Science of Microscopy

Science of Microscopy

Edited by

Peter W. Hawkes
John C.H. Spence

Volume I

Peter W. Hawkes
CEMES CNRS
Toulouse
France

John C.H. Spence
Department of Physics
Arizona State University
Tempe, AZ
and
Lawrence Berkeley Laboratory
Berkeley, CA
USA

Preface

In these two volumes, we have asked many of the leaders in the field of modern microscopy to summarize the latest approaches to the imaging of atoms or molecular structures, and, more especially, the way in which this aids our understanding of atomic processes and interactions in the organic and inorganic worlds.

Man's curiosity to examine the nanoworld is as at least as old as the Greeks. Aristophanes, in a fourth-century BC play, refers to a burning glass; the Roman rhetorician Seneca describes hollow spheres of glass filled with water being used as magnifiers, while Marco Polo in the thirteenth century remarks on the Chinese habit of wearing spectacles. Throughout this time it would have been common knowledge that a drop of water over a particle on glass will provide a magnified image, while a droplet within a small hole does even better as a biconvex lens. By the sixteenth century magnifying glasses were common in Europe, but it was Anthony van Leeuwenhoek (1632–1723) who first succeeded in grinding lenses accurately enough to produce a better image with his single-lens instrument than with the primitive compound microscopes then available. His 112 papers, published in *Philosophical Transactions of the Royal Society*, brought the microworld to the general scientific community for the first time, covering everything from sperm to the internal structure of the flea. Robert Hooke (1635–1703) developed the compound microscope, publishing his results in careful drawings of what he saw in his *Micrographia* (1665). The copy of this book in the University of Bristol library shows remarkable sketches of faceted crystallites, below which he has drawn piles of cannon balls, whose faces make corresponding angles. This strongly suggests that Hooke believed that matter consists of atoms and had made this discovery long before its official rediscovery by the first modern chemists, notably Dalton in 1803. (Greeks such as Leucippus (450 BC) had long before convinced themselves that a stone, cut repeatedly, would eventually lead to "a smallest fragment" or fundamental particle; Democritus once said that "nothing exists except atoms and empty space. All else is opinion" (!))

This atomic hypothesis itself has a fascinating history, and is intimately connected with the history of microscopy. It was Brown's observation in 1827 of the motion of pollen in water by optical micro-

scopy which laid the basis for the modern theory of matter based on atoms. As late as 1900 many chemists and physicists did not believe in atoms, despite the many independent estimates that could be made of their size. These were summarized by Kelvin and Tait in an appendix to their *Treatise on Natural Philosophy*, together with an erroneous and rather superficial estimate of the age of the earth, to be used against Darwin. (This text was the standard English language physics text of the late nineteenth century, despite its failure to cover much of Maxwell's work.) Einstein's 1905 theory of Brownian motion, and Perrin's (1909) more accurate repetition of Brown's experiment, using microscope observations to estimate Avogadro's number, finally settled the matter regarding the existence of atoms. Einstein does not reference Brown's paper, but indicates that he had been told about it. But as Archie Howie has commented, it is interesting to speculate how different the history of science would be if Maxwell had read the Brown paper and applied his early statistical mechanics to it. By the time of Perrin's paper, Bohr, Thomson, Rutherford and others were well committed to atomic and even subatomic physics.

In biology, the optical microscope remained an indispensable tool from van Leeuwenhoek's time with many incremental improvements, able to identify bacteria and their role in disease, but not viruses, which were first seen with the transmission electron microscope (TEM) in 1938. With Zernike's phase contrast theory in the thirties a major step forward was taken, but the really dramatic and spectacular modern advances had to await the widespread use of the TEM, the invention of the laser and the CCD detector, the introduction of the scanning mode, computer control and data acquisition, and the production of fluorescent proteins.

The importance of this early history should not be underestimated—in the words of Feynman "If in some cataclysm, all scientific knowledge were to be destroyed and only one sentence passed on to the next generation of creatures, what statement would contain the most information in the fewest words? I believe it is the atomic hypothesis—that all things are made of atoms."

Images of individual atoms were first provided by Muller's field-ion microscope in the early 1950s, soon to be followed by Albert Crewe's Scanning Transmission Electron Microscope (STEM) images of heavy atoms on thin-film surfaces in 1970. With its subångström resolution, the modern transmission electron microscopes (TEM) can now routinely image atomic columns in thin crystals. For favorable surface structures, the scanning tunneling microscope has provided us with images of individual surface atoms since its invention in 1982, and resulted in a rich spin-off of related techniques.

Probes of condensed and biological matter must possess a long lifetime if they are to be used as free-particle beams. For the most part this has limited investigators to the use of light, X-rays, neutrons and electrons. The major techniques can then often be classified as imaging, diffraction, and spectroscopy. These may be used in both the transmission and reflection geometries, giving bulk and surface information respectively. Chapter 8 (Bauer) reviews both the low-energy electron

microscope (LEEM) and spin-polarized LEEM methods which, using reflected electrons, have recently revolutionized surface science and thin-film magnetism. Here the high cross-section allows movies to made of surface processes at submicrometer resolution, while Auger electron spectroscopy is conveniently incorporated. Chapter 9 (Feng and Scholl) deals with the closely related photoelectron microscopy, where a LEEM instrument is used to image the photoelectrons excited by a synchrotron beam. Here the superb energy selectivity of optical excitation can be used to great advantage. Chapter 3 (Reichelt) describes advances in scanning electron microscope (SEM) research, where the lower-energy secondary electrons provide images with large depth of focus for the most versatile of all electron-optical instruments. The numerous modes of operation include X-ray analysis, cathodolumines-cence, low-voltage modes for insulators and the controlled-atmosphere environmental SEM (ESEM). Turning now to the transmission geometry, we review the latest work in atomic-resolution TEM in Chapter 1 (Kirkland, Chang and Hutchison), the technique which has transformed our understanding of defect processes in the bulk of solids such as oxides, and the STEM mode in Chapter 2 (Nellist). STEM provides an additional powerful analytical capability, which, like the STM, can provide spectroscopy with atomic-scale spatial resolution. An entire chapter (Chapter 4, Botton) is then devoted to analytical TEM (AEM), with a detailed analysis of the physics and performance of its two main detectors, for characteristic X-ray emission and energy-loss spectroscopy. The remarkable recent achievements of *in-situ* TEM are surveyed in Chapter 6 (Ross), including transmission imaging of liquid cell electrolysis, observations of the earliest stages of crystal and nanotube growth, phase transitions and catalysts, superconductors, magnetic and ferroelectric domains and plastic deformation in thin films, all at nanometer resolution or better. Again, the large scattering cross-section of electron probes provides plenty of signal even from individual atoms, so that movies can be made. Chapter 5 (King, Armstrong, Bostanjoglo and Reed) summarizes the dramatic recent revival of time-resolved electron microscope imaging, which uses laser-pulses to excite processes in a sample. The excited state may then be imaged by passing the delayed pulse to the photocathode of the TEM in this "pump-probe" mode. Single-shot transmission electron diffraction patterns have now been obtained using electron pulses as short as a picosecond. Most of these techniques are undergoing a quiet revolution as electron-optical aberration corrector devices are being fitted to microscopes. The dramatic discovery, that, after 60 years of effort, aberration correction is now a reality, was made about ten years ago, and we review the relevant electron-optical theoretical background in Chapter 10 (Hawkes). Finally, in biology, potentially the largest scientific payoff of all is occurring in the field of cryo-electron microscopy, where single-particle images of macromolecules embedded in thin films of ice are imaged, and a three-dimensional reconstruction is made. While the structure of the ribosome and purple membrane protein (among many others) have already been determined in this way, the grand challenge of locating every protein and molecular machine in a single cell remains

to be completed. We summarize this exciting field in Chapter 7 (Plitzko and Baumeister).

Electrons, with the largest cross-section and a source brighter than current generation synchrotrons, provide the strongest signal and hence the best resolution. They do this in a manner that can conveniently be combined with spectroscopy, and we now have aberration-corrected lenses for them. But multiple scattering and inelastic background scattering often complicate interpretation. X-ray imaging of nanostructures, even at synchrotrons, involves much longer data acquisition times but the absence of background and multiple scattering effects greatly improves quantification of data, and thicker samples can be examined. (It is easy to show that the small magnitude of the fine-structure constant will almost certainly never permit imaging of individual atoms using X-rays. We should also recall that in protein crystallography, about 98% of the X-ray beam hits the beam-stop and does not interact with the sample. Of the remaining 2%, 84% is annihilated in production of photoelectrons, and 8% in Compton scattering, while only the remainder produces Bragg diffraction. For light elements the inelastic cross-section for kilovolt electrons is about three times the elastic.) Success with X-rays has thus come mainly through the use of crystallographic redundancy to reduce radiation damage in protein crystallography. However, soft X-ray imaging with zone-plate lenses provides about 30 nm resolution in the "water window" with the advantages of thick samples and an aqueous environment. Applications have also been found in environmental science, materials science and magnetic materials. In addition, the equivalent of the STEM has been developed for soft X-rays: the scanning transmission x-ray microscope (STXM), which uses a zone-plate to focus X-rays onto a sample that can be translated by piezo motors. This arrangement can then provide spatially-resolved X-ray absorption spectroscopy. This work is reviewed in Chapter 13 (Howells, Jacobsen and Warwick).

Both X-ray and electron-beam imaging methods are limited in biology by the radiation damage they create, unlike microscopy with visible light, which also allows observations in the natural state. Optical microscopy is undergoing a revolution, with the development of super-resolution, two-photon, fluorescent labeling and scanning confocal methods. These methods are reviewed in Chapters 11 and 12. Chapter 11 (Diaspro, Schneider, Bianchini, Caorsi, Mazza, Pesce, Testa, Vicidomini, and Usai) discusses two-photon confocal microscopy, in which the spot-scanning mode is adopted, and a symmetrical lens beyond the sample collects light predominantly from the excitation region, thereby eliminating most of the "out-of-focus" background produced in the normal full-field "optical sectioning" mode. Three-dimensional image reconstruction is then possible. Two-photon microscopy combines this with a fluorescence process in which two low-energy incident photons are required to excite a detectable photon emitted at the sum of their energies. This has several advantages, by reducing radiation damage and background, and allowing observation of thicker samples. The method can also be used to initiate photochemical reactions for study. Chapter 12 (Hell and Schönle) describes the

latest super-resolution schemes for optical microscopy, which have now brought the lateral resolution to about 28 nm and, by the symmetrical lens arrangement (4-π confocal), increased resolution measured along the optic axis by a factor of up to seven. The lateral resolution can be improved by modulating the illumination field or by using the stimulated emission depletion microscopy mode (STED), in which saturated excitation of a fluorophor produces nonlinear effects allowing the diffraction barrier to resolution to be broken.

For the scanning near-field probes new possibilities arise. Although restricted to the surface (the site of most chemical activity) and requiring in some cases complex image interpretation, damage is reduced, while the subångström resolution normal to the surface is unparalleled. The method is also conveniently combined with spectroscopy. Early work was challenged by problems of reproducibility and tip artifacts, but Chapters 14–17 in this book show the truly remarkable recent progress in surface science, materials science and biology. Chapter 14 (Nikiforov and Bonnell) describes the various modes of atomic force microscopy which can be used to extract atomic-scale information from the surfaces of modern materials, including oxides and semiconductors. Work-functions can be mapped out (a Kelvin probe with good spatial resolution) and a variety of useful signals obtained by modulation spectroscopy methods. In this way maps of magnetic force, local dopant density, resistivity, contact potential and topography may be obtained. Chapter 15 (Sutter) describes applications of the scanning tunneling microscope (STM) in materials science, including inelastic tunneling, surface structure analysis in surface science, the information on electronic structure which may be extracted, atomic manipulation, quantum size and subsurface effects, and high temperature imaging. Weierstall, in Chapter 16, reviews STM research at low temperatures, including a thorough analysis of instrumental design considerations and applications. These include measurements of local density-of-states oscillations, energy dispersion measurements, electron confinement, lifetime measurements, the Stark and Kondo effects, atomic manipulation, local inelastic tunneling spectroscopy, photon emission, superconductivity and spin-polarized tunneling microscopy. Finally, in Chapter 16, Amrein reviews the special problems that arise when the atomic force microscope (AFM) is applied to the imaging of biomolecules; much practical information on instrumentation and sample preparation is provided, and many striking examples of cell and macromolecule images are shown.

We include two chapters on unconventional "lensless" imaging methods—Chapter 18 (Dunin-Borkowski, Kasama, McCartney and Smith) deals with electron holography and Chapter 19 with diffractive imaging. Gabor's original 1948 proposal for holography was intended to improve the resolution of electron microscopes, and only recently have his plans been realized. Meanwhile, electron holography using Möllenstedt's biprism and the Lorentz mode has proved an extremely powerful method of imaging the magnetic and electrostatic fields within matter. Dramatic examples have included TEM movies of superconducting vortices as temperature and applied fields are varied, and ferroelectric and magnetic domain images, all within thin self-supporting films. Chapter 19 (Spence) describes the recent develop-
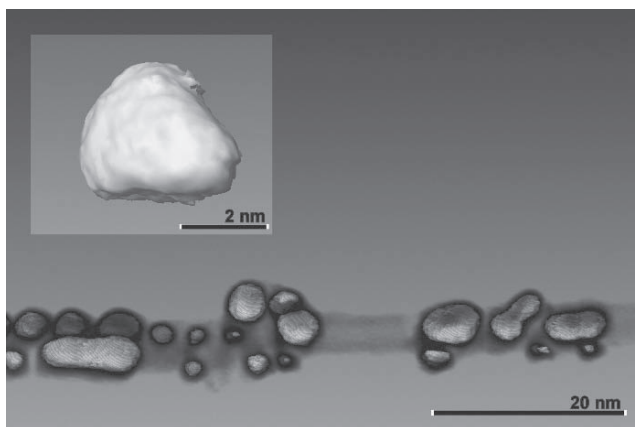
A projection from a three-dimensional image of a carbon nanotube with gold clusters attached. This was reconstructed by taking a series of projected STEM-ADF images at different tilt angles. A faceted gold cluster is shown in the inset. Electron tomography makes it possible to study the three-dimensional nanotube–metal contact geometry which determines the electrical contact resistance to the nanotubes (courtesy of J. Cha, M. Weyland, and D. Muller, 2006).

ment of new iterative solutions to the non-crystallographic phase problem, which now allows diffraction-limited images to be reconstructed from the far-field scattered intensity distribution. This has produced lensless atomic-resolution images of carbon nanotubes (reconstructed from electron microdiffraction patterns) and phase contrast images from both neutron and soft X-ray Fraunhofer diffraction patterns of isolated, non-periodic objects. In this work, lenses are replaced by computers, so that images may now be formed with any radiation for which no lens exists, free of aberrations. Our volumes end with a chapter by van Aert, den Dekker, van Dyck and van den Bos on the definition of resolution in all its forms.

Coverage has been limited to high-resolution methods, with the result that some important new microscopies have been omitted (such as magnetic resonance imaging (MRI), projection X-ray tomography, acoustic imaging etc.). Field-ion microscopy and near-field optical microscopy are also absent. Conversely, although there is no chapter on tomography in materials science, we must mention the rapid progress of these techniques, which has culminated in a remarkable near-atomic reconstruction by J. Cha, M. Weyland and D. Muller of a carbon nanotube to which gold clusters are attached (see figure). For further information on this branch of tomography, see Midgley and Weyland (2003), Midgley (2005), Weyland et al. (2006) and Kawase et al. (2007).

The ingenuity and creativity of the microscopy community as recorded in these pages are remarkable, as is the spectacular nature of the images presented. Neither shows any signs of abating. As in the past, we fully expect major advances in science to continue to result from breakthroughs in the development of new microscopies.

Peter W. Hawkes
John C.H. Spence

References

Kawase, N., Kato, M., Nishioka, H. and Jinnai, H. (2007). Transmission electron microtomography without the "missing wedge" for quantitative structural analysis. *Ultramicroscopy* 107, 8–15.

Midgley, P.A. (2005). Tomography using the transmission electron microscope. In *Handbook of Microscopy for Nanotechnology* (Yao, N. and Wang, Z.L., Eds) 601–627 (Kluwer, Boston).

Midgley, P.A. and Weyland, M. (2003). 3D electron microscopy in the physical sciences: the development of Z-contrast and EFTEM tomography. *Ultramicroscopy* **96**, 413–431.

Weyland, M., Yates, T.J.V., Dunin-Borkowski, R.E., Laffont, L. and Midgley, P.A. (2006). Nanoscale analysis of three-dimensional structures by electron tomography. *Scripta Mater*. **55**, 29–33.

Note on the second printing

Barely a year has passed since we wrote the above Preface but very considerable progress has been made in many forms of microscopy. These are indicated in the additional comments and references at the ends of the chapters. One technique that was not accorded a chapter in the first printing has come of age in 2007. This is scanning ion microscopy. Although attempts to use ions in a scanning instrument go back to the 1960s (e.g., Drummond and Long, 1967; Martin, 1973), it is only very recently that technical progress has resulted in a high-resolution commercial instrument (the ORION helium ion microscope released by Carl Zeiss in 2007). The secret of this advance lies in the atomic-level ion source (ALIS): the tip is precisely shaped with either an atomic triad or a single atom at its apex and operates at liquid-nitrogen temperature. A resolution of 0.7 nm with 45 keV helium ions has been obtained with a prototype; the quoted energy spread is only 0.3 eV and the brightness is of the order of 10^9 A/cm^2 sr. The current is, however, very low, in the femtoampère or picoampère range and thus considerably less than that in a STEM. Earlier scanning ion microscopes used liquid-metal ion sources and their resolution rarely exceeded 10 nm. For the background to this development, see Ishitani and Tsuboi (1997), Sakai et al. (1999), Ishitani and Ohya (2003) and Maclaren et al. (2003) and for recent progress, Ward et al. (2006), Griffin and Joy (2007) and Joy et al. (2007).

References

Drummond, I.W. and Long, J.V.P. (1967). Scanning ion microscopy and ion beam machining. *Nature* **215**, 950–952.

Griffin, B.J. and Joy, D.C. (2007). Imaging with the He scanning ion microscope and with low-voltage SEM – a comparison using carbon nanotube, platinum thin film, cleaved molybdenum disulphide samples and metal standards. *Acta Microscopica* **16**, Suppl. 2, 3–4.

Ishitani, T. and Ohya, K. (2003). Comparison in spatial spreads of secondary electron information between scanning ion and scanning electron microscopy. *Scanning* **25**, 201–205.

Ishitani, T. and Tsuboi, H. (1997). Objective comparison of scanning ion and scanning electron microscope images. *Scanning* **19**, 489–497.

Joy, D.C., Griffin, B.J., Notte, J. and Fenner, C. (2007). Device metrology with high-performance scanning ion beams. *Proc. SPIE* **6518**, to be published.

Maclaren, D.A., Holst, B., Riley, D.J. and Allison, W. (2003). Focusing elements and design considerations for a scanning helium microscope (SheM). *Surface Rev. Lett.* **10**, 249–255.

Martin, F.W. (1973). Is a scanning ion microscope feasible? *Science* **179**, 173–175.

Sakai, Y., Yamada, T., Suzuki, T. and Ichinokawa, T. (1999). Contrast mechanisms of secondary electron images in scanning electron and ion microscopy. *Appl. Surface Sci.* **144**, 96–100.

Ward, B.W., Notte, A. and Economou, N.P. (2006). Helium ion microscope: a new tool for nanoscale microscopy and metrology. *J. Vac. Sci. Technol.* **B24**, 2871–2874.

PWH and JCES

Contents

VOLUME II

Contributors

Matthias Amrein
Microscopy and Imaging Facility, Faculty of Medicine, Department of Cell Biology and Anatomy, University of Calgary, Calgary, Canada

Michael R. Armstrong
University of California Chemistry and Materials Science Directorate, Lawrence Livermore National Laboratory, Livermore, CA, USA

Ernst Bauer
Department of Physics, Arizona State University, Tempe, AZ, USA

Wolfgang Baumeister
Max Planck Institute of Biochemistry, Martinsried, Germany

Paolo Bianchini
Department of Physics, LAMBS-IFOM MicroScoBIO Research Centre, University of Genoa, Genoa, Italy

Dawn A. Bonnell
Department of Materials Science and Engineering, Nano-Bio Interface Center, University of Pennsylvania, Philadelphia, PA, USA

Oleg Bostanjoglo
Optisches Institut, Sekr. P1-1, Technische Universität Berlin, Berlin, Germany

Gianluigi Botton
Department of Materials Science and Engineering, McMaster University, Hamilton, Canada

Valentina Caorsi
Department of Physics, LAMBS-IFOM MicroScoBIO Research Centre, University of Genoa, Genoa, Italy

Shery L.-Y. Chang
Department of Materials, University of Oxford, Oxford, UK

Arnold J. den Dekker
Delft Centre for Systems and Control, Delft University of Technology,
Delft, The Netherlands

Alberto Diaspro
Department of Physics, LAMBS-IFOM MicroScoBIO Research Centre,
University of Genoa, Genoa, Italy

Rafal E. Dunin-Borkowski
Department of Materials Science and Metallurgy, University of
Cambridge, Cambridge, UK; also at Centre for Electron Nanoscopy,
Technical University of Denmark, Kongens Lyngby, Denmark

Jun Feng
Lawrence Berkeley National Laboratory, Berkeley, CA, USA

Peter W. Hawkes
CEMES-CNRS, Toulouse, France

Stefan W. Hell
Department for NanoBiophotonics, Max Planck Insitute of
Biophysical Chemistry, Göttingen, Germany

Malcolm Howells
Advanced Light Source, Lawrence Livermore National Laboratory,
Livermore, CA, USA

John L. Hutchison
Department of Materials, University of Oxford, Oxford, UK

Christopher Jacobsen
Department of Physics and Astronomy, Stony Brook University, Stony
Brook, NY, USA

Takeshi Kasama
Frontier Research System, Institute of Physical and Chemical
Research, Hatoyama, Japan, and Department of Materials
Science and Metallurgy, University of Cambridge, Cambridge,
UK

Wayne King
Chemistry and Materials Science Directorate, Lawrence Livermore
National Laboratory, Livermore, CA, USA

Angus I. Kirkland
Department of Materials, University of Oxford, Oxford, UK

Davide Mazza
Department of Physics, LAMBS-IFOM MicroScoBIO Research Centre,
University of Genoa, Genoa, Italy

Martha R. McCartney
Department of Physics and Astronomy and Center for Solid-State
Science, Arizona State University, Tempe, AZ, USA

Peter D. Nellist
Department of Physics, Trinity College, Dublin, Ireland; Department of
Materials, University of Oxford, Oxford, UK

Maxim P. Nikiforov
Nano-Bio Interface Center, University of Pennsylvania, Philadelphia,
PA, USA

Mattia Pesce
Department of Physics, LAMBS-IFOM MicroScoBIO Research Centre,
University of Genoa, Genoa, Italy

Juergen Plitzko
Max Planck Institute of Biochemistry, Martinsried, Germany

Bryan W. Reed
Chemistry and Materials Science Directorate, Lawrence Livermore
National Laboratory, Livermore, CA, USA

Rudolf Reichelt
Institut für Medizinische Physik und Biophysik, Westfälische
Wilhelms-Universität, Münster, Germany

Frances M. Ross
IBM Research Division T. J. Watson Research Center, Yorktown
Heights, NY, USA

Marc Schneider
Biopharmaceutics and Pharmaceutical Technology, Campus Saarbrücken,
Germany

Andreas Scholl
Lawrence Berkeley National Laboratory, Berkeley, CA, USA

Andreas Schönle
Department of NanoBiophotonics, Max Planck Institute of
Biophysical Chemistry, Göttingen, Germany

David J. Smith
Department of Physics and Astronomy and Center for Solid-State Science, Arizona State University, Tempe, AZ, USA

John C.H. Spence
Department of Physics, Arizona State University Tempe, AZ, *and* Lawrence Berkeley Laboratory, Berkeley, CA, USA

Peter Sutter
Center for Functional Nanomaterials, Brookhaven National Laboratory, Upton, NY, USA

Ilaria Testa
Department of Physics, LAMBS-IFOM MicroScoBIO Research Centre, University of Genoa, Genoa, Italy

Cesare Usai
National Research Council Institute of Biophysics, Genoa, Italy

Sandra Van Aert
University of Antwerp, Antwerp, Belgium

A. Van den Bos
Faculty of Applied Sciences, Delft University of Technology, Delft, The Netherlands

D. Van Dyck
University of Antwerp, Antwerp, Belgium

Giuseppe Vicidomini
Department of Physics, LAMBS-IFOM MicroScoBIO Research Centre, University of Genoa, Genoa, Italy

Tony Warwick
Advanced Light Source, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

Uwe Weierstall
Department of Physics, Arizona State University, Tempe, AZ, USA

Part I

IMAGING WITH ELECTRONS

Atomic Resolution Transmission Electron Microscopy

Angus I. Kirkland, Shery L.-Y. Chang and John L. Hutchison

1 Introduction and Historical Context

1.1 Introduction

High-Resolution Transmission Electron Microscopy (HRTEM) uses a self-supporting thin sample (typically tens of nanometes) illuminated by a highly collimated kilovolt electron beam. A series of magnetic electron lenses image the electron wavefield across the exit face of the sample onto a detector at high magnification. HRTEM has evolved from initial instrumentation constructed by Knoll and Ruska (1932a–c) to its current state where individual atom columns in a wide range of materials can be routinely imaged (Smith, 1997; Krakow et al., 1984) using sophisticated computer-controlled microscopes (Figure 1–1). For this reason HRTEM now occupies a central place in many characterization laboratories worldwide and has made a substantial contribution to key areas of materials science, physics, and chemistry [for key examples showing its wide ranging influence see the frontispiece in the book by Spence (2002)]. Instrument development for HRTEM also supports a substantial commercial industry of manufacturers (Hall, 1966; Hawkes, 1985; Fujita, 1986).[1]

Numerous HRTEM studies of bulk semiconductors (Smith and Lu, 1991; Smith et al., 1989), defects (Figure 1–2) (Olsen and Spence, 1981), and interface structures (Figure 1–3), (Bourret et al., 1988; Gutekunst et al., 1994) in these materials, of metals and alloys (Penisson et al., 1988; Krakow, 1990; Ishida et al., 1983; Amelinckx et al., 1993; Thomas, 1962), and of ceramics, particularly oxides (Lundberg et al., 1989; Buseck et al., 1989), have been reported in a vast literature spanning four

---

[1] We note that HRTEM has also made substantial contributions to structural biology (see, Burge, 1973; Unwin and Henderson, 1975; Henderson, 1995; Koehler, 1973, 1978, 1986 for reviews of selected representative examples from this field; see also Chapter 7 by Plitzko and Baumeister in this volume). However, due to limitations of space we will not consider this aspect further.
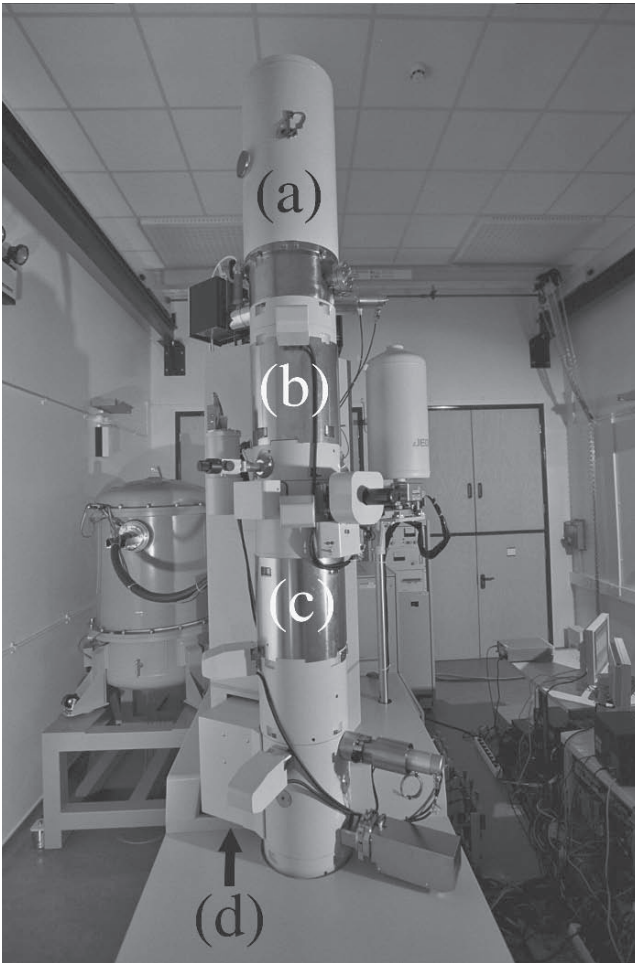
Figure 1–1. A modern 200-kV HRTEM fitted with a (a) field emission gun, (b) probe and (c) image forming aberration correctors, and (d) an omega geometry energy filter.
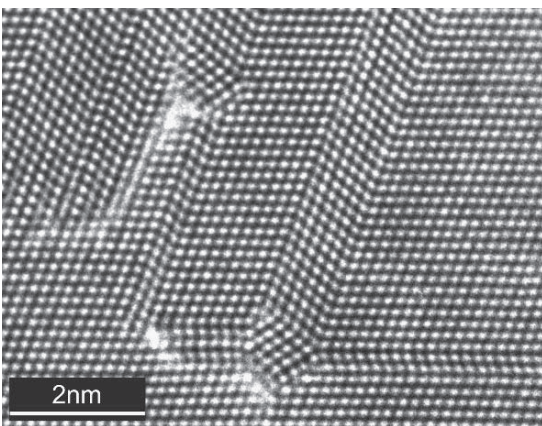

Figure 1–2. HRTEM image of a [110] oriented CVD deposited diamond film showing twins, stacking faults, and nanograins. Note the local disorder at the intersection of the stacking faults and twins.

Figure 1–3. (a) HRTEM image of a lattice matched heterojunction between InP and (Ga, In)As. At the defocus used and for this particular foil thickness differences in scattering between these two isostructural materials allows them to be distinguished. (b) HRTEM image of a heterojunction between InAs and (In, As)Sb that have a significant lattice mismatch. In this case the lattice misfit is accommodated as a regular array of Lomer dislocations marked.
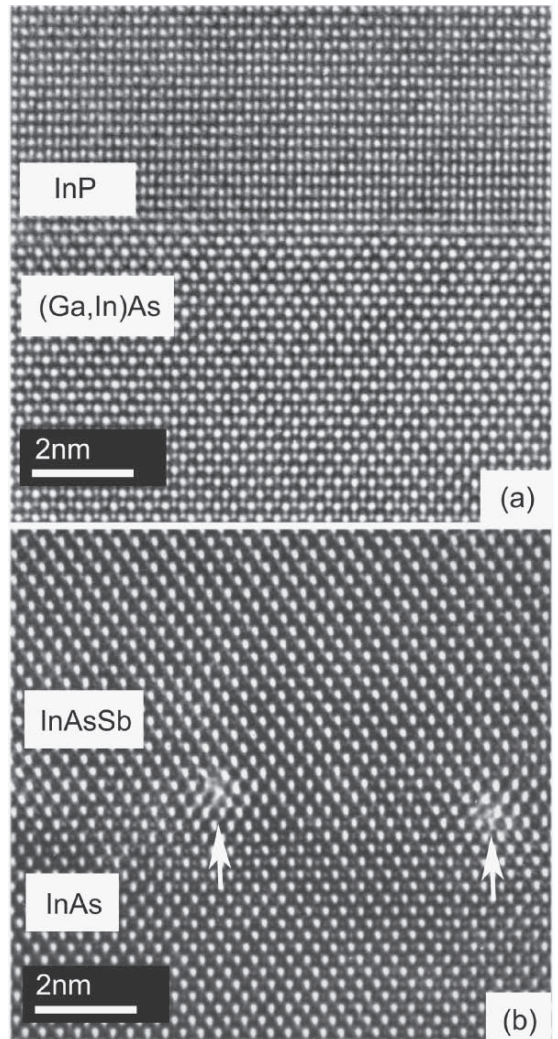
decades (for additional general reviews see Buseck et al., 1989; Smith, 1997; van Tendeloo, 1998; Spence, 1999). An excellent collection of representative HRTEM images can be found in Shindo and Hiraga (1998). More recently HRTEM has become an essential tool in the characterization and discovery of nanoscale materials (Figure 1–4) (see, for example, Iijima, 1991; Yao and Wang, 2005). Of crucial importance in all of these areas is the ability of HRTEM to provide real-space images of the atomic configuration at localized structural irregularities and defects in materials, that are inaccessible to broad-beam bulk diffraction methods and that largely control their properties.

Advances in instrumentation for HRTEM over the same timescale have enabled this information to be recorded with increasing resolution and precision leading to improvements in the quantification of the data obtained.
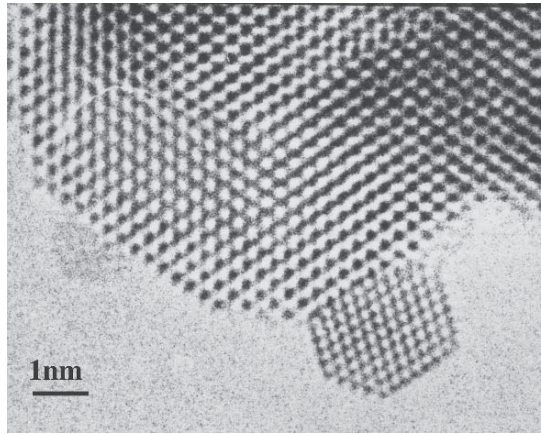
Figure 1–4. HRTEM image of a nanocrystalline gold particle supported on a {111} cerium oxide surface. The gold particle shows an almost perfect cubeoctahedral morphology and both particle and substrate are in an epitaxial [110] orientation despite the large lattice mismatch.

This chapter concentrates on HRTEM at atomic resolution. Following a brief historical overview of the development of HRTEM (for a more detailed article outlining key events in the history of electron microscopy see Haguenau et al., 2003) we begin by outlining some of the theory pertinent to image formation at high resolution and the effects on recorded images of the aberrations introduced by imperfect objective optics. We also provide various definitions of resolution.

The second section surveys the key instrumental components affecting HRTEM and provides an outline of currently available solutions. The final section describes computational approaches to both the recovery of the specimen exit-plane wavefunction (coherent detection) from a series of images and methods available for HRTEM image simulation.

1.2 Historical Summary

Historically, the resolving power of the electron microscope rapidly matched and then exceeded that of the optical microscope in 1934 (Ruska, 1934). However, further improvements in resolution proved relatively slow due to the need to identify and overcome various instrumental limitations (see later).

The first subnanometer lattice-fringe images were obtained in the 1950s (Menter, 1956) from phthalocyanine crystals and this was later extended to lattice images of metal foils showing crossed fringe patterns (Komoda, 1966).

Concurrently, the first published HRTEM images of complex transition metal oxides provided preliminary evidence that many of these (specifically those of Mo, W, Ti, and V) were not perfect structures (Allpress et al., 1969; see also Buseck et al., 1989, for a review). These observations of planar faults in oxides possibly represent the first occa-

sion in which useful information at the atomic level was provided by HRTEM. This work created much interest among solid-state chemists, who for the first time saw a new scientific tool that would enable them to overcome the barriers to structural determinations of these materials imposed by their large unit cells and often extensive disorder. It also immediately provided an explanation for nonstoichiometry in these materials and entirely changed the way in which thermodynamic properties of oxides were modeled.

The work summarized above was possible with the typical instrumental resolutions available in most laboratories at that time. However, it was not until this improved that it became possible to resolve individual cation columns in these and certain other classes of material. In the 1970s the first images showing the component octahedra were published (Cowley and Iijima, 1972) with a resolution of 0.3 nm for a series of mixed Ti–Nb structures that demonstrated a direct correspondence between the lattice image and the projected crystal structure.

The typical spatial resolution (slightly better than ca. 0.5 nm) provided by most commercial instruments in the 1960s and 1970s was largely limited by mechanical and electrical instabilities. Subsequent improvements in instrument design and construction led to a generation of microscopes becoming available in the mid 1970s with point resolutions of less than 0.3 nm operating at intermediate voltages around 200 kV (Uyeda et al., 1972). Toward the end of this period the dedicated 600 kV Cambridge HREM (Cosslett et al., 1979) and several other high-voltage instruments also became operational (Hirabayashi et al., 1982), providing a resolution somewhat better than 0.2 nm.

The following two decades saw further significant improvements in microscope design with dedicated high-resolution instruments being produced by several manufacturers. One outcome of these developments was the installation of commercial high-voltage HRTEMs (operating at ca. 1 MV) in several laboratories worldwide (Gronsky and Thomas, 1983; Matsui et al., 1991). These machines were capable of point resolutions of ca. 0.12 nm, significantly higher than that available in intermediate voltage instruments. Concurrently, commercial instrument development also started to concentrate on improved intermediate voltage instrumentation (at up to 400 kV) (Hutchison et al., 1986) with interpretable resolutions between 0.2 and 0.15 nm.

In the 1990s further progress was made in improving resolution through a combination of key instrumental and theoretical developments. For the former the successful design and construction of improved high-voltage instrumentation (Phillip et al., 1994; Allen and Dorignac, 1998) demonstrated interpretable resolutions close to the long sought after goal of 0.1 nm. Perhaps more significantly, field emission sources became widely available on intermediate voltage microscopes (Honda et al., 1994; Otten and Coene, 1993) improving the absolute information limits of these machines to values close to the point resolutions achievable at high voltage.

This new generation of instruments also led to renewed theoretical and computational efforts aimed at reconstructing the complex specimen exit wavefunction using either electron holograms (Lichte, 1991;

Orchowski et al., 1995) or extended focal or tilt series of HRTEM images (see later). These latter "indirect" approaches extended the interpretable resolution beyond conventional axial image limits and provided both the phase and modulus of the specimen exit wavefunction, free from effects due to the objective lens rather than the aberrated intensity available in a conventional HRTEM image.

The latest instrumental developments have seen the successful completion and testing of the necessary electron optical components for direct correction of the spherical aberration present in all electromagnetic round lenses and these corrected instruments are now capable of directly interpretable resolutions close to or below 0.1 nm at intermediate voltages.

2 Essential Theory

In this section we outline some of the essential theory required for understanding HRTEM image contrast. Many of the topics described here are treated in more detail elsewhere (see, for example, Spence, 2002; Reimer, 1984, 1997; Hawkes and Kasper, 1996; Ernst and Rühle, 2003) (for the latter see Chapter 2 in particular) and only selected frameworks directly relevant to HRTEM imaging using modern instrumentation are discussed further.

We begin with a general review of the essentials of the HRTEM image formation process and subsequently treat the key areas of resolution, and the effects of the objective lens and source in more detail.

2.1 Image Formation

As shown schematically in Figure 1–5 (from a simplified ray optical perspective and from a wave optical perspective as described subsequently) the overall process in the formation of an HRTEM image involves three steps.

1. Electron scattering in the specimen.
2. Formation of a diffraction pattern in the back focal plane of the objective lens.
3. Formation of an image in the image plane.[2]

To understand the relationship between contrast recorded in an HRTEM image and the atomic arrangement within the specimen it is essential to develop theoretical frameworks describing each of these steps.

To describe the general case (of arbitrary specimens) each of the above steps requires a complex mathematical and computational treatment that is outside the scope of the section (comprehensive accounts can be found elsewhere, e.g., Spence, 2002; Buseck et al., 1989; Cowley,

---

[2] Although not formally derived here it can easily be shown that the specimen, back focal, and image planes are mathematically related by Fourier transform operations.
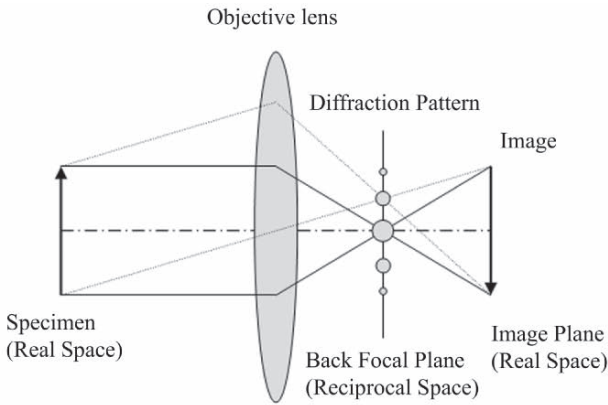
(a)
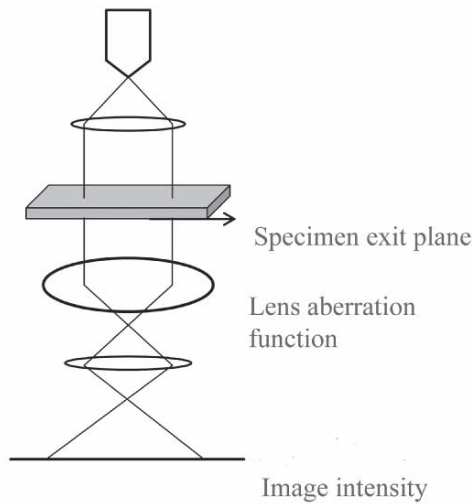
(b)

Figure 1–5. (a) Schematic optical ray diagram showing the principles of the imaging process in HRTEM and indicating the reciprocal relationships between specimen, diffraction and image spaces. (b) Schematic diagram illustrating the wave optical relationship between the recorded image intensity in HRTEM and the specimen exit-wave through the objective lens aberration function.

1975; Reimer, 1997; Ernst and Rühle, 2003; Hawkes and Kasper, 1996). We therefore restrict ourselves to treatment of the simplest cases for illustrative purposes and subsequently follow the nomenclature and sign conventions given in Spence (2002).

For thin specimens, neglecting absorption, the effect of the specimen on an incident electron wave is to alter only its phase leaving the amplitude unchanged. Under this *phase object approximation* (POA), which ignores Fresnel diffraction within the specimen but includes the

effects of multiple scattering, the specimen exit-wave complex amplitude can be written as

$$\psi_e(x, y) = \exp\{-i\sigma\phi_p(x, y)\} \tag{1}$$

where σ is an interaction constant given by

$$\sigma = 2\pi m e \lambda_r / h^2 \tag{2}$$

in which both m and λ_r are relativistically corrected values of the electron mass and wavelength and $\phi_p = \int_{-t/2}^{t/2} \phi(x, y, z)dz$ is the two-dimensional projection of the specimen potential along the beam direction. The interaction constant decreases with accelerating voltage (with values of 0.00729 $V^{-1} nm^{-1}$ at 200 kV decreasing to 0.00539 $V^{-1} nm^{-1}$ at 1000 kV), whereas the specimen inner potential generally increases with atomic number, although this also depends on the density (Shindo and Hiraga, 1998) (Table 1–1).

Equation (1) shows that within this model the effect of the specimen is to advance the phase of the electron wave by $\sigma\phi_p(x, y)$ over the wave in vacuum. For suitably thin specimens of low atomic number the values of the mean inner potential are such that this phase advance is small and hence the exponential term in Eq. (1) can be expanded and approximated as[3]

$$\psi_e(x, y) \approx 1 - i\sigma\phi_p(x, y) \tag{3}$$

in the *weak phase object approximation* (WPOA), which assumes kinematic scattering within the specimen requiring that the intensity of the central unscatttered beam is significantly stronger than that of the diffracted beams.

It is important to note that both of the above formulations are projection approximations such that atoms within the specimen can be moved along the incident beam direction without affecting the exit wavefunction.

Table 1–1. Mean inner potential of representative materials in volts.

Element	Z (atomic number)	Mean inner potential (V)
C	6	7.8 ± 0.6
Al	13	13 ± 0.4
Si	14	11.5
Cu	29	23.5 ± 0.6
Ge	32	15.6 ± 0.8
Au	79	21.1 ± 2

---

[3] Strictly it is the variation in the phase change produced by different parts of the specimen that is important, which supports this approximation.

The complex amplitude of the scattered wave in the back focal plane of the objective lens is given by the Fourier transform of Eq. (3). With $\phi_p(x, y)$ real this gives

$$\psi_d(u, v) = \delta(u, v) - i\sigma F\{\phi_p(x, y)\} \tag{4}$$

Equation (4) is subsequently modified by the presence of a limiting objective aperture and by phase shifts introduced by the objective lens.

The former can be included through the simple function

$$P(\boldsymbol{k}_x, \boldsymbol{k}_y) = 1 \quad |\boldsymbol{k}| \leq r \,.$$
$$P(\boldsymbol{k}_x, \boldsymbol{k}_y) = 0 \quad |\boldsymbol{k}| \geq r \,. \tag{5}$$

The phase shifts introduced by the objective lens are parameterized by the coefficients of a wave aberration function, $W(u, v)$, which is treated in detail in a subsequent section.

Thus, including aperture and lens effects the complex amplitude, under the WPOA, is given by

$$\psi'_d(\boldsymbol{k}_x, \boldsymbol{k}_y) = \delta(\boldsymbol{k}_x, \boldsymbol{k}_y) - i\sigma F\{\phi_p(x, y)\} P(\boldsymbol{k}_x, \boldsymbol{k}_y) \exp[iW(\boldsymbol{k}_x, \boldsymbol{k}_y)]. \tag{6}$$

A further Fourier transform of Eq. (6) finally gives the image amplitude (in the image plane) as

$$\psi_i(x, y) = 1 - i\sigma \{\phi_p(-x, y)\} F\{P(\boldsymbol{k}_x, \boldsymbol{k}_y) \exp[iW(\boldsymbol{k}_x, \boldsymbol{k}_y)]\}. \tag{7}$$

Since the cosine terms in the expansion of equation (7) cancel, the recorded image intensity, to first order, is

$$I(x, y) = \psi_i(x, y)\psi_i^*(x, y)$$
$$\approx 1 + 2\sigma\phi_p(-x, -y) * \mathcal{F}\{\sin[W(\boldsymbol{k}_x, \boldsymbol{k}_y)]P(\boldsymbol{k}_x, \boldsymbol{k}_y)\}. \tag{8}$$

The above expression shows that for this simplest theory the image contrast is proportional to the projection of the specimen potential convolved with an impulse response function arising from the instrument.

Detailed treatment of the latter requires the inclusion of the effects due to the partial coherence of the electron source, which acts to damp higher spatial frequencies (see later), and of the detector, which also modifies the recorded contrast through its modulation transfer function (see later).

A useful modification to the above treatment makes the potential $\phi_p(x, y)$ complex.

This complex projected specimen potential (Cowley and Pogany, 1968) provides a description of the attenuation of the image wavefeld through either scattering outside a limiting aperture or more usefully for unfiltered HRTEM by the depletion of the elastic wavefeld by inelastic processes (Yoshioka, 1957).

A number of further extensions to this basic treatment have previously been proposed to overcome the limitations of a projection approximation in the thick phase grating approximation (Cowley and Moodie, 1962). We do not give a detailed derivation here but note that this approximation successfully accounts for multiple scattering and a degree of curvature of the Ewald sphere (Fresnel diffraction within the specimen) and is thus more generally applicable to HRTEM imaging under less restrictive conditions than the WPOA.

The *projected charge density* (PCD) approximation is an alternative extension that provides a tractable analytic expression for the image intensity including the effects of multiple scattering (unlike the weak phase object) but retaining the restriction of a projection approximation.

Starting from the expression for the specimen exit-wave complex amplitude given by the POA in the absence of an objective aperture and with no wave aberration function we can write

$$\psi_e(x, y) = \exp\{-i\sigma\phi_p(x, y)\} \tag{9}$$

The amplitude in the back focal plane of the objective lens is given by Fourier transformation of the above as

$$\psi_d(\boldsymbol{k}_x, \boldsymbol{k}_y) = F\{\exp[-i\sigma\phi_p(x, y)]\}\exp[i\pi\Delta C_1\lambda(u^2 + v^2)]$$
$$\approx \Phi(\boldsymbol{k}_x, \boldsymbol{k}_y)[1 + i\pi\Delta C_1\lambda|\boldsymbol{k}|^2]. \tag{10}$$

if only a small defocus, ΔC_1, is allowed and where $\Phi(k_x, k_y)$ represents the Fourier transform of $\exp[i\sigma\phi_p(x, y)]$.

Thus the image amplitude (in the image plane) is given by

$$\psi_i(x, y) = \exp(-i\sigma\phi_p(x, y)) + i\pi\Delta C_1\lambda F^{-1}\{(\boldsymbol{k}_x^2 + \boldsymbol{k}_y^2)\Phi(\boldsymbol{k}_x, \boldsymbol{k}_y)\}. \tag{11}$$

A standard theorem from Fourier analysis is now used (Bracewell, 1965), which states that if $f(x, y)$ and $\Phi(u, v)$ are a Fourier transform pair then

$$F^{-1}\{(\boldsymbol{k}_x^2 + \boldsymbol{k}_y^2)\Phi(\boldsymbol{k}_x, \boldsymbol{k}_y)\} = -\frac{1}{4}\pi^2[\nabla^2 f(x, y)]. \tag{12}$$

Applying this result the image amplitude is given by

$$\psi_i(x, y) = \exp[-i\sigma\phi_p(x, y)] - i\Delta C_1\lambda/4\pi\nabla^2\{\exp[-i\sigma\phi_p(x, y)]\}$$
$$= \exp[-i\sigma\phi_p(x, y)] + (i\Delta C_1\lambda\sigma/4\pi)\exp[-i\sigma\phi_p(x, y)]$$
$$\times\{\sigma\nabla\phi_p(x, y) + i\nabla^2\phi_p(x, y)\} \tag{13}$$

which yields an image intensity to first order as

$$I(x, y) \approx 1 - (\Delta C_1\lambda\sigma/2\pi)\nabla^2\phi_p(x, y) \tag{14}$$

From Poisson's equation $\nabla^2\phi_p(x, y) = -\rho_p(x, y)/\varepsilon_0\varepsilon$ we finally obtain

$$I(x, y) = 1 + (\Delta C_1\lambda\sigma/2\pi\epsilon_0\epsilon)\rho_p(x, y). \tag{15}$$

in which $\rho_p(x, y)$ is the projected total charge density including the nuclear contribution.

Historically, the restriction of limited defocus and no spherical aberration (or other uncorrected aberrations) meant that the application of the PCD approximation was restricted to relatively low resolution imaging. However, this approximation would now seem to be ideal for the interpretation of aberration corrected images (see later) in which these restrictions can be experimentally achieved at high resolution (O'Keefe, 2000). Finally, we note that this theory has also been modified (Lynch et al., 1975; Chang, 2000) to include the effects due to higher order lens aberrations and the presence of a limiting objective aperture.

Further extensions to the models outlined above require solution of the dynamic electron diffraction problem using one of several possible computational algorithms, a description of which is outside the scope of this section (Goodman and Moodie, 1974; Cowley, 1975; Self et al., 1983; Jap and Glaeser, 1978; van Dyck, 1983; van Dyck and Coene, 1984; Coene and van Dyck, 1984a,b; Hirsch et al., 1965; Stadelmann, 1987, 1991; Spence and Zuo, 1992; Chen and van Dyck, 1997; Kirkland, 1998; Ernst and Rühle, 2003). However, for generalized HRTEM image simulation the multislice algorithm (Cowley, 1959a–c, 1975; Goodman and Moodie, 1974) has been most commonly employed for the simulation of HRTEM images, to a large extent due to its computational efficiency compared to alternative methods (van Dyck and Op de Beeck, 1994) such as Bloch wave calculations, and this is therefore outlined in a subsequent section.

2.2 Resolution Limits

Unlike their optical equivalents there is no simple measure of resolution for the electron microscope, as the resolution depends on both the instrument and also on the scattering properties of the sample used.[4]

The ultimate resolution of any optical system is the diffraction limit imposed by the wavelength of the radiation, λ, and the aperture angle of the objective lens, α, and the refractive index, n, which can be formalized through Abbe's equation as[5]

$$r_d = k\lambda/n\sin(\alpha) \qquad (16)$$

However, due to imperfections in the objective lens and limited coherence (as discussed subsequently) experimental resolution limits are far lower than that set by Eq. (16) and hence a "single figure" definition of resolution for HRTEM is not possible.

Two independent definitions of attainable resolution are commonly used, defined, respectively, by the key optical properties of the objective lens and those of the source.

---

[4] For a more detailed treatment of resolution see Chapter 20 by van Aert et al. in this volume.

[5] The value of the constant k lies between 0.6 and 0.8 depending on the coherence of the illumination.

The first of these is the "directly interpretable" or Scherzer limit (Scherzer, 1949) or point resolution, which defines the maximum width of a pass band transferring all spatial frequencies from zero, without phase reversal, and is determined by the coefficients of the wave aberration function of the objective lens (see later).

Ignoring the phase shifts due to higher order aberrations (see later) the *phase contrast transfer function* (PCTF) (Hawkes and Kasper, 1996) due to defocus and spherical aberration is given by

$$\sin W(\boldsymbol{k}) = \sin\{\pi C_1 \lambda |\boldsymbol{k}|^2 + \tfrac{\pi}{2} C_3 \lambda^3 |\boldsymbol{k}|^4\}. \tag{17}$$

For HRTEM this defines a focus setting (Scherzer, 1949) that offsets the phase shift due to spherical aberration, C_3 through a suitable choice of defocus:[6]

$$C_{1,\text{Scherzer}} = -1.2(C_3\lambda)^{1/2} \tag{18}$$

which leads to a broad band of phase contrast transfer without zero crossings (Figure 1–6) up to a frequency of $\boldsymbol{k}_{\max} = 1.6(C_3\lambda^3)^{-1/4}$.

The reciprocal gives the point resolution as

$$d_1 = 0.625(C_3\lambda^3)^{1/4} \tag{19}$$

Thus, HRTEM images of thin specimens recorded at the Scherzer defocus (or its extended variant) will have components that are directly proportional to the (negative of) the *projected potential* of the specimen extending to spatial frequencies equal to the interpretable resolution limit (Cowley and Iijima, 1972; Hanßen, 1971). For higher spatial frequencies up to that defined by the information limit the contrast is partially reversed as the PCTF starts to oscillate (Figure 1–6).[7]

Given the form of Eq. (19) it is evident that a decrease in electron wavelength has a greater effect than an equivalent decrease in the spherical aberration, and for this reason high-voltage instrumentation (see earlier) has, until recently, been the preferred route to achieving higher interpretable resolutions.

The higher resolution, the *information limit*, defines the highest spatial frequency transferred from the specimen exit wavefunction to the image intensity. This is determined by the effects of spatial and temporal coherence (see later) in the illumination and by mechanical instabilities and acoustic noise that also act to damp the transfer of higher spatial frequencies.

---

[6] In this definition of the Scherzer focus, the passband in the CTF contains a local minimum = 0.7. The original definition, $C_{1,\text{Scherzer}} = (C_3\lambda)^{1/2}$, avoids a local minimum in the passband at the cost of a slightly poorer point resolution $\rho_S = 0.707(C_3\lambda^3)^{1/4}$.

[7] The above definitions of point resolution assume a fixed positive C_3. Variable C_3 in corrected microscopes modifies these results as detailed in a subsequent section.