lots of analysis already exists.

3.2 MECHANISMS OF ELASTIC SCATTERING

In the previous chapter we simply stated that electrons going through a thin specimen are either scattered or not scattered and either lose energy or don't lose energy. It's now time to describe the ways in which this scattering occurs and in this chapter we'll confine our attention to elastic events, saving inelastic scattering for Chapter 4.

It's convenient to divide elastic-scattering mechanisms into two principal forms: electron scattering from single, isolated atoms and collective scattering from many atoms together within the specimen. We'll start in the same way we did in the previous chapter by looking first at the interaction of a single electron with an isolated atom. In this situation, elastic scattering can occur in one of two ways, both of which involve Coulomb forces. As shown in Figure 3.1, the electron may interact with the electron cloud, resulting in a small angular deviation. Alternatively, if an electron penetrates the electron cloud and approaches the nucleus, it will be strongly attracted and may be scattered through a larger angle that, in rare cases in the TEM, can approach 180° (i.e., complete backscattering).

ELASTIC?

You should be aware that either of these two interactions may not be truly elastic, so our separation of scattering into elastic and inelastic is a bit of a simplification.

In fact many electron-electron interactions are inelastic, as we'll see in the next chapter. We'll also see, for example, that the nuclear interaction may result in the generation of a bremsstrahlung X-ray or may even result in the displacement of the atom from its site in the crystal, both of which involve some energy loss for the electron. Indeed, the higher the angle of scattering of an electron emerging from the specimen, the greater the chance that it will have undergone an inelastic event at some time during its passage through the specimen. Despite all this, we'll continue to ignore any inelastic effects in this chapter.

The second principal form of elastic scattering occurs when the electron wave interacts with the specimen as a whole. We've already mentioned the best-known form of this interaction, namely, diffraction, which is particularly important at low-scattering angles. Understanding diffraction involves treating the electron beam as a wave, rather than as a particle as we did in Figure 3.1. Following the original approach of Huygens for the diffraction of visible light, we imagine each atom in the specimen that interacts with the incident plane wave acts as a source of secondary spherical wavelets, as illustrated in Figure 3.2.

INTERFERENCE

These wavelets reinforce one another in certain angular directions and cancel in other directions: both reinforcement and cancellation are extremely useful phenomena.

Thus, the low-angle, elastic scattering distribution is modified by the crystal structure of the specimen, and strong diffracted beams emerge at specific angles. The

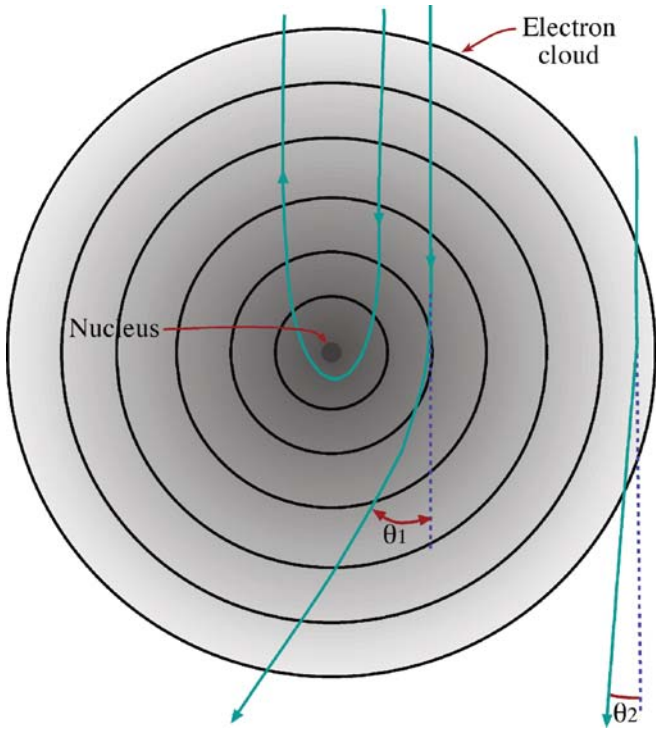


FIGURE 3.1. An isolated atom can scatter a high-energy electron by two mechanisms. Coulombic interaction within the electron cloud results in low-angle scattering; Coulombic attraction by the nucleus causes higher-angle scattering (and perhaps complete backscatter when $\theta > 90°$). The potential within the electron cloud is always positive.

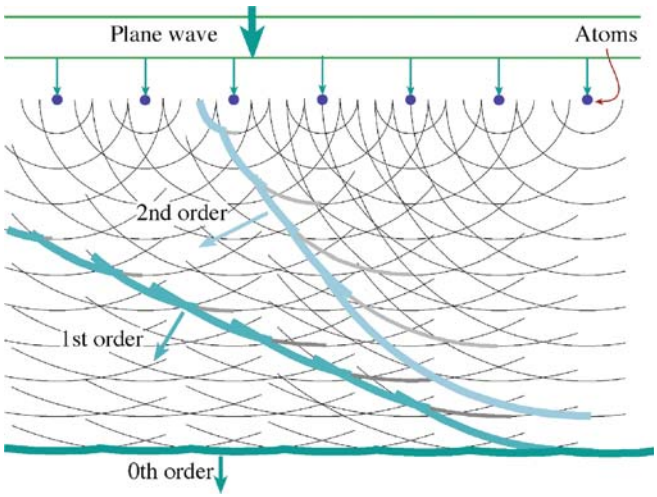


FIGURE 3.2. A plane, coherent electron wave generates secondary wavelets from a row of scattering centers (e.g., atoms in the specimen). The secondary wavelets interfere, resulting in a strong direct (zero-order) beam and several (higher order) coherent beams scattered (diffracted) at specific angles.

diffracted beam scattered through the smallest angle is called the first-order beam and we'll discuss these and higher-order effects in depth in Chapters 11 and 12. We'll now go on to examine these two forms of elastic scattering in more detail, starting with the billiard-ball approach. Then we will briefly describe the scattering of a wave to show how it relates to this particle-based treatment and later we'll use the wave approach as the basis for a full analysis of diffraction.

3.3 ELASTIC SCATTERING FROM ISOLATED ATOMS

So let's look at two possible paths for a beam electron interacting with an isolated atom as shown in Figure 3.1. Whether it interacts more strongly with the nucleus or the electrons, the electron is scattered through an angle θ.

SCATTERING ANGLE

Elastic electron-electron interactions usually result in a relatively low scattering angle, while electron-nucleus interactions cause higher-angle scattering.

If we just consider an electron, charge e, accelerated through a voltage V before being scattered from an isolated atom, the electron-electron and electron-nucleus scattering paths are hyperbolic and can be given by two simple equations (Hall 1953) which are useful because they summarize the principal factors that control elastic scattering

$$r_e = e/V\theta \tag{3.1}$$

$$r_n = Ze/V\theta \tag{3.2}$$

where r is the radius of the scattering field of the nucleus and the electron. The different cross sections for scattering through angles $> \theta$ are given by πr_n^2 for the nucleus and $Z\pi r_e^2$ for the scattering by Z electrons in the cloud. If we sum the two components and (just as back in equation 2.8) multiply by $N_0\rho t/A$ we'll get a sense of the total elastic scattering through a film of thickness t.

This approach is "defective in many respects" as Hall says but gives you a good qualitative sense of the various parameters that affect elastic scattering. You can see that the atomic number Z of the atom controls the elastic interaction with the nucleus, but the electron-electron scattering is more a function of the incident-beam energy V (which has to be in esu if the dimensions of these equations are to be correct). We'll see later in Chapter 22 that the strong effect of Z becomes important when we need to enhance scattering in low-Z materials, such as polymers and biological tissue, in order to get better TEM image contrast. Notice that when the electron passes close to the nucleus (r_n is small) the angle θ will be large. We'll see in Chapter 22 that this dependence on θ directly relates to TEM-image contrast. The electron-beam energy can also control the image contrast to some extent. So Z, V, and θ all affect image contrast and are the three major reasons why you cannot avoid having to study the physics of electron scattering.

3.4 THE RUTHERFORD CROSS SECTION

For the next three sections, we'll ignore the low-angle, electron-electron scattering and concentrate only on scattering by the nucleus. The high-angle, electron-nucleus interaction is analogous to the backscattering of α particles from a thin metal foil. The first observation of such backscattering in 1911 by H. Geiger (of *counter* fame) and a Manchester University *undergraduate*, E. Marsden, enabled their professor, Ernest Rutherford, to deduce the existence of the nucleus (never overlook undergraduate research results!). Rutherford (1911) described backscattering as "the most incredible event that has ever happened to me" (even though he'd already won the Nobel Prize for Chemistry in 1908), and he derived the following expression for the differential cross section for this high-angle scattering by the nucleus alone

$$\sigma_R(\theta) = \frac{e^4 Z^2}{16(4\pi\varepsilon_0 E_0)^2} \frac{d\Omega}{\sin^4 \frac{\theta}{2}} \tag{3.3}$$

All the terms in this equation were defined back in Chapter 2. The expression ignores relativistic effects and assumes that the incident electron does not lose

significant energy through inelastic processes, so that the energy of the electrons, E_0 (in keV), is fixed. As we've already noted, although strictly inaccurate, these assumptions can be tolerated in the TEM (at least at 100 keV or below).

3.5 MODIFICATIONS TO THE RUTHERFORD CROSS SECTION

You'll often see the Rutherford differential cross section in different, but mathematically similar, forms. For example, equation 3.3 neglects the so-called screening effect of the electron cloud. Screening can be thought of as making the nucleus appear somewhat less positive to the incident electron (although the overall charge within the electron cloud is *always* positive). So the differential cross section is effectively reduced and the amount of scattering is lowered. Screening is only important when the beam electron passes far from the nucleus and under these circumstances the scattering angle will be small (say $<\sim 3^\circ$). If we wish to account for screening, we replace the $\sin^2(\theta/2)$ term with $[\sin^2(\theta/2) + (\theta_0/2)^2]$ where θ_0 is called the screening parameter given by

$$\theta_0 = \frac{0.117Z^{1/3}}{E_0{}^{1/2}} \tag{3.4}$$

(Here E_0 is in keV.) What we are saying is that the screening parameter can be described by a particular scattering angle, θ_0. When the scattering angle is greater than θ_0 we can neglect electron-electron interactions and the electron-nucleus interaction is dominant. The value of θ_0 at 100 keV is only $\sim 2^\circ$ for Cu and less for lighter elements, so above a few degrees, all scattering can be approximated to Rutherford high-angle scattering.

As we've noted, so far all our equations are non-relativistic, which is unfortunate since relativistic effects are significant for electrons with energies $>\sim 100$ keV (which is the case for most materials investigations in the TEM). Fortunately, we can easily correct for relativity to give a more accurate cross section by using λ_R, the relativistically corrected electron wavelength (see equation 1.7), and a_0, the Bohr radius of the scattering atom

$$a_0 = \frac{h^2\varepsilon_0}{\pi m_0 e^2} \tag{3.5}$$

where ε_0 is the dielectric constant. Using the other constants listed in Table 1.1 we find a_0 is 0.0529 nm (if you're old enough you can easily remember this as 0.5 Å). The net result of adding screening and relativity corrections is that

$$\sigma_R(\theta) = \frac{Z^2\lambda_R^4}{64\pi^4 a_0^2}\frac{d\Omega}{\left[\sin^2\left(\frac{\theta}{2}\right) + \frac{\theta_0^2}{4}\right]^2} \tag{3.6}$$

This expression describes the screened, relativistic, differential Rutherford cross section. One very important effect of incorporating screening into these equations is that the cross section does not go to infinity as the scattering angle goes to zero which is an important limitation of all the simpler equations that we used initially.

The screened Rutherford cross section is the one most widely used for TEM calculations, although it has particular limitations at the highest operating voltages (300–400 kV) and for heavier elements ($Z>30$) which scatter electrons through large angles. Under these circumstances, you should use another cross section, such as that of Mott, for which you should consult the text by Mott and Massey (which we referred to in Chapter 2) or Newbury (1986).

So, as we did for the basic cross section back in Chapter 2, we can integrate this expression to obtain the total cross section over specific angular ranges. We can substitute appropriate values for the various constants and integrate the differential cross section from θ to π to obtain the total nuclear cross section (in scattering events/electron/atom/m^2) for electrons elastically scattered into angles $> \theta$

$$\sigma_{\text{nucleus}} = 1.62 \times 10^{-24}\left(\frac{Z}{E_0}\right)^2 \cot^2\frac{\theta}{2} \tag{3.7}$$

(From what we told you in Chapter 2 you ought to be able to work out the integration necessary to determine the probability of electrons being scattered into angles $<\theta$.) Again we see that the beam energy (E_0), the angle of scattering (θ), and the atomic number (Z) all affect the probability that an electron will be scattered by the nuclei of atoms in the specimen. If you simplify this last expression by assuming that θ is small you should be able to see some parallels with Hall's less accurate equation 3.2 for nuclear scattering. However, there's much more to this whole scattering process than we have covered here and you should read Newbury (1986) and Jones (1992) for further discussion of these calculations once you really appreciate their significance.

The best way to summarize the characteristics of cross sections is to present some data. Figure 3.3 shows the variation of the screened Rutherford cross section in equation 3.7 with scattering angle for (a) three different elements and (b) two different beam energies. As you can see for Cu, the cross section decreases by several orders of magnitude from $\sim 10^{-22}$ to $\sim 10^{-28}$ m^2 as the scattering angle increases from 0 to 180°; so, as we've already told you, scattering is most likely to occur in the forward ($\theta \sim 0^\circ$) direction and drops off rapidly

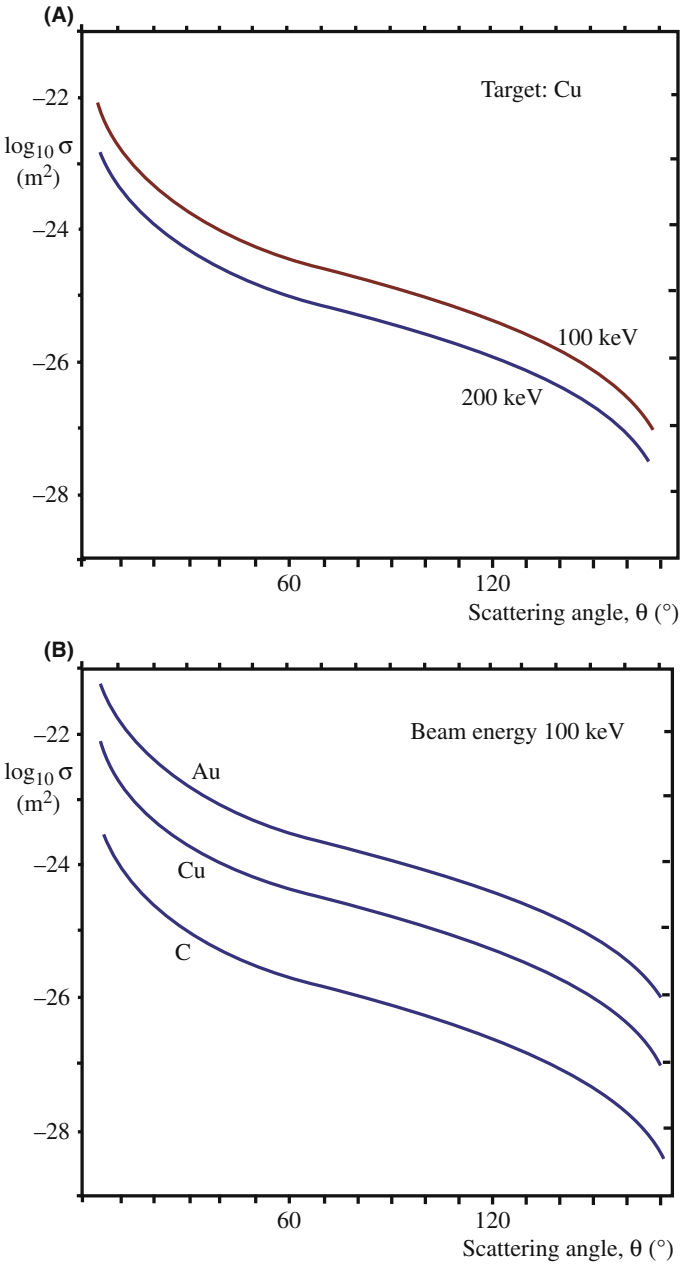


FIGURE 3.3. The variation of the logarithm of the screened relativistic Rutherford cross section with scattering angle from equation 3.7, describing the change in cross section for electrons scattered at angles $> \theta$ (A) for different elements at 100 keV and (B) for scattering from Cu at different accelerating voltages.

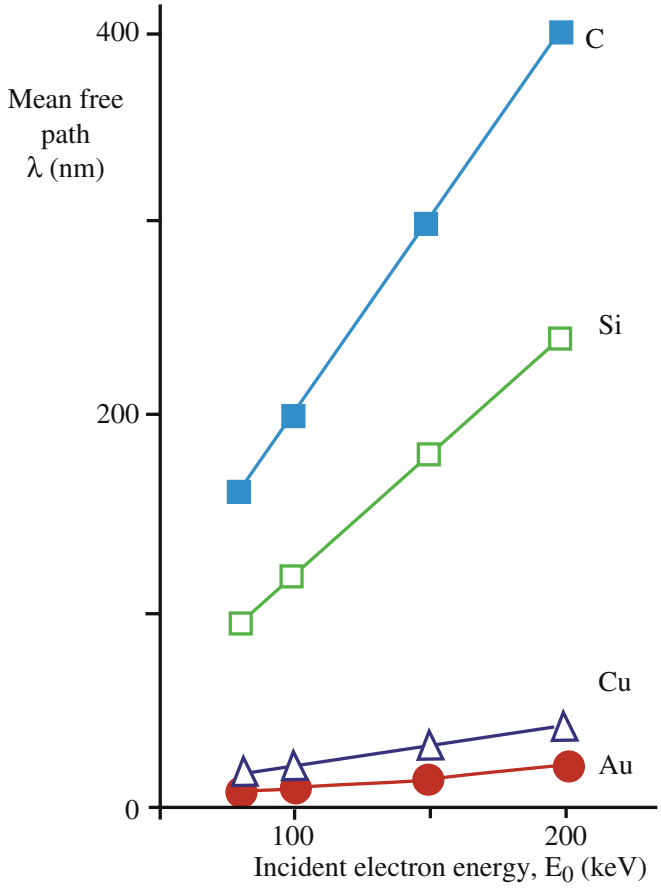


FIGURE 3.4. The variation of the mean free paths of elastic scattering for four different elements as a function of the beam energy, calculated assuming a screened, relativistic Rutherford cross section.

above a few degrees. Increasing Z from carbon to gold can increase the cross section by a factor of ~100, which is why you need proportionately thinner TEM specimens if you want to 'see' through higher-Z materials. Doubling the electron-beam energy can lower the cross section by a factor of two or three, which is why higher-energy electrons are less likely to be scattered by your specimen than lower-energy electrons, all else being equal. Figure 3.4 plots the related mean free paths for elastic scattering. From this graph you can see that very few high-angle elastic scattering events will occur if you can make your specimen < 100 nm in thickness. Within such specimens, most electrons either undergo a single-scattering event or are not scattered and we'll assume that this simplification is a viable approximation to what's actually going on in the microscope many times throughout this text. This approximation is the main reason why, as we've already noted that in almost all TEM studies, the 'thinner is better' criterion applies.

3.6 COHERENCY OF THE RUTHERFORD-SCATTERED ELECTRONS

Up to now, in this chapter, we've treated the electron as a particle, but there is useful insight to be gained if we examine the wave nature of the scattered electron. High-angle Rutherford-scattered electrons are *incoherent*: i.e., there is no phase relationship between them. This is a tricky concept because we are scattering particles. Such incoherent scattering is important in two respects. First, the high-angle, forward scattering can be used to form exceptionally high-resolution

images of a crystalline specimen in which the image contrast is due solely to the value of Z, not the orientation of the specimen (as is the case for low-angle coherent diffraction). Such *Z-contrast images,* as we'll see in Chapter 22, provide qualitative atomic-resolution elemental analysis, in addition to showing atomic-resolution detail at interfaces between regions of different Z. Compared to other image-contrast mechanisms, Z-contrast imaging is a relatively new technique for most microscopists but, particularly since the availability of C_s-correctors, it has consistently held the record for the highest-resolution images and analysis in the TEM (e.g., Varela et al. 2005) and is already beginning to revolutionize our understanding of the atomic-level structure and chemistry of crystal defects.

COHERENCY

Coherency of the scattered electron is a wave property. If the scattered electron waves have a phase relationship they must be coherent.

Second (but much less important), the high-angle backscattered electrons (BSEs) can be used to form images of the beam-entrance surface of the specimen, in which the contrast is not only due to differences in Z, but also to changes in surface topography of the specimen. BSE images are rarely used in the TEM because the BSE signal is small. If you go back and look at the Monte Carlo simulation in Figure 2.4 you'll see that out of 10^3 incident electrons in Cu only about three (0.3%) were backscattered. Therefore, the quality of this signal is very poor, the images are noisy, and the contrast is low. The contrast is much better for bulk specimens in an SEM in which many more electrons are backscattered (e.g., about 30% in Cu) and BSEs provide a stable, high-contrast imaging technique in SEMs in which you can discriminate between the signals from adjacent elements in the periodic table. In principle, Z-contrast should be able to do the same in the TEM.

3.7 THE ATOMIC-SCATTERING FACTOR

The classic Rutherford differential cross section cannot be used to calculate the cross section exactly, because it ignores the wave nature of the electron beam. A full treatment involves wave mechanics and is well beyond the scope of this text. Perhaps the most familiar aspect of the wave approach to cross sections is the concept of the atomic-scattering factor $f(\theta)$, which is related to the differential elastic cross section by a simple equation (more on this in Section 3.8)

$$|f(\theta)|^2 = \frac{d\sigma(\theta)}{d\Omega} \tag{3.8}$$

What we will now do is to highlight some of the important features that lead to this result by outlining the basic arguments.

- $f(\theta)$ is a measure of the amplitude of an electron wave scattered from an isolated atom.
- $|f(\theta)|^2$ is proportional to the scattered intensity.

From these two statements and given the importance of scattered-electron intensity in images and DPs, you can appreciate why $f(\theta)$ is such an important parameter in TEM.

The scattering-factor approach is complementary to the Rutherford differential cross section analysis, because it is most useful for describing the low-angle (i.e., $<\sim 3^\circ$) elastic scattering where the Rutherford model is inappropriate. Usually, $f(\theta)$ is defined in the following manner

$$f(\theta) = \frac{\left(1 + \frac{E_0}{m_0 c^2}\right)}{8\pi^2 a_0} \left(\frac{\lambda}{\sin\frac{\theta}{2}}\right)^2 (Z - f_x) \tag{3.9}$$

All the terms have been previously defined (note that we've dropped the screening term, so remember what this implies). If you need a more detailed approach you could consult the physics-based text by Reimer. Because we're now thinking in terms of waves, we need the wavelength λ (controlled of course by the beam energy E_0), and f_x is the scattering factor for X-rays, which is well known. The most widely referenced source of electron-scattering factors for TEM was usually the classic work of Doyle and Turner (1968), but you can now find values in software packages (see Section 1.6) and you can even do your own calculations using the free software on the NIST database (also in Section 1.6 and URL #1). The appearance of f_x in equation 3.9 is a reminder that $f(\theta)$ is a fundamental result of the wave nature of the electron.

***f*(θ)**

The atomic-scattering factor $f(\theta)$ depends on λ, θ, and Z.

We can plot this angular variation for a single isolated atom. Figure 3.5 summarizes graphically what we already know about the magnitude of elastic scattering (see equations 3.1 and 3.2)

- It decreases as θ increases ($\theta = 0^\circ$ for the incident-beam direction).
- It decreases as λ decreases (i.e., as the accelerating voltage (V) increases).
- It increases with Z for any value of θ.

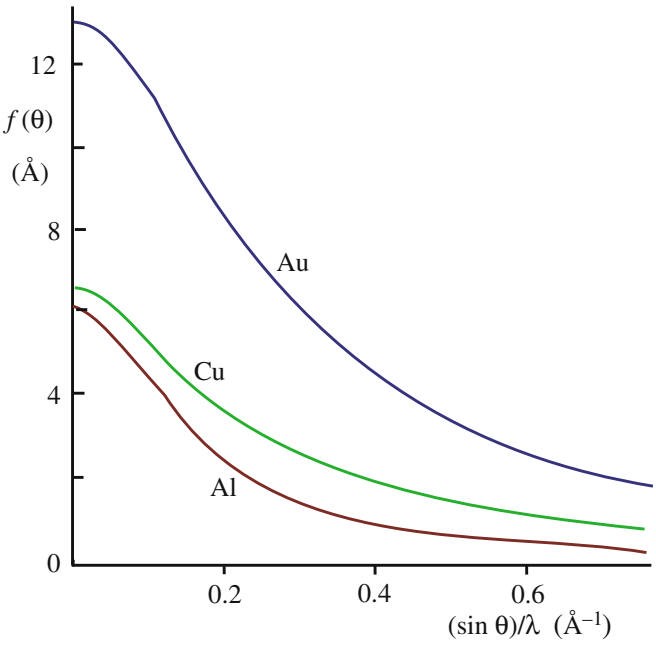


FIGURE 3.5. Change in the atomic scattering factor $f(\theta)$ with scattering angle θ (calculated from equation 3.9) showing that elastic scattering decreases with angle away from the incident beam direction ($\theta = 0°$) and increases with Z.

This expression (equation 3.9) for $f(\theta)$ contains components of both elastic nuclear scattering (the Z term) and elastic electron-cloud scattering (the f_x term). We'll see later in the chapters on diffraction in Part 2 that the $f(\theta)$ approach is used exclusively and, if we neglect the f_x term, then it can be shown that $|f(\theta)|^2$ is mathematically equivalent to the high-angle Rutherford differential cross section, as we defined it in equation 3.6. So now we've tied together the particle and wave approaches to elastic scattering.

ANGLE VARIATION

The important point to remember is that both the differential cross section and the scattering factor are simply measures of how the electron-scattering intensity varies with θ.

3.8 THE ORIGIN OF $f(\theta)$

Since $f(\theta)$ relates to the amplitude of a scattered wave, we'll consider briefly how it arises. The following analysis is not intended to be completely rigorous, but only to give the fundamental ideas behind the meaning of $f(\theta)$ and its relation to the differential scattering cross section. You can safely delay studying this topic until curiosity wins, then you can go and read the really thick physics textbooks.

To find the total elastic-scattering cross section, we have to integrate $d\sigma/d\Omega$. Note that this is a particle model, but you should also be aware of how the wave nature of the electrons is brought in. We can consider the wave nature by looking at Figure 3.6 (which you should realize is closely related to Figures 2.3 and 2.12).

We can describe the incident beam as a wave of amplitude ψ_0 and phase $2\pi kr$

$$\psi = \psi_0 e^{2\pi i k r} \tag{3.10}$$

In this definition of phase, k is the magnitude of the wave vector and r is the distance that the wave has propagated, as we'll discuss in detail later in Chapter 11. When the incident plane wave is scattered by a point charge, a spherical scattered wave is created which has a different amplitude ψ_{sc} but keeps the same phase apart from a $\pi/2$ addition which we return to in a moment

$$\psi_{sc} = \psi_0 f(\theta) \frac{e^{2\pi i k r}}{r} \tag{3.11}$$

In this equation, $f(\theta)$ is the amplitude we would have if $\psi_0 = 1$, i.e., it is the *atomic-scattering amplitude*.

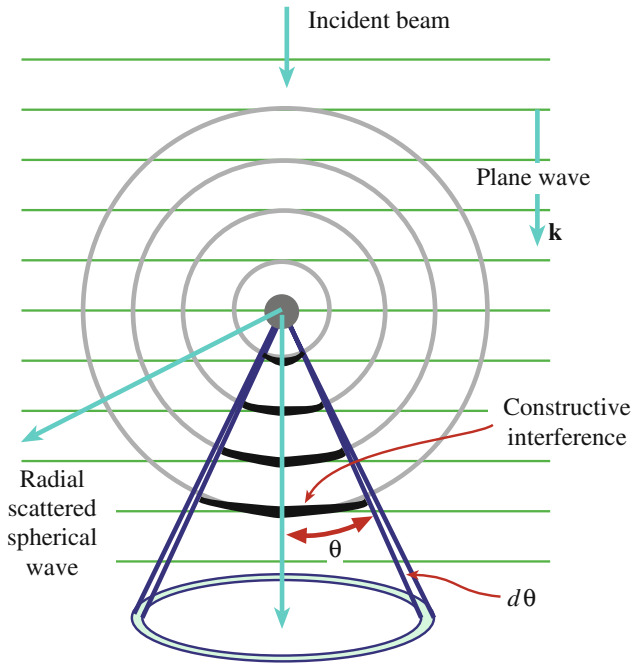


FIGURE 3.6. Generating a scattered wave by the interaction of a plane wave (horizontal line, wavelength λ) with a point charge. The circles represent the scattered spherical wavefronts which are in phase and retain the original λ. The in-phase, constructive interference between the plane and spherical waves is shown by the dark areas. The angles θ and $d\theta$ are the same as in Figure 2.3.

So obviously we need to know $f(\theta)$ and an acceptable model is essential to make the problem manageable. Up to this point, our treatment has been quite rigorous and ideally, the model would distinguish between a neutral atom in a metal, a covalently bonded atom, and an ion. If you're desperate, the quantity $f(\theta)$ can always, in principle, be calculated from the Schrödinger equation. In practice, however, we usually use a simple approximation which we'll now describe.

If we write down the expression for the scattering process shown in Figure 3.6, then we have

$$\psi_{sc} = \psi_0 \left[e^{2\pi i \mathbf{k}_I \cdot \mathbf{r}} + i f(\theta) \frac{e^{2\pi i k r}}{r} \right] \qquad (3.12)$$

You should note first of all that, as usual for Huygens wavelets, there is a 90° phase shift (shown by the inclusion of 'i' in the second term) between the incident and scattered beams and secondly, that $f(\theta)$ can be expressed as

$$f(\theta) = |f(\theta)| e^{i\eta(\theta)} = |f(\theta)|(\cos\eta(\theta) + i \sin\eta(\theta)) \qquad (3.13)$$

which means that the phase, $\eta(\theta)$, of $f(\theta)$ also depends on the angle of scatter, θ.

First aside: In writing equation 3.12, we have introduced two wave-propagation parameters: the vector $\mathbf{k}_I$ for the incident plane wave and the scalar k for the spherical scattered wavelet. By writing the 2π factor separately as part of the phase term, we have implicitly defined k to be $1/\lambda$. Many physics textbooks include the 2π in k so they have k given by $2\pi/\lambda$. Just be careful when you compare similar formulas in two textbooks.

1/λ AND 2π/λ

Sometimes $k = 1/\lambda$ and sometimes $k = 2\pi/\lambda$ and it's sometimes difficult to find out which definition is being used.

Second aside: The 90° phase change for the scattered-wave component in equation 3.13 can be easily understood if you consider the following. If the amplitude of the wave is initially $\psi_0 \sin(2\pi kr)$ then, after it has passed through the specimen, it will be ψ_{tot}. After scattering, the phase is increased by ϕ, so we can express the new ψ_{tot} as

$$\psi_{tot} = \psi_0 \sin(2\pi kz + \phi) = \psi_0 \sin(2\pi kz)\cos\phi + \psi_0 \cos(2\pi kz)\sin\phi \qquad (3.14)$$

Now if ϕ is small, then $\cos\phi \approx 1$ and $\sin\phi \approx \phi$; $\cos\theta$ is always the same as $\sin(\theta + \pi/2)$, hence

$$\psi_{tot} = \psi_0 \sin(2\pi kz) + \psi_0 \phi \sin\left(2\pi kz + \frac{\pi}{2}\right) \qquad (3.15)$$

The $\pi/2$ term would arise if we used the exponential rather than the sine to denote phase, so we can now write equation 3.15 as

$$\psi_{tot} = \psi + i\psi_{sc} \qquad (3.16)$$

This equation has the same form as that given in equation 3.12.

3.9 THE STRUCTURE FACTOR *F*(θ)

The next introductory step in discussing electron scattering is to take the idea of individual atoms scattering electrons (i.e., $f(\theta)$), which we've just discussed in some detail, and consider what happens when the atoms are stacked together regularly in a crystal structure. (We can, in principle, also do this for an amorphous solid but we'll stick to crystals for simplicity.) We will deal with this approach in great detail in Chapter 13, but for now we can introduce the structure factor $F(\theta)$, which is a measure of the amplitude scattered by a unit cell of a crystal structure. Because $F(\theta)$ is an amplitude like $f(\theta)$, it also has dimensions of length. We can define $F(\theta)$ as the sum of the $f(\theta)$ terms from all the i atoms in the unit cell (with atomic coordinates $x_i\, y_i\, z_i$) multiplied by a phase factor. The phase factor takes account of the difference in phase between waves scattered from atoms on different but parallel atomic planes with the same Miller indices (hkl). The scattering angle θ is the angle between the incident and scattered electron beams. So we can write

$$F(\theta) = \sum_i^{\infty} f_i e^{2\pi i(hx_i + ky_i + lz_i)} \qquad (3.17)$$

The amplitude (and hence its square, the intensity) of scattering is influenced by the *type* of atom ($f(\theta)$), the *position* of the atom in the cell (x,y,z), and the specific *atomic planes* (hkl) that make up the crystal structure. None of this is very surprising, but it turns out that this equation predicts that in certain circumstances the amplitude of scattering is zero. This behavior is intrinsic to the scattering process, is implicit back in Figure 3.2, and is often a very useful diagnostic test when determining crystal structures in the TEM.

ZERO SCATTERING

Under specific conditions, electrons scattering in a crystal may result in ZERO scattered intensity. Why might this occur?

We'll return to this point in Chapter 13 in much more detail.

3.10 SIMPLE DIFFRACTION CONCEPTS

As we mentioned earlier, electron diffraction is by far the most important scattering phenomenon in the TEM. The reason for this importance, as we'll show you in Chapters 11 and 12, is that we can use diffraction to determine the spacing of planes in crystals and, as you'll see later in Chapters 20, 21, and in the companion text, there is a whole field termed *electron crystallography* which gives an unprecedented amount of crystallographic information from space-group symmetry data right down to the dimensions of single unit cells. At the most basic level, the interplanar spacings in different crystal structures are characteristic of that structure.

We'll see that the *positions* of the diffracted beams of electrons are determined by the size and shape of the unit cell and the *intensities* of the diffracted beams are governed by the distribution, number, and types of atoms in the specimen. We'll also show you in Part 3 how diffraction leads to contrast in TEM images which is controlled by the orientation of a crystal with respect to the electron beam and which you can control simply by tilting your specimen.

DP PLUS IMAGE

We can distinguish different crystal structures by observing and measuring DPs. The combination of the DP and the electron image(s) is a most powerful tool for characterizing crystals and particularly their defects.

It's easy to see, in a qualitative manner, how diffraction modifies the distribution of the low-angle scattering, described by $f(\theta)$, and shown for a single atom in Figure 3.5. When we consider the effect of the arrangement of atoms in the specimen, then Figure 3.5 has to be modified. For an amorphous specimen, the atoms are almost (but not quite) randomly arranged. A random arrangement would result in a plot similar to Figure 3.5, but there are certain interatomic spacings that tend to occur in an amorphous structure (e.g., first- and second-nearest neighbor spacings are usually relatively well defined). As a result, the amplitude (and hence the intensity) of diffraction is stronger at some angles than at others, so we see diffuse, bright rings on the TEM screen. If the specimen is crystalline, then the intensity of the diffracted beams is a maximum at specific angles because the interplanar spacings are very well defined. The variation of $f(\theta)$ with θ plotted in Figure 3.7A and B is equivalent to the radial intensity variation across the DPs in Figure 2.13A and C, respectively, and thus emphasizes the strong relationship between $f(\theta)$ and diffracted intensity. We'll describe this important relationship mathematically in Section 3.10.B below.

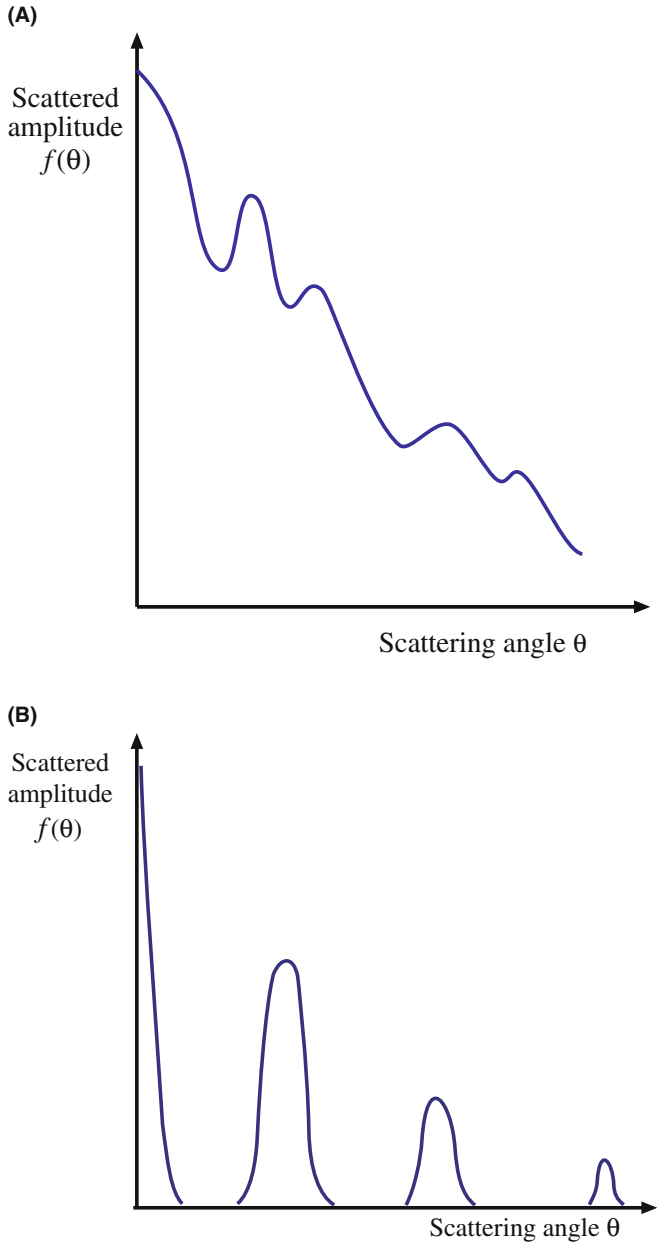


FIGURE 3.7. Change in $f(\theta)$ with θ for (A) an amorphous specimen and (B) a crystalline specimen. The amplitude (and therefore the intensity) of scattering generally decreases with increasing θ but the smooth decrease is modified at certain scattering angles (compare these curves with the intensity variation along a radius of the DPs in Figure 2.13A and C, respectively).

3.10.A Interference of Electron Waves; Creation of the Direct and Diffracted Beams

To interpret low-angle elastic scattering (which is primarily from the electron cloud) it is best to think in

terms of electron waves and not in terms of particle-particle interactions that characterize high-angle Rutherford scattering. If you go back and look at Figure 3.2 you see a periodic one-dimensional array of scattering centers (slits), and a monochromatic wave (i.e., fixed λ) is advancing toward these centers. Each center acts as a new source of a wave of the same λ. Thus many new waves are created and, when more than one wave is present, the waves can interfere with one another. This process happens from even the thinnest specimens and is entirely a wave phenomenon that doesn't need concepts such as cross section, which we apply when we think of the electron as a particle.

A rule of wave theory is that waves reinforce one another (this is constructive interference) when they are in phase. Waves also cancel one another (destructive interference) when they are out of phase. What you see in Figure 3.2 is that the diffracted waves are in phase with one another only in certain directions. There is invariably a *zero-order wave* that proceeds in the same direction as the incident wave, which in the TEM we'll refer to as the direct beam of electrons, as we defined at the start of Chapter 2. There are also *higher-order waves* that propagate in forward directions that are at some fixed (but very small) angle to the incident wave and we'll call these the diffracted beams.

So diffraction creates many electron beams traveling at specific angles relative to a single monochromatic incident beam. In the chapters on diffraction in Part 2, we'll find ways to measure these angles and relate them to the spacing of the scattering planes.

DIRECT AND DIFFRACTED

The direct beam consists of electron, scattered in the same direction as the incident beam. Often in TEM terminology these electrons are called the transmitted beam but this term is ambiguous since, in fact, all forward-scattered beams are 'transmitted' through the specimen.

3.10.B Diffraction Equations

Here we'll introduce the mathematical relationships that describe the diffraction process. The idea of using diffraction to probe the atomic structure of materials was credited to von Laue (1913) in Germany, although others such as Ewald were working on similar ideas at the same time. Von Laue's crucial idea was that much shorter electromagnetic rays than light would cause diffraction or interference phenomena in a crystal. Although his colleague Sommerfeld, with whom he discussed the idea while skiing, disagreed, Friedrich, one of Sommerfeld's assistants, and Knipping tested the idea experimentally by irradiating a copper sulfate crystal and became the first to observe diffraction from crystal planes. In fact it was a remarkable stroke of luck that the $CuSO_4$ diffracted the X-rays at all because of the strict equations that govern diffraction.

Von Laue used the well-known light-optics approach to argue that the diffracted waves are in phase if the path difference between waves scattered by adjacent scattering centers is a whole number of wavelengths, $h\lambda$ (h is an integer). Thus, as shown in Figure 3.8, if the scattering centers (B and C) are spaced some distance a apart and the incident beam (wavelength λ) makes an angle θ_1 with the line connecting the scattering centers and is diffracted at an angle θ_2, then the path difference AB – CD is

$$a(\cos\theta_1 - \cos\theta_2) = h\lambda \tag{3.18}$$

Now in three dimensions, two more Laue equations can be written for two more distances b and c and appropriate angles θ_n

$$b(\cos\theta_3 - \cos\theta_4) = k\lambda \tag{3.19}$$

$$c(\cos\theta_5 - \cos\theta_6) = l\lambda \tag{3.20}$$

These three simultaneous equations bear von Laue's name and for his original suggestion and the analysis of the experiments of Sommerfeld's students, he received the Nobel Prize in Physics in 1914 (nice work: three equations). We'll show in Chapter 11 that

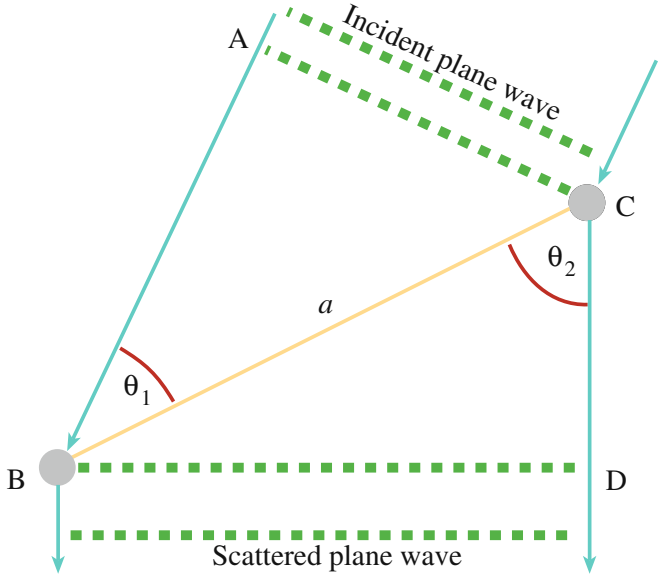


FIGURE 3.8. The approach used by von Laue to calculate the path difference for a wave (wavelength λ). In this one-dimensional figure the wave is incident at an angle θ_1 and scattered at an angle θ_2 from two atoms (B and C) spaced distance a apart. The path difference between scattered waves is AB – CD.

in a TEM specimen, when all three Laue equations are satisfied simultaneously a diffracted beam is produced. We'll also show you in Chapters 11 and 12 that the letters *hkl* are the indices of the diffracted beam and are equivalent to the Miller indices (*hkl*) of the diffracting crystal plane (or some multiple thereof).

Usually in TEM, we use a simpler approach to describe diffraction. Von Laue's approach was simplified by the family team of Sir William H. (obviously the dad) and Mr. W. Lawrence Bragg (the son) in England who proposed (Bragg and Bragg 1913) that the waves behaved as if they were reflected off atomic planes as shown in Figure 3.9.

In parallel with von Laue's optical approach, the Braggs argued that waves reflected off adjacent scattering centers must have a path difference equal to an integral number of wavelengths, if they are to remain in phase. So, in the TEM the path difference between electron

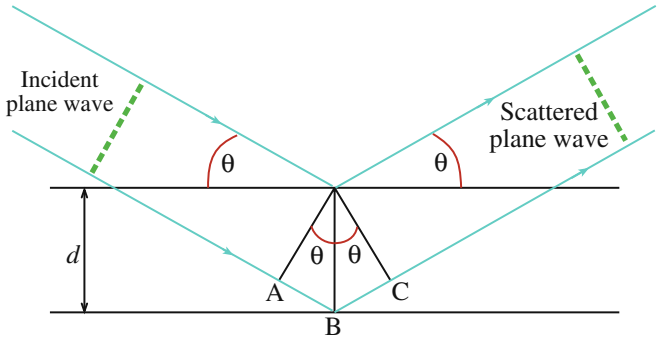


FIGURE 3.9. The Bragg description of diffraction in terms of the reflection of a plane wave (wavelength λ) incident at an angle θ to atomic planes of spacing *d*. The path difference between reflected waves is AB + BC.

waves reflected from the upper and lower planes in Figure 3.9 is (AB + BC). Thus, if the 'reflecting' *hkl* planes are spaced a distance *d* apart and the wave is incident and reflected at an angle θ_B, both AB and BC are equal to $d \sin \theta_B$ and the total path difference is $2d \sin \theta_B$. So we can write what is known as Bragg's (although grammatically and historically it should be Braggs') law

$$n\lambda = 2d \sin \theta_B \tag{3.21}$$

We'll reserve θ_B for the Bragg angle, which is *the* most important scattering angle (remember we really mean *semi*-angle) in TEM and you'll come across it many more times in this text. The Braggs also received a Nobel Prize in Physics a year after von Laue but this time for only one equation (even nicer work!) and despite the fact that the idea of reflected electrons, while mathematically correct, is physically wrong. We'll continue to use the term Bragg reflection to describe diffraction in the TEM because everyone does so, even though it's inaccurate, and because it is extremely useful. However, we'll demonstrate to you, in a rigorous fashion, the mathematical equivalence of the Bragg and von Laue approaches in Chapter 12.

It is simple to see from the Bragg equation that atomic planes which are closer together give rise to larger angles of scatter. This reciprocal relationship (*d* is proportional to $1/\theta$; see Chapter 12) is very important in diffraction-pattern interpretation. So, if you know λ for the incident electron (which you control by choosing the accelerating voltage) and you can measure θ experimentally, you can work out the interplanar spacings in your specimen. It is this crystallographic information that makes diffraction such an important aspect of the TEM.

CHAPTER SUMMARY

What should you remember from this chapter? Until you have time to study this material very carefully you may find it difficult, so here are a few suggestions:

- Know the words! In particular, we can describe the scattering process by three parameters

  $\sigma(\theta)$ the scattering cross section

  $\dfrac{d\sigma(\theta)}{d\Omega}$ the differential scattering cross section

  $f(\theta)$ the atomic-scattering amplitude

In particular, don't be put off because 'differential scattering cross section' sounds difficult. All three terms are *very* important in different parts of TEM.

- The relationships between $f(\theta)$ and $\sigma(\theta)$ are very important (as a principle, but not much used in practice).

- The relationship between $f(\theta)$ and the intensity in a DP is very important.

Remember that, although we often write $\sigma(\theta)$ as σ, there is an angle involved in any σ

- The fact that the electron is a charged particle is critical to the whole scattering process.
- The strength of the scattering, $f(\theta)$, depends inversely on the scattering angle, θ.

Yes, a really rigorous treatment of scattering would take into account the wave nature of the electron (wave mechanics), relativity, and the electron charge, all at the same time. Because we're good guys we won't inflict this on you or ourselves. Fortunately, if required, we can do very well using compiled tables of cross sections and scattering data, which are available on the web (e.g., URL #1).

We can describe the effect of the crystal structure on the electron scattering by one more parameter, the structure factor $F(\theta)$

- $F(\theta)$ is a measure of the amplitude scattered by a unit cell and $|F(\theta)|^2$ is proportional to the scattered intensity.

The diffraction process from a TEM specimen is usually described by the Bragg equation which tells us the important reciprocal relationship between atomic-plane spacings and scattering angles.

A final point to think about: remember that $f(\theta)$ is the property of a 'scattering center.' We usually think of this center as being an atom. What happens if the scattering center is an ion (i.e., if it is charged)? Is the scattering process affected by how this atom is bonded to its neighbors? What changes if the atom has a covalent rather than a metallic bond? These are important questions (otherwise we wouldn't ask them) and we'll teach you the answers as we go on.

DIFFRACTION AND SCATTERING

Andrews, KW, Dyson, DJ and Keown, SR 1967 *Electron Diffraction Patterns* The Institute of Physics Bristol UK. The original text focused on diffraction in the TEM.

Bragg, WH and Bragg, WL 1913 *The Reflection of X-rays by Crystals* Proc. Roy. Soc. Lond. **A88** 428–438. The paper that led to the Nobel Prize.

Doyle, PA and Turner, PS 1968 *Relativistic Hartree-Fock X-ray and Electron Scattering Factors* Acta Crystallogr. **A24** 390–397.

Mott, NF and Massey, HSW 1965 *The Theory of Atomic Collisions* Oxford University Press New York.

Reimer, L 1997 *Transmission Electron Microscopy; Physics of Image Formation and Microanalysis* 4th Ed. Springer New York. Rigorous thorough treatment of the scattering process, especially as used in Section 3.8.

Rutherford, E 1911 *The Scattering of* α *and* β *Particles by Matter and the Structure of the Atom* Phil. Mag. **21** 669–688. His scattering.

von Laue, M 1913 *Kritische Bemerkungen zu den Deutungen der Photoframme von Friedrich und Knipping* Phys. Z. **14** 421–423. Paper that led to the Nobel Prize.

Wang, ZL 1995 *Elastic and Inelastic Scattering in Electron Diffraction and Imaging* Plenum Press New York. Much more detailed than the approach used here.

SCATTERING APPLIED TO EM

Hall, CE 1953 *Introduction to Electron Microscopy* p 229 McGraw-Hill New York.

Jones, IP 1992. *Chemical Microanalysis Using Electron Beams* Institute of Materials London.

Newbury, DE 1986 *Electron Beam-Specimen Interactions in the Analytical Electron Microscope* in *Principles of Analytical Electron Microscopy* p 1 Eds. DC Joy, AD Romig Jr and JI Goldstein Plenum Press New York.

Varela, M, Lupini, AR, van Benthem, K, Borisevich, AY, Chisholm, MF, Shibata, N, Abe, E and Pennycook, SJ 2005 *Materials Characterization in the Aberration-Corrected Scanning Transmission Electron Microscope* Annu. Rev. Mat. Sci. **35** 539–569. Review of Z-contrast imaging.

URLs

(1) www.nist.gov/srd/nist64.htm NIST Standard reference database #64 provides values of the differential elastic-scattering cross sections, total elastic-scattering cross sections, phase shifts, and transport cross sections for elements with $Z = 1$ to 96 and for beam energies from 50 eV to 300 keV (in steps of 1 eV).

SELF-ASSESSMENT QUESTIONS

Q3.1 What are the primary causes of elastic scattering?
Q3.2 What do we mean by the term 'wave-particle duality'?
Q3.3 What forces act on an electron as it interacts with atoms?
Q3.4 What term describes the strength of the scattering process?
Q3.5 What factors control the interference between waves?
Q3.6 How is the scattering amplitude related to the intensity of the scattered beams that we see in the microscope?
Q3.7 What are the two principal forms of elastic scatter?
Q3.8 Relate the general form of the Rutherford differential cross section (equation 3.3) to the equation describing the cross section for nucleus scattering proposed by Hall (equation 3.2).
Q3.9 What is a screening parameter and why do we need to incorporate it in the equations that describe scattering?
Q3.10 Why is it important to include a screening parameter in the Rutherford cross section?
Q3.11 Why do elastic electron-electron interactions usually result in a relatively low scattering angle, while elastic electron-nucleus interactions cause higher-angle scattering?
Q3.12 From your answer to the previous question describe the different information that might be contained in low-angle and high-angle scattered electrons and how you might obtain that information
Q3.13 How thin should your specimen be so that scattering within it approaches the ideal of a single event per electron?
Q3.14 What is the relationship between the atomic scattering factor $f(\theta)$ and the structure factor $F(\theta)$?
Q3.15 Why do crystalline and amorphous specimens give rise to different scattering distributions?
Q3.16 What are the fundamental differences between the von Laue and Bragg approaches to diffraction and what are the similarities?
Q3.17 Put some reasonable values for d and λ into equation 3.21 and calculate a typical Bragg angle in a TEM.
Q3.18 Why is the Bragg approach fundamentally incorrect?
Q3.19 What do we mean by the term 'scattering center'?
Q3.20 What is the relationship between the spacing of the lattice planes and the angle of scatter?

TEXT-SPECIFIC QUESTIONS

T3.1 In Figure 3.1, why are the electrons interacting with both the nucleus and the electron cloud shown to deviate in the same directions (i.e., both are bent through an angle θ) when the nucleus and the electron clouds in fact have opposite electrical charges?
T3.2 Why do we show the electron close to the nucleus in Figure 3.1 as being turned around rather than being pulled directly into the (highly positively charged) nucleus?
T3.3 Look again at Figure 3.1 and explain why elastic electron-electron interactions usually result in a relatively low scattering angle, while elastic electron-nucleus interactions cause higher-angle scattering.
T3.4 In Figure 3.2 why don't we see a third-order scattered beam?
T3.5 Relate Figure 3.3 to equations 3.1 and 3.2.
T3.6 Can you show that the data in Figures 3.3 and 3.4 are consistent? (Hint: assume that θ is small (i.e., $\sim 0^\circ$) for elastic scatter.)
T3.7 Relate Figure 3.5 to equations 3.1 and 3.2.
T3.8 Is Figure 3.5 plotted for a screened or unscreened atomic potential? Explain your answer.
T3.9 In Figure 3.6 why don't we show constructive interference of waves going back in the direction of the incident beam?
T3.10 Figures 3.3, 3.5, and 3.7 all have the same general form. Why is this?
T3.11 Discuss the advantages and disadvantages of the Rutherford cross section for elastic scattering. Put in some values into equation 3.7 and determine the value of the cross section.
T3.12 Write down concise definitions of coherent, incoherent, elastic, and inelastic as we use them. (Hint: first take a look at Webster's.)
T3.13 Explain in a paragraph the relationship between scattering cross section and atomic scattering factor mentioning the important factors that influence them.
T3.14 In Figure 3.6, if this process were Bragg diffraction, how would the Bragg angle relate to θ?
T3.15 Use equation 3.21 to determine the value of θ if $n = 1$ and $d = 2$ for each of the wavelengths in Table 1.2. Thus, discuss whether or not relativistic corrections are important.
T3.16 Copy Figure 3.8 and draw on it where other atoms in the diffracting planes might be positioned. (Hint: look at Figure 3.9.)
T3.17 Explain why we talk about the Bragg angle in Figure 3.9 as being a semi-angle of scattering.

4

Inelastic Scattering and Beam Damage

CHAPTER PREVIEW

In the previous chapter, we discussed elastic scattering of the electron beam in which the incident electrons lost no energy as they traversed the specimen. Inelastically scattered (often termed energy-loss) electrons are equally important and we'll discuss many of these processes here, but leave the applications till later. Why are we interested in inelastic scattering? Well, such scattering generates a whole range of signals, each of which can tell us more about the chemistry of the specimen than we can find out from the elastic electrons. In addition to the energy-loss electrons themselves, the most important signals are the characteristic X-rays, secondary electrons, and, occasionally, visible light (cathodoluminescence (CL)) and so we'll emphasize how these arise. We will also tell you why all these signals are useful, to varying degrees, to materials scientists, engineers, and nanotechnologists.

So how do we use these inelastic signals? First we have to detect them and we'll describe electron detection in general in Chapter 7 and the spectrometers that disperse electrons according to their energy in Chapter 37. Then in Chapters 38–40 we'll talk about analyzing these energy-loss electrons. We will discuss how to detect X-rays in Chapter 32 and how we get quantitative, elemental information from the spectra in Chapters 33–36. In all cases we get complementary information to that gained in TEM images and DPs. We'll briefly discuss CL images and spectra in Chapter 29. Obviously there's a lot of useful information in these signals and this is a major advantage to using ionizing radiation. However, the other side of the coin is that all the inelastic processes deposit energy in your specimen, which can be damaged if it is beam sensitive. So we must also look at the downside of the inelastic processes and we end the chapter by discussing this problem under the general topic of beam damage or radiation damage.

A warning: This chapter is based on some quite difficult theoretical, physical concepts. However, these concepts form the basis of AEM, which constitutes Part 4 of the book, so we have to address the material. You can safely delay studying much of this material in detail until you reach Chapter 32 and beyond.

- What you need to get out of this chapter is an appreciation that a modern TEM is primarily a signal-generating and detecting instrument, not just a tool for producing high-magnification images.

One redeeming feature is that in all these processes we can treat the electron as a particle. So this is (almost) a wave-free chapter, which generally means that it is more easily understandable unless you have an in-depth physics background.

4.1 WHICH INELASTIC PROCESSES OCCUR IN THE TEM?

Historically, the conventional TEM used only two *elastic* signals, namely, the direct beam and the diffracted beam(s). As we've seen, these signals constitute the DP and we'll see in due course how they can be used to produce images. In operating a TEM in this classical manner we are being extraordinarily inefficient; we throw away a vast amount of information about our specimen which is contained in the signals that result from *inelastic* scatter. Some of these signals are shown back in Figure 1.3 and are often used predominantly in related instruments such as the SEM and the Auger electron spectrometer (AES), but we can also use TEMs to detect many of these signals,

thus allowing for a more complete characterization of the specimen.

Because some of the incident-beam electrons lose energy, all these signals are related to the general topic of electron energy-loss spectrometry (EELS). The EELS signals and the accompanying X-ray signals constitute analytical electron microscopy (AEM), which we cover in Part 4. In seeking to detect more signals from the specimen, we find that practically we cannot do everything at once, nor can we do it all with equal efficiency. Nevertheless, various kinds of analytical TEMs exist which, in one form or another, can detect all the signals shown in Figure 1.3. With the advent of aberration correction, the spatial resolution and the detection limits of the various techniques have reached or closely approach the single-atom level and so are very well suited to the characterization of nanostructured materials.

In this chapter we'll cover all the signals that are detectable and what use (if any) they are to the nanotechnologist. We need to know

- What are the inelastic-scattering interactions?
- What is the range of energy losses associated with each process?
- What is the likelihood that each energy-loss process will occur?
- What is the scattering angle for the various energy-loss electrons?

When a high-energy electron encounters an atom, it first penetrates the outer, loosely bound electron cloud, then it passes the more tightly bound inner (or core) shell electrons, and finally it may encounter the nucleus.

RULE OF THUMB

The deeper the electron penetrates into the atom, the greater the energy that may be lost. It is possible (but very rare) for the electron to lose all its energy in a single nuclear interaction.

This range of inelastic scattering produces a range of scattering angles, but there is no simple relationship between the energy lost and the scattering angle. We'll separate the inelastic processes into three components

- Processes that generate X-rays
- Processes that generate other (secondary) electrons
- Processes that result from collective interactions with many atoms or electrons

We know the first two rather well, but the third is usually poorly defined. Figure 4.1 shows the cross sections for the most important inelastic processes

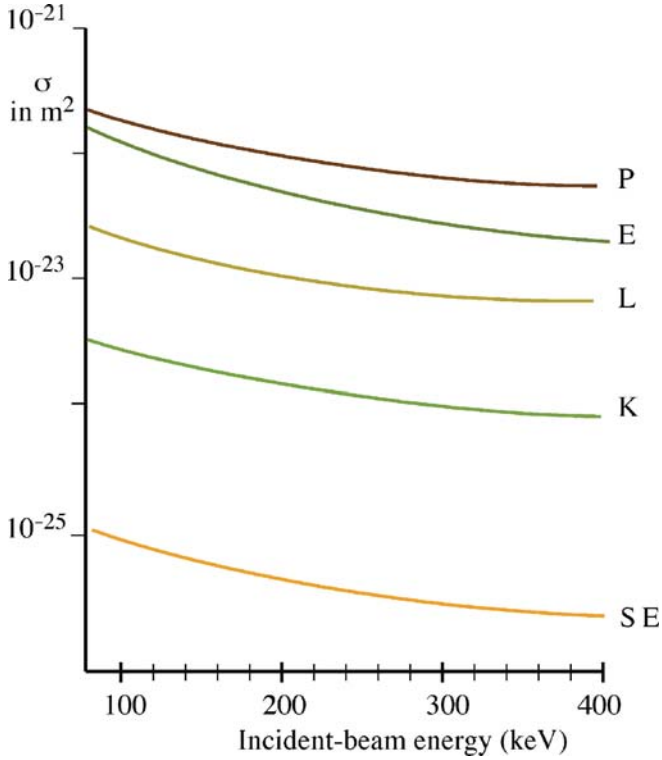


FIGURE 4.1. Cross sections for the various inelastic scattering processes in Al as a function of the incident electron energy, assuming a small angle of scattering ($\theta \sim 0°$); plasmon (P), K- and L-shell ionization (K,L), secondary-electron generation (SE). For comparison purposes the elastic cross section (E) is also included. The values are relatively insensitive to the beam energy.

that we'll talk about. As you can see, the elastic and one of the inelastic (plasmon excitation) processes are by far the most likely events and together account for almost all of the *total scattering cross section* that we discussed back in Chapter 2 (with the caveat that the data in this figure are for small-angle scattering only). These cross sections vary over several orders of magnitude and this fact alone should give you some feel for the relative generation probability of each signal. We'll discuss the specific cross sections for inelastic scattering in more detail as we describe each individual scattering event.

We'll see throughout the book that energy-loss processes are both useful and damaging. For example, in Chapter 19 we describe how energy-loss electrons cause Kikuchi lines to arise in DPs and these are extraordinarily useful. In contrast, some of those same energy-loss electrons cause diffuse scatter that lowers the signal to background information in all DPs and images. If your specimen is thick enough, the energy-loss electrons hide all the useful contrast information, but we'll see in Part 4 how to use EELS to filter out those electrons from images and DPs. This filtering improves the quality of both image and DP and allows the study of much thicker specimens.

4.2 X-RAY EMISSION

We'll consider X-ray emission first because it's the most important secondary signal generated in the specimen. From X-rays we can quickly find out what elements constitute the part of the specimen interacting with the electron beam and we can also quantify the amount of each element in quite a straightforward manner. (The way to do all of this is described in Part 4.) Two kinds of X-rays are produced

- Characteristic X-rays; we'll see these are very useful for local elemental analysis of nano-structured materials and crystal defects.
- Bremsstrahlung X-rays, which are useful to the biologist, but generally regarded as a nuisance by most materials scientists (nano or otherwise).

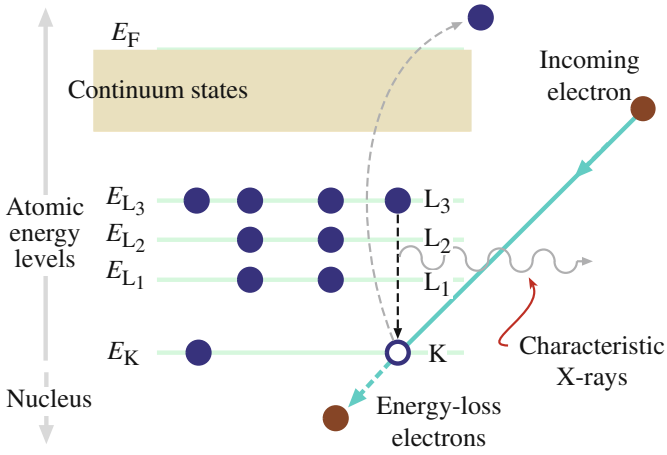


FIGURE 4.2. The ionization process. An inner (K) shell electron is ejected from the atom by a high-energy electron. When the hole in the K shell is filled by an electron from the L shell, characteristic (K_α) X-ray emission occurs. The beam electron loses energy but continues on through the specimen.

4.2.A Characteristic X-rays

How do we produce characteristic X-rays and of what are they 'characteristic'? First of all, a high-energy beam electron must penetrate through the outer conduction/valence bands and interact with the inner-shell (or core) electrons. If more than a critical amount of energy is transferred to an inner-shell electron, that electron is ejected; i.e., it escapes the attractive field of the nucleus, leaving a hole in the inner shell. In an isolated atom, the electron is ejected into the vacuum while in a solid it escapes above the Fermi level into the unfilled states. The atom is then left in an excited state because it has more energy than it would like, and we describe it as ionized.

The ionized atom can return almost to its lowest energy (ground state) by filling in the hole with an electron from an outer shell. This transition is accompanied by the emission of either an X-ray or an Auger electron. This latter process was first described by Auger (1925) and won him the Nobel Prize for Physics. (Since the discoverer was French, we pronounce his name "Ozhay" with a soft *g* as in beige.) In both the X-ray and Auger cases, the energy of the emission is *characteristic* of the difference in energy between the two electron shells involved and this energy difference is unique to the atom. The process of X-ray emission is shown schematically in Figure 4.2. We'll cover Auger emission in Section 4.3.B

Note that characteristic X-rays can also be produced if an atom is ionized by a process other than electron irradiation. For example, ionization can occur as a result of X-ray bombardment also, in which case we use the term fluorescence. It is customary *not* to refer to electron-induced X-ray emission as fluorescence, although you may occasionally come across such usage in the literature.

We've been able to detect X-rays in electron microscopes for many years, but Auger electron detection is rather specialized and usually carried out in a dedicated AES. More recently, however, we've found ways to detect the Auger signal in ultrahigh vacuum (UHV) TEMs and so we'll discuss this in Section 4.3.B below.

You need to know several aspects of the ionization process to understand why the characteristic X-rays are so useful and what it takes to generate them

- What are *electron shells*?
- Why do we use the terms X-ray *lines, families, and weights*?
- What is the *critical ionization energy* and the *ionization cross section*?
- What controls the *X-ray energy* and *wavelength*?
- What is the *fluorescence yield*?

Electron shells: We use a specific terminology to identify the different characteristic X-rays. To understand the terminology you must be familiar with the simple Bohr theory of atomic structure in which the electrons are circling the nucleus in specific shells. (The electrons stay in their shells rather than spiral into the nucleus because of the constraints imposed by quantum theory.)

Aside: For historical reasons, the innermost electron shell is called the K shell and the next innermost is the L shell, the next the M, and so on, as used in Figure 4.2. All the shells (except the K shell) may themselves have subshells (e.g., L_1, L_2, etc.). We name the characteristic X-rays in terms of the shell being filled and the shell from which the electron comes. (The K, L, M, etc., terminology was first introduced by Charles Barkla, an early X-ray spectroscopist and has nothing to do with Royal

Dutch Airlines. The reason Barkla chose K as the first shell may have been because he wasn't sure if he'd need a J shell but knew he'd need an L shell!)

Remember that the difference between the two shell energies equals the energy of the characteristic X-ray. Thus, if we fill a K-shell hole from the L shell we get a K_α X-ray, but if we fill it from the M shell we get a K_β X-ray. If the hole is in the L shell and we fill it from the M shell we get an L_α X-ray, and if we fill it from the N shell we get an L_β X-ray. The notation is in fact much more complex because we differentiate the α X-rays in terms of α_1 and α_2 depending on which subshell of the outer shell the electron falls from to fill the hole. The α_1 X-ray is from the outermost subshell (e.g., the L_{III} or M_V), the α_2 from the next innermost (the L_{II} or M_{IV}). To make this a bit clearer you can look at the diagram in Figure 4.3, although it may in fact confuse you more since it's not at all obvious why, with all these possible electron transitions, only a small fraction generates sufficient X-rays for us to use. Suffice it to say that X-ray physics is an arcane discipline. Fortunately, for X-ray detection in the TEM

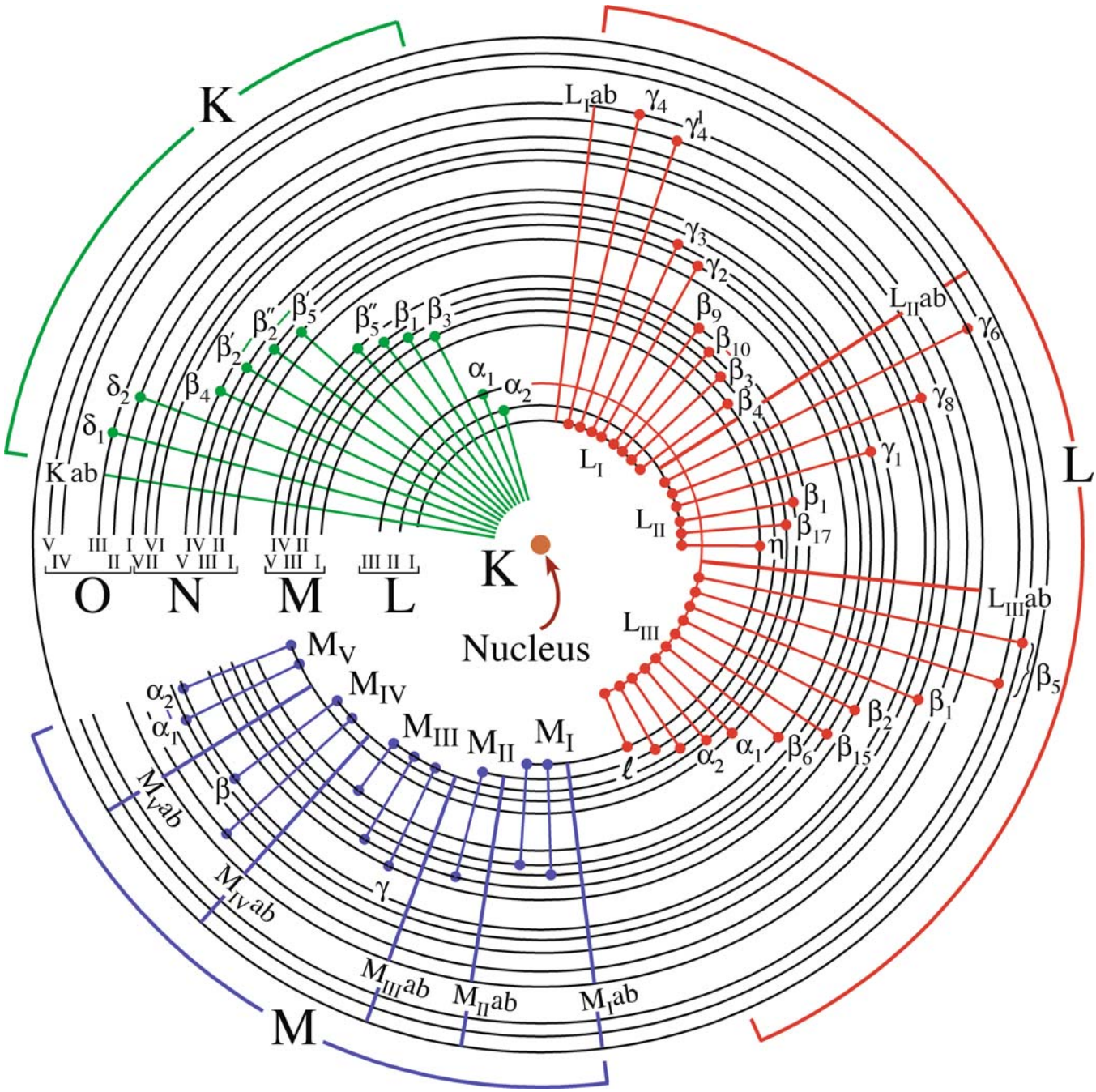


FIGURE 4.3. The complete range of possible electron transitions that give rise to K, L, and M characteristic X-rays. Not all these X-rays are detectable by the XEDS in the TEM.

you don't need to worry about such details because, as you'll see later, the detectors we use can't usually discriminate between the X-rays from different subshells, except at the highest X-ray energies, so K, L, and M and α and β are about all you'll need to remember.

X-ray lines, families, and weights: Often we refer to specific characteristic X-rays as 'lines' because they originally appeared as lines on photographic plates from early spectrometers. Each characteristic X-ray line has a specific wavelength or energy. Groups of lines that arise from transitions to a specific K, L, or M shell are often called 'families.' Much more detail can be found in books on X-rays and X-ray spectrometry, and we'll tell you more in Chapter 34.

Not all electron transitions are equally probable (think of those cross sections again) and this is taken into account by the different 'weights' (i.e., relative intensity compared with the most intense line) of the X-ray lines that are given in Table 4.1. These weights are only important within a given K, L, or M family and do not relate families (e.g., K to L), because experimental conditions affect X-ray generation in each family differently. In X-ray analysis in the TEM we only use the most intense lines, usually the α lines (or, if the spectrometer can't resolve them, both α and β lines). This will become more obvious when you've learned about X-ray qualitative analysis in Chapter 34.

K,L,M,...
A family is a group of X-ray lines and each member of the family has a different relative intensity, which unfortunately is called its weight.

Critical ionization energy: The electron beam has to transfer an amount of energy greater than a certain value to the inner-shell electron in order to ionize the atom. This energy is called the critical ionization energy (E_c). If we want to generate a useful number of X-rays, then the beam energy E_0 must be significantly greater than E_c. The value of E_c increases as the electrons are more tightly bound to the nucleus so, for a given element, the innermost shell (K) has a higher E_c than the next (L) shell, and so on. Atoms with higher Z have more protons binding the core electrons to the nucleus and, therefore, have a higher E_c. You can see this effect if you go and look at an X-ray spectrum, e.g., in Figure 1.4A (or one of many others throughout Part 4) in which the energy of the X-ray peaks increases with increasing atomic number. Since there's a lot of shells and a lot of atoms, the list of critical ionization energies is long, as you'll find out if you consult any X-ray textbook. A similar list is also invaluable in EELS since the E_c obviously corresponds to a critical amount of energy lost by a beam electron and thus gives rise to peaks (usually called edges) in the energy-loss spectrum. As we'll see in Chapter 39, EELS edges, like characteristic X-ray peaks, can also be used to identify uniquely the presence of a particular element in the specimen.

TABLE 4.1 Relative Weights of X-ray Lines

K_α (1)	K_β (0.1)		
$L_{\alpha 1,2}$ (1)	$L_{\beta 1}$ (0.7)	$L_{\beta 2}$ (0.2)	$L_{\gamma 1}$ (0.08)
	$L_{\gamma 3}$ (0.03)	L_l (0.04)	L_η (0.01)
M_α (1)	M_β (0.6)	M_ζ (0.06)	M_γ (0.05)
	$M_{II}N_{IV}$ (0.01)		

The cross section for ionization (σ): It is shown in Figure 4.1 for K- and L-shell electrons. It is not a strong function of energy and has a relatively large value, and so we expect to see X-rays generated in all TEMs. In low-voltage SEMs we have to worry about another parameter called the overvoltage, U, which is the ratio of the beam energy E_0 to the ionization energy E_c. The cross section varies with U as shown in Figure 4.4. What this figure tells you is that if E_0 is close to E_c then there isn't much chance of ionization occurring. However, in the TEM E_0 is $\geq$ 100 keV and E_c is generally $<$ 20 keV so U is usually $>$ 5. Therefore, X-ray generation is expected and the cross section is pretty constant with energy. Despite this relatively simple behavior, there is considerable uncertainty about the absolute value of the ionization cross sections because few reliable experimental measurements have been made at TEM voltages. Most models are variations on the original expression given by Bethe (1930) which describes the total (not the differential) ionization cross section as

$$\sigma_T = \left(\frac{\pi e^4 b_s n_s}{E_0 E_c}\right) \log\left(\frac{c_s E_0}{E_c}\right) \qquad (4.1)$$

where the only new terms are n_s, which is the number of electrons in the ionized subshell, and b_s and c_s, which are constants for that shell. We are not particularly concerned with any angular variation in the ionization process. The differential form of the Bethe expression shows two features

- The electron that ionized the atom is deviated through a really small angle (<∼10 mrads).

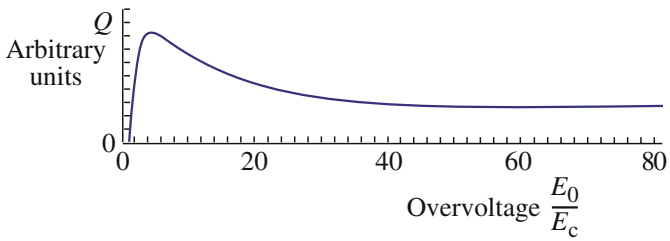


FIGURE 4.4. The variation of the ionization cross section with overvoltage. Ionization is most probable if the beam energy is ∼5× the critical ionization energy. The cross section decreases, but not substantially, at higher overvoltages, typical of a TEM.

- The resultant characteristic X-ray is a spherical wave emitted uniformly over 4π sr.

Remember the Bethe expression is an *inelastic* cross section but, like the Rutherford (elastic) cross section, it also needs to be corrected for the effect of relativity at TEM beam energies. So we substitute the term $m_0v^2/2$ for the beam energy and introduce a standard relativistic factor, β ($= v/c$) (Williams 1933)

$$\sigma = \left(\frac{\pi e^4 b_s n_s}{\left(\frac{m_0 v^2}{2}\right) E_c}\right)\left[\log\left[c_s\left(\frac{m_0 v^2}{2E_c}\right)\right] - \log(1-\beta^2) - \beta^2\right] \tag{4.2}$$

This modified Bethe cross section can be manipulated to fit almost any X-ray data just by altering b_s and c_s, although such parameterization is not always justified. Several cross section models have been developed, all of which are modifications to Bethe's approach (e.g., Newbury 1986, Goldstein et al. 1986). A good source of cross section data can be found on the NIST Web site (URL #1).

The X-ray energy/wavelength: X-rays are electromagnetic radiation and so we usually think of them of as waves with a specific wavelength λ. But, just like electrons, X-rays can show particle-like characteristics and then we describe them as photons with a specific energy such as E_K or E_L, where the subscript refers to the shell from which the core electron was ejected.

There is a similar inverse relationship between the X-ray wavelength and its energy, as we saw for electrons back in Chapter 1. However, there are a couple of important differences which you *must* remember.

- An X-ray is a photon of electromagnetic energy, so the concepts of rest mass and momentum embodied in the electron energy are irrelevant; an X-ray has *no mass*.
- X-rays, like all electromagnetic radiation, travel at the speed of light (c) in vacuum and consequently we don't have to make increasing relativistic corrections as their energy increases. So the quantized X-ray energy is just $h\nu$ where h is Planck's constant and ν is the frequency and in order to express this energy in eV we equate it to E, where E is the X-ray energy.

Thus

$$E = h\nu = \frac{hc}{\lambda} \tag{4.3}$$

Now since h and c are constants we can substitute values with appropriate units into the equation and find that the X-ray wavelength is given by

$$\lambda = \frac{1.24}{E} \tag{4.4}$$

where λ is in nm and E in keV. This expression is *very* similar to the expression for the uncorrected *electron* wavelength ($1.22/E^{1/2}$) where E is the electron energy in eV (not keV) that we gave back in Chapter 1. You can easily confuse the two, so beware!

SPEED

Don't confuse electromagnetic waves (e.g., X-rays), which always travel at the speed of light, with electron waves whose speed depends on their energy.

Because the X-ray energy depends on the difference in the inner-shell energies and these differences increase monotonically with Z, we can use the detection of a characteristic X-ray with a specific energy as an unambiguous sign of the presence of an element in the specimen (although it doesn't necessarily mean that that element is intrinsic to your specimen, as we'll see in Chapter 33). The concept of the atomic number (Z) of the specimen and its relationship to the X-ray energy/wavelength was reported by the brilliant young physicist, HGJ Moseley. Soon after his discovery, Moseley volunteered for the British army and, despite his talents, was dispatched to the trenches of Gallipoli in 1915 where he was promptly killed before he could be nominated for the Nobel Prize, which would undoubtedly have been his. He is remembered by Moseley's law which states

$$\lambda = \frac{B}{(Z-C)^2} \tag{4.5}$$

where B and C are constants. So we can also generate a list of X-ray energies which are associated with each atomic transition. As with E_c the complete list is enormous and given in Bearden's tables. More compact lists are given in the software attached to the XEDS system on your TEM or in handy 'slide rules' by the manufacturers of X-ray spectrometers, in textbooks, review articles, or on trusted Web sites such as NIST (URL #2) or NPL (URL #3).

If you compare the value of E_c and the relevant characteristic X-ray energy you'll see that they are not quite identical. The X-ray energy, E_K or E_L, is invariably less than E_c. Table 4.2 lists a comparison of critical ionization energies and corresponding X-ray energies for a range of elements. Note how the differences in energy increase with increasing Z. The differences arise because the atom doesn't return completely to ground state when the X-ray is emitted. If the electron that fills the hole in the ionized inner shell comes from an outer shell then this process will leave a hole in that outer shell.

TABLE 4.2 Difference Between E_c and E_K

Element	Critical ionization energy E_c (keV)	X-ray energy E_K (keV)
C	0.282	0.277
Al	1.562	1.487
Ca	4.034	3.692
Cu	8.993	8.048
Ag	25.531	22.163

Note that the energies may be affected by bonding states but shifts will only be a few eV.

This hole must also be filled by another electron with perhaps the emission of another X-ray (much more likely as Z increases) and so on until eventually a free electron from the conduction or valence band fills the last hole in the outermost core shell.

An example: A Cu K-shell electron requires 8.993 keV of energy for ionization (E_c = 8.993 keV). One possible sequence by which this extra energy within the atom is lost is first by the creation of a Cu K_α X-ray (8.048 keV), then an L_α X-ray (0.930 keV). These X-ray energies total 8.978 keV and the remaining few eV could come from the hole in the M shell being filled from the conduction band with the emission of a photon or the generation of phonons (see below).

CASCADE TO GROUND

So the ionized atom returns to ground state not via a single event but by a cascade of transitions, depending on the complexity of the electronic structure of the atom.

The possible variations are enormous and affected by such events as Coster-Kronig transitions, in which the atomic shells rearrange their energies after the electron transition. The situation is further complicated if the ionized atom is bound to a different atom, in which case the energy of the X-ray can be shifted slightly (<~5 eV). Such detail is well beyond what you need to know now but a textbook on X-ray spectrometry will give you more, if you so wish.

Fluorescence yield: Remember that an ionized atom does not have to lose energy by giving off a characteristic X-ray but can emit an Auger electron instead. The probability of X-ray versus Auger emission is described by the fluorescence yield, ω, which is the ratio of X-ray emissions to inner-shell ionizations. The fluorescence yield is a strong function of atomic number as shown in Figure 4.5, decreasing at a rate proportional to Z^4 as Z decreases. One common expression for ω gives

$$\omega = \frac{Z^4}{a + Z^4} \tag{4.6}$$

where $a \sim 10^6$ for the K shell. While an approximation, this equation still describes a formidable dependence on Z. For carbon (Z = 6), ω is $\sim 10^{-3}$ and for Ge (Z = 32), ω is ~ 0.5. So you have to ionize 1000 carbon atoms before you get a single C K_α X-ray but only two ionizations are needed to produce a Ge X-ray. So if you ionize low-Z atoms, the chances are you won't see an X-ray and therefore XEDS is *not* the best way to analyze light elements; you should use EELS (see Part 4) because we can always detect the energy-loss electron whether or not it has generated an X-ray. If you want to know more about the chances of generating X-rays in your microscope then the database by Hubbell et al. is a good source.

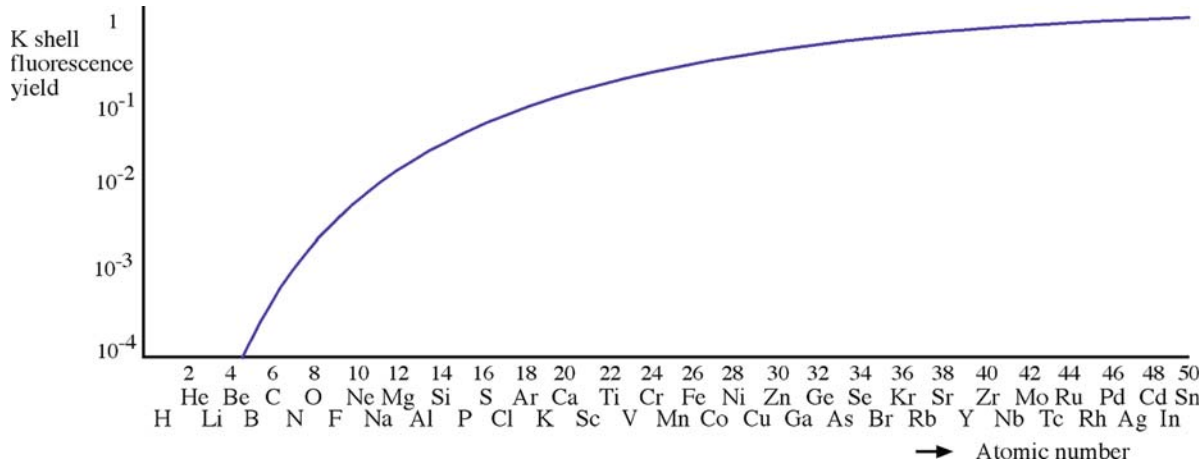


FIGURE 4.5. Fluorescence yield for K-shell X-rays as a function of atomic number. Note the rapid decrease at low atomic numbers. X-rays from elements below Be are not detectable.

4.2.B Bremsstrahlung X-rays

If the electrons in the beam penetrate completely through the electron shells they can interact inelastically with the nucleus. If the electron interacts with the Coulomb (charge) field of the nucleus, it can suffer a substantial change in momentum and during this process it may emit an X-ray. Since the electron can suffer any amount of energy loss, depending on the strength of its interaction, then these X-rays can have any energy up to the beam energy. Such X-rays are known by their original German name of *bremsstrahlung* which can be translated as 'braking radiation.'

The likelihood of bremsstrahlung creation is usually described by the cross section derived by Kramers. This expression is often used for thin TEM specimens, although it was originally derived for bulk samples. It is common to use the Kramers cross section to predict the bremsstrahlung production rather than the probability of interaction. The approximate expression used is

$$N(E) = \frac{KZ(E_0 - E)}{E} \tag{4.7}$$

where $N(E)$ is the number of bremsstrahlung photons of energy E (i.e., the intensity) produced by electrons of energy E_0, K is Kramers' constant, and Z is the atomic number of the ionized atom. This relationship predicts that it is far more likely that the interaction causes a small loss of energy and exceedingly rare that the electron loses all its energy in one deceleration at the nucleus. So the bremsstrahlung intensity as a function of energy is shown in Figure 4.6. In contrast to the isotropic emission of the characteristic X-rays, the bremsstrahlung is highly anisotropic, showing strong forward scattering which increases as E_0 increases. This anisotropy is very useful since it allows us to design spectrometers that collect many more useful characteristic X-rays than relatively useless bremsstrahlung X-rays.

The bremsstrahlung has a continuous energy spectrum on which the characteristic X-rays that we just talked about are superimposed, as also shown schematically in Figure 4.6 and experimentally in the spectrum back in Figure 1.4. Since the characteristic X-rays have a narrow energy range, they appear as sharp peaks in the spectrum centered at specific energies, indicated by computer-generated lines on the display (now another reason to call them 'lines'). The bremsstrahlung intensity depends on the average Z of the specimen and this is useful to biologists or polymer scientists who are interested in this aspect of their specimens. But materials scientists generally dismiss the bremsstrahlung as a signal, which only succeeds in obscuring characteristic lines. We'll come back to the X-ray spectrum in more detail in Chapters 32–36.

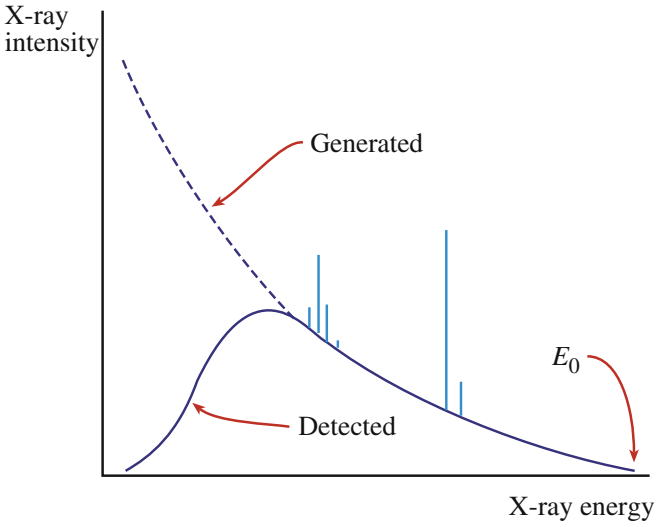


FIGURE 4.6. The bremsstrahlung X-ray intensity as a function of energy. The generated intensity increases rapidly with decreasing X-ray energy but at energies $<\sim$ 2 keV the bremsstrahlung is absorbed in the specimen and in any detector being used so the observed intensity in the detected spectrum drops rapidly to zero. E_0 is the energy of the electrons that cause the X-ray emission. Two families of characteristic lines at specific energies are also shown superimposed on the bremsstrahlung.

4.3 SECONDARY-ELECTRON EMISSION

Secondary electrons (SEs) are electrons within the specimen that are ejected by the beam electron.

- If the electrons are in the conduction or valence bands then it doesn't take much energy to eject them and they typically have energies $<\sim$ 50 eV.
- If the electrons are ejected from an inner shell by the energy released when an ionized atom returns to the ground state, then these SEs are called Auger electrons. The process is often termed a non-radiative transition (since no X-ray emerges from the atom) and the energy undergoes an 'internal conversion' (which is not quite a religious experience).

Historically, SEs were usually considered only in relation to the SEM where they are used to form (often stunning) images which are sensitive to surface topology. We'll now discuss each of these SE signals and their relative importance in the TEM.

4.3.A Secondary Electrons

SEs are ejected from the conduction or valence bands of the atoms in the specimen. The actual emission process can be quite complex and no simple cross section model covers all production mechanisms. The data in Figure 4.1 indicate that SE emission is a far less likely process than all the other inelastic processes we've discussed, but enough are generated for them to be useful in the TEM.

Usually, SEs are assumed to be free electrons, i.e., they are not associated with a specific atom and so they contain no specific elemental information. Because SEs are weak they can only escape if they are near the specimen surface. So we use them in SEMs for forming images of the specimen surface. While SEs are the standard signal used in SEMs, they are also used in STEMs where they can provide very high resolution, topographic images of the specimen surface. We'll discuss ways to detect SEs in Chapter 7 and we'll talk about the images themselves in Chapter 29.

We'll discuss several reasons for the improved (SE) resolution in STEM in Chapter 29. However, recent developments in high-resolution field-emission gun (FEG) SEMs have produced SE image resolution <0.5 nm (close to surface atom resolution) at 30 kV. (We discuss FEGs in the next chapter.) A STEM at 100 kV can offer similar or better resolution even without an FEG, so the SEs are very useful. Aberration correction in STEM naturally brings about even higher-resolution SE images, close to the atomic level.

SE RESOLUTION

SE images in a STEM have much better resolution than SE images in the (relatively) low-kV SEMs.

The number of SEs does depend on energy; it rises to a maximum at about 5 eV and drops close to zero with energies >~50 eV. (You should know that, on rare occasions, strongly bound inner-shell electrons can be ejected with energies up to about 50% of the beam energy. Such fast SEs are generally ignored because they do not seem to limit the resolution of XEDS in the TEM.) The SE yield (number of SEs/incident-beam electron) is generally regarded as being independent of E_0; if there is any Z dependence (which is still a matter of some debate) then it is very small. The angular distribution of emitted SEs is not important since the detector uses a strong field to gather SEs emerging from the surface at any angle. But the *number* of SEs increases with specimen tilt because SEs escape more easily as the surface is tilted parallel to the beam. This behavior is a critical aspect of SE emission because it mimics Lambert's cosine law of visible-light reflection, accounting for the great similarity between SE images of rough specimens and the everyday, reflected-light images we are accustomed to seeing with our eyes.

4.3.B Auger Electrons

Remember we said at the start of this chapter that the emission of Auger electrons is an alternative to X-ray emission as an ionized atom returns to its ground state.

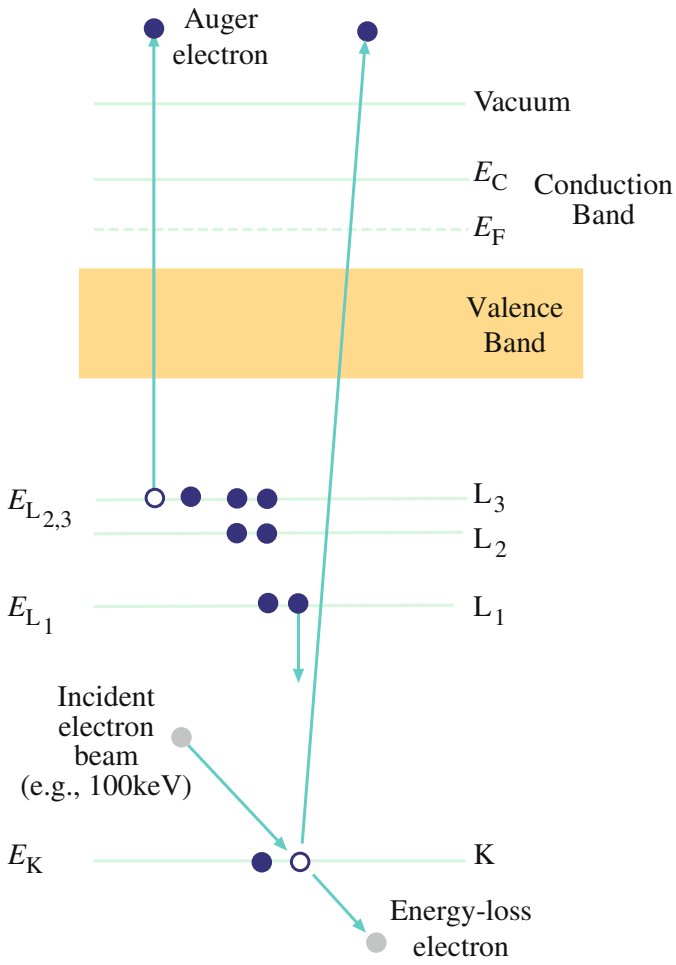


FIGURE 4.7. The process of inner (K) shell ionization and subsequent Auger-electron emission. The energy released when the L_1 electron fills the hole in the K shell is transferred to an electron in the $L_{2,3}$ shell which is ejected as a $KL_1L_{2,3}$ Auger electron.

Figure 4.7 shows how such an atom ejects an outer-shell (Auger) electron; it's instructive to compare with Figure 4.2 for X-ray emission. The ejected electron has an energy given by the difference between the original excitation energy (E_c) and the binding energy of the outer shell from which the electron was ejected. This explains the rather complex nomenclature used to describe each Auger electron (see the caption to Figure 4.7). So the Auger electron has a characteristic energy that is dependent on the electronic structure of the ionized atom and is almost identical to the energy of the alternative, characteristic X-ray.

Because they have such low energies, the Auger electrons that do escape come from very close to the specimen surface. They contain chemical information and consequently AES is a recognized surface-chemistry technique. Because of the similarity in energy between Auger electrons and characteristic X-rays, you might ask, why is light-element X-ray analysis in the TEM not just a surface technique? What you have to

remember is that characteristic X-rays are much less strongly absorbed in the specimen than electrons of similar energy. So most X-rays generated in a thin TEM specimen can escape and be detected. (So it's all to do with the cross sections for interaction in the first place.)

AUGER

The Auger process is favored in atoms having small binding energies, i.e., the lighter elements. Typical Auger electron energies are in the range of a few hundred eV to a few keV and are strongly absorbed within the specimen.

Because Auger emission is a surface phenomenon, the state of the specimen surface is paramount. Oxidation or contamination will prevent interpretable Auger analysis of the true surface chemistry and so we only carry out AES in a UHV system. As a result, the Auger signal has traditionally been ignored by electron microscopists and confined to the realm of surface chemistry, along with such techniques as ESCA and SIMS. However, as TEMs are being built with better vacuums and UHV STEMs become more common, the Auger signal may receive more interest. Unfortunately, it is not simple to attach an Auger system to a STEM unless you build a dedicated instrument in which routine AEM is difficult, so such studies are still very rare.

4.4 ELECTRON-HOLE PAIRS AND CATHODOLUMINESCENCE (CL)

These two signals are closely related. We'll see in Chapter 7 that one way to detect electrons is to use a semiconductor that creates electron-hole pairs when hit by high-energy electrons. So if your specimen happens to be a direct-gap semiconductor then electron-hole pairs will be generated inside it.

Cathodoluminescence is explained schematically in Figure 4.8. The emitted photon has a frequency (i.e., color) equal to the energy of the gap (E_G) divided by Planck's constant (h). If the band gap varies for some reason, there will be a spectrum of light given off or the color of the light will vary depending on what part of the specimen is being observed. So CL spectroscopy has applications in the study of semiconductors and impurity effects therein. While the spatial resolution of CL is not down to the nanometer level like X-rays or secondary electrons, it is still well within the nano-scale range, typically defined as <100 nm.

CATHODOLUMINESCENCE

Electrons and holes will recombine and in doing so give off light; this process is referred to as CL.

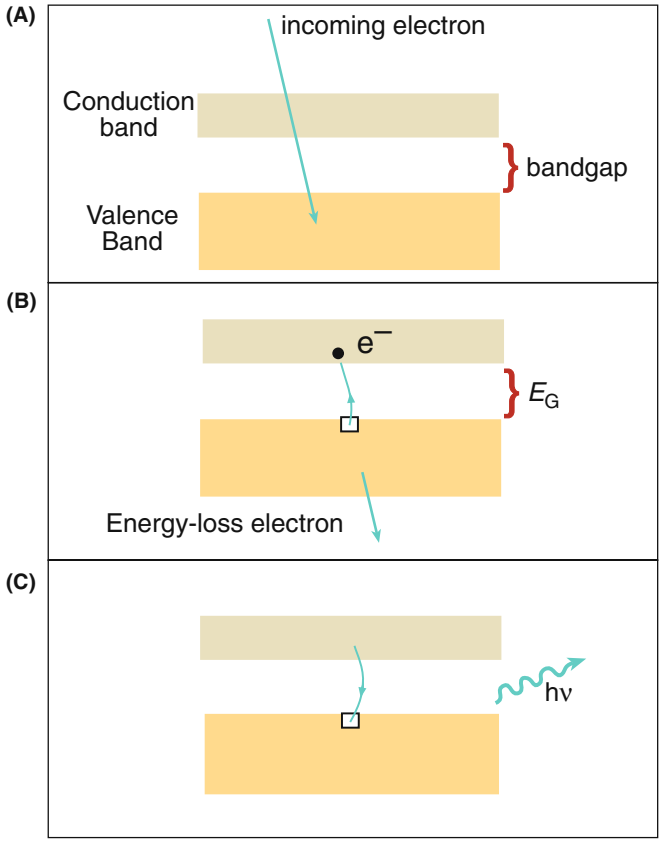


FIGURE 4.8. Schematic illustration of CL. (A) Initial state before a beam electron interacts with valence-band electrons. (B) A valence-band electron is excited across the gap into the conduction band, leaving a hole in the valence band. (C) The hole is filled by a conduction-band electron falling back into the valence-band hole. Upon recombination a photon of light is emitted with a frequency determined by the band gap.

Now if you apply a bias to your specimen or if it happens to be a p-n junction or a Schottky-barrier diode, then the electrons and holes can be separated by the internal bias. You can pick up this charge if you ground the specimen through a picoammeter. In this situation, your specimen is acting as its own detector! The current you then detect is sometimes called the electron-beam-induced current (EBIC) signal. If you detect this signal and use it to form an image then you are doing charge-collection microscopy (CCM).

The CL and CCM modes of operation are standard methods of characterizing bulk samples in the SEM. In principle, there is nothing to prevent us doing the same in a STEM, and a few people have built dedicated instruments, but generally the space available in the TEM stage limits the efficiency of signal collection. This may improve with C_s correction, but, in general, these two techniques are rare and mainly limited to studies of semiconductors (e.g., Boyall et al.) although some minerals also exhibit CL. We'll describe

CL detectors in Chapter 7 and show you an image in Chapter 29. Just remember that CL and CCM are potentially powerful, but highly specialized, techniques.

4.5 PLASMONS AND PHONONS

We can link these two phenomena because they are both examples of what we call collective oscillations.

We can consider plasmons as analogous to sound waves, since they are longitudinal oscillations of the free-electron gas, which create regions of varying electron density as shown schematically in Figure 4.9. These oscillations are damped out in less than a femtosecond and the wave is localized to <10 nm. If you go back to Figure 4.1, you'll see that the plasmon process has the largest cross section so it's by far the most common inelastic interaction occurring in materials and, as we'll see in Chapter 38, plasmon peaks are strong features of EEL spectra. Plasmons can occur in any material with weakly bound or free electrons, but they occur predominantly in metals, particularly ones like aluminum which have a large Fermi surface and, therefore, a high free-electron density. The plasmon oscillation is quantized and the mean free path for plasmon excitation is ~100 nm. As we'll also see in Section 38.3.C this quantization makes the number of plasmon excitations a useful way to measure your specimen thickness. Also, the plasmon energy is a function of the free-electron density and this changes with composition (see Chapter 38 also), so the plasmon excitation is chemically dependent, although we rarely use it for elemental analysis.

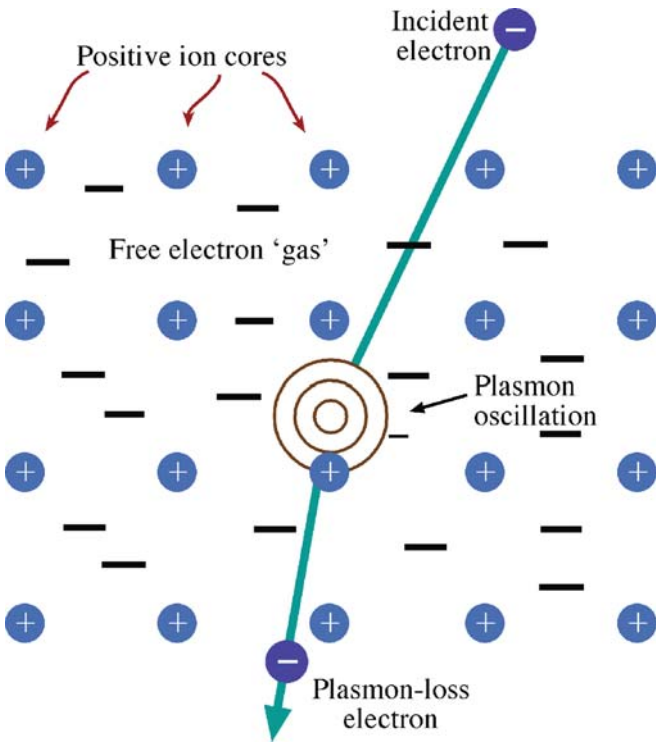


FIGURE 4.9. Schematic diagram of a high-energy beam electron exciting a plasmon oscillation in a free-electron gas that permeates the ion cores in a metal.

PLASMONS AND PHONONS

Plasmons are collective oscillations of free electrons that occur when a beam electron interacts with the free electron 'gas.'

Phonons are collective oscillations of atoms in a solid that arise when the atomic lattice is struck by a beam electron.

The differential cross section for plasmon excitation has a general Lorenztian form

$$\frac{d\sigma_\theta}{d\Omega} = \frac{1}{2\pi a_0}\left(\frac{\theta_E}{\theta^2+\theta_E^2}\right) \tag{4.8}$$

where a_0 is the Bohr radius, θ is the scattering angle, and θ_E is the so-called characteristic scattering angle given by $E_P/2E_0$ (which is always going to be small given the large value of E_0 in TEMs). Since E_P, the plasmon energy, is almost fixed (~15–25 eV), the cross section is a strong function of θ, dropping rapidly to zero at values much above 10 mrads, indicating once again the strong forward scattering of such energy-loss electrons.

When a high-energy electron strikes an atom in the specimen, the lattice shakes, just like hitting a chain-link fence with a stick. This process occurs because, as shown in Figure 4.10, all the atoms are linked together elastically. Phonons can also be generated by other inelastic processes occurring within the atom; for example, the energy of Auger or X-ray emission or an interband transition is sometimes converted internally to lattice vibrations. Any shaking of the atoms is equivalent to heating up the specimen and the net result of all phonons is that the specimen gets warmer. As we will see, this is particularly damaging to some specimens.

The incident electron can generate phonons in any solid specimen, even amorphous ones in which there is no periodic crystal structure. Typically, a phonon vibration causes a very small energy loss of < 0.1 eV but the phonon-loss electrons are scattered out to quite large angles (5–15 mrads), and these electrons account for the diffuse background intensity present between the Bragg intensity maxima in DPs. Phonon-scattered electrons carry no useful chemical information nor do they carry contrast useful to the microscopist.

It is not important to know the phonon-scattering cross section exactly, but it is useful to remember that

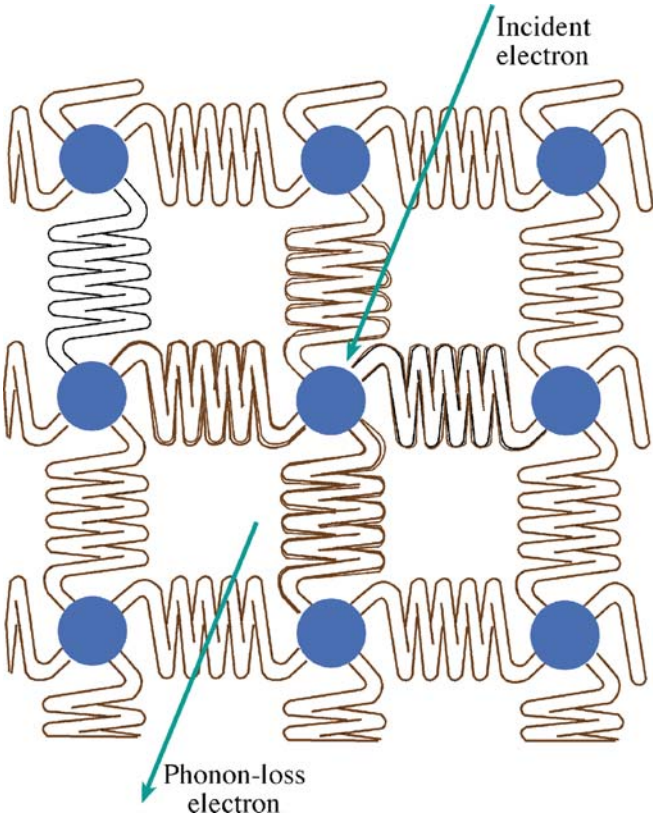


FIGURE 4.10. An illustration of the crystal lattice as a group of atoms linked elastically by springs. The bonds vibrate when struck by a high-energy electron creating lattice oscillations or phonons and these vibrations are equivalent to heating the specimen.

phonon scattering increases with Z with a dependence $\sim Z^{3/2}$, which is a somewhat weaker dependence than for true elastic scattering. Also, because of the effect of temperature on atomic vibration, phonon scattering increases as the temperature rises. This accounts for the increase in thermal-diffuse scattering with temperature and is the major reason why we cool specimens if we want to obtain good, sharp DPs. The mean free path for phonon scattering at room temperature varies from a couple of nm for Au up to about 350 nm for Al, and at liquid-He temperatures these values increase $\sim$ 2–3X.

We don't use either plasmons or phonons directly to form images (although in principle this is possible), but we do detect the electrons that caused them, and we'll discuss the (rather limited) uses of plasmon energy-loss electrons in Chapter 38.

PHONONS

These oscillations involve all the atoms in the crystal lattice vibrating collectively. Such vibrations are equivalent to specimen heating. You can reduce the number of phonons by cooling your specimen.

There are other inelastic processes that can occur as the beam electron traverses the specimen, such as inter-band and intra-band transitions within the valence or conduction bands. The bonding between the atoms and local atomic arrangements within the specimen can also affect many of the inelastic events that we have described in ways that we can both predict and discern. Other electronic features such as the band gap can affect the possibility of certain inelastic interactions occurring. In principle, any atomic or electronic characteristic of your particular specimen that reduces the energy of a beam electron during its passage through the specimen can be detected, measured, and perhaps quantified and simulated. These signals can all be used to form images containing information which complements that contained in the elastic-electron (mainly diffraction-contrast) images that constitute traditional TEM studies of materials. We'll give you plenty of examples of such inelastic images in Part 4.

4.6 BEAM DAMAGE

The inelastic collisions that give us all the useful signals we've just discussed bring with them an unfortunate side effect, electron-beam damage. We are often less precise and call this phenomenon radiation damage. The damage, which affects the structure and/or the chemistry of the specimen, depends, in some form or other, on the incident-beam energy. Certain materials are more susceptible than others but, in the end, you can damage virtually anything that you put into the TEM, particularly now that aberration correction permits even more electron current to be focused into even smaller beams. Therefore, damage represents a real physical limit on what the TEM can do and may be regarded as the microscopists' analog of the Heisenberg uncertainty principle in that the very act of observing your specimen can change it.

DAMAGE

Once its structure or chemistry is changed, your thin specimen is not representative of its parent material and interpreting any of your TEM images, DPs or spectra becomes more difficult.

On the other hand, we can sometimes use beam damage to aid certain in-situ transformations that are speeded up by the damage process or we can use electron damage to emulate other forms of radiation damage. Generally, however, beam damage must be considered undesirable.

Damage takes one of three principal forms

- *Radiolysis:* Inelastic scattering (mainly electron-electron interactions such as ionization) breaks the chemical bonds of certain materials such as polymers and alkali halides.

- *Knock-on damage or sputtering:* Knock-on damage is the displacement of atoms from the crystal lattice and creates point defects. If atoms are ejected from the specimen surface we call it sputtering. These processes are ubiquitous if the beam energy (E_0) is high enough.
- *Heating:* Phonons heat your specimen and heat is a major source of damage to polymers and biological tissue.

We will see that, paradoxically, radiolysis is reduced at higher E_0 while knock-on damage is increased; so there is sometimes no way around the damage problem; you can get it at all energies. Cooling your specimen can obviously help if it is likely to be damaged by heating.

All these processes occur in the voltage range available in commercial TEMs and so you must be aware of the dangers. The actual processes can be very complicated and are also specimen-specific, so we could get bogged down in an enormous amount of detail. What we'll do, however, is describe the fundamental processes in different materials, explain how you can determine if your specimen is being damaged and describe how you can minimize or eliminate the problem.

If you find you need to know more about radiation damage, the text by Jenkins and Kirk is the place to start. But we should also note upfront that whole areas of TEM such as in-situ studies and environmental TEM actually take advantage of radiation damage to enhance particular reactions.

We'll start our brief overview of damage in different kinds of TEM specimens by explaining the terms we use to measure damage.

4.6.A Electron Dose

In the TEM we define the electron dose as the charge density (C/m^2) hitting the specimen. It is easy to convert this to the number of electrons/unit area (usually e/nm^2) knowing that $e = 1.6 \times 10^{-19}$ C. This term is *not* the same as for radiation effects on the human body, for which we define dose as the energy absorbed per unit volume. This human dose is defined by the Gray (Gy) which is the absorption of 1 J of ionizing radiation/kg of material and 1 Gy = 100 rads (in pre-SI units). If we convert the incident electron dose to an absorbed dose it can easily be shown that typical electron exposures inside the TEM are well above the lethal limit for human tissue and early microscopists occasionally found this out to their cost (although we hasten to add that, to our knowledge, while no deaths occurred, digits were apparently lost). While this fact is another warning about the dangers inherent in TEM it is more pertinent as a reminder to you that we put an enormous amount of energy into our specimens. This latter point is well illustrated if you calculate the total power input into the specimen, as we do in the next chapter. Fortunately, such a small fraction of the beam energy is transferred to a thin specimen that most specimens survive this, otherwise hostile, environment.

4.6.B Specimen Heating

Specimen heating is difficult to measure experimentally because of the many variables that can affect the result, such as the thermal conductivity, thickness, and surface condition of the specimen as well as the beam size, energy, and current. Hobbs has calculated the effects of beam current and thermal conductivity on the specimen temperature, as shown in Figure 4.11. From these results we can say that, as a rule for metals and other good conductors, beam heating is negligible under standard TEM conditions but, for insulators, it can be quite substantial. To minimize heating, follow the instructions given at the end of the next section.

You will often hear questions about beam heating from people who don't use TEM.

- If thermal conduction is very high, heating is negligible.
- If thermal conduction is poor, heating can be quite substantial.

So, beam heating for metals is usually minimal but small ceramic particles may be heated by the beam to temperatures of ~1700°C. If a good thermal conductor is thermally insulated from its surroundings, considerable

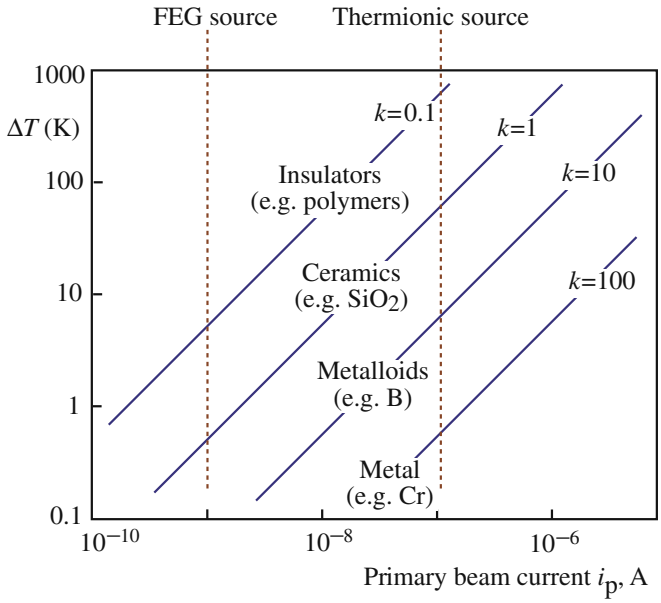


FIGURE 4.11. The increase in specimen temperature as a function of the beam current and the thermal conductivity of the specimen (k, in W/m K). Typical materials are noted, but should not be considered representative, since k varies substantially in any class of materials.

beam heating can occur. Most electrons go through a thin specimen and lose very little energy, as we'll see in our discussion of EELS.

4.6.C Beam Damage in Polymers

Polymers are particularly sensitive to the electron-electron interactions since these can break chemical bonds creating new structures; we call this process radiolysis.

- Electrons can cause the main polymer chain to break, thus changing its basic structure.
- Electrons can cause side groups to break off, leaving reactive free radicals which may crosslink to form a new structure.

Breaking a polymer chain this way is known as scission. Generally, polymers show a tendency either to break down or to crosslink under electron irradiation. In the former case, the polymer will continue to lose mass while, in the latter, the polymer eventually becomes mainly carbon. Mass loss can sometimes be measured directly by EELS in the TEM and it can also manifest itself as a major dimensional change in your specimen because, ultimately, a hole appears in the damaged area; if you're watching carefully you'll see that the image contrast will usually change before the hole appears!

If your polymer specimen was originally crystalline, then radiation damage results in a loss of crystallinity, which you can measure quantitatively from the loss of diffraction contrast in the image or the loss of sharp peaks in the DP (which are gradually replaced with diffuse scattered intensity characteristic of amorphous structures (go back and look at Figure 2.13A). Sometimes you can preserve the crystal structure by staining with a heavy metal such as Pb or U. However, whenever you stain your specimen you affect its structure and change the chemistry, so this isn't ideal.

There are several methods you can use to minimize beam damage in polymers

- Use low-dose imaging techniques (see Chapter 31).
- Cool the specimen to liquid N_2 temperatures or lower T, if possible.
- Coat the specimen with a conducting metal film.
- Use STEM imaging (Section 22.3).
- Do all of the above, if necessary.

In addition to these practical steps, any contribution to damage from beam heating is generally minimized by reducing the cross section for inelastic scatter, i.e., by using the highest available voltage. So HVEMs are better for the study of heat-sensitive materials. If the specimen is thinner than the mean free path for inelastic interaction then less energy is transferred to the specimen and the result is less damage due to heating effects.

4.6.D Beam Damage in Covalent and Ionic Crystals

In covalent and ionic materials such as ceramics and minerals, radiolysis can change the specimen chemistry and possibly its structure through a series of reactions driven by the electron beam. The inelastic interaction primarily responsible for radiolysis is the interband transition, similar to that which causes CL. The transition of a mobile, valence-band electron to the conduction band leaves a hole in the original energy level. Rather than emitting a photon, the electrons and holes may partially recombine via an intermediate, metastable state called an exciton which, through a rather complicated sequence of events, can create an anion vacancy and a cation interstitial. Crystalline quartz (although a very hard material) can be amorphized by a similar process. Often radiolysis can result in the formation of new compounds, which can be identified *in situ* as they form, by electron diffraction and AEM. The formation of Ag from Ag halides in a photographic emulsion is a (rare) example of useful radiolysis. So, somewhat paradoxically, if we use photographic film we depend on radiation damage to record the information we generate in our TEMs (more about this in Chapter 7).

We can't stop radiolysis simply by cooling or coating our specimen, since it isn't affected by heat-transfer considerations. The best way is to lower the cross section for the initiating electron-electron interactions and we can do this by using higher voltages and thinner specimens (thinner is better again). Nevertheless, radiolysis remains a major limitation in the TEM when looking at certain ceramics and minerals and most polymers.

4.6.E Beam Damage in Metals

The primary way that metals are damaged is by knock-on or displacement/sputtering damage. This process occurs by the direct transfer of the beam energy to atoms in the solid, knocking them out of their atomic site and creating a combination vacancy and interstitial (or Frenkel pair). For an atom to be kicked out of its very comfortable and stable crystal lattice site, the beam electron has to penetrate close to the nucleus and be effectively stopped in its tracks by the Coulombic attraction, thus transferring most, if not all, of its energy to the atom.

KNOCK-ON DAMAGE
It is directly related to the incident beam energy.

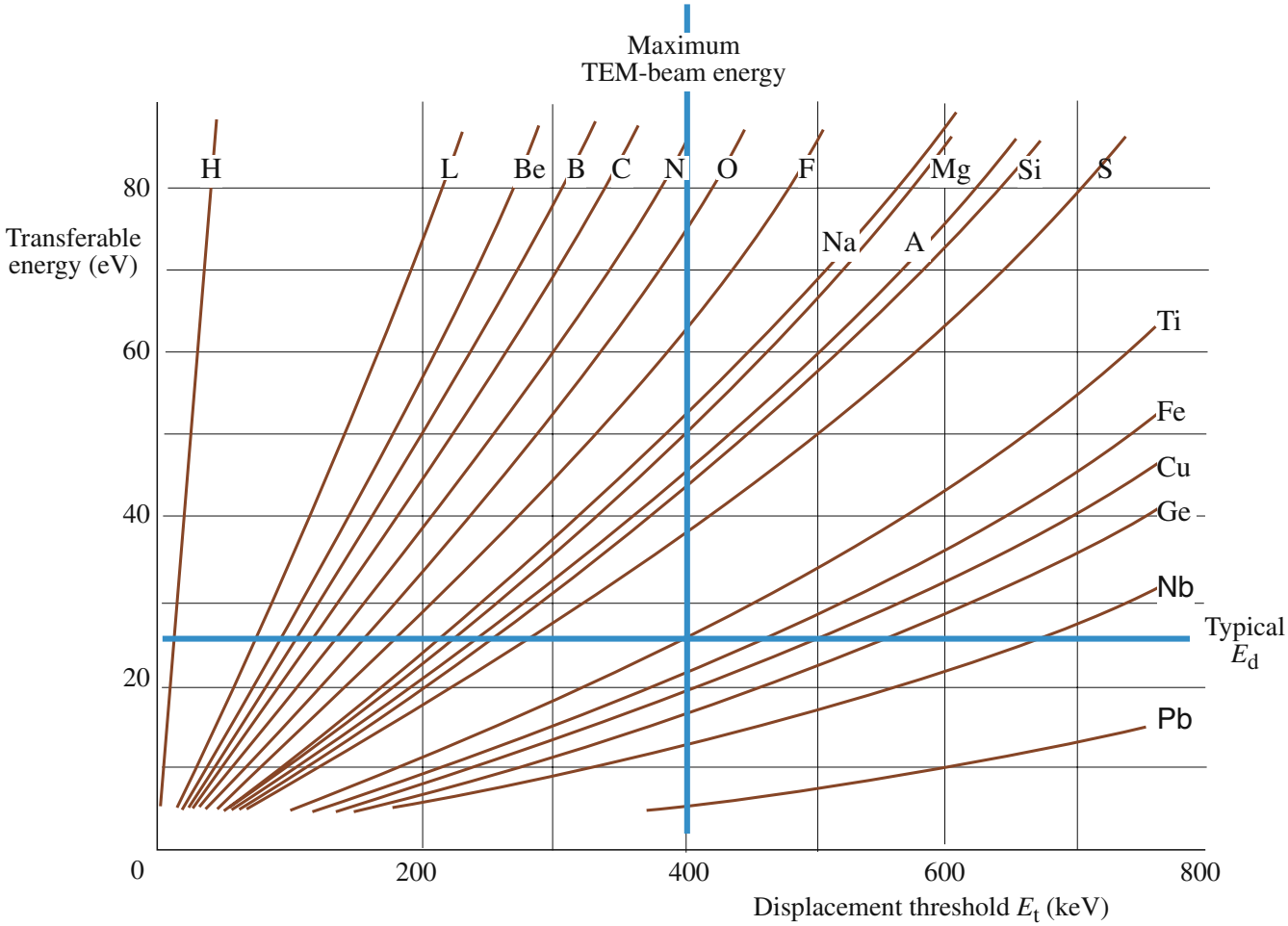


FIGURE 4.12. The maximum transferable energy for a range of atoms as a function of the displacement-threshold energy. A maximum IVEM beam energy is indicated (400 keV) and a typical E_d is shown as ~25 eV, but it can vary substantially with bond strength in different materials. In regions of the graph above 400 keV and below E_d, damage will not occur.

How strongly the atoms are bonded to their neighbors will also be a factor. A simple expression given by Hobbs for the displacement energy E_d allows us to determine the threshold energy (E_t) for displacement of atoms of atomic weight A

$$E_t = \frac{\left(\frac{100 + AE_d}{5}\right)^{1/2} - 10}{20} \tag{4.9}$$

where E_t is in MeV and E_d is in eV. E_d is typically in the range from 5 to 50 eV, but varies with bonding type. If you can transfer more than the threshold energy to an atom then you will displace it from its site. This concept is summarized in Figure 4.12. So, for example, at 400 keV, $>$ 80 eV can be transferred to a carbon atom, about 45 eV can be transferred to an Al atom, and about 25 eV to a Ti atom. If we assume an average displacement energy of ~ 25 eV then it is quite evident that if you have a 400-kV intermediate voltage TEM, you can displace any atom with an atomic weight below about Ti (unless it happens to be bonded so that E_d is much greater than the typical value in Figure 4.12). If you're using an HVEM with beam energies of 1 MeV or more you will *invariably* cause displacement damage, except perhaps in the heaviest elements. The only way to avoid displacement damage is to operate below threshold and you should determine this energy experimentally given the very approximate nature of Figure 4.12.

How can you identify displacement damage? The only sure way is to record images of the same area before and after radiation and compare the contrast under the same imaging conditions. Knock-on damage often manifests itself as small vacancy clusters which appear as black-white lobe contrast or dot contrast, as we showed back in Figure 1.8. Sometimes damage is discernible as dislocation loops and stacking fault tetrahedra caused by the agglomeration of vacancies. Such crystal defects could of course be easily confused with defects that are intrinsic to the material rather than introduced by the act of observing the materials in the TEM. Displacement damage can also occur in polymers

and minerals, of course. The problem here is that we just suggested going to higher voltages as one way of minimizing thermal effects and radiolysis and in doing so, we may well induce knock-on damage. So depending on your material, there may in fact be no way to avoid damaging your specimen in one way or another except, perhaps, by becoming a metallurgist.

Perhaps the only bright side to displacement damage is that we can study it for its own sake. It can be argued, though by no means conclusively, that electron-beam damage in materials can be equivalent to neutron damage, such as that occurring in nuclear reactors. A general rule of thumb is that a few minutes' exposure in an HVEM is equivalent in terms of damage to many years in a nuclear reactor and so accelerated studies of materials degradation in reactor environments were possible. With this justification, an enormous amount of work was carried out in the 1960s when nuclear power was in vogue. Three Mile Island and Chernobyl seriously reduced the number of such studies but there are many reviews extant in the older literature and an occasional more modern reference. Given that the current political climate (!) is more favorable toward nuclear energy than for some decades, it is not unreasonable to suspect that beam damage might see a resurgence in importance. So, if you haven't settled on a career yet, you may wish to spend a little time studying this field because all known radiation-damage microscopists have retired or are close, and the hard-earned knowledge from the early halcyon years is in danger of being forgotten!

Vacancies caused by displacement damage can enhance diffusion processes which, in turn, can speed diffusional transformations when they're being studied in situ in the IVEM/HVEM. There are many other problems that can arise during in-situ observations, so interpretation isn't always straightforward. The book by Butler and Hale is recommended for more facts, practical hints, and many beautiful images of reactions occurring, in real time, in thin foils. More recently, Gai has edited a volume of contributed chapters on in-situ TEM and there is much excitement in the field about the possibilities for growth of this important area due (not surprisingly) to the advent of C_s correction bringing about TEM stages with larger pole-piece gaps, thus permitting easier insertion of gas-reaction stages.

4.6.F Sputtering

The displacement of surface atoms, or sputtering, occurs in the TEM at voltages which are $<0.5E_t$. If your specimen is quite thick then this problem is minor since the average, through-thickness, characteristics of the specimen are not changed significantly by any changes on the surface. But, as we've already noted many times, your specimen should really be very thin (thinner is better yet again) if you want to get the best images and the best analytical information. In these specimens, modifications of the surface structure and chemistry may be sufficient to affect the image interpretation and/or change the average through-thickness composition enough to affect the accuracy of any quantitative analysis. Table 4.3 lists typical sputtering threshold energies (E_s) compared with displacement thresholds (E_d) and their relative values compared to how much energy can be transferred by electrons of different energies. As you can see, there is cause for concern even with 100-keV beams. A STEM can easily drill a hole through MgO.

TABLE 4.3 Comparison of Maximum Transferable Kinetic Energy (T) at 100, 200, 300 and 400 keV with Displacement (E_d) and Sputtering (E_s) Energies

Element	T (eV)				E_d (eV)	E_s (eV)
	100 keV	200 keV	300 keV	400 keV		
Al	8.93	19.5	31.6	45.3	16	4–8
Ti	5.00	11.0	17.8	25.5	15	4–8
V	4.73	10.3	16.72	24.0	29	7–14
Cr	4.63	10.1	16.38	23.5	22	5–11
Fe	4.31	9.40	15.25	21.8	16	4–8
Co	4.08	8.91	14.45	20.7	23	5–12
Ni	4.10	8.94	14.5	20.8	22	6–11
Cu	3.79	8.26	13.4	19.2	18	4–9
Zn	3.69	8.03	13.03	18.7	16	4–8
Nb	2.59	5.65	9.17	13.2	24	6–12
Mo	2.51	5.47	8.88	12.7	27	7–14
Ag	2.23	4.87	7.90	11.3	28	7–14
Cd	2.14	4.67	7.58	10.9	20	5–10
Ta	1.33	2.90	4.71	6.75	33	8–16
Pl	1.23	2.69	4.37	6.26	33	8–16
Au	1.22	2.67	4.32	6.2	36	9–18

CHAPTER SUMMARY

Inelastic scattering transfers energy to your specimen, generating a lot of useful signals with which we can form different images or get spectroscopic information about the chemistry and electronic structure of the specimen. Much of this information is localized at the nanometer level or below.

- The generation of many different characteristic signals all localized with nanometer-scale (or below) spatial resolution is a most powerful aspect of TEM.

- Unfortunately, these same inelastic processes create beam damage and heat which can be disastrous under certain conditions for all kinds of TEM specimens. To minimize heat transfer, cool your specimen and use higher voltages and thinner specimens but beware:
- If the accelerating voltage is high enough, knock-on and sputtering damage will occur in all materials, creating non-characteristic crystal defects and changing the surface chemistry.

On the brighter side, beam damage can be a positive help in the simulation of nuclear-radiation effects and can also enhance in-situ transformation studies and environmental TEM.

In addition to the following references, many more will be found in the companion text.

IN-SITU

Butler, EP and Hale, KF 1981 *Dynamic Experiments in the Electron Microscope* Practical Methods in Electron Microscopy **9** Ed. AM Glauert Elsevier Amsterdam. Radiation damage plays a major role in many of the topics discussed.

Gai, PL Ed. 1997 *In-Situ Microscopy in Materials Research* Springer-Verlag New York.

Inokuti, M 1971 *Inelastic Collisions of Fast Charged Particles with Atoms and Molecules—The Bethe Theory Revisited* Rev. Mod. Phys. **43** 297–347. A classic and most comprehensive review of progress over the 40 years following Bethe's first cross section description.

Wang, ZL 1995 *Elastic and Inelastic Scattering in Electron Diffraction and Imaging* Plenum Press New York.

BEARDEN'S TABLES AND MORE

Bearden, JA 1964 NYO-10586 US Atomic Energy Commission Oak Ridge TN.

Deslattes, RD, Kessler, RD Jr, Indelicato, P, de Billy, L, Lindroth, E and Anton, J 2003 *X-ray Transition Energies: New Approach to a Comprehensive Evaluation* Rev. Mod. Phys. **75** 35–99.

SOME HISTORY

Auger, MP 1925 *Sur L'Effet Photoélectrique Composé* J. Phys Radium **6** 205–208.

Bethe, HA 1930 *Zur Theorie des Durchgangs Schneller Korpuskularstrahlen Durch Materie* Ann. der Phys. Leipzig **5** 325–400.

Kramers, HA 1923 *On the Theory of X-ray Absorption and Continuous X-ray Spectrum* Phil. Mag. **46** 836.

Moseley, HGJ 1914 *High Frequency Spectra of the Elements* Phil. Mag. **26** 1024–1032.

Williams, EJ 1933 *Applications of the Method of Impact Parameter in Collisions* Proc. Roy. Soc. London **A139** 163–86.

SPECIMEN DAMAGE

Egerton, RF, Li, P and Malac, M 2004 *Radiation Damage in the TEM and SEM* Micron **35** 399–409. More recent paper on radiation damage.

Hobbs, LW 1979 in *Introduction to Analytical Electron Microscopy* Eds. JJ Hren, JI Goldstein and DC Joy Plenum Press New York p437 gives calculations of beam heating.

Jenkins, ML and Kirk, MA 2000 *Characterization of Radiation Damage by Transmission Electron Microscopy* Institute of Physics Bristol and Philadelphia. The only text on this topic; helpful reading if specimen damage affects your work. Includes discussion of knock-on damage.

Sawyer, LC, Grubb, DT and Meyers, GF 2008 *Polymer Microscopy* 3rd Ed. Springer New York. For polymers, of course.

SPECIAL TECHNIQUES

Boyall, NM, Durose, K and Watson IM 2003 *A Method of Normalizing Cathodoluminescence Images of Electron Transparent Foils for Thickness Contrast Applied to InGaN Quantum Wells* J. Microsc. **209** 41–46.

Goldstein, JI, Williams DB and Cliff, G 1986 *Quantitative X-ray Analysis* in *Introduction to Analytical Electron Microscopy* Eds. JJ Hren, JI Goldstein and DC Joy p 155 Plenum Press New York.

Hubbell, JH, Trehan, PN, Singh, N, Chand, B, Mehta, D, Garg, ML, Garg, RR, Singh, S and Puri, S 1994 *A Review, Bibliography and Tabulation of K, L and Higher Atomic Shell X-ray Fluorescence Yields.* J. Phys. Chem. Ref. Data **23** 339–364.

Markowicz, AM and van Grieken, RE 2002 *Handbook of X-ray Spectrometry* Marcel Dekker New York. Excellent text for when you really need more detail on topics such as families of X-rays.

Newbury, DE 1986 in *Introduction to Analytical Electron Microscopy* Eds JJ Hren, JI Goldstein and DC Joy p 6 Plenum Press New York.

Venables, JA, Hembree, GG, Drucker, J, Crozier PA and Scheinfein MR 2005 *The MIDAS Project at ASU: John Cowley's Vision and Practical Results* J. Electr. Microsc. **54** 151–162. Rare example of attaching an Auger system to a STEM.

URLs

1. www.physics.nist.gov/PhysRefData/Ionization/Xsection.html
2. www.physics.nist.gov/PhysRefData/XrayTrans/index.html

SELF-ASSESSMENT QUESTIONS

Q4.1 Distinguish background, continuum and bremsstrahlung X-rays.
Q4.2 'Characteristic' X-rays are characteristic of what?
Q4.3 Why shouldn't we refer to an 'ionized' electron?
Q4.4 Approximately how large is an electron shell relative to an electron?
Q4.5 What is the critical ionization energy?
Q4.6 What is the ionization cross section?
Q4.7 What controls the energy and wavelength of the characteristic X-ray?
Q4.8 What is the fluorescence yield?
Q4.9 What do we mean by the term 'weight of an X-ray line'?
Q4.10 What is overvoltage?
Q4.11 What is the difference in the angular distribution of the ionizing electrons and the emitted X-rays?
Q4.12 Does an X-ray have kinetic energy like an electron?
Q4.13 Why is all the energy transferred to the atom during ionization (E_c) not recovered by the emission of the characteristic X-ray(s)?
Q4.14 How is energy conserved in the overall electron-atom interaction?
Q4.15 Why should you be concerned about fast secondary electrons in TEM?
Q4.16 What is cathodoluminescence?
Q4.17 How can you minimize electron-beam damage to your specimen?
Q4.18 Why does sputtering of atoms from the surface of the specimen take less energy than displacing atoms in the interior?
Q4.19 What is radiolysis?
Q4.20 What's the best reason for using an HVEM to intentionally displace atoms in a specimen?

TEXT-SPECIFIC QUESTIONS

T4.1 Look at Figure 4.1 and explain why
   A. We need to be concerned about plasmon energy losses whenever inelastic scatter compromises the operation of the TEM.
   B. All scattering processes show approximately the same dependence on beam energy.
T4.2 What does Figure 4.1 tell us about the advantages and disadvantages of operating at 400 kV rather than 100 kV?
T4.3 From Figure 4.2
   A. Explain what kind of characteristic X-ray has actually been emitted. (Hint: go to Figure 4.3.)
   B. What happens to the hole left by the electron falling out of the L_3 energy level?
   C. Under what conditions is the presence of the vacuum energy level expected?
T4.4 If, as stated in the text, K-shell X-rays have higher energy than L-shell X-rays and so on, why does the Ag L line (2.3keV) have a higher energy than the Al K line (1.5keV)?
T4.5 Use equation 4.6 to determine the relative fluorescence yields of (a) Be and N and (b) Si and Ag. Cross-check your answer with Figure 4.5. Use your answers to explain why X-ray analysis is challenging at lower atomic numbers and quantification is relatively straightforward for higher atomic numbers. Do Auger electrons have an equivalent 'yield'? If so, derive the approximate dependence on Z.
T4.6 Can you explain why we talk about 'braking' radiation when the nucleus and the electron have opposite charges that one might expect to attract one another very strongly?
T4.7 Looking at Figure 4.6, the generated bremsstrahlung intensity increases for lower X-ray energies, but we have just learned that the characteristic X-ray intensity decreases at lower energies (from lower atomic number elements). Explain this difference.
T4.8 Why are the conduction and valence bands shown separately in Figure 4.8 but not in Figure 4.2?
T4.9 Calculate the electron dose in a 1-nm electron probe containing 1 nA of current.
T4.10 Using Equation 4.9, compare the displacement threshold energies for Li, Al, Cu and Au and estimate the sputtering energies. Compare your answers where possible with the data in Table 4.3 and Figure 4.12.

T4.11 Inelastic scattering ultimately limits the information we can obtain from the TEM because the transfer of energy to the specimen can change its structure or chemistry (the two characteristics that the TEM is best at discerning) either temporarily or permanently. So, why do we take great pains to both maximize the generation and detection of inelastic scattering?

T4.12 Relate the amount of energy lost by the electron to the typical potential-well model of the atom.

T4.13 If you look at Table 4.1 you'll see that the cross sections are smaller for interactions that result in larger energy transfer. Why is this a good thing for TEM operators?

T4.14 Why don't we routinely try to use Auger electrons for surface analysis in the TEM?

T4.15 Based on Figures 4.1, 4.4 and 4.5 explain the most likely signal(s) generated when (a) a 100-kV electron beam strikes specimens of Be and Si and (b) when a 300-kV beam does the same.

T4.16 Compare and contrast plasmons and phonons. (Hint: consider their generation processes, energies, scattering angles, cross sections, range, effects on the specimen.) From your answer discuss their relative importance to the TEM user in terms of the information they contain or destroy.

T4.17 (Extra challenge) If the electron beam sputters Au from the surfaces of a thin film, which surface is affected most?

5

Electron Sources

CHAPTER PREVIEW

A reliable source of electrons to 'illuminate' the specimen is one of the most important parts of a TEM. Fortunately, electron sources are plentiful, but to get the best images and other signals out of our expensive microscope, we need to use the best available source. There are stringent requirements to produce the beam of electrons with the necessary properties and these are best met by only two types of source: thermionic and field-emission sources (or 'guns' as they are often called). Thermionic sources are (now rarely) tungsten filaments or (now commonly) lanthanum hexaboride (LaB_6) crystals, and field emitters are fine tungsten needles. In this chapter we'll first explain briefly the physics of the different electron-emission processes because then you'll understand why we operate the sources in certain ways. Next we'll tell you the characteristics we need from our electron beam. Then we'll compare the sources and show you that no one source is best for all aspects of TEM, but all have their roles. Finally, we'll explain ways to check that a particular source meets your needs.

Because the source is so critical to the performance of the microscope, the technology is advancing rapidly with the aim being to have complete computer control, which would leave you, the operator, with very little to do except push the 'on' button. This state of affairs is most advanced for the field-emission source, and since these are both delicate and expensive, it is just as well. But the majority of TEMs still use thermionic sources, and these may need a fair bit of operator control. In these circumstances, you should know how these sources work and why you do certain things to them. So we'll spend much of this chapter talking about thermionic sources, although field emission is essential for the best performing TEMs of any kind (imaging, analytical, etc.), so there's a good chance that field emission will expand to be the source of choice in the future.

5.1 THE PHYSICS OF DIFFERENT ELECTRON SOURCES

We use two kinds of electron sources in TEMs: the first kind is called a thermionic source, which, as the name suggests, produces electrons when heated, and the second type is a field-emission source, which produces electrons when a large electric potential is applied between it and an anode. Schottky sources combine both heat and field emissions. These sources are part of an assembly which we refer to as the electron 'gun.' Now, from a physics standpoint, it is really quite interesting to know the details of how electron sources work and there's a great deal of active research into new and improved sources, and carbon nanotubes show some real promise. (Like future TV screens but using higher voltages.)

From a practical standpoint, you don't have to know too much about the physics, and we can summarize the essential points very briefly, using two simple equations. Keep in mind two points as you read about sources

- Your TEM will use a thermionic source (W or LaB_6) or a field-emission (W) source and the two cannot be interchanged.
- Field-emission sources give more monochromatic electrons; thermionic sources are less monochromatic and give 'whiter' electrons.

The analogy here is to X-rays or visible light. The 'color' of electrons depends on their energy spread (which translates into a frequency or a wavelength range using equation 1.6); we'll discuss this in Section 5.2. You don't always need to use monochromatic electrons, even if your field-emission TEM did cost twice as much as a conventional microscope would with a thermionic source.

5.1.A Thermionic Emission

If we heat any material to a high-enough temperature, we can give the electrons sufficient energy to overcome the natural barrier that prevents them from *leaking out* from the surface. This barrier is termed the work function (Φ) and has a value of a few eV.

The physics of thermionic emission can be summarized in Richardson's law which relates the current density from the source, J, to the operating temperature, T in Kelvin

$$J = AT^2 e^{-\frac{\Phi}{kT}} \tag{5.1}$$

Here k is Boltzmann's constant (8.6×10^{-5} eV/K) and A is Richardson's constant ($A/m^2 K^2$), which is only constant for a given source material. From this equation you can see that we need to heat the source to a temperature T such that energy $> \Phi$ is given to the electrons; then they will escape from the source and be available to form an electron beam. Unfortunately, when we put a few eV of thermal energy into most materials they either melt or vaporize. So the only viable thermionic sources are either refractory (high melting point) materials or those with an exceptionally low Φ. The source used for the first several decades of TEMs (and still used in some SEMs) was tungsten which melts at 3660 K and the only thermionic source used by modern TEMs is lanthanum hexaboride (LaB_6) which has a low Φ. If you look at Table 5.1, you'll see the relative values of J_c, T, and Φ for tungsten and LaB_6.

We use several different words to describe the sources. We called tungsten sources *filaments*, because tungsten was drawn into fine wire similar to the filament used in an incandescent light bulb. LaB_6 crystals (which should not be called filaments) are usually grown with a <110> orientation to enhance emission. Sometimes we call both tungsten and LaB_6 sources *cathodes* because, as we'll see, the complete gun assembly acts as a triode system in which the source is the cathode.

So all you need to know from the physics is that heating up a thermionic source gives you a higher J. But there is a limit, because higher temperatures shorten the source life through evaporation and/or oxidation. So we seek a compromise operating temperature. We thus operate under a condition called 'saturation' which we'll discuss in 'Thermionic Guns.'

5.1.B Field Emission

Field-emission sources, usually called FEGs (for field-emission guns: pronounced either as 'F-E-Gs' or as 'fegs') operate in a fundamentally different way to thermionic sources. The principle behind FE is that the strength of an electric field E is considerably increased at sharp points because, if we have a voltage V applied to a (spherical) point of radius r, then

$$E = \frac{V}{r} \tag{5.2}$$

We call the fine needles 'tips.' The technique of atom-probe field-ion microscopy (APFIM) is another well-established experimental tool for materials characterization. APFIM uses *specimens* with a very fine needle shape, and so there's a lot of expertise available to help produce FE tips. One of the easiest materials to produce with a fine needle point is tungsten wire which can readily be given a tip radius of < 0.1 μm. If we apply a 1-kV potential to this tip then E is 10^{10} V/m and this lowers the work-function barrier sufficiently for electrons to tunnel out of the tungsten. The tunneling process is the same as you've met in semiconductor devices. Applying such high fields imposes a severe stress on the tip and the material must be mechanically strong to remain intact. FE, like thermionic emission from LaB_6, varies with the orientation of the tungsten crystalline tip; the <310> orientation is found to be best.

For FE to occur, the surface has to be pristine, i.e., it must be free of contaminants and oxide. We can achieve this by operating in ultra-high vacuum (UHV) conditions ($<10^{-9}$ Pa). In this case the tungsten is operated at ambient temperatures and the process is called 'cold' FE. Alternatively, we can keep the surface in a pristine condition at a poorer vacuum by heating the tip. The thermal energy assists in electron emission

TABLE 5.1 Characteristics of the Principal Electron Sources

	Units	Tungsten	LaB_6	Schottky FEG	Cold FEG
Work function, Φ	eV	4.5	2.4	3.0	4.5
Richardson's constant	A/m^2K^2	6×10^9	4×10^9		
Operating temperature	K	2700	1700	1700	300
Current density (at 100 kV)	A/m^2	5	10^2	10^5	10^6
Crossover size	nm	$> 10^5$	10^4	15	3
Brightness (at 100 kV)	A/m^2sr	10^{10}	5×10^{11}	5×10^{12}	10^{13}
Energy spread (at 100 kV)	eV	3	1.5	0.7	0.3
Emission current stability	%/hr	<1	<1	<1	5
Vacuum	Pa	10^{-2}	10^{-4}	10^{-6}	10^{-9}
Lifetime	hr	100	1000	>5000	>5000

so much that, in fact, the electrons don't tunnel through the barrier. For such 'thermal' FE, surface treatments with ZrO_2 improve the emission characteristics, particularly the stability of the source, and such Schottky emitters are the most popular. There are pros and cons for both CFE and TFE, which we'll talk about later in the chapter.

5.2 THE CHARACTERISTICS OF THE ELECTRON BEAM

The electron beam in a TEM requires certain characteristics which are controlled by the source itself and how we integrate the source into a gun assembly. We describe the performance of an electron source by such words as brightness, coherency, and stability. While these words mean something to you already, they have very precise meanings in TEM terminology, so we'll go through the various characteristics, tell you what they mean and why they are important in the TEM. We'll then compare the properties of the various sources that you may have in your microscopes. You'll see that there's no best source for all applications, but for specific applications one source or other is usually better.

Before we define the electron-beam characteristics needed in a TEM, it is worth summarizing a few of the properties of electron beams in general and how these vary with kV.

5.2.A Brightness

The word *brightness* is often confused with *intensity* and indeed the two terms are related. For instance, when we look at the viewing screen of a TEM, we may say how 'bright' it is, when we are really referring to the intensity of light coming from the screen. When we think of the intensity of any radiation source, it is in terms of the flux emanating from it. For a light bulb, it would be the number of photons per unit area per unit time. For electron sources we talk about the current density, which is the number of electrons (or charge) per unit area per unit time.

BRIGHTNESS
While current density can be a useful term, it is more important to define the brightness. Brightness is the current density per unit solid angle of the source.

Electron sources differ considerably in their size and, as a result, the electrons leave the source with a range of divergent angles, so we can't ignore the angular distribution of the electrons. Brightness is particularly important when we are using very small convergent probes, as we do in AEM and STEM. The concept of brightness is less important in conventional TEM where we use a relatively large, defocused beam, but it is still relevant to the intensity we see on the screen, and so it affects how easy it is to operate the microscope and see our images and DPs.

So we can consider an electron source as having the following characteristics

- Diameter d_0
- Cathode emission current i_e
- Divergence angle α_0 (remember when we say angle we mean semi-angle)

We'll describe the actual way in which these characteristics are achieved in Section 5.3, where we discuss the complete gun assembly, but if you look at Figure 5.1 you'll see that i_e, d_0, and α_0 are actually defined at the gun crossover, that is, the point at which the electrons are focused *after* leaving the source. The current density

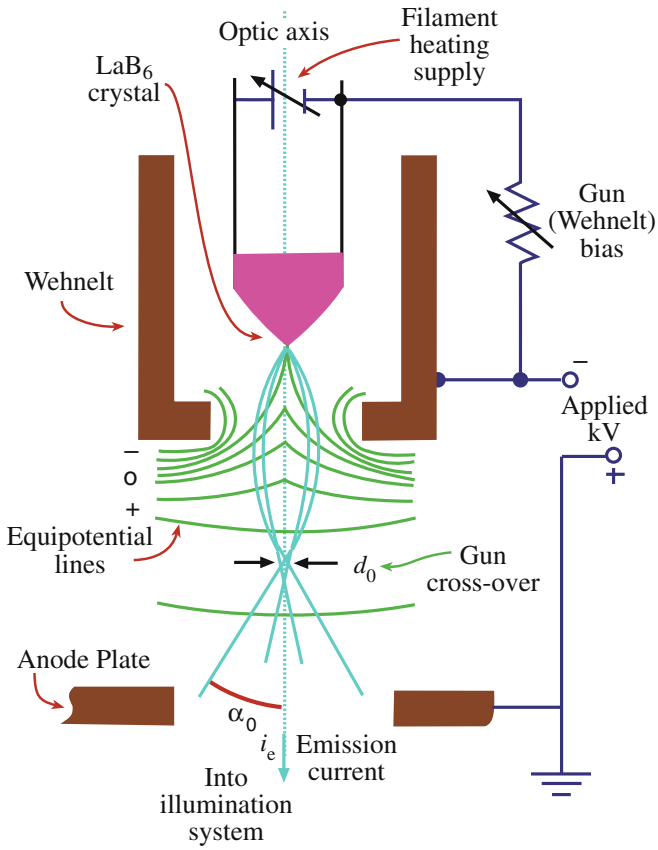


FIGURE 5.1. Schematic diagram of a thermionic electron gun. A high voltage is placed between the cathode and the anode, modified by a potential on the Wehnelt which acts as the grid in a triode system. The electric field from the Wehnelt focuses the electrons into a crossover, diameter d_0 and convergence/divergence angle α_0 which is the true source (object) for the lenses in the TEM illumination system.

(current per unit area) is $i_e/\pi(d_0/2)^2$ and the solid angle of the source is $\pi\alpha^2$, so we define the brightness β as

$$\beta = \frac{i_e}{\pi\left(\frac{d_0}{2}\right)^2 \pi(\alpha_0)^2} = \frac{4i_e}{(\pi d_0 \alpha_0)^2} \tag{5.3}$$

This brightness equation is an important one which you should remember. What is not shown in this equation is the important fact that β increases linearly with increasing accelerating voltage for thermionic sources. This was one reason for the development of intermediate-voltage (300–400 kV) instruments.

Obviously, the higher the value of β, the more electrons we can put into a beam of a given size, and so the more information we can generate from the specimen and also the more we can damage sensitive specimens. The beam current is an important part of the brightness equation. Measuring the beam current in situ can be a very good diagnostic tool. We'll talk about this later in the chapter when we discuss measuring β, but for the time being you can again look at Table 5.1 to see how the various sources compare in brightness.

UNITS FOR BRIGHTNESS

β is in units of $A/m^2\,sr$.

Now we can consider some real numbers. With a cold FEG at 100 keV, we can put 1 nA into an area of diameter 1 nm. If you convert this current density to units of power (1 watt = 1 J/s), you'll find that the energy the electron beam puts into this small area of the specimen is nearly $150\,MW/mm^2$. By comparison, the output of a typical electric power-generating turbine is anywhere from 350 to 1000 MW. If all this energy were in fact absorbed by the TEM specimen the technique would be useless since the specimen would vaporize. We'll find out later why this doesn't happen but clearly we can alter our specimen when we look at it in the TEM, as we discussed in relation to beam damage in the previous chapter. The energy density we just calculated is such that a TEM electron source is the brightest, continuously radiating source known; it is considerably brighter than a supernova.

The brightness is particularly important in AEM, which is the quantitative analysis of the many signals that come from a specimen irradiated by an electron beam, shown back in Figure 1.3. As you'll see in Part 4, we need to put the most beam current into the smallest probe to optimize both spatial resolution and analytical sensitivity. Similarly, as we go to higher magnifications in HRTEM, the intensity of light coming from the viewing screen (see Chapter 7) becomes less because we are viewing only a small fraction of the illuminated area of the specimen. The electron density can be increased by using the brightest source. Then images can be recorded with reasonably short exposure times minimizing image drift and other instabilities. So brighter is better for AEM and HREM.

5.2.B Temporal Coherency and Energy Spread

The coherency of a beam of electrons is a way of defining how well the electron waves are 'in step' with one another. You know that white light is incoherent, because it consists of photons with a range of frequencies (colors), and so to get a coherent beam of electrons we must create one in which all the electrons have the same frequency (i.e., wavelength) just like monochromatic light. We refer to this aspect of coherency as temporal coherency, which is a measure of how similar the wave packets are. If the packets are all identical they have the same coherence length. A definition of the coherence length λ_c is

$$\lambda_c = \frac{v\mathrm{h}}{\Delta E} \tag{5.4}$$

where v is the electron velocity, ΔE is the energy spread of the beam, and h is Planck's constant. This means we must have stable power supplies to the source and a stable high-voltage supply (or high tension, as it is sometimes called for historical reasons) so that all the electrons have a small ΔE, thus giving a well-defined wavelength. We'll show in detail in Section 37.7 how we can add an energy-selecting spectrometer to the gun to choose electrons with an energy-spread of as little as 10 meV. Such monochromators are expensive and cut down the total beam current tremendously but, for certain very specialized applications, they are invaluable. This loss of current can be offset somewhat by C_s correction (even more expense) and instruments with combinations of monochromators and C_s corrector are the most expensive (and rare) TEMs in the world. If you look at Table 5.1 you'll see that, without a monochromator, typical ΔE values for the three sources are in the range 0.3–3 eV (which is still remarkably small compared with a total energy of 100–400 keV). So it isn't really correct to imply as we did at the start of the chapter that thermionic sources give 'white' electrons since ΔE is so very small. From these values of ΔE, if you take care to get the units consistent, you can calculate typical coherence lengths, which turn out to be a few hundred nanometers.

Temporal coherency is important when the energy spread of the electrons that are *incident* on the specimen affects the microscopy. Because we can make such good high-voltage power supplies, the incident electron-energy spread rarely limits any aspect of TEM except the highest energy-resolution EELS (see Chapters 37–40). In other words, for most practical purposes our electron sources are stable enough. However, we'll see that it's a very different matter when we have to consider the electrons that have come *through* the specimen because they may have lost substantial amounts of

energy and that's where energy-filtered TEM technology really comes into its own (see Section 37.6).

5.2.C Spatial Coherency and Source Size

Spatial coherency is related to the size of the source. Perfect spatial coherence would imply that the electrons were all emanating from the same point at the source. So source size governs spatial coherence and smaller sources give better coherency (just as they give higher brightness). The spatial coherence is strictly defined by looking at electron-interference fringes in the equivalent of a Fresnel biprism experiment in light optics, with which you may be familiar. We can define the distance d_c, the effective source size for coherent illumination, to be

$$d_c = \frac{\lambda}{2\alpha} \tag{5.5}$$

where λ is the electron wavelength and α is the angle subtended by the source at the specimen. We can control α by inserting an aperture in the illumination system, as we'll see when we describe the construction of a TEM in Chapter 9. But if this aperture is not limiting then it is the smallest source which subtends the smallest angle, and thus has the highest spatial coherence. Putting reasonable values for 100-keV electrons into equation 5.5 we find that the spatial coherence is at best only about a nanometer. To maximize the coherency, you can choose several approaches

- Make the source size, d_c, smaller, e.g., by using a FE source. This explains why research into using nanotubes as electron sources is ongoing.
- Use a smaller illumination aperture, thus reducing α.
- If your source size is large (e.g., a W hairpin) decrease the accelerating voltage and thus increase λ.

A small electron source subtends a small angle at the specimen, and we can help by using small limiting apertures. Small beams are more spatially coherent than large beams and give better spatial resolution of analysis (see Part 4). The more coherent and parallel the beam, the better the quality of the phase-contrast images (Part 3), the sharper the DPs (Part 2), and the better the diffraction contrast in images of crystalline specimens (Part 3). So that's why spatial coherence is important. The whole concept of coherence is rather more complex than we have described here and an in-depth and rather mathematical description of electron coherency in the TEM is given in the review by Hawkes and in the companion text.

COHERENCY

Spatial coherency is more important practically than temporal coherency; smaller source –> higher β, better spatial coherency, but lower stability.

5.2.D Stability

In addition to the stability of the high-voltage supply to the source, it is also important that the electron current coming from the source is stable. Otherwise the screen intensity would vary, making it difficult for you to take correctly exposed images and also making quantitative analytical measurements impossible. Thermionic sources are generally very stable except when they are first installed or when they are about to fail. Typically, you can expect variations of $< \pm 1\%$/hr in the current and TFE sources are similarly stable. For CFE sources, however, the emission current is not very stable, and electrical feedback circuits are required to maintain stability to $<\pm 5\%$. Stability improves with better UHV conditions in the gun.

To summarize: the important properties of electron sources are their brightness, temporal coherency, energy spread, spatial coherency and stability. A smaller source size gives higher β and better spatial coherency, but less stability.

Now that we know the critical characteristics required of electron sources, let's examine those used in commercial TEMs.

5.3 ELECTRON GUNS

It's no good just having a source. We need to be able to control the electron beam and direct it into the illumination system of the TEM. We do this by incorporating the source into a gun assembly which, in effect, acts as a lens to focus the electrons coming from the source. The design of the gun is different for thermionic sources and FE sources.

5.3.A Thermionic Guns

LaB_6 is the only thermionic source used in modern TEMs so we'll just describe these. The LaB_6 crystal is used as the cathode in a triode gun shown in Figure 5.1. In addition to the cathode, there is a grid called a Wehnelt cylinder and an anode at earth potential with a hole in its center. What these three components look like in practice is shown in Figure 5.2, where they are all separated. The cathode is attached to the high-voltage cable that, in turn, connects to the high-voltage power supply. The LaB_6 crystal is bonded to a metal wire such as rhenium, which is resistively heated to cause thermionic emission.

When the electrons leave the cathode they have a negative potential of whatever accelerating voltage you have chosen (say 100 kV) with respect to the earthed anode. So they accelerate through this potential difference, acquiring an energy of 100 keV and a velocity greater than half the speed of light.

Now to get a controllable beam of electrons through the hole in the anode and into the microscope itself, we apply a small negative bias to the Wehnelt cylinder. The

FIGURE 5.2. The three major parts of a thermionic gun, from top to bottom: the cathode, the Wehnelt cylinder and the anode shown separated. The Wehnelt screws into the cathode support and both are attached to the high-voltage cable which also contains power supplies for heating the cathode and biasing the Wehnelt. The anode sits just below the Wehnelt and the whole assembly sits on the top of the column of lenses that make up the rest of the TEM.

BEAM CURRENT

As the cathode heating current (i_f) increases, T increases until thermionic emission occurs; then an emission current from the cathode i_e can be measured. Sometimes you'll find this current referred to as the beam current. This is misleading; the true beam current is that which enters the specimen after the electrons have left the gun and gone through the illumination system of the TEM.

electrons coming off the cathode see the negative field and are converged to a point called a crossover between the Wehnelt and the anode as shown in Figure 5.1. We could operate the controls for the cathode heating and the Wehnelt bias independently, but the electronic circuitry of the gun is designed so that as the emission current increases, the Wehnelt bias increases; this arrangement is called a self-biasing gun. The result is shown in Figure 5.3, which plots the emission current (i_e) against the current used to heat the cathode (i_f). As you can see, i_e reaches a maximum such that a further increase in i_f doesn't increase the current going into the TEM column. This is the *saturation condition* and all thermionic sources should be operated at or just below saturation. Operating above saturation reduces the source life without any compensating advantage; operating significantly below saturation reduces the current into your specimen, thus reducing the intensity of all the signals coming out of your specimen although on occasions, as we'll see later, undersaturation can be useful.

WEHNELT

The Wehnelt acts as a simple electrostatic lens: the first lens the electrons go through in the TEM.

In addition to optimizing the source life, operating at saturation also optimizes brightness. If you look at Figure 5.1, the crossover is the source size d_0 that we used back in the brightness equation (equation 5.3) and the divergence angle at the crossover is α_0 in that same equation. The current in the crossover is the emission current i_e. Now, as shown in Figure 5.4A, if the Wehnelt bias were too low (diagram i) d_0 would not be very small and if the bias were too high (diagram iii) the cathode-emission current would be suppressed. In either case β would be low. The optimum β is at an intermediate-bias setting (diagram ii), as summarized in Figure 5.4B. You would be right if you thought that the small bias on the

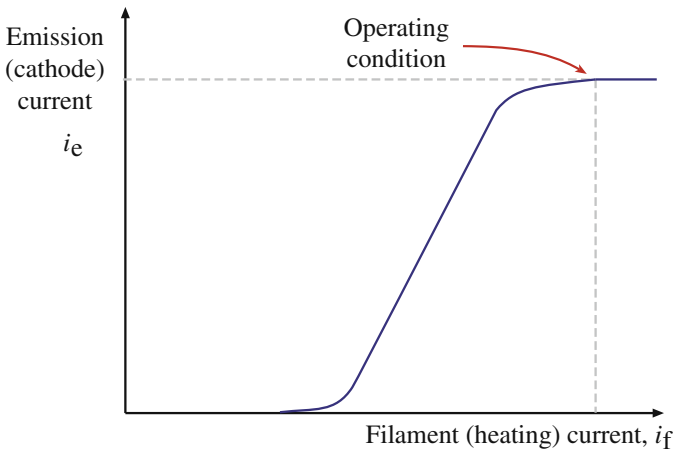


FIGURE 5.3. The relationship between the current emitted by the electron source and the source heating current for a self-biasing gun. Increasing the source current results in a maximum emission current termed saturation.

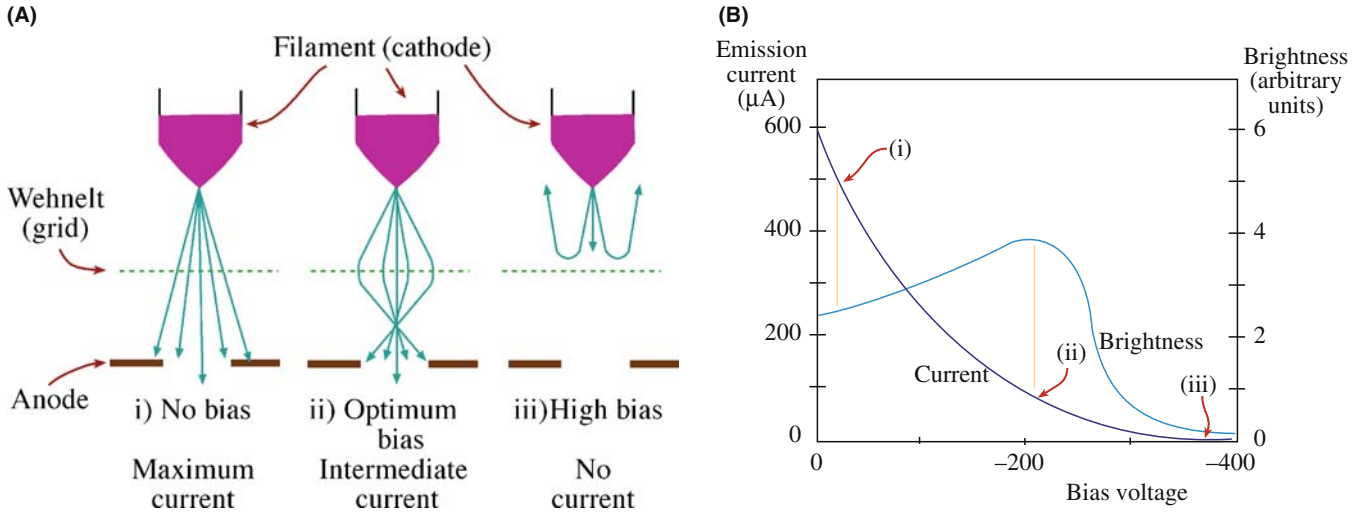


FIGURE 5.4. (A) The effect of increasing Wehnelt bias (i–iii) on the distribution of electrons coming through the anode. (B) The relationship between the bias and the emission current/gun brightness. Maximum brightness (ii) is achieved at an intermediate Wehnelt bias, and an intermediate emission current (ii).

Wehnelt acts against the accelerating voltage, so the true beam voltage should be the applied kV minus the Wehnelt bias (which may be up to 2 kV), but this is compensated for in the design of the gun, so don't worry.

So how do we achieve saturation? One way is to look at the meter which displays i_e and watch it rise to a maximum as i_f is continuously increased. This method may not be easy because the appropriate readouts may not be available or, if they are, they may not be very sensitive. So the standard way is to look at the image of the source crossover on the TEM screen. This image shows you the distribution of electrons coming off the source. As thermionic emission starts, the electrons may come from both the central tip and/or a region surrounding the tip of the crystal. Since LaB_6 sources have well-defined crystal facets (Figure 5.5A) the undersaturated image is as shown in Figure 5.5B. With increasing emission the halo collapses in on the central bright disk, although some structure may still be visible. The cathode is truly saturated when no structure is visible (Figure 5.5C).

It is best to operate an LaB_6 source just below saturation, since this will extend the source life without undue loss of signal. The electrons in the halo are more coherent than those in the central bright region and have reduced energy spread. LaB_6 crystals are susceptible to thermal shock which can cause them to break, and so you should take care when *heating* and *cooling* an LaB_6 source (often the TEM computer takes control of this). If you have to switch the source on and off manually then increase/decrease the heating current slowly, with 10–20 seconds pause between each setting. This is particularly critical after you've installed a new LaB_6 source.

The appearance of the image of the source, as in Figure 5.5, can also be used to align the gun assembly so that the beam is aligned along the optic axis of the TEM. This is the only other adjustment you have to do to the gun, apart from saturating it. The source is usually pre-aligned by the manufacturer, so alignment should be simple when it is put inside the Wehnelt. Typically, a misaligned, undersaturated source image is asymmetrical as in Figure 5.5B and, in those circumstances, all you have to do is adjust the gun to make it symmetrical prior to final saturation. Detailed instructions will be in the manufacturer's handbook. Most modern TEM guns are so well constructed that slight electronic corrections are all that you need to ensure alignment.

SMALL PROBES - FINE BEAMS

Achieving optimum β is critical in any operations that require a fine beam (<0.01 μm).

In an SEM, which always requires a small probe rather than a broad beam, the gun is carefully adjusted by the manufacturer to produce optimum β at saturation; you may not have any external control of the Wehnelt. In a TEM, particularly when you are operating in a broad-beam mode, there is no need to optimize β, but you may need to increase the current density and make the image appear brighter. You can do this by decreasing the Wehnelt bias, using the gun-emission control. When you decrease the bias, you should go back and adjust i_f to ensure you're at saturation, since the saturation condition will change with changing bias. So now you will have a greater current density falling on the screen, but the crossover size will have increased, thus decreasing β. This

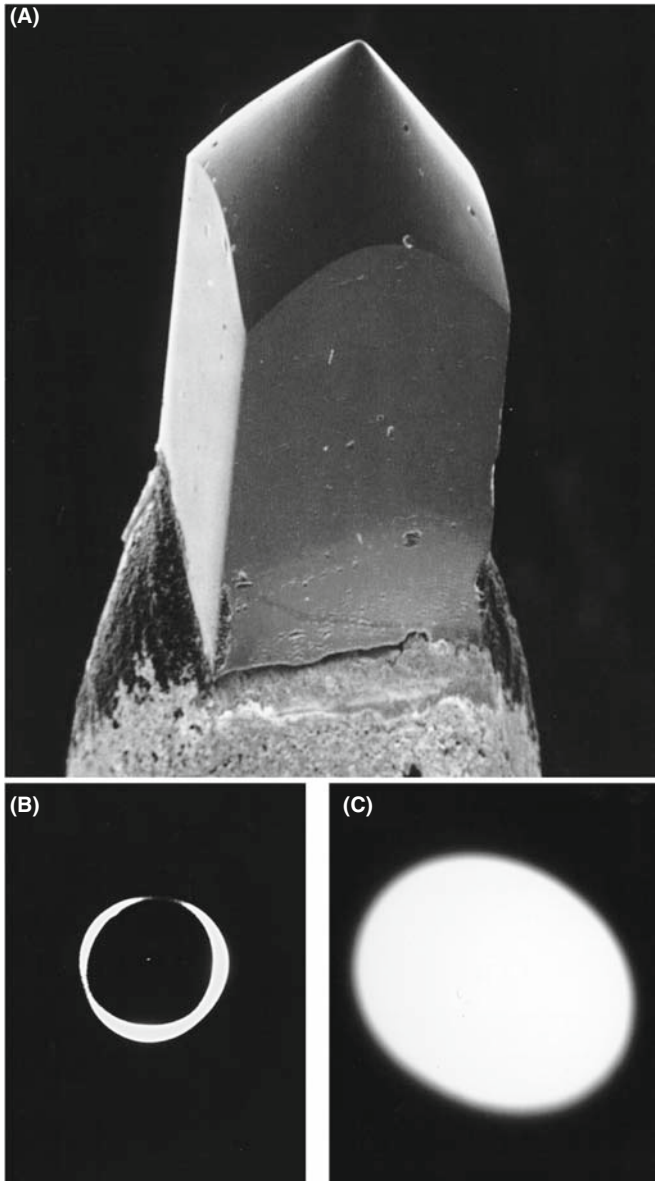


FIGURE 5.5. (A) An LaB_6 crystal and the electron distribution when the source is (B) undersaturated and (C) saturated.

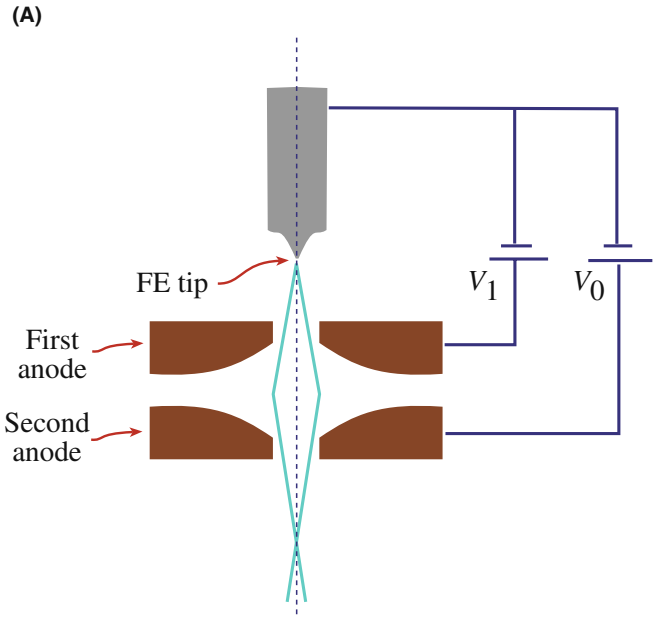
is not important when you're operating with a broad beam. However, if you want to operate at maximum β with a focused beam, as is the case for AEM, then you need to be able to measure β; we'll show you how to do this in Section 5.5.

5.3.B Field-Emission Guns (FEGs)

In many ways, FEGs are much simpler than thermionic guns. In order to get a FEG to work we make it the cathode with respect to *two* anodes. The first anode is positively charged by several kV with respect to the tip. This charge produces the extraction voltage since it generates the intense electric field which enables electrons to tunnel out of the tip. Increasing the extraction voltage when you first switch on the gun has to be done slowly, so the thermo-mechanical shock doesn't fracture the tip. This is the only practical step you have to carry out to run a FEG; it is invariably computer controlled.

- Anode 1 provides the extraction voltage to pull electrons out of the tip.
- Anode 2 accelerates the electrons to 100 kV or more.

The electrons are accelerated through the appropriate voltage by the second anode. The combined fields of the anodes act like a refined electrostatic lens to produce a crossover, as shown in Figure 5.6A. This



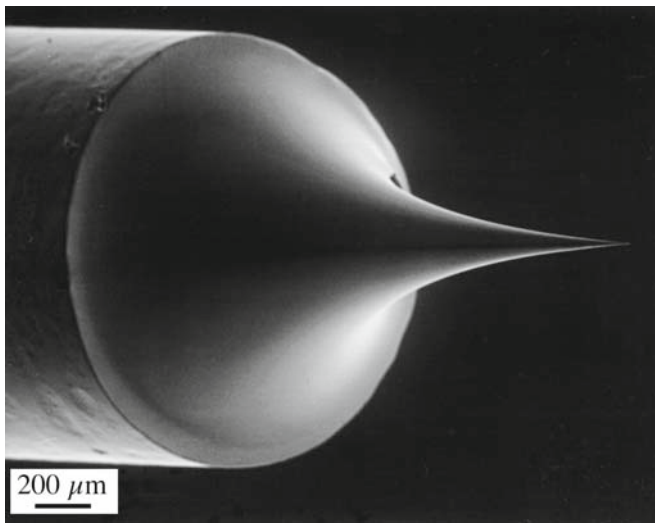


FIGURE 5.6. (A) Electron paths from a field-emission source showing how a fine crossover is formed by two anodes acting as an electrostatic lens. Sometimes an extra (gun) lens is added below the second anode. (B) A FEG tip, showing the extraordinarily fine W needle.

lens controls the effective source size and position, but it isn't very flexible. Incorporating a magnetic lens into the gun gives a more controllable beam and larger β. The faults (known as lens aberrations) in the gun lens are very important in determining the source size; we'll talk extensively about lens aberrations in Chapter 6.

VACUUM IS IMPORTANT

In a vacuum of 10^{-5} Pa, one monolayer of contaminants will form on a substrate in less than a minute. At 10^{-8} Pa, it will take 7 hours to form a monolayer.

We have already noted that CFE requires a pristine surface and, even in UHV conditions, surface contaminants build up on the tip. With time, the emission current falls and the extraction voltage must be increased to compensate. Eventually it becomes necessary to remove the contamination by 'flashing' the tip. This just means reversing the potential to the tip and 'blowing off' the surface layer of atoms, and/or heating the tip quickly to ~5000 K to evaporate the contaminants. In most CFE guns flashing occurs automatically, when the extraction voltage increases to a certain predetermined level. Because they are continuously heated, Schottky FEGs do not form the same surface contamination layer and so don't need flashing. A typical FEG tip is shown in Figure 5.6B.

A future generation of ultra-bright field-emission electron sources might well be based on carbon nanotubes which have already demonstrated a brightness in excess of 10^{14} A/m^2 sr although such a laboratory demonstration is a long way from reliable service in a commercial IVEM in a student-oriented lab. It is perhaps more likely that this technology will be developed for tips in scanning-probe microscopes or electron emitters for flat-panel display technology well before it becomes an electron source in a TEM.

5.4 COMPARISON OF GUNS

The important characteristics of the three guns we've talked about are summarized in Table 5.1. For historical reasons, we've included tungsten sources which are the worst in most respects (except price).

LaB_6 is a much more useful source for several reasons. While it is not as refractory as tungsten, LaB_6 has a much lower value of Φ and, since Φ appears in the exponential in the Richardson equation, its effect on the current density is dominant. LaB_6 crystals can be produced with a fine tip about 1 μm in radius, which accounts for the smaller crossover size. As a result, LaB_6 current densities are considerably higher than for tungsten and the brightness is typically 10 times greater, even though LaB_6 is usually operated at a much lower T to increase operating life. The decreased source size also results in improved coherency and the energy spread approaches 1 eV.

Because LaB_6 is highly reactive, the gun vacuum has to be good enough to minimize oxidation during operation, thus ensuring a reasonable life expectancy. Anything that requires improving the vacuum is good since better vacuums improve most aspects of TEM performance, but this improvement comes at a price.

The increased brightness, higher coherency, and longer life are tremendous advantages and explain why the only thermionic sources we should now use in TEM are LaB_6. You, as the operator, may have considerable control over its performance and unless the computer control overrides you, careless heating, cooling, and oversaturation can easily destroy a LaB_6 crystal. So treat LaB_6 sources gently and you will be well rewarded. If users are not careful, your TEM supervisor may extend the life of the LaB_6 to the point where it behaves no better than a W filament. LaB_6 sources don't die, they fade away.

In FEGs, the current density is enormous and β is correspondingly high. The values in Table 5.1 are all for 100 kV-accelerating voltage and you should remember that for thermionic sources, β increases linearly with kV, so there are advantages to using 300- and 400-kV instruments. However, the LaB_6 source brightness at 400 kV still does not approach β of a FEG at 100 kV. So for all applications that require a bright, coherent source, the FEG is best. This is the case for AEM, HRTEM, and such special applications as electron holography and Lorentz microscopy (for looking at magnetic domains). However, as we'll see later, the coherence of the source may produce a new complication: we must interpret the image!

There are significant differences between a CFE and a Schottky (TFE) sources and, depending on what you need from your TEM, one or the other may be significantly better. First of all, the extremely small source size of a cold FEG means that the beam is highly spatially coherent and the resulting energy spread is the smallest available without monochromation. Thermally assisted Schottky FEGs have a somewhat larger source size and larger energy spread but they provide greater stability of beam current and lower noise. A CFE requires UHV; such technology is expensive and generally requires a higher level of operator competence. However, a UHV brings other advantages such as a cleaner specimen stage and reduced contamination from the microscope system. Cleaning the tip by heating as we do for a Schottky rather than flashing as in a CFE lowers the stress on the tip, ensuring a longer life. But the cleaner UHV system for a CFE gun also

ensures a long life and so both FE sources offer similar lifetimes measured in the thousands of hours (if you're careful!). In summary, if you are doing EELS and need the lowest possible energy spread or you are doing the highest resolution STEM imaging and X-ray analysis in which the highest brightness and smallest probe size are required, then a cold FEG has advantages. For routine FEG work a Schottky gun is better, more reliable, and easier to operate.

Lastly, we should note that for routine, relatively low magnification (<50–100,000×) TEM imaging, a FEG is far from ideal because the source size is too small. It is thus not possible to illuminate large areas of your specimen at low image magnifications without losing current density, and therefore intensity, on the screen. So you can't see your image clearly at low magnifications. Under these circumstances, a LaB_6 source is better. However, increased computer control and the need for optimal performance means that FEG TEMs are increasingly popular and for the best high-resolution imaging and analytical performance, as we've made clear, there is no alternative. If you still don't get it, there's a good interactive summary of both thermionic and FE electron guns and how they work on URL #1.

5.5 MEASURING YOUR GUN CHARACTERISTICS

This section requires that you know how to operate a TEM. If you're a novice, you should skip this part of the chapter for now because we are going to refer ahead in the book for much of what you need to know.

For conventional TEM imaging and diffraction and many other routine uses, all you need to do is saturate and align the (thermionic) gun or just switch on the FEG and then ignore it. In many instruments the computer takes care of this. There are, however, times when we need to be able to measure the brightness and coherency. The source brightness is a most important parameter to measure in an AEM since, if the gun is not operating at its maximum β, then the quality of the analytical information that is generated will be poorer. Similarly, knowing the energy spread of your source is important for EELS and having a measure of the beam coherency can be important for some more advanced techniques that we've just mentioned. So let's see how we can measure the various parameters that we've just discussed. We'll start with β, then ΔE, and finally the coherency.

By measuring the three variables in equation 5.3, i.e., the beam current, the beam diameter, and the angle of convergence, we can determine β. However, while we can easily get a measure of the emission current at the gun, it is impossible to measure d_0 and α_0 there (think about why this is so). So we make the approximation that, if we neglect lens aberrations, β is constant throughout the electron-optical system so it doesn't matter where it is measured. It is easiest, practically, to determine β at the plane of the specimen and we'll now show you how to do this. (Neglecting lens aberrations is reasonable but you should be aware that C_s correction in TEMs can effectively increase the brightness of the electron beam *at the specimen* since we can use larger apertures to permit higher currents in the probe without broadening the probe dimensions.)

5.5.A Beam Current

You can measure the beam current at the specimen i_b directly using a Faraday cup in a specimen holder. A Faraday cup consists of a small aperture above a relatively deep hole in an earthed metal block. If the aperture is small enough (e.g., ~50 μm) and the metal block deep enough (~2 mm), and made of something light like Al to minimize backscatter, then it is a reasonable assumption that no electrons escape back out of the entrance aperture. All the electrons going into the aperture therefore go to earth, and you can measure the electron current using a picoammeter in the earth line. (Ideally a Faraday cup would be available permanently in the column of a TEM, and this would permit constant monitoring of the beam current but no TEM manufacturer offers such.) You should calibrate your Faraday-cup measurement against the TEM screen exposure meter or the electron energy-loss spectrometer shield current. Carrying out this procedure allows you to make a rapid estimate of i_b at any time you need it.

In modern TEMs equipped with a Schottky FEG, the beam-current fluctuation should be less than a few percent over many hours of operation. The stable beam current in Schottky-FEG TEMs does not need to be monitored frequently and can be calibrated easily through the readout from the viewing screen. Conversely, the beam current decreases with time in cold-FEG TEMs. Figure 5.7 shows the time dependence of the emission and the beam current measured after tip-flashing in the cold-FEG STEM. While the emission current decreases almost linearly, the beam current drops parabolically up to 3 hours after flashing. In thermionic-source TEMs, the beam current also decreases after stabilization. This variation would not be nearly so large for a Schottky FEG which, of course, never needs flashing.

MEASURING THE BEAM

A Faraday cup is a black hole for electrons and a very useful diagnostic tool for your TEM performance.

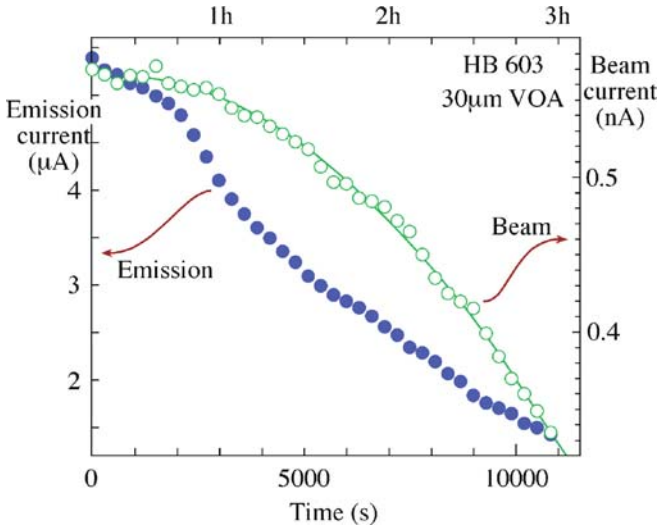


FIGURE 5.7. Time dependence of the emission and the beam currents of a cold FEG in a 300-keV VG HB603 STEM. Both the currents were measured after flashing the tip.

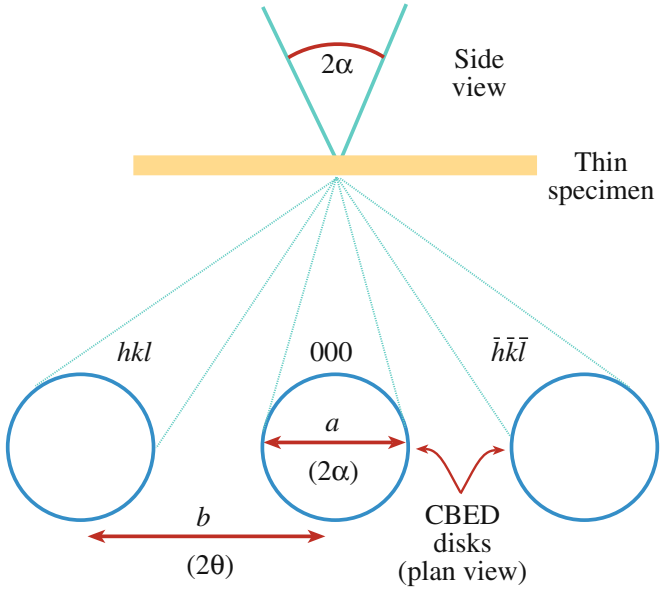


FIGURE 5.8. The distances on a convergent beam DP from which you can measure the beam convergence angle, α, which is proportional to the width of each diffraction disk.

As we'll show in Chapter 6, i_b is a strong function of the beam size. Therefore, the current is controlled by the first condenser (C1) lens strength and the size of the final beam-limiting aperture in the second condenser (C2) lens. If you look ahead to Figures 9.10 and 9.11 you will see the variation of i_b as a function of C1 lens strength and the effect of C2 aperture size on α_0.

- The beam current is usually in the range from nanoamps to picoamps.
- The emission current is typically several microamps.

So the current decreases by three to six orders of magnitude between the gun and the specimen: most is lost in the illumination system, as we'll see in Chapter 9.

5.5.B Convergence Angle

You can easily measure the convergence angle α from the convergent-beam electron diffraction (CBED) pattern, which you can see directly on the TEM screen. (You will need to read Chapter 21 in order to find out how to generate CBED patterns.) In the schematic diagram in Figure 5.8, the total convergence angle 2α is proportional to the width of the diffraction discs, *a*. This width can easily be calibrated if the specimen has a known Bragg angle $2\theta_B$ (see Chapter 11), since $2\theta_B$ is proportional to the distance, *b*, from the 000 disc to the *hkl* disc. Thus

$$2\alpha = 2\theta_B \frac{a}{b} \tag{5.6}$$

The convergence angle, α, at the specimen is not only important in the brightness equation, but we'll see that it also plays a major role in CBED, STEM imaging, XEDS, and EELS. So you must know how to measure and control α because it is essential in many aspects of TEM. The value of α is controlled by the size of the final limiting aperture in the illumination system and we'll see how this works in Chapter 6.

5.5.C Calculating the Beam Diameter

While it is a relatively simple matter to measure i_b and determine α, the measurement of *d*, the beam diameter, is not so straightforward. However, *d* is a major factor in all aspects of TEM such as AEM and STEM imaging where we use a fine focused beam. We can either calculate *d* or measure it experimentally. The former is easy but imprecise, the latter is difficult and can be equally imprecise. The first problem with determining *d* is that there is no universally accepted definition of the beam diameter. The manufacturer will give you a list of nominal beam sizes for each setting of the C1 lens. These values are *calculated* and may differ from the actual beam size by large amounts. The calculation assumes that the electron-intensity distribution in the beam is Gaussian, and the beam diameter is defined as the full-width at half-maximum (FWHM) of the Gaussian distribution, defined in Figure 5.9. To approach a Gaussian intensity distribution, the beam must be well aligned, any astigmatism in the condenser lenses corrected (see Chapter 9), and all apertures in the illumination system accurately centered. Even under these conditions you cannot obtain Gaussian conditions for

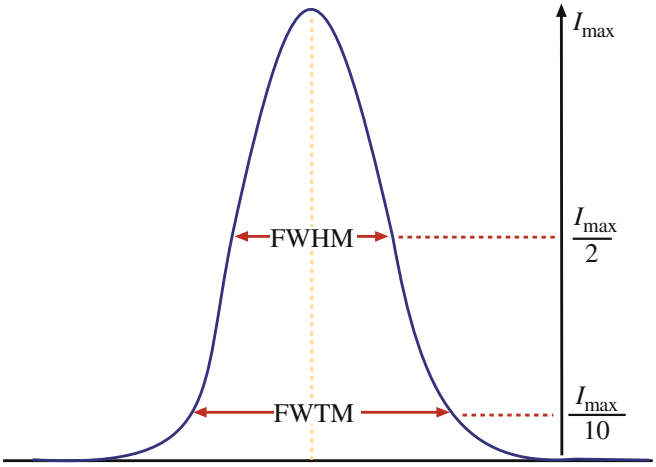


FIGURE 5.9. The definition of the full width at half maximum (FWHM) and the full width at tenth maximum (FWTM) of a Gaussian intensity distribution which is typical of a well-aligned electron beam. Ideally the beam hitting your specimen should always approximate to this kind of intensity distribution.

every possible beam size. For example, there may be six different C1 lens excitations, each of which gives a different calculated beam size, but there are invariably fewer than six C2 apertures available, so each beam size cannot be correctly apertured; spherical-aberration (C_s) effects will broaden the beam size beyond a true Gaussian (see Chapter 6). If you select too small an aperture, then the intensity distribution will be truncated at a fraction of the full Gaussian curve. If you use too large an aperture, the actual beam will extend out well beyond the calculated size and this has very important implications for XEDS in the TEM as we describe in Section 33.3.A.

To make a complete calculation of the beam size, we assume that it is determined by an initial Gaussian diameter at the gun (d_g). This diameter is broadened by the effects of spherical aberration in the beam-forming lens (d_s) and diffraction at the final aperture (d_d). All these terms can be added in quadrature (although for no better reason than that it seems reasonable) to give a total, calculated beam size, d_t

$$d_t = (d_g^2 + d_s^2 + d_d^2)^{1/2} \tag{5.7}$$

This equation gives us only a first-order estimate, since the contributions are not all Gaussian. We'll now briefly discuss the origin of each of these terms.

The value of d_g is a function of β, and a value of β has to be assumed for the purposes of calculation. The expression for d_g is

$$d_g = \frac{2}{\pi}\left(\frac{i}{\beta}\right)^2 \frac{1}{\alpha} \tag{5.8}$$

We have already defined i, β, and α.

The disc of minimum confusion caused by spherical aberration has a diameter given by

$$d_s = 0.5 C_s \alpha^3 \tag{5.9}$$

where C_s is the spherical-aberration coefficient, which we discuss in detail in Chapter 6. This is the full diameter containing 100% of the beam current. Clearly, this term is not Gaussian unless the beam is correctly apertured, which, as we just discussed, is not always possible. However, C_s correction is now available in probe-forming TEMs and then, of course, this contribution to the beam broadening goes away. The calculated diameter due to diffraction is

$$d_d = 1.22 \frac{\lambda}{\alpha} \tag{5.10}$$

which is the Rayleigh criterion that we discussed in Section 1.2.B and refers to a spacing between two overlapping images of the probe. Although all these definitions clearly do not define the same diameter of the electron distribution, the summation in quadrature is still assumed to give a first approximation of the FWHM of the beam. Figure 5.10 shows the result of calculations of the three contributions to the beam diameter in a VG HB501 STEM and the comparison with two experimental measurements carried out, as we'll now describe.

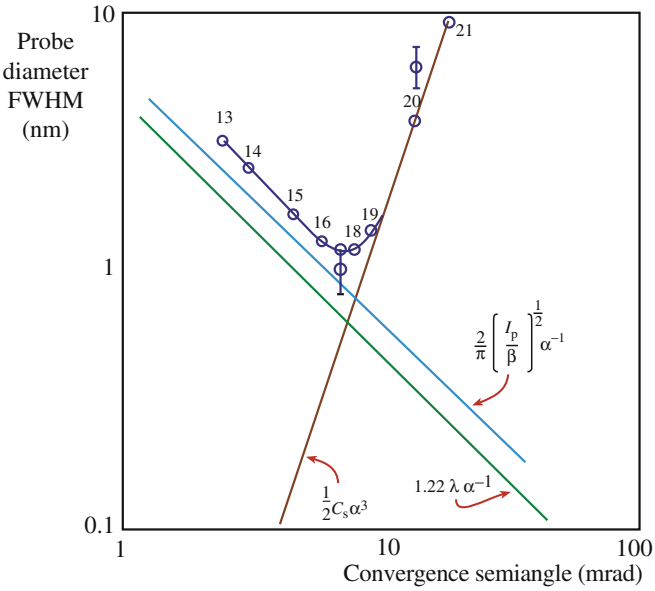


FIGURE 5.10. Calculations of the various contributions to the beam size as a function of the convergence angle α, in a FEG STEM with a probe current I_p of 0.85×10^{-8} A. Two experimental points are shown (with error bars) at first-condenser lens settings 17 and 20 (corresponding close to the minimum and the maximum probe sizes, respectively). The minimum diameter is ~ 1 nm with $\alpha < 10$ mrads.

Given all the (somewhat inaccurate) assumptions we made, the agreement with experiment is quite reasonable.

5.5.D Measuring the Beam Diameter

Given the uncertainties in the calculation of the probe size which we just described, it would seem much more reliable to measure d experimentally. To measure the beam size in a TEM you must form an image of the beam on the viewing screen or computer display under conditions where you know, or can calibrate the magnification. This is a non-trivial exercise and you may need to consult the manufacturer's handbook to be sure that you are doing it correctly. You can then *photograph* the beam and determine the intensity distribution from a microdensitometer trace across the film, or an electronic scan across the readout from the CCD detector (see Chapter 7) as shown in Figure 5.11. What you'll learn later on in Chapter 9 is that the first-condenser lens (C1) is responsible for controlling the beam size (and hence the current) and that is why different settings of C1 are mentioned in Figures 5.10–5.12. From Figure 5.11 a couple of important points can be drawn

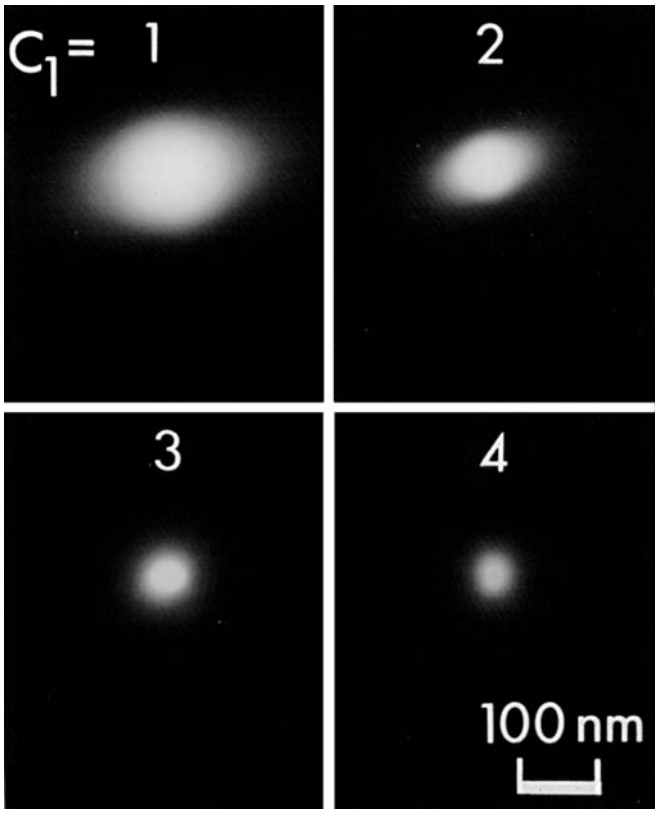


FIGURE 5.11. Four images of the beam formed on the TEM screen at different C1 lens settings. Spot #3 most closely corresponds to the Gaussian intensity distribution shown in Figure 5.9.

- The FWHM contains 50% of the integrated intensity. It is the value used by the manufacturers when they calculate beam sizes. It is also the important dimension when considering the effect of d on the (S)TEM image resolution.
- The full width at tenth maximum (FWTM) contains 90% of the integrated intensity. It is a more relevant dimension generally because the Faraday cup (or any probe-current measuring method) measures the current in the total beam, which is much closer in size to the FWTM. This dimension is also more relevant to measurement of the XEDS spatial resolution. (Think about it and see Chapter 36 later.)

When you insert the beam diameter in the brightness equation, either the FWHM or the FWTM can be used. The FWTM is equal to 1.82 × FWHM and this is also shown in Figure 5.9. You should note, therefore, that you overestimate β if you use the smaller FWHM.

In a dedicated STEM you can't image the beam directly, since there are no post-specimen lenses to magnify its image, no screen on which to project it, and no photographic film to record it. The value of d must be determined indirectly, as in other scanning instruments. The worst method (apart from all the rest) involves scanning the beam across a knife-edge specimen and monitoring the intensity change that occurs, for example, by recording the output from the annular dark-field detector (see Chapter 9). The specimen should be atomically sharp and not transparent to electrons until right at the edge. Such specimens don't exist. This approach yields an integrated intensity profile, as shown in Figure 5.12. In order to extract a value of the FWHM or FWTM from the profile, you must make measurements between various points determined by integrating the intensity from one side of a two-dimensional Gaussian to the other. Nevertheless, the two experimental beam-size measurements in Figure 5.10 show reasonable agreement with the values calculated from the brightness equation. The measurement of d is clearly not a simple procedure.

5.5.E Energy Spread

Remember that the energy spread (ΔE) of the electron beam is a measure of the temporal coherency. This spread is important in EELS and, in fact, the only way to measure the energy spread is to use an electron spectrometer. Under conditions where the spectrometer itself is not limiting the resolution of the spectrum, the value of ΔE can be simply measured by collecting a spectrum of electrons without a specimen in the way of the beam as we describe in Section 37.3.C. The spectrum then consists of a single Gaussian peak and the resolution of the spectrum is defined as the FWHM of this peak. Typical values of ΔE for the various electron

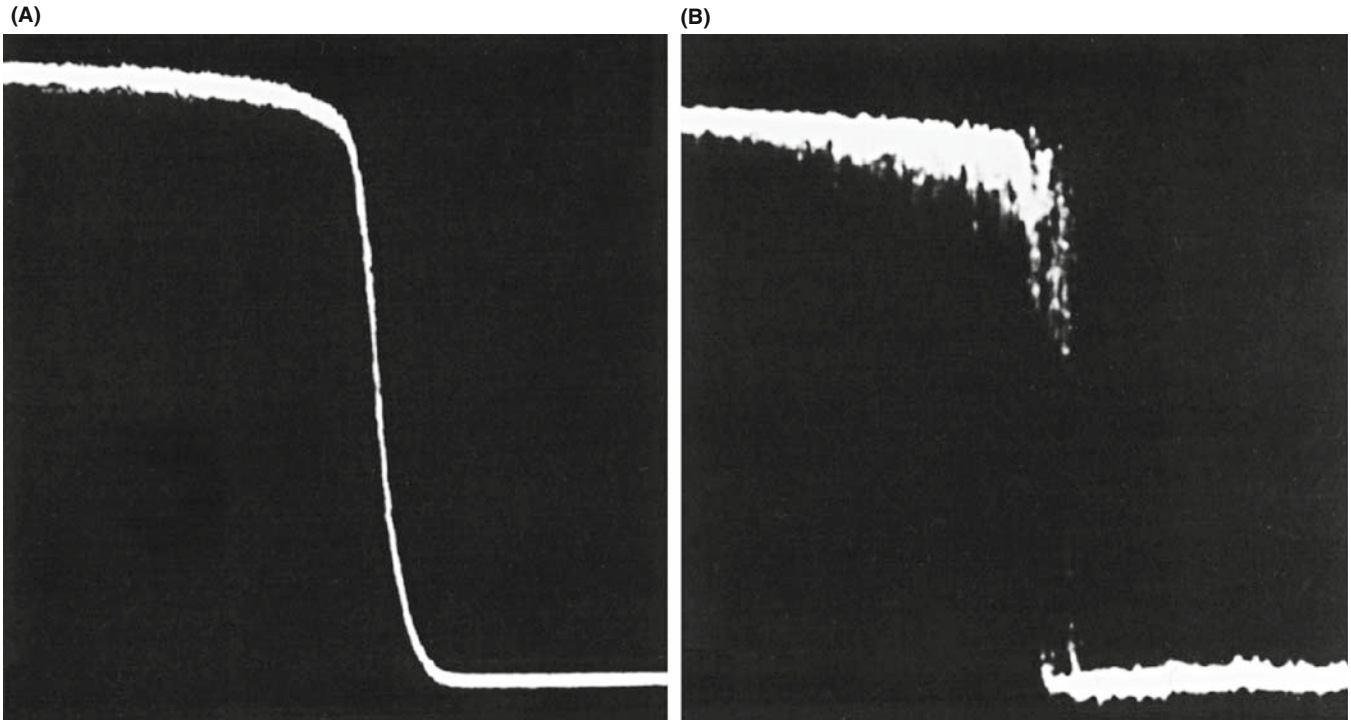


FIGURE 5.12. Intensity profiles obtained by scanning a fine beam across a sharp edge of a cube of MgO. The measured probe size (FWTM) in (A) is 7.4 nm (magnification 1×10^6) and in (B) 1.8 nm (magnification 11×10^6). The smaller probe contains a much smaller current and is thus a much noisier trace.

sources are given in Table 5.1; with a monochromator (see Section 37.7) ΔE for any source can be reduced to $<\sim$100 meV.

5.5.F Spatial Coherency

It's difficult to measure the coherency of the beam experimentally although, as we've discussed, small sources ensure spatial coherency. One practical way of measuring the coherency is to form an image of the edge of a hole in a specimen such as a thin holey carbon film. When you operate slightly out of focus you see alternating dark and bright fringes called Fresnel fringes as shown in Figure 5.13A. Typically, for a thermionic source, only one or two fringes are visible. These fringes are a phase-contrast effect (which we cover in great detail in Part 3). We can also use the fringes to correct astigmatism in the objective lens, as we'll see in Chapter 6. The number of visible fringes is a measure of the beam coherency and Figure 5.13B shows the enormous number that can be generated by a FEG.

5.6 WHAT kV SHOULD YOU USE?

For the materials scientist and nanotechnologist, this is usually an easy question to answer: we'll call it the kV axiom.

THE kV AXIOM

You should always operate at the maximum available kV (unless you shouldn't).

However, there are exceptions to this axiom and the most obvious is avoiding knock-on beam damage, but we'll see others later in the book. So don't forget that you can always operate a modern 300-kV TEM at 100 kV. It's like being able to change the wavelength of a monochromatic light source in a visible-light microscope. As we saw in Chapter 4, the threshold for displacement damage for most metals is less than 400 kV, which is the highest available voltage on 'off-the-shelf' TEMs. For lighter and more beam-sensitive materials such as some ceramics and polymers, lower kV may be better. For most materials specimens, there is not much use going below 100 kV since the images will be rather dim and you'll have to make *very* thin specimens to see anything useful. However, when studying a crystalline specimen by diffraction contrast, 100 kV is better than 200 kV is better than 300 kV, providing you can still see through the specimen! For biological specimens, there is significant advantage to the increased contrast available in lower-kV images and STEM in (30-kV) SEMs is an increasingly useful imaging tool.

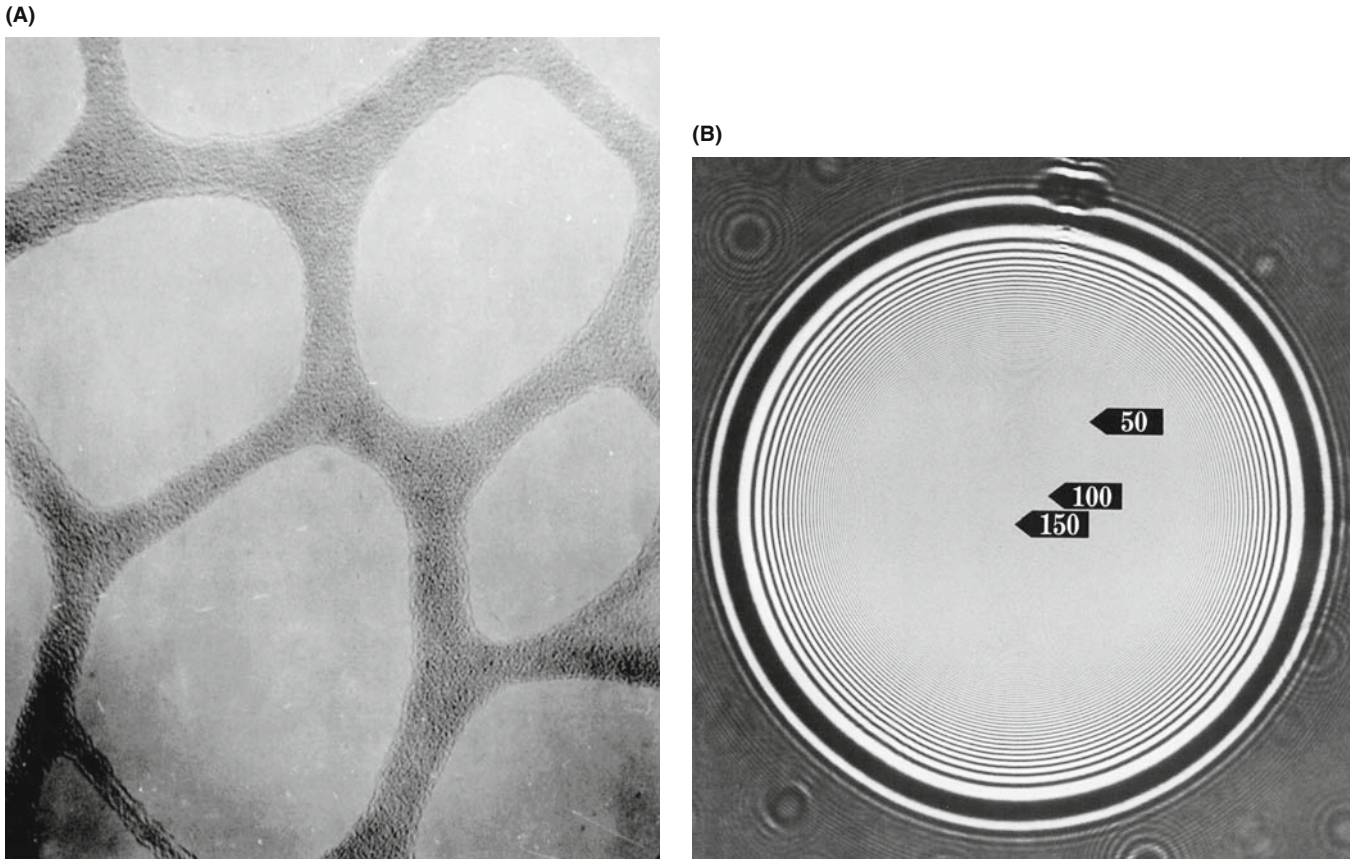


FIGURE 5.13. Fresnel fringes from (A) a thermionic source with poor coherency and (B) a FEG with high coherency.

Apart from these exceptions, the reasons for choosing the highest kV are

- The gun is brightest so you get the most signal to put into your specimen.
- The wavelength is shortest; the image resolution is potentially better.
- The cross section for elastic scatter is also reduced so beam broadening is reduced and analytical spatial resolution is enhanced.
- The cross section for inelastic scatter is smaller, so heating effects may be smaller.
- You can 'see' through thicker specimens.
- The peak to background ratio in X-ray spectra is improved (see Chapter 36).

When you've learned about EFTEM later on, return to this chapter and ask: why not 80 kV?

CHAPTER SUMMARY

Most TEMs use LaB_6 thermionic sources. Take care when heating and cooling your LaB_6 crystal; always operate just below saturation to maximize the lifetime of the source and (almost) always operate at the highest kV. If you're doing high-resolution work of any kind, find a FEG TEM and for high-resolution imaging, then the degree of coherency is important too. If you're doing XEDS, get some idea of the beam current that you can get from your FEG under typical operating conditions. Also measure the beam size and convergence angle to give a measure of β. If you're doing EELS then the energy spread is essential information and you may need to find one of the rare, monochromated TEMs. Treat your source particularly carefully if your TEM is of such a vintage that you have to change it, align it, saturate it, or switch it off. There's nothing more annoying than losing your source since it usually happens at some critical point during your work. Fortunately, computer control is making such events much rarer.

SOME HISTORY

Why include a note on history? Because you should/will study old papers. The sources used in the 1950s and 1960s were so-called W hairpins. The simple hairpins were replaced by pointed W filaments (hairpins with an atom-probe-like tip attached). In the 1970s, LaB_6 filaments replaced W as more efficient emitters. Other materials might be more efficient, but... In the late 1980s, FEGs began to be used in several labs. By the 2000s, these are the sources of choice, especially for those who can afford them. So, we are back to the pointed filament. Now, consider how the TEM techniques had to be different in those old papers.

SOURCES

Broers, AN 1974 in *Recent Advances in Scanning Electron Microscopy with Lanthanum Hexaboride Cathodes* SEM 1974 9–18, Ed. O Johari IITRI Chicago IL.

de Jonge, N and van Druten, NJ 2003 *Field Emission from Individual Multiwalled Carbon Nanotubes Prepared in an Electron Microscope* Ultramicroscopy **95** 85–91. Demonstrated a brightness in excess of $10^{14}A/m^2sr$.

Hawkes, PW 1978 *Coherence in Electron Optics* Adv. Opt. Electr. Microsc. **7** 101–184.

Orloff, J 1989 *Survey of Electron Sources for High-Resolution Microscopy* Ultramicroscopy **58** 88–97.

Veneklasen, LH 1972 *Some General Considerations Concerning the Optics of the Field Emission Illumination System* Optik **36** 410–433.

THE PROBE

Michael, JR and Williams, DB 1987 *A Consistent Definition of Probe Size and Spatial Resolution in the Analytical Electron Microscope* J. Microsc. **147** 289–303. Details on how to measure the diameter of the probe.

Mook, HW and Kruit, P 1999 *On the Monochromatisation of High Brightness Electron Sources for Electron Microscopy* Ultramicroscopy **78** 43–51. Application to EFTEM.

THE COMPANION TEXT

There is much more in-depth treatment of electron guns in the companion text. Electron coherence is a very tricky topic that the most experienced of us can still confuse. It is examined in the discussion of lenses and of holography in the companion text.

URLs

1) http://www.matter.org.uk/tem/electron_gun/electron_sources.htm

SELF-ASSESSMENT QUESTIONS

Q5.1 State the two types of electron sources currently used in TEMs and explain how they work.
Q5.2 Name two thermionic sources and the properties that make them useful.
Q5.3 What is the difference between a field-emission and thermionic source TEMs?
Q5.4 What is brightness and how does it change with kV?
Q5.5 When is high brightness most useful? When is low brightness useful?
Q5.6 Name the five most important properties or characteristics of an electron beam?
Q5.7 What are some reasons to choosing the highest kV when operating a TEM?
Q5.8 What is the purpose of the Wehnelt in a thermionic source? Why don't we need one in a FEG?
Q5.9 Describe the purpose of the two anodes in a FEG.
Q5.10 Why is a FEG operated under high-vacuum conditions?
Q5.11 If faced with the urge to 'crank up the temperature' on the LaB_6, how long between each heat setting should a wise microscopist wait?
Q5.12 What are the limitations of a field-emission TEM?
Q5.13 How would you maximize the coherency of the source?
Q5.14 What is the 'saturation condition' for a thermionic source?
Q5.15 Why would you operate your thermionic filament just under the saturation condition?
Q5.16 How do you know you have achieved gun saturation while using the TEM?
Q5.17 What is spatial coherency and why is it important?
Q5.17 What is temporal coherency and how is it measured?
Q5.19 Name three ways to increase the coherency of the beam.

TEXT-SPECIFIC QUESTIONS

T5.1 Carefully redraw Figure 5.1 to scale.
T5.2 Sketch Figures 5.5A and 5.5C with scale bars.
T5.3 By considering equation 5.3 explain
A. Why a LaB_6 source brightness varies as a function of the orientation of the LaB_6 crystal?

B. Why adjusting the voltage on the Wehnelt can change the apparent brightness of a thermionic source? (Hint: look at Figure 5.4.)
C. Why the brightness of a cold FEG is generally higher than that of a (thermally assisted) Schottky FEG?
D. Why a tungsten source gets brighter if you sharpen the tip of the hairpin?

T5.4 Why is the concept of gun brightness generally not relevant when observing a specimen in TEM mode? Under what operating conditions does the brightness become crucial and why?

T5.5 Estimate the approximate power density in the probe when a FEG puts 1nm of current into a spot of diameter 1nm and a LaB_6 gun puts 10 pA into 1nm. Show all steps and justify any approximations.

T5.6 If the emission current from a thermionic source is several hundred μA (see Figure 5.4B), why does the beam at the specimen contain only a few hundreds or thousands of picoamps?

T5.7 Calculate the beam size for a 100-keV FEG source from the data in Chapter 1, Table 5.1, and equations 5.7, 5.8, 5.9, and 5.10. State any assumptions. Compare your data with Figure 5.10.

T5.8 From Figure 5.10, explain why it would be useful to be able to use larger apertures in the probe-forming system and what prevents us from doing so in most TEMs until recently? (Hint: go back to the questions for Chapter 1.)

T5.9 From the data in Figure 5.10 calculate the beam brightness.

T5.10 Why do you need a different C2 aperture for each possible C1 lens setting? Do you have such a range of apertures on your microscope? (Hint: look at Figure 5.11.)

T5.11 Why is the probe size measured in different ways in a TEM and a dedicated STEM (compare Figures 5.11 and 5.12).

T5.12 Can you think of any use for the much larger number of Fresnel fringes in FEG-TEM images compared with thermionic-source TEM images, such as in Figure 5.13?

T5.13 How would the data in Figure 5.7 change if you had not flashed the tip?

T5.14 Can you think of any other suitable specimens for measuring the experimental probe size by the method shown in Figure 5.12?

T5.15 Can you think of what effect a non-Gaussian probe shape might have on (A) your STEM images and (B) your analyses in AEM?

T5.16 Which line in Figure 5.10 would change for a TEM with a spherical-aberration corrector? Which way would it move and what would be the consequence for the minimum probe dimension?

T5.17 Why is a Schottky field emitter so called?

6

Lenses, Apertures, and Resolution

CHAPTER PREVIEW

Electron lenses are the TEM's equivalent of the glass lenses in a visible light microscope (VLM) and, to a large extent, we can draw comparisons between the two. For example, the behavior of all the lenses in a standard TEM can be approximated to the action of a convex (converging) glass lens on monochromatic light. The lens is basically used to do two things

- Take all the rays emanating from a point in an object and recreate a point in an image
- Focus parallel rays to a point in the focal plane of the lens

The lens can't collect *all* the rays from the object and we often deliberately limit the collection angle with an aperture. We can draw ray diagrams showing how electron lenses control beams of electrons. These diagrams correspond directly to the ray diagrams used in physical optics. Of course the analogy with light fails for certain aspects, but basically it will pervade this chapter. So we'll start by reminding you of the principles of light optics insofar as they relate to electron optics. Then we'll discuss the electron lens in more detail, showing how an electron behaves as it passes through such a lens. We'll describe some actual lenses and tell you how we use different kinds of electron lenses to do different things in the microscope.

A major limit to the use of electron lenses is the fact that we aren't very good at making them. They suffer from rather severe spherical and chromatic aberrations which we usually control by inserting limiting apertures to select electrons nearest to the optic axis since these are least affected by the lens aberrations. Recent technical developments have permitted these aberrations to be largely overcome, but aberration-corrected TEMs are both rare and expensive; most microscopists still have to live with these limitations. So you need to understand lens aberrations, since they play a major role in deciding what we can and cannot do with the microscope. In particular, lens aberrations (rather than the wavelength of the electrons) limit the resolution of the TEM (unlike in the VLM). Since resolution is often the single most important reason for buying a TEM, you need a firm understanding of this concept. Unfortunately, we electron microscopists aren't always very precise in our definitions of resolution. Finally, we describe how the apertures we use aid both the depth of field and the depth of focus of the instrument.

6.1 WHY LEARN ABOUT LENSES?

Why should we learn about electron lenses? As in a VLM, the lenses in a TEM control all the basic operational functions of the instrument. As you are well aware, we have to physically move glass lenses up and down in a VLM to control the intensity of the illumination and to focus the image. The focal length of a glass lens is fixed so we have to change lenses to change the magnification. We choose stronger lenses for higher magnification. By contrast, in a TEM, the positions of the lenses are fixed but we can change the strength of the lens at will.

CHANGING THE LENS

We change focus, change the intensity of illumination or change magnification by changing the strength of the lenses.

As you'll see, in most cases the lenses we use are electromagnetic, so we change their strength by

changing the current through a coil around a soft-iron core which changes the strength of the resultant magnetic field. Almost any operation we carry out on the TEM involves changing magnification or focus; we use electron lenses to magnify and focus the electron beam, the images, and the DPs.

These factors are critical in the principal functions of a TEM: imaging, diffraction, and analysis which, respectively, comprise the next three parts of this book. An aperture is used to select different electron beams to form different images, thus manipulating the image contrast. Another aperture is used to select different regions of the specimen to contribute to the DP as we'll see in Chapter 9.

So knowing how these aperture/lens combinations work allows you to understand how we control the TEM and why we do certain operations on the microscope.

APERTURES

We use apertures in the lenses to control the divergence or convergence of electron paths through the lenses which, in turn, affects the lens aberrations and controls the current in the beam hitting the specimen.

An understanding of electron lenses will help us to answer such questions as

- Why can we see finer detail with an electron microscope than with a light microscope?
- Why can't we see as much detail as we might expect from physics?
- Why does the TEM have a better depth of field and depth of focus than the VLM?

We'll see that the answers to these questions lie in the quality of the lenses, and how we use them. In this chapter we'll discuss the basics of how a lens/aperture combination works. Throughout the book you'll come across different uses and combinations of lenses and apertures. So this is a central chapter for the serious microscope operator but it is only an introduction to the important aspects of electron optics, which is a field in itself. For this you need to explore Chapter 2 in the companion text and the electron-optics texts in the references. Apart from the tremendous advances brought about by aberration correction, electron optics is relatively static these days but traditional light optics is undergoing a renaissance from which electron optics can only benefit. If you are interested you should consult any of the optics textbooks that we referenced back in Chapter 2 and check out URL #1.

6.2 LIGHT OPTICS AND ELECTRON OPTICS

You are already familiar with the action of a magnifying glass lens on light rays. The magnifying glass is a convex lens. It can be used in two ways to control the light rays coming through it. First, it can produce a magnified image of the object you're looking at. Second, it can focus a parallel beam of light to a point, in the focal plane of the lens. (When younger, we've all used this latter property to set something or someone on fire by focusing the (parallel) sun's rays.) These two actions, forming an image of an object and focusing parallel rays to a point, are all we need in order to understand how the lenses in a TEM work. The reason that we can get away with this simple approach is because the electron lenses act, to a reasonable approximation, like convex glass lenses; in detail they're often equivalent to more complex combinations of convex lenses and aberration correction involves the equivalent of a divergent or concave lens. We introduce the practical use of lenses in the TEM in Chapter 9.

6.2.A How to Draw a Ray Diagram

In traditional light optics it's customary to draw diagrams of the paths of light rays through the lens and these ray diagrams are usually drawn horizontally because the traditional optical bench on which light-optical experiments are carried out is a horizontal setup. Likewise, we draw diagrams of the electron trajectories through electron lenses but, since the TEM is a vertical instrument, we will draw all our ray diagrams vertically assuming the gun is at the top of the column of lenses (although this isn't invariably so, as we describe in Chapter 9).

Let's start by drawing ray diagrams to illustrate the two fundamental lens actions of image formation and the focusing of parallel rays. In these and all subsequent diagrams we'll draw all the lenses in the TEM as convex lenses. We will draw all electron ray paths as straight lines outside the lens, and we'll start by assuming that the lenses are perfect. We'll also draw the lenses as so-called 'thin' lenses, which means their thickness is small compared to their radii of curvature. Actually, we'll make the lenses *very* thin. We'll see that these assumptions are all precisely wrong, yet sufficiently reasonable that traditional ray diagrams are nonetheless very useful.

The first thing we need to do is to have a base line on which to draw our diagrams; this line is called the optic axis (also called the rotation axis in the TEM because, as you'll learn, the electrons actually rotate through the lens even though we draw the ray paths as straight lines).

THE OPTIC AXIS

An imaginary line down the column of the TEM passing through the center of each lens.

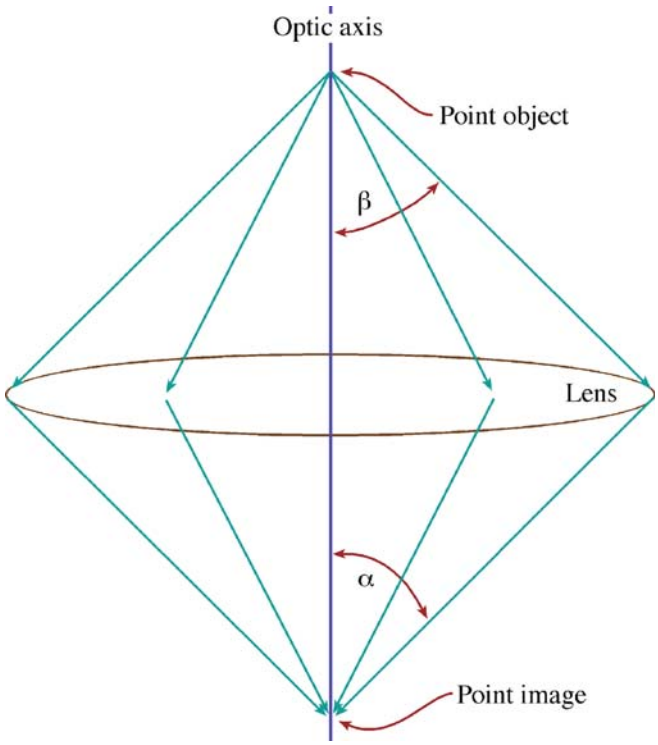


FIGURE 6.1. Image formation by a convex lens. A point object is imaged as a point and the collection angle of the lens is defined relative to the object (β) or the image (α).

Now the first action of a lens that we want to show is how it produces an image of an object. In a TEM the object will usually be the specimen itself or an image of it, but it may also be the electron source, which is an object for the illumination system. If we assume the object is a point and the radiation is emanating from that point (a so-called 'self-luminous object'), then a perfect lens will gather a fraction of that radiation and form a point image. This action is shown in Figure 6.1 in which the point is on the optic axis. The fraction of the rays from the object gathered by the lens is an important variable, defined by the angle β in Figure 6.1. Ultimately, as you can see, β is governed by the size of the lens, but we often choose to limit β by inserting an aperture, as we'll discuss later in this chapter. You'll often see the angle of collection defined as α, but we will reserve α for convergence angles (see Section 2.7). From now on, as we did in earlier chapters, we'll talk about angles when we actually mean semi-angles.

LENSES ARE FINITE

All lenses are imperfect insofar as they cannot gather all the radiation emitted by an object and so can never create a perfect image.

However, as you know from Chapters 2 to 4, most electrons are strongly forward scattered, so we can in practice gather a high fraction of the scattered electrons.

The angles in Figure 6.1 and in the other ray diagrams we'll draw are all greatly exaggerated.

EXAGGERATE ANGLES

In practice, a typical value of β is maybe a few tens of milliradians (10 mrad ~ 0.57°) so if the diagrams were drawn to scale they would be many times longer than they were wide and all the ray paths would be exceedingly narrow. Since drawing to scale is impractical, we always exaggerate the angles considerably in all electron ray diagrams.

If the object has a finite size, we can illustrate this by an arrow, asymmetrically positioned with respect to the optic axis, as in Figure 6.2. Then the lens creates an image of the arrow, rotated by 180°. To draw this figure, the first step is to draw line 1 from the arrowhead through the center of the lens, because rays crossing the optic axis in the lens (or on-axis rays which travel down the axis) are *not* affected by the lens at all and remain as a straight line.

The second step is to draw line 2 which is a ray from the arrowhead that is parallel to the optic axis. We could draw such a ray from any point along the arrow and the further away rays are from the optic axis, the more strongly they are bent by a convex lens. So we take line

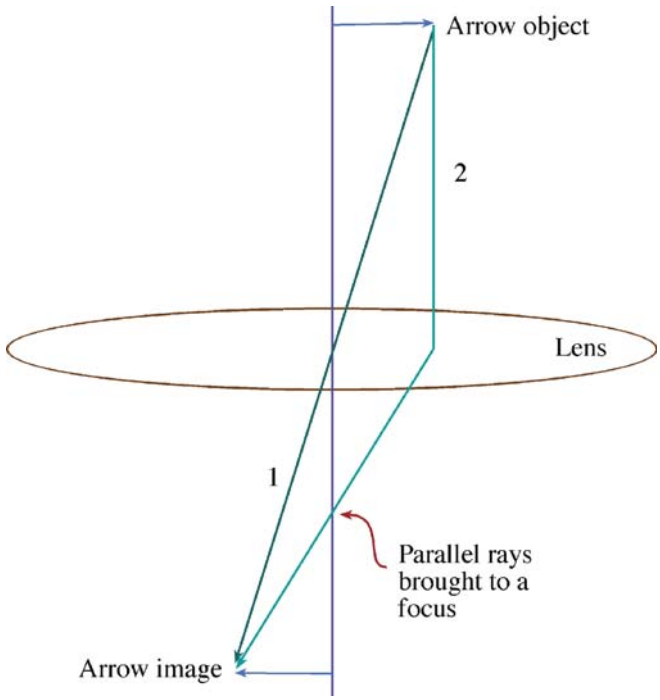


FIGURE 6.2. How to draw a ray diagram: first construct ray path #1 through the middle of the lens, then draw ray path #2 (initially parallel to the optic axis) to determine the lens strength. Where path #2 intersects the optic axis defines the focal plane.

2 and bend it toward the optic axis as it passes through the lens. We can choose to make the lens as strong as we wish, and the strength determines how much the ray is bent and where lines 1 and 2 meet to recreate an image of the arrowhead. Where ray path 2 intersects the optic axis (and thus intersects a ray path along the axis) defines the focal plane of the lens and thus illustrates the second fundamental action of a convex lens, i.e., the lens brings rays that are initially parallel to a focus.

FOR A THIN LENS

A fundamental principle of how a lens works is that an electron passing through the middle of the lens is unaffected so we can draw its path as a straight line. All other electron paths are bent when they pass through the lens.

Some important points on electromagnetic lenses

- The strength of the lens determines where the parallel electrons are focused: stronger lenses having shorter focal lengths.
- The focal plane is where initially parallel rays intersect after passing through the lens.
- The image formed by the lens is rotated by 180° with respect to the object.

Now a full ray diagram for an object of finite size, symmetrically positioned about the axis, combines aspects of Figures 6.1 and 6.2, as shown in Figure 6.3. In Figure 6.3, all rays from a point in the object are brought back to a point in the image and all parallel rays (whether parallel to the optic axis or not) are brought to a focus in a plane at a position depending on their angle to the axis.

Note that on-axis parallel rays are focused on axis and off-axis parallel rays are focused off axis.

This is a most important property, since it allows the lens to create DPs in the focal plane. We'll use this diagram to introduce you to the principal terms used in lens optics.

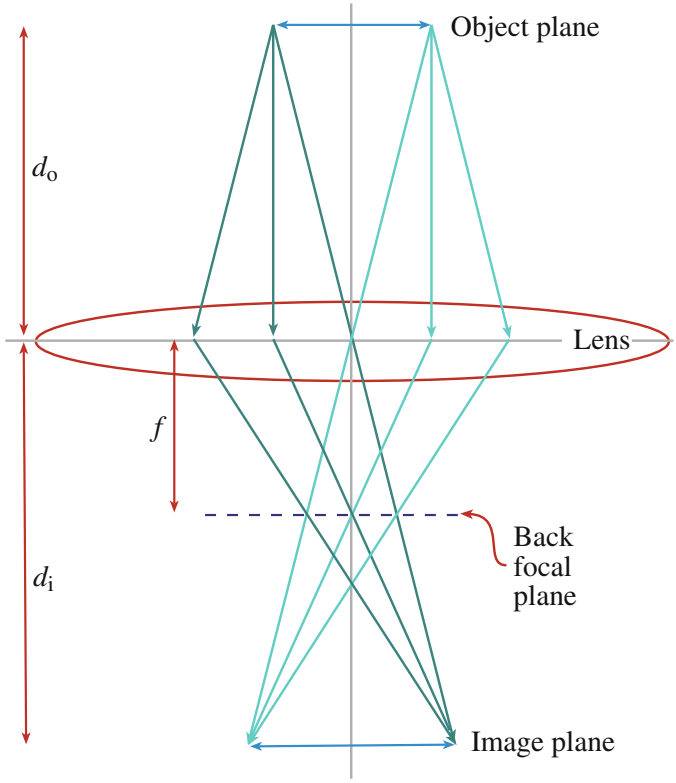


FIGURE 6.3. A complete ray diagram for a finite object, symmetrically positioned around the optic axis. All rays emerging from a point in the object (distance d_o from the lens) that are gathered by the lens converge to a point in the image (distance d_i from the lens) and all parallel rays coming from the object are focused in the focal plane (distance f from the lens).

6.2.B The Principal Optical Elements

From the above diagrams, we can define several *principal planes* to which we will often refer. The first plane is the plane of the lens. In a thin lens this plane can be imagined as a line through the middle of the lens. The object plane is the plane containing the object point in Figure 6.1 or the object arrow in Figures 6.2 and 6.3. The object plane always lies above the lens in question in the diagrams in this text. The image plane (sometimes called the Gaussian-image plane) is the plane containing the image point or arrow and it always lies below the lens. These two planes are said to be conjugate, which means optically equivalent. Rays leaving a point in one plane are brought to a point (if the lens is perfect) in a conjugate plane and vice versa. In other words, the electron doesn't care which way it goes through the lens and this is the basis for the theorem of reciprocity which we'll discuss when we compare TEM and STEM imaging in Chapter 9. The focal plane of the lens is the plane in which the parallel rays are brought to a focus as shown in Figures 6.2 and 6.3. In the image-forming process in a TEM, the focal plane lies after or 'behind' the lens and so the plane is sometimes called the back-focal plane (BFP). There is also an equivalent front-focal plane (FFP) and a convex lens would take all the rays coming from a point in the front-focal plane and create a parallel beam of radiation, in exactly the reverse manner to Figures 6.2 and 6.3.

6.2.C The Lens Equation

From the above diagrams we can define three important distances, labeled in Figure 6.3: the distance from the object plane to the lens (the object distance d_o), the distance from the lens to the image plane (the image distance d_i), and the distance from the lens to the back-focal plane (the focal

length f). Now if the lens is symmetric in strength either side of the lens plane (i.e., the front and back-focal planes are the same distance from the lens) then we can write the following basic equation

$$\frac{1}{f} = \frac{1}{d_o} + \frac{1}{d_i} \tag{6.1}$$

THE PLANES
The principal planes of a lens comprise the object, image and focal planes.

which is known as Newton's lens equation. You'll find a proof in any standard optics text (several were referenced back in Chapter 2). The distances d_o and d_i are measured from the two different principal planes in a thick lens, but from the same plane in the middle of a thin lens, which we are assuming here. In all cases that we'll consider, the object distance (and therefore the image distance) is greater than the focal length. Thus a real image is produced on the other side of the lens beyond the back-focal plane. If the object were within the (front) focal length, then a virtual image would be produced on the same side of the lens as the object, and this is often the case in light optics. Since we don't deal with virtual images in the TEM we'll ignore this aspect.

6.2.D Magnification, Demagnification, and Focus

We can use Newton's lens equation to define the magnification of a convex lens as

$$M = \frac{d_i}{d_o} \tag{6.2}$$

M is also approximately equal to the ratio of the collection angles of the lens subtended at the object (β) and at the image (α) as shown in Figure 6.1, assuming that these angles are small, as they invariably are in a TEM. In this example the magnification is unity.

STRENGTH VERSUS MAGNIFICATION
Under conditions normally found in the TEM, strong lenses *magnify less* and *demagnify more*. In VLMs stronger lenses produce greater magnifications.

Now we may sometimes want to *demagnify* an object (for example, when we want to form a small image of the electron source, to create the smallest possible probe at the specimen). If that is the case, we define the demagnification as $1/M$. In a VLM we could change the magnification by moving the object relative to the lens or vice versa, and adjusting our eyes accordingly, but generally we rotate in another objective lens of different strength (curvature). In a TEM we change magnification in this latter way by changing the strength of the lens, but you'll see that we can do this without changing the lens itself. So electron lenses differ fundamentally from glass lenses in that one lens can be adjusted to a range of strengths.

If we make the lens stronger, then the focal length is shortened as shown in Figure 6.4. If f is shortened but d_o is unchanged, then d_i must be correspondingly shorter and the image magnification is smaller, or the demagnification is larger.

How do we get the high magnifications that we need to form images of atomic columns such as Figure 1.2? Since, as you'll see in Chapter 9, we tend to operate the objective lens of the TEM at a fixed strength, we move the object plane close to the lens thus making d_o small and M correspondingly large (see equation 2). We then make the image plane of the first lens, the object plane for the next lens and repeat this for several lenses in tandem one after the other. So we end up with a multi-lens system like a compound VLM. We'll discuss many more details of lens combinations in the illumination and imaging systems of the TEM in Chapter 9.

Now, in principle, there's nothing to stop us magnifying as much as we wish. However, above a certain magnification, we will see no more information because

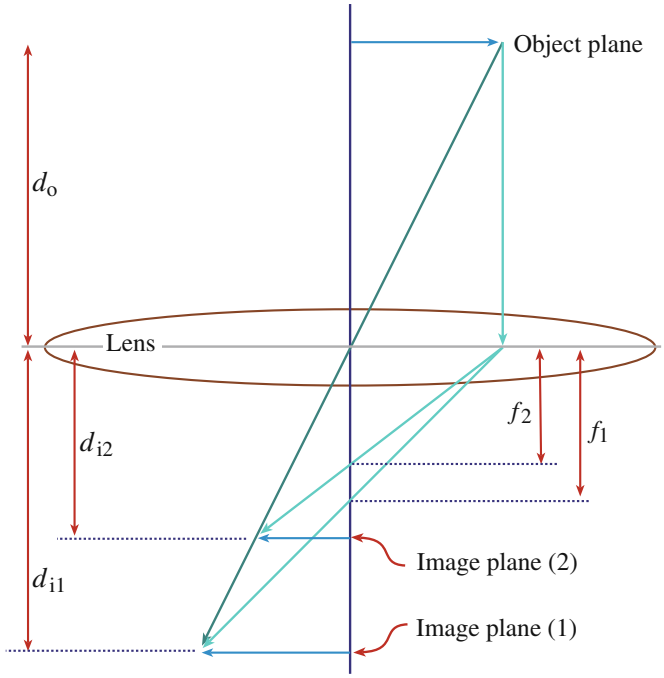


FIGURE 6.4. Strengthening the lens shortens the focal length f. So a weaker lens ($f1$) produces a higher magnification of the object than a stronger lens ($f2$) since the image distance d_i increases, but the object distance, d_o, is unchanged.

other factors limit the image detail and therefore the resolution of the microscope. We'll discuss this point later in Section 6.6. We'll also see that there are times when we want to look at an image of the focal plane (because this contains the DP). To do this, the back-focal plane of the upper lens must become the object plane for the subsequent lenses in the imaging system.

MAGNIFICATION VERSUS RESOLUTION
Don't confuse the two.

When discussing the focus of images we need another convention because we'll find that there is much useful information to be gained and certain technical advantages to operating out of focus. This situation is somewhat different to almost any other form of microscopy wherein out-of-focus images are generally less useful or, more likely, completely useless. However, in TEM we need to define the following two conditions relative to the plane in which a focused image is formed

- If the lens strength is increased such that the image forms above (i.e., before the rays get to) the image plane, then the image will be out of focus and we say the lens is *overfocused*.
- If the lens is weakened and the image forms below (i.e., after) the image plane, the image will be out of focus and the lens is said to be *underfocused*.

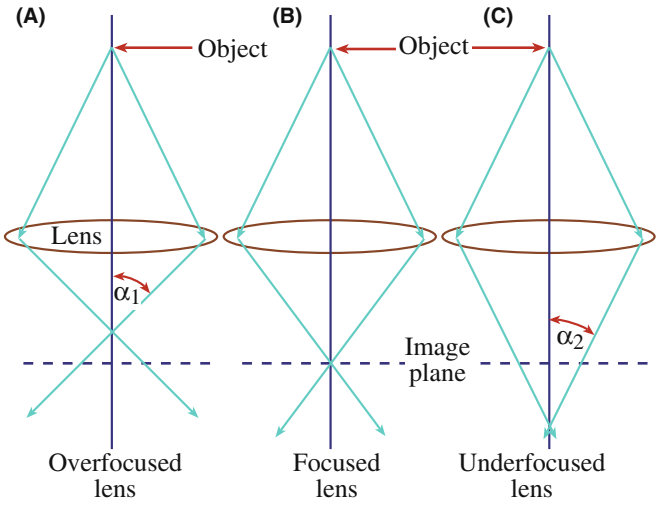


FIGURE 6.5. (A) The concept of overfocus in which a strong lens focuses the rays from a point in the object above the normal image plane where a focused image (B) of the object is usually formed. At underfocus (C) the lens is weakened and focuses the rays below the image plane. It is clear from (C) that at a given underfocus the convergent rays are more parallel than the equivalent divergent rays at overfocus ($\alpha_2 < \alpha_1$).

It's very easy to confuse these two terms unless you think in terms of the vertical frame of the microscope as shown in Figure 6.5. One point to note from Figure 6.5, which we'll find useful, is that the electrons are closer to being parallel to the optic axis when the lens is underfocused than when it is overfocused.

WEAK LENS
A weak, underfocused lens gives a more parallel electron beam. Remember α_1 and α_2 are very small.

We'll exploit underfocused imaging conditions on many occasions in the future. We'll also find there are times when we should operate with our DPs out of focus and also get different information to when it is in focus. So even dexterously challenged TEM operators or those with aging eyes can still do well!

6.3 ELECTRON LENSES

Electrons were first successfully focused by Busch in 1927; he used an electromagnet of the sort that Ruska later incorporated into the first TEM shown in Figure 1.1. Busch also showed that it was possible to focus electrons using electrostatic fields and we've already seen how this works in thermionic electron guns in Chapter 5. In practice, magnetic lenses are superior in many respects, particularly because they are not susceptible to high-voltage breakdown. The TEMs that we're discussing in this text all use magnetic lenses, so we won't discuss electrostatic lenses further here but they are examined in the companion text.

6.3.A Polepieces and Coils

To make a magnetic electron lens we need two parts. Both are drawn schematically in cross section in Figure 6.6. First there is a cylindrically symmetrical core of soft magnetic material such as soft iron, with a hole drilled through it. We call this soft iron a *polepiece* and the hole is called the *bore* of the polepiece. (Soft refers to the magnetic not the mechanical behavior.) In most lenses there are two polepieces (upper and lower), which can be part of the same piece of soft iron as in Figure 6.6 or they may be two separate pieces. The distance between the polepiece faces is called the *gap* and the bore-to-gap ratio is another important characteristic of such lenses, controlling the focusing action of the lens. Some polepieces are machined to a cone shape; the cone angle is then an important variable in the lens performance.

The second part of the lens is a coil of copper wire which surrounds each polepiece. When we pass a current through the coil, a magnetic field is created in the bore.

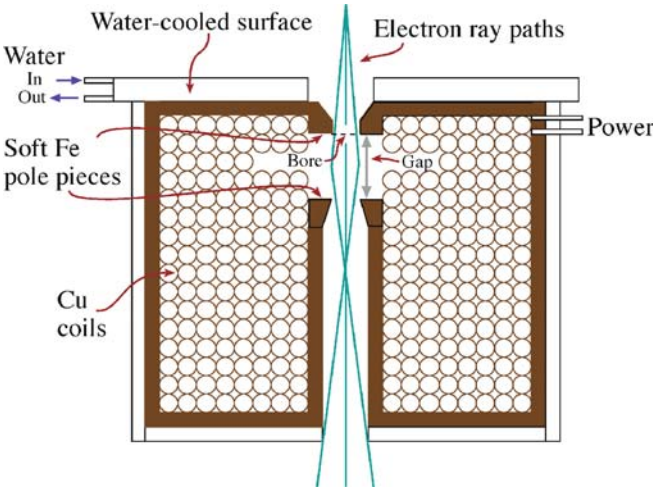


FIGURE 6.6. Schematic diagram of a magnetic lens. The soft-iron polepieces sit in the hole down the middle of the lens and are surrounded by the copper coils through which the current runs to magnetize the polepieces. When viewed in cross section, the bore and the gap between the polepieces are visible. The magnetic field is weakest on axis and increases in strength toward the sides of the polepiece, so the more the electrons travel off axis the more strongly they are deflected.

FIGURE 6.7. A real lens: the cylindrical shape conceals the copper wire coils. The two conical polepieces beside the lens sit inside the central hole in the lens. The three-pin electrical connections provide current to the coil to magnetize the polepieces, and cooling water is circulated in and out of the two holes in the top plate of the lens to dissipate the resistive heat generated in the coils. Compare this picture with the schematic in Figure 6.6.

This field is inhomogeneous along the length of the lens, but axially symmetric. It is the strength of the field in a magnetic lens that controls the electron trajectories or ray paths. As you can see, the electron path through the lens is a reasonable approximation to the schematic diagram back in Figure 6.1.

The resistive heating of the coil means that the lenses have to be cooled and a water recirculating system is an essential part of TEM lenses. A real lens removed from the column of a TEM is shown in Figure 6.7.

6.3.B Different Kinds of Lenses

The principles that we've just described are incorporated into different kinds of lenses used in the TEM. Most lenses in the microscope are weak lenses with large gaps. Either they act to demagnify the source image onto the specimen or they magnify the image or DP from the specimen and project it onto the viewing screen or CCD in ways that we'll see in Chapter 9. Typically these lenses are of the sort shown schematically in Figure 6.6. An aperture can be introduced into the bore of the lens, as we'll discuss later.

PRACTICAL HINT

You should be able to get a readout (on the TEM computer display) of the current through any lens coil. It is a useful thing to know the standard lens currents for your common operating modes such as imaging and diffraction and for creating various beam sizes.

Compared to the other lenses in a TEM, the objective lens is a very strong lens. Several types exist, depending on the needs of the particular TEM. The most flexible objective lens is that in which the upper and lower polepieces are separated and have their own coils as shown in Figure 6.8A. This geometry gives the space needed to allow us to insert both the specimen and the aperture between the polepieces. With this type of polepiece, other instruments such as X-ray spectrometers can have relatively easy access to the specimen. For the same reason, it is straightforward to design specimen holders that do a variety of tasks such as tilting, rotating, heating, cooling, straining, etc. This versatility accounts for the popularity of the split-polepiece lens in TEMs.

With split polepieces it is possible from to make the upper polepiece behave differently from the lower polepiece. The most common application of this is to excite the upper-objective polepiece very strongly. This kind of (asymmetrical) lens is ideal for an AEM/STEM because it can produce both the necessary broad beam of electrons for TEM and a fine beam of electrons for AEM and STEM. We'll see how this is accomplished in more detail in Chapter 9.

If high resolution is a major requirement, then we'll see that it is essential to keep the focal length of the objective lens short and this means a very strong lens is

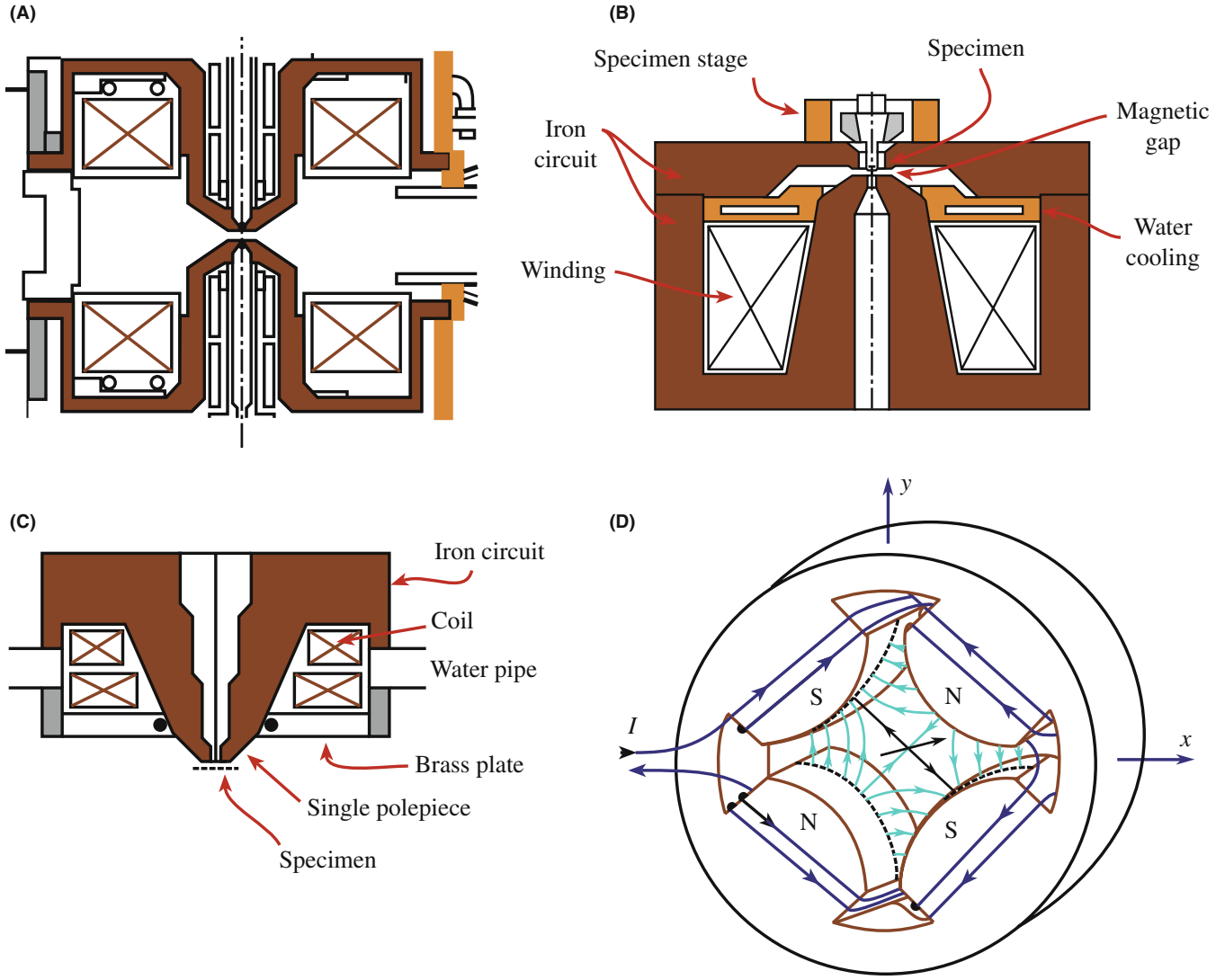


FIGURE 6.8. A selection of different lenses; (A) a split polepiece objective lens, (B) a top-entry immersion lens, (C) a snorkel lens, and (D) a quadrupole lens.

needed. This is traditionally accomplished by using an immersion lens. The specimen is dropped into (i.e., immersed in) the center of the lens field as shown in Figure 6.8B. In such a top-entry stage the specimen is surrounded by the objective lens and so it is a more difficult engineering feat to manipulate, heat or cool the specimen and it is not possible to get X-ray detectors near the specimen, so analytical microscopy is very inefficient. If the focal length is kept really short to give the highest resolution, then it becomes difficult to tilt the specimen more then a few degrees. So in the highest-resolution TEMs you can't do much apart from imaging and diffraction over a restricted range of tilt (see Chapter 8 on stages). This limitation can be overcome by designs such as the snorkel lens as shown in Figure 6.8C, which is a single polepiece lens with a small bore to give a strong lens. Spherical-aberration correction also reduces the need to have the strong lenses for high resolution, so larger gaps are feasible in aberration-corrected TEMs without compromising resolution.

THE OBJECTIVE LENS

The most important lens in the TEM. It forms the images and DPs that are magnified by the other lenses. It is also the most difficult to construct since the specimen must be located close to the plane of this lens.

The limitations of ferromagnetic polepieces can be overcome using superconducting lenses. We cannot make soft-iron polepieces stronger than their saturation magnetization and this limits the focal length and the probe-forming capability of the lens. Superconducting lenses can overcome these limitations but

since a superconductor generates a fixed field, it cannot be varied in the same way as a conventional ferromagnetic lens and so it is not very flexible. Periodically there are increased fluxes of papers describing superconducting lenses because they are small, they don't need water cooling, and they cool the area around the specimen which improves the vacuum, helps minimize contamination, and preserves biological or polymeric specimens. Such lenses also saw a brief flurry of activity after the discovery of high-Tc superconductors. These lenses can generate intense fields (>100 T compared to the maximum of ~2 T in electromagnetic lenses) which are very promising for forming fine probes with high-energy electrons (useful in AEM). Superconducting lenses are so strong that their aberrations (which we'll get to in Section 6.5) are inherently small and they could feasibly be used to construct very compact TEMs.

In addition to these variations on the theme of a single or double polepiece, it is also possible to design a quadrupole, sextupole, or octupole lens in which the focusing action is achieved by four, six, or eight polepieces, respectively. Adjacent polepieces are of opposite polarity as shown in Figure 6.8D. These lenses are not used in TEMs as magnifying lenses but are used to correct lens defects such as astigmatism (see Chapter 9), are used as lenses in aberration correctors (see Section 6.5.A) and also in electron energy-loss spectrometers (Chapter 37). These lenses require less power, and they don't introduce any rotation into the image, which as we'll now show, is a characteristic of standard, electromagnetic lenses.

6.3.C Electron Ray Paths Through Magnetic Fields

We need a bit of mathematics to explain how magnetic lenses actually work. When an electron with charge q (= –e) enters a magnetic field with a strength **B** (Tesla) and an electric field of strength **E**, it experiences a force **F**, known as the Lorentz force, which depends on the velocity of the electron, **v**. All these factors are related through the equation

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) = -\mathrm{e}(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \qquad (6.3)$$

where the term in parentheses is a vector cross-product. Since we are not applying an electric field within the lens, the resulting (Lorentz) force **F** is a vector normal to **v** and **B,** which are inclined to one another at an angle θ. You can easily work out the relative directions of **E**, **v**, **B**, and **F** using the right-hand rule in which your thumb represents the direction of the force acting on a *positive* charge moving in the direction of the middle finger through a field in the direction of the index finger. So the force on the electron acts in the *opposite* direction to your thumb.

RIGHT-HAND RULE
Field: Forefinger
Velocity (Speed): Second finger
Thrust: Thumb

The force on an electron entering a uniform magnetic field, nearly 90° to **B** is

$$F = \mathrm{e}vB\sin\theta = \mathrm{e}vB = \frac{mv^2}{r} \qquad (6.4)$$

where r is the radial distance of the electron from the optic axis (sometimes called the cyclotron radius for historical reasons which you should be able to recognize) and m is the mass of the electron. We can rearrange equation 6.4 to give an expression for

$$r = \frac{mv}{\mathrm{e}B} \qquad (6.5)$$

Since v is a relativistic velocity, we should write this equation as

$$r = \frac{\left[2\mathrm{m}_0 E\left(1 + \frac{E}{2E_0}\right)\right]^{1/2}}{\mathrm{e}B} \qquad (6.6\,\mathrm{A})$$

where m_0 and E_0 are the rest mass and energy of the electron, respectively. This form of the equation allows us to substitute known constants to estimate r (in meters)

$$r = \frac{3.37 \times 10^{-6}\left[V\left(1 + 0.9788 \times 10^{-6}V\right)\right]^{1/2}}{B} \qquad (6.6\,\mathrm{B})$$

In deriving equation 6.4, we made a rather gross oversimplification. If θ equals 90°, the electron is traveling straight down the optic axis and is not focused; in fact it doesn't even notice that a lens is there! It is the *deviation* from θ = 90° that gives the lens effect. The next step, therefore, is to separate the electron velocity **v** in a magnetic field into two components, $\mathbf{v}_1$ perpendicular to, and $\mathbf{v}_2$ parallel to the magnetic-field direction **B**, as shown in Figure 6.9, where $v_1 = v\sin\theta$ and $v_2 = v\cos\theta$. The parallel component, $\mathbf{v}_2$, results in motion parallel to the optic axis in the z direction, with $z = v_2 t$, while the perpendicular component produces circular motion with a radius given by equation 6.5.

THE FIELD
For V = 100 kV and B = 1 Tesla, from equation 6.5 the radius, r, is < 1 mm.

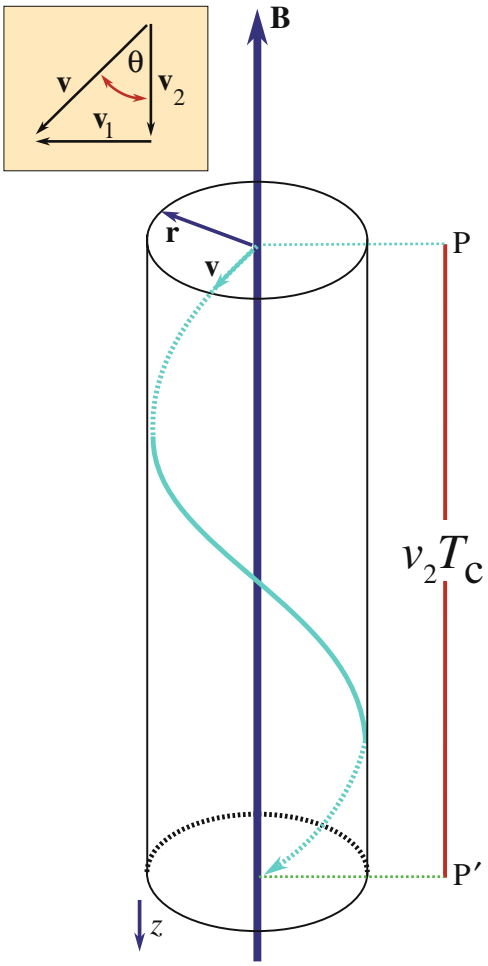


FIGURE 6.9. Electron trajectories in a homogeneous magnetic field, strength **B**. The electrons have velocity components parallel and perpendicular to the field, so long as they are not traveling at 90° to the direction of **B**. The Lorentz force causes electrons passing through point P on the optic axis to spiral through the field and intersect the axis again at P′. The electron's helical path defines the cyclotron radius, *r*.

So all the ray diagrams that we draw ignore this complicating factor which also explains why the optic axis is sometime referred to as the rotation axis. The period of rotation (T_c) through the field gives rise to the (cyclotron) frequency ω_c

$$\omega_c = \frac{2\pi}{T_c} = \frac{eB}{m} \tag{6.7}$$

From these various relationships, we can calculate the complete ray paths through the lens. The most important equations are called the *paraxial* (i.e., near-axis) ray equations. These equations determine both *r* and the angle of rotation (θ) about the axis as the electron moves around the axis in the direction *z*: it rotates under the influence of the rotationally symmetrical field, *B*. These equations, which neglect electron trajectories far off axis, are derived in texts on electron optics. As Hawkes succinctly states "a straightforward, but quite lengthy calculation yields"

$$\frac{d^2r}{dz^2} + \frac{\eta^2 B^2 r}{2\,V^{1/2}} = 0 \tag{6.8}$$

$$\frac{d\theta}{dz} = \frac{\eta B}{2\,V^{1/2}} \tag{6.9}$$

where V is the accelerating voltage of the microscope and η is $(e/2m_0c^2)^{1/2}$. You can see from equation 6.8 that the rate of change of *r* along the optic axis is smaller for more energetic electrons (larger V) and larger for more intense field strengths (larger **B**). Likewise, from equation 6.9, the angular rotation rate increases with increasing field strength and decreases for more energetic electrons.

SPIRAL

The electron spirals through the lens field: a helical trajectory. For electrons with higher keV, we must use stronger lenses (larger **B**) to get similar ray paths.

While these conclusions might be intuitively obvious, the implication is often missed. When we change the TEM accelerating voltage, we change the lenses in the microscope! (Think what this would mean in a VLM.) Therefore, the calibration of the TEM and the lens 'constants' change as we change the kV. Remember the initial paraxial assumption; we'll use non-paraxial rays to explain the effect of spherical aberration on resolution a little later in Section 6.5.A.

While all these ray equations are approximations, they form the basis for more detailed mathematical models of electron motion through lenses (see Chapter 2 in the companion text and URL #2). The more complete models are used in advanced software which simulates the effects of new lens shapes, bore/gap ratios, etc., and has permitted significant advances in the design of lenses to meet the more stringent demands of the latest TEMs.

PITCH OF THE HELIX

When we increase **B**, the pitch of the helical path becomes less if we do not change the energy, because the electrons rotate round the axis more often per unit path length along the axis (*z*).

6.3.D Image Rotation and the Eucentric Plane

So the electrons follow a helical path as they traverse the field along the axis of the lens. This rotation is rarely

shown on standard ray diagrams. You'll see this effect as you operate your TEM because the image or DP rotates on the display screen as you try and focus or if you change magnification. This rotation may require calibration as we'll see in Chapter 9, unless the manufacturer has compensated for it by including an extra lens.

We've already seen in Figure 6.4 that if we change the strength of the lens while keeping d_o fixed, the position of the focal plane and the image plane will also change. Because of this, we have to define a standard object plane for the main imaging lens of the microscope and we call this the *eucentric* plane. Your specimen height should always be adjusted to the eucentric plane because an image of an object in this plane will not move as you tilt the specimen around the primary tilt axis of the holder. (The image will still move if you tilt orthogonally, unless the TEM stage is completely computer controlled to compensate for this.) All other planes in the imaging system are defined with reference to the eucentric plane.

We'll tell you much more about this very important reference plane in Chapter 9.

EUCENTRIC PLANE

If your specimen is in the eucentric plane, then the objective lens strength is always the same when the image on the screen is in focus.

6.3.E Deflecting the Beam

There are many occasions during the operation of the TEM when we want to deflect the beam entering the lens. We may wish to deflect the beam laterally off axis or tilt it to a certain angle with respect to the optic axis. In STEM, these operations are essential to the whole process of forming a scanning image. It is also useful in AEM to be able to blank the beam, i.e., deflect it off axis so it goes into a Faraday cup to measure the current, or to prevent the beam from hitting the specimen when no useful spectroscopic data are being gathered. The way we do this is to apply an electromagnetic field to tilt or traverse the beam or an electrostatic field to blank it. Electromagnetic scan times are of the order of milliseconds while electrostatic blanking can occur in fractions of a microsecond.

Although we are assuming that the lens is thin and has effectively zero thickness along the optic axis, the magnetic field actually acts over a length L. The angle of deflection ε is (for small ε)

$$\varepsilon = \frac{eLB}{mv} \tag{6.10}$$

From this equation we can show that to tilt the beam by 5° we need a coil carrying about 0.2 A and ~100 turns applied along a length of 10 mm, giving a field of 0.01 T. For electrostatic blanking we need about 2 kV/mm.

6.4 APERTURES AND DIAPHRAGMS

We mentioned earlier that an aperture is often inserted into a lens. The aperture limits the collection angle (β) of the lens as shown schematically in Figure 6.10 and such an aperture in the objective lens allows us to control the resolution of the image formed by the lens, the depth of field and the depth of focus, the image contrast, the collection angle of the electron energy-loss spectrometer, the angular resolution of the DP, and so on. In other words, this aperture is important! Physically, the aperture may reside above, in, or below the plane of the lens as we draw it in ray diagrams (but it doesn't really matter since the actual effect will be the same and we've already seen that the electron doesn't care which way it is going). Apertures can also perform other functions, which we'll come across later, such as protecting the specimen from stray radiation in the illumination system, measuring the current in the beam or changing that current.

Usually the apertures are circular holes in metal disks and the disks are made of either Pt or Mo, which are both refractory metals.

A quick word on terminology: While the aperture is the hole in the disk, the metal surrounding the aperture is called the diaphragm (like the variable iris diaphragm in a VLM or your camera). We use the aperture to allow certain electrons to pass through the lens and exclude others by causing them to hit the surrounding diaphragm. This 'aperture/diaphragm' wording, while strictly correct English, is a bit cumbersome, and microscopists tend to be lazy and use 'aperture' in both the correct sense of a hole, but also incorrectly to describe the action of the diaphragm. So we might say that "the objective aperture was used to exclude high-angle scattered electrons from the image" or, as we indicated above, "the aperture protects the specimen from stray radiation" while, strictly speaking, the diaphragm did the excluding and protecting. We'll try to be both consistent and correct in our usage of both terms, but sometimes the precise terminology is awkward.

Diaphragms come in several forms, depending on their function and the particular microscope. They can be either individual disks, each with a particular aperture diameter, or they can be a series of different apertures in a single metal strip (as shown in Figure 6.10). The diameter can be as small as 10 μm, which is about

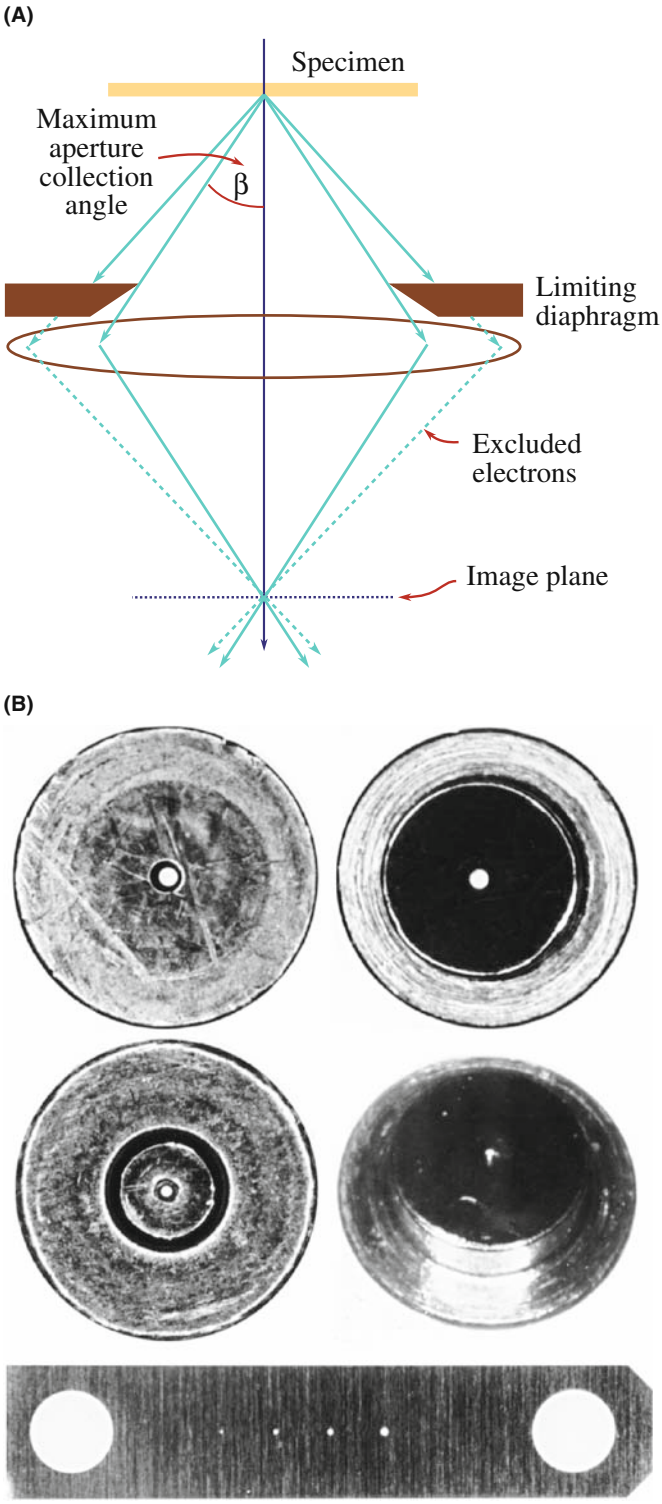


FIGURE 6.10. (A) Ray diagram illustrating how a diaphragm restricts the angular spread of electrons entering the lens. Only electrons emerging from the specimen scattered through angles less than that subtended by the aperture at the object (β) are included in the image-forming process (full lines). The excluded electrons are scattered at angles >β and are stopped by the diaphragm (dotted lines). (B) A selection of diaphragms: the top and middle left are upper and lower views, respectively, of a conventional objective diaphragm; the top/middle right are views of a 'top-hat' (thick) C2 diaphragm; below is a metal strip containing several apertures. Each diaphragm is ~3 mm across.

the smallest circular aperture we can make consistently, or up to ~ 0.3 mm (300 μm). The individual diaphragms or the strips are usually a heavy metal such as Mo or Pt and ~ 25–50 μm thick, but if their job is also to prevent X-rays from hitting the specimen they may be several millimeters thick (see Chapter 33), which means they can be quite expensive if they're made of Pt.

Often the diaphragm collects contamination caused by the electron beam cracking residual hydrocarbons in the vacuum (as we describe in Chapter 8). The contamination tends to accumulate on the edges of the aperture, destroying their circular shape and causing astigmatism. So the diaphragms need occasional cleaning, which can be done by heating them to red heat in the central blue part of a butane flame. In some TEMs, this problem is eliminated by making the diaphragm from very thin metal foil (e.g., Au or Mo). The foil gets hot in the electron beam and any contamination boils off. But such 'self-cleaning' diaphragms are delicate and often crack, thus allowing electrons through other gaps, which defeats the object of the exercise of producing a well-defined aperture.

A SAFETY NOTE ON X-RAYS

X-rays with energies up to the beam energy are generated within the lens wherever the electron beam hits a surface (particularly a limiting diaphragm). So substantial, carefully designed lead shielding is incorporated into the column of the TEM to prevent irradiation of the operator. Obviously, it could be very dangerous to tamper with the lenses or diaphragms of the microscope in any way and only qualified engineers should dismantle, take apart, or repair the lenses or remove the diaphragms.

6.5 REAL LENSES AND THEIR PROBLEMS

It might appear from what we've discussed so far that the analogy between electromagnetic lenses and convex glass lenses is complete, but that is not so. Over the 300 years since van Leeuwenhoek first constructed a light microscope, glass lenses have developed to a point where perfect lenses can be fabricated. In the 80 years since Busch's first magnetic lens, we haven't progressed so far and our lenses are still very imperfect. We've already compared the best electromagnetic lens to the equivalent of using the bottom of a well-known soft-drink bottle as a magnifying glass. Another common description is that if the lenses in your own eyes were as good as our best electromagnetic lens, then you'd be legally blind! So we have to modify all the ideal ray diagrams we've drawn to take into account the imperfections of the lenses. These imperfections all limit the resolution of the microscope but, paradoxically, help us

to get better depth of focus and depth of field from the microscope.

There are many kinds of lens defects (see Chapter 2 in the companion text) and, at one time or another, the effects of all the various defects can be seen in an image or DP. In practice, however, most of us don't need to know about all of them and we'll emphasize the ones that limit the microscope performance in substantial ways. These comprise spherical aberration, chromatic aberration, and astigmatism.

6.5.A Spherical Aberration

The term 'spherical aberration' has almost entered the popular vocabulary since its presence was discerned in the main optics of the Hubble Space Telescope (unfortunately after launch). This defect occurs when the lens field behaves differently for off-axis rays. For our electromagnetic lenses, the further off axis the electron is, the more strongly it is bent back toward the axis. As a result, a point object is imaged as a disk of finite size, which limits our ability to magnify detail because the detail is degraded by the imaging process. As we've told you many times, we can now correct for this aberration, but it still limits the resolution of most TEMs so we need to examine it carefully.

The effects of spherical aberration are shown in Figure 6.11. A point object P is imaged at P′ in the Gaussian image plane. The image is not a point but is instead a central high-intensity region surrounded by a halo of decreasing intensity (similar to Figure 2.11). Spherical aberration is most important in the objective lens because it degrades the detail that we can resolve in TEM images: all the other lenses magnify any error it introduces. It is equally deleterious in the condenser lenses in an AEM or STEM which we use to form the smallest probe with the most current. What we can accomplish at the resolution limits of all forms of TEM is almost always limited by spherical aberration; which is why we're so excited that we can now correct for it.

From Figure 6.11 you can see why we use the term 'spherical' to describe the aberration. The effect of this aberration is to take the curved (spherical) wavefront from the source and increase the curvature. Now if you go back and look at Figure 6.9, you'll see that electrons traveling through a point P on axis intersect the axis again at point P′ where the distance PP′ is given by

$$\mathrm{PP}' = v_2 T_c = v T_c \cos\theta = 2\pi \frac{mv}{eB}$$
$$\left(1 - \frac{\theta^2}{2} + \ldots\right) = L_0\left(1 - \frac{\theta^2}{2} + \ldots\right) \qquad (6.11)$$

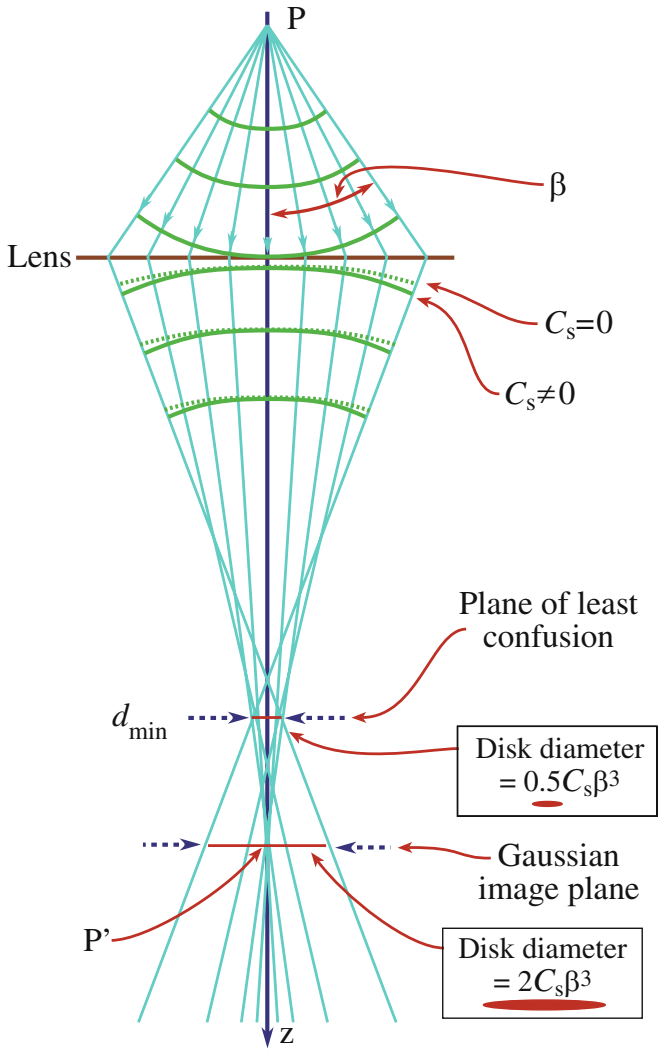


FIGURE 6.11. Spherical aberration in the lens causes wavefronts from a point object P to be spherically distorted by bending the rays at the outside of the lens more than those close to the axis. The point is thus imaged as a disk with a minimum radius in the plane of least confusion and a larger disk at P′ in the Gaussian-image plane. The plane of least confusion is where the smallest image of the object is formed. Schematic intensity distributions at these two important planes are shown beside the ray diagram.

In this relationship, $L_0 = \mathrm{PP}'_0$, where P'_0 is the Gaussian image of the point P for very small θ (i.e., paraxial (near-axis) conditions). As θ increases, the distance PP′ decreases because of spherical aberration and we can write

$$PP' = PP'_0 = -\Delta z \qquad (6.12)$$

where $\Delta z = 0.5 L_0 \theta^2$. Thus we get an expression describing the error, δ, in the Gaussian image position due to spherical aberration

$$\delta = \Delta z \tan\theta \sim \Delta z \theta = 0.5 L_0 \theta^3 \qquad (6.13)$$

So the diameter of the Gaussian image of a point *formed by paraxial rays* is given by this expression, which we will write as

$$\delta = C_s\theta^3 \qquad (6.14)$$

where C_s is a constant (a length) for a particular lens called the spherical aberration coefficient. As you will have noticed, we often use 'C_s' to mean spherical aberration as in 'C_s correction,' 'C_s corrector' and 'C_s corrected.' We'll see in a while that this equation is *very* important because of its effect on the resolution of the TEM and so we need to make a few points of clarification here.

- Equation 6.14 was for paraxial rays only. In a real TEM, the apertures are usually large enough that paraxial conditions do not apply and the sharp image is made more diffuse. As a result the spherically aberrated Gaussian image under non-paraxial conditions is broadened to a diameter of $\delta = 2C_s\theta^3$ (see Figure 6.11).
- You'll often see equation 6.14 written as $\delta = C_s\theta^3 M$ when referring to the image plane but because most discussion of TEM resolution refers back to the minimum distance that we can resolve in the object plane (i.e., the specimen) the magnification term is sometimes left out and we will use this approach.
- When referring to resolution in TEM images it is the *radius* of the point that is more important than the diameter.
- In a real lens the value of θ in equation 6.14 which describes the angle of the electron to the optic axis is replaced by the maximum angle of collection of the (objective-lens) aperture, β.

So, in the forthcoming discussion of resolution, we'll use the radius, we'll refer to the object plane, and we'll use β to define the objective-lens angle of collection. This latter approach is compared to our use of α when discussing beam size in Chapter 5, since α defined the angle of beam convergence. Be careful: many other TEM texts use α somewhat more indiscriminately for both collection and convergence angles.

So we end up with an expression for the radius of the spherically aberrated disk of intensity r_{sph} in the Gaussian image plane, referred back to the specimen plane, under non-paraxial (i.e., realistic) conditions, given by

$$r_{sph} = C_s\beta^3 \qquad (6.15)$$

Because β (in radians) is small, then β^3 is a very strong dependence. The units of r and C_s have to be the same and since C_s is typically a few mm, then we can measure r in (very small fractions of) mm. From this derivation (and linking equations 6.13 and 6.14) you can see that C_s has the dimensions of length and typically it is approximately equal to the focal length of the lens, which for objective lenses in most TEMs is 1–3 mm, but in high-resolution instruments may be well below 1 mm (or not, if they have a C_s corrector).

If you look at Figure 6.11, you will see that the smallest dimension of the cone of rays formed by the lens does not occur at the Gaussian image plane. As we note in the figure, the smallest dimension is formed at a plane closer to the lens which goes by the delightful term 'plane of least confusion' or sometimes 'plane of minimum confusion'; this disk has a radius of $0.25C_s\beta^3$ and a diameter of $0.5C_s\beta^3$. As we'll discuss in Section 6.6.C, some texts use this smaller dimension to define the resolution limit imposed by spherical aberration and it is popular with the TEM manufacturers since it is smaller than the disk in the Gaussian image plane and thus the resolution of the microscope appears better!

CONFUSION

Beware when reading about TEM image resolution because of the confusion between the definition that refers to the Gaussian image plane and that referring to the plane of least confusion.

The way that a corrector compensates for C_s in a magnetic lens is in effect to create a diverging (i.e., concave) lens which spreads out the off-axis beams such that they re-converge to a point rather than a disk in the Gaussian-image plane. In practice this correction is achieved by a highly complex, computer-controlled set of quadrupoles and hexapoles or octupoles. There are two main approaches to the solution of C_s correction. The first is due to the work of Rose and colleagues in Germany embodied in the CEOS commercial system and is used for both probe correction in STEMs and image correction in TEMs. The second is due to Krivanek et al. and is used in Nion dedicated STEMs and has also been retrofitted to several VG STEMs. Figure 6.12 shows schematic ray diagrams for the Nion corrector and the CEOS system. We'll leave a more in-depth discussion of C_s correction to the companion text.

6.5.B Chromatic Aberration

This term is related to the 'color' (i.e., frequency, wavelength, or energy) of the electrons. We've assumed so far that the electrons are monochromatic, but they aren't really. However, we can make very good high-voltage

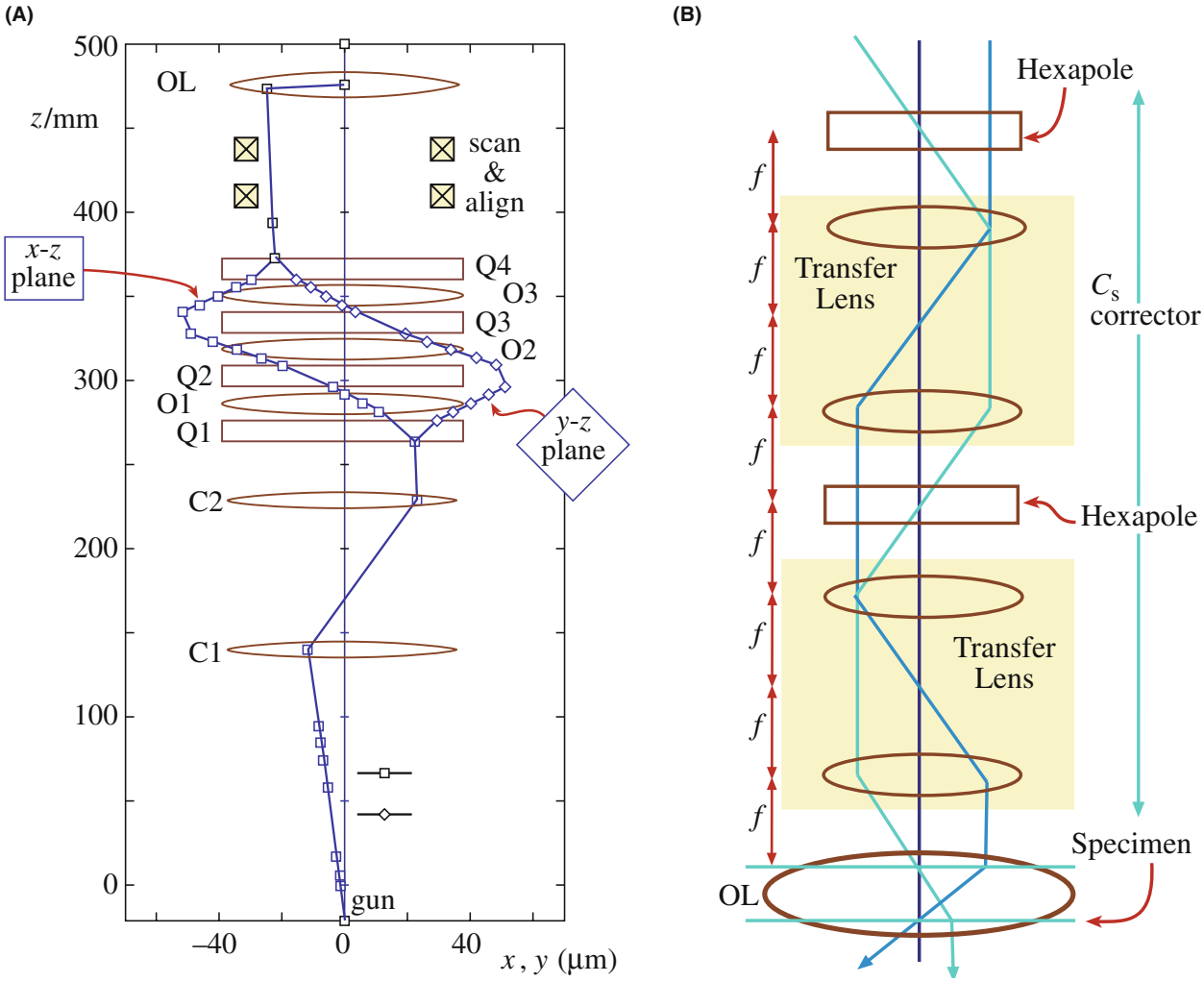


FIGURE 6.12. Ray diagrams showing how the two different commercial systems use (A) multiple quadrupole (Q) and octupole (O) lenses (Nion) or (B) hexapole and other transfer lenses (CEOS) to correct for C_s.

supplies and the variation of the electron energy due to the power supplies is usually smaller than one part in 10^6, which is 0.1 eV for a 100-keV beam. As we discussed in Chapter 5, depending on the electron source the actual energy spread in the beam may vary from ~0.3 eV (cold FEG) to ~1 eV (LaB_6). This range is still so small that we generally don't have to worry about chromatic aberration affecting the image resolution. The exception is if you happen to have a C_s corrector, in which case, after compensating for C_s, C_c is the next most-persistent aberration. Lens elements that can correct for C_c are being developed.

USING A MONOCHROMATOR

Correcting for C_c effects only makes sense if you are dealing with specimens that are thin enough such that specimen-induced chromatic effects do not dominate the resolution. (Correcting C_s is similar.)

Chromatic aberration could be almost completely ignored if we didn't put a specimen into the beam. Unfortunately, this rather essential action creates electrons of a whole range of energies emerging from the thin foil (for reasons we described in Chapter 4). The objective lens bends electrons of lower energy more strongly and thus electrons from a point in the object once again are blurred to form a disk in the Gaussian-image plane (Figure 6.13) (and a smaller disk in the plane of least confusion). The radius r_{chr} of this disk (referring to the object plane) is given by

$$r_{chr} = C_c \frac{\Delta E}{E_0} \beta \tag{6.16}$$

where C_c is the chromatic-aberration coefficient of the lens, ΔE is the energy loss of the electrons, E_0 is the initial beam energy, and β is the angle of collection of the lens. C_c, like C_s, is a length, approximately equal to the focal length. While ΔE in the incident electron beam

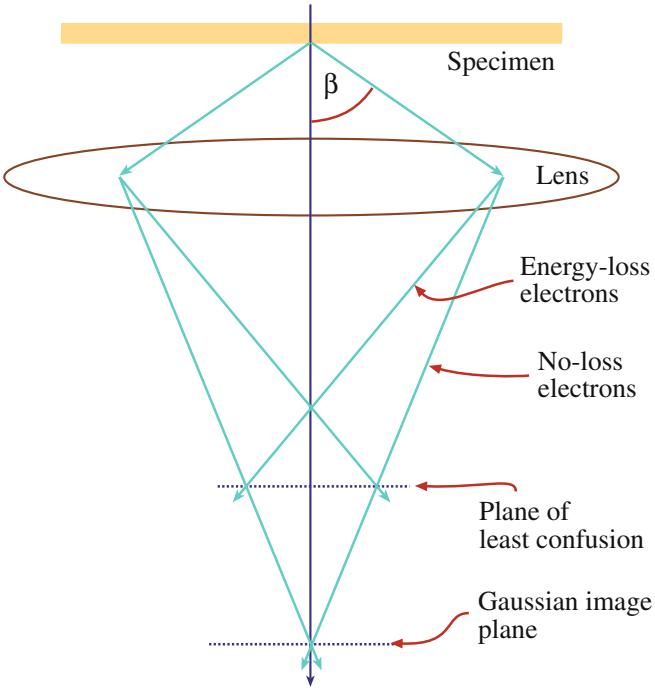


FIGURE 6.13. Chromatic aberration results in electrons with a range of energies being focused in different planes. Electrons emerging from the specimen with no loss of energy are less strongly focused than those that suffered energy loss within the specimen. So, as in Figure 6.11, a point in the object is imaged as a disk in the Gaussian image plane and there is a plane of least confusion.

is $<\sim$1 eV as we just noted, it is typically 15–25 eV for a good fraction of the electrons coming through a typical 50–100 nm thick foil so, as you can easily calculate, r_{chr} is quite a large number (compared to atomic dimensions). Chromatic aberration gets worse for thicker specimens and is better for thinner ones (remember the almost ubiquitous 'thinner-is-better' criterion). So you can do something cheaply to minimize the effects of this aberration; make thin specimens!

The mechanics of C_c correction depend on whether we're trying to compensate for beam effects or specimen-induced effects. We just reminded you that the energy range of electrons coming from the gun is governed by the type of electron source so there are limits depending on which source you have in your TEM. Monochromating the source is a (very expensive) solution. We discuss monochromators at various times in Chapters 37–40 on EELS because that is where such correction pays the greatest dividends.

Unfortunately, for the vast majority of TEM studies, our specimens are not thin enough and, when we have to live with C_c limitations due to having a thick specimen, energy-filtering (EF) is the best solution. EFTEM can correct for the poor resolution that arises when we form images or DPs with electrons that have lost substantial amounts of energy in the specimen, as we'll also discuss in detail in the EELS chapters.

6.5.C Astigmatism

Astigmatism occurs when the electrons sense a non-uniform magnetic field as they spiral round the optic axis. This defect arises because we can't machine the soft-iron polepieces to be perfectly cylindrically symmetrical down the bore. The soft iron may also have microstructural inhomogeneities which cause local variations in the magnetic field strength. Even if these difficulties were overcome, the apertures we introduce into the lens may disturb the field if they are not precisely centered around the axis. Furthermore, if the apertures are not clean, the contamination charges up and deflects the beam. So there are a variety of contributions to astigmatism, which distorts the image by an amount r_{ast} where

$$r_{ast} = \beta \Delta f \tag{6.17}$$

and Δf is the maximum difference in focus induced by the astigmatism. Fortunately, astigmatism is easily corrected using stigmators, which are small octupoles that introduce a compensating field to balance the inhomogeneities causing the astigmatism. There are stigmators in both the illumination (condenser lenses) system and the imaging system (objective lens) and we'll describe how to use them in Chapter 9.

In summary, spherical and chromatic aberration and astigmatism are the three major defects in electromagnetic lenses. There are several minor defects, such as barrel and pincushion distortion, which are self-explanatory in terms of how they distort the image. They are occasionally seen at very low magnification where electrons traveling well off axis and close to the bore of the polepiece appear in the image. Other defects such as coma, and field curvature we will ignore for now.

Again, if you want to learn more about any of these defects they are covered in Chapter 2 in the companion text.

6.6 THE RESOLUTION OF THE ELECTRON LENS (AND ULTIMATELY OF THE TEM)

Another note on terminology: We electron microscopists tend to be rather imprecise in our definition and use of the words 'resolution' and 'resolving power' and related expressions. We've borrowed these terms from classical VLM, which is concerned with the imaging of incoherent light waves through amplitude contrast. High-resolution performance in the TEM is a different matter and traditionally involves phase-contrast imaging of reasonably coherent electron waves, so perhaps we shouldn't be surprised if a different usage has developed. But we should at least define the terms we use. Now in VLM the word *resolution* strictly applies to the act of displaying fine detail in an *image*. The

resolving power of the microscope is the ability to distinguish in the *image* two points, which are closely adjacent in the object. The minimum distance apart of these points in the *object* is the *minimum-resolvable distance*. Since electron microscopists customarily talk about the resolution of the TEM in terms of distances in the *object* (usually a fraction of a nanometer), we should then use the term minimum-resolvable distance but instead everyone says resolution.

Because the lens defects that we've just discussed cause a point object to degrade into a Gaussian image with a finite radius (some combination of r_{sph}, r_{chr}, r_{ast}) they limit the resolution of the electron lens, and hence that of the microscope. The image resolution in the TEM is governed by the ability of the objective lens to image the object, while in the STEM the image resolution is governed by how much beam current we can put into a small probe which is an image of the electron source demagnified onto the specimen. In either case, aberrations limit the resolution.

RESOLUTION AGAIN

We will use the word resolution, but we define it to mean (usually) the minimum-resolvable distance in the object!

6.6.A Theoretical Resolution (Diffraction-Limited Resolution)

If there are *no* aberrations at all, the resolution of *any* lens (glass, electromagnetic, electrostatic. . .) is customarily defined in terms of the Rayleigh criterion, which we introduced back in Equation 1.1 for light optics. Rayleigh's criterion for resolution is arbitrary in the sense that it is not a fundamental physical rule but more a practical definition. This criterion gives us a figure of merit in terms of our eyes' ability to distinguish separate images of two self-luminous, incoherent point sources.

POINTS BECOME DISKS

A single point source will not be imaged as a point, even if no aberrations or astigmatism are present. The finite size of the lens results in diffraction of the rays at the outermost collection angle of the lens, usually defined by the limiting aperture.

This diffraction results in a point being imaged as a disk (called the Airy disk) which has a schematic cross section intensity profile as shown in Figure 6.14A (and also in Figure 2.11). This effect should be familiar to anyone who has encountered basic physical optics. If the two disks overlap so much that they cannot be resolved as in Figure 6.14B, then the points in the object cannot be resolved. Rayleigh assumed that if the maximum from one source lies over the first minimum from the other source, as shown in Figure 6.14C, then your eye can discern this dip as two overlapping images, thus indicating the presence of two separate objects. Under Rayleigh conditions, when the overall intensity profile exhibits a dip in the middle that is above 80% of I_{max}, the two points cannot be resolved. The separation of the two incoherent point sources is then defined as the theoretical resolution of the lens r_{th} and is given by the radius of the Airy disk, which is similar in form to equation 1.1

$$r_{th} = 1.22\frac{\lambda}{\beta} \tag{6.18}$$

BEWARE!

Sometimes in EM texts you'll find the diameter rather than the radius is used. Reason: the beam diameter defines image resolution in STEM and SEM; in TEM, the radius controls the image resolution. (Hence the factor 1.22 in equation 5.10.)

Any standard text on physical optics (which we've already referenced) will show you how to derive this criterion.

Strictly speaking we should not use this equation for electron sources because they are not incoherent. When dealing with high-resolution images, a different approach is used (see Chapter 28). But for our introductory purposes here, we will be content with this approximation.

From equation 6.18 we see that we can get higher resolution if we lower λ or increase β. This terminology can initially be confusing because, as we just did,

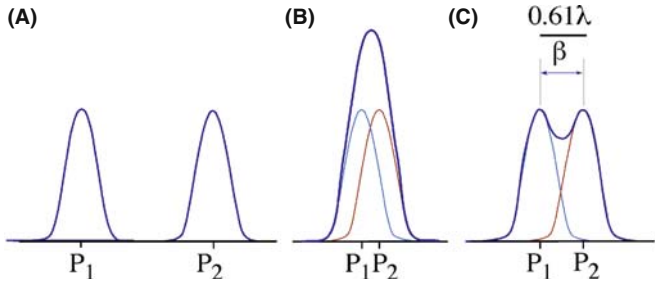


FIGURE 6.14. (A) The Airy-disk intensity profiles from two clearly separated point sources P_1 and P_2. In (B) the two Airy disks are so close that they cannot be distinguished, but in (C) the two are separated such that the maximum in the image of P_1 overlaps the minimum in P_2. This latter situation is the definition of resolution defined by the Rayleigh criterion and is the best (diffraction-limited) resolution.

microscopists often use the expression 'higher resolution' when in fact they mean 'better resolution.' The word higher is then associated with a lower number! It's a smaller r in equation 6.18. It's not just microscopists; a vacuum is also 'higher' if its magnitude is smaller. The improvement in resolution with lower λ is a major reason why there are intermediate and high-voltage TEMs since λ decreases with keV, as we saw back in equation 1.6. The obvious question is, why don't we just increase β (i.e., use a bigger lens aperture or remove it altogether). Well, we could do this if we had perfect lenses, but that isn't the case. All the lens aberrations increase as we increase β (see equations 6.15–6.17); which is why C_s correction is so interesting.

6.6.B The Practical Resolution Due to Spherical Aberration

Let's assume first of all that we have corrected for any astigmatism and our specimen is thin enough that chromatic aberration is negligible. Under these circumstances, the spherical aberration error r_{sph} limits the resolution. Now if you go back and look at equation 6.15 you'll see that r_{sph} increases with the cube of β, a *very* strong dependence. The resolution in the object, then, is given by some combination of the Rayleigh criterion and the aberration error. Hawkes gives a particularly clear description of how this combination leads to a value for the resolution of the microscope. Since this is very often *the* principal figure of merit used when investing hundreds of thousands or even millions of dollars in a TEM, it is essential that you understand that the definition is not exact.

We'll start by summing the radii of the Rayleigh disk and spherical-aberration disk (in the Gaussian image plane) in quadrature (remember it's radii for image resolution, diameters for probe-limited resolutions)

$$r = \left(r_{th}^2 + r_{sph}^2\right)^{1/2} \tag{6.19}$$

Therefore, since both these terms are approximate

$$r(\beta) \approx \left[\left(\frac{\lambda}{\beta}\right)^2 + \left(C_s\beta^3\right)^2\right]^{1/2} \tag{6.20}$$

Since the two terms vary differently with the aperture collection angle β, a compromise value exists when the differential of $r(\beta)$ with respect to β is set to zero and we find that

$$\frac{\lambda^2}{\beta^3} \approx C_s^2\beta^5 \tag{6.21}$$

So we come up with an optimum expression for β which Hawkes (1972) gives as

$$\beta_{opt} = 0.77\frac{\lambda^{1/4}}{C_s^{1/4}} \tag{6.22}$$

The exact value of the numerical factor depends on the assumptions made about the various terms included in the definition of resolution and so is often written simply as ***A***. Sometimes, this compromise value is determined by simply equating the equations for r_{th} and r_{sph} rather than going through the summation in quadrature. A quick calculation for 100-keV electrons (λ = 0.0037 nm) for an instrument with C_s = 3 mm gives a β_{opt} value of ~4.5 mrads.

If this expression for β_{opt} in equation 6.22 is substituted into equation 6.20 we get a minimum value of $r(\beta)$

$$r_{min} \approx 0.91\left(C_s\lambda^3\right)^{1/4} \tag{6.23}$$

This is the expression we want; it gives the *practical* resolution of the TEM.

The numerical factor in equation 6.23 is often written as ***B***. Typically, the value for r_{min} is ~0.25–0.3 nm and the best high-resolution instruments have r_{min} ~0.1–15 nm; 1-Å TEMs are about the best available without C_s correction and about 0.07 nm is (at the time of writing) the best reported resolution with C_s correction. So, as we showed back in Figure 1.2, we can resolve rows of atoms, which in most crystalline materials have a separation close to r_{min} (although low-index planes in some metals are still below this resolution). It's worth noting that since your eyes can resolve a distance of ~0.2 mm, then the maximum useful magnification of the best high-resolution TEM is $\sim 3\times10^6$. Above this magnification, no more detail will be revealed.

Hawkes (1972) reminds us that the decision to add in quadrature back in equation 6.19 was arbitrary, and simply summing r_{th} and r_{sph} is another possible way to determine r_{min} (as we'll see in Section 28.7). But any way you combine the two terms for r (or diameter if you're discussing a probe-limited resolution) leads to expressions that have the same general form as equation 6.22 for β_{opt} and equation 6.23 for r_{min}. In some cases, ***A*** and ***B*** are put equal to unity and not even included, and if you're not pushing any limits in your calculations or experiments this latter approach is a very reasonable approximation.

As we indicated right at the beginning of our discussion, electron microscopists are rather imprecise in our definition of the resolution. However, the resolution is often given as a very precise number!

6.6.C Specimen-Limited Resolution Due to Chromatic Aberration

Remember that we assumed in the previous section that there was no contribution from chromatic aberration. However, if you have a thick specimen then a significant number of electrons will lose 15–25 eV of energy (a typical value of the most probable (plasmon) energy loss; see Figure 4.1). If you put 20 eV into the chromatic-aberration resolution limit given by equation 6.16 you'll find that, at 100 keV with β_{opt} of 4.5 mrads from equation 6.16, the value of r_{chr} is ~2.5 nm.

C_C-LIMITED RESOLUTION

This is typically 10× larger than the C_s-limited resolution. When you're looking through a thick specimen the performance of your TEM is 10× worse than its specified resolution.

If you have a thick specimen, it doesn't matter what voltage you use or how low your C_s is; it doesn't matter if you've got a 1-MeV TEM or access to a C_s corrector; you'll have an image resolution in the 1–3 nm range and you can see all the available information in your specimen at a magnification as little as $\sim 10^5\times$. In fact, the vast majority of all recorded TEM images have C_c-limited resolution: your images will too!

So how thick is thick? Well, it depends on TEM voltage and the mean free path for elastic and inelastic scatter in your specimen, which increases with Z (see Chapter 4). For good high resolution at 100 keV your specimen should be $<\sim$ 30 nm, while at 300 keV you can probably get away with ~50 nm before C_c effects begin to control resolution, assuming $Z <\sim 30$. So for higher-Z specimens the 'thinner is better' axiom is even more important. A more restrictive rule of thumb given by Sawyer and Grubb is that, for biological and polymeric specimens, the resolution limit is about one tenth the specimen thickness. As we noted when we first talked about chromatic aberration, the solution to this problem is in your hands (and in Chapter 10).

6.6.D Confusion in the Definitions of Resolution

If you're new to the subject, you don't have to read this section because it may confuse you still further, but if you've read other TEM texts you may have noticed discrepancies in the definitions of resolution.

We used the expression for r_{sph}, measured at the Gaussian-image plane. Strictly speaking, it is only under ideal conditions (i.e., if $C_s = 0$) that we should use the Gaussian image as a measure of the resolution limited by the lens and it is only really correct to use the Gaussian image under *paraxial* conditions, that is with a *very* small objective aperture. As we've already noted, in the TEM β is usually large enough that paraxial conditions do not apply. So the disk in the plane of least confusion is the relevant feature from which to define the best image resolution, as shown back in Figure 6.11.

If this is so, why did we choose the definition of r_{sph} as the radius of the disk in the Gaussian image plane?

The answer to this question is discussed by Hawkes. Defocusing the image slightly, to bring the plane of least confusion to the Gaussian image plane, will indeed lead to a decrease in the value of the numerical factor in equation 6.23 from 0.91 to 0.43. Hawkes also comments that since this latter value is smaller, manufacturers tend to use it to define the resolution of their instrument! However, this whole treatment of resolution assumes incoherent illumination, which is *not* the case in the TEM. Also, the resolution depends on the contrast in the image and how the lens transfers information from the object to the image. As a result, Hawkes concludes that (see equation 6.23) $\boldsymbol{B} \sim 1$ (from the Gaussian image) is "a more prudent choice" (i.e., closer to reality) than $\boldsymbol{B} = 0.43$ (from the disk in the plane of least confusion) even though, strictly speaking, the plane of least confusion refers to the conditions operating in the TEM.

So it is basically a matter of opinion whether to use the diameter or the radius of the disk in the Gaussian image plane or that in the plane of least confusion. Fortunately, it doesn't really matter too much since, in the end, the choice only alters the value of the numerical terms $\boldsymbol{A}$ and $\boldsymbol{B}$, which we've already mentioned are often approximated to unity anyhow. For example, the value of $\boldsymbol{A}$ will depend on exactly which of the several quoted expressions was used for r_{sph}, e.g., if there was 0.25, 0.5, or 1 in front of $C_s\beta^3$. After these various terms are fed into the equations and the value of β_{opt} is extracted, $\boldsymbol{A}$ only varies by about ±15%. A small variation in $\boldsymbol{B}$ will occur also, for the same reason.

BEWARE!

1. There are inconsistencies in the definition of the terms used to describe the effects of C_s on TEM resolution
2. We use the Gaussian image radius referred back at the object plane, i.e., we use $r_{sph} = C_s\beta^3$.

We have tried to be consistent in our use of the radius of the Airy disk and the radius of the aberration/astigmatism error. Obviously, it doesn't really

matter whether you use the radius or the diameter, so long as you are consistent. Occasionally, however, you may find the Airy disk *radius* is used in combination with the *diameter* of the disk in the plane of least confusion or the Gaussian image plane, so this also contributes much to the discrepancy between various TEM texts.

The question to any student learning HRTEM is: do you know what your resolution really is?

6.7 DEPTH OF FOCUS AND DEPTH OF FIELD

You should have got the message that, because of the poor lens quality we have to use small apertures to minimize their aberrations. This generally means that we cut out many of the electrons that would otherwise be gathered by the lens. Fortunately, our electron sources are so bright that we can live with substantially reduced beam currents hitting our specimen. In fact there are advantages to using small apertures, despite the price we pay in image intensity, probe current, and diffraction-limited resolution. These advantages come in the form of better depth of focus and better depth of field. These terms can be confusing, and once again, the TEM literature is variable. So we need to go back to physical optics to find the correct definition of these terms.

Basically, we are trying to find out how much of the object (the specimen) is in focus at the same time and over what range the image is in focus. (This latter question is irrelevant in SEM and dedicated STEMs without post-specimen lenses where we don't use conventional lenses to form the image, so both terms are equivalent.) In TEM both terms are important.

The depth of field, D_{ob}, is measured at, and refers to, the *object*. It's the distance along the axis on both sides of the object plane within which we can move the object without detectable loss of focus in the image. The depth of focus, D_{im}, is measured in, and refers to, the *image plane*. It is the distance along the axis on both sides of the image plane within which the image appears focused (assuming the object plane and objective lens are fixed). Note that we say "appears" in both cases and this of course also depends on how good your eyes are.

We can derive expressions for these definitions using Figure 6.15. Imagine that ray 1 originates at the highest point up the column where the object can appear to be in focus within the resolution and that this ray arrives at the farthest point down the column where the image can appear to be in focus. Ray 2 represents the other extreme but travels at the same inclination to the optic axis. If these two rays appear to come from the same point (to within the resolution of the lens) the distances d_{ob} and d_{im} correspond to the smallest distances which we can resolve in the object or image, respectively. Note immediately that d_{im} is greater than d_{ob}. Now we can show that angles α_{im} and β_{ob}, which are both small, are given by

$$\alpha_{im} \approx \tan \alpha_{im} = \frac{d_{im}/2}{D_{im}/2} \tag{6.24}$$

and

$$\beta_{ob} \approx \tan \beta_{ob} = \frac{d_{ob}/2}{D_{ob}/2} \tag{6.25}$$

The angular magnification is thus

$$M_A = \frac{\alpha_{im}}{\beta_{ob}} \tag{6.26}$$

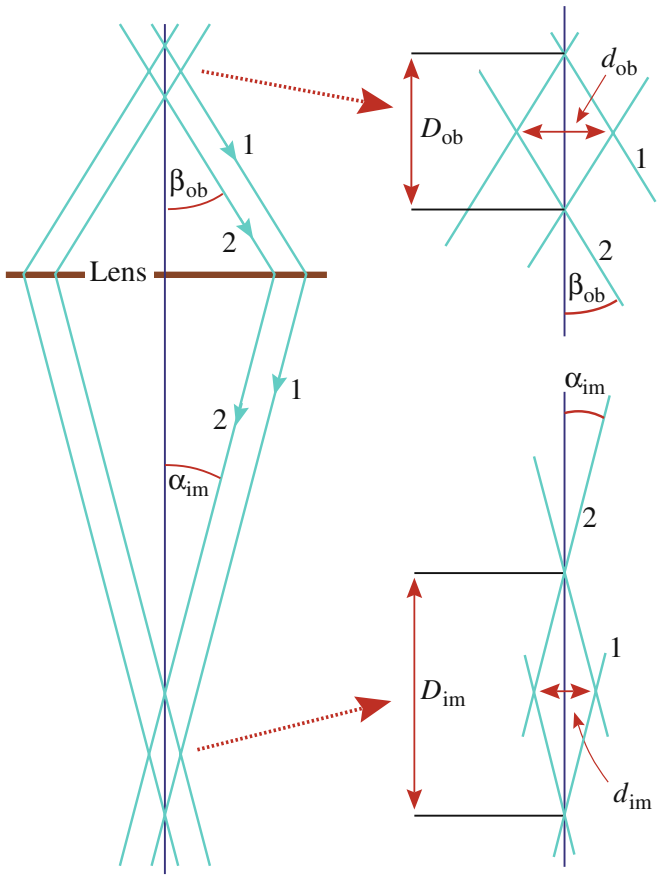


FIGURE 6.15. The definition of the depth of field and the depth of focus. Rays 1 and 2 represent the extremes of the ray paths that remain in focus when emerging $\pm D_{ob}/2$ either side of a plane in the specimen. Typically D_{ob} is greater than the specimen thickness. The same rays define the depth of field over which the image is in focus $\pm D_{im}/2$ either side of the image plane. The resolution in the object is d_{ob} and that in the image is d_{im}.

and the transverse magnification (which we simply call the magnification) is

$$M_T = \frac{d_{im}}{d_{ob}} \tag{6.27}$$

If these two magnifications are related in the usual way by

$$M_T = \frac{1}{M_A} \tag{6.28}$$

Then we can say that the depth of focus is given by

$$D_{im} = \frac{d_{ob}}{\beta_{ob}} M_T^2 \tag{6.29]}$$

and the depth of field is

$$D_{ob} = \frac{d_{ob}}{\beta_{ob}} \tag{6.30}$$

Notice that for a correct calculation of either D_{ob} or D_{im} you must be careful to select the right value of β. Under different circumstances, the limiting angle is defined by the illumination aperture α (in the C2 lens) or the objective aperture β_o (in the objective lens). In thin specimens, which scatter weakly, most electrons emerge from the specimen in a cone closer to that defined by α_{im}, which is often very small ($\sim 10^{-4}$ rad). In a thicker, more strongly scattering specimen, the objective aperture defines the angle and β_o is usually about 10^{-2} rad.

FOCUS AND FIELD

So we get a much greater depth of focus and field by using small apertures (small β).

For a collection angle, β_{ob}, of 10 mrad and a d_{ob} of 0.2 nm, equation 6.30 tells us that the depth of field will be 20 nm, i.e., a specimen of this thickness can all be in focus at the same time. If you only need 2-nm detail in your image, then you can use a specimen which is 200 nm thick and it will still all be in focus.

If we want to see detail at the 0.2 nm level we need to use a magnification of about 500,000×. Equation 6.29 tells us that, under these conditions, the depth of focus will then be 5 km! If we only need to see 2 nm, we can use a magnification of 50,000× and the depth of focus is 5 m. In either case, we have tremendous latitude in where we put the photographic negative or CCD camera because it would still record a focused image many meters either side of the screen. This explains why we can use a CCD camera which can be inserted just below the final projector lens, and still get a focused image with a TV camera well below the standard film camera. In fact, the TEM image would be in focus on the floor under the microscope (or maybe even the floor below if your TEM lab has the misfortune not to be on the lowest floor) if you projected it there but M_T would be different!

Now things get a little more complicated if you're using a C_s-corrected TEM because, for example, in STEM we can use much larger condenser apertures which give much larger convergence angles and thus define the probe a lot more strictly, over distances less than the specimen thickness. In this case, rather than worrying about the reduced depth of field it becomes feasible to think about focusing the probe at different levels within the specimen to explore any structural or chemical variations through the foil thickness. So now we can think about overcoming the projection limitation of TEM images. (Remember the rhino?) The first attempts at imaging single atoms at specific depths within a thin specimen have already been reported.

CHAPTER SUMMARY

We've introduced you to the principles of how an electromagnetic lens works, and how we describe its functions in simple ray diagrams. There are two principal operations: either we use the lens to create an image of an object or we use it bring parallel rays to a focus. We'll see in later chapters that the former operation is used to create magnified images of the specimen on the screen of the TEM. It is also used to create small electron probes (demagnified images of the electron source) at the plane of the specimen in a STEM or AEM. The latter operation is used to create DPs in the back-focal plane of the objective lens.

Our lenses are rather abysmal in their performance, resulting in the need for small limiting apertures. The lens aberrations limit the resolution of the microscope and we usually need an optimum aperture to get the minimum resolution. The small apertures cut down the electron beam intensity, but also give us remarkable depth of focus and depth of

field in our images and specimen, respectively. The recent development of aberration correctors for TEM will revolutionize much of what we've written in this chapter. However, very few TEMs are equipped with correctors and so, for the vast majority of users, it is important to understand the limits imposed on the resolution performance of TEMs by their lenses and by your specimens.

You don't need much skill to use a magnifying glass. The manufacturers may say the same for today's TEM. We say, the more you know about the TEM, the more you'll get out of it and the less likelihood you'll have of making embarrassing errors.

SOME HISTORY

Busch, H 1927 *Über die Wirkungsweise der Konzentrierungsspule beider braunschen Röhre* Arch. Elektrotech. **18** 583–594. The original paper on focusing electron beams.

Hawkes, PW (Ed.) 1997 *Advances in Imaging & Electron Physics Vol. 100: Partial Cumulative Index* Academic Press New York (now published by Elsevier). Essential reference for the historically minded.

Hawkes, PW 2004 *Recent Advances in Electron Optics and Electron Microscopy* Ann. Fond. Louis de Broglie **29** 837–855. An overarching yet concise review of recent advances in electron optics and microscopy, with a great collection of references, both historical and recent.

LENSES AND ELECTRON TRAJECTORIES

Reimer gives a summary of lens defects and more on the derivation of equation 6.11.

Grivet, P 1972 *Electron Optics* Pergamon Press New York.

Hawkes, PW 1972 *Electron Optics and Electron Microscopy* Taylor & Francis Ltd. London. This account is particularly clear if you have an interest in the physics of electron lenses. An important discussion of how to take account of many aberrations when giving a figure of merit. In 'Confusion in the Definitions of Resolution,' we follow Hawkes' clear reasoning regarding the plane of least confusion.

Hawkes, PW (Ed.) 1982 *Magnetic Electron Lenses* Springer New York. A collection of review articles in true Peter Hawkes style; thorough, sound, erudite, and informative.

Hawkes, PW and Kasper, E 1989, 1994 *Principles of Electron Optics* **1**–**3** Academic Press New York. Comprehensive but advanced. Volume **3** includes imaging in the TEM. If by now you're getting the idea that Hawkes is *the* source of electron optical information, then you are right.

Klemperer, O and Barnett, ME 1971 *Electron Optics* Cambridge University Press New York.

Munro, E 1997 *Electron and Ion Optical Design Software for Integrated Circuit Manufacturing Equipment* J. Vac. Sci. Technol. B **15** 2692–2701. More on electrons moving through the lens.

Rempfer, GF 1993 *Electrostatic Electron Optics in the 1940s and Today* MSA Bull. **23** 153–158. By an expert in the use of electrostatic lenses.

ABERRATION CORRECTION

The companion text goes into this in much more detail. In particular, you'll find there that C_s is actually better written as C_3. There are many more 'C_s' terms. These references give an introduction.

Chang, LY, Kirkland, AI and Titchmarsh, JM 2006 *On the Importance of Fifth-Order Spherical Aberration for a Fully Corrected Electron Microscope* Ultramicroscopy **106** 301–306.

Krivanek, OL, Delby, N and Lupini, AR 1999 *Towards Sub-Å Electron Beams* Ultramicroscopy **78** 1–11. Used in the Nioen STEM.

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. Electr. Microsc. **48** 821–826.

RESOLUTION

All texts on TEM will include a discussion of resolution. Particularly useful are those in Reimer 1997, Edington 1976, Fultz and Howe 2002 and Hirsch et al. 1977.

❑ Points to be wary of when reading about definitions of C_s-limited resolution: (see references in Chapter 1)

Sawyer and Grubb (2008) and Egerton 2005 use the Gaussian image radius referred back at the object plane, just as we do; i.e., $r_{sph} = C_s\beta^3$. Reimer 1997 and Fultz and Howe 2001 use the diameter of the disk in the plane of least confusion; i.e., $d_{sph} = 0.5C_s\beta^3$ although both also describe the radius at the Gaussian image plane as we do. Beware: Edington 1976 implies, and Hirsch et al. 1977 state, that $C_s\beta^3$ is the radius *of the disk in the plane of least confusion, which it is not, since by definition it must be less than the Gaussian-image radius (see Figure* 6.11).

Sawyer, LC, Grubb, DT and Meyers, DT 2008 *Polymer Microscopy* 3rd Ed. Springer New York. Rule of thumb for polymers.

❑ Points to be wary of when reading about depth of field and depth of focus

Bradbury et al. 1989 give a particularly clear discussion of the topic. Reimer 1997 uses the term depth of focus for the depth of field and uses depth of image for depth of focus; a rare inconsistency! The terms are used interchangeably in SEM because there is no lens between the object and the image.

Bradbury, S, Evennett, PJ, Haselmann, H and Piller, H 1989 *Dictionary of Light Microscopy* Royal Microscopical Society Handbook #15 Oxford University Press New York. For the VLM comparison.

SPECIAL TECHNIQUES

Borisevich, AY, Lupini, AR, Travaglini, S and Pennycook, SJ 2006 *Depth Sectioning of Aligned Crystals with the Aberration-Corrected Scanning Transmission Electron Microscope* J. Electr. Microsc. **55** 7–12. Moving to confocal imaging in the TEM.

THE COMPANION TEXT

The concepts of depth of field and depth of focus are explored in greater depth in Chapter 2 of the companion text. There, you'll also find more on the principles of lens optics and on Newton's lens equation in particular. Lenses may also be used in other components of the TEM such as electron spectrometers or, in effect, the electron gun; you'll find more on these topics in the same chapter. Quadrupoles, sextupoles, and octupoles are critical components in the correction of aberrations and in some spectrometers; we decided to leave even an introduction to the details until the specialized chapters in the companion text.

URLs

1) http://www.opticsinfobase.org/default.cfm. Optics information base courtesy of the Optical Society of America; lists of papers and journals.
2) http://www.mebs.co.uk/about_us.htm. Munro's company site; provides commercial software for electron optics; essential for the serious designer.

SELF-ASSESSMENT QUESTIONS

Q6.1 How do you focus an image in a TEM?
Q6.2 What kind of visible-light lens does the behavior of a magnetic lens resemble?
Q6.3 Name the main components of a magnetic lens and state their functions.
Q6.4 What are the back and front-focal planes of a magnetic lens?
Q6.5 What do we mean by the term 'optic axis'?
Q6.6 What force acts on an electron in a magnetic field and how can we control this force?
Q6.7 What effect does the magnetic lens have on the trajectory of the electron with respect to the optic axis?
Q6.8 To achieve the highest magnification, where should the specimen be located relative to the objective lens?
Q6.9 Define 'underfocused' and 'overfocused.'
Q6.10 Why is the objective lens the most important lens in a TEM?
Q6.11 Define the eucentric plane.
Q6.12 Explain the difference between a diaphragm and an aperture.
Q6.13 Why do we use apertures in the TEM?
Q6.14 What causes spherical aberration and how can we minimize it?
Q6.15 Define chromatic aberration and describe how to minimize it.
Q6.16 What causes astigmatism and how do we correct it?
Q6.17 Define resolution (strictly speaking, the resolving power) of the TEM?
Q6.18 What ultimately limits the TEM resolution?
Q6.19 In practice, what often limits the practical TEM resolution?
Q6.20 In the TEM, what is depth of field, what controls it, and why is it important?
Q6.21 In the TEM, what is depth of focus, what controls it, and why is it important?

TEXT-SPECIFIC QUESTIONS

T6.1 Estimate the limit of resolution in a 100-kV TEM if the specimen is very thin. Assume $C_s = 1$ mm and $\beta = 10$ mrads.
T6.2 Under the same conditions, estimate the limit of resolution if the specimen is thick enough so that each electron on average undergoes a plasmon loss of ~ 15 eV.
T6.3 If your specimen is pure Al how thick does it have to be such that each electron typically suffers a single plasmon loss. (Hint: go back to Chapter 4.)
T6.4 Go on the Web to find the image resolution offered by commercial manufacturers for a typical 200-kV TEM. (Find Web pages for FEI, Hitachi, JEOL, Zeiss.) Compare the resolution with your answers to questions 1

and 2. What does this exercise tell you about the assumptions that are being made when a TEM resolution is specified?

T6.5 Use suitable values of β_{ob} and α_{im} to deduce values for D_{ob} and D_{im} under the following two conditions: (a) 100-keV electrons, 20 k× magnification, looking for 1 nm detail in the image; (b) 200-keV electrons, 800 k× magnification, looking for 0.2 nm detail in the image.

T6.6 Examine Figure 6.9. Why is B exactly parallel to the optic axis? Is the electron traveling exactly parallel to B? Show, with diagrams, that you understand the reason for these questions.

T6.7 In deriving equation 6.14 we say this is used only for paraxial rays. Why?

T6.8 Explain why the electrons in Figure 6.13 are 'overfocused' if they have lost energy.

T6.9 Discuss the accuracy of using a Faraday cup to measuring beam current.

T6.10 Why do we always draw electron lenses as convex and why haven't we been able to build a concave lens for electrons using cylindrically symmetric lenses? (Hint: recently this problem has, in effect, been solved and has resulted in a solution to the long-standing problem of reducing spherical aberration.)

T6.11 Just for the heck of it try to draw Figure 6.1 to scale assuming that the focal length and object distance are about 3 and 1.5 mm, respectively, and the aperture in the bore of the lens is 60 μm diameter. Estimate values of α and β and explain which of the various lens properties wouldn't be such a good choice for a real lens in TEM.

T6.12 Calculate the radius of the spiral trajectory of 100- and 300-kV electrons in a magnetic field of 1 T.

T6.13 Use ray diagrams to distinguish the terms underfocus and overfocus. Usually it is good to operate any kind of microscope with the lenses in focus. Can you think of any occasion when underfocus (or overfocus) conditions might be useful? (Hint: refer to Figures 6.4 and 6.5.)

T6.14 Why do we use soft magnets and not permanent magnets for electron lenses? If we were in fact to use a permanent magnet what advantages might this bring to the design of a TEM?

T6.15 Why do we have to cool the electron lenses? List as many drawbacks as you can to having to cool the lenses. Can you think of any lenses that might not require cooling? Could we design a TEM in such a way that lens cooling is not required?

T6.16 How could you compensate for the rotation introduced into an electron beam by the action of the lens field? Why would you want to do this?

T6.17 Distinguish the plane of least (or minimum) confusion, the Gaussian image plane, the back-focal plane and the front-focal plane of a lens, using diagrams where necessary. (Hint: Figure 6.11 is a good place to start.)

T6.18 Explain (using diagrams where necessary) why paraxial rays from a point in an object are not subject to significant spherical aberration, yet are still brought to a focus in the Gaussian image plane as a disk rather than as a point.

T6.19 Calculate the radius of the image disk in the plane of least confusion and in the Gaussian image plane under spherical aberration conditions. Assume reasonable values for all terms and justify your assumptions.

T6.20 Calculate the optimum semi-angle of collection for the objective aperture to minimize the contributions of spherical aberration in a 200-keV microscope. State any assumptions. What is the practical resolution of the TEM under these conditions?

T6.21 How is the practical resolution of a TEM further compromised if thick specimens are used? Calculate the expected resolution (i.e., the radius of the disk in the plane of least confusion in the Gaussian image plane) assuming all the electrons suffer an energy loss of $\sim$ 15 eV. Assume reasonable values for all terms and justify your assumptions.

T6.22 Why are the terms *depth of field* and *depth of focus* distinctly differently in TEM, but used interchangeably in SEM?

7

How to 'See' Electrons

CHAPTER PREVIEW

If we are studying the structure of a material, when all is said and done, all we have to show for learning how to operate our expensive TEM, the many hours spent in specimen preparation, etc., is an image or a DP. These images and DPs, which are just different distributions of electron intensity, have first to be viewed in some manner. After viewing, we have to decide if we want to save the results for future reference, perhaps so we can print out the data for a presentation, technical report, or scientific publication. Since, as we noted in the opening chapter, our eyes are not sensitive to electrons, we have to find ways to translate the electron-intensity distributions generated by the specimen into visible-light distributions, which we can see. This chapter will explain how we 'see' electrons.

We'll break the process down into two parts: first, detection (and display) of the image, and second, recording of the image. Both these areas are undergoing rapid change because of ongoing advances in electronic imaging and storage technology, and so this chapter will undoubtedly contain anachronisms by the time you read it. In particular, numbers are favored over photographic data; how can we quantitatively compare two photographs? Comparing two sets of numbers is routine.

7.1 ELECTRON DETECTION AND DISPLAY

As we saw back in Figure 2.1, images and DPs are different kinds of two-dimensional, electron-density distributions which are produced when a thin specimen scatters electrons. We detect and display them in different ways depending on whether we are using a TEM or STEM, as we'll explain in Chapter 9. In a conventional TEM, the images and DPs are static, because the incident beam is fixed, and so we can easily project them onto a viewing screen within the microscope column. TEM images, for example, are *analog* images of electron-density variations in the image plane of the objective lens. We cannot manipulate the image or its contrast in any way between the electrons leaving the image plane and being projected onto the viewing screen. So we will briefly discuss the properties of the viewing screen. The manufacturer controls the choice of screen materials so you might think there's not much need to understand this aspect in any depth. You might be surprised by the limitations you don't need to accept or the improvements which could be made.

When we operate our TEM as a STEM, or we use a dedicated STEM, the image is not static; it is built up over time as the small probe is scanned across the area of interest. Under these circumstances, we can detect the electron signals in several ways. If we are seeking secondary electron (SE) or backscattered electron (BSE) signals, then these detectors sit in the specimen stage area. If we are seeking the same forward-scattered electrons that we view on the TEM screen, the detectors are in the viewing chamber of the TEM. After we've detected any one of these signals, it is usually digitized and the *digital* scanning image is presented on a fluorescent screen as an analog image. You may hear this fluorescent screen referred to as the CRT, which are the initials for cathode-ray tube and a relic from the early days of electron physics. It is becoming much more common for the image or DP to be displayed on a flat-panel screen beside the main TEM column (or even on a plasma or LCD screen on the wall of the EM lab) controlled by the TEM's computer.

We should point out that the sequential or serial nature of the scanning image makes it ideal for on-line image enhancement, image processing, and subsequent image analysis. The signal from any electronic detector can be digitized and electronically manipulated prior to display on the CRT or computer screen, in a way that is impossible with analog images. We can adjust the digital signal to enhance the contrast or to reduce the noise. Alternatively, we can store the digital information and process it mathematically. The availability of cheap memory and

fast computers permits on-line image processing and the rapid extraction of quantitative data from the scanning image; we discuss all this and more in Chapter 31. Because of developments in computer technology, there is great interest in recording analog TEM images via a TV camera in order to digitize them; charge-coupled device (CCD) cameras are readily available for on-line viewing and processing, particularly of HRTEM images. CCD technology is advancing rapidly, driven largely by the digital-camera market and microscopists will continue to benefit from the availability of ever-larger CCD detectors. So we'll spend part of this chapter on CCDs which you'll have now worked out are equally sensitive to visible light and high-energy electrons.

In attempting to compare the properties of detection and recording devices we often use the concept of the 'detection quantum efficiency' or DQE. If a detector is linear in its response then the DQE is defined simply as

$$\mathrm{DQE} = \frac{\left(\frac{S_{\mathrm{out}}}{N_{\mathrm{out}}}\right)^2}{\left(\frac{S_{\mathrm{in}}}{N_{\mathrm{in}}}\right)^2} \qquad (7.1)$$

where S/N is the signal-to-noise ratio of the output or input signal. So a perfect detector has a DQE of 1 and all practical detectors have a DQE <1.

Note on terminology: We use several different terms, often imprecisely, to describe how we 'see' electrons. Since our eyes can't in fact see electrons, we have to resort to the phenomenon of cathodoluminescence (CL) (which we introduced back in Section 4.4) in order to provide an interface between electrons and our eyes. Any electron display system that we look at relies on CL at some point. The CL process converts the energy of the electrons (cathode rays) to produce light (luminescence). As a result, any electron display screen emits light in proportion to the intensity of electrons falling on it. A few definitions are in order

- *Light emission* caused by ionizing radiation is *scintillation*.
- The process of *fluorescence* implies *rapid emission*.
- *Phosphorescence* implies that the wavelength and the *delay time* are longer than for fluorescence.

All these terms are used in electron microscopy (interchangeably and often inaccurately) because the 'fluorescent' screen is coated with a long-delay phosphor (see Chapter 9).

7.2 VIEWING SCREENS

The viewing screen in a TEM is coated with a material such as ZnS, which emits light with a wavelength of ~450 nm. The ZnS is usually modified (doped) to give off green light at closer to 550 nm; hence you'll see screens of different shades of green which, being in the middle of the visible spectrum, is most relaxing for your eyes. As long as sufficient light is emitted, the main requirement of the viewing screen is that the ZnS particle (grain) size be small enough so that your eye cannot resolve individual grains. This means that grain sizes < 100 μm are acceptable (although you can see the grain size if you look at the screen through the auxiliary focusing binoculars). Typical screen coatings are made with a ZnS grain size of ~50 μm, although they may be as small as 10 μm for the highest-resolution screens.

As we've seen in Chapter 4, the cross section for inelastic interactions (and hence the emission intensity of most signals, including CL) decreases with increasing beam voltage. You would thus expect the light intensity to degrade at higher voltages, but this is offset by the increase in gun brightness. In some HVEMs the support for the small focusing screen is made of a heavy metal such as Pt to enhance backscatter and increase screen intensity, Of course, this backscattering will broaden the volume where light is generated and blur the image, so we don't gain very much. In fact most TEMs have very similar screens. Other signals are also given off by the viewing screen, such as X-rays, and whenever you look at the screen you are protected from this lethal radiation flux by lead glass, which is carefully selected to reduce transmitted radiation to levels at or below ambient background. In HVEMs this can amount to several tens of millimeters of glass and, invariably, the optical transmission capabilities are degraded as the glass gets thicker, but obviously we have no alternative if we want to view the screen directly.

A FEW WORDS OF CAUTION ABOUT YOUR SCREEN

There isn't much you can do about choosing the best material for the viewing screen since the manufacturer selects it for you, but you can extend its life substantially by taking care to minimize overexposure. The greatest source of screen damage is the intense direct beam that comes through thin specimens and constitutes the central spot in DPs. Using what you'll learn about operations of the TEM in Chapter 9, you can minimize burning of the screen by (a) only going to diffraction mode with the selected-area aperture inserted, (b) only going to diffraction mode with the C2 lens overfocused, and (c) if the spot appears exceptionally intense despite these precautions, then insert the beam stop while you're observing the pattern on the screen (but not when recording it).

While it is surprising that a modern TEM still relies on an analog screen, the end is perhaps already in sight. One of the latest TEM models (go back and look at Figure 1.9) is built without an operator's viewing screen; all the information is shown on flat-panel computer displays on a console that is separate from the column. Such a design breaks away from more than 70 years of TEM design but has the distinct advantages that

- Anyone in the room (or indeed anyone connected via the Internet) can see the images and DPs, which creates a much better teaching environment.
- The lights don't have to be out to view and record the information.
- The TEM column can be placed in a room that is separated from the operator, whose presence invariably reduces the resolution capabilities of the highest-performance microscopes.

Moving to digital display and recording brings with it the possibility of processing the image or DP to enhance or suppress information prior to publication or presentation. There are obvious ethical considerations here since the scientific community expects that published data be presented with sufficient background information that others would be able to reproduce and cross-check the experiment. So if you process digital images it is wise to publish the unprocessed data at the same time so others can see what data processing has been used. We'll talk a lot more about such ethical issues and related topics when we discuss image processing in Chapter 31.

7.3 ELECTRON DETECTORS

We have several alternatives to the fluorescent screen for detecting electrons. These other electron detectors play a major role in STEMs and AEMs (as well as in SEMs). They are actually essential to the STEM image-forming process that we'll describe in Chapter 9. Such detectors are usually one of two kinds: semiconductor (Si p-n junction) detectors or scintillator-photomultiplier systems. We'll examine the pros and cons of each of these two types and end with a section on CCDs.

7.3.A Semiconductor Detectors

A full understanding of how semiconductor detectors work requires a fair knowledge of solid-state physics. We'll just give a brief outline of the principles as they affect the use of the TEM but if you want to dig deeper the place to start is the excellent text by Pierret.

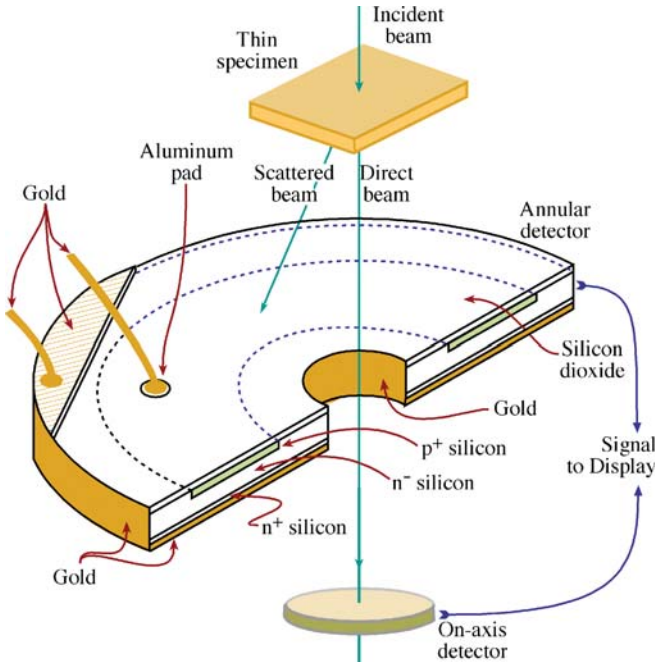


FIGURE 7.1. Semiconductor detector of the surface-barrier type, shown in a configuration where it would be used to detect high-energy, forward-scattered electrons. The direct beam is detected by a small circular detector on the optic axis of the microscope surrounded by a concentric wide-angle annular detector, which detects any scattered electrons.

The semiconductor detector, shown schematically in Figure 7.1, is a doped single-crystal sheet of Si (often inaccurately described as a solid-state detector). We make the Si into an electron-sensitive detector by creating a p-n junction beneath the Si surface in one of two ways. In one type of detector, we create the junction by doping the Si (e.g., by ion implantation of n-type impurity atoms into p-type Si or vice versa). This doping disturbs the equilibrium charge carrier concentration and creates a region across the p-n junction that is free of majority carriers which we call a 'depletion region.' A conducting metal layer is evaporated onto both surfaces to provide ohmic contacts. The alternative type of detector is called a surface-barrier detector (or sometimes a Schottky diode) and we fabricate this by evaporating a thin layer of Au on the surface of high-resistivity n-type Si, or evaporating Al onto p-type Si. This surface layer acts as an electrical contact and also creates a depletion layer and a p-n junction just inside the Si.

When we put either of these detectors into a beam of high-energy electrons, most of the beam energy is transferred to valence-band electrons in the Si which are excited across the band gap into the conduction band thus creating electron-hole pairs (see Figure 4.8). We can separate the electrons and holes most efficiently by applying an external reverse bias to the detector; that is, we put a negative bias on the p side of the junction and a positive bias on the n side. In practice, however, so many electron-hole pairs are created at TEM beam energies

that an external bias is not usually necessary, and the internal bias of the p-n junction acts to separate the electrons and holes. Because the electrons and holes move quite quickly in Si, it takes only a few nanoseconds to gather most of the carriers over an area of $\sim$1 μm^2. So the semiconductor detector is remarkably responsive to electrons. The net result is that the incoming electron signal is converted to a current in the external circuit between the surface contacts, as shown in the surface-barrier detector in Figure 7.1.

Since it takes approximately 3.6 eV to produce an electron-hole pair in Si at room temperature, a 100-keV electron can theoretically produce $\sim$28,000 electrons. This represents a maximum detector gain of close to 3×10^4 but in practice there are losses due to electron absorption in the metal contact layer and recombination of the electrons and the holes close to the Si surface (in a region called the dead layer), and we actually get a gain of closer to 2×10^4.

These semiconductor detectors are very efficient at picking up and amplifying electron signals. Unfortunately, they have an inherently large capacitance, so they are not very responsive to rapid changes in signal intensity. Such changes are quite likely to occur during the rapid scanning process of STEM imaging. In other words, the detector has a narrow bandwidth (typically 100 kHz); this is not a good property for a detector which is subject to widely varying signal intensities. We could lower the capacitance by decreasing the detector area, but if we do this, the signal-to-noise ratio will be lowered. It is the S/N ratio that ultimately limits the quality of all scanning images.

Semiconductor detectors have several advantages

- We can easily fabricate them.
- They are cheap to replace.
- They can be cut into any shape, as long as it is flat.

This latter advantage makes them ideal for squeezing into the confines of TEM stages and columns. For example, we can make the semiconductor detector in annular form so that the main electron beam goes through the hole in it, but the scattered electrons are very efficiently detected. This produces a dark-field (scattered electron) detector. We can also make detectors that are divided into halves or quadrants and each segment is insulated from the other(s). These detectors are very useful for discriminating directional signals such as those coming from magnetic specimens.

There are also some drawbacks to semiconductor detectors

- They have a large dark current (the current registered when no signal is incident on the detector). This dark current arises from thermal activation of electron-hole pairs, or from light falling on an uncoated detector. Since the detectors in a TEM invariably have a metal ohmic contact, the light problem is minimal because light can't penetrate the metal film. Now we could minimize thermal activation by cooling the detector to liquid-nitrogen temperatures but that step is impractical and introduces a cold surface into the vacuum which would simply collect contamination, so we live with noise due to the thermal activation.
- Because noise is inherent in the semiconductor detector, its DQE is poor for low-intensity signals, but rises almost to unity for high-intensity signals.
- The electron beam can damage the detector, particularly in intermediate voltage microscopes. In these circumstances, a doped p-n detector is less sensitive than a surface-barrier detector, because the depletion region is deeper in the Si.
- They are insensitive to low-energy electrons such as secondary electrons.

Despite these drawbacks, both types of Si detector are far more robust than the alternative scintillator, which we will now describe.

7.3.B Scintillator-Photomultiplier Detectors/TV Cameras

A scintillator emits visible light when struck by electrons because of the same CL process that occurs in fluorescent screens. While we are viewing a static TEM image, we want the fluorescent screen to continue emitting light for some time after the electrons hit it, so we use a long-delay scintillator. Of course, when we are using a scintillator to detect rapid changes in signal as in scanning beam imaging, we want the light emission to decay rapidly. So we don't use ZnS in scintillator detectors but rather materials such as Ce-doped yttrium-aluminum garnet (YAG) and various doped plastics and glasses. These materials have decay times on the order of nanoseconds rather than the microseconds needed for ZnS. Once we've converted the incoming electron signal to visible light, the light from the scintillator is amplified by a photomultiplier (PM) system, attached to the scintillator via a light pipe. Figure 7.2 shows a schematic diagram of a scintillator-PM detector setup to detect secondary electrons in a TEM, but the design used to detect primary scattered electrons in the STEM is essentially identical.

The scintillators that we use in STEMs or SEMs are often coated with a 100-nm-thick layer of Al to reflect any light generated in the microscope and stop it from entering the PM tube where it would add noise to the signal. If the detector is in the stage of the microscope, this light could come from the specimen itself if it is cathodoluminescent, or it could be light coming down

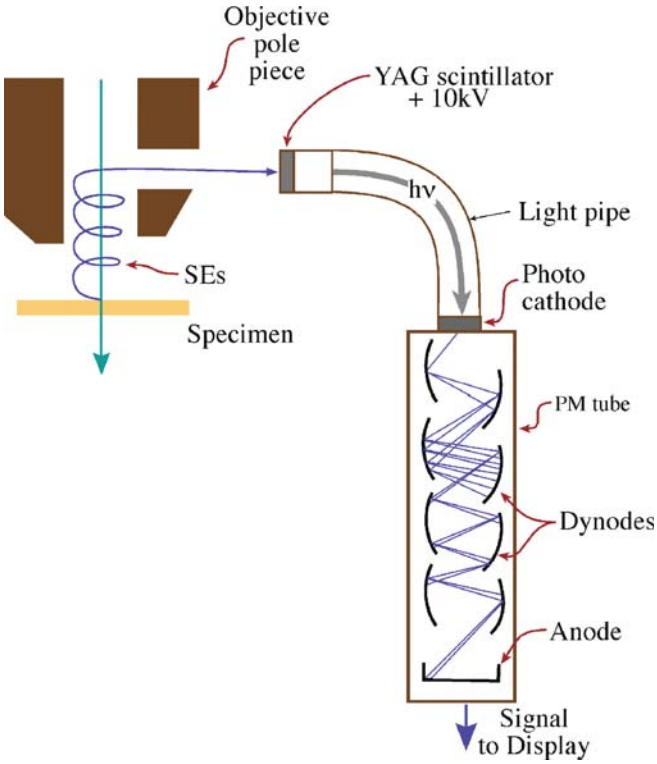


FIGURE 7.2. Scintillator-photomultiplier detector system for SE detection in a TEM. SEs from the specimen spiral back up through the objective lens polepiece and are accelerated by the high voltage onto the scintillator, generating visible light which travels via fiber optics to a photocathode. There the light is re-converted to electrons. The electron signal is then multiplied by several electrodes (dynodes) in the PM tube before being used to modulate the display screen.

the column from a thermionic source and reflected from the polished surface of the specimen. If you have an uncoated scintillator detector in the viewing chamber, then room light may also hit the detector, so you should cover the windows of the viewing chamber.

The advantages of the scintillator-PM system are

- The gain of the system is very high. The gain for the total detector system is of the order of 10^n, depending on the number (n) of electrodes (often called dynodes) in the PM. A value of 10^8 is not unusual (compare with $\sim 10^4$ for the semiconductor detector). This performance is reflected in a typical DQE of close to 0.9 for several commercial scintillators.
- The noise level in a scintillator is low compared with semiconductor detectors, and the bandwidth of the scintillator is in the MHz range. As a result, both low-intensity images and TV-rate images are easily displayed. There is a tremendous practical advantage to TV-rate imaging of digital signals, because such images, when suitably processed and displayed can be viewed, stored, and recorded under normal conditions of room illumination. So you don't have to work in the dark while operating your (S)TEM.

The disadvantages of the scintillator-PM system are

- The scintillator is not as robust as the semiconductor detector, being even more susceptible to radiation damage, particularly after longtime exposure to the beam.
- The scintillator-PM combination is also substantially more expensive and bulky compared to semiconductor detectors and therefore it does not fit well within the TEM stage, nor is it easily manufactured into multi-detector configurations; it is also more expensive. However, plastic scintillators can be shaped to give large collection angle, such as the Robinson BSE detector used in many SEMs.
- The energy-conversion efficiency of a scintillator is also rather low ($\sim$ 2%–20%) compared to a semiconductor detector and, typically, we only get about 4000 photons per incident 100-keV electron, $\sim 7\times$ less than for the semiconductor detector. This low efficiency is offset by the gain in the PM tube.

On balance, the scintillator-PM detector is preferred over the semiconductor detector for most general electron detection in TEM/STEM systems. However, you must take care to minimize any high-intensity beams that may damage the detector and lower its efficiency. Therefore, you need to take more care when operating scintillator detectors.

We've already mentioned that you can view the TEM image directly through a TV camera, rather than looking at the fluorescent screen. There are real advantages to TV cameras, e.g., for on-line viewing of faint HRTEM images (see Chapter 28) or for recording of dynamic in-situ events (see Chapter 29). Also, from a teaching standpoint, anything that pulls the TEM image, in real time, out of the viewing chamber (which only the operator can see into clearly) and onto a classroom or laboratory computer screen or plasma display makes life so much easier. With the increased interest in telemicroscopy, TV cameras and webcams are becoming much more common within the TEM room (e.g., URLs #1 and #2). TV cameras attached to TEMs come in both analog and digital forms. Typically the cameras are placed below the viewing screen so you have to lift up the screen to detect the TV image. The camera may have to be offset if there is another post-column attachment such as an EELS system. Sometimes the camera is placed within the column and is then moved on axis when needed.

Analog YAG-based scintillator TV cameras may be used for applications requiring image intensification if you are dealing with faint images (e.g., if you are using low-dose techniques because your specimen would otherwise be damaged by the beam). Also, wide-angle

cameras are available that collect a much greater area of the image or DP than the standard TEM photographic plate. However, the most widely used TV cameras use only digital-detection technology, which we'll now discuss.

7.3.C Charge-Coupled Device (CCD) Detectors

Electronic technology for recording images and spectra is rapidly closing in on the more traditional analog methods. CCD cameras are becoming the norm for real-time TV recording of images and DPs. They are also being used for two–dimensional arrays for parallel-collection EELS and energy-filtered images, as we describe in Chapter 37.

CCDs are metal-insulator-silicon devices that store charge generated by light or electron beams. CCD arrays consist of several mega- (millions of) pixels which are individual capacitors electrically isolated from each other through the creation of potential wells under each CCD cell, so they can accumulate charge in proportion to the incident radiation intensity, as shown in Figure 7.3A. The largest CCD arrays, as of writing this text are giga-pixels (10^9). Because such systems are so expensive, they are typically developed for use in major astronomical telescopes for detecting very faint light sources. In fact there isn't a good reference text for CCD use in electron microscopy but there is a great one by Howell for astronomy that gives excellent background material if you are interested in digital-image recording.

The maximum size of CCD currently available for TEM use is 4k×4k but the size will only increase with time. (It is also possible to stitch together multiple CCD images using software so, if you have the time, the size of the CCD itself is not a serious issue.) The individual cells currently can be as small as 6 μm although a more typical size range is ~10–15 μm. To create a picture, we have to read out the array. We do this by changing the applied potentials to transfer the charge serially from each potential well along a line in the array into an output amplifier, as shown in Figure 7.3B. With good design of the electrodes, charge-transfer efficiencies of 99.999% can be achieved. Once all the cells are empty the array can be re-exposed. This so-called 'full-frame' design is simple and robust and offers the highest resolution and highest pixel density.

Rather than serial readout of full-frame CCDs, it is also possible to have frame transfer CCDs in which the whole frame is transferred to an adjacent storage array leaving the main array free to collect a new signal flux. This method allows for shorter frame times, and thus a faster image acquisition, but is much more complex and, since more of the device is taken up with the storage, the CCD has lower resolution and much higher cost. The typical frame times needed to record TEM images are long enough that high-speed frame transfer is not usually needed.

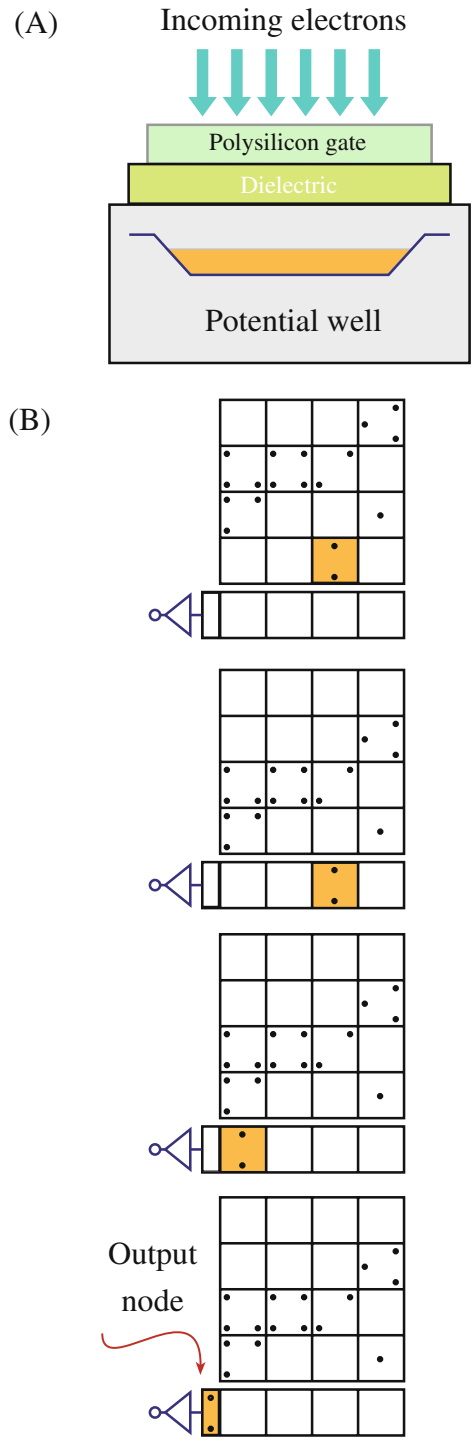


FIGURE 7.3. (A) A single cell in a CCD array showing the storage of charge in the potential well under one pixel. If we vary the applied potential to rows of pixels in sequence as in (B), one pixel row is shifted to the parallel register and is read out pixel by pixel, after which the next row is moved to the parallel register, and so on. The stored charge in each pixel is thus fed into an amplifier and digitized.

The frame time for reading the CCD depends on the size of the image and the specific technology used to readout the detected signal. Ultrahigh-speed CCD cameras are available with $>10^5$ frames/second, but such high speeds are not essential in standard TEM. It is worth noting, however, that time-resolved TEM is an area of growing importance and, in such dedicated instruments, ultra-fast recording is required. Routinely, frame times of <0.001 seconds, well below standard TV rates of 0.033 seconds, are available for in-situ recording of fast events. But the frame time can equally well be several minutes (e.g., for picking up diffuse scattering in faint DPs). Obviously, the longer the exposure, the more the image is susceptible to external vibration, drift, etc.; so long exposures, e.g., for HRTEM images, are not good.

CCD detectors have several advantages

- When cooled, they have very low noise and a good DQE (>0.5) even at low input signal levels.
- The dynamic range is high, making them ideal for recording DPs which can span an enormous intensity range.
- They respond linearly to changes in input signal and show a fairly uniform response across many pixels.

There are some drawbacks to CCDs, not least of which is their cost, but that is always decreasing, as with any Moore's-law-based technology. However, 'blooming,' which occurs when too much signal fills up the pixel and the signal overflows into surrounding pixels, can be a problem. This problem can be overcome substantially by building anti-blooming or overflow drain structures within the device. Apart from these minor factors it is clear that, in the end, CCDs, or other electronic technology, will eventually record and store all TEM images, DPs, and spectra.

7.3.D Faraday Cup

In conventional TEM there isn't much need to know the beam current, but for X-ray analysis in the AEM, it is essential, since there is often a need to compare analytical results obtained under identical beam current conditions. A Faraday cup is a detector that simply measures the total electron current in the beam. We don't use it for any imaging process, but rather as a way of characterizing the performance of the electron source as we saw in Chapter 5. Once the electrons enter the Faraday cup, they cannot leave except by flowing to ground through an attached picoammeter that measures the electron current.

You can easily construct a Faraday cup to go in an SEM, but it is more difficult to design one that fits in the stage of a TEM. A dedicated Faraday cup holder is shown in Figure 7.4A. The entrance aperture is small and the chamber is relatively deep and lined with a low-Z material to minimize backscatter. If you tilt it slightly, the electrons have little chance of being scattered directly back. With such a holder you can only find the hole if you can image the upper surface with SE or BSE detectors, and if these are not available then you must have a cup with a hole in the lower surface too. When the cup is not tilted, the electrons go straight through; if you tilt the cup, then all the electrons are trapped as shown in Figure 7.4A. The way to ensure that you are measuring the maximum current is to look at the picoammeter reading as you tilt the cup. Some manufacturers now incorporate a Faraday cup in the specimen holder. You can measure the current by deflecting the beam into the cup or partially extracting the holder so the beam falls into the cup (Figure 7.4B).

FARADAY CUP

Remember: A Faraday Cup is a Black Hole for Electrons.

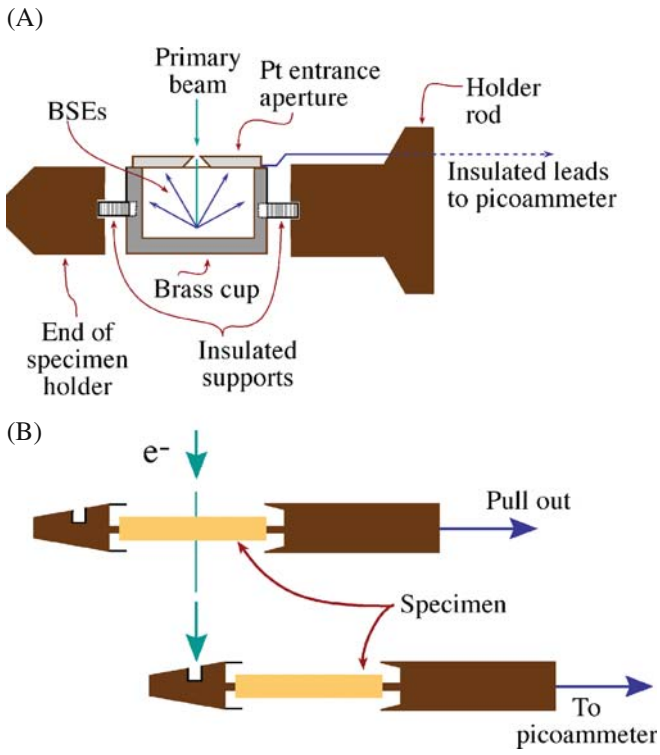


FIGURE 7.4. (A) Schematic diagram of a dedicated Faraday cup in the end of a side-entry specimen holder (more details about these holders in Chapter 8). The entrance aperture has to be found by imaging the top surface using SEs or BSEs. In (B) the specimen holder is retracted so the electrons fall into a cup on the tip of the holder (in which position of course the TEM image of your specimen cannot be seen). In either case, the electron current is measured as it goes to ground through a picoammeter attached to the outside of the holder.

If you don't have a Faraday cup, it is possible to get an approximate reading of the current by just measuring the current through an insulated line from a bulk region of your specimen and correcting for electron backscatter. Backscattering is independent of the accelerating voltage and approximately linear with atomic number up to about $Z = 30$. For example, the backscatter coefficient for Cu is about 0.3 and for Al is about 0.15. It is also possible to deflect the beam onto the last beam-defining diaphragm (see Chapters 6 and 9) and measure the current via an insulated feed-through (also correcting for backscatter).

7.4 WHICH DETECTOR DO WE USE FOR WHICH SIGNAL?

As we mentioned at the start of the chapter, the principal electron signals that we can detect are the forward-scattered electrons (which as we'll see in Chapter 9 form the most common TEM images) and the BSE and SE signals from the beam-entry surface of the specimen.

Semiconductor detectors are only sensitive to electrons with sufficient energy (>5 keV) to penetrate the metal contact layer. So we use these detectors mainly for *high-energy* forward-scattered imaging and *high-energy* BSE imaging. Because of the surface contact layer we don't use semiconductor detectors for *low-energy* SEs but instead use a scintillator-PM system. Remember that the scintillator may also be coated with Al to prevent visible light from generating noise. This coating would also prevent low-energy SEs from being detected. So for SE detection, either there must be no coating or the electrons must be accelerated to an energy high enough to penetrate the coating; we achieve the latter by applying a high kV (>10 kV) positive bias to the scintillator.

The capacitance is relatively high for semiconductor detectors so they are not the detector of choice in dedicated STEMs where high scan-rate TV images are the normal viewing mode, i.e., you need a quick response. The scintillator-PM system is again preferred under these circumstances. As most microscopes move toward TV-rate display of scanning images it is likely that the scintillator-PM will be used increasingly for forward-scattered TEM imaging. Semiconductor detectors may only be used for BSEs, which is not a major imaging mode in TEMs. A summary of all the various electron detectors in a TEM/STEM is given in Figure 7.5. We'll talk more about the methods of imaging in Chapter 9 and later in Part 3 but the primary detectors in (S)TEM pick up the forward-scattered electrons on-axis (called the bright-field (BF) detector), forward-scattered through small angles $<\sim 3^\circ$ (called the annular dark-field (ADF) detector), or scattered out to higher angles (called the high-angle annular dark-field (HAADF) detector).

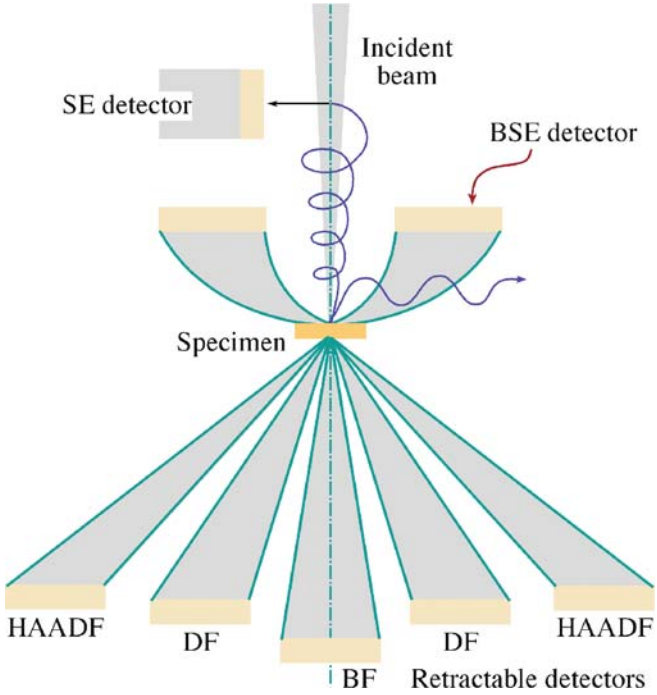


FIGURE 7.5. The various electron detectors in a STEM. Scintillator-PM detectors are invariably used for SE detection and semiconductor detectors for the BSE. The on-axis and annular forward-scattered and high-angle dark field detectors may be either type, depending on the microscope. The SE detector is rare, the BSE detector is a waste of time: only the forward-scattered electron detectors are standard.

Sometimes we examine specimens which themselves exhibit cathodoluminescence under electron bombardment. We discussed CL back in Chapter 4 and we'll give an example of why CL imaging might be useful in Chapter 29. A mirror is used to focus the light from the specimen into a scintillator-PM system; one design is shown in Figure 7.6. In this setup, the specimen must be moved and tilted until, in combination with the collimating lens, the maximum signal is detected by the PMT. This setup effectively prevents detection of most other signals, including X-rays, because the mirror occupies all the free space in the TEM stage. So you have to dedicate the TEM to CL detection alone and ignore other signals. There are only a few such CL-TEMs in the world.

7.5 IMAGE RECORDING

7.5.A Photographic Emulsions

Although photographic film is the oldest recording medium, it still retains sufficient advantages that we continue to use it in some TEMs (and most that are >10 years old). Photographic emulsions are suspensions of silver halide grains in a gel. Electrons strike the

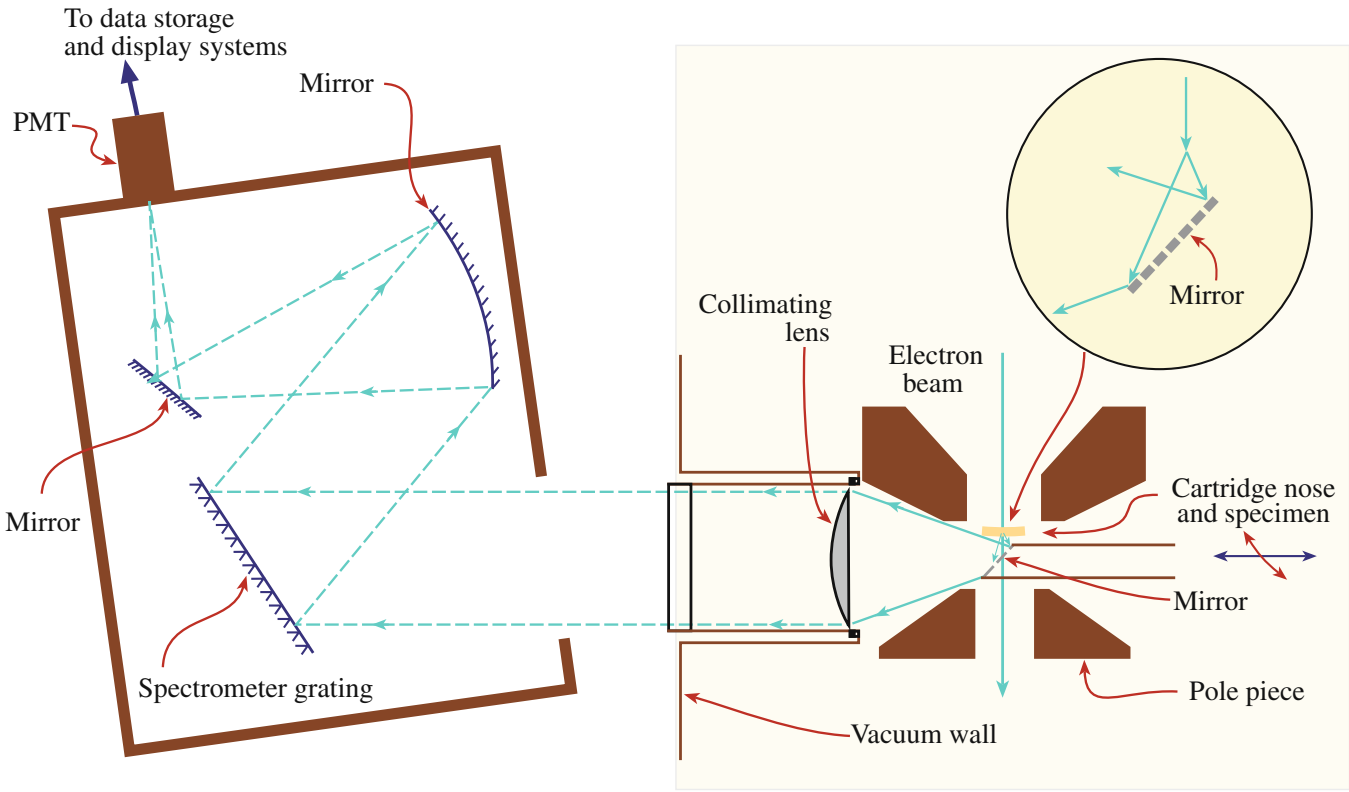


FIGURE 7.6. Cross-sectional diagram of a mirror detector below a thin cathodoluminescent specimen that collects a small fraction of the emitted light and focuses it via a collimating lens into a spectrometer-PM system. The CL signal is usually very weak to start with and so the detector has to be as large as possible, and it takes up much of the free volume in the TEM stage making it impossible to have any other detectors in the stage.

halide, ionize it, and transform it to silver. The emulsions are supported on a polymer film which, unlike the earlier glass plates, outgas and shrink during pre-pumping, or processing. However, glass plates were heavy, occupied an enormous volume compared to film, and Murphy's law meant that your best plates invariably broke because you spent more time handling them. Many aging microscopists still call their EM films 'plates.'

If you are still using film, you have a choice of photographic emulsions, just as you do for your camera (since you'll probably still have a film camera if you're using TEM film). Different speed emulsions are available, with the usual compromise that faster film means a larger grain size and therefore less resolution.

- In principle, for the highest-resolution images, the slowest (finest grain) film is best.
- In practice, we usually minimize the exposure time and go for the fastest film.

We usually want to minimize beam damage and blurring due to movement (drift) of the specimen/stage, so we keep the exposure short. In fact grain sizes for the faster film are about 5 μm compared to about 4 μm for the slowest film, so we don't lose much resolution. The loss of resolution is more than offset by the shorter exposure times which minimize the overall dose to the specimen. The only time you may need to use slower film is if you have a problem with poor image contrast. This problem is more common when imaging amorphous, biological, or polymer specimens.

Although the grain size of the emulsion may be as small as a few micrometers, the actual resolution of the recorded image is worse than this because of electron spreading in the emulsion. The practical resolution may only be about 20–50 μm. Despite this degradation we still have more than 10^7 picture elements or pixels available to store information in a standard 100 mm × 100 mm image (Kodachrome™ film has the silver halide equivalent of 1.8×10^7 pixels). Film has a high DQE, although its dynamic range is rather limited. What this means is that you can easily saturate the film (change all the halide to silver) and you lose any linear relationship between electron intensity and gray scale in the film. As we've already noted, CCDs have a high dynamic range, are linear in response, uniform in output, have anti-blooming technology and the latest CCDs boast $>10^9$ pixels. Perhaps the end is indeed in sight!

Removing photographic film from the TEM will be a major operational improvement, because the absorbed water degrades the vacuum.

PHOTOGRAPHIC FILM

A photographic emulsion on a polymer support is one of the worst things you can put into a high-vacuum instrument.

Both the emulsion and the support outgas, which is a major contribution to the residual high pressure of hydrocarbons and water vapor in the instrument which, in turn, cause contamination of your specimen.

We used to use Polaroid instant film for recording scanning images from the CRT display in STEM mode and there may be a few older instruments around that still require such. But it will generally be good to see the end of any kind of film because they were analog, expensive, used nasty chemicals, and created a lot of mess around the TEM lab. Also there was always the chance that someone before you had loaded the film incorrectly, thus rendering your precious time on the microscope useless and subsequently provoking occasional melees in graduate-student hangouts.

7.5.B Other Image-Recording Methods

Digital images can be stored and retrieved, for example, magnetically on hard drives, or optically on compact discs, DVDs, or, most recently, holographic storage disks. These devices are cheaper and easier to use than photographic recording and images on an optical disc will not degrade with time even after years of storage. However, the usual drawback to digital-storage technology is the question of how long will the technology be around to read any particular stored image. How many images do you still have on a 3.5" floppy? Probably not many! Or a zip drive? Perhaps lots more but you can't read them anymore.

An alternative is the image plate, manufactured by Fuji, which is a kind of re-usable, digital photographic plate. The plates can be used, read, and re-used and this capability permits on-line image processing of the data. While this technology is seeing a lot of use in medical X-ray laboratories, it has not caught on in the TEM field, mainly because it is so expensive.

To present a stored image for publication you still have to print it in some way; photographic methods are still occasionally used. However, laser printers now have the quality required for publishing the highest resolution images (1200 dpi or 48/mm) means that the dot size is well below the resolving power of human eye. Likewise the advent of artificial coloring of electron images, DPs, and spectral images over the last decade or so means that photographic methods of printing are becoming completely obsolete because color photography was never a satisfactory individual lab process.

7.6 COMPARISON OF SCANNING IMAGES AND STATIC IMAGES

We have a choice of creating analog static images in conventional TEM mode or digital scanning images via electronic detection and display. Which is best? While we can only form BSE and SE images in a scanning mode, the answer is not clear for conventional BF and DF images, and the answer depends somewhat on the contrast mechanism that is operating in the specimen, as we'll see in Chapter 22. Regardless of which detector you use, scanning images are always displayed on a computer screen, and this limits the amount of information in the image. Typically the viewing screen will have up to 10^3 lines with a maximum of 10^3 pixels per line, giving a total of 10^6 pixels in each frame. Currently, high-definition TV displays offer at least 1920 pixels per line and 1080 lines per frame giving a total of $2\times\ 10^6$ pixels. In contrast, as we just noted, a TEM image recorded directly onto photographic emulsion will have a higher information density, with $>10^7$ pixels of information available in a 100 mm $\times$ 100 mm image. Furthermore, if a scanning image is to be recorded in a reasonable time, the electron beam can only stay on each point in the image (i.e., each pixel on the display) for a very short time. Typical dwell times per pixel are $<<1$ ms and this means that the signal-to-noise ratio in a scanning image is liable to be quite low. The combination of the lower pixel density compared to a photographic emulsion and the short dwell times means that, almost invariably, STEM images are poorer in quality than static TEM images. However, with the increasing availability of FEG (S)TEMs, increased probe current through C_s correction and improved detection and display technology, STEM digital picture quality compares quite well with analog TEM images. The former technology will continue to improve with the main challenge being our ability to read/access data; the latter will continue to fade away, slowly turning to shades of sepia with increasing time.

CHAPTER SUMMARY

Although it might seem surprising for such a high-end scientific instrument, the TEM is still in the age of analog images. We look at fluorescent screens and computer displays and we still record some of our pictures on photographic film. But darkrooms have disappeared from many labs and the whole area of electron detection is in a state of rapid flux as new electronic technology develops. Semiconductor detectors, scintillators, and CCDs all bring with them the advantage of digital signal collection and therefore the images can be processed and subsequently stored either magnetically or optically. As we said back in 1996, "anything we say about this technology will probably be obsolete before it is published." It is probably safe to speculate that most analog detection, recording, and storage of images and DPs will eventually be replaced by digital methods and the CCD manufacturers are already pronouncing 'the end of film.' So, TEM will produce numbers, but remember that we can all interpret images on film from the 1880s. Can you read data from computer punchcards from the 1970s or from zip disks of the 1990s? A final thought: we encourage you to read the original papers on TEM. All those images were recorded using a photographic emulsion.

REFERENCES

The general references for SEM are the standard book by Goldstein et al. (3rd Ed.) and Reimer's SEM text. The other references here are interesting (often challenging) reading.

Chapman, JN, Craven, AJ and Scott, CP 1989 *Electron Detection in the Analytical Electron Microscope* Ultramicroscopy **28** 108–117.

Howell, SB 2006 *Handbook of CCD Astronomy* 2nd Ed. Cambridge University Press NY.

Knoll, GF 2000 *Radiation Detection and Measurement* 3rd Ed. John Wiley & Sons NY.

Pierret, RF 1996 *Semiconductor Device Fundamentals* Addison-Wesley Boston MA

Reimer, L 1985 *Scanning Electron Microscopy* Springer Verlag New York.

URLs

1) http://tpm.amc.anl.gov/
2) http://telescience.ucsd.edu/gts.shtml

SELF-ASSESSMENT QUESTIONS

Q7.1 How do we 'see' electrons?
Q7.2 What are viewing screens usually coated with?
Q7.3 How do you prolong the lifetime of your screen?
Q7.4 Why is green chosen as the color of the light emitted by the screen?
Q7.5 What is DQE?
Q7.6 What is CL?
Q7.7 How do semiconductor detectors work?
Q7.8 What are the advantages and disadvantages of semiconductor detectors?
Q7.9 How does a scintillator-photomultiplier detector work?
Q7.10 What are the advantages and disadvantages of scintillator-photomultiplier detectors?
Q7.11 What are the advantages of CCD cameras?
Q7.12 What is a Faraday cup?
Q7.13 Which detector should be used for which signal in the TEM?
Q7.14 What is a photographic emulsion, and how does the chemistry work?
Q7.15 What is the difference between scintillation, fluorescence, and phosphorescence?
Q7.16 Why are ZnS-based viewing screens doped?
Q7.17 Why aren't semiconductor devices cooled with LN_2 to reduce the large dark current as they are when used as X-ray detectors (see Chapter 32)?
Q7.18 Give a major disadvantage of photographic film when used in a TEM.
Q7.19 How do you deal with the problem of recording images of specimens with poor image contrast?
Q7.20 What does a 'static' TEM image mean, as opposed to a 'scanning' image?

TEXT-SPECIFIC QUESTIONS

T7.1 Why do we still use analog viewing screens in nearly all TEMs rather than viewing all images and DPs digitally on a computer screen? What would be the biggest advantages of removing the viewing screen?

T7.2 Why do we still sometimes record TEM images and DPs on photographic plates rather than always capturing and storing them digitally?

T7.3 Examine Figure 7.1 and explain why this kind of detector is good for imaging high-angle elastically scattered electrons?

T7.4 Why is this kind of detector (Figure 7.1) not very good at imaging DPs such as those in Figure 2.13B and D?

T7.5 Why would you not use the scintillator-PM detector for the on-axis detector to pick up the direct beam signal in Figure 7.1?

T7.6 Why is the scintillator-PM detector even better than the semiconductor detector for imaging very high-angle, elastically scattered electrons?

T7.7 Go on the Web and find the smallest pixel size available in a commercial CCD camera. How does this dimension compare with the emulsion size in a typical high-resolution photographic film?

T7.8 Go on the Web and find the largest number of pixels available in a commercial CCD camera. How does this number compare with the effective number of pixels in a typical high-resolution photographic film used in a TEM?

T7.9 A problem with recording DPs on CCD cameras is that the intense direct beam can cause 'blooming.' What is blooming and how can it be corrected? (Hint: go on the Web and look for non/anti-blooming cameras.)

T7.10 Why are CL detectors much less common in TEMs than in SEMs? (Hint: take a look at Figure 7.6 and consider the relative sizes of the CL source and the CL detector.)

T7.11 Even if you can, why should you not buy (or bother to use) a BSE detector in your TEM?

T7.12 Go online and see how many TEMs are available for you to watch in someone else's lab? Can you think of specific advantages that could be gained by this accessibility? Can you think of experiments that you could do on remotely accessible TEMs that you could not do in your own laboratory?

T7.13 Why is CL imaging a relatively common imaging method in the SEM but rarely used in the TEM?

8

Pumps and Holders

CHAPTER PREVIEW

In the past three chapters we've described the sources, lenses, and detectors that make up a TEM. The only other parts of the instrument you need to know about in detail are those that, if you are not careful, can seriously degrade the quality of the information you generate even if the rest is perfect. These two parts are the holder in which you put your specimen and the vacuum that surrounds it. While there isn't much you can do to improve the vacuum, beyond buying a better microscope, there is a lot you can do that will degrade the quality of the vacuum in the column and, in doing so, contaminate your specimen. So we'll tell you a few basics about how the vacuum pumps work, and how the vacuum system is put together. Although the vacuum system is under computer control in most TEMs, you still affect the vacuum by what you put in the microscope. Consequently, you need to know what not to do on those occasions when you might otherwise degrade the vacuum.

The vacuum in the stage of a typical TEM is $\sim 10^{-5}$ Pa, compared with atmospheric pressure of $\sim 10^{5}$ Pa. It is quite remarkable that we can transfer a specimen into the TEM, reducing the ambient pressure at its surface by 10 orders of magnitude in a matter of a few seconds. This rapid transfer is a testament to the skills of TEM designers, and particularly the construction of the specimen holder and the airlock system. Specimen holders are the physical contact between you and your specimen across this extraordinary vacuum range. You must transmit all the experimental variables that you want to inflict on your specimen by way of the holder. The most basic requirement is that you should be able to move the specimen laterally to look at different areas; to optimize the imaging you should also be able to move the specimen vertically. In addition we'll describe how you can tilt, rotate, heat, cool, strain, and bias the materials that you are studying. Unfortunately, the holder also transmits vibrations, drift, and contamination to the specimen and may be a source of X-rays that can degrade any analysis that you want to perform. Care of your specimen holders is extremely important since damaged or worn holders reduce the quality of the data generated by the microscope. If you are not careful, a $10,000 holder can easily limit the information generated by a million-dollar TEM.

8.1 THE VACUUM

You know already that electrons are strongly scattered by atoms, which accounts for the versatility of TEM, and the need for thin specimens. Strong scattering also occurs in gases and we can't transmit coherent, controlled electron beams very far through air, so all EMs operate under vacuum. This means that your specimen has to go through an airlock into the TEM. Therefore, you can only control your specimen remotely, not directly, and this makes TEMs more expensive to build. In addition to permitting the electron beam to travel through the instrument undisturbed, the vacuum also plays a role in keeping the specimen clean (or making it dirty). Contamination of the specimen by vacuum-borne contaminants such as hydrocarbons and water vapor can be a problem in many aspects of TEM. Generally, the better the vacuum, the less contamination, but it is the partial pressure of contaminants, not the absolute pressure, which is important. Fortunately, the vacuum systems in most TEMs today are reasonably clean, fully automated and their operation is transparent to the user. Despite this, you should have some understanding of vacuums and how to control them, so this chapter will cover, very superficially, the principles of vacuum systems and pumps.

First of all, a word on units which, as usual, are in disarray. The SI unit of pressure is the pascal; other non-SI units are the torr and the bar. You'll come across all three units in TEM texts and in manufacturers' handbooks, so you need to know the conversions.

We'll mainly use the pascal, but since the torr is still very common terminology, we'll occasionally put approximate torr values in parentheses to remind you of the conversion. Since we deal with very low pressures, the numbers are small, although we perversely use the expression 'high vacuum' for these low pressures. We think of vacuum in terms of rough, low, high, and ultrahigh. A roughing pump gives a pressure between 100 and 0.1 Pa; 0.1–10^{-4} Pa is low vacuum and 10^{-4}–10^{-7} Pa is high vacuum (HV). If the pressure is $<10^{-7}$ Pa you have an ultrahigh vacuum (UHV). These values are approximate, not standardized definitions. A typical modern TEM has a pressure inside the column of $\sim 1.3 \times 10^{-5}$ Pa (10^{-7} Torr), which is in the HV range. UHV TEMs operate below 10^{-7} Pa and the gun region of an FEG TEM operates at $\sim 10^{-9}$ Pa (10^{-11} Torr). To have an electron beam inside the TEM that is not scattered significantly by the air molecules in the column, the pressure must be $<\sim$0.1 Pa. This was achievable with simple mechanical pumps in the early days of the instrument, but there are good reasons to operate at much lower pressures (higher vacuums), for which you need more sophisticated and more expensive apparatus.

PRESSURE

1 Torr is $\sim$ 130 Pa
1 Pa is 7.5×10^{-3} Torr
One bar is atmospheric pressure ($\sim$760 Torr) and is equivalent to $\sim 10^5$ Pa.
The name is the torr; the unit is the Torr, but either way the torr is not an accepted SI unit.
100–0.1 Pa ($\sim$1–10^{-3} Torr) is a rough vacuum.
0.1–10^{-4} Pa ($\sim 10^{-3}$–10^{-6} Torr) is low vacuum.
10^{-4}–10^{-7} Pa ($\sim 10^{-6}$–10^{-9} Torr) is high vacuum (HV).
$<10^{-7}$ Pa ($\sim 10^{-9}$ Torr) is ultrahigh vacuum (UHV).
Be careful when you hear a phrase like "the vacuum in the gun is 10^{-8}" and remember the pascal unit is Pa and the torr unit is Torr.

Generally we use one type of pump to create a rough vacuum and another type to create the higher vacuum. The TEM is kept permanently under vacuum, unless it's being repaired or serviced. If you need access to the inside of the column to change specimens, electron sources, or photographic plates, you do this via an airlock system, which can be pumped separately, as we'll explain later. There are many different kinds of pumps used in TEMs, and you often have a choice when purchasing an instrument. As with most things, you get what you pay for; a clean UHV system is very expensive. We can divide pumps into roughing pumps and HV/UHV pumps, as we'll now discuss.

8.2 ROUGHING PUMPS

The most common roughing pump is a mechanical (rotary) pump in which a belt-driven, eccentrically mounted reciprocating mechanism sucks air through an inlet valve into a chamber and expels it through an exit valve, as shown in Figure 8.1. Such pumps are very reliable, relatively inexpensive, noisy, and dirty and only lower the pressure to $\sim 10^{-1}$ Pa ($\sim 10^{-3}$ Torr). Mechanical pumps should be housed outside your TEM room, and connected to the column through a line that doesn't transmit their vibration. These pumps use a hydrocarbon oil as a medium. If you have such a pump, the line from the pump to the vacuum should contain a foreline trap to condense out oil vapor before it is deposited in the column. Also, the exhaust line from the pump must also be trapped to prevent (possibly carcinogenic) oil vapor escaping into the room where you are working. There are alternative 'dry' roughing pumps which do not use oil. These are more expensive and somewhat less reliable; they do not pump to such low pressure.

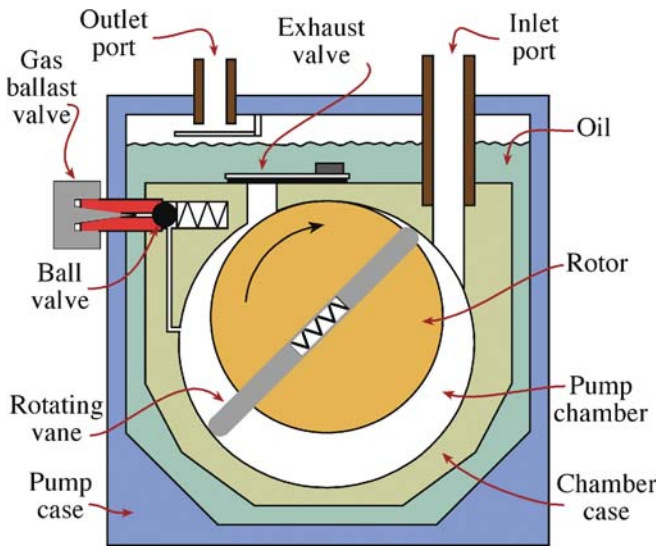


FIGURE 8.1. A mechanical pump for roughing vacuums. The eccentric motion of the pump creates a vacuum in the RH side when it rotates and the vacuum sucks air into the inlet valve. As the cylinder rotates further, it cuts off the inlet and forces the air through the outlet on the LH side, creating a vacuum again on the inlet side as it does so. Because of the constant contact between the rotating cylinder and the inside of the pump, oil is needed to reduce frictional heating.

THE ROOM

The microscope room should be acoustically quiet. You should never feel a draft. All computers with fans should be replaced or moved out of the room—the fan just moves heat around the room. All monitors should be flat panels.

8.3 HIGH/ULTRAHIGH VACUUM PUMPS

8.3.A Diffusion Pumps

These pumps use a hot plate to boil oil, which then forms a series of concentric vapor jets. The jets drag air molecules out of the microscope as shown in Figure 8.2, then condense onto a cold surface, freeing the air molecules which are extracted by the mechanical pump 'backing' the diffusion pump. While this may seem an inefficient way to move air, diffusion pumps can in fact transport more than a hundred liters of air per second, which is quite sufficient to pump out a TEM column. With no moving parts, diffusion pumps are inexpensive and very reliable, but they need external water cooling to aid condensation of the vapor. Failure of the cooling water supply and burnout of the hot plate are about the only possible causes of failure. The absence of moving parts ensures vibration-free operation. As with the mechanical pump, the oil diffusion pump would contaminate the vacuum in the TEM if oil vapor were to escape into the column. To minimize this you must use synthetic nonhydrocarbon oils with low vapor pressures, such as Fomblin™ or Santovac™. (Never use a silicone-based oil, of course.) A liquid-N_2 cold trap sits on top of the pump and condenses out any residual oil molecules. If you have diffusion pumps you must keep the cold traps full of liquid N_2 to maintain a clean system.

Diffusion pumps are capable of very efficient pumping from $\sim 10^{-1}$ to $\sim 10^{-9}$ Pa (10^{-11} Torr) and, if properly trapped, will provide a clean UHV system that is very reliable. The VG series UHV DSTEMs use only oil diffusion pumps to attain UHV conditions.

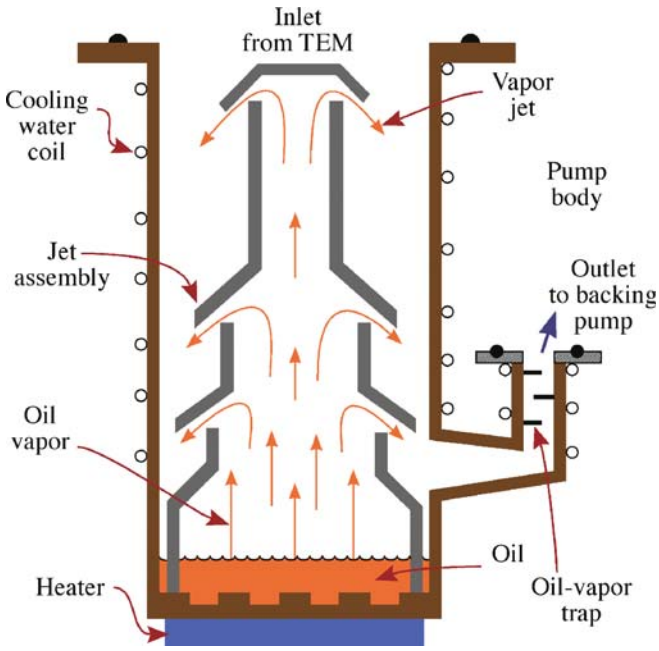


FIGURE 8.2. Principles of diffusion pump operation. A heater plate at the base of the pump boils synthetic oil. The expansion of the oil vapor on boiling creates a pressure which forces the vapor up the central column and out of several holes . The stream of oil vapor pulls gas molecules out of the top of the pump down to the base where the oil condenses and the air is pumped out of the base by a mechanical backing pump.

8.3.B Turbomolecular Pumps

Turbomolecular pumps, or turbopumps, as the name implies, use a turbine to force gases from the microscope. They have many parts moving at high speeds (in excess of 20,000–50,000 rpm is common), so they are more liable to fail than diffusion pumps. The mechanics of the pump are very simple as you can appreciate from Figure 8.3. They do not use oil so they don't introduce hydrocarbons to contaminate the microscope, and the best models (unlike earlier versions) are very quiet and almost vibration-free. In fact, modern turbopumps are being used to prepump the specimen chamber when this is critical, as in the cryotransfer technique (see Section 8.10). If you buy a turbopump, make sure to specify that its use will not transmit vibrations to the TEM column, where it would destroy the image resolution. The turbopump can start (slowly) at

FIGURE 8.3. A turbopump (with and without its casing), which is nothing more than a small turbine that rotates at high speed. Like a jet turbine it pulls air in at the front end and forces it out of the back. The blades are designed like airfoils to enhance the flow of gas through the system.

ambient pressures, increasing speed as the pressure is lowered, ultimately providing UHV conditions at high enough speeds. It is usual, however, to back the turbopump with a dry mechanical pump.

EXHAUST PUMPS

Mechanical, diffusion, and turbopumps are all exhaust pumps; they pull in air from one end and expel it from the other.

8.3.C Ion Pumps

Ion pumps do not contain oil, so they cannot contaminate the TEM column. They also have no moving parts, relying solely on the ionization process to remove air. The ion pump emits electrons from a cathode. These ions spiral in a magnetic field (see Section 6.3) and ionize air molecules, which are then attracted to the cathode. The energetic gas ions sputter Ti atoms from the cathode and they condense throughout the pump chamber, mainly on the cylindrical anode, trapping gas atoms. Thus ion pumps remove gas atoms in two ways; by chemisorption on the anode surfaces and by electrical attraction to the cathodes. The smaller the ion current between the electrodes, the lower the vacuum, so the pump acts as its own vacuum gauge. Ion pumps are only efficient at high vacuums, so they are usually switched on after a diffusion pump has lowered the pressure to $<\sim 10^{-3}$ Pa (10^{-5} Torr). It is common to add ion pumps directly to the stage or gun chambers of TEMs to focus their pumping action on these important regions. Since these pumps are very common on TEMs, we include a diagram (Figure 8.4) showing how they operate.

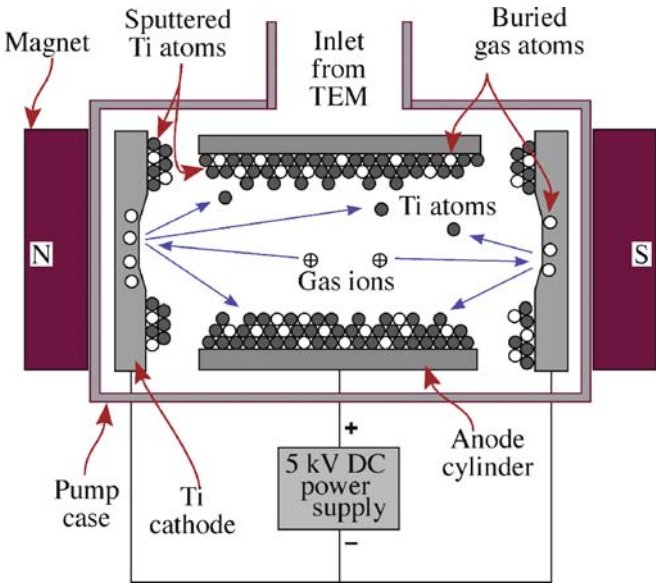


FIGURE 8.4. Schematic diagram showing how ion pumps trap ionized gas atoms by layers of Ti atoms at electrodes. Once trapped, the ions cannot escape until the pump is turned off.

8.3.D Cryogenic (Adsorption) Pumps

As the name implies, cryogenic pumps (cryopumps) rely on liquid N_2 to cool molecular sieves with large surface areas. The cold surface efficiently removes air molecules from ambient pressure down to $\sim 10^{-4}$ Pa (10^{-6} Torr). Because they are oil-free, cryopumps are also used to back ion pumps and prevent their accidental contamination through backstreaming from oil-bearing pumps.

We also use cold surfaces to enhance vacuums in the stage of most non-UHV TEMs. Such 'cold fingers' or 'anticontaminators' provide an alternative site (rather than your specimen) for condensation of residual components in the vacuum.

The same is true if the anticontaminator in your stage is allowed to warm up; then it will degrade the vacuum around your specimen. So you must use another pump such as a diffusion or mechanical pump to remove the air molecules as they are released from captivity. Otherwise, this outgassing will degrade the quality of the vacuum around your specimen, increasing contamination.

TRAPPING PUMPS

Ion pumps and cryopumps are trapping pumps. They keep the air molecules within them and release them when turned off or warmed up, respectively.

8.4 THE WHOLE SYSTEM

As shown schematically in Figure 8.5 the modern TEM has at least two separate pumping systems: one that evacuates the column and one that pumps the camera and screen chamber. We pump the camera separately because the film is one of the primary causes of vacuum degradation since outgassing occurs from the emulsion that contains the AgI grains. So this part of the TEM is usually pumped by a combination mechanical/diffusion pump. The stage is often pumped by a separate ion pump, turbopump, or cryopump, or some combination of these. If the instrument has an FEG, then there is a separate UHV pumping system for the gun region, which often consists of several ion pumps. Each part of the vacuum system consists of roughing pumps (mechanical or turbo) that pump out the appropriate part of the microscope to give the vacuum where the HV/UHV pumps can start to operate.

Looking at Figure 8.5, there are three valves, which are now all computer controlled.

- #1 connects the mechanical pump to the column (the roughing valve).

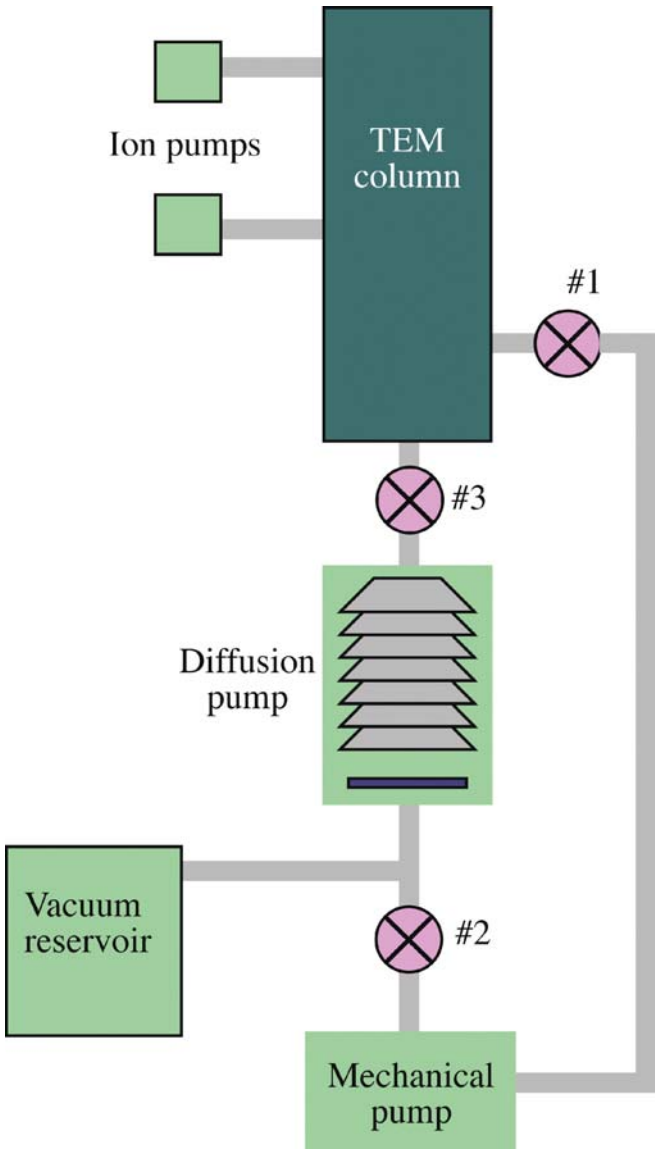


FIGURE 8.5. The principles of the TEM vacuum system. Often, the console display on the TEM will show a similar diagram. The mechanical pump can pump the column directly or back out the diffusion pump, which is connected directly to the base of the microscope. Ion pumps are often interfaced directly to the stage and gun areas. Computer-controlled valves separate the pumps from the column and from each other.

- #2 connects the mechanical pump to the bottom of the diffusion pump (the backing valve).
- #3 connects the diffusion pump directly to the TEM column (the butterfly valve).

If you're pumping down from atmospheric pressure, you first use the mechanical pump to back out the diffusion pump, till it gets to a low enough pressure so its heater can be safely switched on without oxidizing. So close #1, open #2, and close #3.

When the diffusion pump is warmed up, you rough out the column: open #1, close #2, and #3, until the column is at a low enough pressure that the diffusion pump can be used.

At this point, close #1, open #2, and then #3, so the diffusion pump is open to the TEM and may be continuously backed by the mechanical pump. The better approach incorporates a vacuum reservoir between the mechanical and diffusion pumps. When the reservoir is pumped to < 0.1 Pa, the mechanical pump is closed off and the diffusion pump exhausts into the reservoir. When the pressure builds in the reservoir, the mechanical pump will automatically switch on and lower the pressure.

Similar arrangements work for other pumps; e.g., a diffusion pump may be used to lower the pressure in the stage and gun sufficiently for the ion pumps to be switched on, and so on. In most TEMs the stage and gun have significantly better vacuums than the camera region, so the camera/screen is isolated from the rest of the column by a differential pumping aperture (not shown in Figure 8.5). This aperture often coincides with the BFP of the projector lens, since all the electrons have to pass through it and the DP in the BFP localizes all the electron trajectories close to the optic axis. A similar arrangement exists between the stage and gun in FEG systems to preserve the tip in case of a vacuum leak in the stage.

The advent of high-quality digital recording which will remove the need for film in the camera will do more to improve the quality of vacuums in TEMs than any advances in pumping technology.

8.5 LEAK DETECTION

"Nature abhors a vacuum," as Franois Rabelais said in 1534. That's the reason why the pumps must keep pumping: the TEM leaks. But some leaks are too large for the pumps to handle, and then the instrument performance degrades. If you can't run the electron gun, your TEM is useless. Under these circumstances, you have to find the leak, cure it, and repump the instrument, (this is usually a job for your service engineer) but some labs or older TEMs don't have service contracts. Leak detection involves using a mass spectrometer, which can be put into the pumping lines of the microscope. You then release helium gas close to the various parts of the TEM where you suspect a leak (the stage airlock, which sees a lot of use, is a common point of failure). The small He atoms are relatively easily sucked into the column through any leak and register on the mass spectrometer. When a leak is isolated, the TEM may have to be opened to the atmosphere to permit replacement of the defective part, such as the O-ring seals.

The most common cause of a leak is your specimen holder. The O-ring seal on the shaft of a side-entry holder (see the second half of this chapter) is easily contaminated with dust or a hair since it is continually inserted and extracted from the column, and left on the bench while you pore over it. Never touch the O-ring, make sure it doesn't dry out, but if it does (because you don't have someone else looking after the TEM), lubricate it with a very thin film of vacuum grease.

After repairing a leak, when you've pumped down again, it is often useful to 'bake' the column. Baking means heating the internal surfaces to >100°C (or >150–200°C in UHV TEMs) to boil off residual water vapor and hydrocarbons that may have entered the system when it was down to air. Usually, you can achieve the bakeout by leaving the lenses running without their cooling water (check this *very* carefully with the manufacturer before proceeding). In some cases, special heating panels are constructed around the column. Baking can also introduce other leaks as the whole system expands and then contracts, so sometimes leak detection and cure is an iterative process. For UHV systems, you *must* bake to reach the ultimate vacuum, and the higher the temperature the better.

Be wary, however, since sometimes the TEM accessories, such as XEDS and EELS systems, are not designed to be baked to the same high temperature as the column.

8.6 CONTAMINATION: HYDROCARBONS AND WATER VAPOR

As we said right at the start of the chapter, the vacuum (or rather what is still present in the column) can be a source of contamination. Residual hydrocarbons from the pump oil crack under the electron beam. Carbonaceous material then deposits on your carefully thinned specimen, making it difficult to do sensible high-resolution imaging or microanalysis. So a clean vacuum (one in which the hydrocarbon partial pressure is $< 10^{-9}$ Pa) is essential. Fortunately, most modern TEMs are relatively contamination-free, particularly if you use synthetic oils and appropriate traps on the pumps. (See also Section 8.12.)

However, even if you've paid dearly for a clean vacuum system, contamination often occurs; it comes primarily through the airlock with your specimen. You can minimize this by heating the specimen to >100°C in a heating holder or with a halogen lamp in the prepump chamber, or cooling the specimen to liquid-N_2 temperatures in a cooling holder. It may help if the prepump chamber is pumped with an oil-free pump. More recently, plasma cleaning of the specimen holder and specimen prior to insertion in the TEM has proven a very successful way to ensure a clean specimen (more on this in Section 8.12).

Polymers and biological specimens can easily introduce hydrocarbon contaminants, as they outgas in the vacuum, so it is sensible to cool such specimens (since heating or plasma cleaning destroys them). However, when you cool your specimen, it attracts water vapor which condenses as ice on the surface; so load your specimen first, then cool it down in the TEM before you switch on the beam. A low partial pressure of H_2O in the vacuum is obviously essential. Also, warm up any cooled specimens in the TEM before bringing them out to ambient atmosphere, otherwise they will immediately ice up (unless it's a *very* dry winter's day). There will be more about this in the sections on specimen holders.

In addition to the specimen, you personally can be a major source of contamination. Take care never to touch anything that will enter the vacuum, i.e., the specimen itself, the grids, specimen holder (beyond the O-ring seal on the rod), clamping rings, replacement diaphragms, new filaments, replacement Wehnelt, components of XEDS and EELS systems, etc. Use latex gloves whenever you load a specimen, and don't breathe on it. Store specimen holders and specimens in a dry box containing a desiccant such as silica gel, which should be replaced regularly. *Always* prepump fresh film in a vacuum desiccator (which is sometimes integrated into the TEM itself but probably best done elsewhere). Better still, never use film if you can avoid it. Simple precautions like this will minimize contamination of your specimens and the microscope in general and bring a much greater return in terms of good data per TEM session.

8.7 SPECIMEN HOLDERS AND STAGES

To look at your specimen, place it in a specimen holder and insert this assembly into the TEM stage. Therefore, there are two key components which are often not separated, namely, the holder and the stage. In this part of the chapter, we will emphasize the holder but the stage is also critical. Suitable design of the stage is the essential precursor to computer-controlled, remotely accessed TEM, which is already happening.

The cold trap, cold finger, or cryo-blades are a critical part of the stage. Ideally, this cold finger will completely surround the specimen: it cryo-pumps the region around the specimen. However, the cold surfaces, usually brass, provide a source of stray electrons and X-rays which is undesirable for AEM (see Chapter 33), so these blades should be removable for AEM.

X-ray diffractometers use goniometers to hold and tilt the specimen; so do TEMs. Conventional SEMs use a stub on which you mount the specimen so that you can bring the specimen close to the objective lens. However, some high-resolution SEMs use a specimen holder which

is very similar to those used in the TEM, because the specimen is inserted inside the lens, rather than underneath and outside it.

The reason the specimen holder is so important in TEM is that your specimen must invariably be located within the objective lens and the aberrations associated with the objective lens determine the resolution of the TEM.

Historically, microscopists have used two different designs and a lot of what you'll read here or elsewhere has a strong historical background.

- The traditional side-entry holder is a rod with a motor attached to tilt and/or rotate the specimen and a lead connecting it to a power supply and control box, or liquid-N_2 dewar.
- The traditional top-entry holder is a cartridge which you load into the TEM but is detached from the outside world when you use the microscope.

The actual cup that holds your specimen is either 2.3 or 3.05 mm in diameter, so the specimen disk or support grid has to be the same dimension, as we'll see in Chapter 10. The reasons for these dimensions are again partly historical. In the top-entry holder the specimen and part of the holder fit through the bore of the upper polepiece (see Figures 6.7 and 6.8). Clearly, the specimen must be smaller than the bore diameter. So the original top-entry holders used small specimens.

THE WORD 'TRADITIONAL'

Holders are changing; new manufacturers and new capabilities are emerging. Ideally, the holder should not be moved after it is inserted in the TEM.

Side-entry holders are more versatile and larger specimen dimensions first appeared when they were introduced. However, side-entry holders connect the specimen directly to the outside world via a long lever arm, which is undesirable, unstable, and also not necessary in many cases! Ideally, the side-entry holder should leave the specimen in the stage, not connected to the outside world, and all manipulations should be conducted through the stage itself, not the holder. This ideal is being approached as stages become more computer-controlled.

8.8 SIDE-ENTRY HOLDERS

Side-entry holders are now the standard, although their design has changed quite radically in recent years. The traditional design is shown in Figure 8.6. The key parts of this holder are:

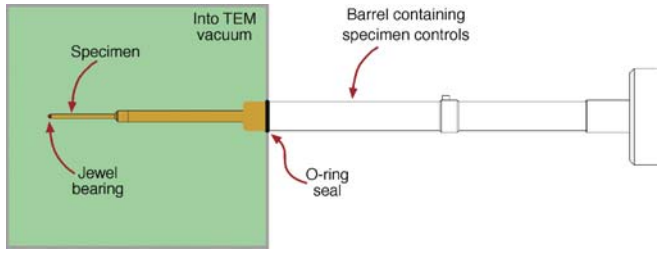


FIGURE 8.6. Principal parts of a side-entry holder that is held in the goniometer stage. The specimen is clamped into the cup at the end of the rod. A small jewel at the end of the rod (usually sapphire) fits into another jewel bearing in the stage to provide a stable base for manipulating the specimen. The O-ring seals the end of the holder inside the vacuum. Manipulating the specimen is accomplished from outside the column via controls within the rod. (See Figures 8.8–8.11 to see just where the specimen goes.)

- *The O-ring,* which is one mechanical link to the microscope column. Some holders have two O-rings and the gap between the O-rings is pumped separately to improve the vacuum.
- *The jewel bearing,* which is the other mechanical link to the microscope column. You push on this bearing to move your specimen back and forth and from side to side. Like the O-ring, you must keep the bearing clean otherwise the specimen will not be stable.
- *The cup,* which actually holds your specimen and thus provides the immediate environment which is seen by stray electrons and any X-rays coming down the column. So cups in holders for AEM are made of Be to minimize the generation of X-rays that would interfere with microanalysis.
- *The clamping ring or screw,* which holds the specimen in the cup. This ring (not shown in the figure), which may also be Be, must be carefully designed. It must hold your specimen firmly (so, e.g., magnetic disks cannot be pulled out of the cup by the lens field). However, the ring must not be so difficult to tighten that you put undue pressure on your specimen. Brittle disks may break as you are loading them. There are two kinds of retaining rings: screw-thread rings, which are easier to control and do not damage metals, but you'll find they may break ceramics because they transfer shear stresses to the disk; spring clips are difficult for the novice to master, but with practice you'll find they offer more control over the load that you put on the specimen, so we recommend them for the experienced ceramist. Unfortunately, no one we know makes Be spring clips!

In a more modern design, the jewel bearing is omitted so that the holder is supported at just one pivot point.

8.9 TOP-ENTRY HOLDERS

Top-entry holders are becoming less common because they essentially preclude XEDS analysis in the TEM. Also, it is more difficult to design such holders so that the specimen can be manipulated (e.g., rotated or strained). Their great advantage was that they were much less susceptible to drift since they were not connected directly to the outside, so early HRTEM required top-entry holders. Today, however, nearly all TEMs up to 400 kV use side-entry holders; only the DSTEMs retain the top-entry (bottom entry?) design.

Another drawback of such holders is that the bore of the objective lens must be asymmetric (using an upper and a lower polepiece), which actually limits the ultimate resolution by constraining the lens designer. Figure 8.7 shows a schematic diagram of such a holder.

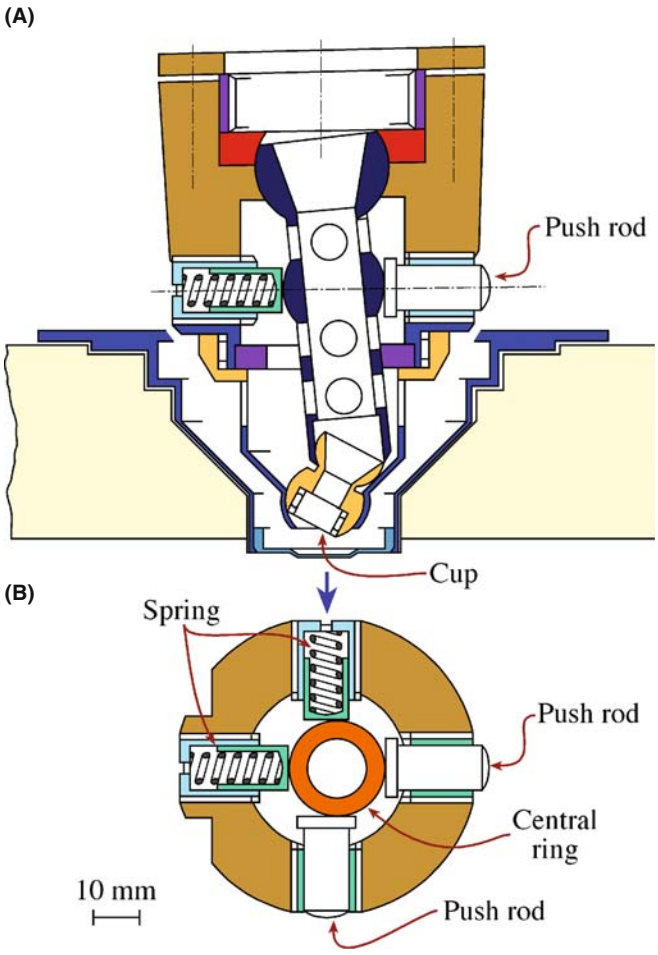


FIGURE 8.7. Top-entry holder: (A) cross section; (B) top view. The cartridge has a cone shape which fits into the tapered bore of the objective lens polepiece. The specimen sits in a cup at the base of a column through the cone down which the incident beam travels. Simple manipulations such as tilting or rotating require complex micromechanical design, since the specimen is at the base of the cartridge and completely surrounded by the polepiece. To tilt, e.g., as shown in (A), push rods are pressed against springs in two orthogonal directions, displacing a central ring around the column (see B), thus tilting the specimen cup.

8.10 TILT AND ROTATE HOLDERS

One feature of TEMs which may surprise you if you are a new user is that a wide variety of holders is available for the TEM. Figure 8.8 shows illustrations of these different designs for the side-entry holder:

- *Single-tilt holder:* This is the basic holder with which any novice should start practicing. You can only tilt around the axis of the rod. It is relatively cheap, robust, and can at least give you some idea of the usefulness of tilting a specimen for diffraction-contrast studies.
- *Quick change holder:* This is also a single-tilt holder that clamps the specimen with a lever arm which you raise and lower onto your disk or grid. It doesn't put a high stress on the specimen, but it doesn't hold it very strongly either. Don't use it for magnetic specimens, but it can be great for ceramics. Different retainers can be substituted for the clamp as shown in Figure 8.8 (bottom), creating a more versatile multipurpose holder
- *Multiple-specimen holder:* This is usually a single-tilt holder, but you can load up to five specimens into the column at one time as shown in Figure 8.9A. A two-specimen, double-tilt version is also available (Figure 8.9B). Such holders can be useful if you are not very good at specimen preparation, or you want to compare different specimens under identical conditions without turning off the beam. However, in

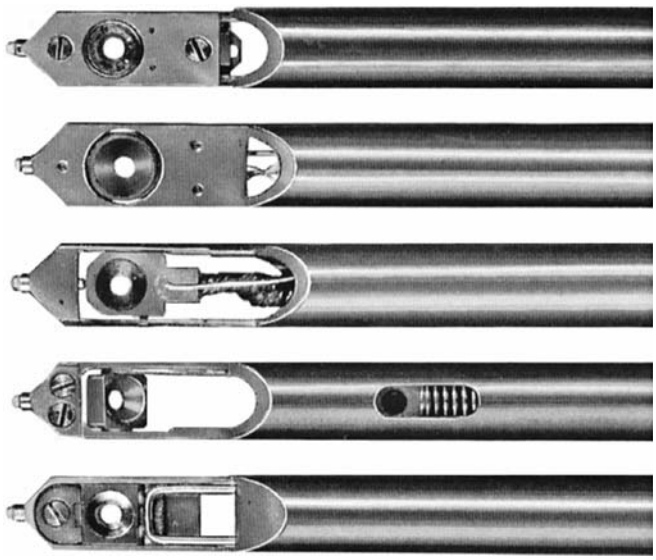


FIGURE 8.8. Examples of different designs for the side-entry holder. From the top, they are: a rotation holder, a heating holder, a cooling holder, a double-tilt holder, and a single-tilt holder.

FIGURE 8.9. Multiple-specimen holders: (A) two-specimen double-tilt and (B) five-specimen single-tilt.

modern TEMs, specimen exchange is relatively quick, except in UHV instruments where the multi-holder would probably be more useful although it is less common.

- *Bulk specimen holder:* This holder is used for surface imaging and diffraction, e.g., using SE or BSE in a STEM or for reflection diffraction and imaging in a TEM (see Chapter 29 for more about these techniques). The bulk specimen is larger than the traditional 3-mm disk (usually ~10 mm × 5 mm) so if you can create a thin specimen of these dimensions, the bulk holder will allow you to sample more of your material at one time (Figure 8.10).

  So don't always think that you are limited to 3-mm specimens!

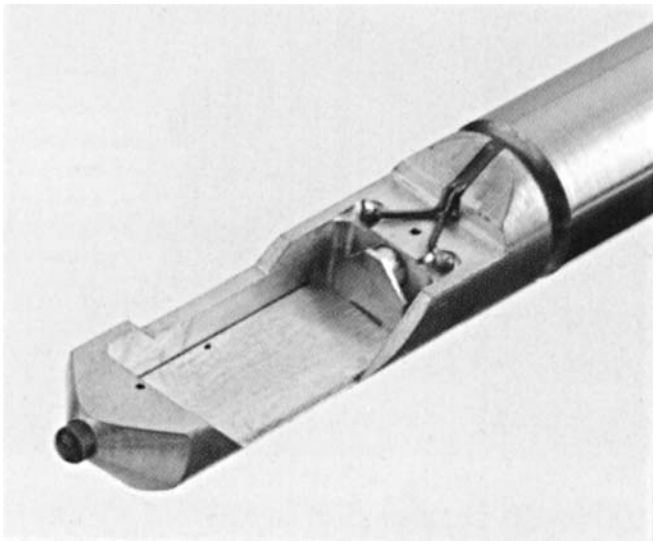


FIGURE 8.10. A bulk holder for large specimens.

- *Double-tilt holder:* This is the most popular holder since it gives you the most flexibility in orienting the specimen. It is absolutely essential for imaging and diffraction studies of crystalline specimens. The tilt axes are fixed as two orthogonal directions. In some designs, you can remove the cup while the specimen is in place which means that you can reinsert your specimen in the same orientation. This feature is extremely useful if your specimen is robust.
- *Tilt-rotate holder:* You would often like to be able to orient your specimen parallel to the tilt axis (along the rod). This holder lets you do just that. It's a strength for the side-entry holder: one tilt axis is always parallel to the rod of the holder which also gives the largest tilt angle.
- *Low-background holder:* The cup and clamping ring are made of Be to minimize the generation of bremsstrahlung X-rays and characteristic X-rays. So they are required for XEDS studies. They can be double or single tilt and may be cooled also.
- *Tomography holder:* This is a new design that allows you to tilt the sample through a full 360°. It's ideal for looking at needles (like an AFM or atom-probe tip).

8.11 IN-SITU HOLDERS

Special holders have been developed that allow you to change your specimen while you observe it in the TEM; in other words you can do experiments (heat, cool, strain, twist, compress, etc.) on a specimen in the TEM.

- *Heating holder:* Such holders in a conventional TEM can go to ~1300°C which is measured by a thermocouple attached to the cup. In HVEMs, the temperature can go higher because of the larger gap between the polepieces. You have to be careful to calibrate the temperature and remember that the temperature may be different for different specimens. You should also be sure that the material you are studying does not form a eutectic alloy with the material forming the holder! If the eutectic does form it will have a lower melting point, so you may deposit part of your specimen and the holder on the objective lens, or down onto the screen, if the microscope is well aligned.
- *Cooling holder:* This is available for either liquid-N_2 or liquid-He temperatures. These holders, which can be single or double tilt, are a great asset for XEDS, EELS, and CBED studies since they minimize surface-borne contamination. They are also essential for in-situ studies of superconducting materials and ideal for polymers or biological tissue. However, you should remember that the cold holder can also act as a small cryo-pump so that it actually attracts contamination.

Since you are necessarily changing the temperature at the specimen relative to its surroundings, be prepared for specimen drift. It takes time for the whole system to stabilize. The Polara from FEI is a modification of the cooling holder which fits in a special dedicated stage; the tip of the holder detaches from the rod with obvious advantages.

- *Cryo-transfer holder:* Certain specimens are prepared at cryogenic temperatures such as liquids, latex emulsions, and tissue in general. This holder permits you to transfer such cold specimens into the TEM without water vapor from the atmosphere condensing as ice on the surface.
- *Straining holder:* This holder clamps the specimen at both ends then applies a load to one end, via a load cell or screw-thread mechanism, as shown in Figure 8.11. The sample can be in the shape of a small tensile specimen and it is thinned in the middle of the gauge length (see inset). The motion of dislocations, cracks, etc., are then easily monitored, so a video camera is an essential accessory. You can vary the load, to study cyclic as well as tensile loading, and the strain rate is another variable that is easily controlled. In Figure 8.11 a furnace is present, so the specimen can be heated while under load. The use of piezoelectric drives is leading to great improvements in this type of holder.
- *Probing holders:* These are like AFM holders for the TEM. You can use them to 'poke' your specimens—just like an AFM, STM, or an indenter—while you observe the effect in the TEM.
- *EBIC and CL holders:* The essential feature is the electrical feed-through that allows you to control the charge recombination in a semiconductor or certain mineral specimens by applying a bias across the specimen surface.

Beware: heating and straining holders, in particular, can produce effects in thin foils that are totally uncharacteristic of your bulk sample. So you must use these holders carefully and interpret your results cautiously. Often, surface reactions will dominate internal reactions when you are trying to induce a phase transformation by heating. The surface may also stop grain boundaries from migrating at temperatures where they would do so in the bulk material. Obviously, defect motion under applied stress may also be strongly affected since the 3D stress field will be very different in bulk specimens compared to thin foils.

These problems can be overcome to some extent if you use thicker specimens and examine them in an HVEM, or at least an IVEM, and the whole field of in situ studies, particularly, heating and/or straining, is best performed in such microscopes (Butler and Hale 1981 and Section 29.12). However, the high-energy electrons in these microscopes may introduce lattice defects that affect the very phenomenon that you want to study, e.g., beam-induced vacancies can change diffusional phase-transformation kinetics very easily.

It is also possible, but much more difficult and expensive, to manipulate specimens in top-entry stages. The top-entry holder shown in Figure 8.12 is a heating-

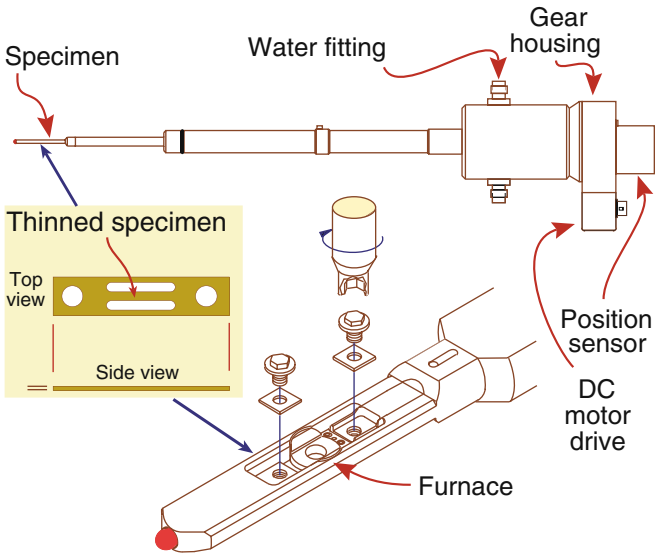


FIGURE 8.11. A side-entry combined straining and heating holder. The specimen looks like a miniature tensile specimen (inset) and is clamped at either end by hex screws. There is a screw-thread arrangement for pulling the specimen contained within the rod. The furnace surrounds the central thin portion of the specimen.

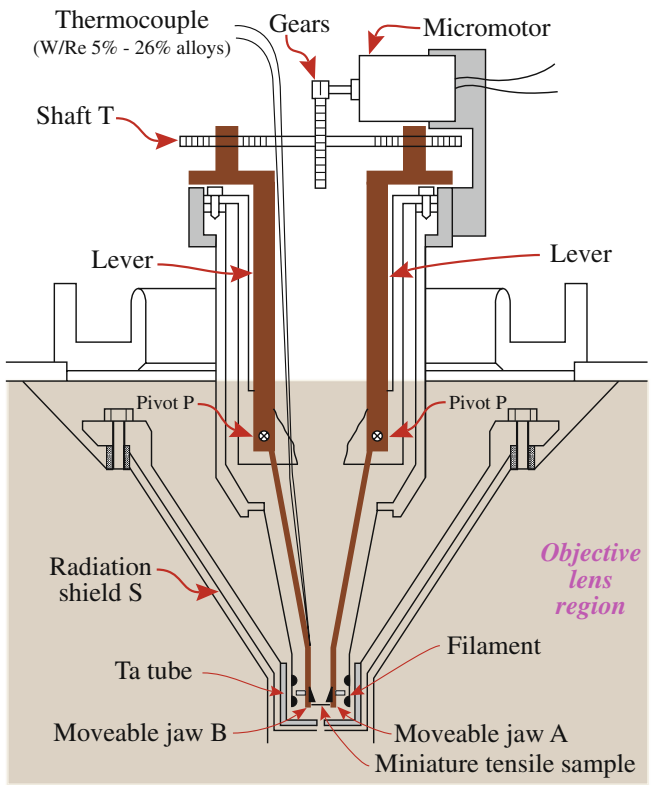


FIGURE 8.12. A top-entry, heating-straining holder which can be used at temperatures up to 2300 K in a 3-MV HVEM.

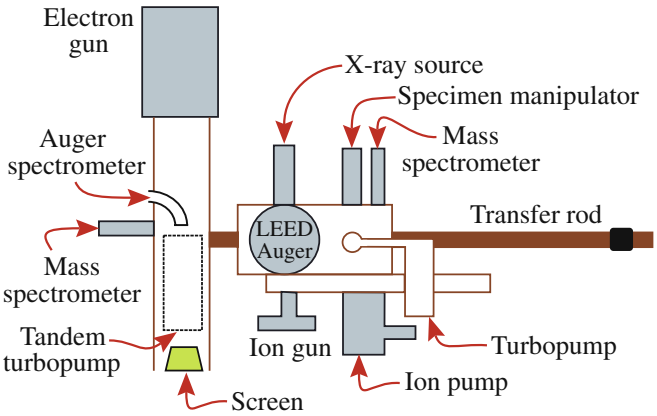


FIGURE 8.13. Schematic diagram showing the Hitachi H900 UHV TEM. This instrument is equipped with a prechamber with LEED, Auger, and an ion gun which can be used to clean the specimen, allowing UHV surface analysis to be carried out on the TEM specimen. The holder has to transfer the specimen through a prepump chamber where it is ion-cleaned before going into the column.

straining holder, which is capable of operating at temperatures up to 2300 K. The heat is provided by a coaxial Ta tube that supports the W heater filament as shown in the figure. The holder is used in a 3-MV microscope where the specimen diameter is 5 mm. The larger specimen diameter means that the disk can be shaped as a small tensile specimen and still be quite robust.

There are also special combinations of holders and stages which have been optimized for particular applications. The example shown in Figure 8.13 has been optimized to combine surface studies using low-energy electron diffraction (LEED) and Auger analysis with TEM. The prechamber is fitted with an ion gun to clean the sample before the surface is analyzed. The specimen can then be moved into the TEM column for transmission studies. A similar prechamber has been

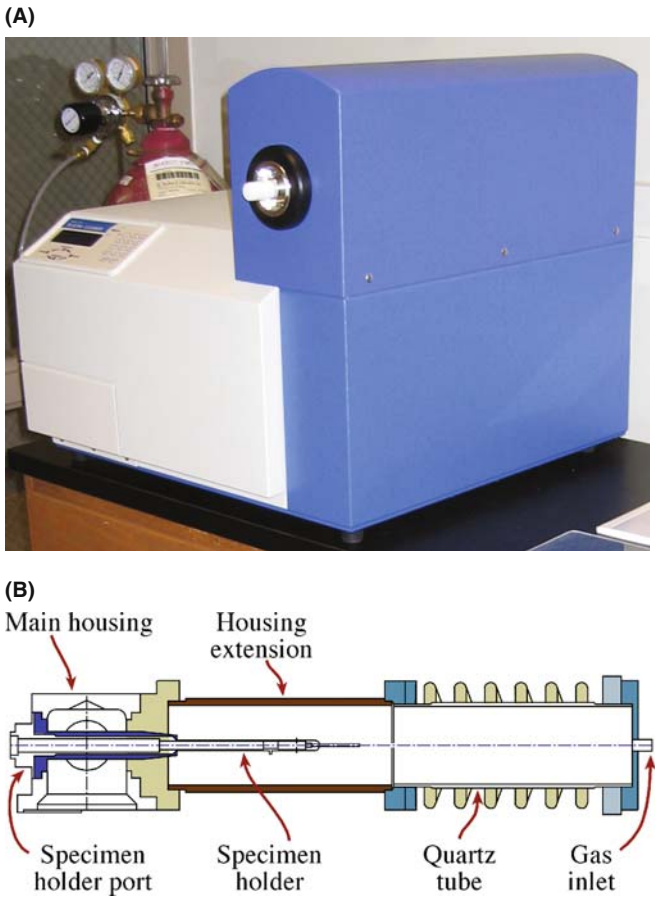


FIGURE 8.14. Plasma cleaner: (A) example and (B) schematic.

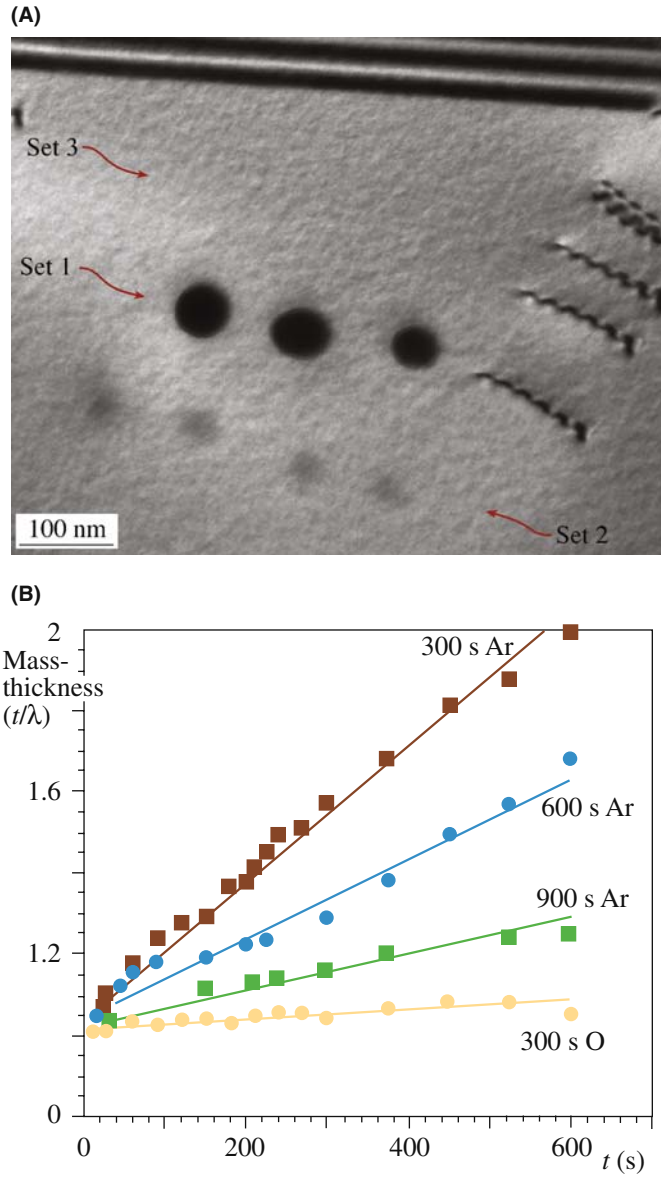


FIGURE 8.15. Cleaning the surface of the specimen using a plasma cleaner reduces the contamination produced by a focused electron beam. (A) Set 1 was produced before the specimen was cleaned. Sets 2 and 3 (which you can't see) were produced after 5 minutes of Ar cleaning and then an additional 5 minutes of oxygen cleaning, respectively. (B) The rate at which the specimen contamination builds up; the additional cleaning in pure oxygen always reduced the contamination rate.

used elsewhere to provide a method to clean the sample before growing thin films on the sample by molecular-beam epitaxy (MBE) or thermal evaporation.

One of the reasons for using higher accelerating voltages is that this gives more room in the specimen-stage region. Thus, even 400-kV microscopes can be fitted with a small, differentially pumped environmental chamber. Such a chamber allows in situ studies of corrosion, degradation of catalysts, etc., especially when combined with a heating holder.

Development of new holders and their use for in-situ studies is going to be one of the most exciting topics in TEM. We'll expand on this topic in the companion text.

8.12 PLASMA CLEANERS

Plasma cleaners have been available for removing surface contamination and modifying surfaces (for example, changing wettability) for over 30 years. The unit is now marketed as a small box (e.g., as shown in Figure 8.14A) that can usually accommodate one specimen holder; the rod of the holder is placed inside a plasma chamber, as shown in Figure 8.14B, just before it goes into the TEM. Plasma cleaners have long been used to modify/clean surfaces of glass, semiconductors and other ceramics, metals, and even polymers and biomaterials. The development of plasma cleaners for cleaning both the TEM holder (the part that goes into the vacuum of the TEM) and more specifically the specimen that has already been placed in the TEM holder, is more recent, but the possibilities are clear. This cleaning is generally regarded as essential for small-probe AEM. Most of the contamination that occurs in the electron beam originates on your specimen, not from a 'dirty' vacuum.

Although the processes occurring in the plasma are complex (like the plasma itself) the basic idea is illustrated by the removal of hydrocarbon contamination on the specimen shown in Figure 8.15A. The plasma consists of a mixture of energetic electrons and ions that bombard the surface and break the C–H bonds. With short duration exposure, the surface of the specimen itself is essentially unaffected. The hydrocarbon is thus gradually reduced in molecular weight and pumped away in the vacuum of the cleaner. The user has some flexibility in the choice of the plasma gas although the manufacturer usually limits this. Although many gasses can be used in principle, O_2, N_2, and Ar are the most common. The graph in Figure 8.15B shows the effect of using an oxygen plasma—it reduces the rate at which the contamination grows under the electron probe; the effect of the oxygen-reactive gas is similar to that found when using iodine in the ion mill.

CHAPTER SUMMARY

The vacuum and the holder are the two parts of the TEM that most closely affect your specimen. You have to treat both carefully if you want to be sure of getting the most out of your TEM. The vacuum is usually automated, so you don't have too much control over it. However, you can degrade the vacuum easily if you are a careless operator; e.g., if you don't bother to prepump your film. You don't want to touch any part (including the specimen holder) that will go into the vacuum. In fact, you should treat the specimen holder as if it were a rare jewel; it may actually contain a couple of synthetic ones and it certainly costs as much as a diamond of several carats! (It's also probably worth more than its weight in gold.)

With the range of holders available today, you can conduct many materials science experiments on your thin specimen while observing it in the TEM. However, if you're looking at crystalline material, the most common manipulation is still tilting the specimen in two orthogonal directions to orient different crystal planes parallel to the electron beam. You'll understand more fully why this is important after you've finished reading Parts 2 and 3 (and even in Part 4).

REFERENCES

There is a great deal of history for you to explore when you have time. Why are specimens 2.3 or 3.05 mm diameter? Why do we talk about loading the plates? In many cases, this history holds us back but we often don't recognize why!

VACUUM

A more detailed list of references is given in the chapter on in situ TEM in the companion text. For a full exposition of vacuum technology for TEMs, read Bigelow or the equally informative user's guide by O'Hanlon.

Bigelow, WC 1995 *Vacuum Methods in Electron Microscopy* in *Practical Methods in Electron Microscopy* **15** Ed. AM Glauert Portland Press London. An essential reference.
O'Hanlon, JF 1981 *A User's Guide to Vacuum Technology* John Wiley and Sons New York.

HOLDERS AND IN SITU

Butler, EP and Hale, KF 1981 *Dynamic Experiments in the Electron Microscope* in *Practical Methods in Electron Microscopy* **9** Ed AM Glauert Elsevier Amsterdam.
Watt, IM 1985 *The Principles and Practice of Electron Microscopy* Cambridge University Press New York. See appendix 1.
Komatsu, M, Mori, H and Iwasaki, K 1994 *Design of a Hot Tensile Stage for an Ultrahigh-voltage Electron Microscope and Its Application to In Situ Deformation of Sapphire at 1620 and 1720* K J. Am. Ceram. Soc. **77** 839–842. Illustrates the use of temperatures as high as 2300 K.
Valdrè, U and Goringe, MJ 1971 in *Electron Microscopy in Material Science*, 208–254 Ed. U Valdrè Academic Press New York. This article gives a detailed description of several TEM holders.

THE COMPANION TEXT

A full chapter on holders is included in the companion text with particular attention being paid to holders for in-situ experiments.

SELF-ASSESSMENT QUESTIONS

Q8.1 How does a diffusion pump work?
Q8.2 How does a mechanical pump work?
Q8.3 How does a turbomolecular pump work?
Q8.4 How does an ion pump work?
Q8.5 How does a cryogenic pump work?
Q8.6 If you are concerned about contaminants in the TEM from pumps, which pump(s) should you use (consider where and why)?
Q8.7 Where does the most common vacuum leak occur in a TEM and why at this point?
Q8.8 Name one way to localize the source of a leak.
Q8.9 Which types of specimens introduce contaminants and how can such introduction be prevented?
Q8.10 How can you personally introduce contamination into the microscope and what should you do to stop doing it?
Q8.11 Summarize the advantages and disadvantages of side-entry holders.
Q8.12 Why are TEM top-entry holders out of fashion?
Q8.13 What is a high vacuum (give a number with SI units)?
Q8.14 What is an ultrahigh vacuum and why is it worth paying a lot of money to buy a TEM with such a vacuum?
Q8.15 Which type of holder is the most flexible for TEM?
Q8.16 Why is the column baked? Why should this be done carefully?
Q8.17 What size is the typical specimen cup in the holder?
Q8.18 Why don't we make larger-diameter thin samples?
Q8.19 What are the drawbacks when using a straining holder?
Q8.20 Why would you cool a specimen?
Q8.21 What do we sometimes deliberately insert into a TEM that creates even more contamination and vacuum degradation than the specimen?

TEXT-SPECIFIC QUESTIONS

T8.1 How would you notice that there is a vacuum leak in the TEM (without watching the gauges)?
T8.2 Why can baking both improve the vacuum and cause more leaks?
T8.3 Explain with a schematic diagram (like Figure 8.5) how the combination of low and high vacuum pumps is used in tandem to pump down the microscope after: (a) it has been let down to air (e.g., for repair); (b) after it has been degraded by the insertion of a specimen; (c) after it has been degraded by changing the film.
T8.4 Copy Figure 8.1 and cut out the circle attached to the rotating vane. By placing that circle and the vane in different positions within the pump chamber show how air is alternately pulled out of the microscope (into the inlet port) then expelled from the pump (through the outlet port).
T8.5 Estimate the number of atoms in a cubic millimeter of air at 10^{-3} and 10^{-6} bar and thus estimate the number of Ti atoms that need to be ionized to reduce the vacuum in that cubic millimeter by that amount in each case. State any assumptions and be brief.
T8.6 Compare and contrast top-entry and side-entry holders from the points of view of stability (mechanical, thermal, etc.) versatility, cost, etc.

T8.7 Under what circumstances might you prefer to use a holder with just one tilt axis but the ability to rotate the specimen?

T8.8 Why do we choose to use Be for analytical holders rather than a lighter metal, or a heavier one such as Mg? What are the potential problems with using Be?

T8.9 What are the problems with manufacturing a stage and holders for UHV operations rather than for the generally poorer vacuum in most TEMs?

T8.10 Why is it difficult to manufacture holders that operate at extremely high or very low temperatures that do not compromise the routine operation of the microscope?

9

The Instrument

CHAPTER PREVIEW

Over the preceding four chapters we've now introduced all the essential components of the TEM and it's time to see how the guns (Chapter 5), lenses (Chapter 6), detectors/screens (Chapter 7), and specimen holders (Chapter 8) are combined to form the instrument. Just as we do for the VLM, it's convenient to divide up the TEM into three components: the illumination system, the objective lens/stage, and the imaging system. The illumination system comprises the gun and the condenser lenses and its role is to take the electrons from the source and transfer them to your specimen. You can operate the illumination system in two principal modes: parallel beam and convergent beam. The first mode is used primarily for TEM imaging and selected-area diffraction (SAD), while the second is used mainly for scanning (STEM) imaging, analysis via X-ray and electron spectrometry, and convergent-beam electron diffraction (CBED).

The objective lens and the specimen holder/stage system is the heart of the TEM. This critical region usually extends over a distance of ~10 mm at the center of the TEM. Here is where all of the beam-specimen interactions take place and the two fundamental TEM operations occur, namely, the creation of the various images and DPs that are subsequently magnified for viewing and recording. Clearly the objective lens is the most important lens in a TEM because its quality determines the quality of all the information about the specimen that we seek. Because of the various defects inherent in magnetic lenses, our specimens must be placed close to the center of the objective lens and this restricts our ability to manipulate the specimen and gain access to the various signals generated. However, as we've mentioned many times already, we can now overcome many of the defects that limit lens performance by inserting complex aberration-correction systems into the column, but only the most expensive TEMs are so equipped.

The imaging system uses several lenses to magnify the image or the DP produced by the objective lens and to focus these on the viewing screen or computer display via a detector, CCD, or TV camera. We'll refer to the magnifying lenses as the intermediate and diffraction lenses and the final lens as the projector lens (since it projects the final image or DP onto the viewing screen or detector). All TEM operations involve observing the electrons on a screen of some form, with or without a specimen in place. In many current TEMs you will have a button for focus, another for magnification, and another for diffraction (or a slide on the computer screen). These three components, the illumination system, stage, and imaging system, are often called the 'column', for obvious reasons. Throughout this book, the electrons will invariably be shown as traveling 'down' the column because this is the construction of most (but not all) TEMs.

The purpose of this chapter is to go through the principal functions of the many lenses in the column and give you some feel for what is happening in the microscope when you press a button, rotate a knob or trackball, or drag your mouse. The better you understand the operation of a TEM, the more you can be sure that you are getting the most out of it.

We do NOT intend to teach you how to operate a specific TEM since individual manufacturers choose different ways to manipulate the many lenses, apertures, deflectors, etc., in the TEM column to achieve the same ends. What we will do is introduce specific *operational procedures* that we hope are generic enough such that you should, in concert with the operating manual of your own TEM, be able to carry out all the basic procedures.

9.1 THE ILLUMINATION SYSTEM

The illumination system takes the electrons from the gun and transfers them to the specimen giving either a broad beam or a focused beam (often called a 'probe' or 'spot') entering the specimen. We can think of these two cases as analogous to a floodlight or a spotlight. In Chapter 5 we described how the gun produces an image of the source (called a crossover). This crossover acts as the object for the first lens in the illumination system that consists of several condenser lenses (C1, etc.). We will discuss the two different ways to use the illumination system: formation of a parallel electron beam (although it is almost never truly parallel) or a convergent beam (which may be divergent).

9.1.A TEM Operation Using a Parallel Beam

In the *traditional* TEM mode, the first two condenser lenses (C1, C2) are adjusted to illuminate the specimen with a parallel beam of electrons, typically several micrometers across at reasonable magnifications (20,000×–100,000×). As shown in Figure 9.1, the C1 lens first forms an image of the gun crossover. In the case of a thermionic source, the original crossover may be several tens of micrometers across, and this crossover is demagnified by an order of magnitude or more. In the case of a FEG, the source size may be less than the desired illumination area on the specimen so it may be necessary to magnify the crossover (so the condenser lenses don't always condense!). The simplest way to produce a parallel beam is shown in Figure 9.1A in

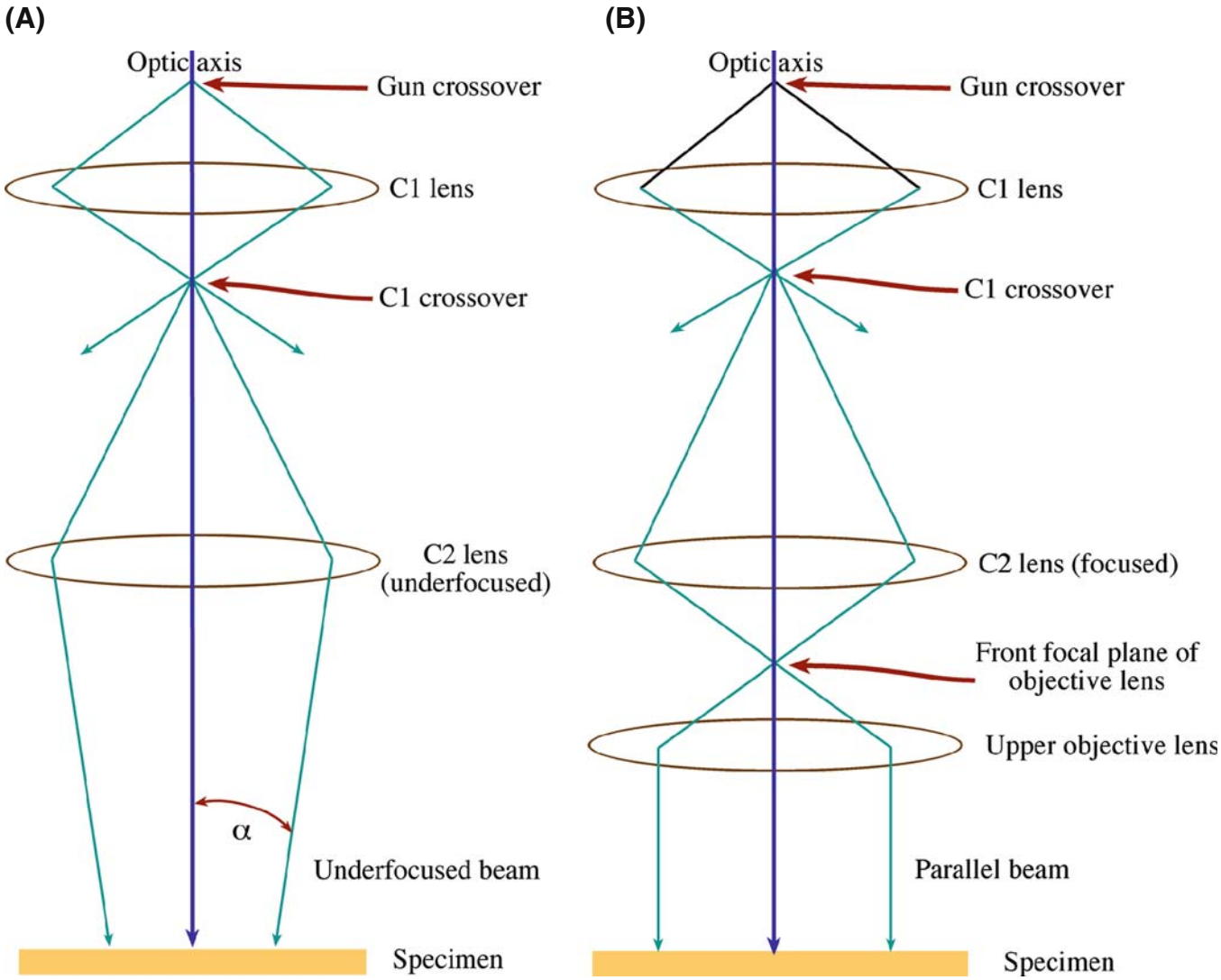


FIGURE 9.1. Parallel-beam operation in the TEM: (A) the basic principle illustrated, using just the C1 and an underfocused C2 lens. (B) The practical situation in most TEMs; using the C1 and C2 lenses to image the source at the FFP of the condenser-objective lens, thus creating a parallel beam at the specimen plane. Therefore, the upper objective is sometimes termed the C3 lens.

which the C2 lens is weakened to produce an underfocused image of the C1 crossover (although no TEM today has only two condenser lenses). Remember from Section 6.2.D and Figure 6.5 that any lens operates either in focus or out of focus and there are times when each mode is useful. In its out-of-focus condition, a lens is said to be overfocused if it is strong and the crossover occurs before the image plane and underfocused if the lens is weak and the crossover occurs after the image plane. Remember also that underfocus gives a more parallel beam at the specimen than overfocus.

CONVERGENCE ANGLES

α is small. All ray diagrams in this book are drawn with exaggerated (large) angles.

While the beam in Figure 9.1A is not exactly parallel to the optic axis, α under these conditions is $< 10^{-4}$ rads (0.0057°), which is effectively a parallel beam.

When generating the small probes we need in STEM and AEM, the upper polepiece of the objective lens acts as the C3 lens (see the next two sections) and controls the beam hitting the specimen as shown in Figure 9.1B. If the C2 lens is focused to produce an image (of the crossover) at the front-focal plane (FFP) of the upper-objective polepiece then a parallel beam of electrons is formed by the lens. The so-called condenser-objective (c/o) lens system is standard on almost all TEMs used for materials characterization but there is no c/o lens in dedicated STEMS, TEMs made before ~1980 and those designed primarily for imaging of biological specimens. The c/o lens system complicates the electron optics, as we shall see in Section 9.1.C below.

We'll see that parallel illumination is essential to get the sharpest selected-area DPs (SADPs) (Chapter 18) and, in principle, the best *classical* image contrast (Part 3). In fact, we usually assume that the incident beam is parallel when interpreting our *classical* images. Usually you should underfocus C2 until the illuminated area on your specimen fills the viewing screen. The higher the magnification the more you have to strengthen C2 (making the beam *less* parallel) in order to keep the illuminated area just filling the screen and thus illuminating less of the specimen.

CLASSICAL? PARALLEL?

Classical (or traditional or conventional) means... thinking in terms of waves and beams. Parallel illumination isn't really parallel, it's just not very convergent.

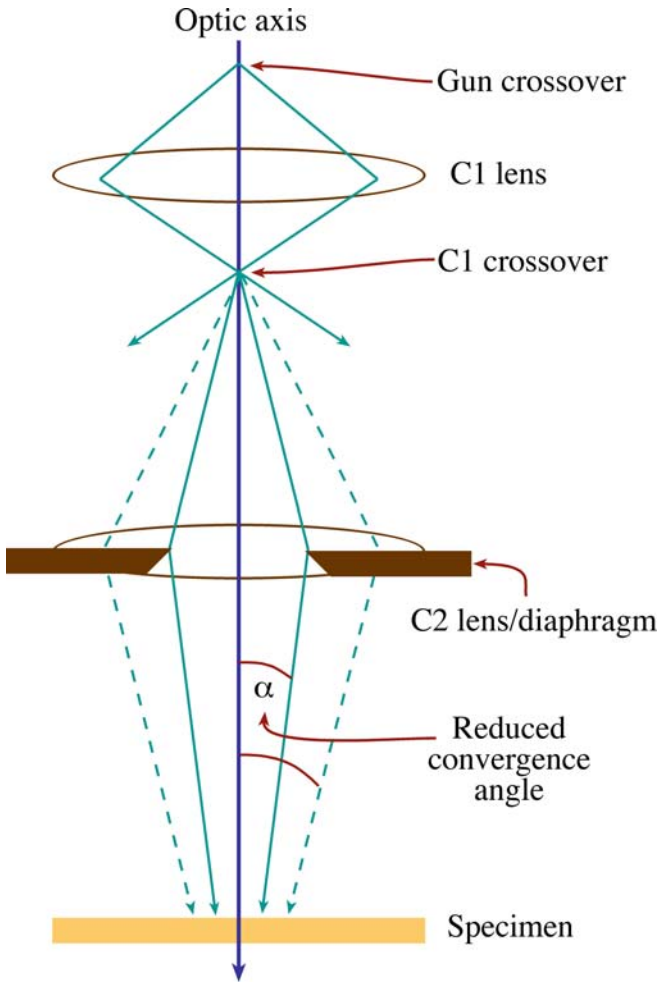


FIGURE 9.2. Effect of the C2 aperture on the parallel nature of the beam: a smaller aperture creates a more parallel beam at the expense of total number of electrons (i.e., reduced probe current) reaching the specimen.

In parallel-illumination TEM mode, such as you would use for diffraction-contrast imaging and SADP formation, there is usually no need to change C1, which is therefore kept at some intermediate setting, recommended by the manufacturer. The only other variable is the C2 aperture. A small aperture reduces the electron current reaching your specimen. However, if you use a smaller aperture, you decrease the angle of beam convergence and therefore make the beam more parallel, as is evident from Figure 9.2.

9.1.B Convergent-Beam (S)TEM Mode

Now, there are times when you may wish to focus the beam more, so that the intensity of the beam on a specific area of the specimen is increased. Let's look at various ways to do this.

If you want to *minimize* the area of the specimen that you are illuminating, you simply change the C2 lens so it is focused rather than defocused and you form an

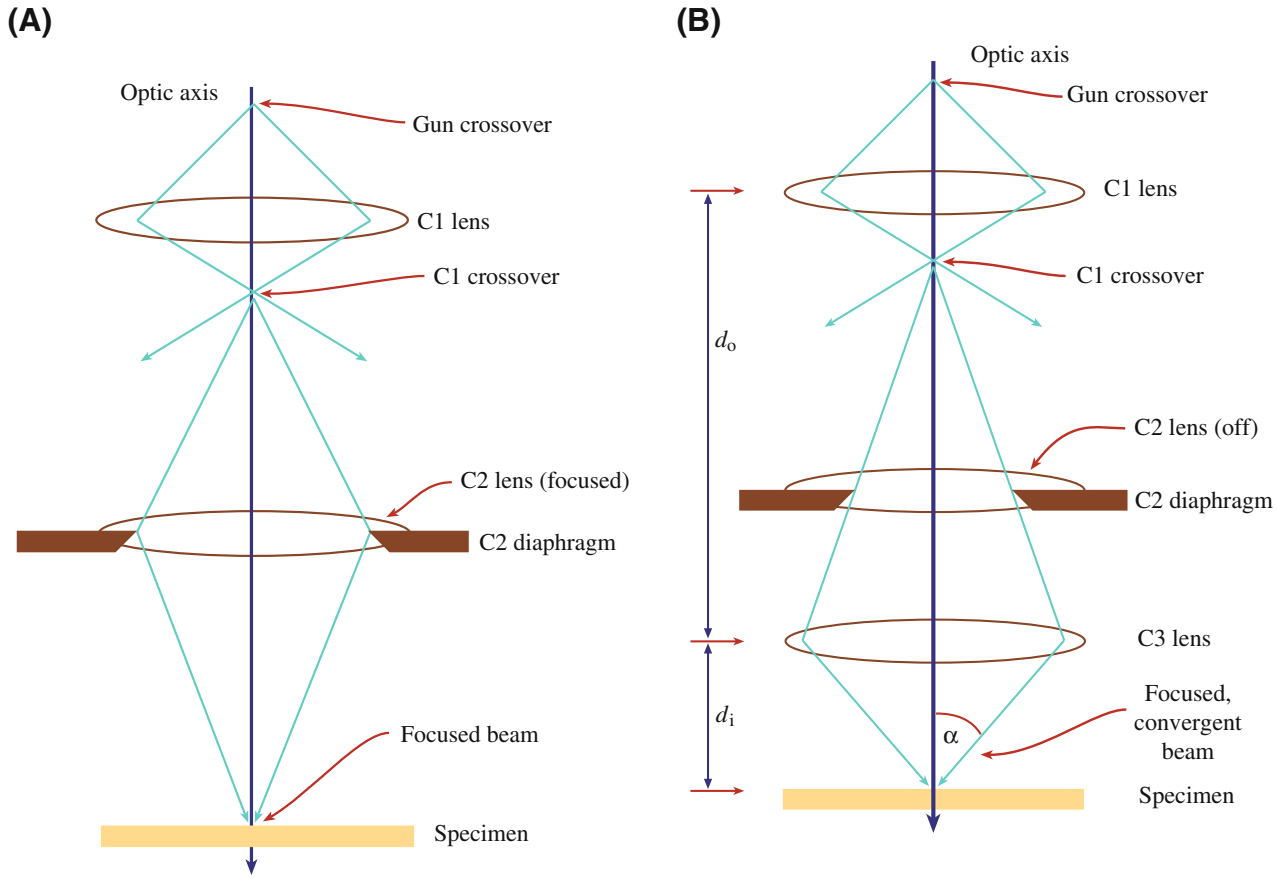


FIGURE 9.3. Convergent-beam/probe mode in the TEM. (A) The basic principle: a focused C2 lens illuminates a small area of the specimen with a non-parallel, convergent beam. (B) The practical situation in most TEMs: use of the upper-objective polepiece as the C3 lens gives the smallest possible probe and large convergence angles. Note that C2 is shown here as having no effect on the ray paths (i.e., it is effectively switched off). The large d_o/d_i ratio gives the maximum demagnification of the image of the gun crossover.

image of the C1 crossover at the specimen, as shown in Figure 9.3A. This is the condition under which you can view the source image to adjust its saturation (go back and look at Figure 5.5B and C) or to measure the dimensions of the beam (see Section 5.5.D). When C2 is focused like this, the beam is at its least parallel and most convergent. While the intensity of illumination on the viewing screen will be greatest, your image contrast will be reduced and any SADP will be distorted. Ideally, for routine TEM work, your specimen should always be thin enough so that you never have to operate with C2 focused but, in practice, you'll often find yourself focusing C2 to compensate somewhat for poor transmission through a thick portion of your specimen.

However, there are times when we do need to create a focused, convergent beam at the specimen. We then use the other principal way to operate the illumination system: the convergent-beam (or probe/spot) mode. When you use this mode you won't immediately see a useful image; the convergence destroys the parallelism and reduces the image contrast. So to see an image we have to scan the beam; this mode of operation of the illumination system is standard for STEM and AEM.

PROBE

The convergent beam is a probe. We use a probe to localize the signals coming from the specimen, as in XEDS, EELS, CBED, or HAADF.

Now, if you have a FEG it's possible to use the C1 and C2 lenses to produce Å-level probes but, with a thermionic source, it isn't possible to use just those two lenses (as in Figure 9.3A) to demagnify the relatively large thermionic-gun crossover to less than several nanometers. So to get the desired probe sizes of $\ll$ 1 nm for analysis, etc., the usual solution, as we just noted, is to introduce a C3 or c/o lens. We can only do this if the objective lens is split into two polepieces with separate coils and then we can make the upper polepiece of the objective lens much stronger than usual and weaken C2 or even turn it off, as shown in Figure 9.3B. In addition, C1 must be strongly excited so the image of the gun crossover is a long way from C3. Thus the C3 image distance (d_i) is much less than the object distance (d_o), which gives a large demagnification of the C1 crossover (see equation 6.2).

From Figure 9.3B you can see that, although C2 is switched off, the C2 aperture still controls the convergence angle (α) of the beam on the specimen. As was the case for parallel-beam mode, a smaller C2 aperture gives a smaller α. You'll see later in Chapter 20 that the correct choice of C2 aperture is important in CBED and also in defining the exact dimensions of the probe for XEDS or EELS analysis (see Part 4). Some TEMs such as the FEI and older JEOLs use a double-gap C1 lens (check the situation for your own TEM in the manufacturer's handbook). In these TEMs there is a single C1 lens winding but effectively four polepieces and hence two crossovers occur in the bore of the lens. In some other TEMs (e.g., the Akashi 002B, the JEOL 2010, and Hitachi 8100 and later instruments) there are actually three independent lenses. Without this additional crossover, it's not possible to get a probe $< \sim$10 nm. An additional condenser lens will yield a probe size about 10× smaller. Even though you only get to use two controls (C1 and C2) there are at least three crossovers. So, as you can see, condenser-lens configurations can get much more complicated than the basic principle shown in Figure 9.3B. The more lenses you have, the more flexible the optics and therefore the more operations that a single TEM can perform.

So, in probe-forming TEMs, the role of C1 is fundamentally different from its role in traditional, parallel-beam TEM. The C1 lens is used to control the probe size (and thus the current) at the specimen. As shown in Figure 9.4, a strong C1 gives you a small probe while a weak C1 creates a large probe. This difference occurs because increasing the strength of C1 shortens its d_i, thus lengthening d_o for the probe-forming C2 (or C3) lens and, therefore, increasing the demagnification (but remember that this is a very simplified diagram).

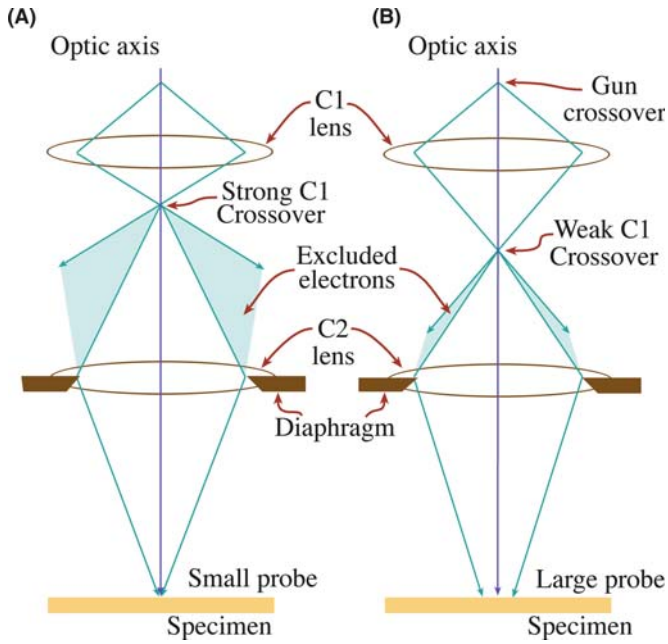


FIGURE 9.4. Effect of the C1 lens strength on probe size: a stronger C1 lens (A) results in greater demagnification by any subsequent lens (C2 and/or C3), giving a smaller electron beam at the specimen. A weaker lens (B) gives a broader probe. Note how changing C1 also changes the number of electrons that hit the C2 diaphragm and therefore, do not contribute to the current in the probe hitting the specimen.

9.1.C The Condenser-Objective Lens

When convergent-beam TEMs were first constructed, it was not possible to design an objective lens that would give both a parallel or a convergent beam at the specimen. So you had to change polepieces in order to change operating modes, which was highly inconvenient. To solve this problem, the condenser/objective (c/o) lens was developed in the mid-1970s. In a c/o lens, the magnetic field between the polepieces of the objective lens is very strong ($\sim$ 2 Tesla). This strong field has a dual function. First, it acts as a condenser lens and effectively causes the beam to converge on the specimen. Thus, as we mentioned above, we can refer to the condenser part of the lens as the C3. Second, once the beam has penetrated the specimen, it is focused just as it would be in a conventional lens. The beam is almost never parallel in a c/o lens and so, as shown in Figure 9.5, yet another lens is introduced into the column between the C2 lens and the objective lens. This extra lens is a condenser mini-lens (called a twin-lens in Philips/FEI instruments) and it is standard on TEMs that also operate as STEMs or analytical TEMs (i.e., it's standard on all modern TEMs used in materials characterization). Without the condenser mini-lens, the action of the c/o is always similar to the objective lens in a dedicated STEM because, in such instruments, parallel-beam imaging conditions are never required. All imaging and analyses are accomplished with a convergent probe. However, in TEM where both parallel and convergent beams are needed, when the mini-lens is activated, the convergent effect of the c/o lens can be offset by forming a crossover at the FFP of the objective. Here the FFP is actually acting as the condenser part of the c/o lens. Thus, unlike in conventional objective lenses (see Figure 6.5), the parallel-beam condition is met in the *overfocus* condition in the c/o lens. If you're feeling confused, take another look at Figure 9.5.

THE c/o LENS

In older TEMs without a c/o lens, underfocus gives a more parallel beam at the plane of the specimen than overfocus. This is not the case if the objective lens is a c/o lens and the beam is more parallel at the specimen when the c/o lens is overfocused.

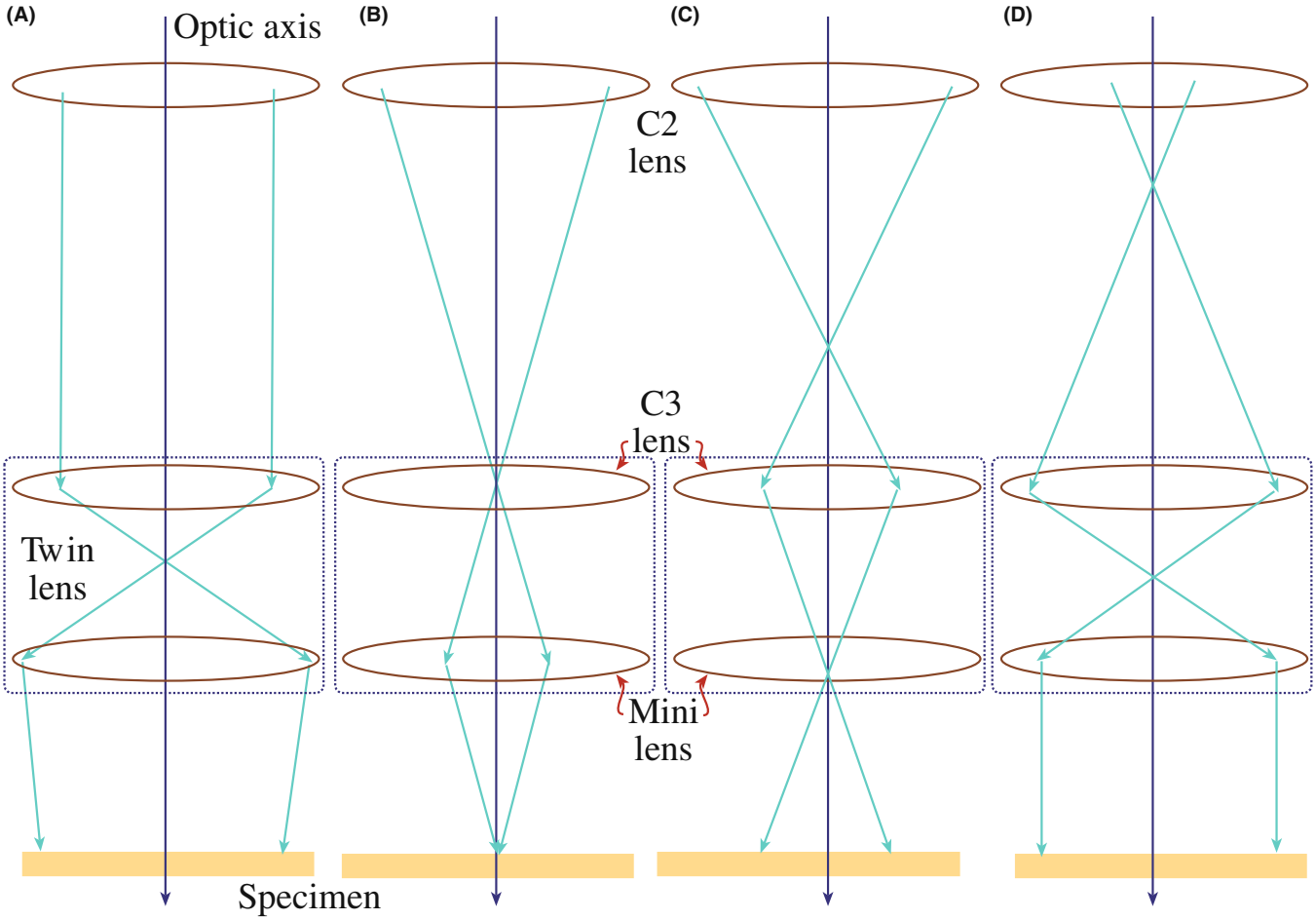


FIGURE 9.5. Condenser/objective lens operation as a function of the strength of the C2 lens. As the C2 strength is increased the beam on the specimen varies from (A) almost parallel, through (B) convergent, and (C) divergent, then (D) back to parallel. You can see that the crossover moves up the column as the C2 strength increases and the C3 lens introduces its own crossover prior to the minilens depending on the convergence of the ray paths entering it.

Figure 9.5 shows the various ray paths through a c/o lens, with a mini/twin lens, as the strength of the C2 lens is changed. As you can see, the beam at the specimen varies from parallel through convergent, divergent, and then back to parallel, as C2 is strengthened. Therefore, because of this interaction between C2 and the c/o lens, changing C2 changes the focus of the probe at the specimen and thus changes the focus of the STEM image and the CBED pattern and also changes the spatial resolution of any XEDS or EELS analyses, as we'll describe in detail in Part 4. So you have to be aware that, when you are operating in probe mode, two lenses control the focus, which can really mess things up unless you are fully aware of the situation. If you are operating in conventional parallel-beam TEM mode then you can simply use the objective lens (lower polepiece) to focus the image (see Section 9.3) and you don't need to adjust the illumination-system optics.

All recent probe-forming TEMs operate at a very specific c/o lens excitation where the STEM probe can be imaged in TEM mode on the viewing screen/computer display. This c/o condition for simultaneous focus and image is a very limited one. (Look in the manufacturer's instructions to find out how to set this condition exactly on your own TEM.) If you don't do this, you're not at the correct objective-lens excitation and the beam can act in contrary ways; the beam convergence can decrease when you expect it to increase and the electron distribution in the probe can be spread into a broad halo rather than a fine focused point. So find out how to do this whenever you're operating in probe-forming mode.

In a c/o lens, while the convergence angle can be effectively reduced to zero, the helical path of the electrons through the strongly excited upper-objective polepiece never permits truly parallel illumination. In contrast, in the older conventional C1/C2 optics, the objective pre-field is not strong and the spiraling is greatly reduced. Therefore, image contrast and other effects arise as a result of c/o lens operation. If you read Christenson, and Eades all of this is very well explained.

9.1.D Translating and Tilting the Beam

There are certain operations where we need to translate the beam laterally on the specimen (e.g., to position a fine probe on a feature of interest for analysis). Similarly, there are times when we need to tilt the beam off the optic axis so it impinges on the specimen at a specific angle (e.g., for centered-DF imaging using a specific diffraction spot which we describe in Section 9.3.C or for hollow-cone imaging/diffraction (Section 9.3.D) and precession diffraction (Section 18.8)). Both translating and tilting are also essential for aligning the beam down the column and are accomplished by varying the current through potentiometers (which we call 'scan coils') to generate a local magnetic field which deflects (rather than focuses) the beam. There are several sets of scan coils in the column, some of which tilt and others which translate the beam. Ray diagrams to explain translating and tilting the beam are shown in Figure 9.6A and B.

When we create a scanning beam for STEM imaging, the beam must always move parallel to the optic axis to mimic the parallel-beam illumination conditions of standard TEM. Such scanning is accomplished by tilting the beam twice with two sets of scan coils (one above the other) to ensure that the beam crosses the optic axis at the FFP of the upper-objective-lens polepiece. Then, wherever the beam enters the upper-objective-lens field, it is tilted to follow a path parallel to the optic axis (so sometimes you have to tilt in order to translate!). You need computer control to do this rather complex adjustment (look ahead to Figure 9.17). Like many other procedures on a modern TEM, this adjustment is made automatically when you select a particular (in this case STEM) operating mode.

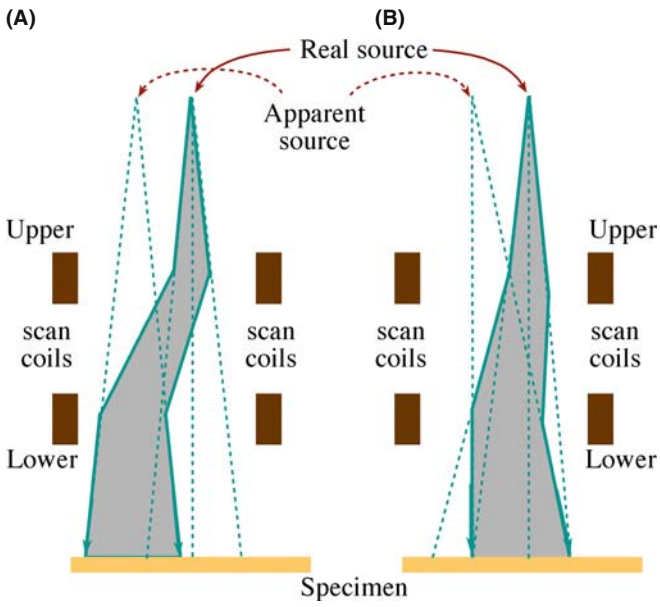


FIGURE 9.6. The use of pre-specimen scan coils for (A) traversing the beam and (B) tilting the beam. Traversing moves the beam to a different area of the specimen but it stays parallel to the optic axis. Conversely, tilting the beam illuminates the same area of the specimen, but from a different angle of incidence.

9.1.E Alignment of the C2 Aperture

Instructions for alignment vary for different TEMs so we'll simply describe the principles. Even if you won't be doing the alignment you will want to check that the microscope is correctly aligned; you can recognize if the wheels are not aligned on your car and you know it is important to balance them for best performance of the vehicle, even if you have to ask someone else to perform the task. If you want the best out of your machine, you'll want to be able to fine-tune this alignment.

If your illumination system is correctly aligned, the gun crossover is on the optic axis and the electrons can then travel down the optic axis in a straight line through the lenses and apertures until they hit the specimen. We'll assume that your gun is aligned so that the source crossover is on the optic axis of the column, as we described back in Chapter 5. Alignment of the illumination system used to be a tedious, manual affair involving tilting and translating each of the condenser lenses and centering the apertures on axis. Now most of the lens components are machined accurately enough so that only minor electronic tilts and translations are needed and usually these are computer controlled. Under these circumstances, centering the C2 aperture remains the most critical step in obtaining the best performance out of your TEM, particularly if you intend to operate in scanning mode for STEM imaging and analysis. (Other apertures in the illumination system are usually left untouched.) The C2 alignment remains a manual operation on most TEMs, although undoubtedly this will not always be the case.

Operational Procedure #1 You must always have the C2 aperture accurately centered on the optic axis of the TEM. If the aperture is misaligned, the image of the beam on the screen moves off axis and distorts as you underfocus or overfocus C2, as shown in Figure 9.7. To adjust the diaphragm so the aperture is on axis we have to alternately focus and defocus the C2 lens (we'll arbitrarily start by overfocusing but it actually doesn't matter if you under- or overfocus first). It's best to start at a low magnification so you can find the aperture image then repeat at the magnification range where you wish to operate.

- First, overfocus C2 (strengthen the lens) so the beam is spread on the screen and the image of the C2 aperture is visible (make sure any other apertures in the *imaging* system are out of the column).
- Then use the external manual drives (or computer control if you are lucky enough) to move the diaphragm so the aperture image is centered on the screen.
- Next, weaken C2 so the image of the beam is focused to its minimum diameter.

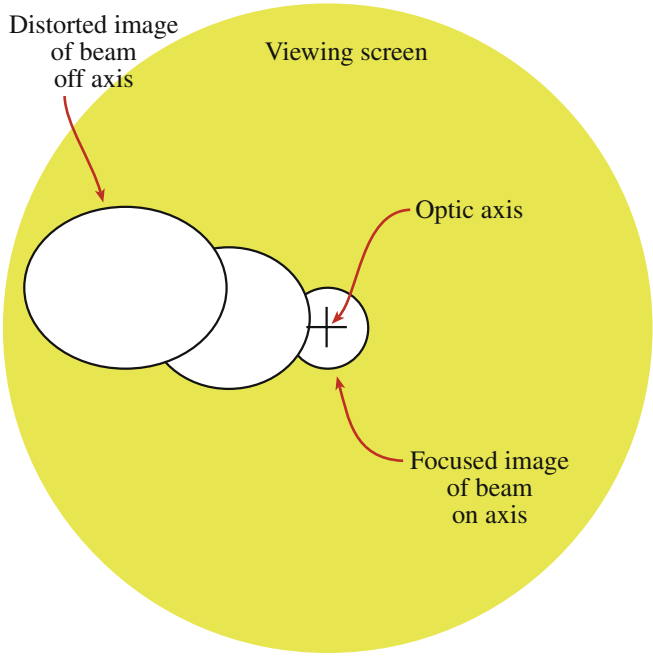


FIGURE 9.7. If the C2 aperture is misaligned, wobbling (alternately underfocusing and overfocusing the C2 lens) causes the image of the beam to sweep off axis (i.e., move across the viewing screen) and to become distorted.

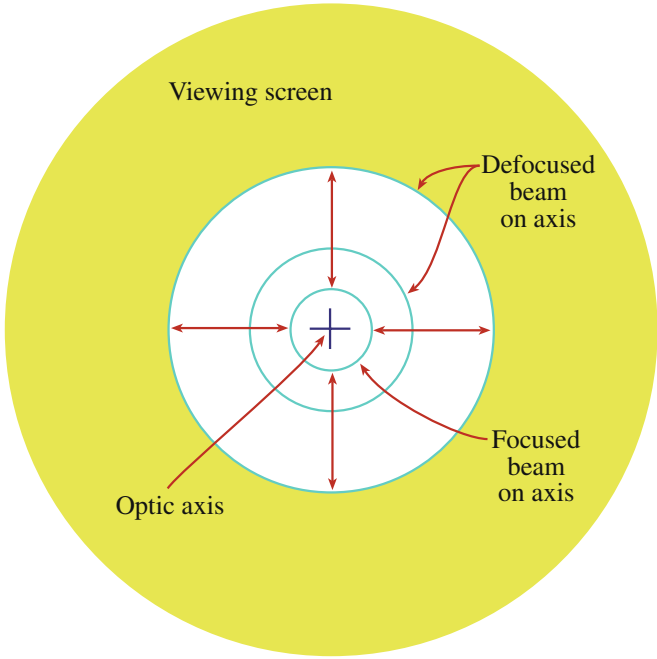


FIGURE 9.8. If the C2 aperture is aligned on axis, the image of the beam remains circular and expands or contracts about the optic axis as C2 is wobbled.

- Then, center the beam with the condenser deflector controls.
- Now underfocus the C2 lens (i.e., continue to weaken it) until you can again see the aperture image and center the image again with the external drives.
- Now strengthen the lens until the spot is focused and center it with the deflector coils.
- You have to repeat this whole operation iteratively until the image of the beam expands and contracts around the center of the screen as shown in Figure 9.8.

WOBBLING

Usually there's a control that will introduce an AC current into the lens coil, in effect 'wobbling' the lens setting either side of focus. This 'wobbler' saves you from manually underfocusing and overfocusing the lens, but think what you are doing.

In TEM mode this operation should only be necessary on the rare occasions when you change the diaphragm or the C1-lens setting but, in STEM, any change of probe size requires re-alignment of the C2 aperture.

9.1.F Condenser-Lens Defects

The illumination-system lenses suffer from all the standard lens defects we described in Chapter 6, such as aberrations and astigmatism. These defects don't really limit the operation of the TEM in parallel-beam mode, but they are crucial if you're intent on forming the finest probe possible for STEM and analytical work. Let's look at the role of each of the major defects.

Spherical Aberration This defect plays no role in limiting parallel-beam formation. However, as we discussed in Chapter 5, in adjusting the illumination system to form the finest possible probe with the maximum available current, spherical aberration in the probe-forming lens (C3) controls the minimum possible probe size. In exactly the same manner as we control the image resolution (see Chapter 6), spherical aberration limits the probe dimensions to a minimum radius (equation 6.22) of $r_{min} \sim 0.91(C_s\lambda^3)^{1/4}$. So the C3 probe-forming lens has a short focal length to minimize C_s. The final probe-limiting aperture in C2 needs to be carefully chosen to be the optimum value (equation 6.23) for the selected probe size $\alpha_{opt} = 0.77\ \lambda^{1/4}/C_s^{1/4}$. In practice, however, there are always more C1 settings than available C2 apertures, so it is not possible to choose the optimum aperture for each probe. This can cause problems if you need a specific probe size for a certain spatial resolution, as we discuss in Chapter 36. Of course, if you are lucky enough to be using a STEM with a C_s-corrected probe then this problem is minimized.

Chromatic Aberration Remember this aberration depends on the energy spread of the electrons. Since the electrons in the beam (particularly from a FEG)

have a very small energy spread we generally regard them as monochromatic and there is no detectable degradation of the probe dimensions. However, for the very best image resolution or the very best EELS, C_c-correction via monochromation of the beam can help, as we'll describe later in Section 37.7.

Astigmatism This is the most common defect in the TEM illumination system and arises either because the final limiting aperture in C2 is misaligned or it is contaminated and charging up, thus deflecting the beam. We'll assume you've centered the C2 aperture as we just described and we'll now tell you how to detect and correct any residual astigmatism due to contaminated apertures. The condenser stigmators introduce a compensating field (like a scan coil) which corrects this distortion.

Operational Procedure #2A — Detecting Astigmatism You can discern astigmatism in the illumination system by looking at an image of the electron source on the display/screen as follows

- Operate in image mode with no specimen inserted (or look through a hole in your specimen).
- Focus C2 so the beam is a minimum diameter (i.e., in spot mode with the size governed by C1).
- Adjust the C2 aperture and the beam traverses so the spot is in the middle of the screen and the image of the beam is circular, as you did when aligning the aperture.
- Wobble the C2 lens either side of the focal setting so the image of the beam expands and contracts about its minimum dimension.
- If there is astigmatism, the image is not circular, but distorts elliptically and rotates through 90° either side of focus, as shown in Figure 9.9.

Operational Procedure #2B — Correcting Astigmatism

- First overfocus the beam so you can see the effect of the astigmatism (i.e., the beam appears elliptical) in one direction.
- Then adjust the stigmators so the image appears circular.
- Now underfocus the beam and repeat the correction.
- Repeat the whole over/underfocus procedure iteratively until the image of the beam remains circular as you expand and contract it on the screen with the C2 lens (see Figure 9.8). Again, a wobbler can help here once you get the hang of the interplay between defocusing and the stigmators.
- If you can't make the image circular, you'll have to increase the range of strength of the stigmators. If you are on maximum strength, then there is too much contamination on the aperture and you either need to replace or flame-clean the diaphragm, as we described in Chapter 6.

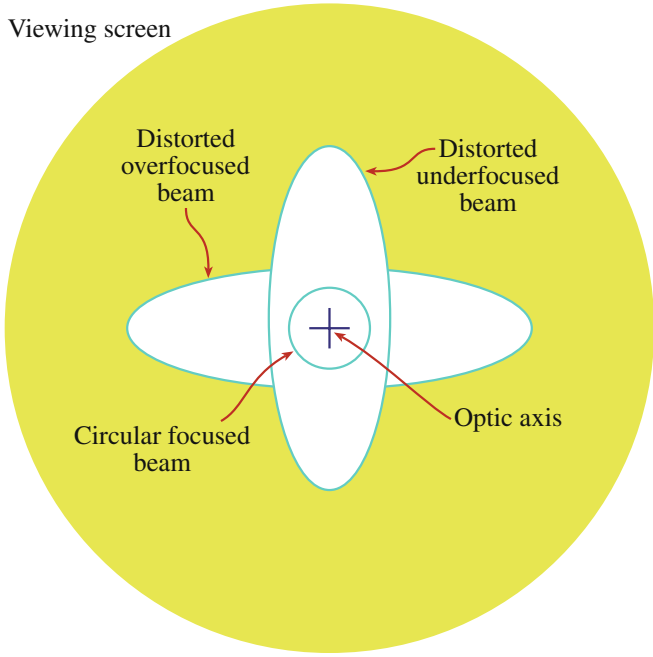


FIGURE 9.9. The effect of astigmatism in the illumination system is to distort the image of the beam elliptically as the C2 lens is wobbled. Correction of this astigmatism results in an image that remains circular as the C2 lens is defocused (as in Figure 9.8).

9.1.G Calibration

We've already described in Section 5.5 what it takes to measure the performance of the electron gun and to optimize the brightness in STEM operation so that the maximum beam current goes into the minimum probe size. We also have to calibrate the illumination system. The major variables are the probe size (and current) for various C1 settings and the convergence angle for various C2 aperture sizes.

The C1 lens strength controls the probe size at the plane of the specimen. We've described in some detail how to measure the beam dimensions at the specimen back in Section 5.5.C. in Chapter 5. Figure 9.10A shows the variation of the calculated (not measured) probe size as a function of the C1 lens setting for a typical TEM. These calculations are approximate, since they define the probe width as the FWHM and assume the C2 aperture correctly limits the Gaussian distribution. Despite these approximations you can clearly see the expected trend of decreasing probe size with increasing C1 strength. The probe current is a strong function of the probe size. If you look in the text by Goldstein et al. you'll find that the maximum probe current is proportional to (almost) the cube of the probe diameter: if you increase the probe size by a factor of 10, the current will increase by almost 1000×. This is an extraordinary dependence and explains why, particularly for probe-related TEM work (XEDS, EELS, STEM imaging, and CBED), we are very

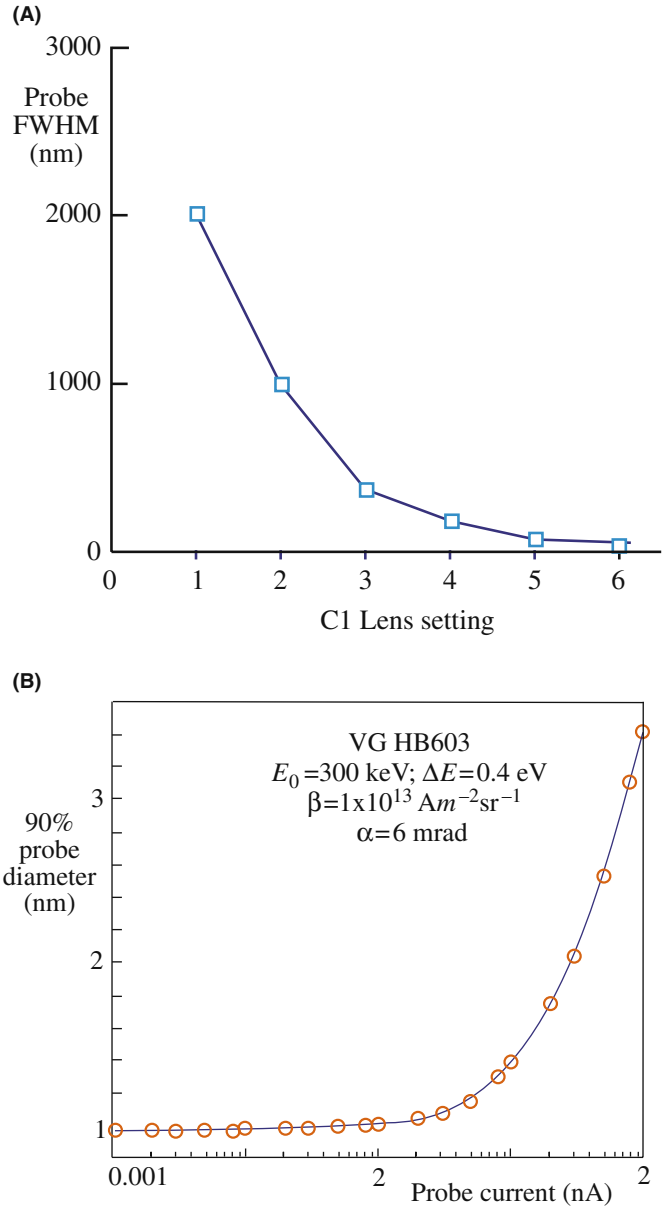


FIGURE 9.10. (A) Calibration of the illumination system requires determining the variation of the probe dimension with C1 lens strength. (B) The probe current is strongly related to the probe size. These data are from a 300-keV FEG-STEM; the dependence would be even greater in a thermionic-source TEM.

concerned about knowing and measuring the probe size and current. (There's much more about this in Part 4.) Figure 9.10B shows the variation of probe current with probe size for a FEG AEM.

The C2 aperture size governs the convergence (semi) angle α, as we also discussed in Chapter 5 when we were determining the gun brightness.

- We measure the total convergence angle 2α from a CBED pattern (Figure 5.8).
- We increase α by increasing the C2 aperture size (Figure 9.11).

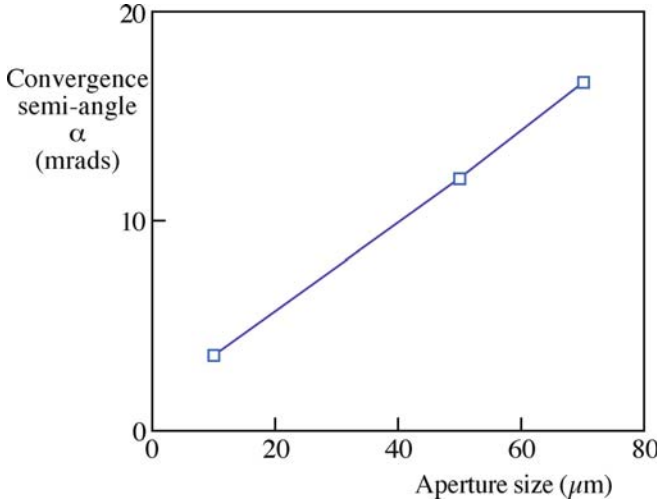


FIGURE 9.11. Increase of the beam-convergence angle, α, with increasing C2 aperture size. Increasing α will also increase the probe current.

Obviously the probe current will also be affected by the aperture size (e.g., if you double the aperture diameter you would expect four times the number of electrons to come through it). But, as you learned back in Chapter 6 (equations 6.22 and 6.23), because of the increasing lens aberrations with increasing α, there is an optimum aperture that gives the maximum probe current (and hence the best image resolution) that we just mentioned. You might well expect that, if there is a C_s corrector in the TEM illumination system, this limit can be overcome and you would be correct.

On some instruments the C2 aperture is virtual (so you have an effective aperture size), which makes it rather difficult to measure 2α. (See the Goldstein et al. text for a detailed description of this problem, which is common in SEMs.) Furthermore, if the C2 lens is excited, it can also change α and then you have to calibrate α both as a function of the aperture size and the C2 lens setting, which is an extremely tedious exercise.

9.2 THE OBJECTIVE LENS AND STAGE

This combination is the heart of the TEM. We use the stage to clamp the specimen holder in the correct position so the objective lens can form images and DPs in a reproducible manner. As we discussed in Chapter 8, there are two different types of holder, top-entry and side-entry and these determine the geometry of the polepiece and the flexibility which you have to make adjustments. We'll only consider side-entry holders here since they are standard but top-entry holders require the same adjustment of the *z*-control or specimen height.

As a practical consideration, you'll find that you would like to be able to tilt the specimen without changing its height on the optic axis. Otherwise you will be continuously using the *z*-control whenever you tilt. Clearly this means that you should ensure that the region of the specimen you want to work on is located close to the tilt axis of the specimen rod.

FOCUSING AND HEIGHT

We cannot emphasize too strongly that we need to fix the height of our specimen on the optic axis so we can always work at the same objective-lens current and thus at a fixed objective-lens magnification.

This is **really** important! This is *usually* ignored!

The central requirement here is the need to define a reference plane so that our calibrations will be reproducible. The reference plane (see Chapter 6) for a side-entry holder is called the eucentric plane. This plane is normal to the optic axis and contains the axis of the specimen-holder rod; clearly there could be many such planes. What is special about the eucentric plane is that when your specimen is located at this plane and the image is in focus, the objective-lens current is a standard reference value. The position of this plane within the objective lens is known as the eucentric height. If you put your specimen in the eucentric plane, then a point on the optic axis does not move laterally when you tilt it around the holder axis, which makes many essential imaging and diffraction operations very easy. Unfortunately, if you tilt your specimen normal to the holder axis, or rotate/traverse it off axis, then the area you're examining almost invariably moves out of the eucentric plane upon tilting and the image can easily disappear from the screen. This limitation can only be overcome by complete computer control of all five axes (three translations (x, y, z) and two tilts (parallel and orthogonal to the axis of the holder)) and such stages are often called double-eucentric. They are becoming more commonly available and are extremely useful. Get one if you can!

Operational Procedure #3A The first thing you must always do when inserting your specimen into the TEM is to ensure that it is in the eucentric plane. To do this, first find the *z*-control, which is usually a knob on the outside of the goniometer stage near the hole where you insert the holder. Then

- Select a recognizable feature in the image and position it at the center of the screen at a reasonable magnification (20–50,000×).
- Tilt the goniometer stage a few degrees clockwise or anticlockwise until the feature moves close the limit of the viewing screen. (If the slightest tilt has this effect then go to a lower magnification until you can tilt a few degrees and still see the feature on the screen.)
- Adjust the height of the specimen holder until the feature returns to the center of the screen.
- Tilt the goniometer the other way and repeat the process.
- Continue to increase the tilt both ways and compensate with the *z*-control until the image of the specimen remains stationary when you tilt the specimen through about ±30° either side of zero tilt.
- Repeat at a higher magnification as needed.

Operational Procedure #3B: Another Way to Set the Eucentricity The eucentric position of the stage is a mechanical point located in the column. As we mentioned earlier, the optimal c/o lens setting in a STEM/TEM is a very well-defined and reproducible excitation of the objective lens. This lens condition also defines a mechanical point somewhere between the polepieces. We need to make the two points (eucentric position and the optimal c/o lens excitation position) the same. To find the eucentric point when these conditions are met, you need to

- Set the lens condition to the optimal c/o operating condition.
- Focus the image using the *z*-control.
- The specimen height will now be set to the eucentric position.

With practice you can set these two points to differ by <10 μm, and < 3 μm is not uncommon.

With computer control and auto-focusing techniques becoming common, this operation can be automated. As a result, in a completely eucentric stage, your specimen doesn't move off the optic axis and remains in focus no matter around what axis it is tilted or rotated. If you don't have such a computer-controlled stage, be cautious and develop the skill to tilt and change the *z*-control simultaneously.

The eucentric plane should also be symmetrically positioned with respect to the upper and lower objective polepiece fields. At this condition, the eucentric plane coincides with the plane at which the electron beam is imaged, in both TEM and STEM modes. If the symmetric plane and the eucentric plane are not coincident, then the images and DPs will appear at different magnifications and different focus settings in TEM and STEM. Obviously this requirement has no meaning in a DSTEM where there is no TEM mode.

Ensuring coincidence of the eucentric and symmetric planes is usually carried out by the manufacturer. You can check it by comparing the focus of a DP or an image in TEM and STEM modes. You should not have to refocus the image or DP with the objective lens when you change from one mode to the other. If you find this is a problem, talk to your technical staff.

9.3 FORMING DPs AND IMAGES: THE TEM IMAGING SYSTEM

You know already that the objective lens takes the electrons emerging from the exit surface of the specimen, disperses them to create a DP in the back-focal plane (BFP), and recombines them to form an image in the image plane (see Figure 6.3). We can use this ray diagram to introduce the basic operations for forming static-beam images and DPs in the TEM. We'll then describe how to do the same thing with a scanning beam in STEM mode.

In this discussion, we will skip many of the details and concentrate on the role of the instrument. In Chapter 11 we'll discuss the details of the diffraction process and then expand these ideas in Chapters 16 through 21. We'll then discuss the images formed in the TEM in Chapters 22 through 31.

PUSH-BUTTON TEM

There are buttons to push on the console or icons to click on the computer screen by which you choose either **image mode** or **diffraction mode**. These are the primary modes, central to TEM operation. Now TEM is so easy it has never been more difficult!

The first operation that you need to master when using the TEM is viewing the DP (diffraction mode). In all the subsequent imaging modes, we'll form our images by using the objective aperture and the DP to select electrons that have been scattered through particular angles.

- *Diffraction mode*: to see the DP you have to adjust the imaging-system lenses so that the *BFP* of the objective lens acts as the object plane for the intermediate lens. Then the DP is projected onto the viewing screen/CCD as shown in Figure 9.12A.
- *Image mode*: if you want to look at an image instead, you readjust the intermediate lens so that its object plane is the *image plane* of the objective lens. Then an image is projected onto the viewing screen/CCD, as shown in Figure 9.12B.

Let's look now at the details of these two fundamental operations from the point of view of the instrument. In subsequent chapters, we will discuss how to understand the images/DPs and why we form them in the ways we do.

9.3.A Selected-Area Diffraction

As you can see from Figure 9.12A, the DP contains electrons from the whole area of the specimen that we illuminate with the beam. Such a pattern is not very useful because the specimen will often be buckled. Furthermore, the direct beam is often so intense that it will damage the viewing screen or saturate the CCD camera. So we perform a basic TEM operation which both selects a specific area of the specimen to contribute to the DP and reduces the intensity of the direct beam in the DP on the screen.

If you look at Figure 9.12A, there are two ways we could reduce the illuminated area of the specimen contributing to the DP.

- We could make the beam smaller
- We could insert an aperture above the specimen which would only permit electrons that pass through it to hit the specimen

The first option involves using C2 and/or C3 to converge the beam at the specimen. We use this approach to form CBED patterns, which we'll discuss in great detail in Chapters 20 and 21. Converging the beam destroys any parallelism, and spots in the pattern are not sharply defined but spread into disks. If we wish to obtain a DP with a parallel beam of electrons, the standard way is to use the selected-area aperture. Now, we can't insert an aperture at the specimen plane, because the specimen is already there! But if we insert an aperture in a plane *conjugate with* the specimen, i.e., in one of the image planes of an imaging lens, then it creates a virtual aperture at the plane of the specimen. This is exactly what we do. This operation is called selected-area diffraction or SAD, and is shown in Figure 9.12A.

The conjugate plane that we choose is the image plane of the objective lens. As shown in Figure 9.13, we create SADPs by inserting what we call the SAD aperture into the image plane of the objective lens and we center this aperture on the optic axis in the middle of the viewing screen. You can see this aperture if you want to change it or center it, by projecting the image plane onto the viewing screen/CCD, as we'll discuss in Section 9.3.B below.

Operational Procedure #4 The specific steps to form an SADP are as follows (while it looks complicated you'll soon find that it only takes a few seconds to do all these steps)

- Choose image mode so you can see an image of your thin specimen on the viewing screen.

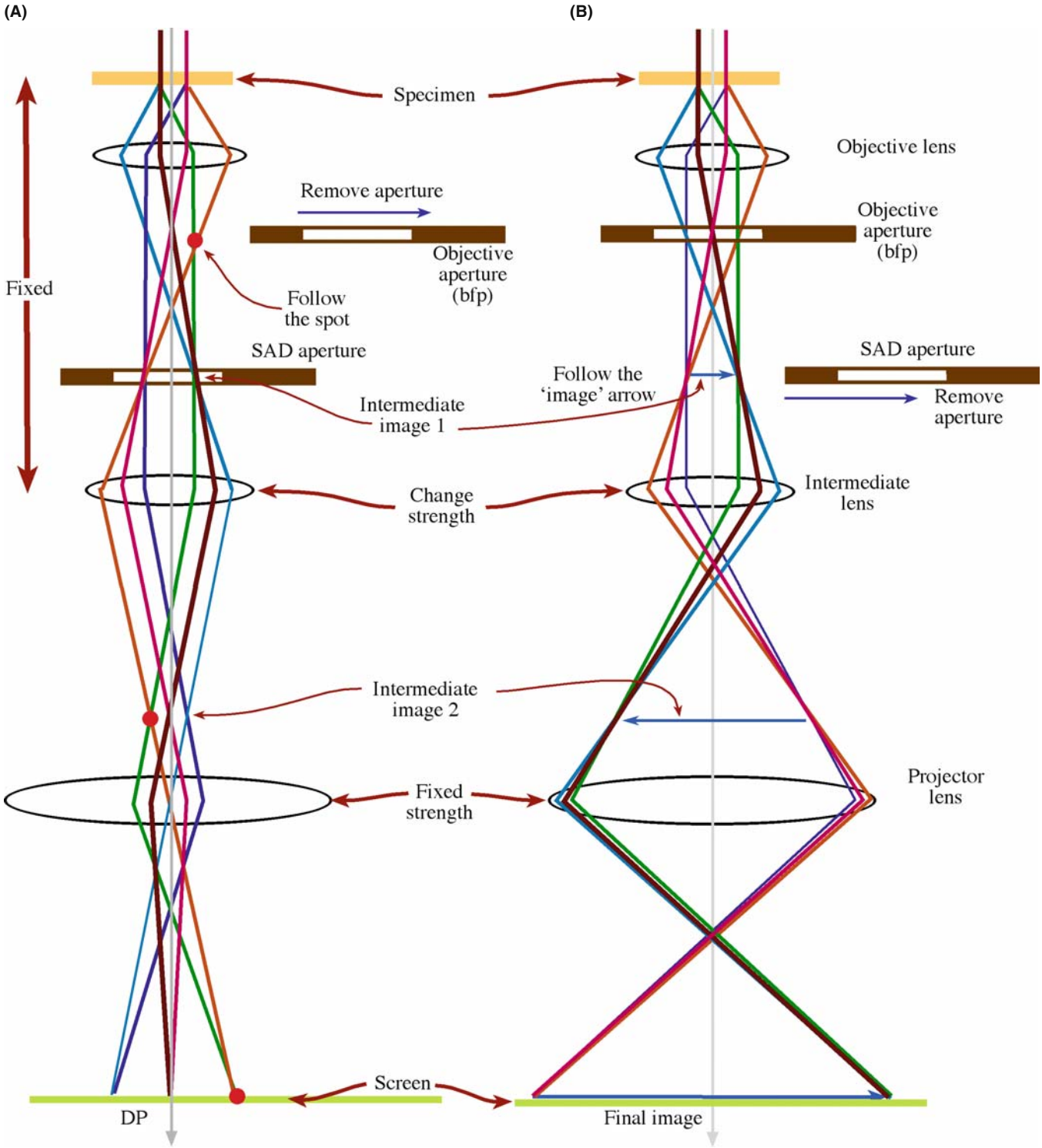


FIGURE 9.12. The two basic operations of the TEM imaging system involve (A) diffraction mode: projecting the DP onto the viewing screen and (B) image mode: projecting the image onto the screen. In each case the intermediate lens selects either the BFP (A) or the image plane (B) of the objective lens as its object. The imaging systems shown here are highly simplified. Most TEMs have many more imaging lenses, which give greater flexibility in terms of magnification and focusing range for both images and DPs. The SAD and objective diaphragms are also shown appropriately inserted or retracted. NOTE: This is a highly simplified diagram showing only three lenses. Modern TEM columns have many more lenses in their imaging systems.

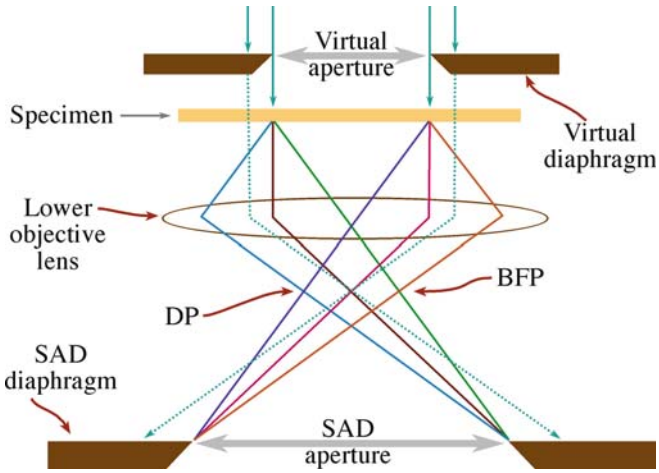


FIGURE 9.13. Ray diagram showing SADP formation: the insertion of an aperture in the image plane results in the creation of a virtual aperture in the plane of the specimen (shown here slightly above the specimen plane). Only electrons falling inside the dimensions of the virtual aperture at the entrance surface of the specimen will be allowed through into the imaging system to contribute to the SAD pattern. All other electrons (dotted lines) will hit the SAD diaphragm.

Under these conditions, with an SAD aperture inserted and the objective apertures removed, any electron that hits the specimen outside the area defined by the virtual aperture will hit the diaphragm when it travels onto the image plane. Such electrons will thus be excluded from contributing to the DP that is projected onto the viewing screen. In practice, we can't make apertures smaller than ~10 μm, and the demagnification back to the plane of the specimen is only ~25× which gives a minimum selected area of ~0.4 μm—which isn't as small as we'd like, particularly in today's nano-world. We'll discuss in Chapter 11 whether or not smaller values would be useful. As always, you really need to know these values for the TEM you are using, so do the calibration.

SAD

It is a basic principle of TEM operation that when you want to look at the DP (i.e., the **BFP** of the objective lens), you put an SAD aperture into the **image plane** of the objective lens.

- Spread the beam by underfocusing C2.
- Make sure that the objective diaphragm is retracted. The image contrast will be poor and it will be difficult to focus but that's all right at this stage.
- Insert the SAD diaphragm and start by choosing the largest available aperture to make finding the aperture easy (there may be three to five different apertures depending on the specific TEM you are using).
- If the screen then goes blank and you can't see anything, the aperture you chose is not on axis and you should lower the image magnification until you can see the aperture and center it on the axis.
- If the illuminated area of the image isn't restricted when you insert the aperture then you are at too high a magnification; so lower it until you can see the image of this aperture on the viewing screen.
- Traverse the specimen if necessary to make sure the specific region you are interested in is visible through the aperture.
- You must focus the SAD aperture by adjusting the intermediate lens so it is conjugate with (i.e., exactly in the plane of) the image of the specimen that we focused with the objective lens.
- Switch to diffraction mode and an SADP will appear on the screen.
- Focus the SADP with the diffraction-focus controls so the spots are sharp. Further underfocusing C2 may also help at this stage.
- If you need a smaller aperture to select a smaller region of the specimen, simply click the diaphragm holder in or out to choose a smaller aperture and repeat the above procedure.

The SADP is often displayed on the viewing screen at a fixed magnification so that you can recognize differences in the magnitude of **g**.

By analogy with an X-ray diffractometer or pinhole camera, we define a distance called the 'camera length' (L). We think of this length as being the distance of the recording plane from the DP (look ahead to Figure 9.23); it's analogous to a real distance in a real camera. We choose the value of L such that inter-spot or ring spacings in the DP are easily discernible on the viewing screen and on the recording medium. This magnification can be changed by adjusting the intermediate lenses. We'll describe how we calibrate this magnification in Section 9.6.B.

At this stage, however, it's worth noting that, with a c/o lens, calibration of L can be a challenge because changing the condenser-lens focus as we just described changes the diffraction crossover and, if you re-focus the SADP with the intermediate lens, then you'll measure a different L.

Beware: in all the early TEM books, SAD is the only standard diffraction technique. As a result, some microscopists only use SAD to obtain diffraction information. However, CBED, which we discuss in Chapters 20 and 21, provides a great deal of complementary diffraction information and *must* be used by all TEM operators in the materials and nanotechnology fields. But there are still certain times when your best option is to form an SADP

- When you need to select a spot from which to form a BF or DF image (see next section)
- When diffraction spots are very close to one another and would overlap in CBED patterns (see examples in Chapters 23 and 24)

- When you are looking for fine structure in the DP such as streaks (see Chapter 17)
- When your specimen is beam sensitive

On all other occasions, when the diffraction maxima provide the most important information in the pattern, you should use CBED.

9.3.B Bright-Field and Dark-Field Imaging

When the SADP is projected onto the viewing screen/CCD, we can use the pattern to perform the two most basic imaging operations in the TEM. No matter what kind of specimen you're looking at, the SADP will contain a bright central spot that contains the direct-beam electrons and some scattered electrons (as shown in Figure 2.13A–C), the distribution of which will depend on the nature of your specimen.

IMAGING
It is another basic principle of TEM operation that if you want to view an image (i.e., the **image plane** of the objective lens), you insert an aperture called the objective aperture into the **BFP** of the objective lens.

Operational Procedure #5 When we form images in the TEM, we either form an image using the central spot or we use some or all of the scattered electrons. The way we choose which electrons form the image is

- Starting with both SAD and objective diaphragm holders retracted, select image mode to observe an image of the desired region of your specimen on the screen. It will not be a good image because there is no objective aperture inserted, but that is OK at this stage.
- Form an SADP of a selected region as we described in operational procedure #4.
- Insert the objective diaphragm so an aperture appears in the BFP of the objective lens, thus blocking out most of the DP, except that small area which is visible through the aperture. As with the SAD operation above, we use the external drives to move the aperture.
- Adjust the external diaphragm drives so the aperture selects the direct beam on the optic axis.
  This centering procedure is absolutely critical to forming the best image. If you have binoculars, view the TEM screen at a much higher magnification to ensure the best possible centering; otherwise do the equivalent using the digital image.
- Return to image mode, remove the SAD aperture, and focus the image with the objective lens. We call the resultant image, formed by the direct-beam electrons, a bright-field (BF) image (Figure 9.14A).
- If instead we choose scattered electrons (e.g., a specific diffracted beam (usually) or a portion of a diffraction ring), then we call the image a dark-field (DF) image (Figure 9.14B).

You can view the BF and DF images at any magnification simply by adjusting the intermediate lenses of the microscope and typical magnification ranges will be ~25,000× to 100,000× for a wide field of view, but up to 10^6× for high-resolution images. Usually you'll have to calibrate the actual magnification and also be able to relate directions in the image at any magnification to directions in the DP at a fixed camera length. These are the two basic calibrations required for any TEM.

BF AND DF
Select the direct beam to form a BF image.
Select only electrons that are not in the direct beam to form a DF image.

The insertion and removal of the SAD and objective apertures can be confusing to the beginner and often the wrong aperture is inserted or not removed when it should be. You have to practice obtaining SADPs and BF/DF images to get used to what aperture should be inserted and when. Both apertures are inserted below the objective lens. The objective aperture goes into the BFP, so it is closer to the lens (i.e., higher up the column) than the SAD aperture, which is in the image plane. Remember that if you're looking at a DP the (lower) SAD aperture should be inserted and the (upper) objective aperture removed. If you want to look at an image, the objective aperture should be inserted and the SAD aperture removed.

OBJECTIVE APERTURE
This is the most important aperture in the TEM. When inserted, its size controls the collection angle (β); hence, it determines the effect of the aberrations of the most important lens in the instrument and thus directly influences the resolution.

9.3.C Centered Dark-Field Operation

If you look at Figure 9.14B, the electrons that are selected by the aperture travel off the optic axis, since we displace the aperture to select the scattered electrons. The more off-axis the electrons are the greater the aberrations and astigmatism they suffer. Therefore, such a displaced-aperture DF (DADF) image is difficult to focus on an older TEM. You'll find that the image will move on the screen as you adjust the objective-lens strength. While there may be situations where you want to use such a DADF image, almost invariably you have to get any

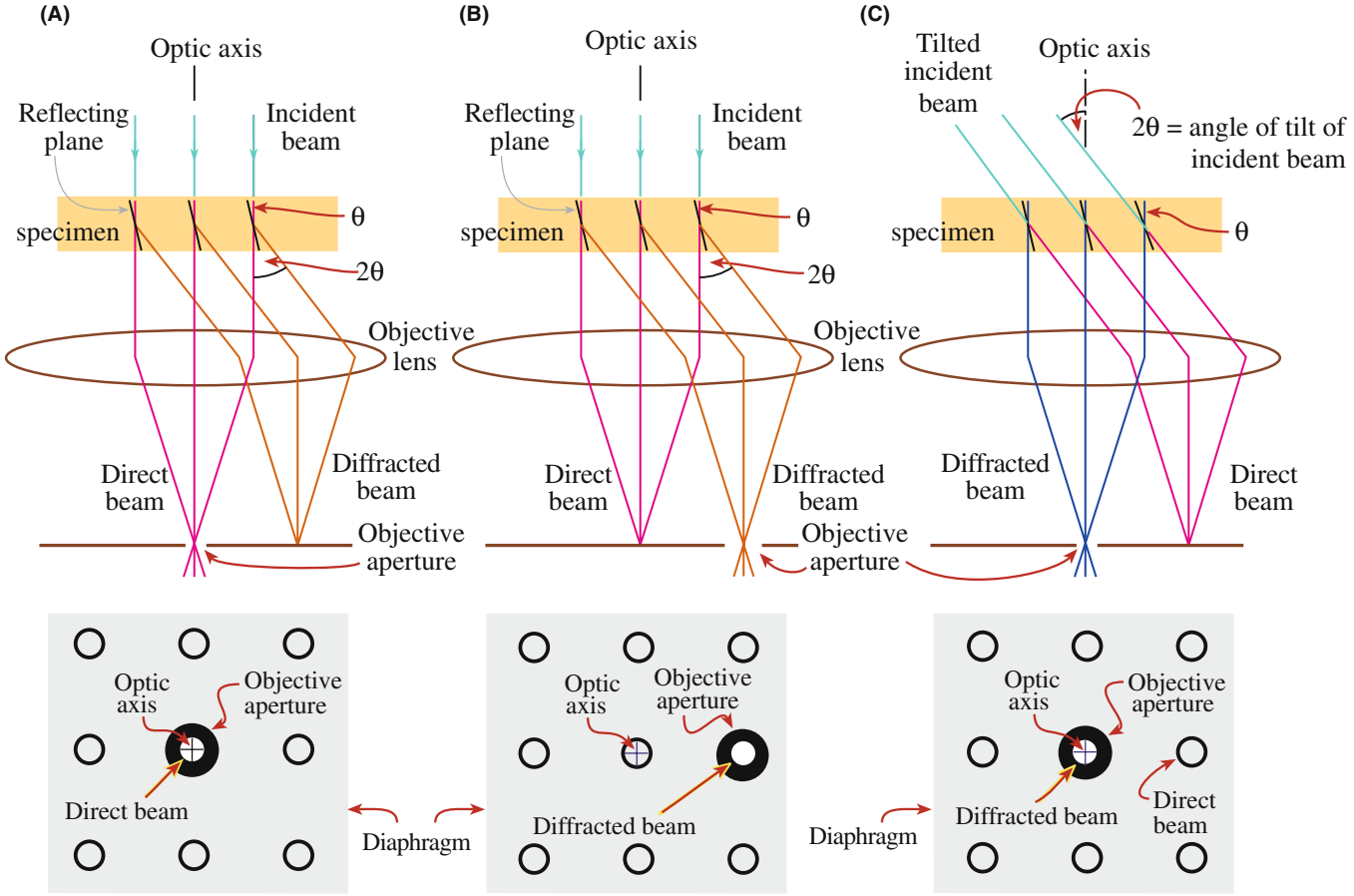


FIGURE 9.14. Ray diagrams showing how the objective lens and objective aperture are used in combination to produce (A) a BF image formed from the direct electron beam, (B) a displaced-aperture DF image formed with a specific off-axis scattered beam, and (C) a CDF image where the incident beam is tilted so that the scattered beam emerges on the optic axis. The area of the DP selected by the objective aperture, as seen on the viewing screen, is shown below each ray diagram. (Images comparing BF and DF are given in Part 3.)

scattered electrons back on the optic axis and then form a DF image under this condition. This operation is called centered dark-field (CDF) imaging. It is the conventional way to do DF imaging so we'll tell you how it's done. If you have a new TEM (with a small C_s) you should consider using only DADF for SAD because it's easier and you've eliminated the reason for forming CDF images!

Operational Procedure #6

- Start with an SADP on the screen and instead of moving the objective diaphragm so the aperture selects scattered electrons, adjust the aperture so it is on axis, as for BF imaging.
- Retract the objective diaphragm drive so you can see the SADP.
- Switch on and adjust the beam-tilt (DF) potentiometers above the objective lens so that the scattered-electron beam that you wish to use to form the CDF image moves toward the central, on-axis position.
- Looking through the binoculars, switch the potentiometers on and off to ensure that the scattered beam exactly superimposes where the direct beam was when the potentiometers are on and the direct beam is back on axis when the potentiometers are off.
- Re-introduce the objective diaphragm and check that the aperture is still centered around the on-axis scattered beam.
  As with BF imaging, this aperture-centering procedure is absolutely critical to forming the best CDF image.
- Select image mode again and focus the CDF image.

What you are doing here is making the incident beam hit the specimen at an angle equal and opposite to the scattering angle. In this way the scattered electrons will now travel down the optic axis, as shown in Figure 9.14C.

We'll return to BF, CDF, DADF, and SAD operations when we discuss specific contrast mechanisms that occur in TEM images in Chapter 22.

9.3.D Hollow-Cone Diffraction and Dark-Field Imaging

One obvious limitation to either DADF or centered CDF imaging is that we are only using a small fraction of the scattered electrons. When that fraction is a specific diffraction spot from a crystalline specimen that can indeed be very useful, as we'll see in Part 3 of the book. However, if, for example, we want to see *all* the portions of the specimen or *all* the precipitate phases that are diffracting into a *set* of diffraction spots (e.g., {111} rather than (111)) then we can either take separate CDF images from *all* the individual (111) reflections in the SADP (which is obviously a pain) or we can carry out hollow-cone (also called conical) diffraction and DF imaging. Likewise, if your specimen is micro/nanocrystalline so it generates a ring pattern, or if it is amorphous so diffuse-intensity rings are formed, then we can use the same technique to maximize the information in the images. There are two ways to do hollow-cone DF imaging: hardware and software. In essence, in either method, a specific set of diffracted beams is collected by the objective aperture when the conical-scanning beam satisfies a particular Bragg reflection condition (see Section 3.10 and Chapter 11).

- *Hardware*: Use an annular condenser aperture such that a cone of electrons illuminates the specimen at a fixed angle to the optic axis. Annular apertures fabricated by FIB techniques (see Chapter 10) are becoming more commonly available. (Think about how to make an annular aperture so that the middle doesn't just fall out of the doughnut shape leaving just a large circular hole?) So then we basically have multiple DPs formed; the direct (000) beam travels through the specimen and emerges off axis and a set of diffracted beams (depending on the cone angle) is always scattered on axis.
- *Software*: Use computer control of the scan coils to spin the incident (and thus the direct) beam around the optic axis. This method is obviously more flexible because, while an annular C2 aperture gives a fixed cone (semi) angle, the scan coils can be adjusted to give a variable cone angle. Thus, with the software approach, all the electrons scattered at a specific angle can be integrated into a single DF image which you can envisage as spinning (Figure 9.14C) around so that a set of diffraction spots always rotates through the optic axis. The principle of hollow-cone diffraction and imaging is shown in Figure 9.15.

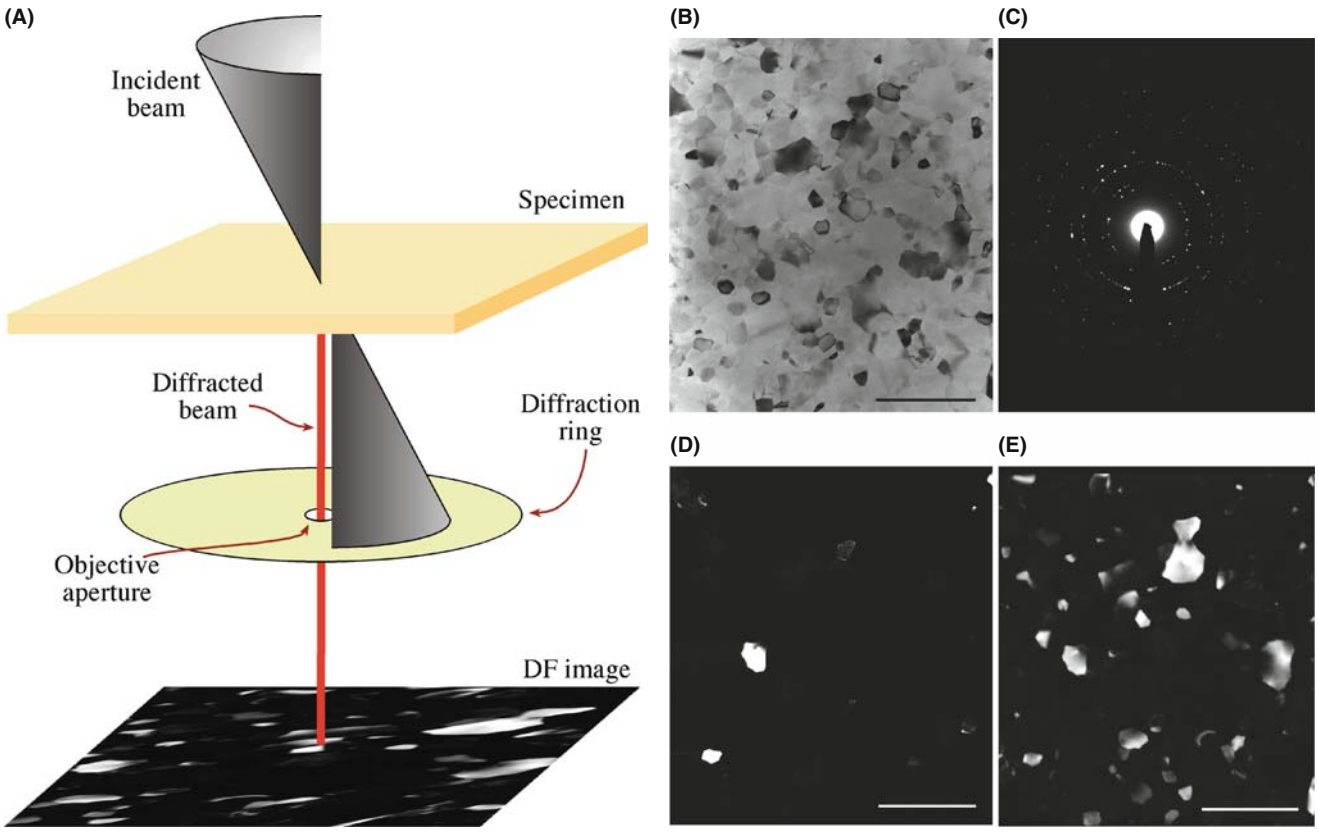


FIGURE 9.15. (A) Ray diagram showing hollow-cone illumination conditions. The direct beam is always off axis but electrons diffracted at the cone angle are always scattered on axis. (B) A BF image of a nanocrystalline Al film. (C) A hollow-cone SAD pattern from the film. Thus, while a single (220) CDF image (D) reveals only a couple of strongly diffracting crystals, a hollow-cone DF image (E) from all the {220} reflections shows diffracted intensity from dozens of grains. The scale bar is 500 nm.

Computer control of the scan coils allows you to select the angle, i.e., the radius of the circle, thus selecting which (*hkl*) ring is to be on the optic axis. You can also control the speed of the rotation so you can synchronize a single rotation with sufficient exposure time to record the DF image. This process is analogous to the normal DF tilt controls that shift the DP in the *x*–*y* plane (by tilting the incident beam as in Figure 9.14C) being replaced with *r*–θ axis controls. If you observe the rotating DP with the objective aperture removed, the direct (000) beam describes a circle, radius *r*, around the optic axis. Any point in the DP that is distance *r* from the (000) spot will at some time pass through the optic axis during the rotation of the DP. An example of a hollow-cone DP (not spinning) is shown in Figure 9.15C. The SAD aperture was about the size of the image (~ 50 μm across) and the objective aperture was small enough to permit diffracted intensity from only one ring to be collected. Figure 9.15B, D and E shows, respectively, a BF image of a polycrystalline metal film from which the DP in Figure 9.15C was obtained, a conventional CDF from a few spots in the DP ring and the hollow-cone DF, revealing many more diffracting grains. You usually have to take a series of time exposures to get a good hollow-cone image. Figure 9.15E was a 20-second exposure during which time the DP rotated about 1000 times through the objective aperture!

Hollow-cone DPs are also discussed in Section 18.6. A closely related technique, precession diffraction, is covered in Section 18.8. The STEM equivalent of hollow-cone DF imaging is annular-DF imaging, which we'll talk about in the next section.

9.4 FORMING DPs AND IMAGES: THE STEM IMAGING SYSTEM

If you want to use a fine probe to form STEM images then the objective-lens optics are a little more complex than in TEM. The key feature to remember is that the scanning beam must not change direction as the beam is scanned (unlike in an SEM where the scanning beam simply pivots about a point above the specimen). If the incident direction varies then the electron scattering (particularly the diffraction) processes would change as the beam intercepts the specimen at different angles. So interpreting the image contrast would be rather difficult to say the least.

STEM

The beam must scan parallel to the optic axis at all times so that it mimics the parallel beam in a TEM even though it's scanning.

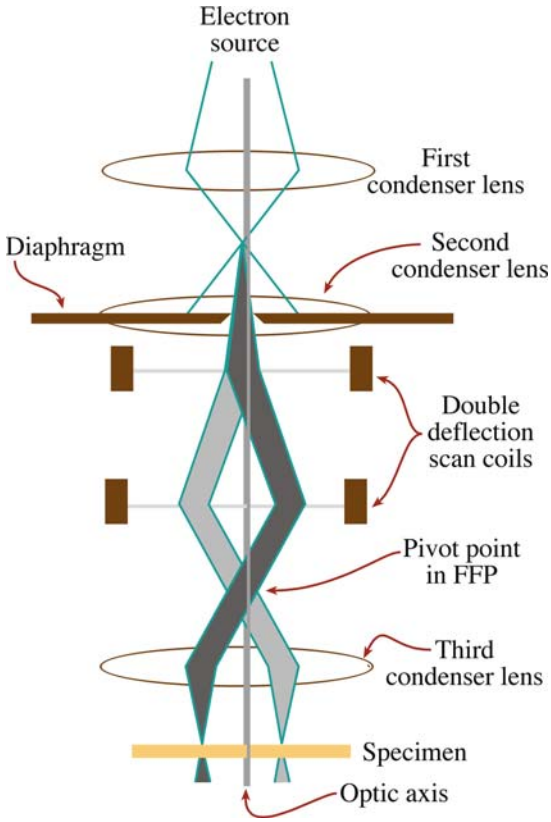


FIGURE 9.16. Scanning the convergent probe for STEM image formation using two pairs of scan coils between the C2 lens (usually switched off) and the upper-objective polepiece. The double-deflection process ensures that the probe remains parallel to the optic axis as it scans across the specimen surface.

As we show in Figure 9.16, the way we achieve parallel incidence is to use two pairs of scan coils to pivot the beam about the FFP of the upper objective (C3) polepiece. The C3 lens then ensures that all electrons emerging from the pivot point are brought parallel to the optic axis and an image of the C1 lens crossover is formed in the specimen plane. Now, if the objective lens is symmetrical, and the lower objective polepiece is similarly strong, then a stationary DP is formed in the BFP (this pattern does not move, even though the beam is scanning, since it is conjugate with the FFP, as shown in Figure 9.17). If we stop the beam from scanning, then we have a CBED pattern in the BFP and we can project that onto the TEM computer screen if we wish. So let's first discuss how to form STEM images.

THE STEM IMAGING LENS?

The STEM image quality depends on the probe. The probe has aberrations because we use a lens to form it. So the STEM image quality does depend on a lens (just not on an imaging lens).

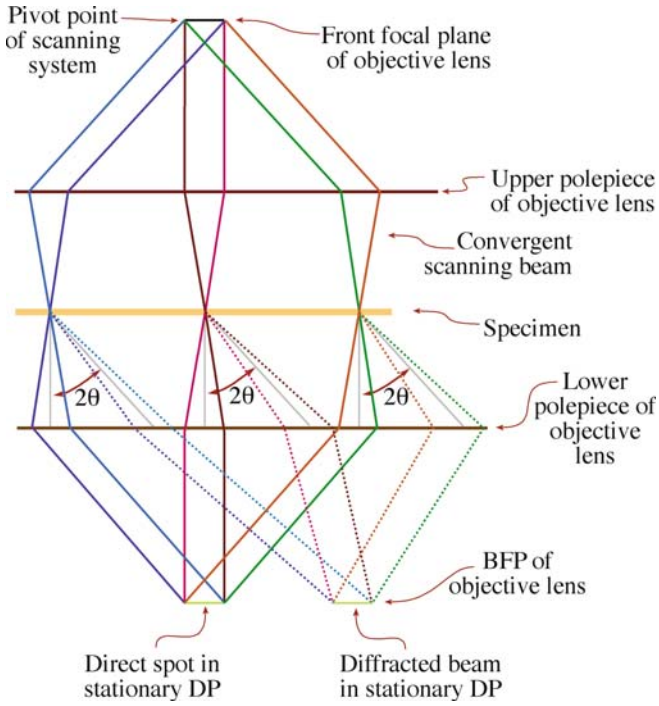


FIGURE 9.17. The creation of a stationary (convergent-beam) DP in the BFP of the objective lens is a necessary prerequisite for STEM imaging. Note that electrons scattered through the same angle (2θ) at different points in the specimen are focused at the same point in the BFP.

One potentially very big advantage for STEM is that, just like in an SEM, we don't use lenses to form the image. So defects in the *imaging lenses* do not affect our image resolution, which is limited by the beam dimensions. Hence chromatic aberration which (as we saw back in Section 6.5.B) can seriously limit TEM image resolution, is absent in STEM images. This is a great advantage if you're dealing with a thick specimen. However, there are drawbacks also, as we'll discuss below and in Part 3, and STEM images aren't widely used, particularly for crystalline specimens.

9.4.A Bright-Field STEM Images

Image formation in the scanning mode is fundamentally different from static-beam TEM image formation. As you've just seen in the TEM, we select a portion of the electrons emerging from an area of the specimen and project that distribution onto a screen. The principle of scanning-image formation is shown in Figure 9.18. Simply stated, we scan the beam on the specimen by adjusting the scan coils; these same coils are used to scan the computer display synchronously. The electron detector acts as the interface between the electrons coming from the specimen and the image viewed on the display screen. Since it takes up to 2048 scan lines to build up an image on the recording screen, the whole process of creating a STEM image is much slower than TEM imaging: it's serial recording instead of parallel recording.

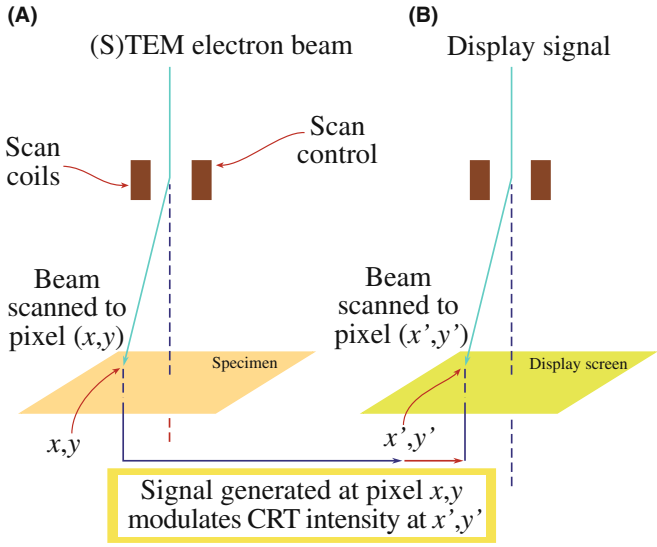


FIGURE 9.18. The principle of forming a scanning image, showing how the same scan coils in the microscope control (A) the beam-scan on the specimen and (B) the beam-scan on the computer display screen of the STEM. Thus no lenses are required to form the image.

THE STEM SIGNAL

The signal is generated at a point on the specimen, detected, amplified, and a corresponding signal displayed at an equivalent point on the computer display. The image builds up over several seconds or even minutes.

This process is exactly the same principle as used in any scanning-beam instrument such as an SEM or an STM (scanning-tunneling microscope). Remember that to form a TEM-BF image, we inserted an aperture into the plane of the TEM DP and only allowed the direct electrons through it into the imaging system. In STEM mode we use an electron detector, in exactly the same way as we use the aperture: we only allow the electrons that we want to contribute to the image to hit the detector. So we insert a BF (either a semiconductor or scintillator-PM) detector onto the axis of the microscope and it intercepts the direct-beam electrons no matter where the beam is scanning on the specimen, as shown in Figure 9.19A. So a variable, direct-beam signal travels from the detector via an amplification system to modulate the signal on the computer display, thus building up a BF image as also shown in Figure 9.19D.

BF DETECTOR

The BF detector picks up the direct beam which varies in intensity depending on the specific point on the specimen illuminated by the probe at that specific time.

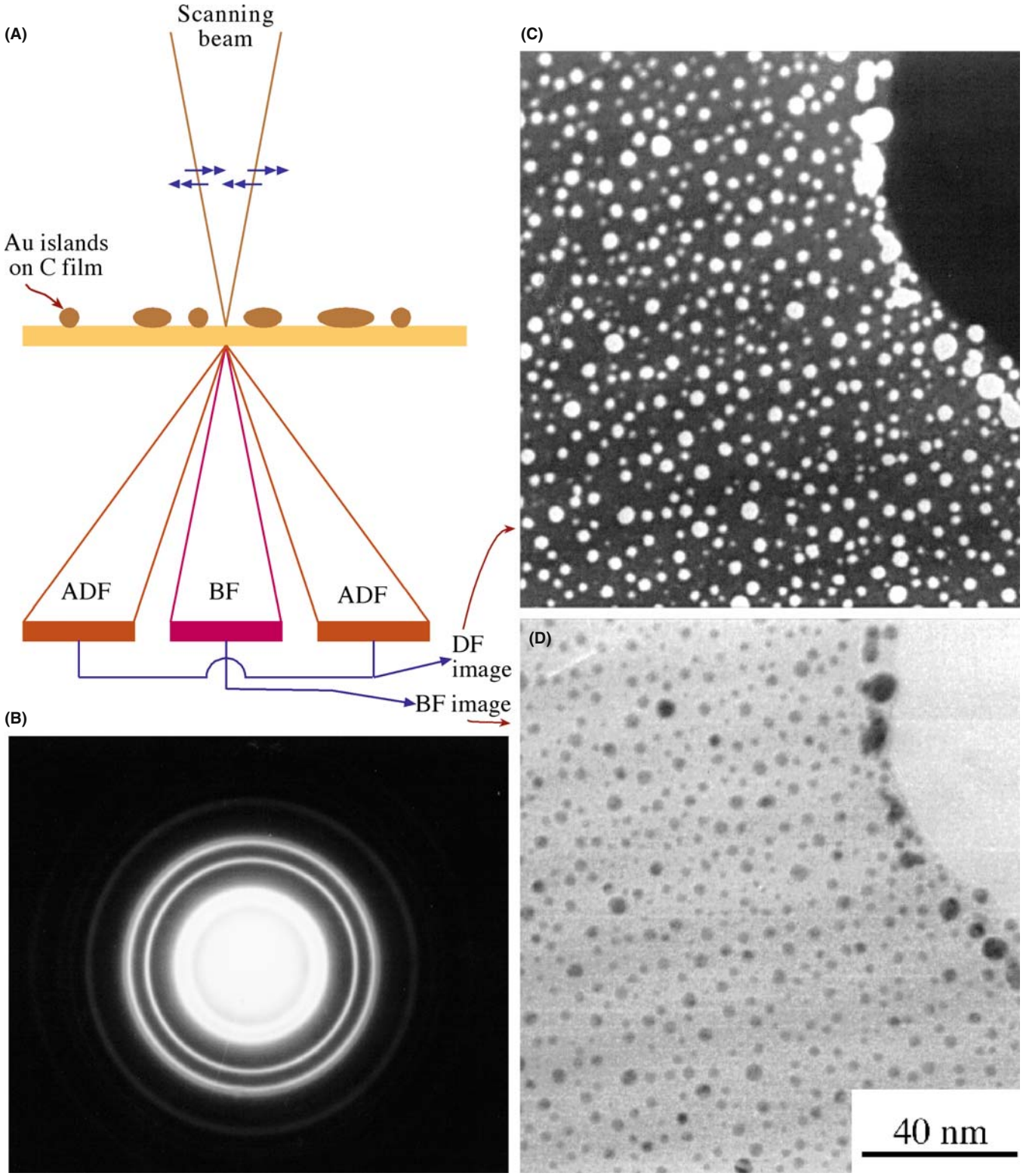


FIGURE 9.19. STEM image formation: A BF detector is placed in a conjugate plane to the BFP to intercept the direct beam (A) and a concentric annular DF detector intercepts the diffracted electrons whose distribution is shown in the SAD pattern in (B). The signals from either detector are amplified and modulate the STEM computer display. The specimen (Au islands on a C film) gives complementary ADF (C) and BF (D) images.

Now in a TEM we can't physically put the detector in the BFP of the objective lens to form a STEM image, because it would interfere with the objective aperture. Therefore, we usually insert the detector into a conjugate plane to the stationary DP (Figure 9.19B). So when you form a STEM image in a TEM, you operate the TEM in diffraction mode and insert a detector into the viewing chamber of the TEM, either above or below (in which case you raise) the screen. The stationary DP falls on the detector and the signal goes to the display. In a DSTEM, there may not be any imaging-system (or post-specimen) lenses, in which case the detector is indeed positioned immediately after the objective lens. Much of what we've just said is automatically done when you 'hit the STEM button'. The message is the same: understand what is happening in your microscope and why.

9.4.B Dark-Field STEM Images

The approach is analogous to that of TEM. We form a DF image by selecting any or all of the scattered electrons, rather than the direct-beam electrons. Remember, in a TEM we tilt the incident beam so the scattered electrons we want to form the image travel down the optic axis and are selected by the objective aperture. In a STEM, we do things rather differently.

DF STEM

If we want a specific beam of scattered electrons to fall on the BF detector, we can simply shift the stationary DP so that the scattered beam is on the optic axis and hits the BF detector.

It's simple to do this with the DP centering controls or you could also displace the C2 aperture. The former is to be preferred since doing the latter misaligns the illumination system.

9.4.C Annular Dark-Field Images

Rather than using the BF detector for DF imaging, we usually use an annular detector, which surrounds the BF detector, and then all the scattered electrons fall onto that detector. We call this process annular dark-field (ADF) imaging and it has certain advantages, depending on the contrast mechanism operating in the specimen, as we'll see in Chapter 22. As we show in Figure 9.19A, the ADF detector is centered on the optic axis and has a hole in the middle, within which the BF detector sits. The resultant ADF image in this simple example (Figure 9.19C) is complementary to the BF image (Figure 9.19D). As you'll see in Chapter 22, we can also use another annular detector that sits around the ADF and picks up the electrons scattered out to even higher angles forming so-called high-angle (HA) ADF (or Z-contrast) images in which Rutherford-scattering effects are maximized and diffraction-contrast effects are smoothed out.

We can take this idea further and make detectors of any size or shape we wish. For example, we could design a detector in which the annulus is split into two halves or four quadrants and electrically isolate each part of the detector. Then we can form different images from electrons that fall on different parts of the detector. It's impossible to do this in a TEM, because the objective aperture that does the selecting is a hole and can't be cut up like a semiconductor detector. We'll talk more about these kinds of detectors when we discuss specific contrast mechanisms in TEM and STEM images in Chapter 22.

9.4.D Magnification in STEM

All the STEM images that we have just described appear on the computer screen at a magnification that is controlled by the scan dimensions on the specimen, *not* the lenses of the TEM. This is a fundamental difference between scanning and static image formation.

STEM IMAGES

STEM images are *NOT* magnified by lenses.

Because scanning images are not magnified by lenses, they are not affected by aberrations in the imaging lenses. They are, however, affected by aberrations of the probe itself and, therefore, can be improved by (illumination-system) aberration correctors.

If the scanned area on the specimen is 10 mm × 10 mm, and the resultant image is displayed on a computer screen with an area 100 mm by 100 mm, then the image magnification is 10×. If the scan dimension is reduced to 1 mm, the magnification on that same screen is 100×, and so on, up to magnifications in excess of $10^7\times$, which are common in dedicated STEMs. As with the TEM, we have to calibrate the STEM magnification and the camera length of the DP we use to create the images.

9.5 ALIGNMENT AND STIGMATION

9.5.A Lens Rotation Centers

You only need to perform two alignments to ensure that the imaging system is operating correctly. By far the most important is the alignment of the objective lens center of rotation and the second is the alignment of the DP on the optic axis. To get the best out of your TEM, you *must* master these two fundamental alignments.

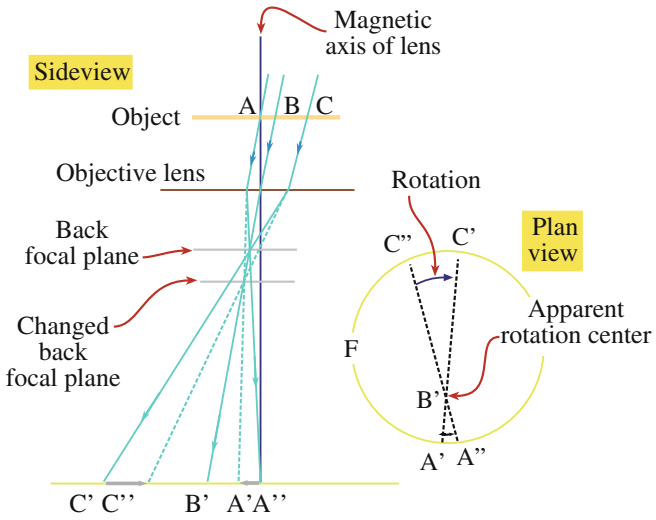


FIGURE 9.20. When the objective-lens center of rotation is misaligned, the image appears to rotate about a point away from the center of the viewing screen when the lens is wobbled about focus. When the rotation center is correctly aligned, the image will rotate about the center of the screen.

Basically, the idea of the objective-lens rotation alignment is to ensure that the objective lens field is centered around the optic axis, so that the direct-beam electrons emerging from the specimen see a symmetric field as they pass through the lens. If the field is off-center, then the electrons will move off axis, suffer more aberrations, and your image will rotate about a position off-axis as you change the objective lens (focus), as shown schematically in Figure 9.20.

Operational Procedure #7

- To center the objective rotation, start at a relatively low magnification (say 10,000×), select an obvious reference point in the image and move it to the middle of the screen, and observe the way the point rotates as you wobble the objective lens from overfocus to underfocus and back again. If the point rotates without moving off center, the lens is aligned, but check the accuracy of the alignment by going to much higher magnifications ($>10^5\times$).
- If the point moves off center at 10,000× then use the beam tilts to move the point in the image that *is* the center of rotation to the middle of the screen while continuously wobbling the objective. Repeat the process at higher magnifications.
- Above $\sim 10^5\times$ the wobbler may introduce too large a rotation, so you may have to defocus the objective lens manually. The actual steps to do this are instrument-dependent, so consult the manufacturer's handbook. This process is also called 'current centering.'
- When the image wobbles (rotates) about the center of the screen at magnifications $>10^5\times$, the objective lens rotation center is well aligned. The higher the magnification at which you can achieve this, the better the alignment and the better quality all your pictures will be.

This operation may be computer controlled in modern TEMs. In some instruments you can also perform 'voltage centering' in which a varying voltage is applied to the gun and the objective lens is aligned to ensure that the electrons remain on axis through the lens as their energy varies. Not all instruments are capable of this alignment.

If you change the DP magnification (i.e., the camera length, L) the whole pattern will move off axis if the diffraction center is misaligned. To align the center you have to adjust the projector lens until the central spot in the DP is on axis and it rotates around the axis as L is changed. Check the manufacturer's handbook on exactly how to do this alignment.

Centering the DP is useful for STEM-image formation, since you have to center the DP such that the direct beam hits the BF detector and the scattered beams hit the ADF detector. Apart from this simple operation, the STEM imaging system needs no lens alignment.

9.5.B Correction of Astigmatism in the Imaging Lenses

After you've centered the image and DP, the main cause of problems in the imaging system is astigmatism in the objective and intermediate lenses.

ASTIGMATISM

Objective-lens astigmatism occurs if the objective aperture is misaligned, so you must always carefully center the aperture on the optic axis, symmetrically around the electron beam used to form the BF or DF image.

Despite careful centering of the objective aperture, residual contamination may also cause astigmatism and then you have to use the objective stigmators to introduce a compensatory field. You'll find that the effects of objective astigmatism are harder to see than condenser astigmatism, which is easily visible on the screen as we just described. Correct objective astigmatism as follows:

Operational Procedure #8 Either look for a small hole in your specimen or look at a rough corner where the specimen edge curves through $> 90^\circ$. Ideally you might use a holey carbon film to correct residual astigmatism, *before* you insert your specimen, especially while learning this procedure, which we illustrate in Figure 9.21.

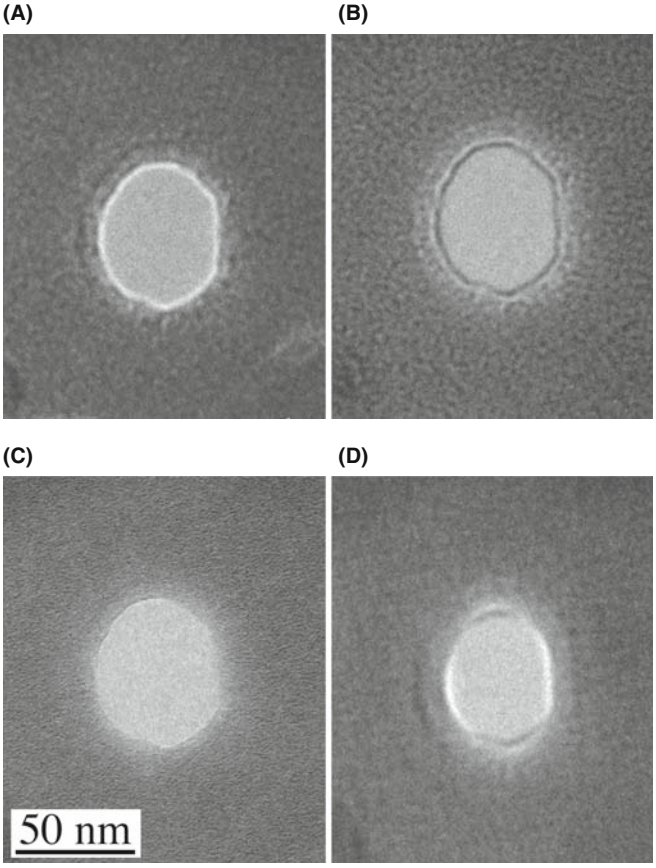


FIGURE 9.21. The image of a hole in an amorphous carbon film illuminated with a parallel beam showing that (A) with the beam underfocused, a bright Fresnel fringe is visible; (B) with the beam overfocused, a dark fringe is visible; (C) at exact focus there is no fringe; and (D) residual astigmatism distorts the fringe. Correcting the astigmatism means changing any image similar to (D) to one similar to (A) or (B).

- Often you can only see objective astigmatism at the highest magnifications so form a BF image of the hole or the specimen edge at high magnification ($>10^5\times$). First, adjust the illumination system to ensure a reasonably parallel beam. (Remember: how to do this depends on whether you have a c/o lens or not.)
- Now defocus the objective lens (either over- or underfocus; it doesn't matter). A Fresnel fringe (which is a phase-contrast effect, as we'll see in Chapter 27) should be visible at the thin edge of your specimen.
- Alternately underfocus and overfocus the objective lens (i.e., use the wobbler again).
- If there is astigmatism, you should see streaking in the image and this streaking effect will rotate through 90° as you wobble either side of exact focus. The streaking is most easily seen by watching the Fresnel fringe at the edge, as we describe below for Figure 9.21.
- Adjust the objective stigmators to compensate for the streaking at overfocus and then again at underfocus (until you're very skilled, you'll probably have to switch off the wobbler to do this and manually defocus the objective).
- Repeat these steps until there is no obvious image streaking as you defocus and the image merely blurs when going out of focus.
- Go to a higher magnification and repeat the procedure. As with the rotation center, the higher the magnification at which you can correct the astigmatism, the better and $>250{,}000\times$ should be possible. But remember that at higher magnifications the image intensity will be reduced. Don't try and compensate for this reduced intensity by condensing the beam because doing so will destroy the parallelism, thus reducing the Fresnel-fringe (phase) contrast.

As shown in Figure 9.21A, when you underfocus the objective lens, there is a bright Fresnel fringe round the edge of the hole. If this fringe is uniform around the hole, then there is no astigmatism. If the fringe varies in intensity, as in Figure 9.21D, then the focus of the lens is changing around the hole because of astigmatism. Then you have to adjust the objective stigmators to make the fringe uniform. The same operation must be repeated at overfocus, when there is a dark fringe around the edge of the hole (Figure 9.21B). At exact focus, you should see no fringe and the image contrast is minimized (Figure 9.21C).

FRESNEL FRINGES
Underfocus = bright fringe
Overfocus = dark fringe

This method of correcting the astigmatism is reasonable at magnifications up to several hundred thousand times.

In practice, if you're working at such high magnifications, you'll probably have to check the astigmatism throughout your TEM session, so you should get used to looking at the Fresnel fringes on a thin, curved edge of your specimen rather than the ideal holey carbon film.

For high-resolution imaging at magnifications of $>300{,}000\times$, we have to use the streaking in the image to correct for astigmatism. We'll talk about this when we discuss HRTEM in Chapter 28.

Intermediate-lens astigmatism is of secondary importance and only affects the DP. Because the DP is at zero magnification in the objective lens, the intermediate lenses are responsible for magnifying it. So if there is residual astigmatism in these lenses, then the DP will show orthogonal distortions as you take it through focus. This effect is small and can only be seen in the binoculars as you focus the DP with the diffraction

focus (intermediate lens) control. Make sure that the incident beam is strongly underfocused to give the sharpest spots. As with objective astigmatism in the image, simply adjust the intermediate stigmators to compensate for any spot distortion at underfocus, and overfocus until the spots expand and contract uniformly in all directions through focus. You should be aware that not all instruments have the requisite intermediate stigmators to carry out this correction.

9.6 CALIBRATING THE IMAGING SYSTEM

Your TEM should be calibrated when it is first installed and then periodically throughout its life, especially if you wish to carry out accurate measurements from images or DPs. If the instrument is modified substantially (e.g., a burned-out lens coil is replaced), then it must be recalibrated. In all cases you must specify a set of standard conditions under which the calibrations are carried out (e.g., objective lens current and other lens settings, eucentric height, etc.). You can find a full description of all the detailed concerns with all the TEM calibrations in Edington (1976). Remember, if you're serious about TEM don't rely on someone else's calibrations.

THE LAST USER

Since you usually will not be the first user, you should take the time to check the existing calibration. Don't assume it is correct.

9.6.A Magnification Calibration

We use standard specimens to calibrate the magnification. The most common specimen we use is a thin carbon-film replica of an optical-diffraction grating of known spacing, as shown in Figure 9.22A. The typical linear density of lines in the replica is 2160/mm (giving a line spacing of 0.463 μm with an error that will depend on how many grating spacings you can measure). Grating replicas enable calibration up to magnifications of ~200,000×. Above this magnification, individual grating spacings are wider than the recording film. So we can then use small latex spheres (50–100 nm diameter) although they are susceptible to beam damage and shrinkage under electron bombardment. At the highest magnifications, phase-contrast images of known crystal spacings, such as the 0002 spacing in the graphite structure (0.344 nm) can be used or the 111 fringes of Si. What we are doing is just using a known periodicity in the crystal; careful consideration of the objective-lens defocus and specimen thickness is required before the phase-contrast lattice image can be directly interpreted, as we discuss in detail in Chapters 27 and 28.

MAGNIFICATION

Magnification calibration is so sensitive to so many variables that some users deposit a standard material on the material they are studying so that the calibration will be done under exactly the same conditions and will appear in the same image as your area of interest.

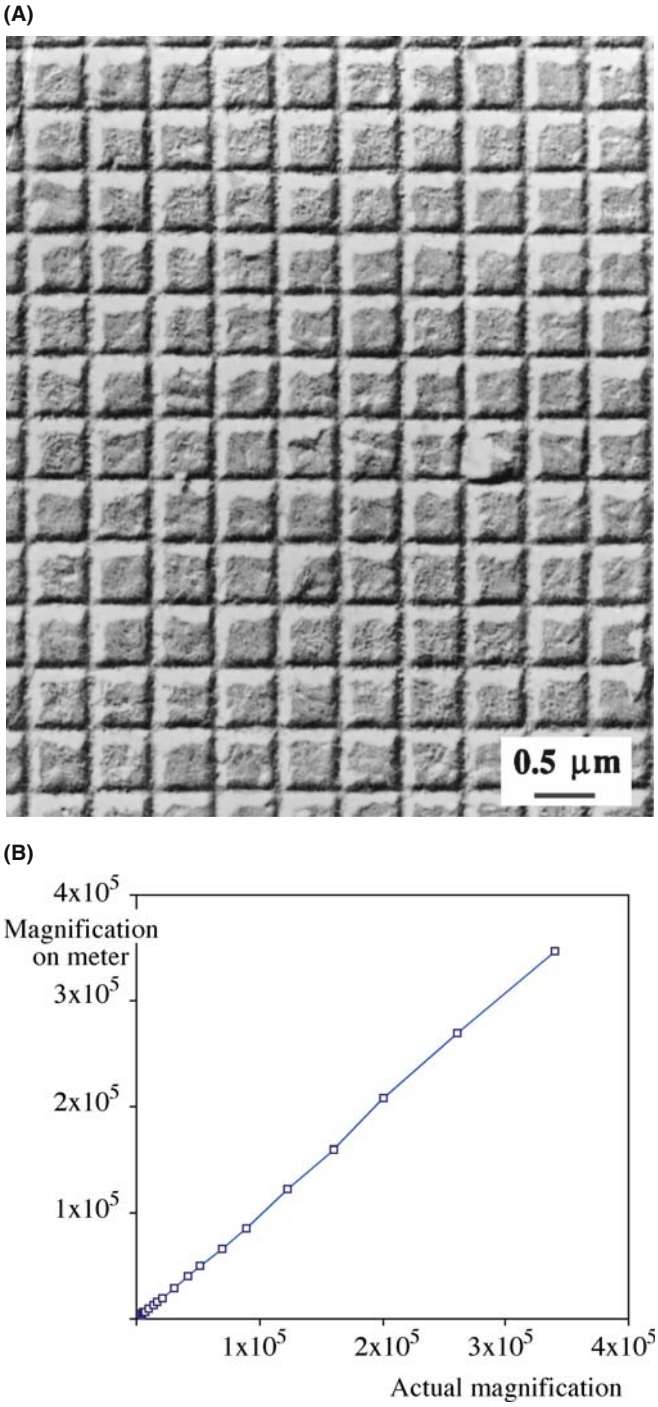


FIGURE 9.22. (A) An image of a diffraction-grating replica for which the actual spacing of the grating is known. (B) The TEM magnification can thus be calibrated, permitting specific magnification settings (usually with approximate/calculated magnifications) to be assigned more accurate, experimental magnifications. If the plot for your TEM is not linear then there are serious problems either with your measurements or with the microscope.

Operational Procedure #9

- Insert the calibration specimen and ensure the holder is at the eucentric height.
- Ensure the illumination is parallel and the objective aperture is well centered on the 000 spot in the SADP.
- Switch back to image mode and then focus the BF image (see below for the best way to be consistent in focusing).
- Record images of the diffraction grating at all magnification settings, taking care to re-check the focus after each change of magnification.
- Calculate the magnifications experimentally from the images knowing the true spacing.

Figure 9.22B shows the classical magnification calibration for a Philips CM30 TEM. Today, you'll use a crystal lattice at ~300 k×, relate it to the same area at say 50 k×, repeat at 10 k×, and then at all magnifications in between.

HYSTERESIS

Electromagnetic lenses have hysteresis. Everything you do must be reproducible if you want your results to be.

You have to calibrate the magnification because the TEM imaging system does not give stable and reproducible lens strengths. The lens strengths will change with ambient temperature (e.g., how many people are in the room), with the efficiency of the cooling system of the lenses and with lens hysteresis. Therefore, if you want to make accurate measurements from your TEM images, you must carry out the magnification calibration *at the time you make the measurements*. In particular, you have to minimize the lens hysteresis by always approaching image focus consistently from overfocus or underfocus and/or reversing the lens polarity several times before finally coming to focus. Also, you must remember that there may be barrel or pincushion distortions in the image particularly at very low magnification (< 5000×) where the lens is not designed to work well. (As students you ought to be able to work out what the first distortion does to your image, but you might not be familiar with the latter distortion at this early stage of your life!)

THE MAGNIFICATION

The image size will change during recording, printing and publishing. *Never* state an image magnification in the caption. Indicate a distance on the recorded image that corresponds to a distance on the actual specimen at the time the image was recorded.

Remember that the area of the specimen you are working on must be at exactly the right 'height' in the column. Because of the magnification error, the TEM is not the best instrument for absolute measurement of particle sizes, etc. However, relative measurement is easily done with reasonable accuracy (±5%), so long as you note the precautions we've just described. Without a calibration, the digital readout is probably no better than ±10% accurate, and so it is unwise to state magnifications to better than ±10%. You should be suspicious of any micrographs that you see in the literature with a magnification that is more precise than three (or even two) significant figures such as 52,550×. It may indicate that the microscopist does not understand the limitations of their instrument and the work should be interpreted with due caution.

We used to call these lines 'micron bars' for obvious reasons but TEM magnifications have become larger and 'nano bars' would often be more appropriate (even though neither one is related to barrel distortion). All modern TEMs automatically put such a marker on the negative. However, its accuracy is only as good as your operational skills.

You can use an identical procedure to calibrate the STEM-display magnification despite the fact that the digital STEM-image magnification is, in principle, easily calculated from the scan-coil strengths. The image magnification differs from the digital readout because of variations in the objective lens. Table 9.1 shows the difference between a typical digital STEM magnification and the experimentally determined magnification using a diffraction-grating replica.

9.6.B Camera-Length Calibration

As mentioned already, we describe the magnification of the DP by the camera length (L), a term that arises from X-ray projection-diffraction cameras which operate without lenses (because focusing X-rays is very difficult). In such cameras, magnification is increased by moving the recording film further away from the specimen, in exactly the same way that you magnify your computer-slide presentations onto a screen by moving the projector further away. (You can always tell non-microscopists who try to *focus* their slides in the same way, or those

TABLE 9.1. Magnification Calibration for a Philips EM400T in STEM Mode at 120 kV

Digital readout	Calculated magnification
3,200	3,420
6,400	6,850
12,500	12,960
25,000	27,000
50,000	54,000
100,000	108,000

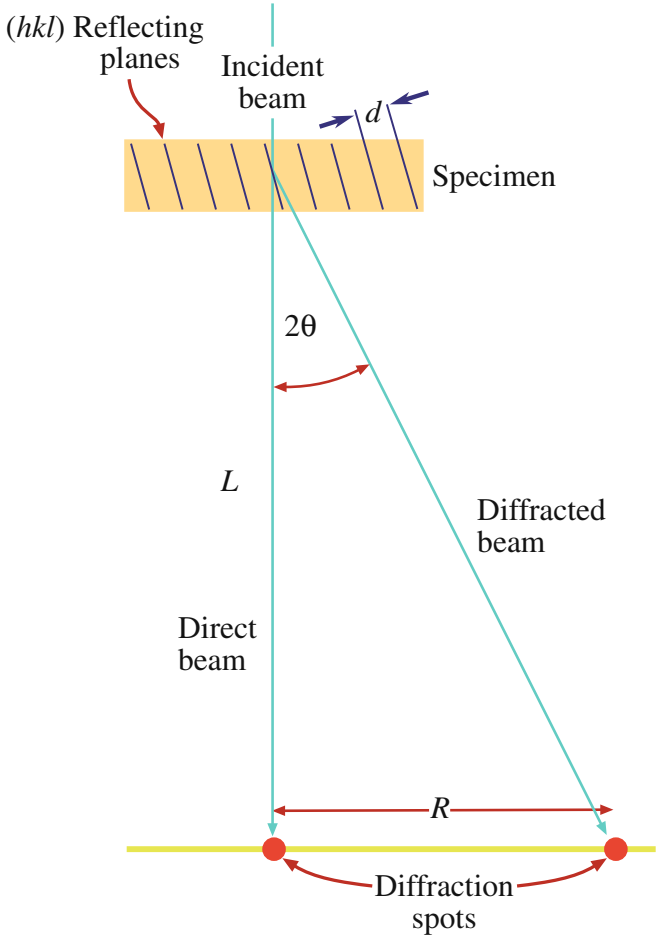


FIGURE 9.23. The spacing R between the direct beam and a scattering maximum (such as a diffracted beam or the radius of a ring of diffracted intensity) is related to the camera length, L. Increased magnification corresponds to effectively increasing L, although in practice, this is accomplished with lenses.

who try to change the magnification by adjusting the projector lens focus!)

The camera-length concept can be applied in the TEM as shown in Figure 9.23. This figure represents the imaging system, but without the lenses drawn in.

CAMERA LENGTH

If we increase the magnification of the lenses between the specimen and the viewing screen, we increase the effective distance L between the specimen and the screen.

The camera length is a calculated value rather than a physical distance. If electrons are scattered through an angle 2θ at the specimen (as in a typical diffraction event described in Chapters 2 and 11), then the distance between the direct and diffracted beams as measured on the screen (R) is determined by L, since

$$\frac{R}{L} = \tan 2\theta \sim \theta \qquad (9.1)$$

From the Bragg equation (equation 3.21) we know that $\lambda/d = 2\sin\theta \sim 2\theta$ and so we can write

$$Rd = \lambda L \qquad (9.2)$$

Operational Procedure #10 To calibrate the magnification of the DP we need to record DPs from a specimen with a known crystal spacing (d) such as a thin film of polycrystalline Au or Al which gives a ring pattern (see Figure 2.11).

- Make sure your specimen is at the eucentric height and the BF image is in focus.
- Insert the SAD aperture and switch to diffraction mode.
- Ensure the beam is parallel so that the DP spots or rings can be focused to the smallest size. If the pattern is not focused then focus with the first intermediate lens (diffraction focus) control. Be careful if you have a c/o lens because changing the illumination conditions (e.g., trying to make the beam more parallel) will change the DP focus. If you then sharpen the DP by changing the diffraction focus, the camera length will change. If in doubt, consult the manufacturer's handbook.
- Since we know the lattice parameter of the metal, we can measure the ring radius R on the photographic film or computer display for any plane that is diffracting (see Chapter 18 to find out exactly how we do this).
- Since we know λ we can easily determine L from equation 9.2.

A typical TEM camera-length calibration is shown in Table 9.2. The STEM camera-length calibration may be different to the TEM if the objective lens setting is not exactly the same in TEM and STEM modes, and this depends on the vintage and make of your instrument. So you should check with the manufacturer before taking the time to perform the calibration.

TABLE 9.2. Comparison of Experimentally Measured Camera Length (and Camera Constant) with the Digital Readout for a Philips EM400T Operating at 120 kV (λ = 0.0335 Å)

Camera length setting	Digital readout (mm)	Measured camera length, L (mm)	Camera constant λL (mm Å)
1	150	270	9.04
2	210	283	9.47
3	290	365	12.22
4	400	482	16.14
5	575	546	18.28
6	800	779	26.08
7	1150	1084	36.29
8	1600	1530	51.22
9	2300	2180	72.99
10	3200	3411	114.20

It's worth explaining in more detail the effect of a c/o lens on the SAD calibration. In a c/o lens, as shown back in Figure 9.5, the beam can go from divergent to convergent conditions at the specimen plane. However, unless the beam is parallel, your DP focus will change with illumination angle. In this case, the diffraction-lens setting at the point of focus of your DP will differ for different illumination conditions. So L will change (by as much as 15%) depending on the angle of convergence/divergence. In order to ensure the same L, you must pick one condition and use it for every calibrated DP. There are several ways to do this.

- Set the diffraction lens to the same excitation for each pattern and focus the DP with the C2 lens.
- Focus the pattern at the objective aperture (especially if the BFP of the objective and the objective aperture share the same position).
- Go to the maximum value of the C2 lens, focus the DP, and use the C1 lens to adjust the brightness of the DP.

There are other possibilities which you'll be able to work out as you get more skilled but, above all, be consistent and always *use the conditions you chose when you did the calibration*.

9.6.C Rotation of the Image Relative to the DP

Anyone studying crystalline materials must determine the angle between directions in the image and directions in the DP. At a fixed camera length, the DP always appears on the screen in a fixed orientation. But if you record images at different magnifications, the images will rotate by an angle ϕ with respect to the fixed DP. (In some TEMs this rotation has been removed by the addition of a compensating projector lens and, in this case, there is always a fixed rotation (ideally 0°) between directions common to both image and DP.)

THE GOLDEN RULE

Always do the calibrations yourself. Do not rely on factory calibrations. The conditions you use in your laboratory may differ from those chosen by the manufacturer.

Operational Procedure #11 To determine this rotation, we often use a specimen of α-MoO_3, because it forms thin, asymmetric crystals with a long edge known to be parallel to the 001 direction in the crystal.

- Take care to ensure that, as usual, the image is focused with the specimen at the eucentric plane.
- Insert the SAD aperture and ensure that it is focused (using the intermediate lenses) to coincide with the image plane.
- Switch to diffraction mode with the beam parallel and adjust the diffraction focus to give sharp diffraction maxima.
- Take a double exposure of the DP and the image as shown in Figure 9.24A.
- Repeat the whole exercise for different magnifications and plot out the variation of the angle ϕ as shown in Figure 9.24B.
- You can do the same, if necessary, for different values of L, which introduce a systematic change in ϕ. For this reason, we recommend that you carry out all your SAD work at a standard value of L; 500–1000 mm is usually optimum.

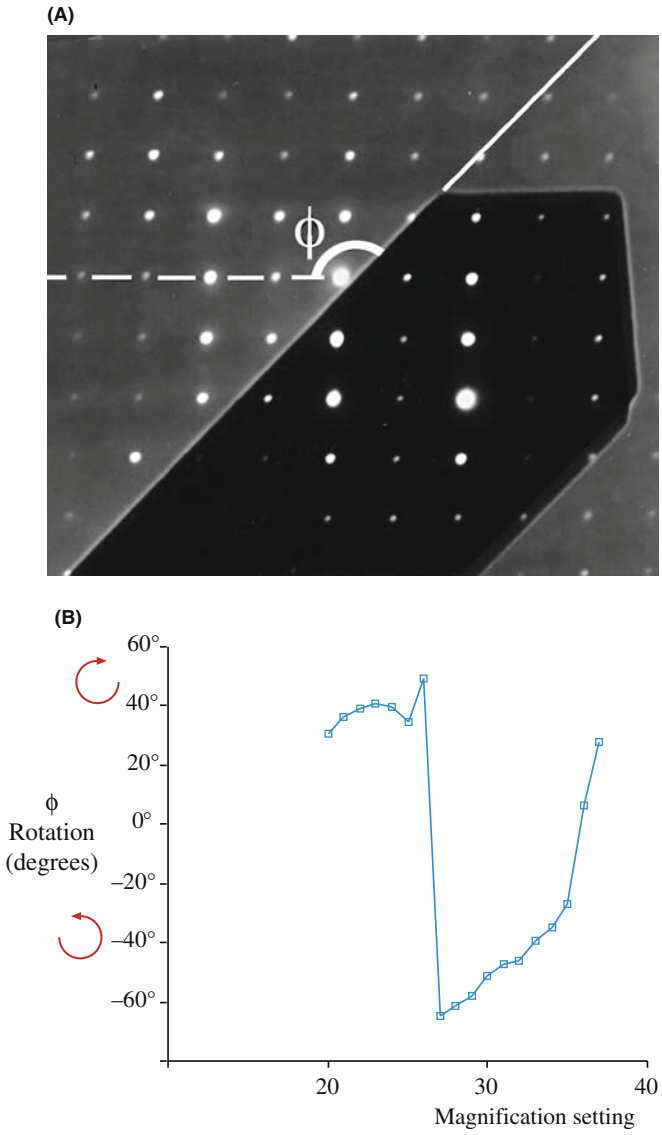


FIGURE 9.24. (A) A double exposure showing the superposition of an image of an α-MoO_3 crystal on a DP from the same crystal, defining the rotation angle ϕ. (B) The rotation calibration gives the angle ϕ between equivalent directions in the image and the DP as the magnification is varied. The calibration assumes a constant camera length. The rotation angle will change significantly if the imaging system switches lenses on or off as the magnification is changed (e.g., between settings 26 and 27).

A further complicating factor is that, as the image magnification is increased, the TEM lens control logic may switch off, or switch on, one of the imaging-system lenses. When this happens, a 180° inversion is introduced into the image. You can see this happen if you watch the image carefully as you change the magnification. This inversion has to be included in the rotation calibration, otherwise a 180° error will be made in the assignment of directions in the image. One way to see if the image has a 180° inversion is to look at the DP and defocus it slightly so the BF image in the direct beam can be seen directly at very low magnification. The 180° inversion is immediately obvious, as shown in Figure 9.25. As we already noted, some manufacturers adjust the imaging-system lenses to compensate for the rotation, thus keeping a constant image-DP rotation angle at all magnifications. Likewise, in STEM mode, since the imaging lenses are not used for magnification, there is a fixed angle between the images and DPs.

McCaffrey and Baribeau used molecular-beam epitaxy to create a cross section TEM specimen made from a single-crystal Si/Ge wafer specimen that can perform these three major calibrations (magnification, camera-length, and image/DP rotation). Similar specimens are available from suppliers of TEM consumables (see URLs 1–3, for example). These suppliers also offer a range of standard specimens for other TEM needs (e.g., XEDS and EELS analysis). If you are serious about your microscopy, buy your own standards, keep them safe and clean, and use them often.

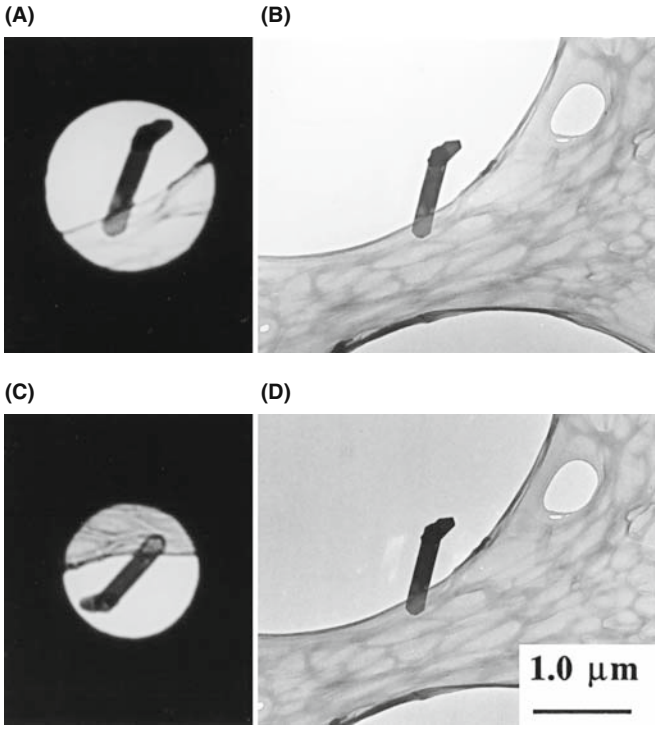


FIGURE 9.25. Defocused direct beam in a DP from α-MoO_3 compared with a BF image, showing how to determine if a 180° inversion exists or not. If the image of the specimen inside the expanded image of the beam is rotated with respect to the image on the screen, as in (C) and (D), then a 180° inversion is required to determine the correct angle between directions in the DP and directions in the image. In (A) and (B), no rotation occurs between the DP and BF image.

9.6.D Spatial Relationship Between Images and DPs

If you don't use a double exposure when comparing images and DPs (or indeed when comparing directions in any two films) you need a fixed reference line. This line must be independent of slight variations that may arise depending on the film size, how you loaded it, etc. The best reference line is the edge of the plate-numbering system that is superimposed on each film.

Yesterday's advice:
Whenever you're comparing images and DPs, it is essential to compare the photographic negatives with the *emulsion side up*. This is contrary to usual photographic practice, but it's necessary to preserve the relationship between manipulations of your specimen and what you see happening on the screen. *If you don't do this,* it is easy to introduce a 180° error into the relationships between images and DPs.
Today's advice:
Don't use film. If you use a CCD, you'll never think of inverting the image!

9.7 OTHER CALIBRATIONS

The accelerating voltage: The selected voltage may differ from the absolute voltage by detectable amounts. There are several ways to determine the actual voltage: First, you can measure the electron wavelength, λ, by measuring the angle, ϕ, between two Kikuchi line pairs (see Chapter 19) that intersect a distance R from the direct beam

$$\tan\phi = \frac{R}{L} = \frac{\lambda}{d} \tag{9.3}$$

Alternatively, you can match simulations of HOLZ lines to experimental lines in CBED patterns (see Chapter 21) and determine which λ gives the best match. Finally, in principle, you could get your XEDS computer system to display the X-ray spectrum (Chapter 34) out to E_0, the beam energy, and the bremsstrahlung intensity vanishes to zero at the exact beam energy (this is called the Duane-Hunt limit). However, XEDS computer displays rarely show energies > 30–40 keV.

The specimen tilt axis and the sense of tilt: In a side-entry stage, the principal tilt axis is parallel to the specimen-holder rod. Since the image is usually rotated on the screen relative to the specimen, how can you locate this tilt-axis direction? Well it's easy if you can move the specimen in a known direction. From this movement, you can determine the direction of the tilt axis for a specimen of known geometry.

Yesterday's advice:
If you gently push on the end of your side-entry specimen holder, the image moves parallel to the principal tilt axis. But don't pull the holder in the other direction since you may inadvertently cause air to leak into the airlock.
Today's advice:
We're not letting you remove the draft-excluder (which covers the end of your specimen holder) to touch the holder after it has equilibrated. Monitor how you load the specimen instead which takes care and practice—but so do most TEM operations.

If you are looking at the DP, defocus the pattern so you can see the BF image in the direct-beam spot, as in Figure 9.26, then carry out the same exercise. If you are using a top-entry holder you will need to calibrate this tilt using a known specimen geometry, which is more challenging.

Focal increments of the objective lens: If you're going to do high-resolution, phase-contrast imaging and you don't have a C_s corrector, then you need to know the value of each defocus step of the objective lens so you can correctly calculate and interpret the image contrast. There is a simple method for determining this step value. Superimpose a focused image and an image defocused by a known number of objective-lens focal increments (Δf). The two images will be separated by a distance Δx which is related to Δf by

$$\Delta f = \frac{\Delta x}{2\,Mm\theta} \tag{9.4}$$

where M is the magnification, m is the number of focal increments, and θ is the Bragg angle for the reflection used to form the image. If you insert some typical values, you'll find that it is difficult to be very accurate with this method. We'll return to this topic in Chapter 28.

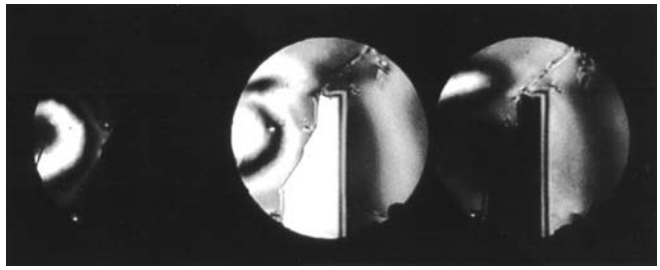


FIGURE 9.26. Defocused multiple-DF image showing how it is possible to determine simultaneously the direction of features in the image (e.g., the vertical twin boundary) and directions in the DP (e.g., the horizontal vector between the diffraction disks). If the specimen holder is moved in the direction of the principal tilt axis, the image will move and identify the relationship between that tilt axis and the crystallographic direction in the DP.

CHAPTER SUMMARY

We've now shown you how a TEM is put together. While the manufacturer does the best job possible, there are still some essential steps for you, the operator, to carry out. You must understand how to align the illumination system so the beam is on axis. You can then create a parallel beam for TEM and a convergent one for STEM. The C2 aperture is a crucial part of the whole illumination system and the most easily misaligned. Astigmatism is not too much of a problem if the instrument in general and the apertures in particular are kept clean. The objective lens/stage combination controls all the useful information that is created as the beam is scattered by your specimen. *Always* start a microscope session by fixing the eucentric height and, before you do any worthwhile imaging, align the objective center of rotation, and minimize any objective astigmatism at high magnification. Diffraction and STEM operations require a centered, focused DP.

If you want to make any quantitative measurements from your images and DPs (and you really ought to do this if you have any aspirations to be a real microscopist), then calibration cannot be avoided. Your images and DPs are relatively useless unless you know their magnification and camera length, respectively, and the angular relationship between the two. So take the time to do this early in your studies. In doing so you will not only ensure that you produce quality data, but you will also learn an enormous amount about how these complex instruments work.

You'll notice that some of the data (e.g., Figure 9.12B and Table 9.1) are for older TEMs. It would be better to use data from even older machines ("Why?" you should ask) but you are unlikely to see such machines in operation.

SOME HISTORY

SAD was invented by LePoole in 1947. Some early machines used a pair of blades with a 90° bend that slid past one another to give a square aperture! See, e.g., LePoole, JB 1947 *A New Electron Microscope with Continuously Variable Magnification* Philips Tech. Rept. **9**(2) 33–45. You can find further discussion of TEM calibration in Edington's text.

OPERATING THE TEM

Chapman, SK 1980 *Understanding and Optimizing Electron Microscope Performance 1. Transmission Microscopy* Science Reviews Ltd. London. An old book but the principles are so clearly outlined that, if you are a serious user, this must be in your laboratory.

Chapman, SK 1986 *Maintaining and Monitoring the TEM* Royal Microscopical Society Handbook No. 8 Oxford University Press New York. If you can't find this on the shelf in the office of the EM technical staff, be sure to buy it for them for Christmas. Your small investment will be repaid many times over.

Chescoe, D and Goodhew, PJ 1990 *The Operation of Transmission and Scanning Microscopes* Royal Microscopical Society Handbook No. 20 Oxford University Press New York. A great in-depth description of the operational principles described briefly in this chapter.

Christenson, KK and Eades, JA 1988 *Skew Thoughts on Parallelism*, Ultramicroscopy **26** 113–132.

Edington, JW 1976 *Practical Electron Microscopy in Materials Science* Van Nostrand Reinhold New York. Now very dated, but still a valuable text for the details of hands-on, diffraction-contrast imaging and SAD; it's especially good on calibrations. However, this text was written before c/o lenses were introduced which, as we've noted, introduce complications into hitherto simple operations like forming parallel beams.

Goldstein, JI, Newbury, DE, Joy, DC, Lyman, CE, Echlin, P, Lifshin, E, Sawyer, LC and Michael, JR 2003 *Scanning Electron Microscopy and X-ray Microanalysis* 3rd Ed, Kluwer New York. Dealing with a virtual C2 aperture, scanning imaging and much more.

Keyse, RJ, Garratt-Reed, AJ, Goodhew, PJ and Lorimer, GW 1997 *Introduction to Scanning Transmission Electron Microscopy* Royal Microscopical Society Handbook No. 39 Bios Scientific Publishers Oxford. Basic introductory level but the only book available that emphasizes the operation of dedicated STEMs which historically were made by one company (VG) but are now available through at least three manufacturers (Hitachi, JEOL, and Nion). The DSTEM may well come to dominate the semiconductor industry in the way the SEMs did in the 1980s. If so, these instruments will become fixtures in any self-respecting EM Center. If, by now, you are getting the idea that the RMS handbook series is useful then you are right (see http://www.rms.org.uk/other-publications.shtml).

McCaffrey, JP and Baribeau, JM 1995 *A Transmission Electron Microscope (TEM) Calibration Standard Sample for All Magnification, Camera Constant, and Image/Diffraction Pattern Rotation Calibrations* Microsc. Res. Tech. **32** 449–454.

Watt, IM, 1997 *The Principles and Practice of Electron Microscopy* 2nd Ed. Cambridge University Press New York. Basic, hands-on, easy-to-read book, full of instructive images and diagrams, and easily understood explanations.

URLs

1) http://www.emsdiasum.com/microscopy/products/catalog.aspx
2) http://www.tedpella.com/calibrat_html/TEM6.htm
3) http://www.2spi.com/catalog/stand.html

COMPANION TEXT

C_c correction will be discussed in the chapters on HRTEM and EELS. The C_c monochromator takes on several forms but each decreases the total beam current.

SELF-ASSESSMENT QUESTIONS

Q9.1 Name the three principal components of the TEM.
Q9.2 How is the beam tilted or translated?
Q9.3 How do you form a parallel beam; why would you want to do this and why is it not exactly parallel?
Q9.4 How do you form a convergent beam; why would you want to do this and why is it sometimes divergent?
Q9.5 How do you align the gun and why should you bother?
Q9.6 How can you tell when your C2 aperture is misaligned and how would you correct it?
Q9.7 How does the probe size change with the C1 setting and why might you want to change the size?

Q9.8 What is SAD and how is it different from CBED?
Q9.9 What is the difference between a BF image and a DF image?
Q9.10 What is the advantage of forming images using STEM rather then TEM?
Q9.11 What causes lens astigmatism? How do you correct for it (a) in the illumination and (b) in the imaging system?
Q9.12 Why does the objective lens rotation need to be calibrated?
Q9.13 Why would you want to form hollow-cone DPs and images?
Q9.14 Explain CDF, DADF, HCDF, SADP, SDP, HCDP, FFP, BFP? If you can't find all these acronyms/initials in the text, suggest what they mean.
Q9.15 What is the most important aperture in the TEM and why?
Q9.16 How do you change the image magnification in a TEM?
Q9.17 How do you change the image magnification in STEM? Explain why this method confers advantages over the answer to the previous question.
Q9.18 Why do we need to do the magnification calibration?
Q9.19 What are the reasons you might need to translate the beam?
Q9.20 How should you always start a microscope session (apart from quiet prayer)?
Q9.21 How should you always finish a microscope session (apart from a loud prayer)?

TEXT-SPECIFIC QUESTIONS

T9.1 To converge the beam, which lenses do you use and how does this differ depending on the electron source? (Hint: go back to Chapter 5 and also look at Figures 9.3–9.5.)
T9.2 Explain the differences between the various methods used to create a parallel beam at the specimen plane in Figures 9.1 and 9.2.
T9.3 A smaller probe has a smaller current in it. Use Figure 9.4 to explain why this is so and state some consequences of reducing the probe current.
T9.4 Why use SAD, instead of just forming DPs without a selecting aperture?
T9.5 What illumination-system operation is essential to form Fresnel fringes for astigmatism correction?
T9.6 What is the most common problem with TEM illumination systems (apart from the previous user of the microscope)?
T9.7 How can you tell when your specimen is in the eucentric plane and why do you want it to be there?
T9.8 Typically, how accurate is the magnification of images of particles in TEM? What do you have to do to get better accuracy?
T9.9 List the operational steps necessary to see a SADP if you have a BF image on the TEM screen.
T9.10 List the operational steps necessary to see a CDF image if you have a DP on the TEM screen.
T9.11 Explain the principle shown in Figure 9.13 of using a virtual rather than a real aperture in the specimen plane for selected-area diffraction.
T9.12 Use Figures 9.16 and 9.19 to explain why we can't change the direction of the incident beam in STEM imaging.
T9.13 Why might we say that STEM images don't suffer from the defects in the imaging lenses and would we be correct?
T9.14 Is Figure 9.21 taken with a thermionic or FEG source? Explain your answer.
T9.15 Discuss the pros and cons of a c/o lens system using Figure 9.5 as a basis for your discussion. Justify why every materials-research TEM must have a c/o system.
T9.16 Distinguish traversing and tilting a beam and, using specific figures in the text, explain how these operations can be used for DADF operation, hollow-cone imaging and diffraction, stationary-DP formation, and scanning-image formation.
T9.17 When might you want to use a wobbler? Sketch a ray diagram explaining how a wobbler works. Sketch an image of how an astigmatic beam changes as it is wobbled and how a misaligned beam changes as it is wobbled.
T9.18 Use the references to find an equation governing the relationship between probe current and probe size.
T9.19 Estimate the relationship between the final aperture radius in the illumination system and the convergence angle of the beam and explain how this is relevant to Figure 9.11.
T9.20 Why does converging the beam expand the spots in the DP and why is there different contrast information (images) in different expanded spots, as in Figure 9.26?
T9.21 Why is it important to calibrate the rotation of your DP with respect to your image?
T9.22 Estimate the accuracy of the magnification and rotation calibrations.
T9.23 Look up the lattice parameters of α-MoO_3, make some reasonable assumptions about the TEM, and calculate an approximate camera length for the DP in Figure 9.24.

10

Specimen Preparation

CHAPTER PREVIEW

Specimen preparation is a very broad subject; there are books devoted to this topic alone. The intention here is to summarize the techniques, suggest routes that you might follow, and above all to emphasize that there are many ways to produce a TEM specimen; the one you choose will depend on the information you need, time constraints, availability of equipment, your skill, and the material. So we'll concentrate on the 'principles of cooking,' but won't try to list all the possible 'recipes.' One important point to bear in mind is that your technique must not affect what you see or measure, or if it does, then you must know how. Specimen preparation artifacts may be interesting but they are not usually what you want to study. Incidentally, we'll make 'specimens' from the 'sample' we're investigating so we'll look at 'TEM specimens,' but sometimes we, and everyone else, will interchange the two words.

The TEM specimen, when you've made it, must be electron transparent (usually) and representative of the material you want to study. In most cases (but not all) you would like your specimen to be uniformly thin, stable under the electron beam and in the laboratory environment, conducting, and non-magnetic (we'll discuss some exceptions as we proceed). Few specimens approach the ideal and usually you have to compromise. In general we can divide specimens into two groups: self-supporting specimens and specimens resting on a support grid or thin washer; the grid or washer is usually Cu but could be Au, Ni, Be, C, Pt, etc. Before discussing these two groups we will briefly review the most important part of specimen preparation, namely, safety. You may damage the microscope later, but this is the stage where you could do much worse to yourself and your colleagues.

It is often assumed that preparation of the TEM specimen will take several hours. Actually this time could be as short as 5 minutes or as long as 2 days even for the same material. For example, as you'll see, if you want to examine a piece of $YBa_2Cu_3O_{6+x}$, the high-temperature superconductor, you could crush the sample in a pestle and mortar using a nonaqueous solvent, catch the small particles on a carbon film, and put the specimen in the TEM; time required is about 10 minutes. Alternatively, you might cut the sample into thin slices using a diamond saw, cut 3-mm-diameter disks from the slice, thin the disk on a grinding wheel, dimple the thinned disk, then ion mill to electron transparency at liquid-nitrogen temperatures, carefully warm the specimen to room temperature in a dry environment, and put it in the TEM; time required is 1 or 2 days. Which method you choose would depend on what you want to learn about your material.

10.1 SAFETY

Either the specimen itself or the best method for preparing it for viewing in the TEM may require extreme care. Even materials which are safe and relatively inert in bulk form may be hazardous in powder form. Four favorite (because they work so well) liquids for polishing solutions are hydrogen cyanide, hydrofluoric acid, nitric acid, and perchloric acid. These liquids may be poisonous, corrosive (HF quickly penetrates the body and then dissolves the bone), or explosive (perchloric acid and nitric acid when mixed with certain organic solvents). It is clearly essential that you check with your laboratory manager, the reference texts, and the appropriate materials safety data sheets (MSDS) before you begin specimen preparation. This checking might also save a lot of time.

SAFETY FIRST

This whole section should, of course, be in a big red box. Some of the chemicals we use are really dangerous. We remember HF, perchloric acid, and HCN all being used concurrently in one small specimen-prep room.

In spite of these restrictions you may still need/want to use these acids and acid/solvent mixtures. The ion thinner may not be available or you may not be able to accept the damage that ions produce. In this event there are five brief points that you should bear in mind.

- Be sure that you can safely dispose of the waste product *before* you start.
- Be sure you have the 'antidote' at hand.
- Never work alone in the specimen-preparation laboratory. Always wear safety glasses when preparing specimens and/or full protective clothing, including face masks and gloves, if so advised by the safety manual.
- Only make up enough of the solution for the one polishing session. Never use a mouth pipette for measuring any component of the solution. Dispose of the solution after use.
- Always work in a fume hood when using chemicals. Check that the extraction rate of the hood is sufficient for the chemical used.

Since these four acids can be so dangerous, we'll mention them specifically, but remember—always seek advice before chemically preparing specimens.

Cyanide solutions: If possible, avoid this solution even though you may see it in the textbooks. The only metal where it really excels is gold and you can thin this by very careful ion milling.

Perchloric acid in ethanol or methanol: If you have to use this 'universal polish' you should be aware that many laboratories require that you use a special dedicated hood which can be completely washed down since crystallized perchloric acid is explosive. The phase diagram in Figure 10.1 for the perchloric-acetic (acid)-water system makes the message clear. If you have to use perchloric-acetic acid mixtures or indeed when using any perchloric-containing mixtures, keep the density below 1.48. If you are very careful, if you *always* add the acid to the solvent, and you make sure that the liquid *never* becomes warm, then perchloric acid solutions can be used to produce excellent TEM specimens of Al, stainless steel, and many other metals and alloys.

Nitric acid: In combination with ethanol, this acid can produce explosive mixtures, especially if left for long periods of time and exposed to sunlight. It is preferable to use methanol rather than ethanol, but in either case, keep the mixture cool and dispose of it properly.

HF: This acid is widely used in the semiconductor industry and in 'frosting' light bulbs; the reason in both cases is that it dissolves SiO_2 leaving no residue. Careful use of dilute solutions can produce specimens that have large thin areas. Remember: if you use HF, completely cover any exposed skin; HF rapidly penetrates the flesh and dissolves bone and you won't even feel it!

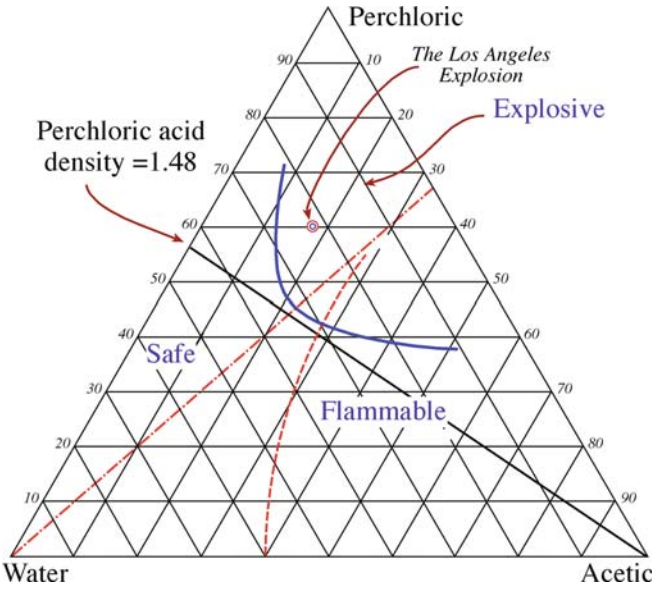


FIGURE 10.1. Perchloric-acetic-water phase diagram showing the hazardous regions and the recommended density line for safe use of all perchloric solutions. Always operate to the left of this line.

10.2 SELF-SUPPORTING DISK OR USE A GRID?

The type of TEM specimen you prepare depends on what you are looking for so you need to think about the experiment that you are going to do *before* you start thinning. For example, is mechanical damage to be avoided at all costs, or can it be tolerated so long as chemical changes don't occur—or vice versa? Is the specimen at all susceptible to heat or radiation? Depending on the answers to these questions, some of the following methods will be inappropriate. A flow diagram summarizing the different preparation philosophies is shown in Figure 10.2.

NANOMATERIALS

Think about your choice of material for the supporting grid.

A self-supporting specimen is one where the whole specimen consists of one material (which may be a composite). Other specimens are supported on a grid or on a Cu washer with a single slot. Several grids are

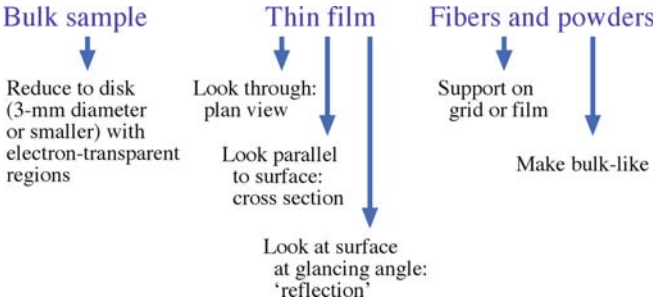


FIGURE 10.2. Flow chart summarizing the different sample geometries you may encounter.

shown in Figure 10.3. Usually the specimen or grid will be 3 mm in diameter.

Both approaches have advantages and disadvantages. Both offer you a convenient way of handling the thin specimen, since either the edge of the self-supporting disk or the grid will be thick enough to pick up with tweezers. If possible, never touch your specimen when it is thin. We recommend vacuum tweezers, but you'll need to practice using them; you can quite easily vibrate the specimen and break the thin area. You can get round this by using mouth-vacuum tweezers but see the section on safety first. Mechanical stability is always crucial. For example, single crystals of GaAs or NiO break very easily, so it is usually an advantage to have your specimen mounted on a grid since you then 'handle' the grid. However, if you are performing X-ray analysis on a specimen the grid may contribute to the signal, because the X-rays can also arise from the grid. Thus you see a Cu peak in the X-ray spectrum where no Cu is present in the specimen. We'll talk in Chapter 33 about how to minimize this artifact. Of course, the self-supporting specimen essentially has the same problem—it's just not as obvious! In fact, the preferred geometry for such analysis is usually the one where the specimen is thinnest.

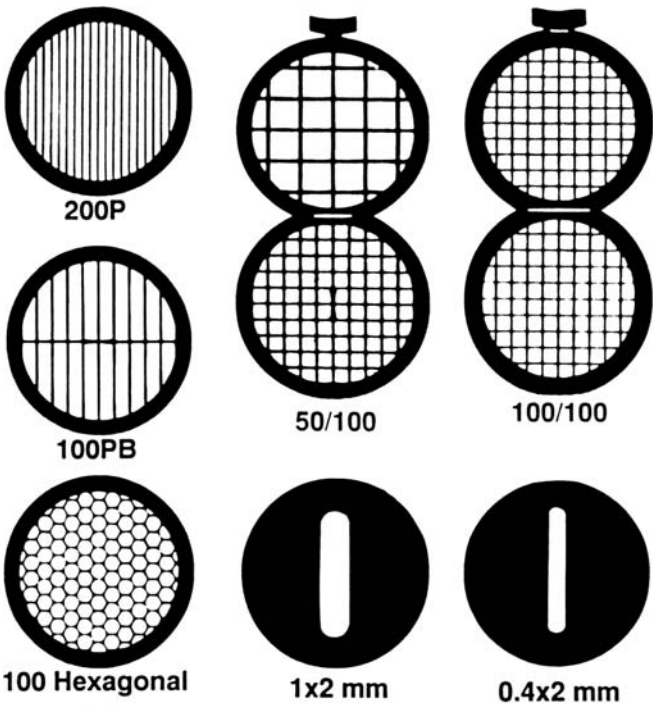


FIGURE 10.3. A variety of specimen support grids of different mesh size and shape. At top right is the oyster grid, useful for sandwiching small slivers of thin material.

THE DIAMETER OF TEM SPECIMENS

Why are the TEM specimens 3.05 mm in diameter? Because the manufacturers say so. Must this always be the case? Only if you need to double-tilt or tilt-rotate.

Why 3-mm disks? The disk diameter is usually a nominal 3.05 mm. We thus refer to the specimen as a 3-mm disk. Occasionally you will encounter a microscope which uses a 2.3-mm disk. The smaller diameter was used in earlier microscopes and has two important advantages, which are not fully exploited by modern machines. Ideally, the region of the specimen which you want to study will be located at the center of your disk, no matter how large the disk is. As we saw in Chapter 9, the reason is that, as you tilt the specimen in the microscope, the region of interest will then stay at the same position (height) above the objective lens and on the optic axis. Since, for a self-supporting disk, the rim of the specimen must be relatively thick and the total area of the material you'll study is small and confined to the center of the disk, you can make more 2.3-mm specimens from a given volume of material. This may be very important if the specimen is particularly special (expensive, rare) or if specimens break easily. A sample which is 5 mm × 5 mm will give one 3-mm disk or four 2.3-mm disks. The second advantage of such specimens relates to tilting; the smaller specimen holder can be manufactured to allow a greater tilt angle. Don't forget that if you need only one axis of tilt you may find the bulk holder useful. Then you can use a specimen which may be up to 10 mm long and 3 mm wide.

10.3 PREPARING A SELF-SUPPORTING DISK FOR FINAL THINNING

Preparation for final thinning involves three parts

- Initial thinning to make a slice of material between 100 and 200 μm thick.
- Cut the 3-mm disk from the slice.
- Prethin the central region from one or both faces of the disk to a few micrometers.

The method you use will depend on what you want to study and the physical characteristics of the material (whether it is soft or hard, ductile or brittle, delicate or robust, single phase or a composite, etc.).

10.3.A Forming a Thin Slice from the Bulk Sample

The materials you may need to thin can vary enormously. Clearly, we have to treat ductile and brittle materials differently.

(a) *Ductile materials such as metals*. Usually you don't want to introduce mechanical damage. For example, you may want to study the defect structure or the density of defects in processed materials. The ideal method is to use a chemical wire/string saw, a wafering saw (not diamond—the soft metal will dull the blade), or spark erosion (electro-discharge machining) to get a thin slice < 200 μm. (A string saw works by passing the string through an acid or solvent and then across the sample until the string 'cuts' through the sample; for example, you can use dilute acid to cut copper.) You could also roll the material to, very thin sheet, then anneal it to remove the defects introduced by rolling but that's a different material-processing route.

(b) *Brittle materials such as ceramics*. Here there are two cases: (i) where you must not introduce mechanical damage, (ii) where you don't mind introducing mechanical damage or the material won't damage. You have several options depending on the material. Some materials (Si, GaAs, NaCl, MgO) can be cleaved with a razor blade; these are materials with a well-defined cleavage plane and it is possible to carry out repeated cleavage to electron transparency (see Section 10.6.E). The ultramicrotome (see Section 10.6.B) allows you to cut very thin slices for immediate examination. If you don't want to cleave the specimen or you want to prepare a specimen parallel to a plane that doesn't cleave, you will need to use a diamond wafering saw. There are special techniques for some materials: you can, for example, use water as the solvent on a string saw to cut rock salt. One of the main limitations with sawing is that the process destroys some of your sample.

10.3.B Cutting the Disk

The same constraints hold for the coring process as for cutting slices: if the material is reasonably ductile and mechanical damage is not crucial, then the disks can be cut using a mechanical punch. A well-designed punch can cut disks with only minimal damage around the perimeter, but the shock can induce shear transformations in some materials. For more brittle materials the three principal methods are spark erosion (also essential when you need to avoid damage in a metal), ultrasonic drilling, and using a grinding drill. In each case the cutting tool is a hollow tube with an inner diameter of 3 mm. Again, you want the wall of the tube to be thin to minimize the amount of material that is wasted. Spark erosion is used for conducting samples and introduces the least amount of mechanical damage. The choice between an ultrasonic drill (vibrating in H_2O) and a grinding (or slurry) drill is often a matter of personal preference or availability. Both remove material mechanically and are widely used for ceramics and semiconductors. The drill may leave small particles in the specimen, and *all* mechanical thinning methods leave some surface damage. As a rule of thumb, abrasives produce damage to 3× their grit size. So a 1-μm abrasive will cause damage to 3 μm below the surface of each side of the specimen. Hence the final disk must be thicker than the 2× the damage depth or else mechanical damage will always be visible in the final specimen. Four different coring instruments from one manufacturer are shown in Figure 10.4.

CORING
Like extracting the core of an apple or a rock sample from the earth.

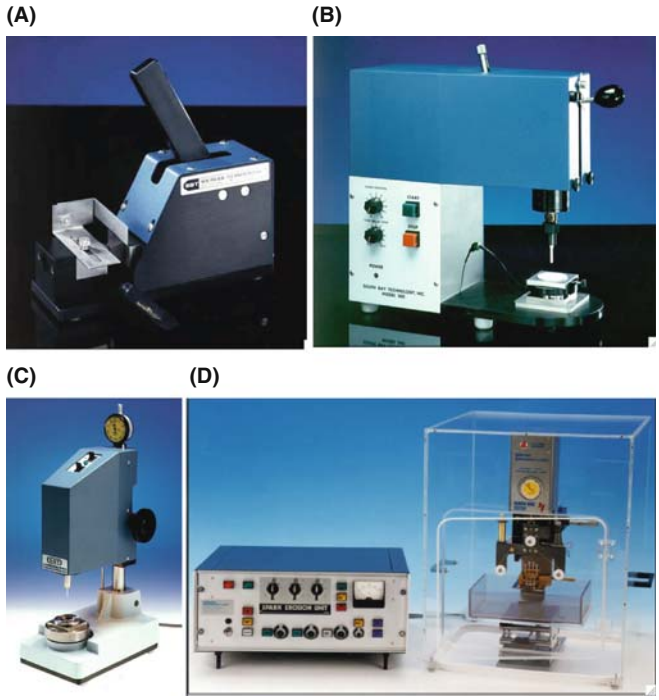


FIGURE 10.4. Four different coring tools from South Bay Technology. (A) A mechanical punch for stamping disks from thin sheets of ductile materials. A sheet sample is placed in the punch and the handle on the right is pushed down, ejecting a 3-mm-diameter disk suitable for thinning. (B) An abrasive-slurry disc cutter uses a rotary motion of the coring tube to drill round the disk. (C) An ultrasonic cutter. (D) A spark-erosion cutter; the erosion takes place under a solvent and behind a safety shield.

Note that there are variations for all these techniques: e.g., for Si, GaAs, and some other materials, you can glue the sample to a support, coat it with a protective layer, and cut circles through the film—then chemically etch the desired region. You need to experiment but the method should introduce no mechanical damage.

10.3.C Prethinning the Disk

The aim of this process is to thin the center of the disk while minimizing damage to the surface of the sample. In general we will refer to this stage as 'dimpling' no matter how the thinning is achieved. Any damage you create at this stage will have to be removed during the final thinning process (if you're interested in defects or if the damage changes your chemistry).

Most commercial mechanical dimplers use a small-radius tool to grind and polish the disk to a fixed radius of curvature in the center. Although the first instruments for dimpling were 'home built' the commercial models (see Figure 10.5) are now well developed. You can control the load, precisely determine the thickness of removed material (the depth of the dimple), quickly change the polishing tool, and interrupt the process to remove the sample for closer examination before continuing. The investment is well justified for materials laboratories. One alternative that has been used successfully is a (recycled) dentist's drill and some imagination. Typically dimpling can be carried out to produce regions ~10 μm thick although, in principle, precision dimpling with microprocessor control can sometimes produce electron-transparent specimens which are <1 μm thick.

For mechanical dimpling, as a general rule, the same guidelines apply as to all mechanical polishing; always gradually decrease the 'grit' size and conclude with the finest available, again ensuring that the final specimen thinness is >2× the damage depth of the smallest grit dimension. The better the polished surface, the better the final specimen. If both sides of a disk are dimpled the chances of final perforation occurring in the center are substantially increased, but in some cases you may wish to preserve one side of the specimen and thin from the other side only. One-sided dimpling is then essential prior to thinning to perforation.

Dimpling can also be performed chemically. Often in the case of Si this is achieved by allowing a jet of HF and HNO_3 to impinge (from below as shown in Figure 10.6) on the Si disk which has the edges lacquered to produce a supporting rim. The HNO_3 oxidizes the Si and the HF removes the SiO_2. Similar approaches use Br and methanol for thinning GaAs. This dimpling method uses dangerous chemicals, but it is very efficient. It can even be carried to final perforation with care.

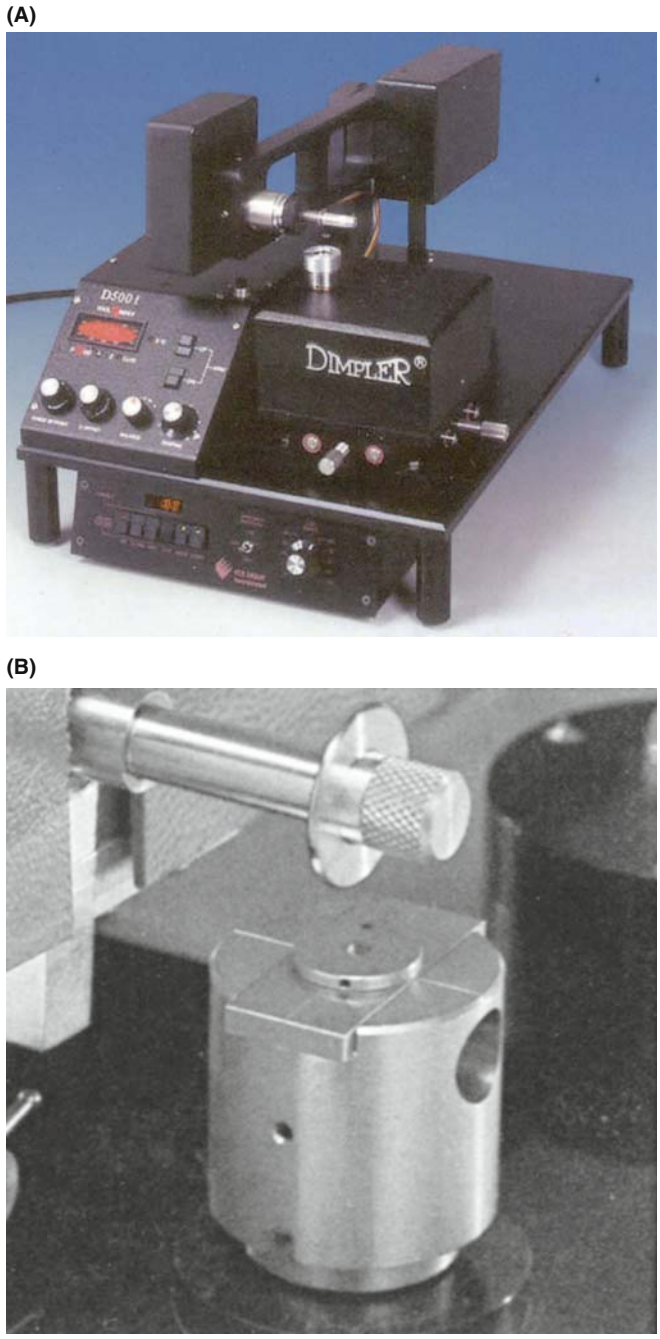


FIGURE 10.5. (A) Dimpling apparatus; (B) the grinding tool and specimen support block.

TEM specimen preparation has been revolutionized through the development of the tripod polisher; this tool can help you to thin your sample mechanically to less than 1 μm. You must consult the general references at the end of the chapter before using this tool. The tripod polisher, so called because it has three feet, is simply a device to hold your specimen while you mechanically thin it on a polishing wheel. You can purchase a tripod polisher commercially or build your own.

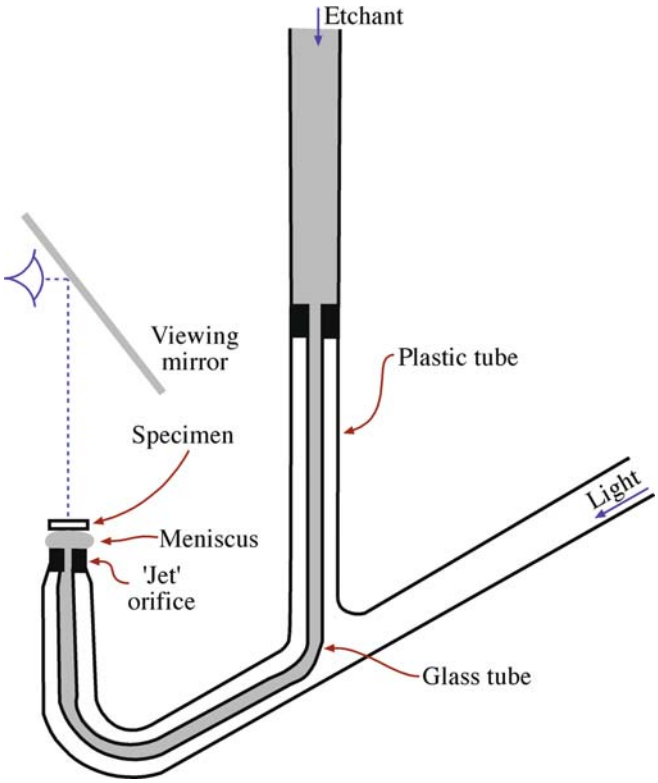


FIGURE 10.6. Surface dimpling using a chemical solution, e.g., to remove Si from one side of a disk. The light pipe permits visual detection of perforation using the mirror.

TRIPOD POLISHER

The tripod polisher has been used for many decades in the gem industry. It just looked different.

For some materials, such as Si, you can use this polisher to thin the specimen to electron transparency.

There are, however, several secrets in using the tripod polisher.

- You must use a very flat polishing wheel; the recommended approach is to use a glass platen. Take the greatest care in adjusting the micrometer to level the tripod.
- You need a supply of fine diamond lapping films; these are not inexpensive but it is a false economy to use them after they are worn. Always use a new sheet for polishing the second side of your sample since it is then particularly vulnerable.
- The diamond lapping films must not have an adhesive backing; you 'fix' them to the glass platen using the surface tension of the water and ensure that they are flat using a wiper blade. Bumps under the films will destroy your specimen.
- Any debris on the film will reduce its useful life; if the pad dries with polishing paste still present you should discard it.
- Minimize the effect of debris, which you produce on the polishing film as you thin your sample, by paying careful attention to where you place the specimen on the polishing wheel; orient interfaces in cross section samples normal to the radius and don't cross the debris trail.

With practice, tripod polishing can dramatically reduce the time required for the final thinning step. This tool has had a major impact on making TEM a quality-control instrument, particularly in the semiconductor industry.

10.4 FINAL THINNING OF THE DISKS

10.4.A Electropolishing

Electropolishing can only be used for electrically conducting samples such as metals and alloys. The method can be relatively quick (a few minutes to an hour or so) and it can produce foils with no mechanical damage. But it can change the surface chemistry of the specimen and it can be hazardous to your health as you can see from the safety section at the start of the chapter.

The basic premise is that there is a certain applied voltage at which the current due to anodic dissolution of the specimen creates a polished surface rather than etching or pitting, as shown in Figure 10.7. The classical jet polish is shown in Figure 10.8A. By keeping the volume of the reservoir constant, the jet falls under constant pressure. The voltage is applied between the tip of the pipette and the specimen. A twin-jet apparatus can be used to pump a jet of electrolyte onto both sides of the dimpled disk, as shown schematically in Figure 10.8B. A laser beam or light sensor detects transparency and a warning sound is given. At the warning, the electrolyte flow must be cut off immediately to prevent loss of thin area and the disk must be rapidly extracted from the electrolyte and washed in solvent to remove any residual film of electrolyte which may etch the surface.

Undoubtedly you get better at electropolishing with practice, but reproducing the correct conditions of temperature, electrolyte solution chemistry, stirring rate, applied voltage, polishing current, etc., can only be achieved through trial and error.

10.4.B Ion Milling

Ion milling involves bombarding your delicate thin TEM specimen with energetic ions or neutral atoms and sputtering material from your film until it is thin enough to be studied in the TEM. A schematic diagram and a commercial model are shown in Figure 10.9. The variables which you control include the voltage,

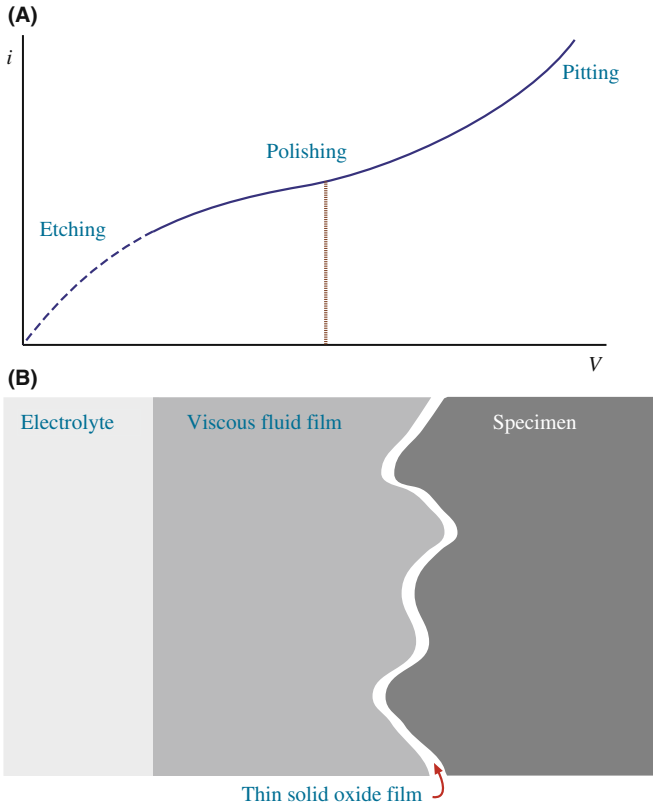


FIGURE 10.7. (A) Electropolishing curve showing the increase in current between the anode and the cathode as the applied voltage is increased. Polishing occurs on the plateau, etching at low voltages, and pitting at high voltages. (B) The ideal conditions for obtaining a polished surface require the formation of a viscous film between the electrolyte and the specimen surface.

temperature of the specimen (e.g., cold milling (liquid N2)), the nature of the ion (Ar, He, or a reactive ion (iodine)) and the geometry (the angle of incidence).

ION THINNING

Variables are ion energy, angle of incidence, vacuum, initial surface topology, initial chemistry, initial orientation, initial crystallography of the surface, beam energy, and beam profile. Note the word *initial*.

An accelerating voltage of 4–6 keV is usually used. The ion beam will always penetrate the specimen to some extent, so we minimize this by inclining the incident ion beam to the surface of the specimen. In the past, we often aligned the ion beam at an angle of 15–25° to the surface. However, Barna has shown that this angle of incidence should be avoided in many cases since it leads to compositional thinning; use an inclination of $\leq 5°$ to avoid preferential thinning and minimize ion implantation. Some implantation will occur so that the chemistry of the near-surface region is changed and the

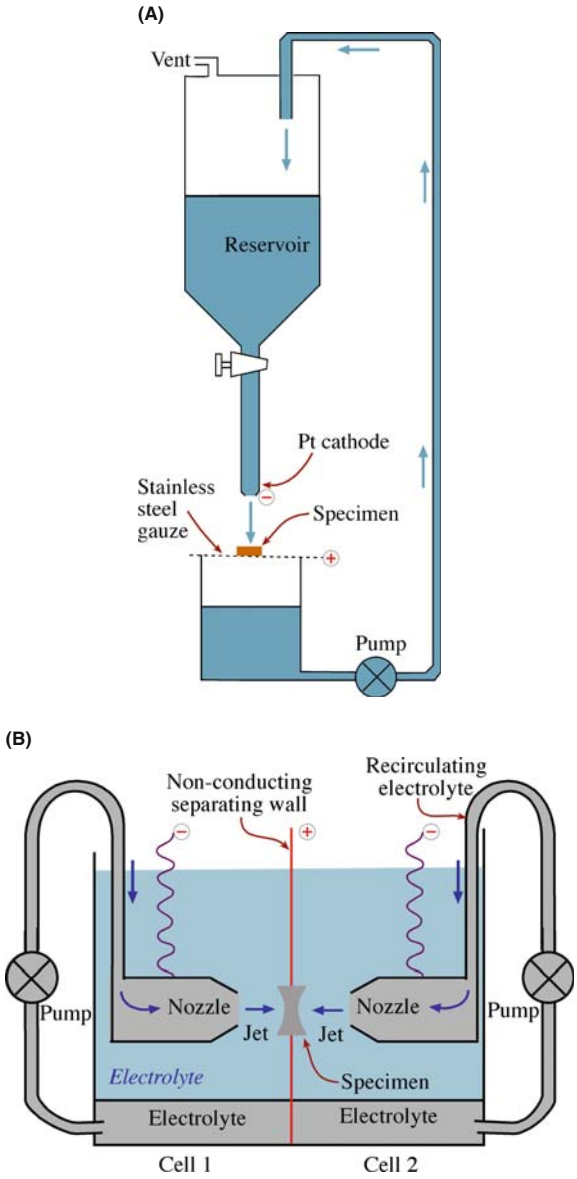


FIGURE 10.8. (A) Jet electropolishing by allowing a single jet of gravity-fed electrolyte to thin a disk supported on a positively charged gauze. The disk has to be rotated periodically. (B) Schematic of a twin-jet electropolishing apparatus. The positively charged specimen is held in a Teflon holder between the jets. A light pipe (not shown) detects perforation and terminates the polishing.

material is physically damaged (the top layer is often amorphized). If you use a low angle of incidence ($< 5°$), you'll deposit the energy of the ion beam in a region close to the surface of the specimen. A lower beam energy or a lower Z ion will also do less damage, but in both cases milling time will increase (Figure 10.10). In principle, you could also control the vacuum around the sample.

One thing you must remember is that ion thinning is closely related to ion-beam deposition. One manufacturer uses a similar arrangement to coat samples for SEM. The result is that material removed from one

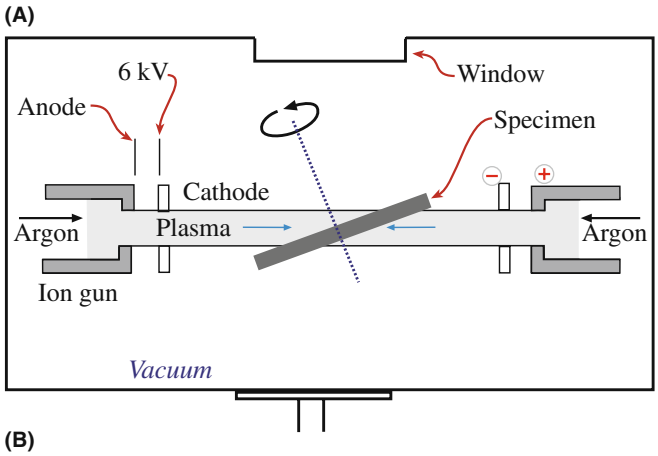


FIGURE 10.9. (A) Schematic diagram of an ion-beam thinning device: Ar gas bleeds into an ionization chamber where a potential up to 6 keV creates a beam of Ar ions that impinge on a rotating specimen. Although not shown, the whole apparatus is under vacuum. The specimen may be cooled to liquid-N_2 temperatures and perforation is detected by the penetration of ions through the specimen. (B) Typical ion mill.

part of the sample can easily be redeposited elsewhere on the sample.

The theory of ion milling is complex. We can define the sputtering yield to be the number of atoms ejected per incident ion; the yield depends on the mass of the incoming ion. The yield also depends on the ion used and the sample being milled. The principal variables are

- *The ion*: mass, energy, charge, and angle of incidence
- *The 'target'*: mass density, atomic mass, crystallinity, crystal structure, and orientation

Ar is used because it is inert, heavy, and not naturally present in most samples. Special applications may use reactive iodine, or add oxygen, etc.; this idea of reactive-ion etching is commonly used in semiconductor processing. The problem is that the reactive ion may contaminate or corrode your thinning device, the diffusion pumps, etc. Heavy ions give less penetration, but create more damage.

Most of the thinning parameters are generally fixed except the ion energy, the angle of incidence and any rotation, and the temperature of the specimen. A typical approach is to start with rapid thinning conditions (heavy ions, high incidence angle) and slow the thinning rate as perforation approaches. The effect of incidence angle on the thinning process is shown in Figure 10.10. Cooling the specimen is recommended for almost all materials; otherwise, it is possible that the ion beam might heat it to 200°C or higher. Even in metals which have good thermal conductivity, the creation of vacancies through ion damage can cause diffusional changes equivalent to heat treatment at such temperatures.

You may encounter discussions of whether to use ions or neutral atoms; one idea is that neutralized ions should not be affected by charging of a non-conducting specimen. It is not clear that neutral atoms remain neutral throughout the thinning process so this may be a moot point.

Ion milling is the most versatile thinning process, being used for ceramics, composites, polyphase semiconductors and alloys, and many cross section specimens. In addition, fibers and powders, which constitute a wide range of important materials, can also be thinned by ion milling. To do this, you have to first embed the particles or fibers in epoxy and transfer the mixture into a 3-mm brass tube for strength. The next step is to saw the tube/epoxy mixture into 3-mm disks and finally dimple and ion mill to electron transparency, as shown in Figure 10.11. A similar method (but without the brass tube) can be used prior to ultramicrotomy of powders and fibers (see Section 10.6.B).

Remember: Always beware of artifacts: some stories best illustrate this. Goodhew reports that Ar bubbles

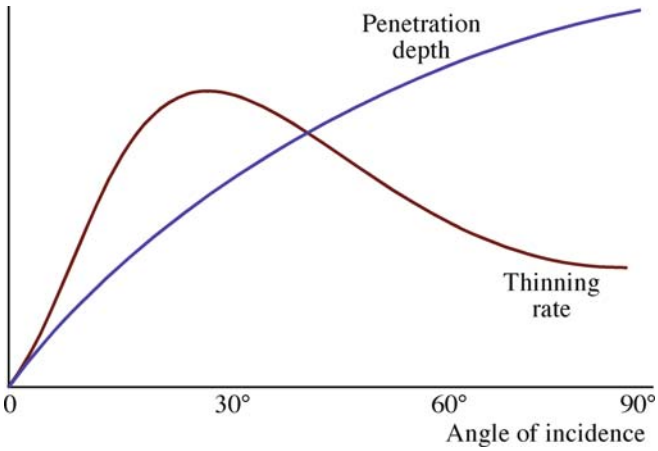


FIGURE 10.10. Variation in penetration depth and thinning rate with the angle of incidence. High-incidence angles promote implantation, which is undesirable. The rate of thinning reaches a maximum at ~20° incidence, after which the beam penetrates rather than sputters the sample surface. Initial thinning should start at 20–30° reducing to <10° as perforation approaches.

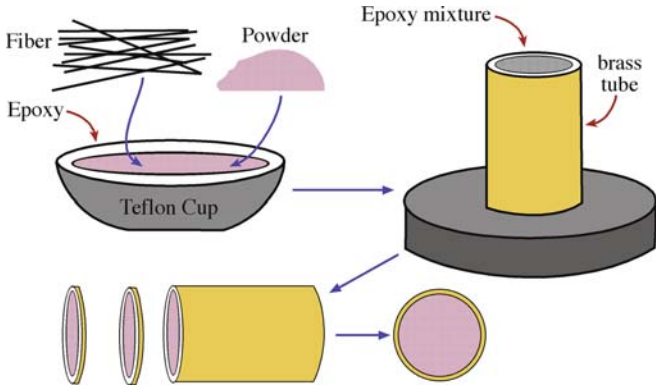


FIGURE 10.11. Sequence of steps for thinning particles and fibers by first embedding them in epoxy and forcing the epoxy into a 3-mm (outside) diameter brass tube prior to curing the epoxy. The tube and epoxy are then sectioned into disks with a diamond saw, dimpled, and ion milled to transparency.

form in silicon at a depth of ~10 nm after 5-keV thinning. Elemental analysis (XEDS) of some β-aluminas which had the correct structure by HRTEM (composition $K_2O \cdot 11Al_2O_3$) gave a composition with the K completely replaced by Ar. Glasses and zeolites can also accommodate large amounts of Ar. Cooling the specimen can often reduce contamination and surface damage. It is best to use two ions guns. If this is not acceptable, because you want to study the surface region, then you may want to coat one side with a polymer-protective lacquer and then dissolve this coating after thinning to remove sputtered material.

Why Rotate and Cool the Specimen? The specimen is usually rotated (at a few rpm) during thinning, otherwise you tend to get surface structure—grooves which run in certain directions; if you see these, check to see that the rotation has not stopped. In the preparation of cross-section specimens, you may use beam blockers and rotation control. In the first, you physically block the sample to shield it in certain directions from the ion beam so that it cannot thin, say an interface, preferentially. In the second, you vary the rate at which you rotate the sample to achieve the same effect. The latter is preferred if it is available since the time spent thinning the specimen is maximized.

Why cool the specimen? You can minimize atom migration in or on the specimen. We noted above that the specimen might be heated to >200°C otherwise. An additional advantage is that the cooling system also cools the surroundings to give a contribution of cryopumping and simple cryotrapping. However, you have to give the specimen time to warm up after milling which can increase preparation times.

Tilting the Specimen This depends on your ion miller but if you're choosing a new machine there may be an advantage in tilting the gun rather than the specimen. If the specimen is inclined, then you need a clamping ring and you may sputter this when you thin the specimen. This has led to the development of ion polishing instruments (see later) where the ion thinner has been optimized to provide a low angle without a retaining clamp. The specimen rests on a support and can be thinned at an angle of 4–5°.

Practical Design of the Ion Miller The schematic diagram in Figure 10.9 doesn't do justice to a modern ion miller which is a highly sophisticated piece of equipment. Two ion guns are available to thin from each side. The operating vacuum is $< 10^{-3}$ Pa without Ar and $10–10^{-1}$ Pa when Ar is bled into the gun. The ion guns are basically hollow chambers into which the Ar is introduced; then it is ionized and accelerated through a hole in the cathode. The hole gradually enlarges due to ion sputtering and cathodes need replacing after some time to maintain a high-intensity ion beam. More advanced gun designs incorporate saddle fields to focus the ion beam at the specimen and increase the thinning rate. The beam can be neutralized in some systems if the charged ions cause too much damage.

Some special phrases you'll encounter

- Reactive ion milling. The classic example is the use of iodine in the work described by Cullis and Chew. Iodine has a clear advantage for InP where the formation of In islands under Ar thinning is suppressed. In CdTe only growth defects were observed in iodine-thinned specimens, but many other defects were found in the same material thinned using Ar ions (Figure 10.12).
- Beam blockers and variable rotation speeds. Often the epoxy in a cross-section specimen thins faster than the specimen. Therefore we want to direct the ion beam at the different materials for different amounts of time. The two approaches used are blocking the beam geometrically using 'beam blockers' or varying the rotation velocity; e.g., you don't want the beam to thin along the interface. The latter approach can be extended further to oscillate the specimen, always keeping the ion beam at the same angle of incidence, so that it is never parallel to the interface.
- Low-angle, low-energy ion mills. Examples include the PIPS (Gatan's precision ion polishing system) and Baltec's Gentle Mill. These ion mills combine high-powered ion guns and a low angle of incidence (4°) to thin one side of a specimen with minimum surface damage and heating. The low incidence angle removes any surface roughness and differential thinning problems, while the high-power guns ensure reasonable thinning rates. The Gentle Mill can even be used to thin FIBbed specimens (when not on a C foil) (see Section 10.7).

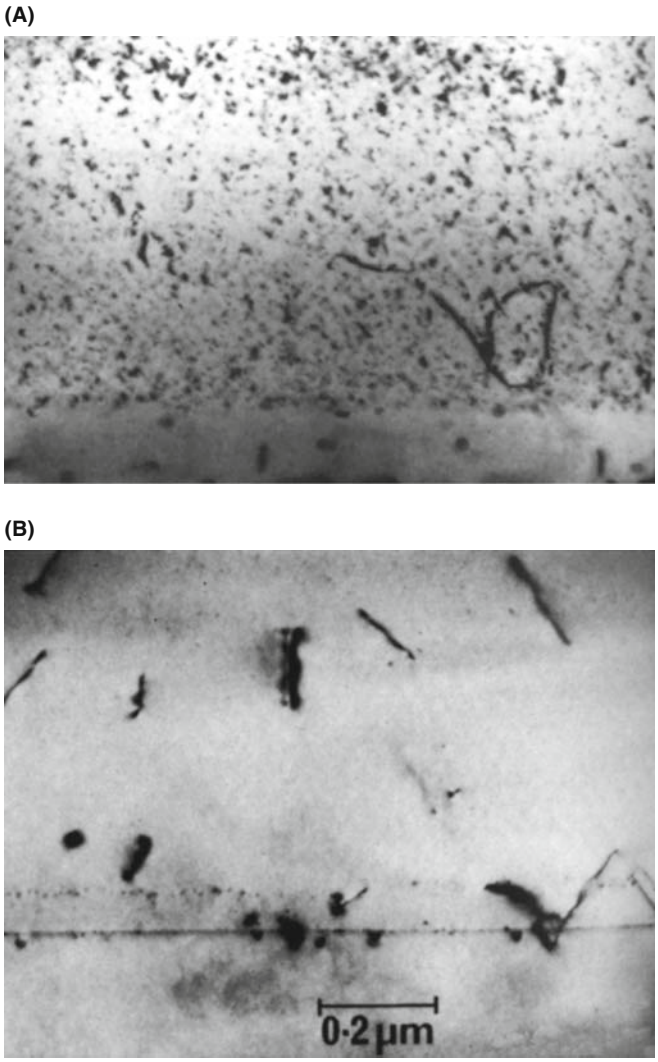


FIGURE 10.12. BF images of CdTe showing (A) defects (dark spots) in Ar-thinned specimen and (B) undamaged crystal thinned by reactive-iodine ion milling. The residual defects in (B) were formed during CdTe crystal growth.

Some final points to remember

- Materials thin at different rates. It's a good idea for the person responsible for the ion millers to run a test specimen periodically with nominally the same conditions, to be sure that the machine is still working optimally.
- Don't start with a thick sample. Always make the surface as smooth as possible before beginning to ion thin.
- Keep a record of what conditions you use: record the beam current, angle of incidence, rotation rate, and kV.
- Ion milling will form a layer on one surface or both which will probably be a combination of amorphous, highly damaged, and implanted material! The chemistry of the layer will differ from the rest of the specimen. The thickness of crystalline material will thus be less than the total thickness.

10.5 CROSS-SECTION SPECIMENS

The cross-section specimen is a special type of self-supporting disk. You must master this preparation technique if you are studying interfaces. We have often stressed that one of the principal limitations of the TEM is its insensitivity to variations in the structure and chemistry of the specimen in the direction of the electron beam. Therefore, if we are to look at structural and chemical variations close to an interface we have to prepare specimens in which the interface is parallel to the electron beam and this involves cross sectioning the sample. The most widely studied cross-section samples are semiconductor devices which often have multiple layers and therefore have multiple interfaces. But any composite materials, samples with surface layers (e.g., oxide-metal interfaces), MBE specimens, quantum-well heterostructures, etc., are candidates for this type of preparation.

There are numerous techniques for preparing cross-section specimens. Many details are reported in four MRS proceedings so we'll only describe a few basic principles. First, rather than trying to thin one interface only, the sample can be cut and glued together to produce several layers, rather like a club sandwich. Then the sandwich is sectioned such that we can see the layers, as shown schematically in Figure 10.13. In this process, a critical step is the gluing of the sections to form the sandwich. Several epoxies are available that cure at low temperatures, so that you won't heat treat the specimen inadvertently. The thickness of the epoxy layer must be such that it is thick enough for good adhesion, but not so thick that it is completely thinned away during final ion milling.

You can then cut the glued sections into 3-mm rods using an ultrasonic drill. Alternatively, you can cut the

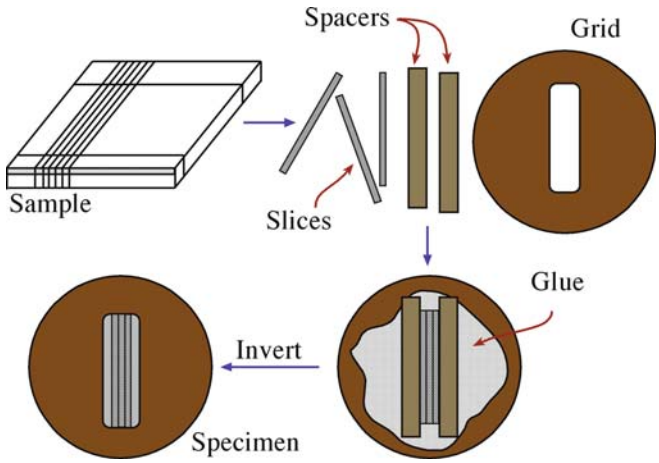


FIGURE 10.13. Schematic sequence for cross-section specimen preparation: the sample is cut into thin slices normal to the interfaces which are glued together between spacers which could be Si, glass, or some other inexpensive material so that they are wider than the slot in the grid. The 'club' sandwich is then itself glued to the grid (over the slot) and ion milled to perforation.

samples smaller and encase them in a 3-mm thin-walled tube. Section the filled tube into disks which you can then ion thin. The advantage of this method is that the final specimen has a thick ring of the tube metal around it, which gives it mechanical stability. With multiple interfaces the final thinning is almost always guaranteed to produce electron transparency at a useful region.

10.6 SPECIMENS ON GRIDS/WASHERS

The alternative to self-supporting disks is to make small electron-transparent portions of the specimen or create particles and support them on a thin film on a grid or washer. We can deposit these small particles on amorphous or crystalline films. The classic example is the amorphous carbon film (the holey carbon film), but this is not always the best choice. Some of the particles of the material of interest will be located partially over a hole so that they do not overlap anything else.

The thin supporting film should have a uniform thickness; the idea is that you are not actually interested in this material and therefore want to minimize its effect on the image of the material you are interested in.

The particles may stick to the film or may have to be clamped between two grids. Special hinged 'oyster' grids (see Figure 10.3) are available which make this very easy. Some of the processes we've already discussed can be used to make these specimens.

10.6.A Electropolishing—The Window Method for Metals and Alloys

Electropolishing is an application of electrochemistry and is regarded by many as a 'black art': a recipe which works one day but might not work the next. We can electropolish a thin sheet of metal. First cut the sheet into a square ~10 mm on the side, then seal the edges with a polymer lacquer to prevent preferential attack. The 'window' of exposed metal is immersed in electrolyte (usually cooled to slow the rate of dissolution), surrounded by a cathode and a voltage is applied, as in Figure 10.14A. The solution may or may not be stirred. The correct voltage will ensure that a viscous layer of electrolyte builds up at the surface of the specimen which results in uniform controlled thinning without pitting or corrosion. After some time, which you have to determine experimentally, the sheet is removed, cleaned, and turned through 180° and replaced in the bath as shown schematically in Figure 10.14B. If this procedure is done correctly (and this might require several rotations) the sheet will finally thin in the center. If final thinning occurs too near the top of the sheet, the edge of the perforation is smooth and relatively thick. After perforation, remove the sheet and cut off slivers of material from around the perforation using a scalpel under an inert solvent such as ethanol. Catch the floating slivers on oyster grids, dry them and they are ready for viewing.

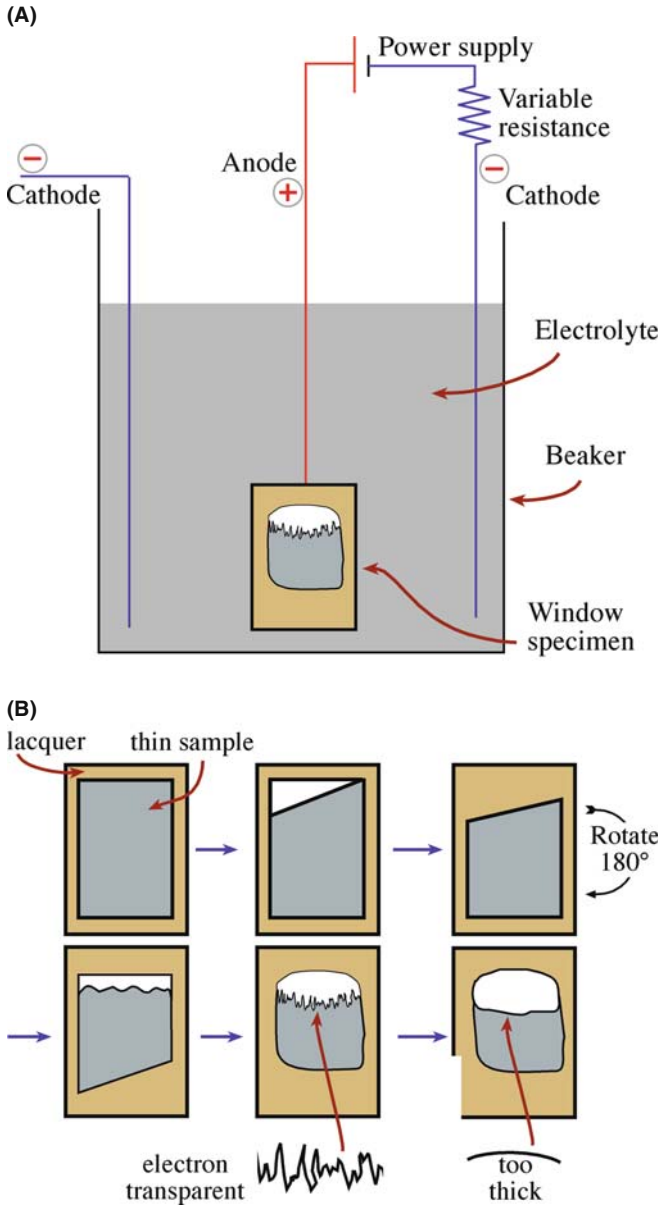


FIGURE 10.14. Window polishing. (A) A sheet of the metal ~100 mm^2 is lacquered around the edges and made the anode of an electrolytic cell. (B) Progress during thinning: the initial perforation usually occurs at the top of the sheet; lacquer is used to cover the initial perforation and the sheet is rotated 180° and thinning continues to ensure that final thinning occurs near the center of the sheet; if the final edge is smooth rather than jagged it is probably too thick.

10.6.B Ultramicrotomy

The microtome has long been used for sectioning biological materials. A tome is a 'piece cut off' so a microtome refers to the instrument used to cut a very thin tome (not like the one you're reading). With care and

much practice the biologist can reconstruct a 3D picture of the specimen. For visible-light microscopy the specimens are usually <0.1 mm thick; for the TEM the slices may be <100 nm thick and the instrument is known as an ultramicrotome. These instruments are routinely used for biological samples or for polymers where the samples tend to be quite soft. More recently, they have been used for many studies of crystalline materials. The principal advantages of the technique are that it leaves the chemistry unchanged and is thus ideal for AEM specimens, and you can use it to create uniform thin films of multiphase material. The main disadvantage, of course, is that it fractures and/or deforms the samples and therefore is most useful in cases where the defect structure is of secondary (i.e., zero) importance.

The ultramicrotome operates by moving the specimen past a knife blade. The blade can be glass (cheap) for soft materials but will be diamond for harder ones. Since there are so many possible applications, we will describe a few and refer to the references at the end of the chapter for more details. Two processes can occur in principle: the knife can cut a soft sample or it can cause a partly controlled fracture in a hard/brittle sample. In either case the limiting process is usually plastic deformation of the sample. The principles of this technique are shown in Figure 10.15.

You may also find ultramicrotomy useful if you want to study particles or fibers which are too small to thin individually but are too large to be electron transparent. You can embed the sample as we saw for the ion-thinned particles but without using the metal sheath (see Figure 10.11). We also use epoxy if the sample contains so many interconnected pores that it cannot be thinned mechanically. For porous materials, place the sample in a vacuum chamber, pump out the chamber, and coat the sample with epoxy using a dropper in the chamber. When the sample is fully encapsulated, admit air to the chamber so as to push the epoxy into the pores. After curing, you can ultramicrotome the sample in the usual way.

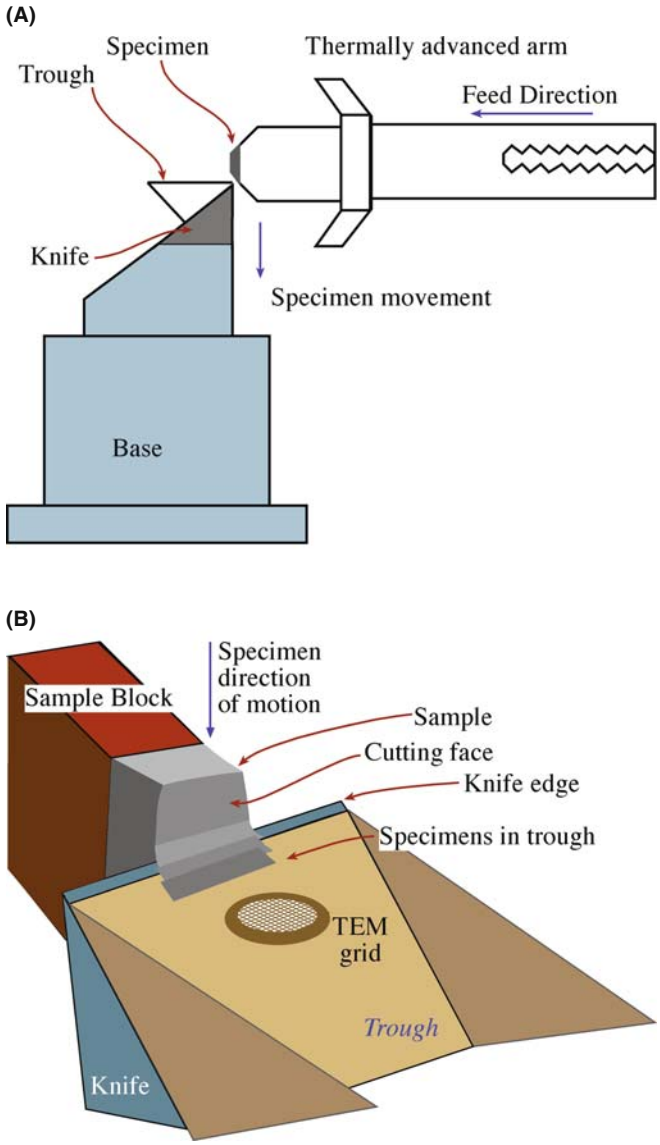


FIGURE 10.15. Ultramicrotomy. (A) The sample is first embedded in epoxy or some other medium or the whole sample is clamped and moved across a knife edge. (B) The thin flakes float off onto water or an appropriate inert medium, from where they are collected on grids.

10.6.C Grinding and Crushing

Many brittle materials such as ceramics and minerals are most easily prepared by crushing in a clean pestle and mortar (preferably in an inert liquid). The liquid containing the particles can then be ultrasonically stirred and allowed to settle. Particles suitable for TEM are too small to be seen by eye and the supernatant liquid in which they remain should appear clear. A drop of this liquid, if placed on a holey carbon film on a grid, will evaporate in a dry environment, leaving a distribution of the particles on the support film. If the particles have to be crushed dry, then agglomeration can be a problem. Electrostatic forces sometimes cause small particles to clump together and distributing them on a grid can be very difficult. In these cases, it sometimes pays to mix up the crushed material in an epoxy, then ultramicrotome the epoxy, as we just described in the previous section.

We can collect dust particles found in airborne pollution by simply exposing a support film (on a grid) to the atmosphere for a period of time. Interstellar dust can be sampled from a spacecraft or a high-flying plane.

10.6.D Replication and Extraction

These methods are among the oldest TEM specimen-preparation techniques. We use direct replication to

study fracture surfaces or surface topography in general. Evaporate a carbon film on the surface of interest, then etch away the underlying surface with an acid so that the carbon film floats off. If you coat this film with a heavy metal at an oblique angle, you will thus produce a sample that shows enhanced mass-thickness contrast (see Chapter 22); support the film on a grid for observation. As an alternative (Figure 10.16A) you can first replicate the surface by softening a plastic, pressing it on the surface, and allowing it to harden. Pull off the plastic replica, coat it with carbon, then dissolve the plastic with a suitable solvent, and pick up the carbon replica on a support grid. If the carbon replica is produced directly from a metal surface, it may be necessary to dissolve some of the metal with acid then float off the carbon onto distilled water before picking up on a grid, as shown in Figure 10.16B. After picking up on a grid it may be useful to coat the replica obliquely with a heavy metal to enhance any topographic (thickness) contrast.

Extraction replication has seen a resurgence of interest since AEM techniques appeared, because we can extract a particle from its surrounding matrix, thus allowing us to analyze that phase alone without interference from electron scattering into the matrix.

The various steps for extraction are shown in Figure 10.17A. The sample is polished metallographically to expose the particles on the surface. An appropriate etching process is used to remove the matrix such that the particles stand proud of the surface. A carbon film is evaporated onto the surface and scored into ~2 mm squares. Then the etching is continued. As the matrix is dissolved, the squares of carbon film float to the surface carrying the particles with them. Catch one of these squares on a grid and you have your specimen ready for the TEM as shown in Figure 10.17B. Again, oblique shadowing may be useful to enhance image contrast, but not if you plan to use AEM.

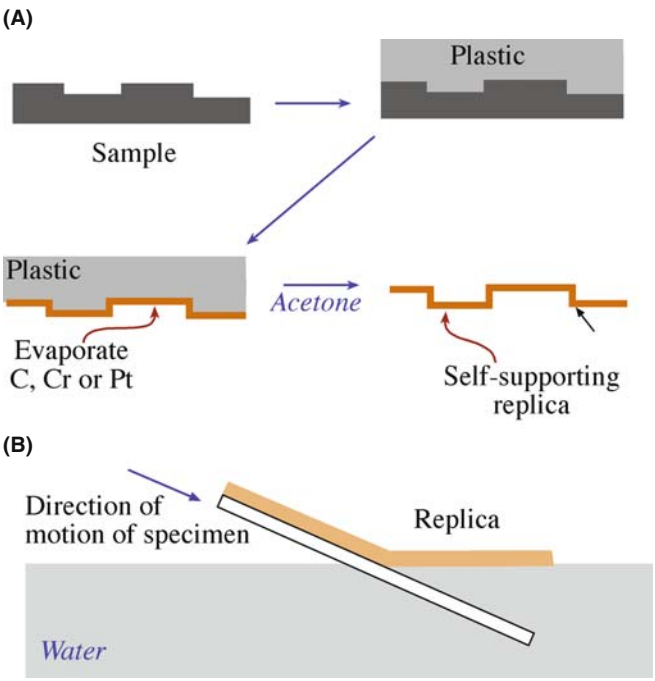


FIGURE 10.16. (A) Replication of a surface by the two-step method: spray acetone on the surface to be replicated before pressing a plastic (usually cellulose acetate) onto the surface which softens in contact with the acetone; the plastic is removed from the surface when it has hardened and a C, Cr, or Pt film is evaporated onto the replicated plastic surface; the plastic is then dissolved with acetone and the evaporated film retains the original topography. (B) Alternatively, the direct carbon replica of a metal surface may be floated off on distilled water after scratching the carbon and etching to free the film, which may subsequently be shadowed obliquely to enhance the topography.

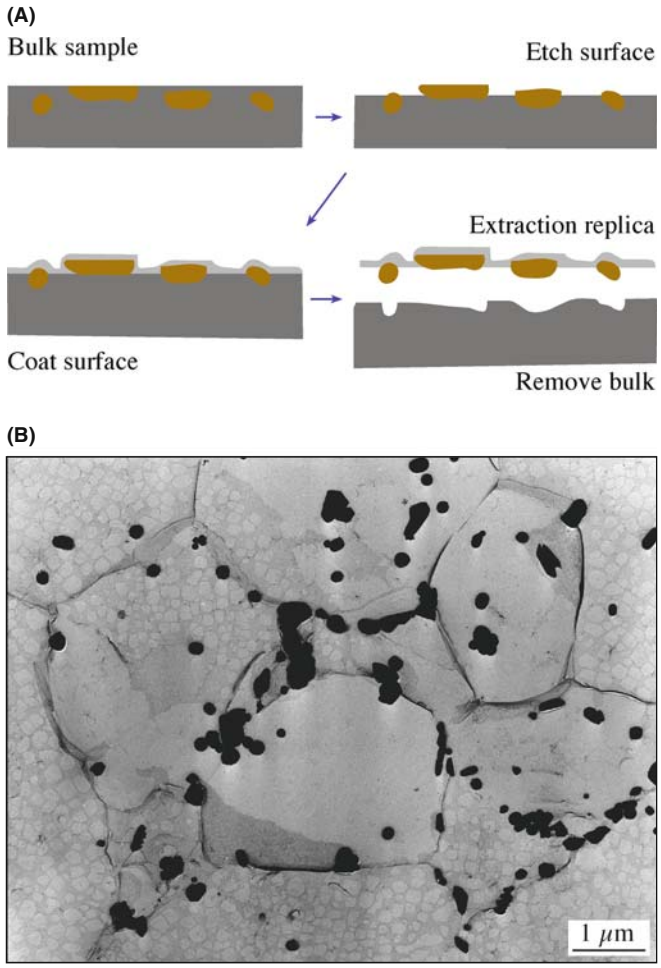


FIGURE 10.17. (A) Making the extraction replication: particles embedded in a matrix are revealed by etching the matrix, which leaves the particles standing proud of the surface; a thin amorphous carbon film is evaporated over the particles, then the rest of the matrix is etched away leaving the particles adhering to the carbon film. (B) Example from a γ/γ' alloy showing not only that the particles are mainly located at the grain boundaries but also the different contrast from γ' grains and two-phase γ/γ' grains.

10.6.E Cleaving and the SACT

Cleaving is one of the oldest techniques and has been used to make thin specimens of graphite, mica, and other layer materials that are weakly bonded along one plane. The classical idea is to attach adhesive tape to both sides of the sample and then pull the two pieces of tape apart. This process is repeated until the specimen is thin enough for TEM. You can really only tell this by experience: as it becomes thinner, graphite becomes a lighter shade of gray in transmitted visible light. Molybdenite (MoS_2) becomes a lighter shade of green as illustrated in Figure 10.18. Place the tape with the thin flake of material in a solvent to dissolve the glue (all traces of glue *must* be removed). This technique is not as easy as it once was. The glues used to be readily soluble in trichlorethylene which is now a known carcinogen.

A special variation on cleaving, known as the small-angle cleaving technique (SACT), is illustrated in Figure 10.19. The idea is to propagate a crack through the sample along a plane that is not a natural (crystallographic) fracture plane and then create another crack that is shallowly inclined to the first fracture surface. The technique can be applied to crystalline samples such as Si coated with thin films or to coated glass samples (as shown in Figure 10.19) that have no preferred fracture plane. In the case of glass, this is particularly attractive since the only alternative is ion milling which tends to implant argon into the open structure of the glass and is thus not suitable for AEM specimens.

Figure 10.19A shows a sample that has been scribed with several parallel lines. Each of the rectangular samples can then be fractured again by pressing on it with a stylus to produce a sharply wedged specimen. With luck, the resulting sample is electron transparent as seen in Figure 10.19D where a small particle of NaCl is imaged between a glass substrate and a coating layer. Although it is a hit-and-miss technique, you can make so many specimens in one sitting that a hit is assured (with lots of practice). You can then put this sample into the FIB and produce an even better sample without spending too much FIB time (money).

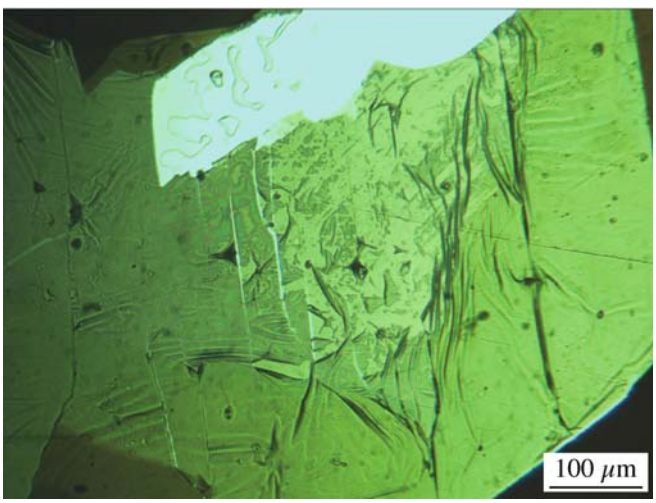


FIGURE 10.18. Image of cleaved MoS_2 showing regions of different shades of green, which correspond to different thicknesses.

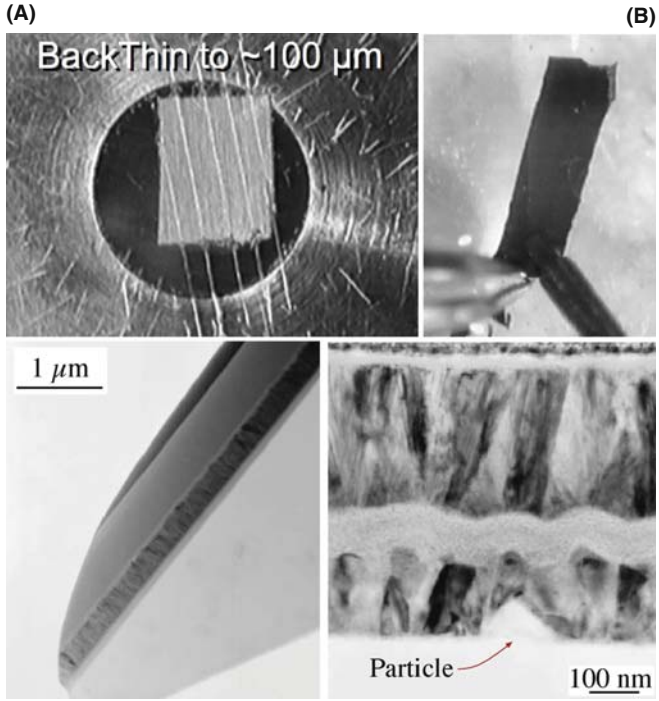


FIGURE 10.19. SACT of a coating on glass. (A) Scratch the sample; (B) cleaving along the scratch; (C, D) TEM images.

SACT

The small-angle cleaving technique is invaluable for films on Si or glass where there is no crystal structure; making the 90° wedge would be a LACT!

10.6.F The 90° Wedge

The 90°-wedge specimen was developed because many compound semiconductors such as GaAs are grown with a (001) surface and can be easily cleaved on the (110) and ($1\bar{1}0$) planes that are perpendicular to this growth surface. When you are practiced at cleaving the sample as shown in Figure 10.20, you can examine a specimen in the TEM within 30 minutes of completing the growth.

Mount the specimen as shown in the figure, preferably so that you won't need to tilt it in the microscope. Although the specimen is only transparent close to the edge of the 'hole,' you will have a long strip of material suitable for viewing. As always, beware of artifacts. If your specimen is perfect, you will know exactly how

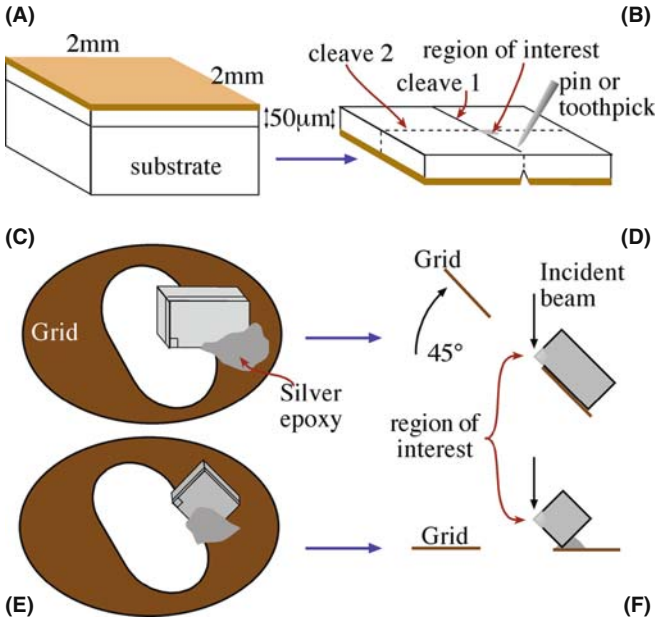


FIGURE 10.20. The 90°-wedge specimen: (A) prethin to create a 2-mm square of the multilayers on a Si substrate; (B) scribe the Si through the surface layers, turn over, and cleave; inspect to make sure the cleavage is clean, giving a sharp 90° edge; reject if not; (C, E) mount the 90° corner over the edge of a hole in a Cu grid; (D, F) then insert in the TEM; note that two different orientations are available from a single cleavage operation.

thick it is at the position you choose for study. We will find this wedge useful when we discuss image contrast in Part 3.

10.6.G Lithography

Here we use a technique developed for advanced engineering applications. Lithography is used in the microelectronics industry to define fine lines of width down to 100 nm. An illustration of how lithography can be used specifically to prepare TEM specimens (as opposed to generating a structure which might best be characterized by TEM) is shown in Figure 10.21. We can draw lines on the layered material using standard lithographic techniques. Material on either side of the lines is then removed by etching (chemical or ion) to give a plateau which is thin in one direction. We then remove most of the remaining substrate and attach the specimen to a support washer. We can then observe the specimen directly in the TEM. Although the width (formerly height) of the electron-transparent region is narrow it can extend across the entire hole in the 3-mm disk. The major disadvantages or limitations of the technique are (i) the dimension in the direction of the electron beam is fixed by the lithographic capabilities and (ii) tilting the specimen may quickly cause the thicker region to block the electron beam.

10.6.H Preferential Chemical Etching

The principle behind this technique is the same as for lithography: we remove part of the sample to leave an area which is electron transparent. The trick is to keep part of the final specimen thick enough for handling, or ideally for supporting, the specimen. Naturally, this approach only works with certain materials although the principle might be extended to other thin films. The technique has been used for III–V compounds where $Al_{1-x}Ga_xAs$ acts as an etch stop for GaAs and for Si where an etch stop can be produced by implanting with boron (Figure 10.22). In both cases, the resulting thin layers may be used as substrate materials for thin-film studies rather than as the subject of study in their own right. An extreme example of this approach is the use of thin films of Si_3N_4 as an 'etch stop' so that a uniform layer of amorphous Si_3N_4 remains across the window. Such specimen supports are commercially available; diamond films can be made in a similar way.

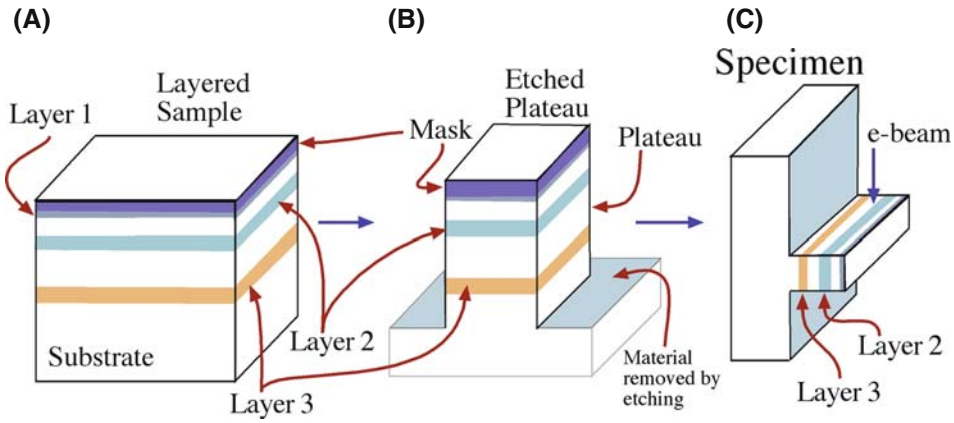


FIGURE 10.21. Etching of a multilayer sample (A). Etch away most of the sample, leaving a small etched plateau (B); mask a region <50 nm across and etch away the majority of the surrounding plateau. If this thin region is turned 90° and mounted in a specimen holder (C), the interfaces are now parallel to the electron beam.

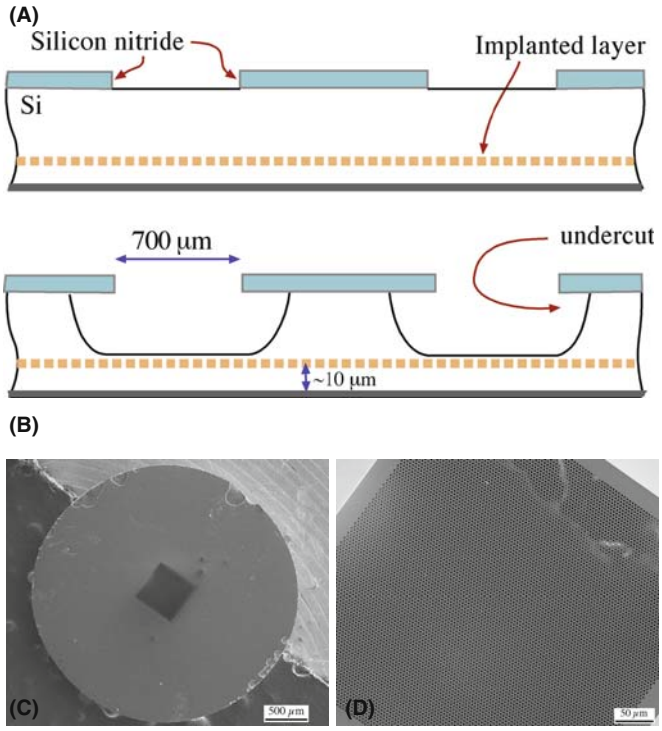


FIGURE 10.22. Lithographic techniques applied to thinning a multilayer specimen: (A) the unthinned sample is shown with a grid of Si_3N_4 barrier layers evident. Etching between the barrier layers, shown in (B), produces an undercutting down to the implanted layer which acts as an etch stop, producing a uniform layer ~10 μm thick. Further thinning with a different solution produces large areas of uniformly thin material (not shown) supported by the Si_3N_4 grid and the remaining unthinned regions. (C, D) A commercial Si_3N_4 thin films support disk; (D) is the enlarged view.

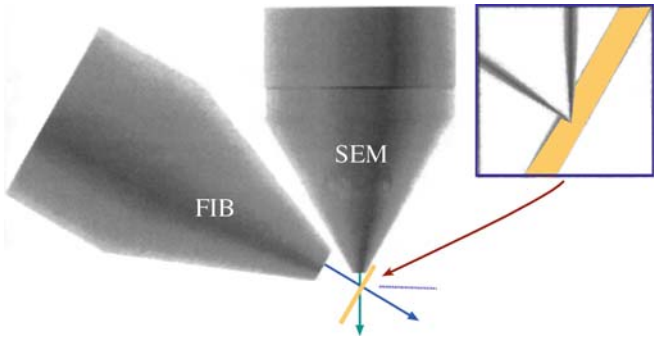


FIGURE 10.23. Schematic of a two-beam (electron and ion) FIB instrument.

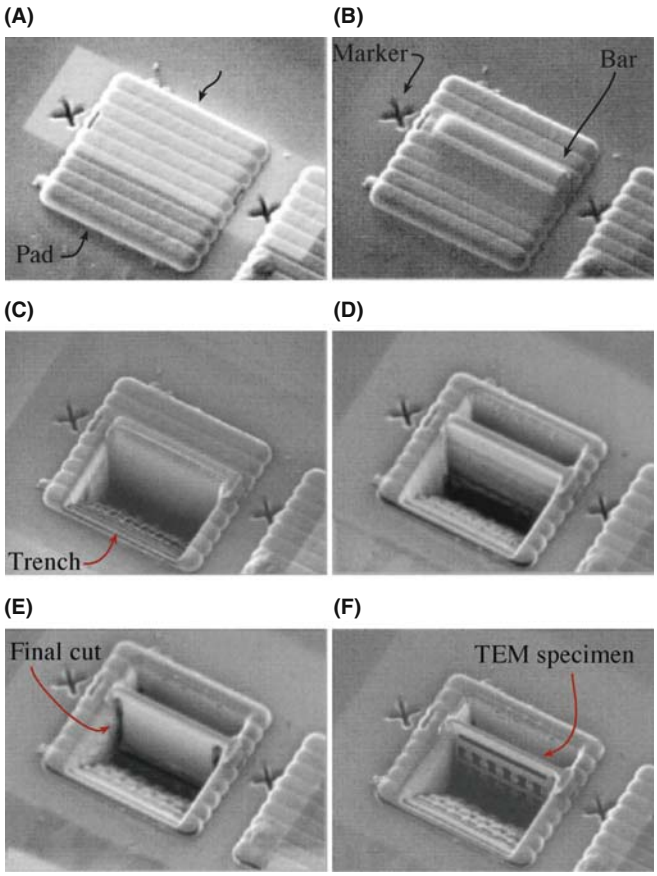


FIGURE 10.24. Stages in making TEM samples using a FIB instrument. (A) The area of interest has been marked. (B) A Pt bar is deposited to protect this area from the Ga beam. (C, D) The two trenches are cut. (E) The bottom and sides of the slice are (final) cut. (F) The TEM specimen is polished in place before extracting it.

10.7 FIB

The focused-ion beam (FIB) instrument is becoming much more readily available as prices moderate and their value is realized; however, a FEG-equipped FIB can still cost more than your TEM. We deal with FIB more extensively in the companion text but include a summary here because you must know about it even if you can't get access to one yet.

When preparing TEM specimens we can think of the FIB essentially as an SEM with a built-in ion mill. (Sometimes it's an ion gun with an SEM attachment.) The single ion gun produces a well-controlled beam of Ga ions (rather than Ar used in the ion mill). In the simplest (cheaper) design the ion beam also acts as the electron beam of the SEM with the secondary electrons being used to form the 'SEM' image of the sample. A schematic of the FIB is given in Figure 10.23. The various stages in the preparation process are shown in Figure 10.24. The pad in (A) is the coating of Pt. The two Xs, to mark the region of interest, have been drawn on the sample using the ion beam, and another Pt strap is deposited between them (B). Next, two staircases are cut out on either side to leave the thin wall shown in (C) and (D). In (E) the 'wall' has been trimmed away at the sides so that it is only supported at the top. The last step

(F) is to ion-polish the thin wall until it is really a thin TEM specimen and finally attach it to a probe for lift out or use static electricity to lift it out and place it on a supporting (usually C) film. Attaching the FIBbed specimen to a probe is becoming the norm since it allows the FIBbed specimen to be further cleaned to remove Ga contamination and/or reduce the thickness further. An FEI version of the instrument is shown in Figure 10.25. We'll discuss the details of FIB in the companion text because not everyone can afford to buy a FIB or can afford to use one even if it is 'available'!

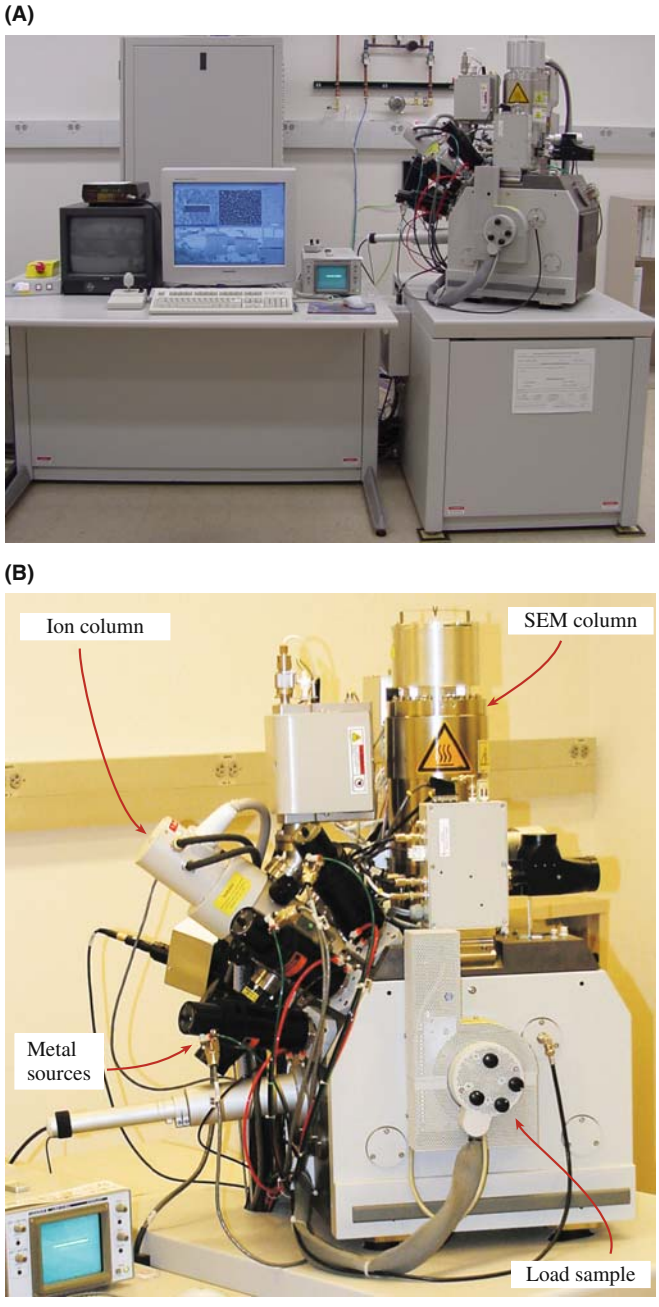


FIGURE 10.25. A dual-beam FIB instrument. (A) Overview; (B) enlarged view to compare with Figure 10.23.

10.8 STORING SPECIMENS

The best advice is to look at your specimen as soon as possible after preparation. If that is not possible, then keep your specimens under optimum conditions. Usually this means keeping them dry (water vapor affects the surface region of most materials), perhaps in an inert atmosphere (dry nitrogen works well, or a dry-pumped, oil-free vacuum desiccator) and in an inert container (a glass petri dish with filter paper).

The next problem is long-term storage; for periods up to 1 month, you can use the above procedure. If you want to keep the specimen longer, your choices can be more difficult. Don't use gelatin capsules for anything resembling 'delicate' material. Don't use slotted grid-holders for anything which might deform (break or bend) during handling; that rules out self-supporting ceramics, metals, and semiconductors. Always use vacuum tweezers to manipulate delicate specimens (remember safety when using mouth-vacuum tweezers). Remember, your most important specimen is the one most likely to break, bend, interact with sharp tweezers, or jump onto the floor.

Last, old specimens can be cleaned by ion polishing or chemical cleaning. This process does thin the specimen further so you may lose the area you originally studied. Ion polishing can also be useful for 'sectioning' specimens. The 'safer' re-cleaning (or refreshing) method is to use a plasma cleaner but this can change your sample.

If your specimen is a collection of nanoparticles, you will realize that such particles can be so reactive that they may have changed before they reach the TEM even for the first time. You may then have to use techniques that are more familiar to cryo-transfer users, such as transferring the specimen from the preparation chamber to the TEM in a controlled environment. Such a procedure is not routine using conventional holders.

10.9 SOME RULES

We stress once again that you must know what you want to study in your specimen before you begin specimen preparation. Figure 10.26 is a second flow chart to summarize the various possible options. Be aware of the limitations of the method you choose, particularly the artifacts introduced. Table 10.1 summarizes the artifacts introduced by various methods.

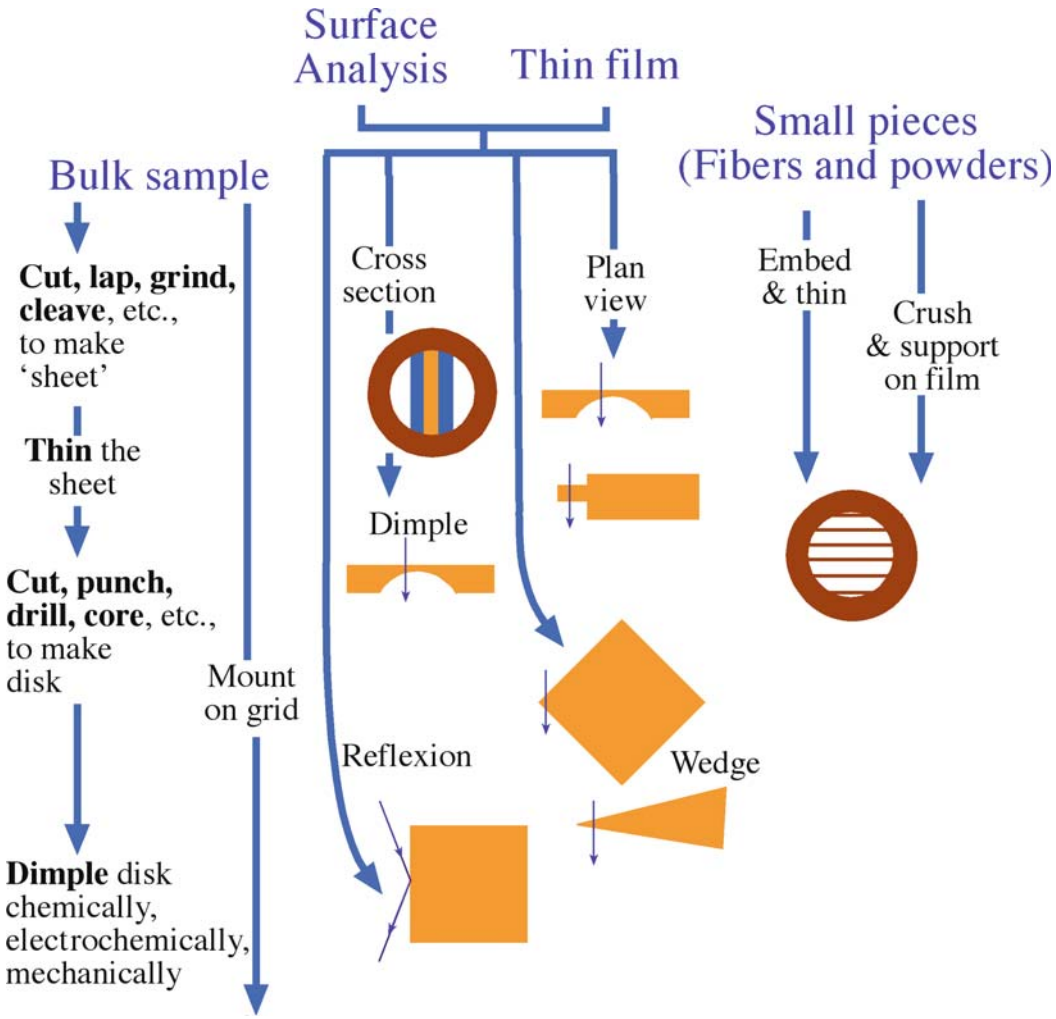


FIGURE 10.26. Summary flow chart for specimen preparation.

TABLE 10.1 Artifacts Produced During Specimen Preparation (after T. Malis)

Artifact/problem	Consequence
Variable thickness	
	▪ Limited local area for chemical mapping (EP, IT, C, CD)
	▪ Very limited area for EELS
	▪ Somewhat limited area for absorption-free XEDS
	▪ Omission of low-density defects
	▪ Distorted defect densities (EP, IT, TP)
Uniform thickness	
	▪ Limited diffraction information (UM)
	▪ Limited microstructure information (UM)
	▪ Handling difficulties (UM)
Surface films	
	▪ Bath residue, spec. dissolution and/or redeposition EP
	▪ Enhanced surface oxide (EP)
	▪ Extremely irregular topographies (IT)
	▪ Faster contamination buildup under beam (EP, R)
	▪ Retention of matrix on extracted particle
	▪ C-redeposition (UM—embedded, UM, C, R—support films)

Table 10.1(continued)

Artifact/problem	Consequence
	■ Cu_2O formation from Cu grids upon heating (R, UM, C) ■ Ion amorphization, diffusion-pump oil, redeposition (IT*)
Differential thinning	
	■ Different phases thin at different rates (EP, IT) ■ Different orientations thin at different rates (IT) ■ Grain/phase boundary grooving (EP, IT) ■ Anodic attack of matrix/particle (UM)
'Selectivity'	
	■ Perforation influenced by local defect structure (EP, IT) ■ Very limited or no microstructure information (C, R) ■ Weak local regions debond and fall out (all)
'False' defects	
	■ Microstructure obscured by high defect density (UM, CD) ■ Deformation-induced defects (EP, TP) ■ Ion-induced loops, voids (IT) ■ Heat-altered defects (EP, IT)

EP: electropolished; UM: ultramicrotomed; CD: controlled dimpling; R: extraction replication; IT: ion thinned; TP: tripod polish; C: cleavage (grinding, crushing).

CHAPTER SUMMARY

Specimen preparation is a craft and there is no substitute for hard work, careful, detailed experimentation and lots of practice as you seek to master it. This is the most tedious aspect of all of TEM work but, if you invest the time, your reward will be the best of times on the TEM itself. The quality of your data is at least directly proportional to the quality of your specimen (and this relationship is often far stronger than the linear nature just implied). You simply have to find the method that works best for your particular material. While there are many cookbooks available, the recipes are often too individualized and not to your specific taste.

There are few rules for specimen preparation except that thinner is usually better, although such specimens are more prone to artifacts. Think about each step and what it might do to change the microstructure or microchemistry of your material. Take care to avoid the physical dangers that are present whenever you use dangerous chemicals, ionizing radiation, or sharp knives. Be clean, use fresh materials, tidy up after yourself, and apply all the other lessons that you learned in kindergarten!

Although all the equipment mentioned here is available commercially, most were originally developed on a shoestring budget in someone's lab so you can always build your own electropolisher or even an ion mill. If you are working with brittle materials, buy or build a tripod polisher and learn how to use it.

A last reminder: The recipe books listed below are a great source of ideas. New recipes are appearing all the time. As is often the case in cooking it helps to see an expert chef in action to realize what is possible. In other words, when you have seen a really good TEM specimen, you'll know what yours should look like.

REFERENCES

These references only give a sampling. More extensive lists, especially for the more specialized techniques like FIB, are given in the companion text.

GENERAL TECHNIQUES

An extensive list of references is included in the chapter on specimen preparation in the companion text. The first four references below are essential: from the MRS Proceedings.

i. Bravman, JC, Anderson, RM and McDonald, ML (Eds.) 1988 *Specimen Preparation for Transmission Electron Microscopy of Materials* Mater. Res. Soc. Symp. Proc. **115** MRS Pittsburgh PA. (Number I in

the series.) We've updated the flow chart in the article on p51 by Goodhew, PJ, The tripod polisher is described by Klepeis, SJ, Benedict, JP and Anderson, RM on p179. (Pictures of the gem variety are shown in Figure 36.3 of *Ceramic Materials* by Carter and Norton.) Brown, JM and Sheng, TJ describe the use of lithography on p229.

ii. Anderson, RM (Ed.) 1990 *Specimen Preparation for Transmission Electron Microscopy of Materials, II* Mater. Res. Soc. Symp. Proc. **199** MRS Pittsburgh PA.

iii. Anderson, RM, Tracy, B and Bravman, JC (Eds.) 1992 *Specimen Preparation for Transmission Electron Microscopy of Materials, III* Mater. Res. Soc. Symp. Proc. **254** MRS Pittsburgh PA. Alani, R and Swann, PR (1992) discuss ion milling on p43.

iv. Anderson, RM and Walck, SD (Eds.) 1997 *Specimen Preparation for Transmission Electron Microscopy of Materials, IV* Mater. Res. Soc. Symp. Proc. **480** MRS Pittsburgh PA.

CHEMICAL POLISHING

Thompson-Russell, KC and Edington, JW (1977) *Electron Microscope Specimen Preparation Techniques in Materials Science* Macmillan Philips Technical Library Eindhoven Netherlands. Many recipes.

ION MILLING AND FIB

Barber, DJ 1970 *Thin Foils of Non-metals Made for Electron Microscopy by Sputter-Etching* J. Mater. Sci. **5(1)** 1–8. First use of ion milling to make a ceramic TEM sample.

Barna, A 1992 *Topographic Kinetics and Practice of Low Angle Ion Beam Thinning* in MRS Proc. **254** 3–22. An early advocate of using low-energy ion beams.

Cullis, AG, Chew, NG and Hutchinson, JL 1985 *Formation and Elimination of Surface Ion Milling Defects in Cadmium Telluride, Zinc Sulphide and Zinc Selenide* Ultramicroscopy **17** 203–211. The paper on reactive-ion milling of TEM samples.

Giannuzzi, LA and Stevie, FA 2004 *Introduction to Focused Ion Beams: Instrumentation, Theory, Techniques and Practice* Springer Verlag NY. The guide to FIB.

Harriott, LR 1991 *The Technology of Finely Focused Ion Beams* Nucl. Instr. Meth. Phys. Res. Section B: Beam Interactions with Materials and Atoms **55B** (1–4) 802–810. One place to start if you're interested in 'Why Ga?'

Medard, L, Jacquet, PA and Sartorius, R 1949 *Sur les dangers d'explosion des bains aceto-perchloriques de polissage electrolytique (Explosion Hazard of Acetic and Perchloric Acid Mixture Used as Solution in Electrolytic Polishing)* Rev. Metall. **46**(8) 549–560. Jacquet has many other papers on specimen preparation if your French is good.

OTHER MATERIALS

Carter, CB and Norton, MG 2007 give some illustrations of the tripod polisher used to facet diamond in *Ceramic Materials: Science and Engineering* Springer Verlag NY.

Malis, TF 1989 *AEM Specimens: Staying One Step Ahead* in Microbeam Analysis-1989 487–490 Ed. P.E. Russell, San Francisco Press San Francisco gives a discussion of using the microtome for hard samples.

Sawyer, LC, Grubb, DT and Meyers, GF 2008 *Polymer Microscopy* 3rd Ed. Springer Verlag NY gives the details for preparing polymer specimens.

THE COMPANION TEXT

A complete chapter in the companion text not only discusses more specialized techniques but also extends this chapter's discussion.

SELF-ASSESSMENT QUESTIONS

Q10.1 Name two ways you can damage yourself more than your specimen while preparing it.

Q10.2 What are the differences between self-supporting specimens and specimens resting on a grid or thin washer?

Q10.3 What is the main problem associated with using a grid?

Q10.4 What does dimpling mean?

Q10.5 What should we keep in mind when using the tripod polisher to pre-thin the specimen?

Q10.6 What is the difference in the use between electropolishing and ion milling?

Q10.7 Why do we need to use a low angle of incidence in ion milling?

Q10.8 Why do the specimens need to be rotated and cooled during the ion-thinning process?

Q10.9 Why are various chemical etching techniques used in TEM specimen preparation?

Q10.10 Why do you want the region of interest of your specimen to be in the center of the grid?

Q10.11 What methods can be used for preparing thin specimens (initial thinning) of ductile materials?

Q10.12 How is the thickness of your specimen affected by the damage done to the surface?

Q10.13 Why might you need to cool the electrolyte during electropolishing?
Q10.14 What are the advantages and disadvantages of ultramicrotomy?
Q10.15 When is it reasonable to use a mechanical punch?
Q10.16 As a 'rule of thumb,' how much damage does a typical abrasive cause to a specimen?
Q10.17 How would you go about creating a GaAs cross-sectional TEM specimen? If the material is heat sensitive (i.e., it reacts) what can be done to minimize the heat load?
Q10.18 If specimen preparation is so important, why do we spend relatively little time discussing this issue in the book, in class, and in our training?
Q10.19 List the four most widely used polishing solvents.
Q10.20 List at least three precautionary measures you should bear in mind while using hazardous chemicals.
Q10.21 What precautions are necessary for short-term storing of specimen?

TEXT-SPECIFIC QUESTIONS

T10.1 When and why would you take the time to use the window method (Figure 10.14) to prepare thin foils of a metal sheet rather than dimpling and jet polishing (Figure 10.8)?
T10.2 What TEM technique would benefit most from ultramicrotomed specimens and which technique would benefit least?
T10.3 List five different particle samples that might be easily prepared for viewing in a TEM. Give a literature (research journal) reference for each one.
T10.4 Nanotubes, nanowires, buckyballs, etc., are nano-scale specimens that have recently received significant attention. How would you prepare such specimens for examination in the TEM?
T10.5 If you had a large piece of a weld from which you wanted to prepare a thin specimen of a very specific region (e.g., the heat-affected zone or a single-pass region) for chemical analysis, how would you go about preparing such a specific thin specimen?
T10.6 List the advantages and disadvantages of making a replica of a specimen surface as in Figure 10.16 to view in the TEM rather than viewing the surface directly.
T10.7 List the advantages and disadvantages of making an extraction replica of particles from a specimen surface as in Figure 10.17 to view in the TEM rather than viewing the particles within a thin foil.
T10.8 What kinds of difficulties might arise when attempting to perform elemental analysis in the AEM when your specimen has been (a) electropolished or (b) ion milled? How might those difficulties be overcome?
T10.9 Why might ultramicrotomy be the ideal specimen preparation method for AEM but totally unsuitable for routine TEM imaging and diffraction?
T10.10 Why should your specimens remain isolated from you and other human beings after they are thinned?
T10.11 Use Figure 10.26 where appropriate to propose a method for getting a TEM specimen from
A. the join in a soldered copper alloy
B. catalyst particles on a substrate
C. a specific junction in a semiconductor device