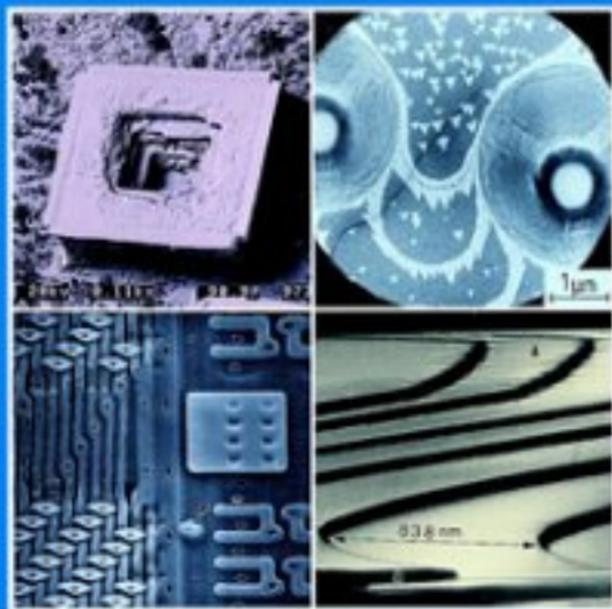


# Electron Microscopy

Principles and Fundamentals

Edited by S. Amelinckx, D. van Dyck,  
J. van Landuyt, G. van Tendeloo



**VCH**  
A Wiley company

S. Amelinckx, D. van Dyck, J. van Landuyt,  
G. van Tendeloo

# **Electron Microscopy**

Principles and Fundamentals



**Handbook of Microscopy**  
**Applications in Materials Science, Solid-State Physics and Chemistry**

Methods I

1997. ISBN 3-527-29280-2.

Methods II

1997. ISBN 3-527-29473-2.

Applications

1997. ISBN 3-527-29293-4.

**Further Reading from VCH**

S. N. Magonov, M.-U. Whangbo

Surface Analysis with STM and AFM

Experimental and Theoretical Aspects of Image Analysis

ISBN 3-527-29313-2

D. A. Bonnell

Scanning Tunnelling Microscopy and Spectroscopy

Theory, Techniques and Applications

ISBN 3-527-27920-2

# **Electron Microscopy**

Principles and Fundamentals

Edited by

S. Amelinckx, D. van Dyck,

J. van Landuyt, G. van Tendeloo



Prof. S. Amelinckx  
Electron Microscopy for  
Materials Science (EMAT)  
University of Antwerp - RUCA  
Groenenborgerlaan 171  
2020 Antwerp  
Belgium

Prof. D. van Dyck  
Electron Microscopy for  
Materials Science (EMAT)  
University of Antwerp - RUCA  
Groenenborgerlaan 171  
2020 Antwerp  
Belgium

Prof. J. van Landuyt  
Electron Microscopy for  
Materials Science (EMAT)  
University of Antwerp - RUCA  
Groenenborgerlaan 171  
2020 Antwerp  
Belgium

Prof. G. van Tendeloo  
Electron Microscopy for  
Materials Science (EMAT)  
University of Antwerp - RUCA  
Groenenborgerlaan 171  
2020 Antwerp  
Belgium

This book was carefully produced. Nevertheless, authors, editors and publisher do not warrant the information contained therein to be free of errors. Readers are advised to keep in mind that statements, data, illustrations, procedural details or other items may inadvertently be inaccurate.

Published by  
VCH Verlagsgesellschaft mbH, Weinheim (Federal Republic of Germany)

Editorial Directors: Dr. Peter Gregory, Dr. Jörn Ritterbusch  
Production Manager: Hans-Jochen Schmitt

Every effort has been made to trace the owners of copyrighted material; however, in some cases this has proved impossible. We take this opportunity to offer our apologies to any copyright holders whose rights we may have unwittingly infringed.

Library of Congress Card No. applied for.

A catalogue record for this book is available from the British Library.

Die Deutsche Bibliothek – CIP-Einheitsaufnahme  
**Electron microscopy** : principles and fundamentals / ed. by S.  
Amelinckx ... - Weinheim : VCH, 1997  
ISBN 3-527-29479-1

© VCH Verlagsgesellschaft mbH, D-69451 Weinheim (Federal Republic of Germany), 1997

Printed on acid-free and chlorine-free paper.

All rights reserved (including those of translation into other languages). No part of this book may be reproduced in any form – by photoprinting, microfilm, or any other means – nor transmitted or translated into a machine-readable language without written permission from the publishers. Registered names, trademarks, etc. used in this book, even when not specifically marked as such, are not to be considered unprotected by law.

Composition: Alden Bookset, England  
Printing: betz-druck, D-64291 Darmstadt  
Bookbinding: W. Osswald, D-67433 Neustadt

# List of Contributors

Amelinckx, Severin (Chap. 1.1.1)  
Electron Microscopy for Materials  
Science(EMAT)  
University of Antwerp – RUCA  
Groenenborgerlaan 171  
2020 Antwerp  
Belgium

Bauer, Ernst (Chaps. 1.6 and 1.9)  
Arizona State University  
Department of Physics and Astronomy  
Tempe, AZ 85287-1504  
USA

Browning, Nigel D. (Chap. 2.3)  
University of Illinois  
Department of Physics  
845 West Taylor Street  
Chicago, IL 60607-7059  
USA

Cherns, David; Steeds, John W.;  
Vincent, Roger (Chap. 1.5)  
H. H. Wills Physics Laboratory  
University of Bristol  
Tyndall Avenue  
Bristol BS8 1TL  
U. K.

Colliex, Christian (Chap. 1.3)  
Lab. de Physique des Solides  
Université Paris-Sud  
Bâtiment 510  
91405 Orsay Cedex  
France

Cowley, John M. (Chaps. 1.2 and 2.2)  
Arizona State University  
Dept. of Physics & Astronomy  
Box 87 15 04  
Tempe, AZ 85287-1504  
USA

Fiermans, Lucien; De Gryse, Roger  
(Chap. 2.4)  
Department of Solid State Sciences  
Surface Physics Division  
University of Gent  
Krijgslaan 281/S1  
9000 Gent  
Belgium

Fujita, Hiroshi (Chap. 1.4)  
Research Center for Science and  
Technology  
Kinki University  
Higashi-Osaka  
Osaka 577  
Japan

Jakubovics, John P. (Chap. 1.7)  
University of Oxford  
Department of Materials  
Parks Road  
Oxford OX1 3PH  
U. K.

Joy, David C. (Chap. 2.1)  
The University of Tennessee  
EM Facility – Programm in Analytical  
Microscopy  
F241 Walters Life Sciences Building  
Knoxville, Tennessee 37996/0810  
USA

Lichte, Hannes (Chap. 1.8)  
Institut für Angewandte Physik  
Technische Universität Dresden  
01062 Dresden  
Germany

Nellist, Peter D. (Chap. 2.3)  
Cavendish Laboratory  
Madingley Road  
Cambridge CB3 0HE  
U.K.

Oleshko\*, Vladimir; Gijbels, Renaat  
(Chap. 2.5)  
Department of Chemistry  
University of Antwerp (UIA)  
2610 Wilrijk-Antwerpen  
Belgium

\*On leave from the Russian Academy  
of Sciences  
N. N. Semenov Institute of Chemical  
Physics  
117421 Moscow  
Russia

Pennycook, Stephen J.; Jesson, David  
E.; Chisholm, Matthew F. (Chap. 2.3)  
Solid State Division  
Oak Ridge National Laboratory  
P.O. Box 2008  
Oak Ridge, TN 37831-6030  
USA

Scheinfein, Mike, R. (Chap. 2.7)  
Arizona State University  
Dept. of Physics and Astronomy  
Tempe, AZ 85287-1504  
USA

Unguris, John; Kelley, Michael H.;  
Gavrin, Andrew; Celotta, Robert J.;  
Pierce, Daniel T. (Chap. 2.7)  
Electron Physics Group  
Nat. Institute of Standards and  
Technology  
Gaithersburg, MD 20899  
USA

Van Dyck, Dirk (Chap. 1.1.2)  
Electron Microscopy for Materials  
Science (EMAT)  
University of Antwerp – RUCA  
Groenenborgerlaan 171  
2020 Antwerp  
Belgium

Van Espen, Pierre; Janssens, Gert  
(Chap. 2.6)  
Micro and Trace Analysis Centre  
Dept. of Chemistry  
University of Antwerpen (UIA)  
Universiteitsplein 1  
2610 Antwerpen  
Belgium

# Contents

## Introduction

### **1 Stationary Beam Methods**

- 1.1 Transmission Electron Microscopy 3
  - 1.1.1 Diffraction Contrast Transmission Electron Microscopy 3  
*S. Amelinckx*
  - 1.1.2 High-Resolution Electron Microscopy 109  
*D. Van Dyck*
- 1.2 Reflection Electron Microscopy 163  
*J. M. Cowley*
- 1.3 Electron Energy-Loss Spectroscopy Imaging 181  
*C. Colliex*
- 1.4 High Voltage Electron Microscopy 203  
*H. Fujita*
- 1.5 Convergent Beam Electron Diffraction 223  
*D. Cherns, J. W. Steeds, R. Vincent*
- 1.6 Low-Energy Electron Microscopy 243  
*E. Bauer*
- 1.7 Lorentz Microscopy 261  
*J. P. Jakubovics*
- 1.8 Electron Holography Methods 271  
*H. Lichte*
- 1.9 Spin-Polarized Low-Energy Electron Microscopy 293  
*E. Bauer*

### **2 Scanning Beam Methods**

- 2.1 Scanning Reflection Electron Microscopy 305  
*D. C. Joy*
- 2.2 Scanning Transmission Electron Microscopy 329  
*J. M. Cowley*
- 2.3 Scanning-Transmission Electron Microscopy: Z Contrast 361  
*S. J. Pennycook, D. E. Jesson, P. D. Nellist, M. F. Chisholm, N. D. Browning*
- 2.4 Scanning Auger Microscopy (SAM) and Imaging X-Ray Photoelectron Microscopy (XPS) 387  
*R. De Gryse, L. Fiermans*

VIII *Contents*

- 2.5 Scanning Microanalysis 427  
*R. Gijbels, V. Oleshko*
- 2.6 Imaging Secondary Ion Mass Spectrometry 457  
*P. van Espen, G. Janssens*
- 2.7 Scanning Electron Microscopy with Polarization Analysis  
(SEMPA) 483  
*J. Unguris, M. H. Kelley, A. Gavrin, R. J. Celotta,  
D. T. Pierce, M. R. Scheinfein*

**Index** 499

# Introduction

Recently VCH published a "Handbook of Microscopy" edited by the undersigned and to which contributed 90 experts, each on their special field of competence on methods and applications of microscopy techniques. Only methods applied or applicable to solids are discussed; they range from the classical optical microscopy to the most recently developed scanning probe methods such as scanning tunnelling microscopy and atomic and magnetic force microscopy.

Each of these techniques allows to solve a number of problems but even for the specialist it is not always evident which technique is best suited to solve a given problem. An answer to this question was given in a separate volume of applications; numerous case studies, classified according to the nature of the solid, illustrate the use of various techniques, focusing the attention on the obtainable information.

Among the numerous methods those making use of an electron beam are the most widely used; they are of particular significance because of their great versatility and their wide range of applications. Moreover a number of these methods can be implemented in a single instrument. It was therefore thought of interest to publish separately the set of contributions referring to electron beam methods. The present book is concerned with the fundamentals only, and refers to part III of the "Handbook of Microscopy" for a full account of the possible applications.

We first introduce the stationary beam methods as opposed to scanning beam methods. Transmission electron microscopic techniques can be applied to a wide range of materials. The main limitations are that it must be possible to prepare a thin foil of the material and that it must be stable against radiation damage by the electron beam. The various forms of electron microscopy have their specific areas of application. Diffraction contrast as used in conventional transmission electron microscopy (CTEM) is the method indicated for the study of defects; it essentially reveals strain fields and orientation differences, for instance due to dislocations, stacking faults, anti-phase boundaries, precipitates, grain boundaries and domain boundaries. It does not reveal the periodic structure of the solid. For this purpose high resolution electron microscopy (HREM) is required; it has found wide application as a method to image complicated crystal structures with a subnanometer resolution. Recent advances in the interpretation and processing of the images have allowed to achieve a resolution of the order of 1 Å. Structure images obtained in this way often allow to propose a structure model which can subsequently be refined by X-ray diffraction. This is especially valuable for materials only available as fine grain powders. Single crystal electron diffraction patterns and images are then used to determine the approximate lattice constants and to produce a preliminary model; both are subsequently refined using the Rietveld method of inter-

preparing X-ray powder diffraction patterns. Recently, computational methods have been developed for quantitative structure refinement from HREM images that can also be applied to crystal defects.

Convergent beam electron diffraction (CBED) and imaging is very sensitive to crystal symmetry and to strain. It allows the determination of the local space group of small crystal areas. It is also a useful tool for the quantitative study of stacking faults and dislocations and it allows the quantitative determination of the lattice potential.

With a high energy electron beam (RHEEM) under grazing incidence, surface features can be revealed such as growth steps, reconstructed surfaces, surface layers with a structure different from that of the bulk, etc.. The method provides a foreshortened image of the surface but this is a minor drawback. It requires excellent vacuum conditions. The interpretation problems are addressed in the text.

Electron energy loss spectroscopy (EELS) has a relatively long tradition as an analytical tool but only more recently have methods become available for the practical application of the method to elemental mapping, making use of characteristic energy losses for different elements. The method constitutes also a sensitive probe for local chemical bonding, for instance along planar interfaces. The recent developments have been possible largely due to the availability of suitable data processing equipment at a reasonable cost. A related feature is the possibility to obtain zero-loss images by allowing only elastically scattered electrons to contribute to the image. Also images produced by electrons which have suffered a well-defined energy loss can similarly be obtained.

One of the intrinsic limitations of transmission electron microscopy is the need to have electron transparent foil specimens. For medium voltage microscopes (100–200 kV) this limits the foil thickness from 50 Å to a few thousand Å, depending on the material and on the applied method of observation. In this thickness range surface effects may significantly alter bulk phenomena such as dislocation movement during plastic deformation, grain growth, etc.. Specimens of larger thickness representative of the bulk can be used in high voltage microscopes (HVEM) due to the larger penetrating power of the electrons. However with increasing voltage also the radiation damage rate increases, which limits the observation time. On the other hand radiation damage can conveniently be studied under controlled conditions using “in-situ” high voltage microscopy. The large size of the specimen chamber of high voltage microscopes makes various “in-situ” studies possible. The specific specimen holders required for such studies are described.

Substantially undeformed images of surface features, such as growth steps, may be obtained by low energy electron microscopy (LEEM) in which the beam is incident normal to the crystal surface and the specularly reflected beam produces the image. The method is very depth sensitive and allows to detect monoatomic steps; it requires a very clean ultrahigh vacuum. The contrast depends sensitively on the acceleration voltage used ( 2–20 V); it does not only reveal surface steps,

but it is also orientation dependent and changes considerably with the surface structure. As a result domains of reconstructed surfaces can be revealed.

Moving charges are influenced by magnetic fields; they are subject to the Lorentz-force. This phenomenon is exploited in transmission electron microscopy to reveal magnetic domain walls along which the magnetisation vector and hence also the Lorentz-force changes its direction. Such walls can be imaged either as dark or as bright lines depending on whether an excess or a deficiency of electrons is caused at the walls. The different imaging modes are discussed in the chapter on Lorentz microscopy and a number of examples illustrate the practical applications.

Already in 1948 Gabor, when developing the foundations of holography, had applications in the field of electron microscopy in mind. However, as yet electron holography has not become a routine technique. Considerable progress has nevertheless been made by applying numerical reconstruction techniques which became feasible thanks to the generalized availability of affordable computer hardware. The contribution included in the book gives an adequate survey of the state of the art and of the perspectives of further development allowing subångström resolution.

Electron beams can easily be manipulated, i.e. deflected, scanned and focused into a very small probe by electromagnetic fields. The best known and most widespread example is undoubtedly the TV set. Benefitting from developments in this highly commercial field a number of scanning beam analytic and imaging methods have been developed. Methods for chemical analysis using initially a stationary beam as a probe were made into imaging methods by scanning the electron beam over the specimen and synchronously scanning another electron beam modulated by the detector signal over a cathode ray screen. This simple principle has been applied to a number of different signals. In the simplest cases the signals of interest consist either of reflected or of transmitted electrons corresponding to two imaging methods: "scanning electron microscopy" (SEM) and "scanning transmission electron microscopy" (STEM).

Scanning electron microscopy (SEM) is the most popular of the microscopic techniques. This is due to the user friendliness of the apparatus, the simplicity of interpretation of most of the images and the ease of specimen preparation. The obvious limitation is that only surface features are easily accessible. However a variety of different imaging modes can be applied: for instance a charge collecting mode (EBIC) and a light collecting mode (Cathodoluminescence).

Scanning transmission electron microscopy is a very powerful technique operating at nanoscale. Due to the small size of the electron probe very small specimen areas can be studied by several parallel modes: nanodiffraction and imaging, X-ray microanalysis and electron loss spectrometry. Of special interest is the Z-contrast mode, operating under incoherent imaging conditions. It allows imaging of individual atom columns combined with their spectroscopic analysis.

Typical methods for the study of surfaces such as Auger spectroscopy (AES) and photoelectron spectroscopy (XPS), which were initially meant for static chemical

analysis have been developed into scanning chemical mapping methods. The different types of electron optical apparatus to achieve this goal are described in detail. Along the same line X-ray microanalysis had been extended into a scanning mapping method, which can be applied in most electron microscopes.

SIMS (secondary ion mass spectrometer) is one of the most sensitive surface techniques using an ion beam as a probe rather than an electron beam. Originally it was operated in a static mode but like the other beam methods of chemical analysis it was extended into a scanning imaging mode. Since the incident ion beam removes surface atoms it can also be used to obtain composition depth profiles.

Two recently developed methods specifically designed to study magnetic phenomena are SEMPA (scanning electron microscopy with polarisation analysis) and SPLEEM (spin polarised low energy electron microscopy). The first of these techniques provides high resolution images (as compared to Lorentz microscopy) of magnetic microstructures by measuring the spin polarisation of low energy secondary electrons generated in a scanning electron microscope. It is of special importance for the study of magnetic recording materials. SPLEEM makes use of the spin-spin interaction between the spins of the incident electrons and specimen regions with preferred spin alignment in the specular reflection of low energy electrons. It is often combined with LEED (low energy electron diffraction).

Each of the techniques briefly mentioned above is discussed in sufficient detail to understand the physico-chemical principles involved and to judge the applicability to solve given materials characterisation problems.

The editors wish to thank the publisher's staff for their continued support and for the smooth collaboration in the editing procedure.

The editors,        S. Amelinckx  
                          D. Van Dyck  
                          J. Van Landuyt  
                          G. Van Tendeloo

# 1 Stationary Beam Methods

# 1.1 Transmission Electron Microscopy

## 1.1.1 Diffraction Contrast Transmission Electron Microscopy

### 1.1.1.1 Introduction

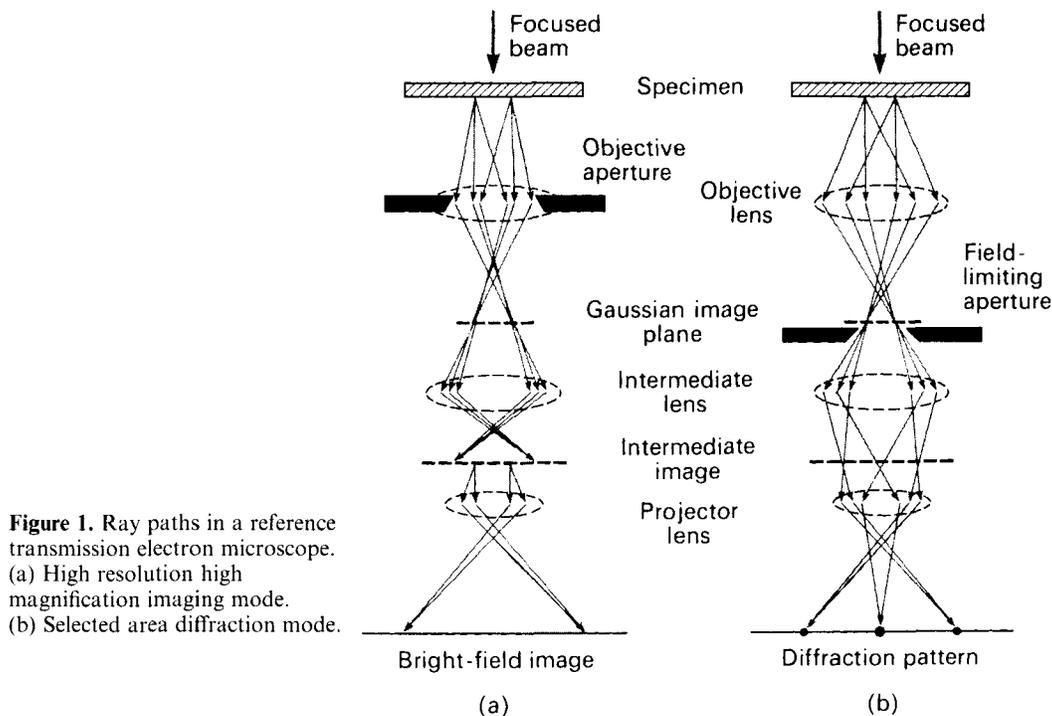
Image formation in transmission electron microscopy is essentially a diffraction phenomenon, normal absorption only playing a minor role. It is therefore necessary to discuss first the basis of electron diffraction. A detailed interpretation of an image requires a knowledge of the corresponding diffraction pattern, adequately oriented with respect to the image. Modern transmission electron microscopes are constructed in such a way that they can be easily switched from the imaging mode to the diffraction mode, and vice versa, without changing the orientation of the specimen. The electron microscope optics are discussed briefly below, emphasizing the general principle rather than the concrete details, since the latter depend on the particular instrument.

### 1.1.1.2 Instrumentation

#### Transmission Microscopes

A modern transmission electron microscope can be schematized as a three-lens

system: an objective lens, an intermediate lens and a projector lens. Each of these lenses is in actual fact often a composite lens, but since the details of this depend on the particular instrument this will not be discussed here. Moreover, condensor lenses are used in the illumination system, but we will consider only the ray paths in a three-lens, image formation system. The system allows easy switching from the high magnification imaging mode to the selected area diffraction mode. The ray paths are shown in Fig. 1. Movable selection apertures are placed: one in the image plane of the objective lens and a second one close to the back focal plane. The front aperture is used to select a small area ( $\leq 1 \mu\text{m}$ ) of the specimen whilst viewing the image. The second one enables us to select either a single beam or a number of image-forming diffracted beams. The image resolution of the system is to a large extent determined by the characteristics of the objective lens, in particular by its spherical aberration constant. Whereas in the high-resolution structure imaging mode the quality of the objective lens is crucial, this is much less the case for diffraction contrast images. In the latter mode the availability of very small beam-selection apertures is important, since this determines to what extent a single beam can be selected for dark-field imaging of materials with a relatively large unit cell.



**Figure 1.** Ray paths in a reference transmission electron microscope. (a) High resolution high magnification imaging mode. (b) Selected area diffraction mode.

The intermediate and projector lenses provide the desired magnification. When using the high-resolution mode a sufficiently large magnification (on the fluorescent screen or on the TV monitor) is necessary to see separate atom columns so as to allow proper focusing. Although magnetic lenses normally rotate the image about the optical axis, in recently designed microscopes these rotations are compensated for by a suitable device and, as a result, the image and diffraction pattern have parallel orientations, which is particularly important in diffraction contrast work. In certain instruments the image and the diffraction pattern may still differ in orientation by  $180^\circ$  for certain lens combinations.

### Lens Configurations

#### *High Resolution, High Magnification Imaging Mode*

The electron beam produced by an electron source (see below) is collimated by the condenser lens system (not shown in Fig. 1a) and scattered by the specimen. An image is formed in the image plane of the objective lens (Fig. 1a). The selector aperture allows us to select one area of the image (i.e., of the specimen) which is then magnified by the intermediate lens. The intermediate lens is focused on the image plane of the objective lens and an intermediate image is formed in the image plane of the intermediate lens. This image is the object for the projector lens which

forms a final image on a fluorescent screen or on the entrance plane of a recording device (see below).

### *Diffraction Mode*

In the diffraction mode (Fig. 1b) the intermediate lens is weakened, that is the focal length is made larger, in such a way that the back focal plane of the objective lens coincides with the object plane of the projector lens. A magnified representation of the diffraction pattern is then produced on the fluorescent screen. In the process the selected area is not changed since only the strength of the intermediate lens has been modified. The diffraction pattern is thus representative of the selected area. However, it should be noted that under high resolution conditions the field of view in the image is much smaller than the selected area in the diffraction mode.

### **Diffraction Contrast Imaging Modes**

As discussed below, diffraction contrast images are maps of the intensity distribution in highly magnified diffraction spots. They are usually obtained under two-beam conditions. The aperture placed close to the back focal plane of the objective lens allows us to select either the transmitted beam or the diffracted beam. The corresponding diffraction spot is subsequently magnified by the rest of the lens system. If the transmitted beam is selected, a bright field image is obtained; that is the area of the image not covered by the specimen is bright. If the diffracted beam is selected, a dark field image is obtained; the background is now dark.

Whereas the beam remains along the optical axis in the case of a bright field

image, it encloses twice the Bragg angle of the active reflection with the optical axis for a dark field image. Non-axial beams suffer angle-dependent lens aberrations and the corresponding image is therefore often blurred by streaking. This can be avoided by tilting the incident beam over twice the Bragg angle; the diffracted beam then travels along the optical axis. Recently developed microscopes have a built-in device that allows the incident beam to be deflected over the required angle to bring a selected diffracted beam along the optical axis.

### **Electron Sources**

In older microscopes the source of electrons is a heated V-shaped tungsten wire, the tip of the V forming a localized emitter of electrons with an effective size of the order of  $10\ \mu\text{m}$ . In subsequent versions the effective size of the emitting area is decreased by attaching a pointed tungsten wire to the tip of the V-shaped heating filament. Also, a pointed  $\text{LaB}_6$  single crystal mounted on a V-shaped tungsten heating filament is often used because of its small work function. It emits electrons at a lower temperature than tungsten and, moreover, the thermal energy distribution of the electrons is narrower ( $\approx 1\ \text{eV}$ ), thus leading to less chromatic aberration. In recent years the thermal spread of the emitted electrons has been reduced further to less than  $0.5\ \text{eV}$  by the use of cold field emission guns. Such guns consist of a very fine point placed on the pointed filament which emits electrons by tunneling. A sufficiently strong electric field (of the order of  $10^6\ \text{V cm}^{-1}$ ) is required in the vicinity of the field point to cause a sufficiently strong

tunneling current density. Tunneling can be thermally assisted by slightly heating the tungsten point. The brightness of such sources is about a thousand times larger than that of the reference tungsten filament but, since the emitting area ( $\sim 5\text{mm}^2$ ) is much smaller, the electron current is smaller. Field emission guns are particularly important in applications where a high degree of coherence is desirable, such as in high resolution atomic imaging.

### **Electron Beam Shaping**

The electrons emitted by the source must be accelerated and a collimated beam must be formed. This is achieved by applying a large negative voltage to the emitter, keeping the anode grounded. Close to the emitter is a cylindrical electrode, the 'Wehnelt cylinder', which is biased at a negative voltage of a few hundred volts with respect to the emitter. The function of the Wehnelt cylinder is to stabilize the beam current and to focus the electron beam so as to form behind the anode a cross-over, which acts as the virtual source of electrons. The incident beam is finally shaped into a parallel beam (or possibly a convergent beam) by the system of condenser lenses and by apertures. Typically, the angular spread may be made as small as  $10^{-4}$  rad.

### **Electrical Supply System**

Very stable supply systems for the high voltage used to accelerate the electrons and for the lens currents are essential to minimize chromatic aberrations. In present

day commercial microscopes, supply systems with a stability of 1 part in  $10^6$  are used.

### **Vacuum Systems**

Clean and vibration-free vacuum systems are essential to provide stability and avoid contamination of the specimen by a carbon film resulting from the cracking of organic molecules present in the residual gas. Anti-contamination devices such as metal blades surrounding the specimen and cooling to liquid nitrogen temperature are available in most microscopes. Use is made of diffusion pumps, turbomolecular pumps, ion pumps, and even sublimation pumps to achieve a vacuum of the order of  $133 \times 10^{-8}$  Pa ( $10^{-8}$  torr).

### **Recording and Viewing Media**

Usually images are made visible on a fluorescent screen and viewed by means of binoculars. The simplest recording medium is still the photographic emulsion which is particularly efficient for recording electron images. A shutter mechanism allows one to remove the fluorescent screen and expose the underlying film. Degassing the photographic material prior to use is strongly recommended. In order to reduce radiation damage to the specimen it is important to keep the electron dose used to view and record the image as small as possible; for this purpose electronic viewing and recording methods have been developed. These techniques are discussed in Chap. 8, Sec. 1 and 2 of this Volume.

### 1.1.1.3 Electron Diffraction

#### Atomic Scattering Factor

Electrons are scattered by atoms as a result of the Coulomb interaction with the positively charged nucleus and with the negatively charged electron cloud. The atomic scattering factor thus contains two contributions of opposite sign [1]

$$f_c(\theta) = \frac{me^2\lambda}{2h^2} \frac{[Z - f_x(\theta)]}{\sin^2\theta} \quad (1)$$

where  $f_x(\theta)$  is the atomic scattering factor for X-rays, which are only scattered by the electron cloud.  $Z$  is the atomic number (or the positive charge on the nucleus),  $\lambda$  is the electron wavelength [see Eq. (2)],  $m$  is the electron mass,  $e$  is the electron charge,  $h$  is Planck's constant, and  $\theta$  is the scattering angle. The electron wavelength is given by the de Broglie relation  $\lambda = h/mv$ . It is related to the accelerating potential  $E$  by the relativistic relation

$$\lambda = h \left[ 2m_0E \left( 1 + \frac{eE}{2m_0c^2} \right) \right]^{-1/2} \quad (2)$$

where  $m_0$  is the rest mass of the electron. With  $E$  in the range 200–400 kV the electrons used in electron microscopy are to a non-negligible extent relativistic since they travel at speeds larger than half the speed of light.

The first term in Eq. (1) clearly relates to the nucleus, whereas the second term is due to the electron cloud. The interaction with matter is much stronger for electrons than for X-rays or neutrons by a factor of about  $10^4$ . Multiple diffraction events will therefore have a high probability. The factor  $1/\sin^2\theta$  in Eq. (1) causes scattering to be oriented mainly in the forward

direction. Values of  $f_c(\theta)$  for different atoms have been tabulated in [2].

#### Diffraction by Crystals

The amplitude diffracted in any given direction by an assembly of atoms is the sum of the amplitudes scattered by the individual atoms in that direction, taking into account the phase differences resulting from the geometry of the assembly. In a crystal the atoms are located on a three-dimensional lattice which can be characterized by its three base vectors  $\mathbf{a}_1$ ,  $\mathbf{a}_2$ , and  $\mathbf{a}_3$ . A general lattice node is then given by

$$\mathbf{A}_L = l_1\mathbf{a}_1 + l_2\mathbf{a}_2 + l_3\mathbf{a}_3 \quad (3)$$

where  $l_j$  are integers. The volume of the unit cell is  $V_a = (\mathbf{a}_1 \times \mathbf{a}_2) \cdot \mathbf{a}_3$ . The sum and difference of two lattice vectors is again a lattice vector.

It is convenient to define the corresponding reciprocal lattice by its base vectors  $\mathbf{b}_1$ ,  $\mathbf{b}_2$ , and  $\mathbf{b}_3$ , which are related to the  $\mathbf{a}_j$  by the relations  $\mathbf{a}_i \cdot \mathbf{b}_j = \delta_{ij}$  where  $\delta_{ij} = 1$  if  $i = j$  and  $\delta_{ij} = 0$  if  $i \neq j$ . A general node point of the reciprocal lattice is then given by

$$\mathbf{B}_H = h_1\mathbf{b}_1 + h_2\mathbf{b}_2 + h_3\mathbf{b}_3 \quad (4)$$

the  $h_j$  are integers, called 'Miller indices' for planes.

A useful property is  $\mathbf{A}_L \cdot \mathbf{B}_H = \text{Integer}$ . The volume of the unit cell of the reciprocal lattice is  $V_b = 1/V_a$ . The spacing of lattice planes with indices  $\mathbf{H}$  ( $h_1, h_2, h_3$ ) is given by  $d_H = 1/|\mathbf{B}_H|$ .

A crystal structure is described with respect to its lattice by specifying the contents of the unit cell, that is by giving the position vectors  $\boldsymbol{\rho}_j$  for all atoms  $j = 1, \dots, N$  in the unit cell. The assembly

of scattering units is thus given by the position vectors  $\mathbf{A}_L + \boldsymbol{\rho}_j$  for the atoms of type  $j$ .

### Diffraction Conditions

The diffraction conditions for a periodic object can be formulated in terms of direct space (i.e., of the lattice) or in terms of diffraction space (i.e., of the reciprocal lattice). The two formulations have the same physical content but the latter is often more convenient and more directly related to the diffraction pattern.

If the attention is focused on the lattice the diffraction condition is Bragg's law [3] which states that the path difference between waves 'reflected' by successive lattice planes is an integer number of wavelengths

$$2d_H \sin \theta_H = n\lambda \quad (5a)$$

where  $d_H$  is the interplanar spacing and  $\theta_H$  is the corresponding Bragg angle (Fig. 2);  $n$  is an integer. This statement might create the wrong impression that 'reflection' takes place. However, the difference with specular reflection is important: only for the discrete angles  $\theta_H$  does 'reflection' take place, whereas in specular reflection all

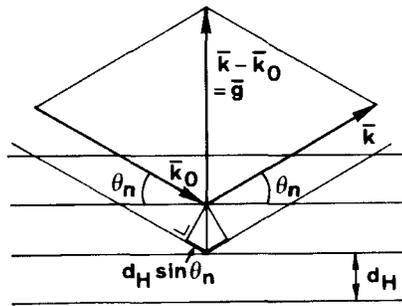


Figure 2. Geometry of Bragg scattering ( $\theta_n = \theta_H$ ).

angles are permitted. This clearly shows that we have, in fact, interference rather than reflection. Since the notion of 'reflection' has nevertheless been extremely useful, especially in structure determination, it continues to be used. Due to the small wavelength of electrons [ $\approx 3 \times 10^{-3}$  nm ( $2 \times 10^{-2}$  Å)] Bragg angles are quite small ( $\approx 10^{-3}$  rad) and Bragg's law can often be approximated by

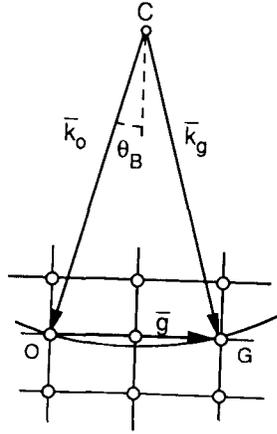
$$2d_H \theta_H = n\lambda \quad (5b)$$

In reciprocal space the diffraction conditions can be formulated in terms of Ewald's sphere

$$\mathbf{k}_g = \mathbf{k}_0 + \mathbf{g} \quad (6)$$

where  $\mathbf{k}_0$  is the wavevector of the incident plane wave [ $\mathbf{k}_0 = (1/\lambda)\mathbf{e}_n$ , where  $\mathbf{e}_n$  is the normal to the plane wavefront] and  $\mathbf{k}_g$  is the wavevector of the scattered wave. The term  $\mathbf{g}$  is a reciprocal lattice vector, called the *diffraction vector*; it is an element of the set  $\mathbf{B}_H$  and is thus specified by three integers  $h_1$ ,  $h_2$ , and  $h_3$ , the Miller indices.

The Ewald condition gives rise to an elegant construction (Fig. 3). Let  $\mathbf{k}_0$  represent the incident wave; its endpoint coincides with the origin of reciprocal space O; its starting point C is then the center of a sphere with radius  $1/\lambda$  (Ewald's sphere) [4]. If this sphere passes through another reciprocal lattice node H, a diffracted beam  $\mathbf{CH} \equiv \mathbf{k}_g$  is produced. Tilting the specimen is equivalent to tilting the reciprocal lattice over the same angle about a parallel axis. Tilting thus permits the 'excitation' of specific nodes of the reciprocal lattice. It should be noted that since the electron wavelength is of the order of 0.001 nm (0.01 Å) (i.e.,  $|\mathbf{k}| = 10 \text{ nm}^{-1}$  ( $10^2 \text{ Å}^{-1}$ ), whereas the mesh size of the reciprocal lattice is of



**Figure 3.** Ewald construction. The Ewald sphere with radius  $|k_0| = |k_g| = 1/\lambda$  passes through the node G.  $k_0$ : wavevector incident beam;  $k_g$ : wavevector diffracted beam;  $g$ : diffraction vector.

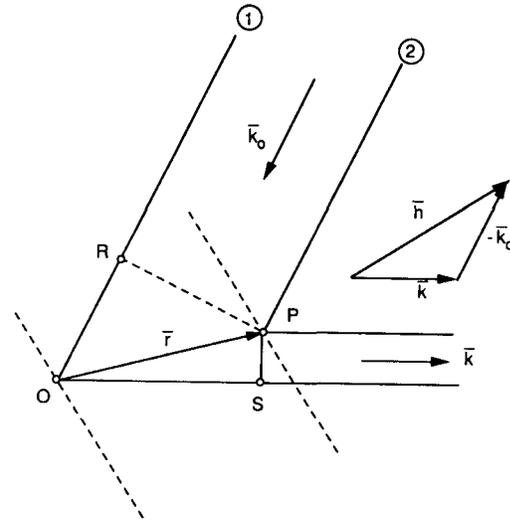
the order of  $0.1 \text{ nm}$  ( $1 \text{ \AA}^{-1}$ ), the sphere radius is quite large and it can be approximated for most practical purposes by a plane normal to the incident wavevector  $k_0$ . The diffraction pattern is thus obtained as a central projection of a planar section of reciprocal space on to the screen (or photographic plate).

### Diffraction Amplitude

Let  $k_0$  be the wavevector of an incident wave  $\exp(2\pi i k_0 r)$  and  $k = 1/\lambda$  the wavevector of the scattered wave. The phase difference between the waves diffracted by an atom at the origin O and an atom P at  $r_j$  is  $(2\pi/\lambda)(OR + OS) = 2\pi(k - k_0) \cdot r_j$  (Fig. 4) and the total scattered amplitude by the assembly of points at positions  $r_j$  is then given by

$$A(\mathbf{k}) = \sum_j f_j \exp[2\pi i (\mathbf{k} - \mathbf{k}_0) \cdot \mathbf{r}_j] \quad (7)$$

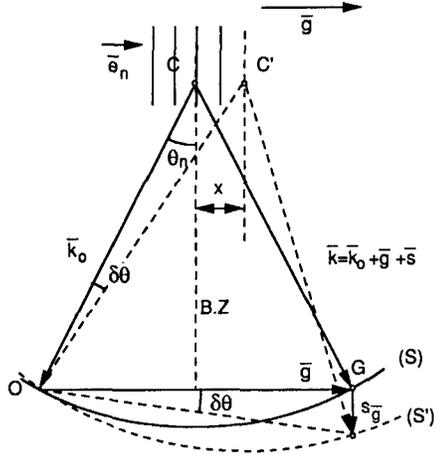
where  $f_j$  is the atomic scattering amplitude



**Figure 4.** Path difference  $OR + OS$  between waves scattered at O and at P.

of the atoms at  $r_j$ . This amplitude will be maximal if all waves are in phase. If the atoms are assumed to be located on a lattice  $r_j = A_L$ ; a maximum of  $A(\mathbf{k})$  will then occur if all exponents are integer multiples of  $2\pi i$ , that is, if  $(\mathbf{k} - \mathbf{k}_0) \cdot A_L = \text{Integer}$ . This will be the case if  $\mathbf{g} \equiv \mathbf{k} - \mathbf{k}_0 = \mathbf{B}_H$ , which is the Ewald condition [Eq. (6)].

The scattered amplitude will also differ from zero even if the scattering vector  $\mathbf{g} \equiv \mathbf{k} - \mathbf{k}_0$  differs somewhat from a reciprocal lattice vector. This deviation from the exact diffraction condition is described by the vector  $s$  (Fig. 5) ('excitation error' also called 'deviation parameter'). It measures the distance of the reciprocal lattice node  $G_H$  to the Ewald sphere; it is normal to the foil surface. The vector  $s$  is positive when the reciprocal lattice node is inside Ewald's sphere; it is negative when it is outside. The positive direction of  $s$  is thus in the direction of propagation of the electrons. Equation (7) can now be



**Figure 5.** The deviation from the exact Bragg condition is characterized by  $s_g$ , which is positive in the sense of the propagating electrons;  $x$  is an alternative deviation parameter. B.Z: Brillouin zone boundary.

generalized to

$$A(\mathbf{g}) = \sum_j f_j \exp[2\pi i(\mathbf{g} + \mathbf{s}) \cdot \mathbf{r}_j] \quad (8)$$

Taking into account that an atom with a scattering factor  $f_j$  is found at the positions  $\mathbf{A}_L + \boldsymbol{\rho}_j$  ( $j = 1, \dots, N$ ), the amplitude can be written as

$$\begin{aligned} A(\mathbf{g}) &= \sum_L \sum_{j=1, \dots, N} f_j \\ &\quad \times \exp[2\pi i(\mathbf{g} + \mathbf{s}) \cdot (\mathbf{A}_L + \boldsymbol{\rho}_j)] \\ &= \sum_L \left[ \sum_j f_j \exp(2\pi i \mathbf{g} \cdot \boldsymbol{\rho}_j) \right] \\ &\quad \times \exp(2\pi i \mathbf{s} \cdot \mathbf{A}_L) \end{aligned} \quad (9)$$

where we note that  $\mathbf{g} \cdot \mathbf{A}_L = \text{Integer}$ , since  $\mathbf{g}$  is an element of the set  $\mathbf{B}_H$  and  $\mathbf{s} \cdot \boldsymbol{\rho}_j$  is negligible compared to the other terms.

The expression

$$F_g \equiv \sum_j f_j \exp(2\pi i \mathbf{g} \cdot \boldsymbol{\rho}_j)$$

describes the amplitude scattered by one unit cell in the direction defined by  $\mathbf{g}$ ; it is called the *structure amplitude* or *structure factor*. Equation (9) then becomes

$$A(\mathbf{g}) = F_g \sum_L \exp(2\pi i \mathbf{s} \cdot \mathbf{A}_L) \quad (10)$$

We assume that the crystal contains  $N_1$ ,  $N_2$  and  $N_3$  unit cells respectively along the  $\mathbf{a}_1$ ,  $\mathbf{a}_2$  and  $\mathbf{a}_3$  directions. The scattering amplitude then becomes

$$\begin{aligned} A(\mathbf{g}) &= F_g \sum_{l_1=0}^{N_1-1} \sum_{l_2=0}^{N_2-1} \sum_{l_3=0}^{N_3-1} \\ &\quad \times \exp[2\pi i(s_1 l_1 a_1 + s_2 l_2 a_2 + s_3 l_3 a_3)] \end{aligned} \quad (11)$$

where  $s_1, s_2$  and  $s_3$  are the components of  $\mathbf{s}$  along the three base directions of the lattice and  $\mathbf{s} \cdot \mathbf{A}_L = s_1 l_1 a_1 + s_2 l_2 a_2 + s_3 l_3 a_3$ . Separating the triple sum into the product of three single sums, one obtains

$$\begin{aligned} A(\mathbf{g}) &= F_g \left[ \sum_{l_1=0}^{N_1-1} \exp(2\pi i s_1 a_1 l_1) \right. \\ &\quad \cdot \sum_{l_2=0}^{N_2-1} \exp(2\pi i s_2 a_2 l_2) \\ &\quad \left. \cdot \sum_{l_3=0}^{N_3-1} \exp(2\pi i s_3 a_3 l_3) \right] \end{aligned} \quad (12)$$

Performing the summation of the geometrical series, and omitting an irrelevant phase factor, one obtains finally the von Laue interference function

$$\begin{aligned} A(\mathbf{g}) &\approx F_g \frac{\sin(\pi s_1 N_1 a_1)}{\sin(\pi s_1 a_1)} \cdot \frac{\sin(\pi s_2 N_2 a_2)}{\sin(\pi s_2 a_2)} \\ &\quad \cdot \frac{\sin(\pi s_3 N_3 a_3)}{\sin(\pi s_3 a_3)} N_1 N_2 N_3 \end{aligned} \quad (13)$$

which describes the dependence of the scattered amplitude on the deviation parameter.

Since  $\pi s_1 a_1$ ,  $\pi s_2 a_2$ , and  $\pi s_3 a_3$  are small, the sine terms in the denominators can be replaced by their arguments. We further note that for large  $N$ ,  $\sin(\pi N s a / \pi N s a) \simeq \delta(s)$  where  $\delta(s) = 0$  for  $s \neq 0$  and  $\delta(s) = 1$  if  $s = 0$ . Taking this into account we can write

$$A(\mathbf{g}) = F_g \delta(s_1) \delta(s_2) \delta(s_3) \frac{\Omega}{V_a} \quad (14)$$

where  $\Omega$  is the volume of the crystal and  $V_a$  is the volume of the unit cell:  $\Omega = N_1 N_2 N_3 V_a$ .

With regard to transmission electron microscopy, thin foil specimens contain a large number of unit cells in the two lateral directions, but along the thickness the number of unit cells is finite and Eq. (13) reduces to

$$A(\mathbf{g}) = F_g \frac{\Omega}{V_a} \delta(s_1) \delta(s_2) \frac{\sin(\pi N_3 a_3 s_3)}{(\pi N_3 a_3 s_3)} \quad (15)$$

where  $s_3$  is measured perpendicular to the foil. Introducing the specimen thickness  $t = N_3 a_3$  we can write

$$A(\mathbf{g}) = F_g \frac{S_a}{V_a} \frac{\sin(\pi s_3 t)}{\pi s_3} \delta(s_1) \delta(s_2) \quad (16a)$$

with  $S_a = N_1 N_2 a_1 a_2$ . This can be rewritten per unit surface area as:

$$A(\mathbf{g}) = \frac{\sin(\pi s t)}{s t_g} \quad (16b)$$

with  $t_g = (\pi V_a / F_g)$ ;  $t_g$  is called the *extinction distance*.

This result is interpreted in terms of diffraction space as meaning that the reciprocal lattice nodes of a thin foil are not mathematical points but are rod shaped (relrods) and perpendicular to the foil plane with a weight given by  $(\sin \pi s t) / \pi s$ . This function is shown in Fig. 6.

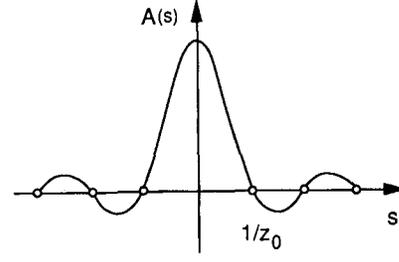
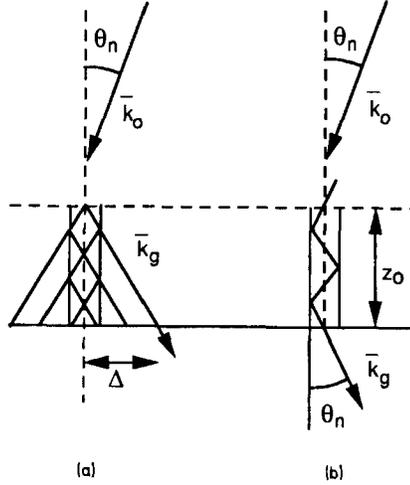


Figure 6. Scattered amplitude versus  $s$  according to the kinematical approximation.  $z_0$ : foil thickness.

The corresponding intensity profile is called the 'rocking curve' according to the kinematical theory (Fig. 6). An amplitude can be associated with each intersection point of the Ewald sphere with this segment, the amplitude being given by the value of this profile at the intersection point. It is customary to describe the diffraction geometry by saying that the reciprocal lattice nodes have become 'relrods' with a length profile in the  $z$  direction, parallel to the foil normal, given by the above-mentioned function, but infinitely sharp in the directions  $x$  and  $y$  in the foil. In a thin foil the vector  $s$  is thus oriented along the normal to the foil plane. By convention,  $s$  is counted positive in the sense of the propagating electrons, that is if the reciprocal lattice node  $G$  is inside Ewald's sphere.

### The Column Approximation

Since Bragg angles are quite small ( $\approx 10^{-3}$  rad) in electron diffraction, the intensity observed at a point in the exit face of a foil is essentially determined by the diffraction events occurring in a narrow column of crystal centered on the considered point extending through the thickness of the foil in a direction parallel

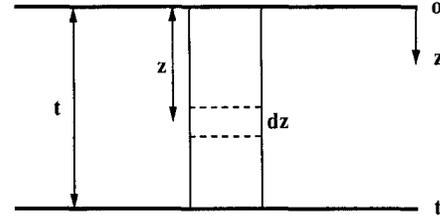


**Figure 7.** Column approximation: (a) kinematical approximation; (b) dynamical approximation.

to the incident beam (Fig. 7) [5]. Such a column diffracts independently of the surrounding columns; its lateral dimensions are, at most, given by  $\Delta = \theta_n z_0$  ( $\theta_n \equiv \theta_B$ ;  $z_0 = t = \text{thickness}$ ). For a strain-free foil limited by parallel faces, all such columns behave in the same manner and a uniform intensity results in the image. However, if defects are present, columns in the vicinity of the defects will produce an intensity which differs from that of the surrounding columns and hence a defect image will result. The amplitude found at the exit point of a column is given by (omitting irrelevant factors)

$$A_g = F_g \int_0^t e^{2\pi i s z} dz \quad (17)$$

where summation over all unit cells in the column has been replaced by integration along the columns. The meaning of  $z$  and  $t$  is shown in Fig. 8. The exponential is the phase factor, relative to the entrance point due to the volume element  $dz$  at level  $z$  behind the entrance face. If  $s$  is a constant,



**Figure 8.** Illustration of the notations used.

as is the case in a defect-free crystal, the result is

$$A_g = \frac{F_g \sin(\pi s t)}{\pi s} \quad (18)$$

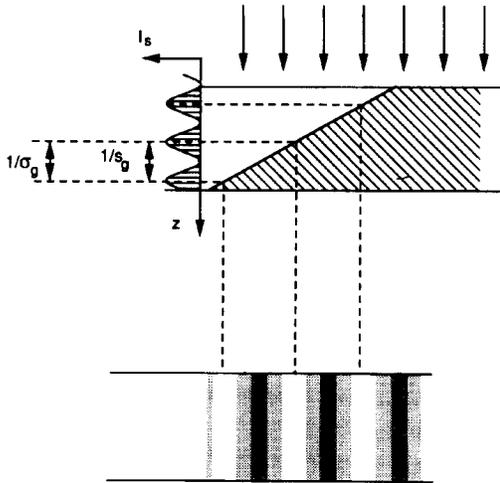
which is consistent with Eq. (16). However, if a defect is present,  $s$  becomes a function of  $x, y$ , and  $z$ . For a column at  $(x, y)$  one obtains

$$A_g(x, y) = F_g \int_0^t e^{2\pi i s(x, y, z)} dz \quad (19)$$

Each column, of which the size is arbitrarily small, now furnishes a picture element of the defect image. For most defects such images will have to be computed numerically since the analytical expressions may become too complicated. Defect imaging is discussed in detail below.

If the strain-free foil has a nonuniform thickness the intensity at the exit face will depend on the length of the column and it will be equal for columns of equal length. The columns that produce extrema in intensity form geometrical loci, called *thickness extinction contours*, which are dark for a minimum and bright for a maximum. In a wedge-shaped crystal such contours form a set of straight fringes parallel to the wedge edge, described by Eq. (18). Their formation is represented schematically in Fig. 9.

When a defect-free foil of uniform thickness is bent the  $s$  value becomes



**Figure 9.** Model for the formation of thickness extinction contours at a wedge-shaped crystal.  $I_s$ : scattered intensity;  $s_g$ : excitation error;  $\sigma_g = \sqrt{1 + (s_g t_g)^2} / t_g$ .

variable along a line perpendicular to the bending axis. The loci of equal  $s$  value are then again imaged as contours of equal brightness; they are called *inclination extinction contours* (Fig. 10a). These image in a sense the rocking curve represented by Eq. (18). Figure 10b shows a cylindrically bent uniformly thick foil of graphite; the inclination contours corresponding to the main and subsidiary maxima can clearly be observed.

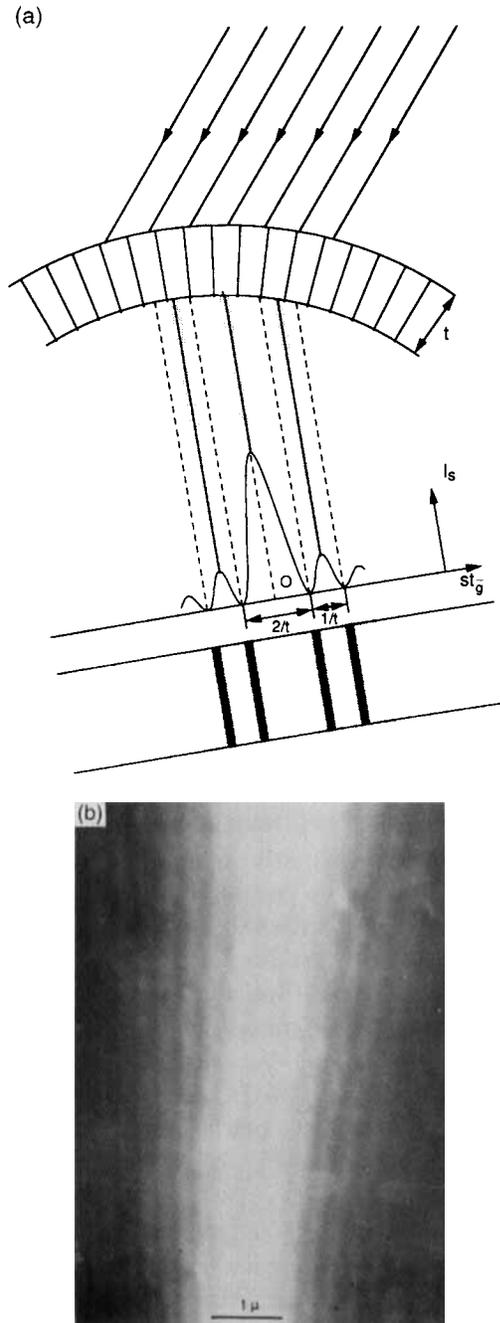
### Amplitude–Phase Diagram

A plane wave represented by

$$A \exp i(kx - \omega t + \psi)$$

$$\equiv A \exp(i\psi) \exp[i(kx - \omega t)]$$

is characterized by a complex amplitude  $A \exp(i\psi)$ , a wavevector  $k$  and an angular frequency  $\omega$ . The interference between two waves of this type, assuming the wavevector  $k$  and the angular frequency  $\omega$  to



**Figure 10.** Diffracted intensity as a function of the angle of incidence in a cylindrically bent foil: (a) formation of bent contours ( $I_s$ : scattered intensity); (b) cylindrically bent graphite foil.

be the same for both, produces a resultant wave with the same  $\omega$  and  $k$  but different  $A$  and  $\psi$ . The problem thus reduces to summing the complex amplitudes  $A \exp(i\psi)$ , the propagation factor being common to all waves.

The complex amplitude is represented in the complex plane by a vector with modulus  $A$  and argument  $\psi$ . It is easy to show that the sum of the two such vectors is again a vector representing the resultant wave. Waves with the same  $k$  and  $\omega$  can thus be summed graphically by adding vectors in the complex plane.

Apart from a phase factor, the amplitude scattered by a column of crystal along the  $z$  axis, in the kinematical approximation, is given by the sum

$$A(\mathbf{g}) = F_g \sum_n \exp(2\pi i s z_n) \Delta z_n \quad (20)$$

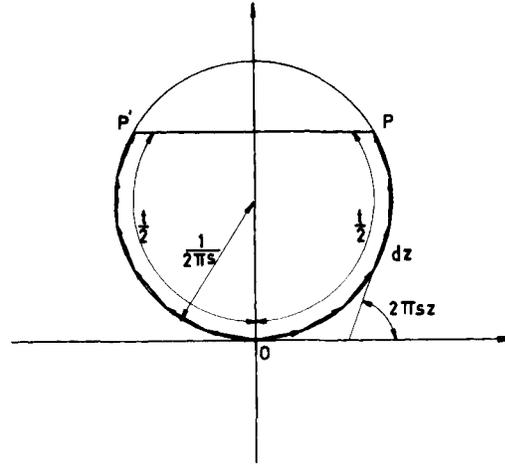
or in the continuum approximation by the integral mentioned above:

$$A(\mathbf{g}) = F_g \int_0^t \exp(2\pi i s z) dz \quad (21)$$

This sum can be considered as consisting of terms  $F_g \Delta z \cdot \exp(2\pi i s z)$ , that is  $A = F_g \Delta z$ ;  $\psi = 2\pi i s z$ , corresponding with slices  $\Delta z$  of the column. The amplitude phase diagram then consists of vectors, all of the same length  $F_g \Delta z$  and enclosing angles of  $2\pi s \Delta z$ . In the limit  $\Delta z \rightarrow 0$ , the locus of the endpoints is a circle with radius (Fig. 11)

$$\begin{aligned} R &= F_g \lim_{\Delta z \rightarrow 0} (\Delta z / 2\pi s \Delta z) \\ &= \frac{F_g}{2\pi s} \end{aligned} \quad (22)$$

The length of the circular arc is equal to the column length, that is to the foil thickness  $t$ . Figure 11 shows that: the diffracted

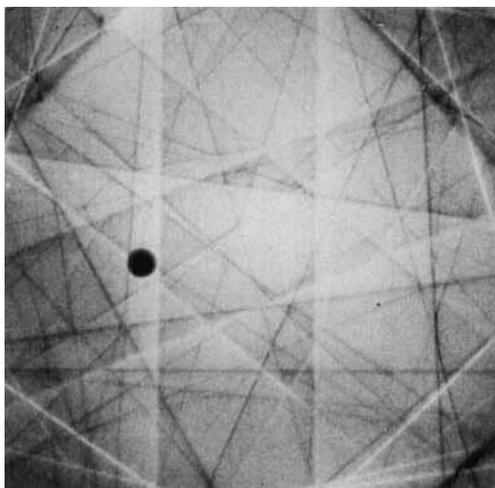


**Figure 11.** Amplitude phase diagram for a perfect crystal foil.

amplitude will be zero if the circular arc is a number of complete circles, that is for  $t = k/s$ ; and there will be maxima if  $t = (1/s)(k + \frac{1}{2})$  ( $k = \text{Integer}$ ), the maximum amplitude being equal to the diameter of the circle [i.e.,  $A_{\text{max}} = (1/\pi s)F_h$ ].

### Kikuchi Lines

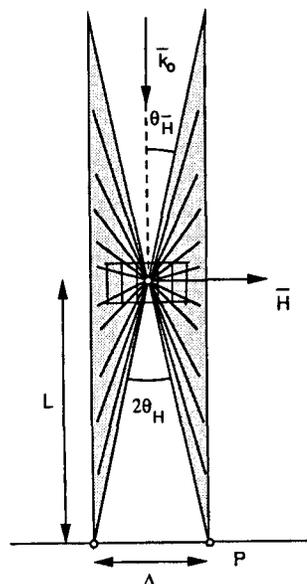
In sufficiently thick and almost perfect specimens, spot patterns are no longer observed; instead a diffraction phenomenon, first discovered by Kikuchi in 1928 [6], is produced. It usually consists in the occurrence of pairs of bright and dark straight lines in the diffraction pattern, as shown in Fig. 12. In foils of intermediate thickness one can observe the Kikuchi pattern superimposed on the spot pattern. The geometry of the Kikuchi pattern can satisfactorily be explained by assuming that not only are electrons Bragg scattered, but that also a substantial fraction, especially in thick foils, is scattered inelastically and incoherently in the crystal, the energy



**Figure 12.** Pattern of Kikuchi lines in a rather thick silicon crystal.

loss being much smaller than the energy of the incident electron; the electron wavelength is then not appreciably changed. Inside the crystal these randomly scattered electrons impinge on the lattice planes from all directions, but preferentially in the forward direction, and can subsequently give rise to Bragg scattering.

A symmetrical situation with respect to the set of lattice planes  $H$ , with spacing  $d_H$ , is shown in Fig. 13. Bragg scattering out of the incident beam is assumed to be weak since the Bragg condition is not satisfied. However, a fraction of the randomly scattered electrons have the correct direction of incidence to give rise to Bragg diffraction by the set of lattice planes considered. The geometrical locus of these Bragg scattered electron beams is a double cone of revolution with an opening angle  $(\pi/2) - \theta_H$  and with its axis along  $H$  (where  $\theta_H$  is the Bragg angle). These cones are therefore rather 'flat' and the intersection lines of the two sheets of this double cone with the photographic plate

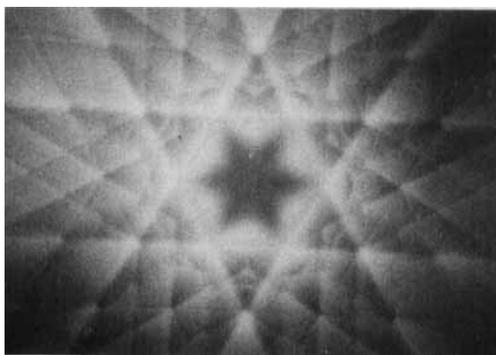


**Figure 13.** Geometry of the Kikuchi cones in the symmetrical orientation.

$P$  looks like two parallel straight lines, although in actual fact they are two branches of a hyperbolic conical section. The angular separation of these two lines is  $2\theta_H$ . The separation  $\Delta$  observed on the plate is thus  $\Delta = 2L\theta_H$ , where  $L$  is the camera length, i.e. the specimen to plate distance. The angular separation does not depend on the crystal orientation.

The geometry of this cone (i.e., the axis of revolution and the opening angle) is entirely fixed by the crystal lattice and the electron wavelength, and is independent of the incident beam direction. Tilting the specimen thus leads to an equal tilt of the double cone, but leaves the geometry of the spot diffraction pattern unchanged, provided the same reflections remain excited, that is as long as the same 'relrods' are intersected by Ewald's sphere. The relative position of the spot pattern and of the Kikuchi line pattern is thus very





**Figure 15.** Kikuchi bands in a symmetrically oriented foil of silicon, along the [111] zone.

origin, cause an excess, compared to a lower background, along the part of the cone containing the coherently scattered Bragg beam. This background is somewhat smaller since the scattering angle is larger. Therefore the excess electrons produce a bright line through the Bragg spot. The angular separation of the bright–dark line pair, is the same as in the symmetrical orientation; however, the linear separation measured on the plate may depend slightly on the tilt angle. The symmetrical situation is represented schematically in Fig. 13. In the symmetrical orientation the Kikuchi lines often form the limiting lines of ‘Kikuchi bands’, the inside of which exhibit a somewhat lower brightness than the outside (Fig. 15) [7, 8]. In this particular orientation the Kikuchi lines can be considered as images of the Brillouin zone boundaries belonging to the different reflections.

#### Determination of Sign and Magnitude of $s$

Starting with a foil in the exact Bragg orientation for the reflection  $G$ , the bright Kikuchi line passes through  $G$  (Fig. 14a),

whereas the dark line passes through the origin of the reciprocal lattice. Tilting the specimen over a small angle  $\delta\theta$  in the clockwise sense (i.e., towards  $s < 0$  about an axis in the foil plane, normal to the  $g$  vector), the position of the bright Kikuchi line moves towards the origin over  $u = L\delta\theta$  (Fig. 14c). The vector  $g$  is then rotated over the same angle  $\delta\theta$  and hence  $s$  becomes negative and equal to  $s = g\delta\theta$ ; the relation between  $u$  and  $s$  is thus:

$$u = \left(\frac{L}{g}\right)s \quad (23a)$$

and

$$\Delta u = \left(\frac{L}{g}\right)\Delta s \quad (23b)$$

This relation allows one to determine the sign and the magnitude of  $s$  from the relative position of a diffraction spot and its associated Kikuchi line (Fig. 14). It also allows one to determine the orientation difference between two crystal parts. The sign of  $s$  is required for a number of applications such as the determination of the sign of the Burgers vector of dislocation, the vacancy or interstitial character of a dislocation loop, and the orientation difference across a domain boundary, as will be discussed below. The magnitude of  $s$  is needed when applying the weak-beam method (see Section 1.1.17.4 of this chapter).

#### Refraction of Electrons at an Interface

Refraction of the incident electron beam takes place at the vacuum–crystal foil interface because the lengths of the wavevectors are different in the two

media:

$$K_0 \equiv K_{\text{vac}} = \frac{(2meE)^{1/2}}{h} \quad (24)$$

$$K \equiv K_{\text{cryst}} = \frac{[2me(E + V_0)]^{1/2}}{h} \quad (25)$$

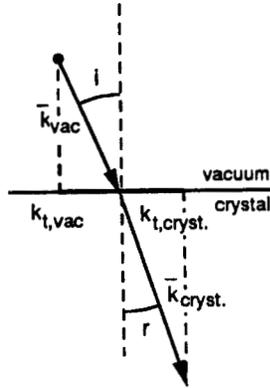
but the tangential components have to be conserved at the interface. Figure 16 shows the relation between the two wavevectors; one has

$$\begin{aligned} n &= \sin i / \sin r \\ &= \left( \frac{K_{t,\text{vac}}}{K_0} \right) / \left( \frac{K_{t,\text{cryst}}}{K_{\text{cryst}}} \right) \\ &= \frac{K_{\text{cryst}}}{K_{\text{vac}}} \\ &= \left( \frac{E + V_0}{E} \right)^{1/2} \end{aligned} \quad (26)$$

The refractive index  $n$  is thus:

$$\begin{aligned} n &= [(E + V_0)/E]^{1/2} \\ &\simeq \left( 1 + \frac{V_0}{E} \right)^{1/2} \end{aligned} \quad (27)$$

Since  $V_0 \ll E$ ,  $n$  is just slightly larger than 1 and the angle of refraction is very small,



**Figure 16.** Refraction of electrons at the crystal–vacuum interface.

especially for quasinormal incidence as is the case in most observations. Refraction nevertheless produces an observable effect for grazing incidence. Small polyhedral particles which are completely embedded in the incident beam may produce diffraction spots consisting of a number of components corresponding to the number of crystal wedges crossed by the beam. Refraction also produces an observable effect on the diffraction pattern of single thin fibers such as chrysotile.

### 1.1.1.4 Kinematical Diffraction Theory

#### Kinematical Diffraction Theory as a Born Approximation

A rigorous diffraction theory is based on Schrödinger's equation, which describes adequately the interaction of the imaging electrons represented by their wave function  $\psi$  with the periodic lattice potential of the crystal  $V(\mathbf{r})$  (a few volts) [9]. The proper equation is thus

$$\left( \frac{\hbar^2}{8\pi^2 m} \right) \Delta \psi + [E + V(\mathbf{r})]e\psi = 0 \quad (28)$$

where  $-e$  is the charge of the electron,  $E$  is the accelerating potential (about 100–400 kV), and the other symbols have their usual meanings. We separate out the constant part  $V_0$  of the lattice potential by introducing  $V'(\mathbf{r}) = V(\mathbf{r}) - V_0$  and  $E + V_0 = \hbar^2 k_0^2 / 2m$ . Furthermore, we define the reduced lattice potential  $U(\mathbf{r})$  as

$$U(\mathbf{r}) = \left( \frac{2me}{\hbar^2} \right) V'(\mathbf{r}) \quad (29)$$

with  $U_0 = 0$  since  $V_0$  was added to  $E$ . We expand  $U(\mathbf{r})$  in a Fourier series:

$$U(\mathbf{r}) = \sum_{\mathbf{g}} U_{\mathbf{g}} \exp(2\pi i \mathbf{g} \cdot \mathbf{r}) \quad (30)$$

The Schrödinger equation can then be rewritten as

$$\Delta \psi + 4\pi^2 k_0^2 \psi = -4\pi^2 U(\mathbf{r}) \psi \quad (31)$$

A solution of the equation without the right-hand side is the plane wave

$$\psi_0 = C \exp(2\pi i \mathbf{k}_0 \cdot \mathbf{r}) \quad (32)$$

Equation (32) represents the incident wave with the wavevector  $\mathbf{k}_0$  corrected for refraction by the constant part  $V_0$  of the lattice potential. The right-hand side of Eq. (31) describes the perturbation caused by the periodic part of the lattice potential which is small compared to  $V_0$ . The kinematical approximation, which is equivalent to the first Born approximation, now consists in substituting  $\psi$  by  $\psi_0$  in the right-hand side of Eq. (31), and solving the resulting equation:

$$\begin{aligned} \Delta \psi + 4\pi^2 k_0^2 \psi \\ = -4\pi^2 \sum_{\mathbf{g}} U_{\mathbf{g}} \exp[2\pi i (\mathbf{k}_0 + \mathbf{g}) \cdot \mathbf{r}] \end{aligned} \quad (33)$$

The solution of this equation is of the general form

$$\psi = \psi_0 + \sum_{\mathbf{g}} \psi_{\mathbf{g}} \quad (34)$$

that is, it consists of the solution of the equation without the right-hand side to which is added a particular solution of the complete equation which is of the same form as the right-hand side. Substituting Eq. (34) in Eq. (33) and remembering that  $\psi_0$  is a solution of the equation without the right-hand side, then  $\psi_{\mathbf{g}}$  must

satisfy the equations

$$\begin{aligned} \Delta \psi_{\mathbf{g}} + 4\pi^2 k_0^2 \psi_{\mathbf{g}} \\ = -4\pi^2 U_{\mathbf{g}} \exp[2\pi i (\mathbf{k}_0 + \mathbf{g}) \cdot \mathbf{r}] \end{aligned} \quad (35)$$

with the boundary condition  $\psi_{\mathbf{g}} = 0$  at the entrance face  $z = 0$ . Each equation is seen to involve one beam only; this is consistent with the kinematical approximation which implies that the scattered beams are all independent. Since we are looking for solutions that represent scattered beams, we use an 'ansatz':

$$\psi_{\mathbf{g}}(\mathbf{r}) = \varphi_{\mathbf{g}}(\mathbf{r}) \exp[2\pi i (\mathbf{k}_0 + \mathbf{g}) \cdot \mathbf{r}] \quad (36)$$

which has the same form as the right-hand side but an amplitude depending on  $\mathbf{r}$ ; it is expected that this dependence on  $\mathbf{r}$  will be a smooth variation.

Gevers [9] has shown that, substituting this expression in Eq. (35) leads, after some transformations, to

$$\begin{aligned} \Delta \varphi_{\mathbf{g}} + 4\pi i (\mathbf{k}_0 + \mathbf{g}) \cdot \text{grad } \varphi_{\mathbf{g}} \\ + 4\pi^2 [k_0^2 + (\mathbf{k}_0 + \mathbf{g})^2] \varphi_{\mathbf{g}} \\ = -4\pi^2 U_{\mathbf{g}} \end{aligned} \quad (37)$$

With the choice of axis adopted in Fig. 17, and denoting by  $\alpha$  the angle between

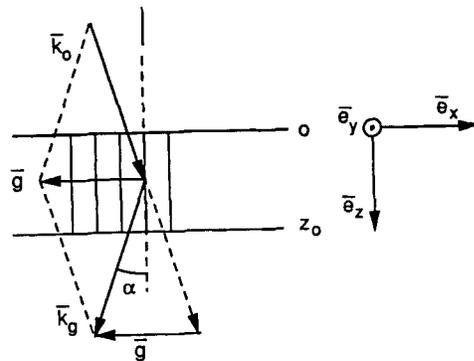


Figure 17. Reference system used in the discussion of the Born approximation.

$\mathbf{k}_0 + \mathbf{g}$  and  $\mathbf{e}_z$ , we obtain

$$\begin{aligned} \frac{\partial \varphi_g}{\partial z} + tg\alpha \frac{\partial \varphi_g}{\partial x} - 2\pi i s_g \varphi_g \\ = \frac{\pi i U_g}{|\mathbf{k}_0 + \mathbf{g}| \cos \alpha} \end{aligned} \quad (38)$$

we have hereby neglected the term  $\Delta \varphi_g$ , which is small compared to the gradient term because  $\varphi_g$  is a smoothly varying function and because the coefficient of the gradient term is large since from geometrical considerations follows that

$$s_g = \frac{k_0^2 - |\mathbf{k}_0 + \mathbf{g}|^2}{2|\mathbf{k}_0 + \mathbf{g}| \cos \alpha} \quad (39)$$

The right-hand side of Eq. (38) has the dimensions of inverse length, which still depends on  $\mathbf{g}$ ; we shall call it  $\pi i/t_g$ , where  $t_g$  is called the *extinction distance*. Noting that  $k_0 \gg \mathbf{g}$  we can put  $|\mathbf{k}_0 + \mathbf{g}| \approx k_0$  and obtain

$$t_g = \frac{h^2 k_0 \cos \alpha}{2me|V_g|} \quad (40)$$

and, reintroducing the accelerating voltage  $E$ , neglecting  $V_0 \ll E$  we obtain

$$t_g = \frac{\lambda E \cos \alpha}{|V_g|} \quad \text{with } \cos \alpha \approx 1 \quad (41)$$

In a plan parallel perfect foil, the scattered amplitude cannot depend on  $x$  and hence  $\partial \varphi_g / \partial x = 0$ . Moreover, even in a deformed foil it is a good approximation to assume that electrons propagate along narrow columns. This 'column' approximation justifies the neglect of the term in  $tg\alpha$ , which is very small. We then finally obtain as the basic equation for the kinematical theory:

$$\frac{d\varphi_g}{dz} - 2\pi i s_g \varphi_g = (\pi i/t_g) \exp(i\theta_g) \quad (42)$$

since  $V_g = |V_g| \exp(i\theta_g)$ .

This equation can readily be integrated between the front ( $z = 0$ ) and exit surface ( $z = t_0$ ) of the foil. Taking into account that  $\psi_g = 0$  for  $z = 0$ , this leads to

$$\begin{aligned} \varphi_g = (\pi i/t_g) [\exp(2\pi i s t_0)] [\exp(i\theta_g)] \\ \times \int_0^{t_0} \exp(-2\pi i s z) dz \end{aligned} \quad (43)$$

or explicitly, after some simplifications,

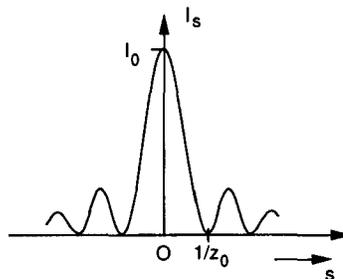
$$\varphi_g = i [\exp(\pi i s t_0)] [\exp(i\theta_g)] [\sin(\pi s t_0 / s t_g)] \quad (44)$$

For the scattered intensity per unit of incident intensity one thus finds

$$I_g = \psi_g \psi_g^* = \varphi_g \varphi_g^* = \frac{\sin^2(\pi s t_0)}{(s t_g)^2} \quad (45)$$

The variation of  $I_g$  with thickness is periodic with the period  $1/s$ , being zero at the entrance face and for  $t_0 = n/s$  (whose  $n$  is an integer), whereas maxima occur for  $t_0 = (n + \frac{1}{2})(1/s)$ .

The dependence of  $I_g$  on  $s$  is represented in Fig. 18. There is a pronounced central maximum (for  $s = 0$ ) given by  $I_g(\text{max}) = (\pi t_0 / t_g)^2$ . There are zeros for  $s = n/t_0$  ( $t_0 \equiv z_0$ ) and maxima approximately half-way between the zeros. These secondary maxima are much smaller than the central peak.



**Figure 18.** Variation of the diffracted intensity with the deviation parameters (foil thickness =  $t_0 \equiv z_0$ ).

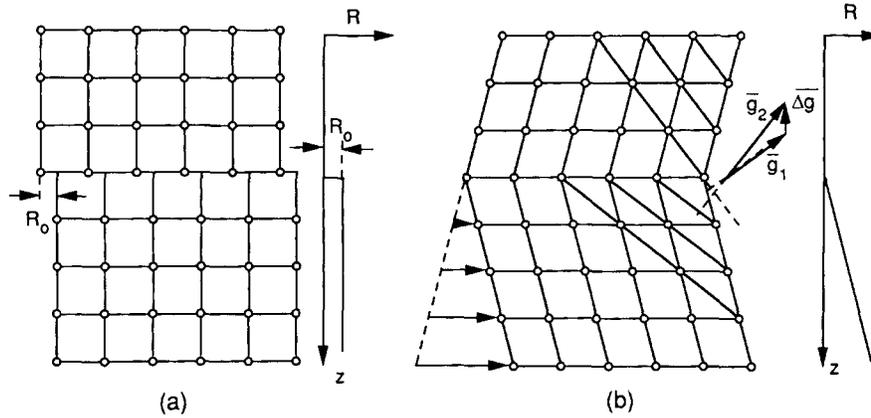


Figure 19. Displacement function of planar interfaces: (a) stacking fault; (b) domain boundary.

### Diffraction by Deformed Crystals: Models

Deformations of the crystal can be modeled by stating that the unit cell which in the undeformed crystal was in  $A_L$  now occupies, after deformation, the position  $A_L + \mathbf{R}(r)$ . The deformation field  $\mathbf{R}(r)$  characterizes the defect. A few simple examples of deformation fields are:

- (i) Planar translation interfaces such as stacking faults, out-of-phase boundaries, and discommensuration walls. All these planar defects have a displacement field of the type  $\mathbf{R} = 0$  for  $z \leq z_1$  and  $\mathbf{R} = \mathbf{R}_0$  for  $z > z_1$ , where  $z_1$  is the level at which the planar defect occurs behind the entrance face (Fig. 19a).
- (ii) Domain boundaries or twin boundaries with a small twinning vector. We now have  $\mathbf{R} = 0$  for  $z \leq z_1$  and  $\mathbf{R} = \mathbf{k}z$  for  $z > z_1$  (Fig. 19b).
- (iii) A pure screw dislocation has a displacement field described by  $\mathbf{R} = \mathbf{b}[\theta/2\pi]$ , where  $\theta$  is the azimuth angle, measured in the plane perpendicular

to  $\mathbf{b}$ . All displacements are clearly parallel to  $\mathbf{b}$ .

- (iv) A spherical inclusion has a radial, spherically symmetric displacement field:

$$\mathbf{R} = \frac{\varepsilon r_0^3 \mathbf{r}}{r^3} \quad (46)$$

for  $r \geq r_0$ ; for  $r < r_0$ ,  $\mathbf{R} = \varepsilon \mathbf{r}$  with  $\varepsilon = (2/3)\delta$ , where  $\delta$  is the lattice mismatch between inclusion and matrix.

Planar interfaces which are inclined with respect to the foil surface can be considered as consisting of 'steps' that are one column wide. Along a line perpendicular to the intersection line of the fault plane and the foil surface, the columns are assumed to contain a planar fault, parallel to the foil plane, at the level where the inclined fault plane intersects the column. Similarly, an inclined dislocation line is assumed to consist of small segments, each one column long, parallel to the foil plane. It is thus sufficient to consider defects in planes parallel to the foil surfaces.

### Scattered Amplitude for a Deformed Foil

#### *Kinematical Formulation*

In the case of a deformed foil, the scattered amplitude corresponding to the scattering vector  $\mathbf{h} = \mathbf{g} + \mathbf{s}$  becomes, from Eq. (8),

$$A(\mathbf{h}) = \sum_L F_g \exp[2\pi i(\mathbf{g} + \mathbf{s})] \cdot [A_L + R(\mathbf{r})] \quad (47)$$

or, replacing the summation by an integration as in Sec. 1.1.3.5 of this Chapter.

$$A(\mathbf{h}) = F_g \int_{\text{Column}} \exp\{2\pi i[\mathbf{g} \cdot \mathbf{R}(\mathbf{r}) + sz]\} dz \quad (48)$$

Hereby we have used the fact that  $\mathbf{g} \cdot \mathbf{A}_L$  is an integer and that  $\mathbf{s} \cdot \mathbf{R}(\mathbf{r})$  is much smaller than the other terms in the exponential. Putting

$$\alpha = 2\pi \mathbf{g} \cdot \mathbf{R}(\mathbf{r}) \quad (49)$$

we can write

$$A(\mathbf{h}) = F_g \int_0^{t_0} \exp[i\alpha(z)] \exp[2\pi isz] dz \quad (50)$$

#### *Stacking Fault Contrast*

Let the fault plane be parallel with the foil planes at  $z = z_1$ , behind the entrance face. Since  $\mathbf{R}$  is constant in this case, the  $\alpha$  is also constant and we can split the integral of Eq. (50) into two parts [10]:

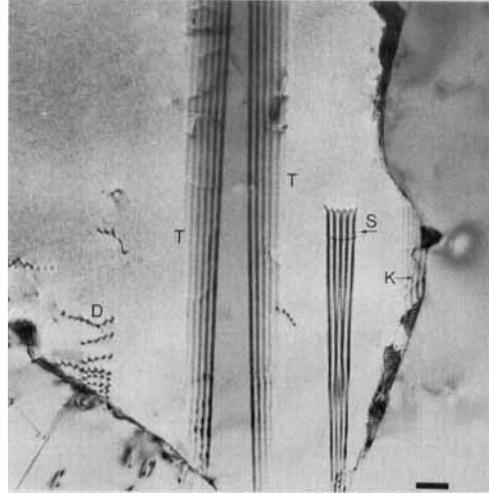
$$\begin{aligned} \frac{A(\mathbf{h})}{F_g} &= \int_0^{z_1} \exp(2\pi isz) dz \\ &+ \exp(i\alpha) \int_{z_1}^{t_0} \exp(2\pi isz) dz \end{aligned} \quad (51)$$

or, after evaluating the integrals and computing  $I(\mathbf{h}) = A(\mathbf{h}) A^*(\mathbf{h})$

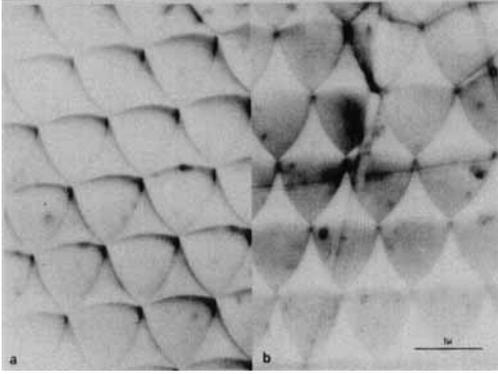
$$\begin{aligned} \frac{I(\mathbf{h})}{F_g^2} &= \{1 - \cos(\alpha + \pi st_0) \cos(\pi st_0) \\ &+ \cos(2\pi su)[\cos(\alpha + \pi st_0) \\ &- \cos \pi st_0]\} / (\pi s)^2 \end{aligned} \quad (52)$$

with  $u = \frac{1}{2}(2z_1 - t_0)$ , that is  $u$  is the distance counted from the central plane of the foil.

The intensity  $I(\mathbf{h})$  clearly depends periodically on the thickness  $t_0$  of the foil as well as on the level of the fault in the foil (i.e. on  $u$ ). For an inclined fault in a foil of constant thickness, the intensity  $I(\mathbf{h})$  is clearly a periodic function of  $u$  with period  $1/s$ ; it is symmetrical in  $u$  since the cosine is an even function. An electron micrograph will produce a projection of this intensity distribution, that is a fringe pattern with depth period  $1/s$ , the lateral extent of which is confined to the projected width of the fault (Fig. 20).



**Figure 20.** Fringe pattern due to a stacking fault (S), wedge fringes at the twin (T) and dislocations (D) (stainless steel). (Courtesy of J. Van Landuyt.)

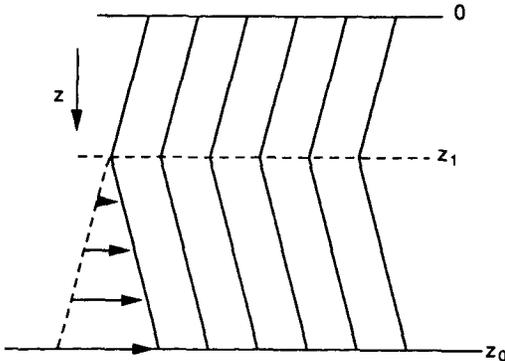


**Figure 21.** Uniformly shaded area due to the presence of a stacking fault parallel to the foil plane (graphite): (a) partial dislocations in contrast; (b) stacking fault areas show up as bright areas. (Courtesy of P. Delavignette.)

If the fault plane is parallel to the foil surfaces a region of uniform shade is produced in the fault area. This shade can be either brighter or darker than the perfect area of the foil (Fig. 21).

#### Domain Boundary Contrast

Let the domain boundary be parallel to the foil surfaces and situated at  $z_1$  (Fig. 22). We then have  $\alpha = 0$  for  $z \leq z_1$  and  $\alpha = 2\pi \mathbf{g} \cdot \mathbf{k} z$  for  $z_1 < z \leq t_0$ . The integral [Eq. (50)] can again be split into two



**Figure 22.** Model for a domain boundary.

parts. Recalling that  $\mathbf{g} \cdot \mathbf{k} = \Delta s$ , we have

$$\frac{A(\mathbf{h})}{F_h} = \int_0^{z_1} \exp(2\pi i s z) dz + \int_{z_1}^{z_0} \exp[2\pi i (s + \Delta s) z] dz \quad (53)$$

Since the two crystal parts on both sides of the surface are perfect, but slightly mis-oriented,  $s$  and  $\Delta s$  are constant. These integrals can easily be evaluated explicitly.

#### Dislocation Contrast

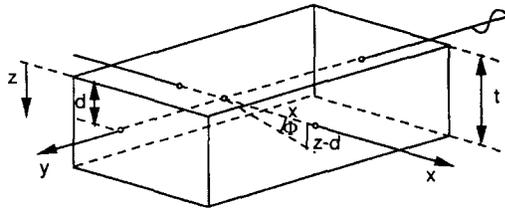
We adopt the geometry shown in Fig. 23, that is the dislocation line is parallel to the foil plane ( $x, y$ ), with the dislocation lying along the  $y$  axis [11]. We then have  $\theta = \arctg(z/x)$  and, since  $\mathbf{R} = \mathbf{b}(\theta/2\pi)$ , from Eq. (49)

$$\alpha = (\mathbf{g} \cdot \mathbf{b}) \arctg\left(\frac{z}{x}\right) \quad (54)$$

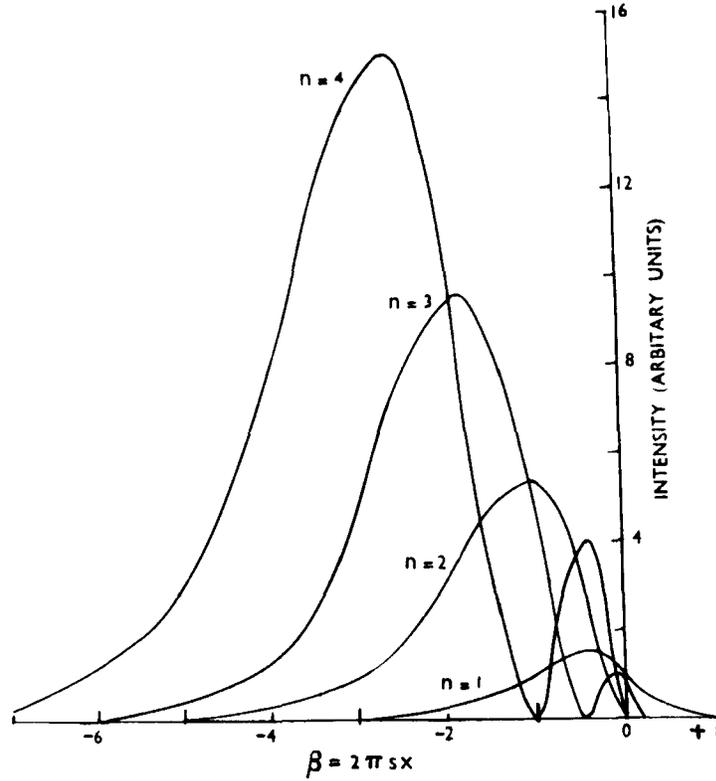
where  $\mathbf{g} \cdot \mathbf{b} = n$  is an integer for a perfect dislocation but a fraction for a partial dislocation. The scattered amplitude is given by

$$\frac{A(\mathbf{h})}{F_g} = \int_{-z_1}^{+z_2} \exp(2\pi i s z) \cdot \exp\left[in \arctg\left(\frac{z}{x}\right)\right] dz \quad (55)$$

The integral can be evaluated for different values of  $n$  but the results are not



**Figure 23.** Foil containing a screw dislocation; the notation used.



**Figure 24.** Image profiles for screw dislocations according to the kinematical theory [13, 14].

elementary functions. The numerical results are represented graphically in Fig. 24 for different values of  $n$ , as a function of  $\beta = 2\pi s x$ . It turns out that the peak height, peak shift and peak width increase with increasing  $n$  [11]. It is clear that no image will be produced if  $n = 0$  because then Eq. (55) reduces to that for a perfect foil. The corresponding profiles for edge dislocations are represented in Fig. 25 [12].

**Lattice Potential of Deformed Crystals**

Deformation can also be introduced in the diffraction equations using the ‘deformable ion approximation’ via the lattice potential [13, 14]. It is assumed that a

displaced volume element of crystal carries along with it the lattice potential it had before the deformation

$$V_{\text{def}}(\mathbf{r}) = V_{\text{undef}}(\mathbf{r} - \mathbf{R}) \tag{56}$$

As a result the Fourier coefficients of the lattice potential in the deformed crystal acquire phase factors since

$$\begin{aligned} V_{\text{def}}(\mathbf{r}) &= V_0 + \sum_g V_g \exp[2\pi i \mathbf{g} \cdot (\mathbf{r} - \mathbf{R})] \\ &= V_0 + \sum_g [V_g \exp(-2\pi i \mathbf{g} \cdot \mathbf{R})] \\ &\quad \times \exp(2\pi i \mathbf{g} \cdot \mathbf{r}) \end{aligned} \tag{57}$$

Strictly speaking, this series is no longer a Fourier series, which is consistent with the fact that the function  $V_{\text{def}}(\mathbf{r})$  is no longer a

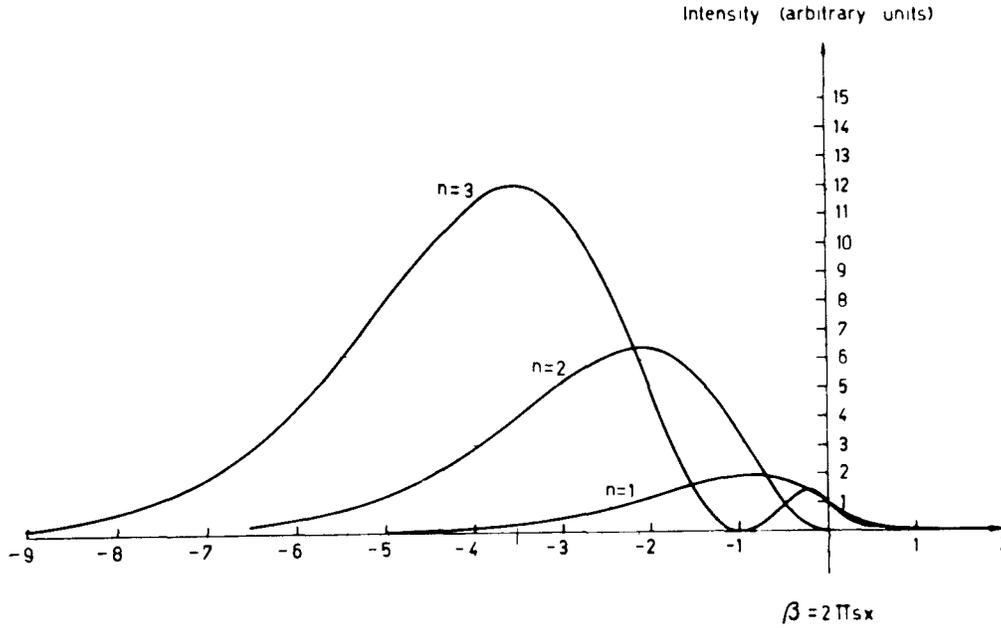


Figure 25. Image profiles for edge dislocations according to the kinematical theory [12].

periodic function. Accepting nevertheless this approximation, the presence of a deformation field can be accounted for by replacing  $V_g$  by

$$V_g \rightarrow V_g \exp[-2\pi i \mathbf{g} \cdot \mathbf{R}(\mathbf{r})] \quad (58)$$

In view of the relations between  $V_g$  and  $1/t_g$  this also amounts to replacing  $1/t_g$  by  $(1/t_g) \exp(-i\alpha_g)$ , with  $\alpha_g = 2\pi \mathbf{g} \cdot \mathbf{R}(\mathbf{r})$ .

For a deformed foil, Eq. (42) thus becomes

$$\begin{aligned} \frac{d\phi_g}{dz} - 2\pi i s_g \phi_g \\ = \left(\frac{\pi i}{t_g}\right) \exp(i\theta_g) \exp(-i\alpha_g) \end{aligned} \quad (59)$$

Introducing a new function  $\varphi_g = \phi_g \exp(-i\alpha_g)$  which does not change the scattered intensity  $I_g = \varphi_g \varphi_g^* = \phi_g \phi_g^*$  this

equation becomes

$$\begin{aligned} \frac{d\varphi_g}{dz} - 2\pi i \left( s_g + \mathbf{g} \cdot \frac{d\mathbf{R}}{dz} \right) \varphi_g \\ = \left(\frac{\pi i}{t_g}\right) \exp(i\theta_g) \end{aligned} \quad (60)$$

This equation is of the same form as Eq. (42), except that  $s_g$  has been replaced by a local value

$$s_{\text{eff}} = s_g + \mathbf{g} \cdot \frac{d\mathbf{R}}{dz} \quad (61)$$

which is, in general, a function of  $x, y$ , and  $z$ .

Depending on the problem to be treated, one has to use Eq. (60) or Eq. (59); if  $\alpha_g$  is a constant, Eq. (60) cannot be used since it reduces to that of a perfect crystal.

### 1.1.1.5 Two-Beam Dynamical Theory [13–15]

#### Basic Equations for Perfect Crystals

Rather than directly deriving the fundamental equations of the dynamical theory from Schrödinger's equation, we shall use a physically transparent method, related to the one originally used by Darwin [16]. We introduce the usual approximations (high energy approximation, column approximation, etc.) already in the model, rather than in the final equations.

The electrons traveling down a column of crystal can be represented by the wavefunction:

$$\psi(\mathbf{r}) = \phi_0(z) \exp(2\pi i \mathbf{k}_0 \cdot \mathbf{r}) + \phi_g(z) \exp(2\pi i \mathbf{k} \cdot \mathbf{r}) \quad (62)$$

where the  $\mathbf{k}$  and  $\mathbf{k}_0$  represent the wavevectors of the scattered and of the incident beam, respectively; the amplitudes  $\phi_g$  and

$\phi_0$  of these beams depend on  $z$ . The column approximation justifies to neglect the  $x$  dependence of these amplitudes. We shall write down the amplitude after interaction with a slice  $dz$  of the column, at level  $z$  behind the entrance face (Fig. 26).

We can write

$$\phi_0(z + dz) = \phi_0(z) \phi_0(dz) + \phi_g(z) \phi_{-g}(dz) \quad (63)$$

where  $\phi_{-g}$  means that scattering with scattering vector  $-\mathbf{g}$  has to be taken into account. This relation states that the transmitted beam results from the interference between the twice transmitted beam of which the amplitude is  $\phi_0(z) \phi_0(dz)$  and the twice-scattered beam with amplitude  $\phi_g(z) \phi_{-g}(dz)$ . Since the slices  $dz$  are arbitrarily thin, it is justified to use the kinematical approximation for  $\phi_0(dz) = 1$  and  $\phi_g(dz)$ . Note that changing  $\mathbf{g} \rightarrow -\mathbf{g}$  implies changing  $s \rightarrow -s$ , since the origin of reciprocal space is now in the node point G. From Eq. (43) we can conclude that

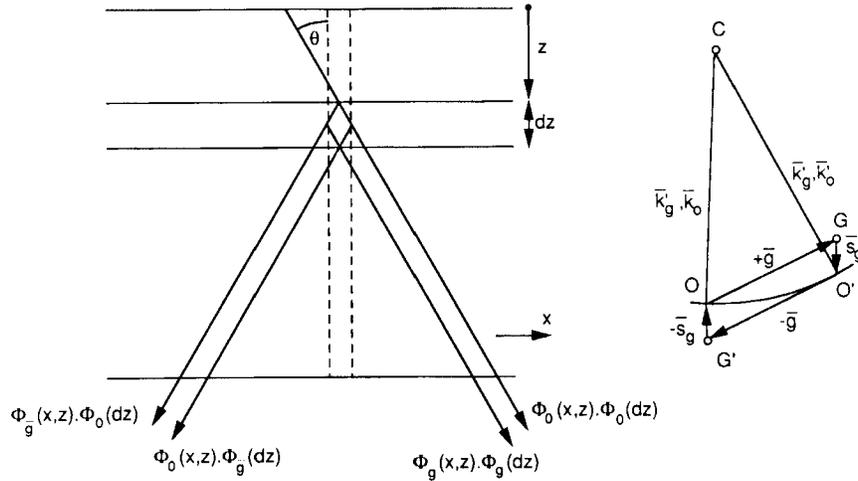


Figure 26. Illustrating the derivation of the fundamental equations of the dynamical theory.

(ignoring the phase factor of the Fourier coefficient)

$$\begin{aligned}\phi_g(dz) &= d\phi_g \\ &= \left(\frac{\pi i}{t_g}\right) \exp(-2\pi i s z) dz\end{aligned}\quad (64)$$

Furthermore, we have  $\phi_0(dz) = 1$  and hence

$$\begin{aligned}\phi_0(z + dz) &= \phi_0(z) + \phi_g(z) \left(\frac{\pi i}{t_{-g}}\right) \\ &\quad \times \exp(2\pi i s z) dz\end{aligned}\quad (65)$$

or

$$\begin{aligned}\phi_0(z + dz) - \phi_0(z) &= \phi_g(z) \left(\frac{\pi i}{t_{-g}}\right) \exp(2\pi i s z) dz\end{aligned}\quad (66)$$

and in the limit for  $dz \rightarrow 0$ :

$$(d\phi_0/dz) = \left(\frac{\pi i}{t_{-g}}\right) [\exp(2\pi i s z)] \phi_g(z) \quad (67)$$

Similarly, for the scattered beam we can write

$$\begin{aligned}\phi_g(z + dz) &= \phi_0(z) \phi_g(dz) \\ &\quad + \phi_g(z) \phi_0(dz)\end{aligned}\quad (68)$$

and taking into account the expressions for  $\phi_g(dz)$  [Eq. (64)] and  $\phi_0(dz) = 1$ , and making the same transformations as for  $\phi_0$ , we obtain:

$$d\phi_g/dz = \left(\frac{\pi i}{t_g}\right) [\exp(-2\pi i s z)] \phi_0(z) \quad (69)$$

Equations (67) and (69) form the Darwin–Howie–Whelan set of coupled differential equations describing the interplay between incident and scattered beams in a perfect crystal. This set can be given a number of alternative forms by introducing amplitudes differing only by phase factors

from  $\phi_0$  and  $\phi_g$ . This does not change the resulting intensity distributions.

One can, for instance, substitute  $T'$  and  $S'$  for  $\phi_0$  and  $\phi_g$

$$\phi_0 = T' \exp(\pi i s z) \quad (70a)$$

$$\phi_g = S' \exp(-\pi i s z) \quad (70b)$$

The system of equations then acquires a symmetrical form

$$dT'/dz + \pi i s T' = \left(\frac{\pi i}{t_{-g}}\right) S' \quad (71a)$$

$$dS'/dz - \pi i s S' = \left(\frac{\pi i}{t_g}\right) T' \quad (71b)$$

or one can make the substitution

$$\phi_0 = T \quad (72a)$$

$$\phi_g = S \exp(-2\pi i s z) \quad (72b)$$

and obtain the following asymmetrical set of equations:

$$dT/dz = \left(\frac{\pi i}{t_{-g}}\right) S \quad (73a)$$

$$dS/dz = 2\pi i s S + \left(\frac{\pi i}{t_g}\right) T \quad (73b)$$

Depending on the problem to be solved, either the symmetrical or the asymmetrical set will have to be used.

### Dynamical Equations for Deformed Crystals

We have seen previously (Sec. 1.1.4.4 of this Chapter) that the deformation of the lattice can be modeled by using a local effective deviation parameter,  $s_{\text{eff}}$ , which is a function of  $r$ , rather than a constant  $s$ .

We found that  $s_{\text{eff}}$  is given by

$$s_{\text{eff}} = s + \frac{d}{dz}(\mathbf{g} \cdot \mathbf{R}) \quad (74a)$$

$$s_{\text{eff}} = s + \frac{d\alpha}{dz} \quad (74b)$$

with  $\alpha = \alpha_g/2\pi$ . One can alternatively replace  $1/t_g$  by  $1/t_g[\exp(-i\alpha_g)]$ . This can be shown directly by intuitive considerations [17].

These substitutions allow us to adapt the different sets of equations to the case of deformed crystals. The presence of a defect is now described by the variation of  $s$  along the integration columns, the diffraction vector  $\mathbf{g}$  being the same all along the column.

### Solution of the Two-beam Dynamical Equations for a Perfect Crystal

The two equations of the system [Eqs. (73a) and (73b)] can be uncoupled by eliminating  $S$  and  $T$ , respectively. For  $S$ , one obtains for instance

$$\frac{d^2 S}{dz^2} - 2\pi i s_g \frac{dS}{dz} + \frac{\pi^2}{t_g^2} \cdot S = 0 \quad (75)$$

and similarly for an equation for  $T$ .

Taking into account the initial values  $T = 1$ ,  $S = 0$  for  $z = 0$ , and requiring that the solutions satisfy both equations of the set, one finds

$$T = \left[ \cos(\pi\sigma_g z) - i \left( \frac{s_g}{\sigma_g} \right) \sin(\pi\sigma_g z) \right] \times \exp(\pi i s_g z) \quad (76a)$$

$$S = (i/\sigma_g t_g) [\sin(\pi\sigma_g z) \exp(\pi i s_g z)] \quad (76b)$$

with

$$\sigma_g^2 = \frac{(1 + s^2 t_g^2)}{t_g^2} \quad (77)$$

The scattered intensity  $I_S$  is then

$$I_S = SS^* = \sin^2 \left[ \frac{\pi\sigma_g z}{(\sigma t_g)^2} \right] \quad (78)$$

and

$$I_T = 1 - I_S \quad (79)$$

since we have, so far, neglected absorption.

The depth periods of both  $I_S$  and  $I_T$  are now  $1/\sigma_g$  as compared to  $1/s_g$ , according to the kinematical theory. For  $s_g = 0$  the depth period remains finite and equal to  $t_g$ . Note that for large values of  $s_g$ ,  $\sigma_g$  reduces to  $s_g$ .

Equations (78) and (79) as well as Eqs. (67) and (69) describe the periodic transfer of electrons from the transmitted beam into the scattered beam, and vice versa (Pendellösung effect). The periodic variations of  $I_S$  (or  $I_T$ ) with the crystal thickness give rise to the formation of thickness extinction contours in wedge-shaped crystals (Fig. 27). The contours are the geometrical loci of equal thickness. According to Eqs. (78) and (79) such fringes should be periodic; in actual fact the fringes are found to be damped with increasing thickness, showing that absorption takes place. We take absorption into account below.

The expressions  $I_S$  and  $I_T$  also describe the dependence on  $s_g$ . The geometrical loci with  $s_g = \text{Constant}$  are the equi-inclination or bent contours. According to the expressions for  $I_S$  and  $I_T$ , this dependence should be symmetrical in  $s$ , i.e.  $I_S(-s) = I_S(+s)$  and  $I_T(-s) = I_T(+s)$ . In fact, one finds that  $I_T(-s) < I_T(+s)$ , whereas  $I_S(-s) = I_S(+s)$ . This effect, called

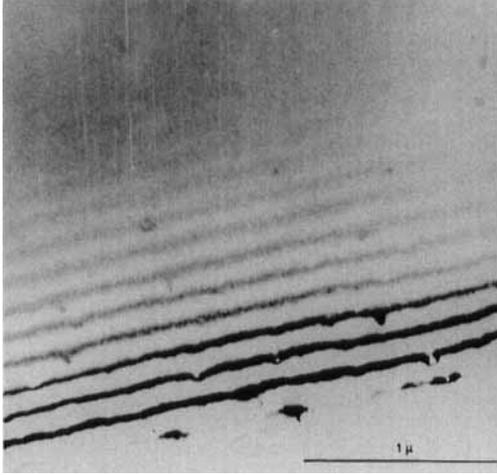


Figure 27. Thickness contours in a silicon wedge.

the *Borrmann effect*, is also a consequence of anomalous absorption (see below).

#### Two-beam Lattice Fringes

In the two-beam case the total wavefunction at the exit face of a crystal foil can be written as [18–20]

$$\Psi = \exp(2\pi i \mathbf{K} \cdot \mathbf{r}) [T + S \exp(2\pi i \mathbf{g} \cdot \mathbf{r})] \quad (80)$$

where  $\mathbf{K}$  is the wavevector of the incident electron beam corrected for refraction by the constant part,  $V_0$ , of the lattice potential. The total wavefunction  $\Psi$  can be imaged in the microscope by collecting the two interfering beams in the objective aperture and focusing on the exit face of the foil. One then observes the intensity distribution

$$\begin{aligned} I &= \Psi \Psi^* \\ &= TT^* + SS^* + T^* S \exp(2\pi i \mathbf{g} \cdot \mathbf{r}) \\ &\quad + TS^* \exp(-2\pi i \mathbf{g} \cdot \mathbf{r}) \end{aligned} \quad (81)$$

$$= I_T + I_S + 2\sqrt{I_T I_S} \sin(2\pi i \mathbf{g} \cdot \mathbf{r} + \varphi) \quad (82)$$

where

$$\text{tg } \varphi = \frac{s_g}{\sigma_g} \text{tg}(\pi \sigma_g z_0) \quad (83)$$

$z_0$  being the foil thickness. The image consists of sinusoidal fringes with a period  $1/|\mathbf{g}|$ , which is equal to the interplanar distance of the lattice planes normal to  $\mathbf{g}$ . The fringes are parallel to the traces of these lattice planes. For  $s_g = 0$ , Eq. (82) simply reduces to

$$I = 1 + \sin\left(\frac{2\pi z_0}{t_g}\right) \sin(2\pi \mathbf{g} \cdot \mathbf{x}) \quad (84)$$

where the  $x$  axis was chosen along  $\mathbf{g}$ .

The contrast of the fringes clearly depends on  $z_0$ , whereas the localization of the fringes depends on  $s_g$ , since this determines  $\varphi$ . For  $s = 0$ ,  $\varphi = 0$ , the bright fringes coincide with the lattice planes  $g x = \text{Integer} + \frac{1}{2}$ . The fringes are therefore called *lattice fringes*; their formation can be understood on a purely geometrical basis [21].

The phase  $\varphi$  in Eq. (82) will, in general, also depend on the exact imaging conditions. Since the beams  $T$  and  $S$  may enclose different angles with the optical axis of the microscope, additional phase shifts will, in general, occur. As long as we consider diffraction contrast images, the image transfer function of the microscope is not important since the intensity distribution in a single beam does not depend on its shape. However, when considering the interference of more than one beam, the function plays an important role and cannot be ignored.

#### 1.1.1.6 Absorption

Absorption can be described phenomenologically by assuming the refractive index

$n = c/v$  (where  $c$  is the velocity in a vacuum, and  $v$  is the velocity in a medium) to be complex. Representing a plane wave as

$$\psi = \psi_0 \exp[i(kz - \omega t)] \quad (85)$$

with  $k = \omega/v = \omega n/c$  and making  $n$  complex  $n_c = n + i\mu$  leads to a complex wavevector

$$k_c = \frac{\omega}{c}(n + i\mu) \quad (86)$$

and to a wave

$$\begin{aligned} \psi &= \psi_0 \exp[i(k_c z - \omega t)] \\ &= \psi_0 \exp(-\omega\mu z/c) \exp[i(kz - \omega t)] \end{aligned} \quad (87)$$

which is clearly damped, the amplitude absorption coefficient being  $\mu\omega/c$ .

Replacing  $V_0$  by  $V_0 + iW_0$  in the lattice potential has the desired effect, since it makes the wavevector  $K$  complex:

$$K^{*2} = \frac{2me[(E + V_0) + iW_0]}{h^2} \quad (88)$$

$$= K^2 + i[2meW_0/h^2 K]K \quad (89)$$

We now define

$$\frac{1}{\tau_0} = 2meW_0/h^2 K \quad (90)$$

by analogy with

$$\frac{1}{t_0} = 2meV_0/h^2 K \quad (91)$$

The complex wavevector  $K^*$  then becomes

$$K^{*2} = K^2 + i\left(\frac{K}{\tau_0}\right) \quad (92)$$

and, since  $1/K\tau_0 \ll 1$ , then

$$\begin{aligned} K^* &= K \left(1 + \frac{i}{K\tau_0}\right)^{1/2} \\ &\simeq K \left(1 + \frac{i}{2K\tau_0}\right) \end{aligned}$$

or, finally,

$$K^* = K + \frac{i}{2\tau_0} \quad (93)$$

The boundary conditions now require that at the entrance face  $z = 0$  the tangential component must be conserved, i.e.  $K_t^* = K_t$ . The imaginary part being zero for  $z < 0$ , this means that for  $z > 0$  the imaginary part must be oriented along  $e_n$ , i.e.

$$K^* = K + \left(\frac{i}{2\tau_0}\right)e_n \quad (94)$$

The expression  $\exp(2\pi i \mathbf{K} \cdot \mathbf{r})$  then becomes

$$\exp(2\pi i \mathbf{K} \cdot \mathbf{r}) \exp\left(\frac{-\pi z}{\tau_0}\right) \quad (95)$$

since  $e_n \cdot \mathbf{r} = z$ . The absorption coefficient for the amplitude is thus  $\mu = \pi/\tau_0$ .

It has been shown by Yosioka [22] that anomalous absorption can be taken into account by assuming the lattice potential to become complex:  $V(\mathbf{r}) + iW(\mathbf{r})$ . We have shown that this applies to the constant term  $V_0 + iW_0$ , and leads to normal absorption in that case. This procedure can be generalized by replacing also the other Fourier coefficients by complex quantities:  $V_g \rightarrow V_g + iW_g$ . In view of the relation between  $1/t_g$  and  $V_g$  [Eq. (41)], this is equivalent to replacing

$$\frac{1}{t_g} \rightarrow \frac{1}{t_g} + \frac{i}{\tau_g} \quad (96)$$

where we have introduced the *absorption lengths*,  $\tau_g$ , which are related to  $W_g$  in the same way as  $t_g$  are related to  $V_g$ :

$$\frac{1}{\tau_g} = \frac{2meW_g}{h^2 k} \quad (97)$$

$\tau_g \gg t_g$ , since  $W_g \ll V_g$ . If  $1/t_g$  becomes complex it is clear that  $\sigma_g$  [defined in

Eq. (77)] must also become complex

$$\sigma_g = \sigma_r + i\sigma_i \quad (98)$$

Omitting, for simplicity, the index  $g$  at  $\sigma_r$  and  $\sigma_i$ , one finds a good approximation, noting that  $\sigma_i^2 \ll \sigma_r^2$  and  $1/\tau_g^2 \ll 1/t_g^2$ :

$$\sigma_r = (1/t_g)[1 + s_g^2 t_g^2]^{1/2} \quad (99a)$$

$$\sigma_i = (\sigma_r t_g \tau_g)^{-1} \quad (99b)$$

We note that  $\sigma_i \cdot \mathbf{r} = \sigma_i z$ .

### 1.1.1.7 Dynamical Equations Including Absorption

Anomalous absorption effects can be taken into account in the basic equations [Eqs. (67) and (69)] or in the equations derived from these, by the above-mentioned substitutions [Eqs. (96) or (98)]. Equations (67) and (69) then become, for instance,

$$\frac{d\phi_0}{dz} = \pi i \left[ \frac{1}{t_{-g}} + i \left( \frac{1}{\tau_{-g}} \right) \right] \exp(2\pi i s_g z) \phi_g(z) \quad (100)$$

$$\frac{d\phi_g}{dz} = \pi i \left[ \frac{1}{t_g} + i \left( \frac{1}{\tau_g} \right) \right] \exp(-2\pi i s_g z) \phi_0(z) \quad (101)$$

The substitution can also simply be performed directly in the final solutions [Eq. (76)]. Although the method is purely phenomenological, the result allows a physically meaningful interpretation. Moreover, in order to obtain tractable analytical solutions one adopts the following approximation:  $\sigma$  is replaced by  $\sigma_r + i\sigma_i$  in the periodic terms of the Eqs. (76a) and (76b), but in the coefficients we

replace  $\sigma$  by  $\sigma_r$ . This approximation preserves the essential features of anomalous absorption. It allows to show that the wavefunction,  $\psi_T$ , of the transmitted beam results from the interference of two waves with slightly different wavevectors ( $\mathbf{e} =$  unit normal to the foil)

$$\mathbf{K} + \frac{1}{2}(s + \sigma_r)\mathbf{e} \quad (102a)$$

and

$$\mathbf{K} + \frac{1}{2}(s - \sigma_r)\mathbf{e} \quad (102b)$$

For the wavefunction,  $\psi_S$ , of the scattered beam, the two interfering waves have wavevectors

$$\mathbf{K} + \mathbf{g} + \frac{1}{2}(s + \sigma_r)\mathbf{e} \quad (103a)$$

and

$$\mathbf{K} + \mathbf{g} + \frac{1}{2}(s - \sigma_r)\mathbf{e} \quad (103b)$$

The beating of these two waves causes the periodic depth variation of  $\psi_T$  and  $\psi_S$  with a period  $1/\sigma_r$ , that is the Pendellösung effect.

The four waves present in the total wavefunction,  $\psi = \psi_T + \psi_S$ , belong to two Bloch wavefields,  $\psi_1$  and  $\psi_2$ . For the waves of the first wavefield,  $\psi_1$ , the sign of  $\sigma_r$  is positive; it contains the two waves with wavevectors  $\mathbf{K} + \frac{1}{2}(s + \sigma_r)\mathbf{e}$  and  $\mathbf{K} + \mathbf{g} + \frac{1}{2}(s + \sigma_r)\mathbf{e}$ . The second wavefield,  $\psi_2$ , corresponds to the negative sign of  $\sigma_r$ ; it contains the two waves with wavevectors  $\mathbf{K} + \frac{1}{2}(s - \sigma_r)\mathbf{e}$  and  $\mathbf{K} + \mathbf{g} + \frac{1}{2}(s - \sigma_r)\mathbf{e}$ . In the simple case  $s = 0$ , the wavefield  $\psi_1$  has a maximum amplitude in planes coinciding with the atomic planes, whereas  $\psi_2$  has maximum amplitude along planes exactly in between the atomic planes [13, 14].

The first type of wavefield which is strongly excited for  $s < 0$  is strongly attenuated, whereas the second wavefield, which is more strongly excited for  $s > 0$ ,

is less attenuated by anomalous absorption. These results are physically meaningful. Electrons propagating close to atomic cores, that is electrons for which the wavefunction is peaked at the atomic positions, will have a larger probability to excite X-rays and hence to be 'absorbed' than will electrons passing between atomic planes.

The damping of the thickness fringes in a wedge-shaped crystal can be understood on the basis of these results. The depth variations of  $I_S$  and  $I_T$  are in both cases caused by the beating of two waves, one of each wavefield. Since one of these waves is attenuated much more than the other, the beating envelope (which has a maximum amplitude when the two beating waves have equal amplitude) decreases with depth in the crystal, even though one of the waves has still an appreciable amplitude.

The Borrmann effect [23] can also be understood on the basis of the same model. For  $s > 0$  the amplitude of the rapidly attenuated wave in  $\psi_T$ , that is  $\frac{1}{2}[1 - (s/\sigma_2)]$ , is smaller than that of the passing wave, that is  $(\frac{1}{2}[1 + (s/\sigma_2)])$ , which is enhanced by anomalous absorption. As a result  $\psi_T$  will have a larger amplitude for  $s > 0$  than for  $s < 0$  for the same absolute value of  $s$ . A similar asymmetry is absent for  $\psi_S$ , where the amplitudes of the two constituent waves are both  $1/(2\sigma t_g)$ , which only depends on  $s^2$ .

### 1.1.1.8 Rocking Curves for Perfect Crystals Taking into Account Anomalous Absorption

Explicit expressions for  $I_T$  and  $I_S$  are obtained by computing  $I_T = \psi_T \cdot \psi_T^*$  and

$I_S = \psi_S \cdot \psi_S^*$ . Since  $\sigma_i \ll \sigma_r$  and  $\tau_g \gg t_g$  one can approximate the expressions by neglecting higher order terms in  $t_g/\tau_g$ . After lengthy but straightforward calculations, one obtains

$$I_T = [\cosh u + (s/\sigma_r) \sinh u]^2 - \left[ \frac{1}{(\sigma_r t_g)^2} \right] \sin^2 v \quad (104)$$

with  $u = \pi\sigma_i z$ ,  $v = \pi\sigma_r z$ ,  $\sigma_i = (\sigma_r t_g \tau_g)^{-1}$ , and  $\sigma_r = [1 + (s t_g)^2]^{1/2}/t_g$ . Similarly,

$$I_S = \frac{\sinh^2 u + \sin^2 v}{(\sigma_r t_g)^2} \quad (105)$$

These expressions are represented in Fig. 28 for a rather thin foil. Note the asymmetry of  $I_T$  with respect to  $s$  (Borrmann effect), even for a rather thin foil.

In the limit  $\sigma_i \rightarrow 0$ , that is as  $u \rightarrow 0$ , one obtains the corresponding expressions for the non-absorption case:

$$I_S = \frac{\sin^2 \pi\sigma_r z}{(\sigma_r t_g)^2} \quad (106a)$$

and

$$I_T = 1 - I_S \quad (106b)$$

For  $s = 0$  one obtains

$$I_S = \sinh^2 u + \sin^2 v \quad (107a)$$

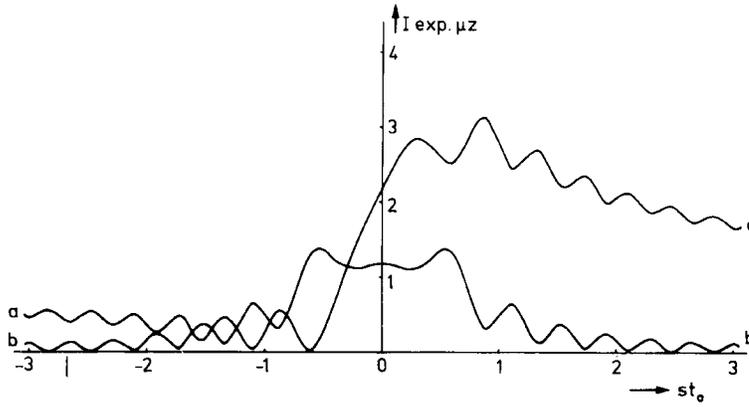
and

$$I_T = \cosh^2 u - \sin^2 v \quad (107b)$$

Note that now

$$I_T + I_S = \sinh^2 u + \cosh^2 u > 1!$$

This apparently absurd result is due to the fact that *normal* absorption, which attenuates both beams to the same extent, has been neglected. It can readily be introduced by multiplying by  $\exp(-\pi z_0/\tau_0)$  [from Eq. (95)].



**Figure 28.** Rocking curves for a perfect crystal taken into account in anomalous absorption. The transmitted intensity  $I_T$  (a) is asymmetrical in  $s \cdot t_g$ , whereas the scattered intensity  $I_S$  (b) is symmetrical in  $s \cdot t_g$  [12].

The expressions for  $I_S$  and  $I_T$  can be rewritten ( $s = 0$ ) as

$$I_{T,S} = \frac{1}{2}[\cosh 2u \pm \cos 2v] \quad (108)$$

where the plus sign refers to the transmitted beam and the minus sign refers to the scattered beam.

### 1.1.1.9 Dynamical Diffraction by Deformed or Faulted Crystals [13, 14, 24]

#### Basic Equations

Quantitative studies of defects in crystals are usually performed under optimized two-beam conditions, since this allows the most straightforward and detailed interpretation. The system of two-beam equations which lends itself most conveniently to the description of diffraction by non-perfect crystals is the Darwin–Howie–Whelan system of coupled differential equations.

Different formulations of this system are available; their solutions for the

amplitudes of scattered ( $S$ ) and transmitted ( $T$ ) beams differ by phase factors only, and hence lead to the same intensity distribution, that is to the same image. We shall use two different forms of this system, depending on the problem to be treated.

The asymmetric form [from Eq. (73)]:

$$\frac{dT}{dz} = \left(\frac{\pi i}{t-g}\right) S \quad (109a)$$

$$\frac{dS}{dz} = 2\pi i s_g S + \left(\frac{\pi i}{t_g}\right) T \quad (109b)$$

and the symmetric form [from Eq. (71)]

$$\frac{dT'}{dz} + \pi i s_g T' = \left(\frac{\pi i}{t-g}\right) S' \quad (110a)$$

$$\frac{dS'}{dz} - \pi i s_g S' = \left(\frac{\pi i}{t_g}\right) T' \quad (110b)$$

The amplitudes  $T'$ ,  $S'$ ,  $T$ , and  $S$  are related as follows:

$$T = T' \exp(\pi i s_g z) \quad (111a)$$

$$S = S' \exp(-\pi i s_g z) \quad (111b)$$

We have shown in Sec. 1.1.4.4 of this Chapter that deformation can be

introduced in the diffraction equations by applying the deformable ion approximation, a defect being modeled by means of a vector field  $\mathbf{R}(\mathbf{r})$ , called its *displacement field*. It was also shown that the Fourier coefficients of the lattice potential then become functions of  $\mathbf{r}$  and acquire a phase factor  $V_{\mathbf{g}} \exp(-i\alpha_{\mathbf{g}})$  with  $\alpha_{\mathbf{g}} = 2\pi\mathbf{g} \cdot \mathbf{R}$ . Strictly speaking, the series given by Eq. (57) is then no longer a Fourier series, which is in agreement with the fact that the lattice potential is no longer periodic. It nevertheless turns out to be a good approximation, especially for small gradients of  $\mathbf{R}(\mathbf{r})$ . The presence of the defect can then be taken into account by substituting in the diffraction equations  $V_{\mathbf{g}} \rightarrow V_{\mathbf{g}} \exp(-i\alpha_{\mathbf{g}})$ . In view of the direct proportionality of  $1/t_{\mathbf{g}}$  and  $V_{\mathbf{g}}$  [Eq. (41)], this implies that  $1/t_{\mathbf{g}}$  has to be replaced by  $(1/t_{\mathbf{g}}) \exp(-i\alpha_{\mathbf{g}})$ . Equations (109a) and (109b) then take the form

$$\frac{dT}{dz} = \left(\frac{\pi i}{t_{-\mathbf{g}}}\right) \exp(i\alpha_{\mathbf{g}}) S \quad (112a)$$

$$\frac{dS}{dz} = 2\pi i s_{\mathbf{g}} S + \left(\frac{\pi i}{t_{\mathbf{g}}}\right) \exp(-i\alpha_{\mathbf{g}}) T \quad (112b)$$

and Eqs. (110a) and (110b) become

$$\frac{dT'}{dz} + \pi i s_{\mathbf{g}} T' = \left(\frac{\pi i}{t_{-\mathbf{g}}}\right) \exp(i\alpha_{\mathbf{g}}) S' \quad (113a)$$

$$\frac{dS'}{dz} - \pi i s_{\mathbf{g}} S' = \left(\frac{\pi i}{t_{\mathbf{g}}}\right) \exp(-i\alpha_{\mathbf{g}}) T' \quad (113b)$$

Both systems of equations can be transformed so as to involve only the gradient of the displacement field.

Putting  $T = T''$  and  $S = S'' \exp(-i\alpha_{\mathbf{g}})$  the first system of equations [Eqs. (112a)

and (112b)] is transformed into the set

$$\frac{dT''}{dz} = \left(\frac{\pi i}{t_{-\mathbf{g}}}\right) S'' \quad (114a)$$

$$\begin{aligned} \frac{dS''}{dz} &= 2\pi i \left(s_{\mathbf{g}} + \frac{d\alpha'_{\mathbf{g}}}{dz}\right) S'' \\ &+ \left(\frac{\pi i}{t_{\mathbf{g}}}\right) T'' \end{aligned} \quad (114b)$$

with  $\alpha'_{\mathbf{g}} = \alpha_{\mathbf{g}}/2\pi$ . Performing the substitution

$$T' = T''' \exp(\pi i \alpha'_{\mathbf{g}}) \quad (115a)$$

$$S' = S''' \exp(-\pi i \alpha'_{\mathbf{g}}) \quad (115b)$$

leads to the system

$$\begin{aligned} \frac{dT'''}{dz} + \pi i \left(s_{\mathbf{g}} + \frac{d\alpha'_{\mathbf{g}}}{dz}\right) T''' \\ = (\pi i / t_{-\mathbf{g}}) S''' \end{aligned} \quad (116a)$$

$$\begin{aligned} \frac{dS'''}{dz} - \pi i \left(s_{\mathbf{g}} + \frac{d\alpha'_{\mathbf{g}}}{dz}\right) S''' \\ = (\pi i / t_{\mathbf{g}}) T''' \end{aligned} \quad (116b)$$

The sets of Eqs. (113) and (116) are of the same form as the corresponding sets for the perfect crystal except for the  $s$  value. They suggest a simple interpretation: the presence of the defect causes the local  $s$  value,  $s_{\text{eff}} \equiv s_{\mathbf{g}} + \mathbf{g} \cdot d\mathbf{R}/dz$  to be, in general, different from its value in the perfect part of the foil. Locally, close to the defect, the Bragg condition is thus better or less well satisfied than in the rest of the foil, and hence the locally diffracted beam has a larger or smaller amplitude than in the perfect part.

If  $\alpha = 2\pi\mathbf{g} \cdot \mathbf{R}_0 = \text{Constant}$ , which is the case for a translation interface, with displacement vector  $\mathbf{R}_0$ , we have  $d\alpha/dz = 0$  and the sets of Eqs. (114) and (116) reduce to those for a perfect crystal.

This is consistent with the fact that the two parts of the crystal on either side of the interface remain perfect. We have therefore to use the sets (113) or (112) to treat this problem, as we shall demonstrate below, rather than the sets (114) or (116).

If the interface is a domain boundary described by  $\mathbf{R} = kz\mathbf{e}_\tau$  (see Sec. 1.1.1.4 of this Chapter) we have  $d\mathbf{R}/dz = k\mathbf{e}_\tau$  and now the sets (114) and (116) can be used. They show that the crystal can be described as an assembly of two juxtaposed crystal parts separated by the interface and having different  $s$  values:  $s$  and  $s + \Delta s$ , with  $\Delta s = k\mathbf{e}_\tau$ .

For the displacement field of a dislocation, neither the factor  $\exp(i\alpha_g)$  nor the gradient  $d\mathbf{R}/dz$  disappear (except for those  $\mathbf{g}$  for which extinction occurs). The sets of Eqs. (114) and (116) as well as (112) and (113) are thus suitable descriptions in this case.

The anomalous absorption has not yet been introduced explicitly in the different sets of equations for faulted crystals. Phenomenologically this can easily be done by replacing  $1/t_g$  by  $(1/t_g + i/\tau_g)$  or  $\sigma$  by  $(\sigma_r + i\sigma_i)$  in the sets of equations. Alternatively, it can be done by making the same substitution directly in the final results. We shall follow the latter approach for a discussion of the fringe patterns at inclined planar interfaces.

### Matrix Formulation for the Amplitudes of Transmitted and Scattered Beams for Perfect Crystals

For a systematic discussion of the contrast at planar interfaces we make use of a matrix formulation for the amplitudes of

the transmitted and scattered beams by a perfect crystal slab [24], which we now derive first.

Let  $T$  and  $S$  represent the transmitted and the scattered amplitudes for an incident wave with unit amplitude. The initial values at the entrance face of the slab are represented by the column vector  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ; at level  $z$  the amplitudes of transmitted and scattered beams are represented by the column vector  $\begin{pmatrix} T \\ S \end{pmatrix}$ . From Sec. 1.1.1.5 of this Chapter we know that, ignoring a common irrelevant phase factor,

$$T \equiv T(z, s_g) = \cos(\pi\sigma_g z) - i \left( \frac{s_g}{\sigma_g} \right) \sin(\pi\sigma_g z) \quad (117)$$

$$S \equiv S(z, s_g) = \left( \frac{i}{\sigma_g t_g} \right) \sin(\pi\sigma_g z) \quad (118)$$

are the solutions of the sets of Eqs. (71) or (73) with initial values  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ . Anomalous absorption is taken into account by assuming  $\sigma_g = \sigma_r + i\sigma_i$  [Eq. (98)] with

$$\sigma_r = \frac{[1 + (st_g)^2]^{1/2}}{t_g} \quad (119a)$$

$$\sigma_i = \frac{1}{(\sigma_r \sigma_g \tau_g)} \quad (119b)$$

In view of the linear character of the system of differential equations, for an incoming wave with an arbitrary amplitude we can write

$$\begin{pmatrix} T \\ S \end{pmatrix}_{\text{out}} = \begin{pmatrix} A & C \\ B & D \end{pmatrix} \begin{pmatrix} T \\ S \end{pmatrix}_{\text{in}} \quad (120)$$

where the elements  $A, B, C$ , and  $D$  of the  $2 \times 2$  matrix remain to be determined.

From

$$\begin{aligned} \begin{pmatrix} T(z, s_g) \\ S(z, s_g) \end{pmatrix} &= \begin{pmatrix} A & C \\ B & D \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &\equiv \begin{pmatrix} A \\ B \end{pmatrix} \end{aligned} \quad (121)$$

we conclude that

$$A = T \quad \text{and} \quad B = S.$$

We now make use of the symmetry of the system of Eqs. (71). We note that this system is mapped on itself by the substitution  $T \rightarrow S$ ,  $S \rightarrow T$ ,  $s_g \rightarrow -s_g$ , since  $t_g = t_{-g}$  in a centrosymmetric crystal. This means that the solution for initial values  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$  is given by

$$\begin{pmatrix} S^- \\ T^- \end{pmatrix} = \begin{pmatrix} A & C \\ B & D \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \equiv \begin{pmatrix} C \\ D \end{pmatrix} \quad (122)$$

where the minus sign means:  $T^{(-)} = T(z, -s_g)$  and  $S^{(-)} = S(z, -s_g)$ . We conclude that  $C = S^{(-)}$  and  $D = T^{(-)}$ . The response matrix is thus completely defined for arbitrary initial values

$$\begin{pmatrix} T \\ S \end{pmatrix}_{\text{out}} = \begin{pmatrix} T & S^{(-)} \\ S & T^{(-)} \end{pmatrix} \begin{pmatrix} T \\ S \end{pmatrix}_{\text{in}} \quad (123)$$

We shall represent the response matrix  $M$  of a perfect crystal slab as

$$M(z, s) \equiv \begin{pmatrix} T & S^{(-)} \\ S & T^{(-)} \end{pmatrix} \quad (124)$$

Imagining a slab of perfect crystal with total thickness  $z_0$  to be sliced in perfect slabs with thicknesses  $z_1, z_2, \dots, z_{n-1}, z_n$  such that  $z_1 + z_2 + \dots + z_n = z_0$  should clearly not influence the final result. We must therefore have

$$\begin{aligned} M(z_1 + z_2 + \dots + z_n, s_g) \\ = M(z_n, s_g) \cdot M(z_{n-1}, s_g) \dots M(z_1, s_g) \end{aligned} \quad (125)$$

This property of the response matrix can be verified by multiplying the matrices.

Equation (125) can formally be generalized to include also the subtraction of a lamella, that is

$$\begin{aligned} M(z_1 - z_2, s_g) \\ = M(-z_2, s_g) M(z_1, s_g) \end{aligned} \quad (126)$$

where

$$M(-z, s_g) = \begin{pmatrix} T(-z) & S^{(-)}(-z) \\ S(-z) & T^{(-)}(-z) \end{pmatrix} \quad (127)$$

### 1.1.1.10 Matrix Formulation for a Foil Containing a Translation Interface

A description of the diffraction effects associated with translation interfaces can be based on the set of Eqs. (113) [24]. The vector  $\mathbf{R}_0$  describes the displacement of the exit part with respect to the entrance part of the foil; it determines the sign of  $\alpha_g = 2\pi\mathbf{g} \cdot \mathbf{R}_0$ . In the front part  $\alpha_g = 0$ , whereas in the exit part  $\alpha_g \neq 0$ . The translation interface at  $z = z_1$  is assumed to be parallel to the foil surfaces. The total foil thickness is  $z_1 + z_2 = z_0$ . The front part being perfect and undisplaced; its response matrix is  $M(z_1, s_g)$ . Let the response matrix of the second part be represented by

$$\begin{pmatrix} U & X \\ V & Y \end{pmatrix} \quad (128)$$

where  $X, Y, U$ , and  $V$  must be determined from the set of Eqs. (113). We note that this set of equations reduces to that for a perfect undisplaced slab by means of the substitution  $T' = T^s$ ,  $S' = S^s \exp(-i\alpha_g)$ .

The solution of this set of equations is thus  $T^s = T(z_2, s_g)$  and  $S^s = S(z_2, s_g)$  since the front and exit parts have the same orientation. For the original set the solution is thus

$$T' = T(z_2, s_g) \quad (129a)$$

and

$$S' = S(z_2, s_g) \exp(-i\alpha_g) \quad (129b)$$

that is

$$\begin{aligned} \begin{pmatrix} T(z, s_g) \\ S(z, s_g) \exp(-i\alpha_g) \end{pmatrix} &= \begin{pmatrix} U & X \\ V & Y \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &\equiv \begin{pmatrix} U \\ V \end{pmatrix} \end{aligned} \quad (130)$$

and hence

$$U = T(z_2, s_g) \quad (131a)$$

and

$$V = S(z_2, s_g) \exp(-i\alpha_g) \quad (131b)$$

We note that the system (113) is mapped onto itself by the substitution  $T' \rightarrow S'$ ,  $S' \rightarrow T'$ ,  $s_g \rightarrow -s_g$ , and  $\alpha_g \rightarrow -\alpha_g$ . The solutions of this new set then also remain the same as those of the original set, except that the interchange of  $S'$  and  $T'$  has caused the initial values to become  $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$  and that  $s_g \rightarrow -s_g$ . We thus find

$$\begin{aligned} \begin{pmatrix} S(z_2, -s_g) \exp(i\alpha_g) \\ T(z_2, -s_g) \end{pmatrix} &= \begin{pmatrix} U & X \\ V & Y \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &\equiv \begin{pmatrix} X \\ Y \end{pmatrix} \end{aligned} \quad (132)$$

and

$$X = S(z_2, -s_g) \exp(i\alpha_g) \quad (133a)$$

and

$$Y = T(z_2, -s_g) \quad (133b)$$

Thus, in a more concise notation, the response matrix of the exit part is

$$\begin{pmatrix} T_2 & S_2^{(-)} \exp(i\alpha_g) \\ S_2 \exp(-i\alpha_g) & T_2^{(-)} \end{pmatrix} \quad (134)$$

The response matrix of the faulted slab can thus be formulated as

$$\begin{aligned} \begin{pmatrix} T \\ S \end{pmatrix} &= \begin{pmatrix} T_2 & S_2^{(-)} \exp(i\alpha_g) \\ S_2 \exp(-i\alpha_g) & T_2^{(-)} \end{pmatrix} \\ &\times \begin{pmatrix} T_1 & S_1^{(-)} \\ S_1 & T_1^{(-)} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{aligned} \quad (135)$$

Matrix (134) can conveniently be written as the product of three matrices:

$$\begin{aligned} &\begin{pmatrix} T_2 & S_2^{(-)} \exp(i\alpha_g) \\ S_2 \exp(-i\alpha_g) & T_2^{(-)} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & \exp(-i\alpha_g) \end{pmatrix} \begin{pmatrix} T_2 & S_2^{(-)} \\ S_2 & T_2^{(-)} \end{pmatrix} \\ &\times \begin{pmatrix} 1 & 0 \\ 0 & \exp(i\alpha_g) \end{pmatrix} \end{aligned} \quad (136)$$

This suggests to introduce as a shorthand the *shift matrix*

$$\mathfrak{S}(\alpha_g) \equiv \begin{pmatrix} 1 & 0 \\ 0 & \exp(i\alpha_g) \end{pmatrix} \quad (137)$$

The final result for the response of the faulted slab can then be written as

$$\begin{pmatrix} T \\ S \end{pmatrix} = \mathfrak{S}(-\alpha_g) M_2 \mathfrak{S}(\alpha_g) M_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (138)$$

with

$$M_j \equiv M_j(z_j, s_j) = \begin{pmatrix} T_j & S_j^{(-)} \\ S_j & T_j^{(-)} \end{pmatrix} \quad (139)$$

The shift matrices have the property

$$\begin{pmatrix} 1 & 0 \\ 0 & \exp(i\alpha_1) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \exp(i\alpha_2) \end{pmatrix} \\ = \begin{pmatrix} 1 & 0 \\ 0 & \exp[i(\alpha_1 + \alpha_2)] \end{pmatrix} \quad (140)$$

that is  $\mathfrak{S}(\alpha_1)\mathfrak{S}(\alpha_2) = \mathfrak{S}(\alpha_1 + \alpha_2)$ ; they commute.

The result can be generalized directly to a succession of overlapping translation interfaces characterized by phase angles  $\alpha_j$ , all referred to the front slab:

$$\begin{pmatrix} T \\ S \end{pmatrix} = \dots \mathfrak{S}(-\alpha'_2)M_3\mathfrak{S}(\alpha'_2)\mathfrak{S}(-\alpha'_1)M_2 \\ \times \mathfrak{S}(\alpha'_1)M_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (141)$$

Introducing the phase angles  $\alpha_j = \alpha'_j - \alpha'_{j-1}$ , which now describe the *relative* displacements of successive lamellae, the rear part being displaced with respect to the front one, we obtain

$$\begin{pmatrix} T \\ S \end{pmatrix} = \dots M_3\mathfrak{S}(\alpha_2)M_2\mathfrak{S}(\alpha_1)M_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (142)$$

#### 1.1.1.11 Matrix Formulation for a Foil Containing a Domain Boundary [20]

It is possible to generalize further by assuming that in successive lamellae the  $s$  values may be different, as in the case of domain boundaries.

For a pure domain boundary the transmitted and scattered amplitudes are

given by

$$\begin{pmatrix} T \\ S \end{pmatrix} = M(z_2, s_{g,2})M(z_1, s_{g,1}) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (143)$$

where now  $s_{g,1} \neq s_{g,2}$ .

Equation (142) also describes the most general case of overlapping mixed boundaries, that is boundaries containing a translation component as well as exhibiting a difference in deviation parameters, provided the  $s$  values in the successive matrices,  $M_j$ , are assumed to be different.

#### 1.1.1.12 Matrix Formulation for a Crystal Containing a Non-reflecting Part: the Vacuum Matrix

A foil may contain lamellae which are very far from any reflecting orientation under the diffraction conditions prevailing in the rest of the foil; except for absorption they behave as if no material was present in that lamellae. This is, for example, the case for a microtwin lamella in a face-centered structure if an uncommon reflection is excited in the matrix. It also applies to a cavity and to a precipitate lamella with a lattice different from that of the reflecting matrix. Even though such parts of the foil do not contribute to the diffraction phenomena, their presence influences the relative phases of the waves diffracted by the foil parts in front and behind these inactive lamellae. This can be accounted for by including the appropriate matrix in the matrix product, describing such a non-diffracting part [24].

In a nonreflecting part the extinction distance,  $t_g$ , is infinite and the system of

Eqs. (113) thus reduces to

$$\frac{dT}{dz} + \pi i s_g T = 0 \quad (144a)$$

and

$$\frac{dS}{dz} - \pi i s_g S = 0 \quad (144b)$$

which integrate to

$$T = T_0 \exp(-\pi i s_g z) \quad (145a)$$

and

$$S = S_0 \exp(\pi i s_g z) \quad (145b)$$

where  $T_0$  and  $S_0$  are the amplitudes at the entrance face of the non-reflecting 'vacuum' lamella. One can thus write

$$\begin{pmatrix} T \\ S \end{pmatrix}_{\text{out}} = \begin{pmatrix} \exp(-\pi i s_g z) & 0 \\ 0 & \exp(\pi i s_g z) \end{pmatrix} \times \begin{pmatrix} T \\ S \end{pmatrix}_{\text{in}} \quad (146)$$

The 'vacuum' matrix is thus

$$V(z, s_g) = \begin{pmatrix} \exp(-\pi i s_g z) & 0 \\ 0 & \exp(\pi i s_g z) \end{pmatrix} \quad (147)$$

where  $z$  is the thickness of the nonreflecting part as measured along the beam path and  $s_g$  is the deviation parameter of the crystal part preceding the 'vacuum' lamella.

### 1.1.1.13 Fringe Profiles at Planar Interfaces [25]

#### General Formulae

Multiplication of the matrices in Eq. (138) results in the following expressions for  $T$

and  $S$ :

$$T = T_1 T_2 + S_1 S_2^{(-)} \exp(i\alpha) \quad (148a)$$

and

$$S = T_1 S_2 \exp(-i\alpha) + S_1 T_2^{(-)} \quad (148b)$$

The minus signs in  $S_2^{(-)}$  and  $T_2^{(-)}$  indicate that the expressions  $S_2$  and  $T_2$  have to be modified by changing  $s$  to  $-s$ . Equations (148) express the fact that the transmitted amplitude results from the interference between the doubly transmitted beam  $T_1 T_2$  and the doubly scattered beam  $S_1 S_2^{(-)} \exp(i\alpha)$ . The minus sign in  $S_2^{(-)}$  indicates that the scattering, by the second part of the foil of the beam which has already been scattered by the first part, takes place from the  $-g$  side of the lattice planes. This implies that the deviation parameter has to be changed from  $s$  to  $-s$ , since for this second scattering event the node  $G$  acts as the origin of reciprocal space (Fig. 29). This second scattering

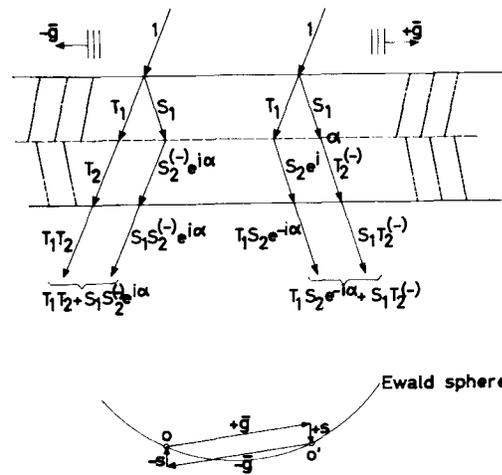


Figure 29. Transmitted and scattered amplitudes for a crystal containing a planar interface.

process is accompanied by a phase shift  $\alpha = 2\pi\mathbf{g} \cdot \mathbf{R}$  due to the translation over  $\mathbf{R}$  of the lattice planes in part II with respect to those in part I. This is taken into account by the phase factor  $\exp(i\alpha)$ . A similar interpretation can be given to the expression for  $S$ . The phase factor is now  $\exp(-i\alpha)$  rather than  $\exp(i\alpha)$  because the phase shifts due to the displacement of part II have opposite signs for  $S_2^{(-)}$  and for  $S_2$ , since  $\mathbf{g}$  has opposite sign for the two processes.

Introducing the explicit expressions for  $T_j$  and  $S_j$  ( $j = 1$  or  $2$ ) given by Eqs. (117) and (118), in Eqs. (148a) and (148b), and noting that  $\sigma = \sigma_r + i\sigma_i$  since anomalous absorption is taken into account, leads to explicit expressions for  $I_T = TT^*$  and  $I_S = SS^*$ . It turns out that it is possible to cast these explicit expressions for  $I_T$  and  $I_S$  in a form which allows a detailed analytical discussion of the fringe profiles. Detailed calculations are given in [25] also for the general case, where the deviation parameters  $s_1$  and  $s_2$  as well as the extinction distances  $t_{g_1}$  and  $t_{g_2}$  in parts I and II, respectively, are assumed to be different. We do not give the details here of these straightforward but tedious calculations, but we will summarize the significant results.

The expressions for  $I_T$  and  $I_S$  can be written as sums of three terms

$$I_{T,S} = I_{T,S}^{(1)} + I_{T,S}^{(2)} + I_{T,S}^{(3)} \quad (149)$$

### Translation Interfaces

We consider first pure translation interfaces [2]. Thus  $s_1 = s_2 \equiv s$ , and limiting ourselves to the case when  $s = 0$ , the

expressions become

$$I_{T,S}^{(1)} = \frac{1}{2} \cos^2\left(\frac{\alpha}{2}\right) \times [\cosh(2\pi\sigma_i z_0) \pm \cos(2\pi\sigma_r z_0)] \quad (150)$$

$$I_{T,S}^{(2)} = \frac{1}{2} \sin^2\left(\frac{\alpha}{2}\right) \times [\cosh(4\pi\sigma_i u) \pm \cos(4\pi\sigma_r u)] \quad (151)$$

$$I_{T,S}^{(3)} = \frac{1}{2} \sin \alpha [\sin(2\pi\sigma_r z_1) \sinh(2\pi\sigma_i z_2) \pm \sin(2\pi\sigma_r z_2) \sinh(2\pi\sigma_i z_1)] \quad (152)$$

where the upper sign corresponds to  $I_T$  and the lower sign to  $I_S$ . The total thickness is  $z_0 = z_1 + z_2$ , where  $z_1$  is the thickness of the front part and  $z_2$  that of the rear part. Furthermore, we have  $u = \frac{1}{2}(z_1 - z_2)$ , that is  $u$  is the distance of the interface from the midplane of the foil. Along a planar interface intersecting the foil surfaces, as is often the case,  $z_1$  and  $z_2$  vary along the foil in such a way that  $z_1 + z_2$  remains constant and equal to  $z_0$ . In the projected area of the interface fringes will be formed which, according to the column approximation, can be considered as being due to the intersection of the depth variation of  $I_T$  (or  $I_S$ ) with the inclined interface.

If  $\alpha = n \cdot 2\pi$ , where  $n$  is an integer, which is the case if there is *no* stacking fault,  $\sin \alpha = 0$  and  $\sin(\alpha/2) = 0$ . The only remaining term is then  $I_{T,S}^{(1)}$ , which as a result must represent the contribution due to the perfect crystal. This term only depends on the total thickness  $z_0$ ; it describes a background onto which the fringes represented by the other terms are superimposed.

The second term,  $I_{T,S}^{(2)}$ , depends on  $u$  and not on  $z_1$  and  $z_2$  separately. It represents a function which is periodic in  $u$ , with a depth period  $1/2\sigma_r$ . The center of the

pattern, at  $u = 0$ , exhibits an extremum; it is a minimum for  $I_S^{(2)}$  and a maximum for  $I_T^{(2)}$ . This pattern consists of fringes which are parallel to the central line of the pattern. We shall see that the amplitude of the fringes is only large enough to be visible in the central part of the pattern.

The dominant features of the pattern, in sufficiently thick foils, are described by  $I_{T,S}^{(3)}$ . Where the interface is close to the entrance face  $z_1$  is small and  $z_2 \simeq z_0$ , and the factor  $\sinh(2\pi\sigma_1 z_2)$  is then large. The term  $\frac{1}{2} \sin \alpha \sinh(2\pi\sigma_1 z_2) \sin(2\pi\sigma_r z_1)$  represents a damped sinusoid with a depth period  $1/\sigma_r$ . This term disappears at the rear surface where  $z_2 = 0$ . If  $\sin \alpha > 0$ , the first extremum is a maximum; as a result the first fringe will be bright at the entrance face. For  $\sin \alpha < 0$ , the first fringe will be dark.

Where the interface is close to the exit face  $z_1 \simeq z_0$  and  $z_2 \simeq 0$ ; the term  $\pm \frac{1}{2} \sin \alpha \sinh(2\pi\sigma_1 z_1) \sin(2\pi\sigma_r z_2)$  is now dominant. It again represents a damped sinusoid. The first extremum, which now refers to the last fringe, is either a maximum or a minimum, depending on the sign of  $\sin \alpha$ . Note that the extremum is different for  $I_T$  and  $I_S$  since in the first case the plus sign applies, whereas for  $I_S$  the minus sign applies.

Figure 30 and Table 1 summarize this discussion. Note that these results, in particular those pertaining to the nature of the edge fringes, imply that anomalous absorption must be sufficiently large to make sure that the dominant behavior is described by the term  $I_{T,S}^{(3)}$ .

For a discussion of the behavior in the central part of the pattern the term  $I_{T,S}^{(2)}$  may become important, since the two terms of  $I_{T,S}^{(3)}$  compensate for certain

thicknesses and  $I_{T,S}^{(2)}$  may then become the dominant term.

For the important case  $\alpha = \pm 2\pi/3$ , which occurs in cubic-close-packed structures, the stacking fault fringes have the following properties, provided the foil is sufficiently thick so as to make anomalous absorption a dominant feature.

The bright field fringe pattern is symmetrical with respect to the line  $u = 0$ . This can be deduced quite generally from the implicit expressions [Eqs. (148)] for  $s_1 = s_2 \equiv s$ , which shows that  $I_S = TT^*$  has the symmetry property

$$I_T(z_1, z_2, s, \alpha) = I_T(z_2, z_1, s, \alpha) \quad (153)$$

On the other hand,  $I_S = SS^*$  has the property

$$I_S(z_1, z_2, s, \alpha) = I_S(z_2, z_1, -s, -\alpha) \quad (154)$$

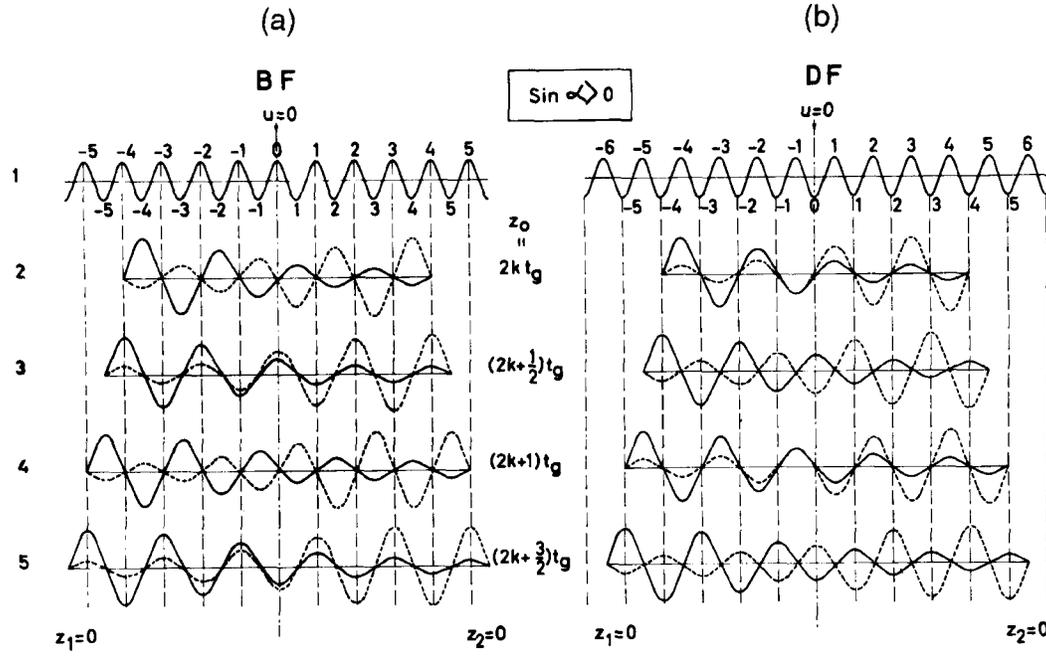
that is the dark field fringe pattern is antisymmetrical with respect to the foil center since interchanging  $z_1$  and  $z_2$  also requires changing the signs of  $s$  and  $\alpha$ , which changes the nature of the edge fringes.

The fringes are parallel to the closest surface; as a result, new fringes caused by an increase in foil thickness are generated by fringe splitting close to the center of the pattern. This result can be understood by noting the relative shift, with increasing thickness, of the curves representing the two terms in  $I_{T,S}^{(3)}$  in the central part of the foil where they overlap.

Close to the entrance face of the foil the fringe patterns are similar, but close to the exit part they are complementary. This property is generally true for diffraction contrast images when anomalous absorption is important; it is also true for dislocation images. The computed profiles shown in Fig. 31, which can be compared

**Table 1.** The properties of fringe profiles due to planar interfaces, showing the nature of the first and last edge fringes for different signs of  $\sin \alpha$  and  $\delta$ . A schematic profile and an observed pattern are given for the two types of fringe.

	$\alpha = 2\pi(\mathbf{g} \cdot \mathbf{R})$				$\delta = s_1 t_{g_1} - s_2 t_{g_2}$				
	Bright field		Dark field		Bright field		Dark field		
	First	Last	First	Last	First	Last	First	Last	
$\sin \alpha > 0$	B	B	B	D	$\delta > 0$	B	D	B	B
$\alpha \neq \pi$	-	-	-	-		-	-	-	-
$\sin \alpha < 0$	D	D	D	B	$\delta < 0$	D	B	D	D



with the observed fringes in a silicon wedge, allow verification of most of these properties.

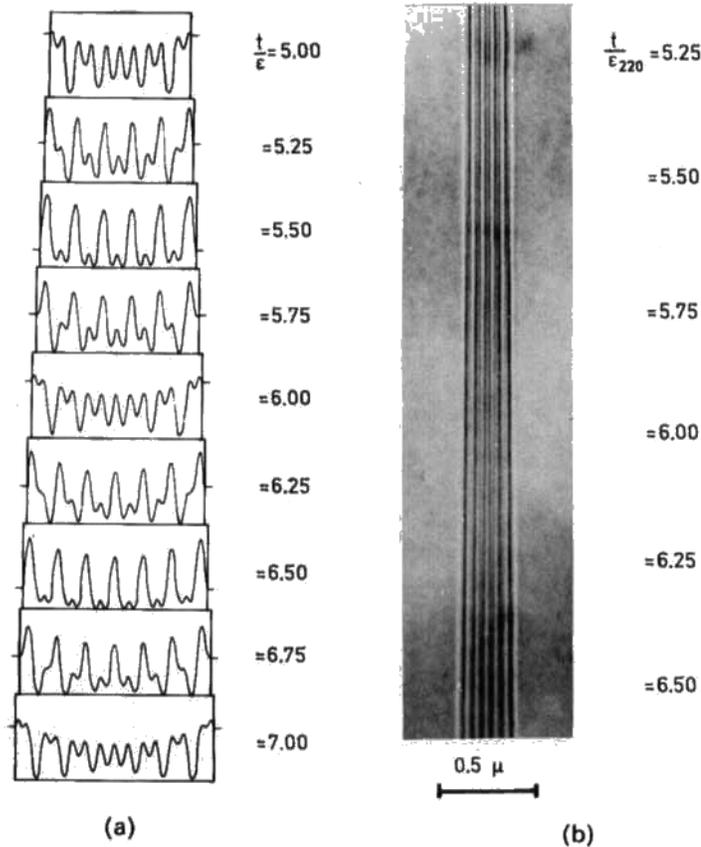
It is clear from the foregoing discussion that the nature of the edge fringes depends on the sign of  $\sin \alpha$ . The case when  $\alpha = \pi$  is singular, because now  $\sin \alpha = 0$  and the term  $I_{T,S}^{(3)}$  (which is mostly dominant) is now absent, as is the background term  $I_{T,S}^{(1)}$  since  $\cos(\alpha/2) = 0$ . The complete

fringe pattern for  $s = 0$  is then represented by

$$I_{T,S}^{(2)} = \frac{1}{2} [\cosh(4\pi\sigma_i u) \pm \cos(4\pi\sigma_r u)] \quad (155)$$

since  $\sin(\alpha/2) = 1$ .

This expression shows that the bright field and dark field images are complementary with respect to the nonperiodic background which is described by  $\cosh(4\pi\sigma_i u)$ . This background exhibits a



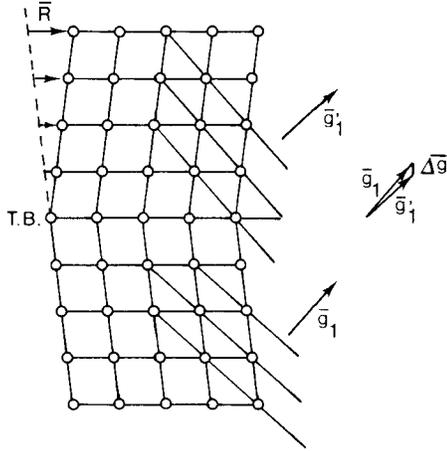
**Figure 31.** (a) Computed profiles for a stacking fault with increasing thickness of the specimen. (b) Fringe pattern due to a stacking fault in a wedge-shaped foil of silicon ( $s = 0$ ;  $g = 220$ ;  $\alpha = 2\pi/3$ ; thickness =  $5t_g - 7t_g$ ). (Courtesy of G. R. Booker.)

minimum for  $u = 0$ , that is it shows a minimum in the central part of the pattern. The fringes with a depth period  $1/(2\sigma_r)$  are parallel to the central lines  $u = 0$ ; they are superimposed on this background. As the thickness increases, new fringes are added at the surfaces.

### Domain Boundary Fringes

We now consider interfaces separating two juxtaposed crystal parts in which the deviation parameters for homologous, simultaneously excited reflections are

slightly different [25]. Such boundaries occur, for instance, in microstructures that result from phase transitions in which rotation symmetry elements are lost. Often the interface is a coherent twin with a small twinning vector. Under these conditions the simultaneously excited diffraction vectors  $g_1$  and  $g_2$  differ by  $\Delta g = g_2 - g_1$ , where  $\Delta g$  is perpendicular to the coherent twin interface (Fig. 32). The difference in the deviation parameters,  $\Delta s = s_2 - s_1$ , is the projection of  $\Delta g$  along the normal to the foil plane. In the 'symmetrical' case, one has  $s_2 = \frac{1}{2}\Delta s$  and  $s_1 = -\frac{1}{2}\Delta s$ , that is  $s_1 = -s_2$ .



**Figure 32.** Schematic representation of a coherent domain boundary. TB, (Twin boundary).

The expressions for the transmitted and scattered amplitudes for a pure domain boundary (i.e., without a translation component) are given by

$$T = T_1 T_2 + S_1 S_2^- \quad (156a)$$

and

$$S = T S_2 + T_2^{(-)} S_1 \quad (156b)$$

of which the interpretation has been discussed in detail for the case of a stacking fault.

The general expressions for the intensities  $I_T$  and  $I_S$ , assuming the extinction distances in the two parts to be different, can again be written as the sum of three terms, and a discussion similar to the one given for translation interfaces is possible [25]. For sufficiently thick foils the behavior is again dominated by the terms  $I_{T,S}^{(3)}$ , which we will now discuss.

The general features of the fringe pattern are adequately exhibited by the 'symmetrical' case defined above. The terms

$I_{T,S}^{(3)}$  are now given by [25]

$$w^4 I_{T,S}^{(3)} = -\frac{1}{2} \delta \{ \cos(2\pi\sigma_r z_1) \times \sinh[2(\pi\sigma_{i,2} z_2 \pm \varphi_2)] \mp \cos(2\pi\sigma_r z_2) \times \sinh[2(\pi\sigma_{i,1} z_1 \pm \varphi_1)] \} \quad (157)$$

with

$$w^2 = 1 + (st_g)^2$$

$$t_g = \frac{1}{2}(t_{g_1} + t_{g_2})$$

$$\delta = s_1 t_{g,1} - s_2 t_{g,2}$$

$$2\varphi_j = \operatorname{argsinh}(st_{g,j})$$

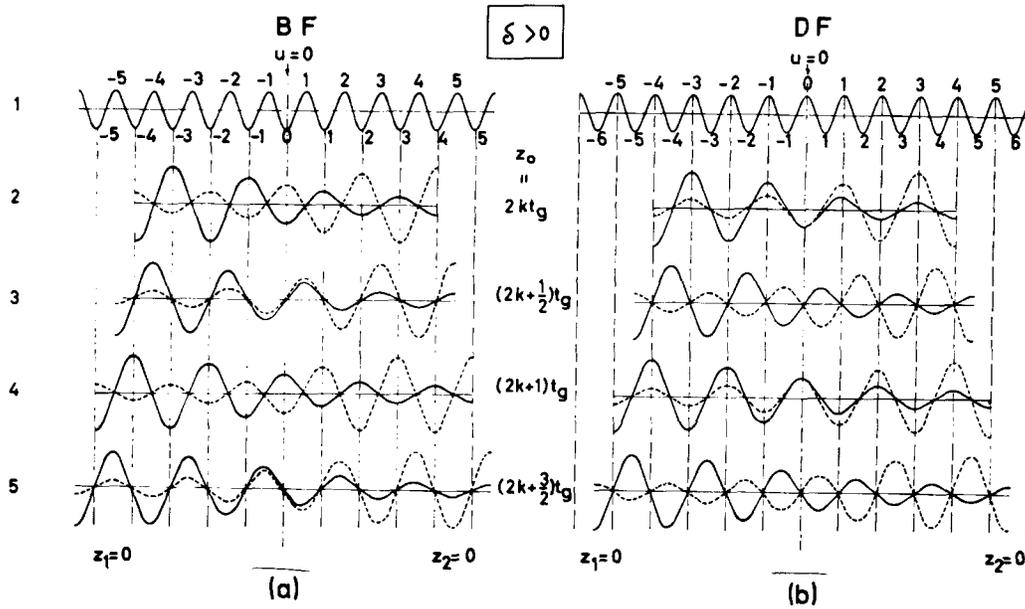
and

$$st_g = \frac{1}{2}(s_1 t_{g_1} + s_2 t_{g_2})$$

The upper sign applies to  $I_T$  and the lower to  $I_S$ . The nature of the fringes is visibly determined by the parameter  $\delta$ , which is assumed to be sufficiently small so that the same  $g$  vector is excited in both crystal parts.

Close to the front surface the first term of  $I_{T,S}^{(3)}$  determines the behavior since  $\sinh(2\pi\sigma_{i,2} z_2 \pm \varphi_2)$  is large for  $z_2 \simeq z_0$  and  $z_1 \simeq 0$ . Close to the exit face the second term is dominant. The dependence of the nature of the edge fringes on the sign of  $\delta$  is summarized in Table 1 and in Fig. 33.

The most striking and useful feature is the difference in the symmetry of the fringe patterns, due to translation interfaces. Whereas the bright field pattern for a translation interface is symmetrical with respect to the central line, it is roughly anti-symmetrical for a domain boundary, the edge fringes being opposite in nature. On the other hand, for the special case  $s_1 = -s_2$  and  $t_{g_1} = t_{g_2}$  the dark field image



**Figure 33.** Schematic representation of the fringe profile due to a coherent domain boundary [25]: (a)  $I_T$ ; (b)  $I_S$ . BF, bright field; DF, dark field.  $I_T$  (Transmitted Intensity),  $I_S$  (Scattered Intensity).

is symmetrical for domain boundary fringe patterns, but anti-symmetrical for translation interfaces. If  $t_{g_1}$  and  $t_{g_2}$  are significantly different, the depth periods close to the front and rear surfaces may be different. As for translation interfaces, the fringes are parallel to the closest surface; this is a consequence of anomalous absorption and it is therefore only true in sufficiently thick foils.

A characteristic feature of domain boundary images is that the domain contrast on either side of the interface (i.e., on either side of the fringe pattern) may be different, which is never the case for translation interfaces. However, for  $s_1 = -s_2$  (i.e., in the symmetrical situation) the domain contrast is the same in both domains in the dark field image, but not in the bright field image. This is a consequence of the symmetry of the rocking

curve for the scattered beam, on the one hand, and its asymmetry for the transmitted beam (the Borrmann effect), on the other.

Along certain interfaces there may be a phase shift as well as a difference in the orientation or length of the excited diffraction vector in the two crystal parts. The fringes produced along such interfaces have properties which are intermediate between those of pure  $\alpha$  fringes and pure  $\delta$  fringes [25].

### Extinction Criteria

It is clear that no  $\alpha$  fringes are produced if  $\mathbf{g} \cdot \mathbf{R} = \text{Integer}$ . In Eqs. (150)–(152) for  $I_{T,S}$ , the terms  $I_{T,S}^{(2)}$  and  $I_{T,S}^{(3)}$  become zero. Only  $I_{T,S}^{(1)}$  is different from zero; however, this term represents thickness fringes,

as it only depends on  $z_0$ . In fact it is easy to verify that for a perfect crystal Eqs. (151) and (107) are identical (for  $s = 0$ ).

If an image is made using a diffraction vector which is common to the two crystal parts, that is if a diffraction spot belonging to the unsplit row or the unsplit plane is selected, the  $\delta$  component of a mixed interface becomes inoperative and only a possible translation component may produce  $\alpha$  fringes. It is also possible to eliminate selectively the translation component from the images of mixed boundaries. In this way it is, for example, possible to image the lattice relaxation along antiphase boundaries or stacking faults with a displacement vector  $\mathbf{R}_0 + \boldsymbol{\varepsilon}$  by exciting only a systematic row of reflections  $(\dots - 2g, -g, 0, +g, +2g \dots)$  for which  $\mathbf{g} \cdot \mathbf{R} = \text{Integer}$ . The presence of relaxation is then revealed by the occurrence of weak residual fringes, due to the additional displacement  $\boldsymbol{\varepsilon}$ , for which  $\mathbf{g} \cdot \boldsymbol{\varepsilon} \neq \text{Integer}$  [26]. Using a number of different reflections for which  $\mathbf{g} \cdot \mathbf{R}_0 = \text{Integer}$  but for which  $\mathbf{g} \cdot \boldsymbol{\varepsilon} \neq \text{Integer}$ , one can obtain a fair idea of the direction, sense, and magnitude of  $\boldsymbol{\varepsilon}$  from observations of the nature and contrast of the edge fringes in the residual fringe patterns.

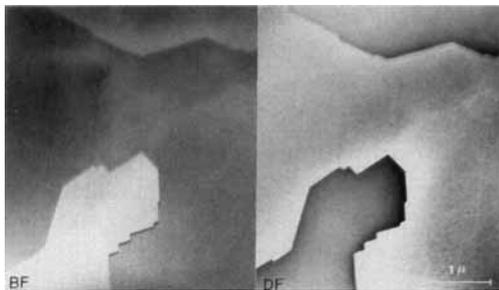
#### 1.1.1.14 Domain Fragmented Crystals: Microtextures

Many phase transformations lead to a decrease in space-group symmetry, the space group of the low-temperature phase being a subgroup of that of the high-temperature phase. As a result, a single crystal of the high-temperature phase usually becomes fragmented into

domains after transformation into the low-temperature, low-symmetry phase. The structures within these domains are then related by symmetry operations lost during the transformation. The lost symmetry elements can be either *translations* or *rotations* [27]. The interfaces resulting from lost translation symmetry are translation interfaces (e.g., antiphase boundaries, stacking faults, discommensurations, and crystallographic shear planes). Lost rotation symmetry elements give rise to twins or domain boundaries. The use of the term *domain boundary* will be reserved for those cases where the lattices of the two domains are only slightly different. The reciprocal lattice nodes belonging to the two domains are then sufficiently close to each other to be excited simultaneously, albeit with different deviation parameters, and produce  $\delta$  fringes.

At the other extreme, if the diffraction spots in a diffraction pattern made across the interface are sufficiently split so as to be able to make a dark field image in one of the components separately, we call the interface a *twin*. The image so obtained then exhibits wedge fringes in the selected domain. It is clear that the distinction between twins and domain boundaries is not very strictly defined in this way and intermediate situations are possible.

In some cases the lattices of the two domains separated by the interface are the same but the structures may be different. This is, for example, the case in noncentrosymmetrical crystals where the structures in the domains may be related by an inversion operation, or by a two-fold axis, the lattice being unperturbed by the interface. The domain structure in  $\alpha$ -quartz provides an example of the latter type. The high-temperature  $\beta$  form of



**Figure 34.** Dauphiné twin domains in  $\alpha$ -quartz revealed by domain contrast [28–30]. BF, bright field; DF, dark field.

quartz has point-group symmetry 622, whereas the low-temperature  $\alpha$  form belongs to the point group 32, that is the six-fold axis of the  $\beta$  phase becomes a three-fold axis in the  $\alpha$  phase. On cooling to below the  $\beta \rightarrow \alpha$  transition temperature (about 573°C), the  $\beta$  phase breaks up into Dauphiné twins,  $\alpha_1$  and  $\alpha_2$ , of the  $\alpha$  phase. The structures of  $\alpha_1$  and  $\alpha_2$  are related by the lost 180° rotation about the three-fold axis, whereas the lattice of  $\alpha_1$  and  $\alpha_2$  remains common (Fig. 34) [28–30].

Inversion boundaries occur in many non-centrosymmetric crystals and, for

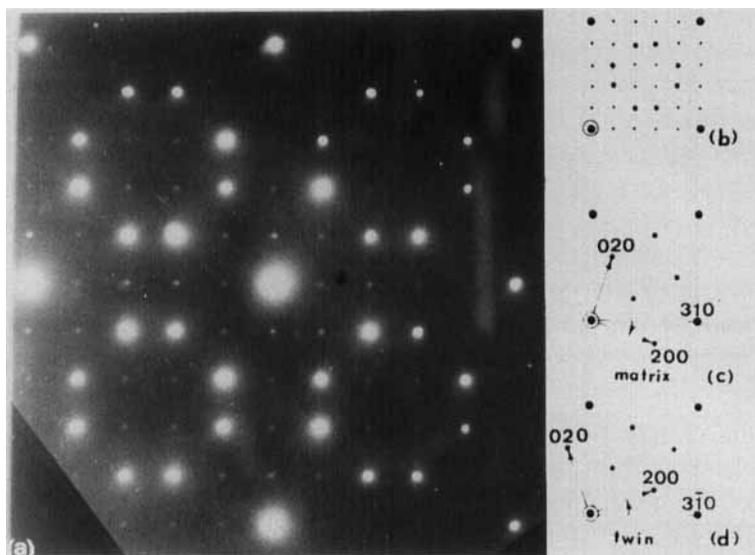
example, in the cubic  $\chi$ -phase alloy Fe-Cu-Mo-Ti (Fig. 35) [31].

Domain textures can conveniently be studied by a combination of diffraction and imaging techniques exploiting different diffraction contrast phenomena.

### 1.1.1.15 Diffraction Patterns of Domain Textures

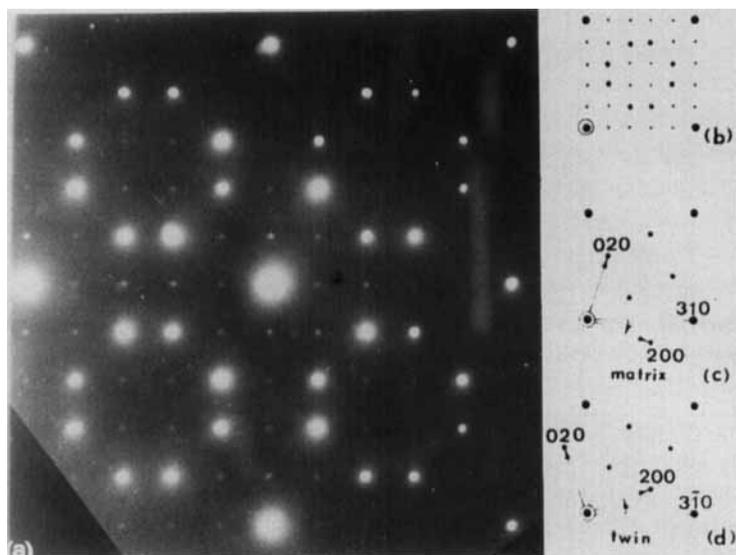
Domain textures produce a composite diffraction pattern which is the superposition of the diffraction patterns of the separate domains. This usually affects the geometry of the diffraction pattern by the occurrence of split spots; in some cases only the intensities are changed compared to those of a monodomain pattern.

Where differently oriented domains overlap when viewed along the zone axis, double diffraction may occur. This may sometimes complicate the interpretation by simulating a diffraction pattern with lower translation symmetry than that of



**Figure 35.** Inversion domains in the  $\chi$  phase of Fe-Cr-Mo-Ti revealed by domain contrast [31].

**Figure 36.** Diffraction pattern of  $\text{Ni}_4\text{Mo}$  exhibiting weak double-diffraction spots. (a) observed pattern; (b) Schematic: small dots represent double diffraction spots; (c) one orientation variant; (d) second orientation variant.



the separate domains, as a result of the double-diffraction spots (Fig. 36).

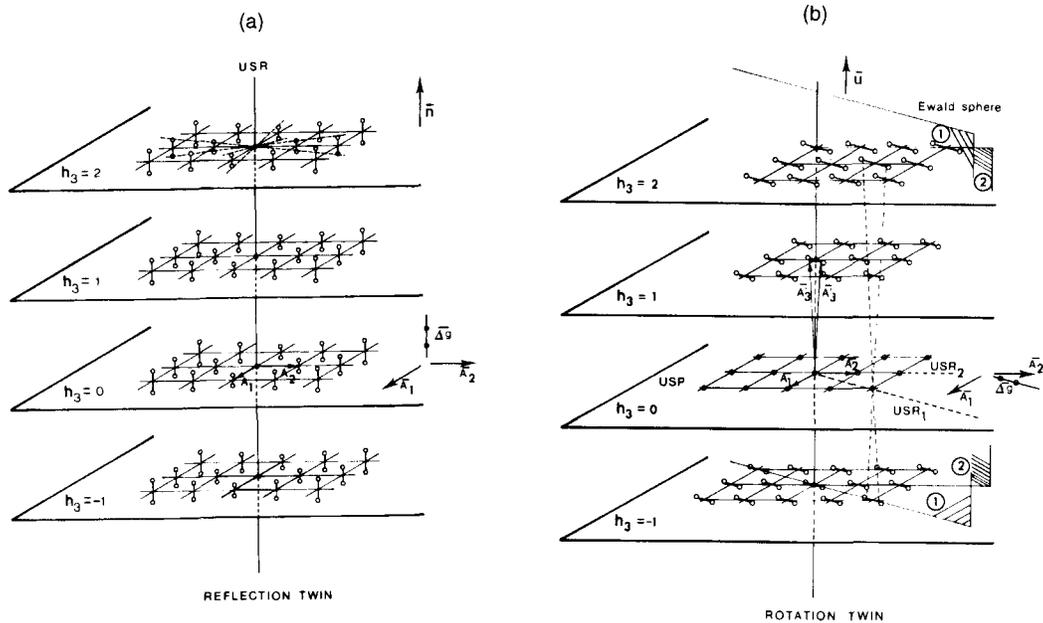
If several orientation variants are present in the selected area the diffraction pattern may become quite complicated and difficult to 'unscramble'. The interpretation can be simplified by first making monodomain diffraction patterns of the domains on both sides of the interface and subsequently from an area across the interface separating the two domains. However, this is only possible if the domains are sufficiently large.

The diffraction patterns across twins have characteristic features which allow one to determine the twinning elements. The reciprocal space of a reflection twin is represented in Fig. 37a; it exhibits a central row of unsplit nodes, perpendicular to the mirror plane in real space. This is a general feature of the relationship between direct and reciprocal space. A common lattice plane in real space (the coherent mirror plane) is represented in reciprocal space as a common lattice row perpendicular to the

mirror plane. A common lattice row in direct space, as is the case for the lattice row along a  $180^\circ$  rotation twin axis, is represented in reciprocal space as a common reciprocal lattice plane perpendicular to the twinning axis. All other spots are split (Fig. 37b) [32].

In the case of a reflection twin the spot splitting is parallel to the unsplit row and its magnitude is proportional to the distance from the unsplit row. The magnitude of the spot splitting is a direct measure for the twinning vector. For a  $180^\circ$  rotation twin, the spots are all split along a direction parallel to the unsplit plane; the magnitude of splitting is proportional to the distance from the unsplit plane.

It is not always obvious how to distinguish between the two cases, since many sections of reciprocal space will look very similar. Tilting experiments exploring the relevant parts of reciprocal space are required in order to differentiate between the two cases. An example of the type of



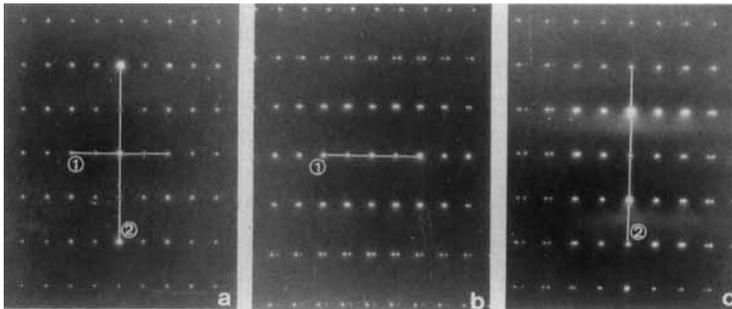
**Figure 37.** Reciprocal lattice of (a) a reflection twin and (b) a  $180^\circ$  rotation twin [32]. USR, (unsplit row); USP (unsplit plane).

experiment to be performed is shown in Fig. 38 [33].

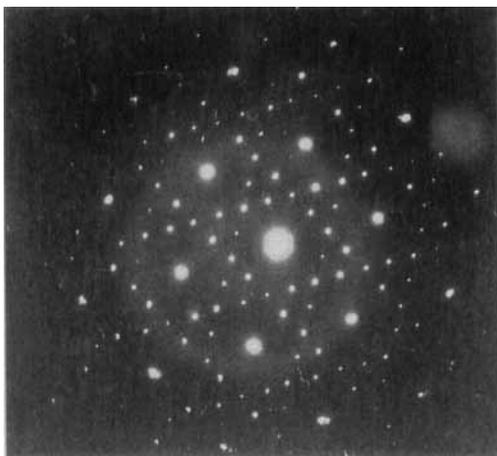
The presence of higher order symmetry elements relating the structures in the different domains is reflected in the symmetry of the diffraction pattern. Figure 39 shows, for instance, the presence of three orthorhombic orientation variants related by  $120^\circ$  rotations along the zone axis. Such microstructures can usually be analyzed in

terms of reflection or  $180^\circ$  rotation twins, by considering pairs of domains.

The distinction between a diffraction pattern produced by a quasicrystal along a noncrystallographic zone (e.g., a five-fold or ten-fold symmetry axis) and a diffraction pattern due to multiply-twinned 'classical' crystals is not always obvious, and has given rise to much debate.



**Figure 38.** Tilting experiment in  $\text{MoTe}_2$  allowing the presence of  $180^\circ$  rotation twins to be shown [33]. 1 and 2 are the tilt axis.



**Figure 39.** Composite diffraction pattern of a foil containing three orthorhombic orientation variants of  $\text{Ni}_3\text{Mo}$  related by  $120^\circ$  rotations. (Courtesy of G. Van Tendeloo.)

The presence of domains which are built on a common lattice is not reflected in the geometry of the diffraction pattern since it causes neither spot splitting nor additional spots compared with a monodomain pattern. The relative intensities of the spots are affected, but this is not easily detected in electron diffraction patterns. Imaging techniques are of considerable help in the study of such textures.

### 1.1.1.16 Imaging of Microtextures

Microtextures can be imaged either by means of domain contrast, by interface contrast, or by both.

#### Domain Contrast of Orientation Variants

Domain contrast usually finds its origin in a small difference in the deviation parameters in adjacent domains leading to a

significant difference in brightness in either the bright field image or the dark field image, made in a split reflection. More pronounced contrast arises if the dark field image is made in one of the components of a split reflection. However, this is only possible if the spot splitting is large enough.

The difference in brightness in the bright field image can be understood with reference to the asymmetric rocking curve for  $I_T$ . In the vicinity of  $s = 0$  the  $s$  dependence of  $I_T$  is quite steep and a small difference in deviation parameter leads to a pronounced difference in transmitted intensity. Optimum domain contrast is thus obtained if the average deviation parameter is close to  $s = 0$ .

In the dark field image optimum contrast is achieved if a single spot can be isolated. If this is not the case the symmetry of the rocking curve for  $I_S$  shows that now the optimum contrast is obtained for an average  $s$  which is different from zero.

Domain contrast can also arise because the moduli of the structural factors, and hence the extinction distances, are different in adjacent domains. This is, for instance, the case for Dauphiné twins in quartz [28–30]. The lattices of  $\alpha_1$  and  $\alpha_2$  coincide, no spot splitting occurs and the above-mentioned contrast phenomena are inoperative. However, a number of coinciding reflections have structure amplitudes of different magnitudes. Dark field images made in such reflections will give rise to domain contrast, often called *structure factor contrast* (see Fig. 34).

It is clear that translation variants cannot give rise to domain contrast since the lattices, as well as the structures, are strictly parallel in the two domains.

### Interface Contrast

It is also possible to image the interfaces rather than the domains. This is the only possibility for translation interfaces. For orientation variants domain contrast and interface contrast are often produced simultaneously.

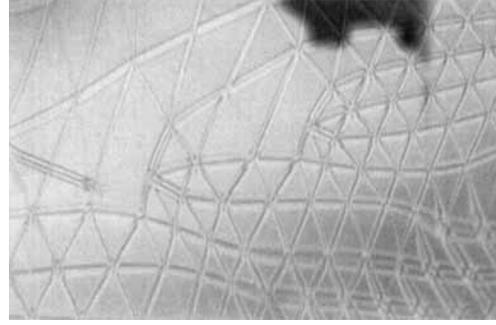
The interfaces separating translation variants such as out-of-phase boundaries, crystallographic shear planes and stacking faults are imaged as  $\alpha$ -type fringes in reflections for which  $\mathbf{g} \cdot \mathbf{R}_0 \neq \text{Integer}$ . This is also the case for interfaces separating structural variants built on a common lattice, but having different structure amplitudes. We have seen above that domain contrast arises as a result of structure factor contrast when the moduli of the structure factors are different. However, it often happens that the structure factors have the same modulus in the two domains, but have a different phase. This is the case for certain reflections in domain fragmented  $\alpha$ -quartz. A dark field image in such a reflection will not exhibit domain contrast but will reveal the interfaces as  $\alpha$ -type fringe patterns.

The phase angle,  $\alpha$ , can be deduced as follows. The structure factors for the structures in the two domains is written with respect to a common origin. The two structure factors are then related as follows:

$$F_{\mathbf{H}}^{(2)} = F_{\mathbf{H}}^{(1)} \exp(i\alpha_{\mathbf{H}}) \quad (158)$$

where  $\alpha$  is the phase angle characterizing the fringe pattern,  $F_{\mathbf{H}}^{(2)}$  and  $F_{\mathbf{H}}^{(1)}$  are the structure factors with indices  $\mathbf{H}$  in the exit and front part, respectively.

Inversion boundaries revealed by means of interface contrast in the  $\chi$  phase of the alloy Fe-Cr-Mo-Ti are visible in Fig. 40 [31, 34–36].



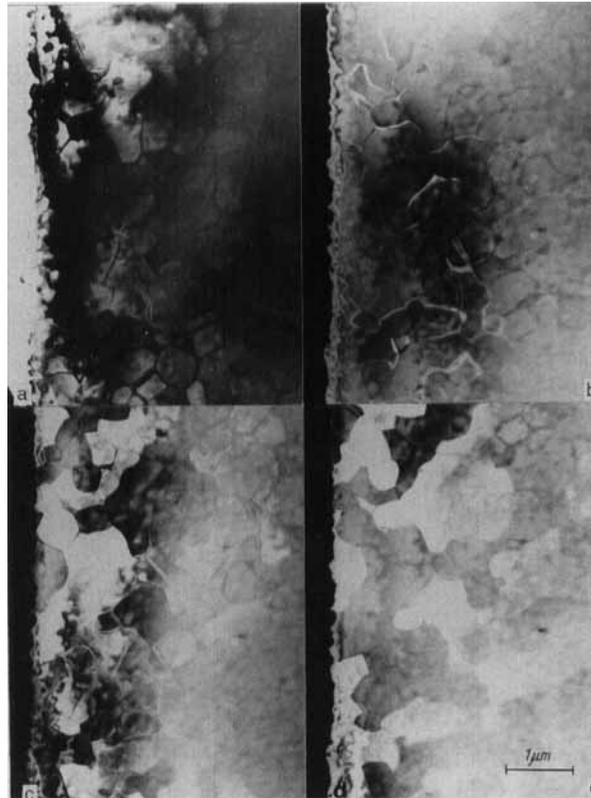
**Figure 40.** Inversion domains in the  $\chi$  phase of Fe-Cr-Mo-Ti, as revealed by interface contrast [34].

A particular type of interface contrast arises in dark field images made in double-diffraction spots caused by overlapping orientation domains. The projected interfacial region will now appear bright since double diffraction is only produced in the regions of overlap along the interfaces (Fig. 41).

The contrast at discommensuration walls is somewhat similar to that at anti-phase boundaries. In the latter case dark field imaging in superstructure spots is used, whereas discommensuration walls are best imaged in dark field images, selecting incommensurate diffraction spots due to the modulation.

### Inversion Boundaries

The first observations of inversion boundaries were made on the cubic noncentrosymmetric  $\chi$  phase in the alloy system Fe-Cr-Mo-Ti [34–36] (see Fig. 35). The contrast at this type of boundary requires some specific discussion. It has been found experimentally that under the appropriate diffraction conditions the domain structure can be revealed by domain contrast as well as by interface contrast. Inversion



**Figure 41.** Interfaces in  $\text{Ni}_4\text{Mo}$  revealed by different imaging modes: (a) BF image; (b) in double-diffraction spot; (c) DF image in 020 of one variant; (d) DF image in 020 of second variant.

domains have a common lattice and hence there is no splitting. The structures are related by an inversion operation, that is the reflections  $\mathbf{H}$  in one domain and  $-\mathbf{H}$  in the other domain are always excited simultaneously and to the same extent. The moduli of the structure factors of simultaneously excited reflections  $\mathbf{H}$  and  $-\mathbf{H}$  are always the same according to Friedel's law  $I_{\mathbf{H}} = I_{-\mathbf{H}}$ . The phases  $\alpha_{\mathbf{H}}$  and  $\alpha_{-\mathbf{H}}$  are different for most reflections since the structure is noncentrosymmetric. For a noncentrosymmetric crystal the phases associated with the Fourier coefficients of the imaginary part of the lattice potential need not be equal to those associated with the Fourier coefficients of the real part.

Serneels et al. [34] have shown that domain contrast arises as a result of the violation of Friedel's law in dark field images in non-centrosymmetric crystals under multiple beam conditions, along a zone which does not produce centrosymmetry in projection. This means that the zone axis cannot be a symmetry axis of even order. For example, in the  $\chi$  phase domain contrast is produced when the zone axis is along the three-fold rotation axis.

Interface contrast arises as a result of the difference in phase of the structure factors associated with the different domains. The interfaces are imaged as  $\alpha$  fringes, the lattices being parallel in the

domains. The difference in phase between the Fourier coefficients related to real and imaginary parts of the lattice potential leads to weak interface contrast, even under two-beam conditions [34].

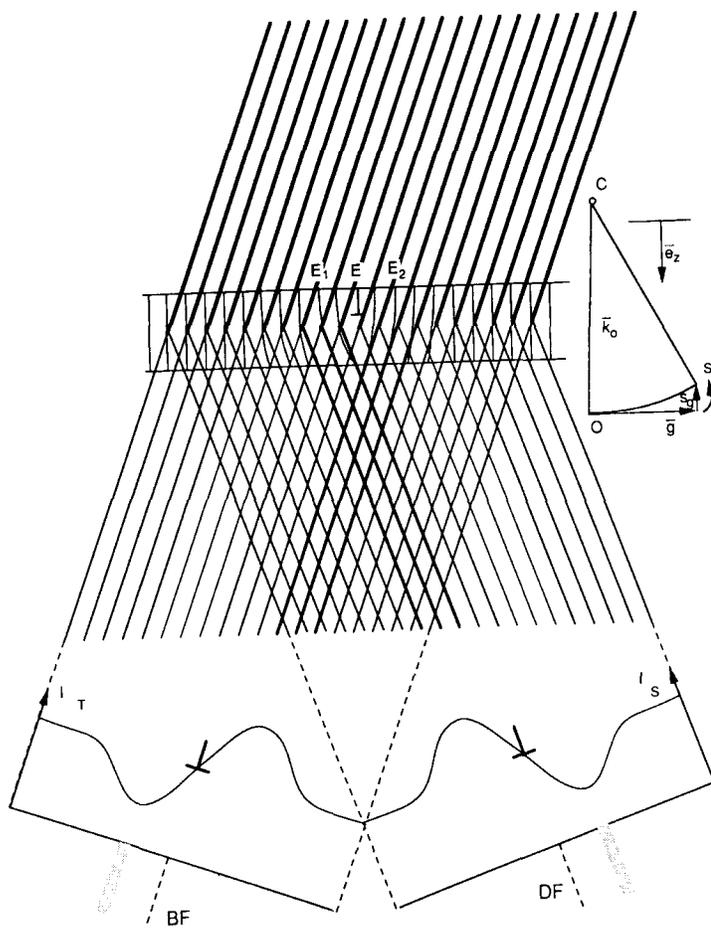
### 1.1.1.17 Dislocation Contrast

#### Intuitive Considerations

Dislocations are usually visible as dark lines in two-beam diffraction contrast

images, made with small values of the deviation parameter. When applying the weak beam method, that is for large values of  $s$ , the dislocations appear as bright lines on a darker background.

The two-beam image formation at dislocations can easily be understood on intuitive grounds [18–20]. The foil represented in Fig. 42 contains an edge dislocation in E. The lattice planes used for imaging are indicated schematically. Due to the presence of the dislocation the lattice planes in the vicinity of E are slightly curved and inclined in opposite



**Figure 42.** Intuitive model for the origin of contrast at an edge dislocation E. The thickness of the lines is a measure of the intensity of the electron beams. BF, bright field; DF, dark field.

senses to the left and right of E. As the specimen is a thin foil, the Bragg condition is relaxed; the reciprocal lattice nodes have become 'relrods'. We can therefore assume that diffraction occurs, even though the Bragg condition is only approximately satisfied with  $s < 0$  in the part of the foil which is not affected by the presence of the dislocations. On the left of the dislocation, at  $E_1$ , the rotation of the lattice planes is then such that, locally, the Bragg condition is better satisfied, that is  $|s|$  is smaller, and hence the diffracted beam will be more intense than in the perfect parts of the foil. On the right of the dislocation in  $E_2$ , the lattice rotation is in the opposite sense and hence the diffracted beam locally will be weaker than in the perfect part of the foil. The relative intensities of the diffracted beams are indicated schematically by lines of different widths in Fig. 42. Since no electrons are lost, the transmitted beam will be depleted where the scattered beam is enhanced.

Selecting the diffracted beam by means of an aperture and magnifying the corresponding diffraction spot will produce a map of the intensity distribution in this beam. This map will reveal a lack of intensity (i.e., a dark line) to the right of the dislocations in  $E_2$  and an excess of intensity over the background in  $E_1$ . The dislocation will thus be imaged as a bright-dark line pair. This image is called a *dark field image*.

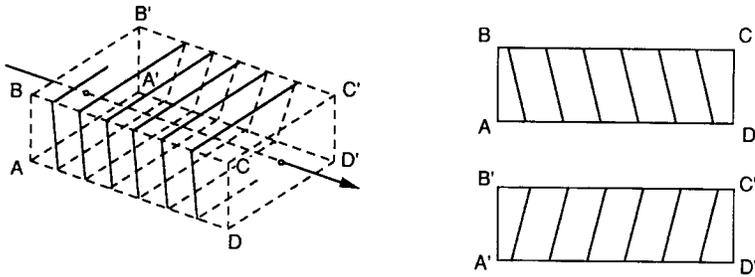
When selecting the transmitted beam a similar intensity map can be produced by magnifying the intensity distribution in the direct beam. Such an image is called a *bright field image*; in this approximation it is complementary to the dark field image, bright and dark lines being interchanged. Bright and dark field images are,

in fact, only small portions of strongly magnified diffraction spots, the intensity distribution being the image.

The possibility of forming such images is a consequence of the 'local' character of electron diffraction. Electrons only sense a narrow column of material because the Bragg angles are small, electron diffraction is strongly peaked forward, and the foil is thin. The columns form, in a sense, the 'pixels' of the image. The assumption that electrons travel in narrow columns is the basis of the 'column approximation' introduced above.

The same type of reasoning can be used to demonstrate that screw dislocations produce a line image. As a consequence of the presence of the screw dislocation, the families of lattice planes intersecting the dislocation line are transformed into helical surfaces. To the left and right of the dislocation the lattice planes are slightly inclined in opposite senses, and hence the local diffraction conditions are different to the left and the right. Again a bright-dark line is produced (Fig. 43).

In both cases, the dark line image is *not* produced at the dislocation core, but in a slightly displaced position called the *image side*. Changing the diffraction conditions so as to make  $s > 0$  in the foil part which is far away from the dislocations, changes the image side, as can be demonstrated by the same reasoning used above. Also, changing  $\mathbf{g}$  to  $-\mathbf{g}$  changes the image side, since now reflection takes place from the other side of the lattice planes. Finally, changing the sign of the Burgers vector changes the sense of inclination of the lattice planes on a given side of the dislocation, and hence also changes the image side. Summarizing, we can say that the image side depends on the sign of  $p = (\mathbf{g} \cdot \mathbf{b})s$ .



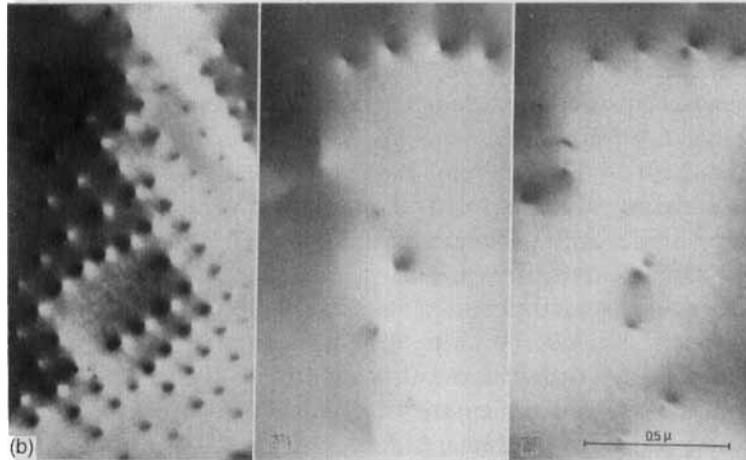
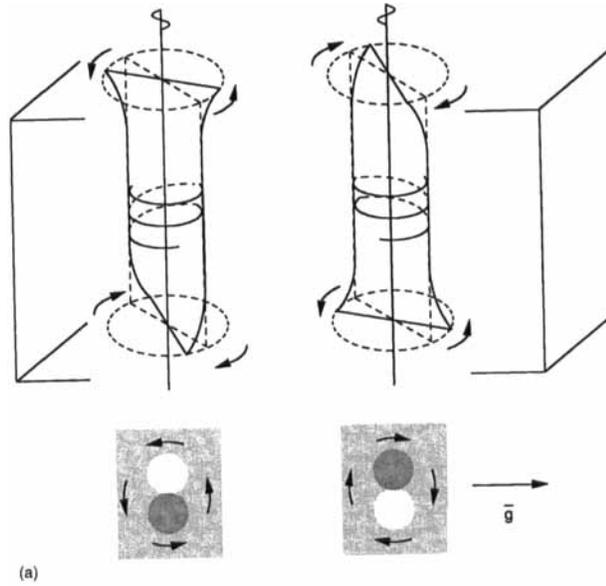
**Figure 43.** Geometry of lattice planes in the vicinity of a screw dislocation, leading to image contrast.

This rule becomes undetermined if  $\mathbf{g} \cdot \mathbf{b} = 0$ . The relation  $\mathbf{g} \cdot \mathbf{b} = 0$  is, in fact, the criterion for the absence of contrast. It expresses the fact that no image is produced when diffraction occurs by the lattice planes which are left undeformed by the presence of the dislocation. To a first approximation, all displacements around a dislocation are parallel to the Burgers vector and are thus parallel to the lattice planes for which  $\mathbf{g} \cdot \mathbf{b} = 0$ . This *extinction criterion* is strictly valid for screw dislocations in an elastically isotropic medium, for which all displacements are parallel to  $\mathbf{b}$ , but it is only a first approximation for edge dislocations. Deviations occur even for screws in strongly anisotropic media, the reason being that the actual extinction criterion is  $\mathbf{g} \cdot \mathbf{R} = 0$ . The displacement field of an edge dislocation contains a component perpendicular to the glide plane, which causes some residual contrast even if  $\mathbf{g} \cdot \mathbf{b} = 0$ , as we shall discuss below.

Some contrast may also result, even though  $\mathbf{g} \cdot \mathbf{b} = 0$ , from the fact that the specimen is a thin foil. The presence of dislocations in a thin foil modifies the displacement field as a result of surface relaxation effects and this may produce contrast. For instance, a pure screw dislocation parallel to the incident beam and perpendicular to the foil surfaces is not expected to produce any contrast since

$\mathbf{g} \cdot \mathbf{b} = 0$  for all active  $\mathbf{g}$  vectors. However, such dislocations produce a dark–bright dot contrast which was attributed to the lattice twist. It was shown by Eshelby and Stroh [35] that, close to the emergence point of a screw dislocation in the foil surfaces, significant elastic relaxation takes place which transforms the lattice planes parallel to the dislocation line into helical surfaces, the sense of the helical twist being determined by the sign of the screw dislocation. This helical twist produces a bright–dark dot pair because on one side of the emergence point the lattice planes are tilted *into* the Bragg condition and on the other side they are tilted *out of* the Bragg condition. The line joining the bright–dark dot pair is perpendicular to  $\mathbf{g}$  (Fig. 44a). Depending on the sense of the helical twist (i.e., on the sign of the screw dislocation), the dot pair is bright–dark or dark–bright. The sign of the screw dislocation can thus be determined from such images. Images of this type observed in a platinum foil are shown in Fig. 44b.

An edge dislocation parallel to the foil surfaces and with its glide plane also parallel to the foil causes a slight misorientation of the two crystal parts separated by the dislocation. The tilt angle,  $\theta$ , depends on the foil thickness and on the position of the dislocation within the foil, being a maximum,  $\theta_{\max} = b/t$ , where  $t$  is the foil



**Figure 44.** (a) Schematic illustration of the surface relaxation around the emergence points of screw dislocations in the foil surfaces. This surface relaxation produces image contrast for  $g \cdot b = 0$ . (b) Screw dislocations viewed end-on in platinum [59].

thickness, if the dislocation is in the central plane. As a result of this slight ‘buckling’ of the foil, a brightness difference is produced between the two crystal parts, separated by the dislocation. The tilt angle,  $\theta$ , can be measured by the displacement of the Kikuchi lines; its sense depends on the sign of the dislocation, and therefore a knowledge of  $\theta$  allows the sign of the

dislocation to be determined (Fig. 45) [18–20].

An edge dislocation viewed end-on along the beam direction produces contrast because in the vicinity of the dislocation the interplanar spacing is slightly modified,  $g$  changes in length and in orientation and, consequently, the diffraction conditions also change. Along a column

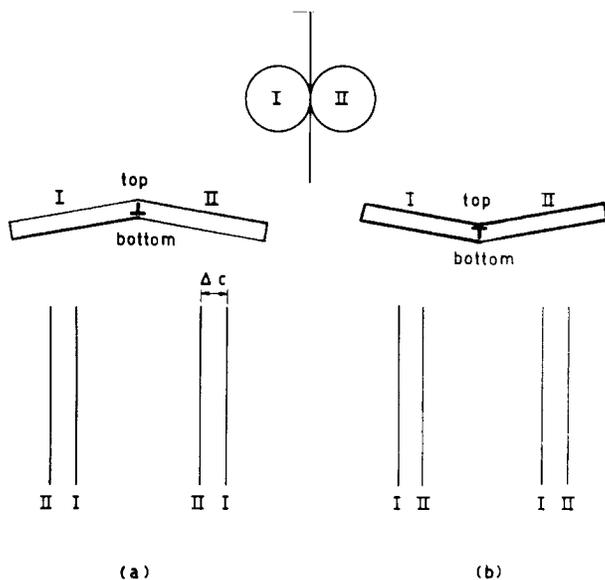


Figure 45. Buckling of a thin foil due to the presence of an edge dislocation parallel to the foil surfaces. The positions of the Kikuchi lines in areas I and II are shown schematically.

parallel to the dislocation (i.e., along  $z$ ),  $s$  remains constant, but  $s$  becomes a function of  $x$  and  $y$ , chosen in the foil plane. As a result, the scattered and transmitted intensities depend on the column positions, that is an image is produced. The contours of equal  $s$  (i.e., of equal brightness) are shown in Fig. 46; they image the strain field around the edge dislocation [37, 38].

According to the simple  $\mathbf{g} \cdot \mathbf{b} = 0$  criterion, an edge dislocation with its supplementary half-plane parallel to the foil plane or, stated otherwise, with its Burgers vector,  $\mathbf{b}$ , along the incident beam, would not produce any contrast. Due to the presence of the 'bump' in the glide plane (i.e., the component of the displacement field) towards the supplementary half-plane, perpendicular to the glide plane,  $\mathbf{g} \cdot \mathbf{R}$  is not zero for all  $\mathbf{g}$  vectors perpendicular to  $\mathbf{b}$  and some contrast is produced. Prismatic dislocation loops in planes parallel to the foil plane have this configuration. The displacement field of such loops

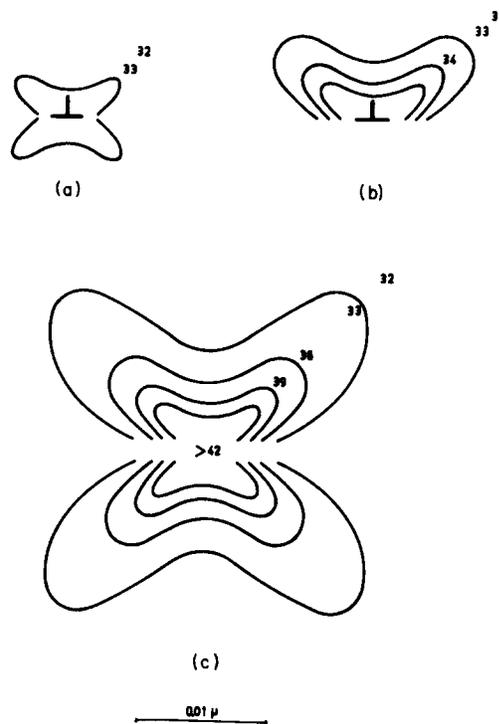
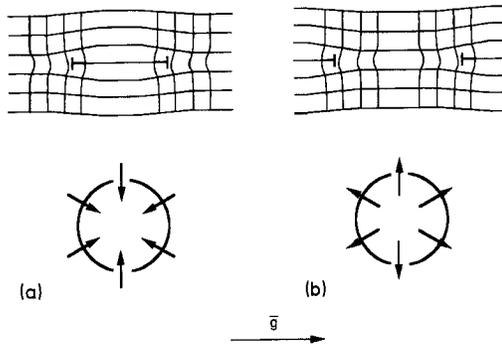


Figure 46. Contours of equal deviation parameter  $s$  in the vicinity of an edge dislocation viewed end-on [37, 38].



**Figure 47.** Radial displacement field around prismatic loops.

now contains a radial component  $R_r$  which is inward or outward for vacancy and interstitial loops (Fig. 47), respectively, as well as a normal component  $R_n = b$ .

For a diffraction vector  $g$  parallel to the foil plane the dot product with the normal component  $g \cdot R_n = g \cdot b$  will be zero everywhere along the loop. However,  $g \cdot R_r$  varies along the loop and vanishes only along the two diametrically opposite segments where  $g$  is perpendicular to  $R_r$ , as represented in Fig. 47. As a result, there will be two short segments only along which complete extinction occurs; the 'line of no contrast' joining these two segments is perpendicular to the active  $g$  vector.

Somewhat against intuition, one finds that parallel dislocation lines with the same Burgers vector do not necessarily exhibit the same contrast, especially when they are close one to the other, as in a ribbon. One of the lines is usually imaged as a darker line than the other(s); which line will exhibit the strongest line contrast depends on the sign of  $s$  and on the sense of  $g$ . The effect is particularly striking in triple ribbons in face-centered-cubic, low stacking fault energy alloys, and in

graphite. An analytical theory, based on the kinematical diffraction theory, allows one to account satisfactorily for the observations, on noting that the total strain field of a triple ribbon is different from that resulting from the mere superposition of the strain fields of three isolated dislocations [39].

### Semi-Quantitative Considerations

It is often useful to be able to predict semiquantitatively, or even only qualitatively, the two-beam image characteristics to be expected for a given defect. For dislocation lines this is possible within the framework of the dynamical theory, including anomalous absorption, by referring to the analytically soluble case of the stacking fault.

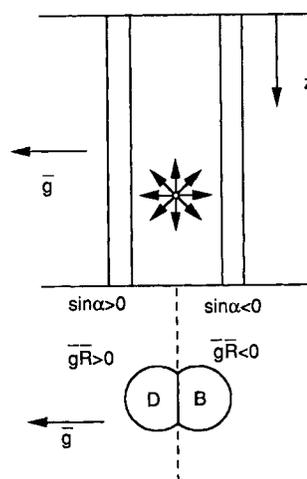
Consider, for example, an inclined edge dislocation with its Burgers vector parallel to the foil plane and an active reflection such that  $g \cdot b = 1$ . A sketch of the displacement field of the dislocation is represented schematically in Fig. 48. The bright field and dark field image profiles can be obtained by considering a row of columns along a line perpendicular to the dislocation line, and computing for each of these columns the amplitude of the transmitted and scattered beams.

We first note that these amplitudes are, to a large extent, determined by the phase relation between the top and bottom ends of the columns and not so much by the details of the variation of this phase along the column. In any case, the phase varies rapidly only in the close vicinity of the dislocation core. We therefore accept as a reasonable approximation that the amplitudes emerging from such a column will be



that the contrast will oscillate. Along the strips where  $\sin \alpha > 0$  (i.e., on the left of the dislocation) the first extremum behind the entrance face will be a maximum in the bright field image, whereas on the left of the dislocation, where  $\sin \alpha < 0$ , the first extremum will be a maximum. Near the exit face the last fringe in the bright field image will be the same as the first, on the left as well as on the right, of the dislocation. The dislocation contrast thus oscillates as a function of depth, since the maxima and minima in brightness on the two sides of the dislocation (i.e., for  $\sin \alpha > 0$  and  $\sin \alpha < 0$ ) are in antiphase. The oscillations will be most pronounced for foil thicknesses equal to  $(2k + \frac{1}{2})t_g$  or  $(2k + \frac{3}{2})t_g$  as for these thicknesses the stacking fault fringes in the bright field image vary strongly in brightness (see Fig. 30). In the dark field image the brightness variation is less pronounced for the same thickness; on the other hand, the variation is now more pronounced for foil thicknesses of  $2kt_g$  and  $(2k + 1)t_g$ . At the same time, the difference between the brightness for  $\sin \alpha > 0$  and  $\sin \alpha < 0$  is more pronounced in the bright field for thicknesses of  $(2k + \frac{1}{2})t_g$  and  $(2k + \frac{3}{2})t_g$  than for thicknesses of  $2kt_g$  or  $(2k + 1)t_g$ . In the former case the contrast is 'oscillating', whereas in the latter case it is 'dotted'; the inverse is true for the dark field image. The stacking fault fringe profiles suggest in the same manner that the bright field image will be similar to the dark field image close to the front surface, but quasi-complementary close to the exit surface.

As a second example, we consider the image associated with the displacement field around a spherical inclusion with  $\epsilon > 0$ , as described by Eq. (46) [40–42]. Also in this case we can deduce from



**Figure 49.** Strain field associated with a spherical inclusion. D, dark; B, Bright.

intuitive considerations the areas in which  $\sin \alpha > 0$  and those in which  $\sin \alpha < 0$ , and hence conclude for a defect close to the surface (within the first extinction distance) which area will be bright and which will be dark. We consider, in particular, the spherically symmetrical displacement or strain field represented in Fig. 49. A line (or plane) of no contrast, along which  $\mathbf{g} \cdot \mathbf{R} = 0$  separates two regions, one in which  $\mathbf{g} \cdot \mathbf{R} > 0$  and one in which  $\mathbf{g} \cdot \mathbf{R} < 0$ . Since  $|\mathbf{R}|$  is small compared to a lattice vector, if  $\mathbf{g} \cdot \mathbf{R} > 0$  then  $\sin \alpha > 0$  and if  $\mathbf{g} \cdot \mathbf{R} < 0$  then  $\sin \alpha < 0$ . That this is so can be deduced from the consideration that, if  $\mathbf{g} \cdot \mathbf{R}$  is positive for all  $z$  values along the column, the integrated phase difference between top and bottom of the column is positive but smaller than  $\pi$  and hence  $\sin \alpha > 0$ , the fastest phase change occurring at the level of the inclusion. The brightness at the exit end of the column is then the same as that of a stacking fault, assuming the effective phase shift,  $\alpha_{\text{eff}}$ , to occur at the level of the defect.

The value of  $\alpha_{\text{eff}}$  decreases with increasing distance from the inclusion and changes sign along the line of no contrast. The image characteristics of stacking faults in sufficiently thick foils, close to the surface, allows us to deduce the dark field image of this kind of defect when close to the back surface. The last fringe in the dark field image of a stacking fault for which  $\sin \alpha > 0$  is dark. We can thus conclude that  $\mathbf{g}$  points towards the dark lobe for an inclusion with  $\varepsilon > 0$  situated close to the back surface. Black and dark are reversed for  $\varepsilon < 0$ . The model also accounts for the periodic interchange with period  $t_g$  of bright and dark lobes with the depth position of the spherical inclusion.

### Kinematical Theory of Dislocation Contrast

Within the framework of the kinematical diffraction theory, image profiles of dislocations are obtained by inserting the adequate expression for the displacement field  $\mathbf{R}(\mathbf{r})$  in Eq. (50) and integrating along columns situated on lines normal to the dislocation line. Due to the symmetry of the displacement field of a dislocation, the profile so obtained is independent of the chosen line of columns for dislocations parallel to the surfaces of the foil.

For example, for a screw dislocation oriented along the  $y$  axis parallel to the foil plane and situated at a depth,  $d$ , behind the entrance face the displacement field is described, according to the isotropic linear elasticity theory by the expressions

$$R_x = 0 \quad (159a)$$

$$R_y = \frac{b\phi}{2\pi} \quad (159b)$$

$$R_z = 0 \quad (159c)$$

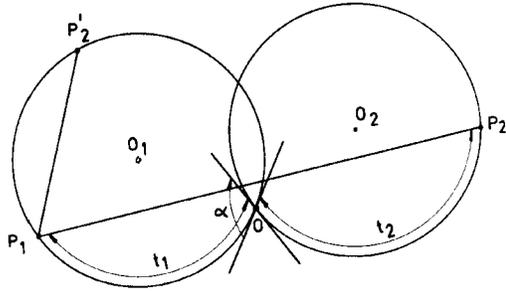
with  $\phi = \text{arctg}[(z - d)/x]$ , that is all the displacements are parallel to  $\mathbf{b}$ . The image profile is then obtained by performing the integration

$$A(\mathbf{g}) = F_g \int_0^{z_0} \exp(2\pi i s_g z) \times \exp(in) \text{arctg}\left(\frac{z-d}{x}\right) dz \quad (160)$$

where  $n \equiv \mathbf{g} \cdot \mathbf{b}$  for various values of the parameter  $x$ , and  $z_0$  is the foil thickness. After a number of approximations, the integrals can be obtained analytically in terms of Bessel functions.

In their discussion of image profiles of dislocations, Hirsch et al. [10, 11] and Gevers [12] made extensive use of amplitude-phase (A-P) diagrams. We shall follow the same type of reasoning, since this allows us to identify more clearly the approximations and limitations of the theory. The integration along a column is represented graphically by the vector sum of the elementary contributions due to the slices  $dz$  along the column. In a perfect crystal we have seen that the vectors representing the amplitudes scattered by successive slices enclose a constant angle  $d\theta = 2\pi s dz$ , as a result of the constant phase difference between successive slices  $dz$ . These small vectors form a regular polygon, which in the limit for  $dz \rightarrow 0$  becomes an arc of a circle with radius  $\frac{1}{2\pi s}$ . The length of the circular arc is equal to the column length, and the amplitude scattered by the column is given by the length of the vector joining the two end-points of the circular arc.

When a stacking fault is present, a discontinuous phase change  $\alpha = 2\pi \mathbf{g} \cdot \mathbf{R}$  occurs in each column at the level of the stacking fault. This is reflected in the A-P



**Figure 50.** Amplitude-phase diagram for a column intersecting a stacking fault. The thickness of the two crystal parts are  $t_1 \equiv z_1$  and  $t_2 \equiv z_2$  [13, 14].

diagram by a relative rotation over an angle  $\alpha$  of the two circles representing the A-P diagrams of the perfect parts, the tangents enclosing an angle  $\alpha$  (Fig. 50). The amplitude diffracted by the faulted crystal is then given by the vector joining the end-points  $P_1$  and  $P_2$ . If we choose the origin of the diagram at the position of the stacking fault, the two circular arcs have lengths equal to the front ( $z_1$ ) and rear part ( $z_2$ ) of the foil:  $z_1 + z_2 = z_0$ , where  $z_0$  is the foil thickness. For an inclined stacking fault the end-points  $P_1$  and  $P_2$  corresponding to successive columns along a profile shift continuously over the same arc length in the same sense. The resulting amplitude thus varies periodically with a depth period  $1/s$ , describing the stacking fault fringes, and is represented analytically by Eq. (52).

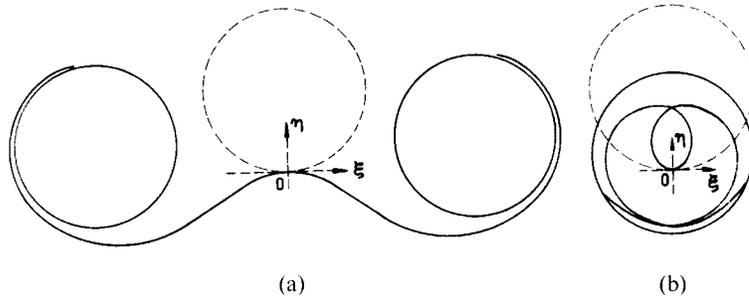
The A-P diagram for a column intersecting a domain boundary also consists of two circular arcs with lengths equal to the thicknesses of front and rear parts; they join smoothly with a common tangent at the level of the boundary, but they have different radii,  $\frac{1}{2\pi s_1}$  and  $\frac{1}{2\pi s_2}$ , since the  $s$  values are different in the two perfect parts. In addition, in this case an inclined domain boundary will produce a fringe

pattern with a depth quasiperiod, which is somewhat variable over the width of the fringe pattern between  $1/s_1$  and  $1/s_2$ . The A-P diagram is the geometrical representation of the analytical expression Eq. (53).

In the A-P diagram for a foil containing a dislocation the phase difference between successive slices of the column at levels  $z$  and  $z + dz$  is no longer a constant  $d\phi = 2\pi s dz$ , since a supplementary phase difference results from the displacements described by  $\alpha(x, y, z)$ . Depending on the signs of  $x$ ,  $s$ , and  $z$ , this additional phase shift will either be added or subtracted; its magnitude depends on  $x$  and  $z$  and is given in the simple case of the screw dislocation by  $n \cdot d[\arctg(z/x)]$ . For  $z \gg x$ , this additional shift becomes zero and the final shape becomes again a circle with radius  $\frac{1}{2\pi s}$  as for the perfect crystal. Close to the dislocation and for  $s$  and  $nx$  having the same sign, that is for  $n\beta > 0$  (with  $\beta = 2\pi s x$ ), the quantity  $n \arctg(z/x)$  has the same sign as  $2\pi s z$  and the angle between two successive vectors is now larger than  $2\pi s dz$ , at least near the depth position of the dislocation, which is chosen as the origin of the diagram (i.e., of the  $z$  axis). As  $z$  becomes larger the angle again approaches  $2\pi s dz$ . The resulting curve will be a wound-up spiral which gradually tends to a circle, approaching it from the interior, the circle being the limiting curve (Fig. 51b). If, on the other hand,  $s$  and  $nx$  have opposite signs (i.e., for  $n\beta < 0$ ),  $\arctg(z/x)$  and  $2\pi s z$  have opposite signs and the resulting angle between successive vectors will be smaller than the value  $2\pi s dz$  in the perfect crystal by

$$n \cdot d[\arctg(z/x)] \equiv n[x/(x^2 + z^2)] dz.$$

Again, as  $z$  becomes large, the additional phase difference tends to zero and the



**Figure 51.** Amplitude-phase diagram for a column passing close to a dislocation core ( $n = 2$ ) [13, 14]: (a) unwound spiral; (b) wound-up spiral.

curve approaches a circle with radius  $\frac{1}{2}\pi s$ . The A-P diagram is now an unwound spiral approaching the limiting circles from the outside, as shown in Fig. 51a. The scattered amplitude is again obtained by taking an arc proportional to  $z_1$  on this curve, in the negative sense leading to  $P_1$  and an arc proportional to  $z_2$  in the positive sense leading to  $P_2$ . The vector  $\overline{P_1P_2}$  is then proportional to the scattered amplitude for the given column, that is for a given  $x$  value. Since  $x$  has different signs on the two sides of the dislocation, the A-P diagram will be an unwound spiral on one side and a wound-up spiral on the other. The vector representing the diffracted amplitude will clearly be larger for those columns for which the distance between the centers of the two limiting circles will be the largest, that is the amplitude will be largest on that side of the dislocation where the A-P diagram is an unwound spiral. This is the side where in the bright field image a dark line will be observed, called the image side (see Sec. 1.1.18.2 of this Chapter). We note that the A-P diagram depends only on  $n \equiv \mathbf{g} \cdot \mathbf{b}$  and on the product  $\beta \equiv 2\pi s x$ , but not on  $s$  and  $x$  separately. This is consistent with the fact that changing the sign of  $s$  changes the image side. In principle, constructing a sufficient number of A-P diagrams allows one to deduce the image profiles. It is clear

that for an inclined dislocation line the length of the arcs to be taken along the spiral-shaped A-P diagrams will vary continuously with the position along the dislocation, one increasing, the other one decreasing. Hereby the end-points of the A-P diagram, which determine the scattered amplitude, will in general vary periodically as the end-points describe the limiting circles. This oscillatory behavior was suppressed in the approximation introduced by Hirsch et al. [11]. The assumption was made that the square of the separation of the centers of the limiting circles is a convenient measure of the scattered intensity. This is a reasonable assumption if  $s$  is sufficiently large so that the limiting circles acquire a small radius compared to the separation of their centers. For screw dislocations Hirsch et al. [11] obtained the computed profiles shown in Fig. 24 for different values of  $n$ . Similar calculations, using the same approximations, have been performed by Gevers [12] for perfect as well as partial dislocations of edge and mixed character. The results for pure edge dislocations are shown in Fig. 25.

### The Weak-Beam Method

The limitations of kinematical theory apply to the results of the preceding

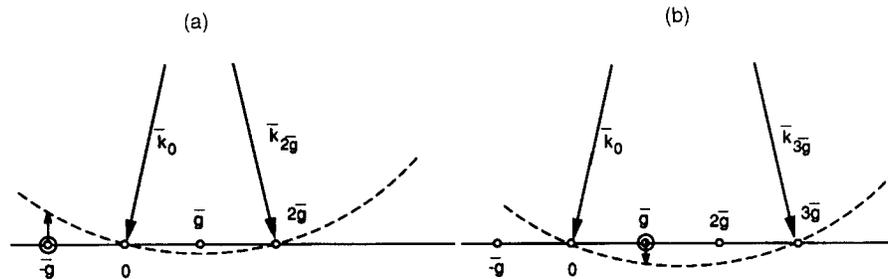
chapter and we therefore conclude that the computed image profiles are only valid for very thin foils and for large  $s$  values.

From the image profiles shown in Figs. 24 and 25, we can deduce that when  $s$  is large the same  $\beta (= 2\pi sx)$  value is reached for small  $x$ . This implies that for large  $s$  values the peak shift and the peak width will be small. This effect, which is consistent with the observations, is systematically exploited in the weak-beam method [43]. It allows one to obtain very well localized and sharp images of the partial dislocations in narrow ribbons, as required for the measurement of stacking fault energies. Unfortunately, with increasing  $s$  value the image contrast decreases and long exposure times are needed to record the image. In practice, a reasonable trade-off between image resolution and exposure time seems to be achieved for  $s$  values of the order of  $0.2 \text{ nm}^{-1}$  for 100 kV electrons.

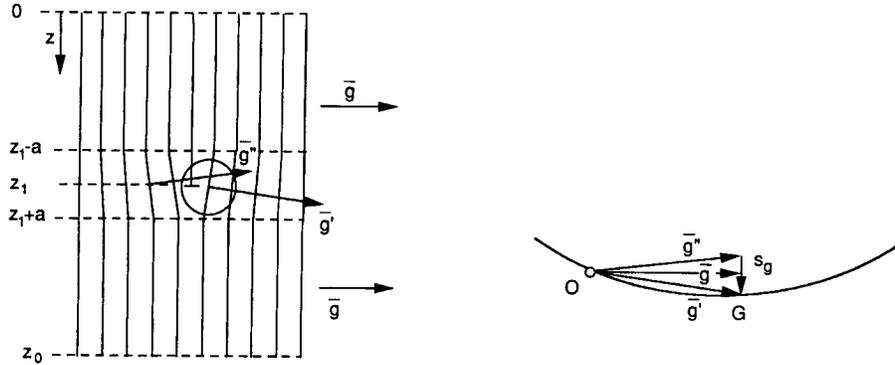
Usually a high order reflection of  $3g$  or  $4g$  is brought in the exact Bragg position and a dark field image is made in the reflection  $g$ . Alternatively, a low order reflection,  $g$  or  $2g$ , may be excited and  $-g$  be used for imaging. These imaging conditions are represented in Fig. 52. In order to realize these diffraction conditions

exactly, the Kikuchi pattern is of great help; moreover, it allows the measurement of  $s$ . In weak-beam images the depth period of extinction contours and of stacking fault fringes is given approximately by its kinematical value  $1/s_g$ . Using such large  $s$  values it is possible to image, for instance, antiphase boundaries in alloys as fringe patterns, even though the extinction distance of the superlattice reflection used is larger than the foil thickness.

The kinematical theory allows us to derive approximate expressions for the peak width and peak positions of weak-beam dislocation images [43–45]. The columns close to the dislocation core can be considered as consisting of three parts (Fig. 53). The central part contains the dislocation, and the two other parts are perfect. In the central part the lattice planes of interest are inclined with respect to their orientation in the perfect parts, in such a way that somewhere close to the dislocation core the local deviation parameter is much smaller than in the perfect part. The scattered intensity will then mainly originate from this region, producing a bright peak on a darker background in the dark field, weak-beam image. The amplitude scattered by a column at  $x$  is



**Figure 52.** Imaging conditions used in the weak-beam method: (a)  $s = 0$  for  $2g$ , image is made in  $-g$ ; (b)  $s = 0$  for  $3g$ , image is made in  $g$  [44].



**Figure 53.** Foil containing an edge dislocation. Model used in discussing dislocation contrast according to the weak-beam method. The foil is assumed to consist of three lamellae: 1 [ $0 \leq z \leq (z_1 - a)$ ] and 3 [ $(z + a) \leq z \leq 70$ ] are perfect; part 2 [ $(z_1 - a) < z < (z_1 + a)$ ] contains the dislocation [45].

given by

$$A \propto \int_0^{z_0} \exp [2\pi i [s_g z + \mathbf{g} \cdot \mathbf{R}(x, z)]] dz \quad (161)$$

This integral can be split in three parts corresponding to the three lamellae in the model shown in Fig. 53:

$$A \propto \int_0^{z_1-a} \exp(2\pi i s_g z) dz + \int_{z_1-a}^{z_1+a} \exp[2\pi i (s_g z + \mathbf{g} \cdot \mathbf{R})] dz + \int_{z_1+a}^{z_0} \exp(2\pi i s_g z) dz \quad (162)$$

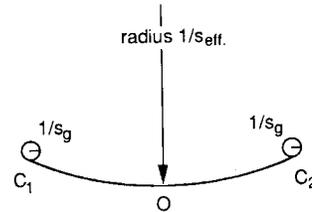
The first and third integral refer to the perfect parts; they do not depend on the presence of the defect. Since  $s$  is large in these parts their contribution is small. Their A–P diagrams consist of small circles with a radius  $\frac{1}{2}\pi s$ . These two circles are connected by a circular arc with a much larger radius,  $1/s_{\text{eff}}$ , which is the A–P diagram of the central part. The amplitude scattered by the column is then given, to a good approximation, by the length of the segment joining the centers of the two small circles. This length is well approximated by the second integral, which we now consider (Fig. 54).

We can write the displacement function  $\mathbf{R}(x, z)$  as a Taylor expansion in the vicinity of the core position  $z = z_1$ :

$$\mathbf{R} = \mathbf{R}(z_1) + (z - z_1) \left( \frac{\partial \mathbf{R}}{\partial z} \right)_{z_1} + \frac{1}{2} (z - z_1)^2 \left( \frac{\partial^2 \mathbf{R}}{\partial z^2} \right)_{z_1} + \dots \quad (163)$$

Retaining only the first two terms, the second integral can be written as

$$\exp \left\{ 2\pi i \left[ \mathbf{R}(z_1) - z_1 \left( \frac{\partial \mathbf{R}}{\partial z} \right)_{z_1} \right] \cdot \mathbf{g} \right\} \times \int_{z-a}^{z+a} \exp \left\{ 2\pi i \left[ s_g + \mathbf{g} \cdot \frac{\partial \mathbf{R}}{\partial z} \right] z \right\} dz \quad (164)$$



**Figure 54.** Amplitude–phase diagram for a column close to the dislocation core according to the weak-beam method [43].

This expression will be a maximum if the modulus of the integrandum is unity, that is for the value of  $x$  given by

$$s_g + \frac{\partial R}{\partial z} \cdot \mathbf{g} = 0 \quad (165)$$

This condition is equivalent to the statement  $s_{\text{eff}} = 0$  [see Eq. (61)].

Introducing the displacement field for edge and screw dislocations, adopting the FS/RH convention (see Sec. 1.1.18.2 of this Chapter) leads to the peak position  $x = x_m$  with

$$x_m = \left(1 + \frac{K}{2(1-\nu)}\right) \left(-\frac{\mathbf{g} \cdot \mathbf{b}}{2\pi s_g}\right) \quad (166)$$

The parameter is  $K = 1$  for an edge dislocation and  $K = 0$  for a screw dislocation;  $\nu$  is Poisson's ratio. In this approximation the peak position does not depend on the foil thickness or on the depth position of the dislocation. The image side, that is the sign of  $x_m$ , is clearly determined by the sign of the product  $(\mathbf{g} \cdot \mathbf{b})s_g$ .

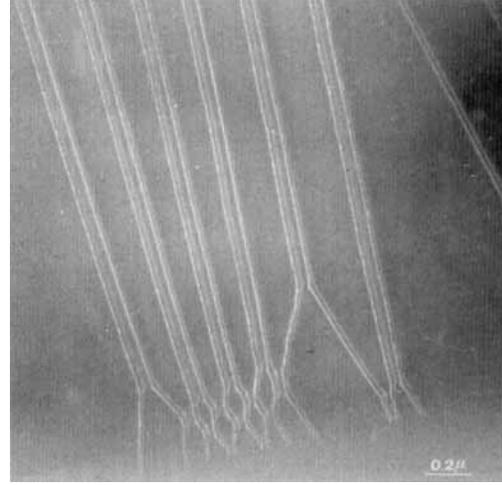
Using the same model, the peak width at half maximum can be deduced from the kinematical approximation. For  $\mathbf{g} \cdot \mathbf{b} = 2$  one finds

$$\Delta x = \frac{0.28}{|s_g|[1 + K/2(1-\nu)]} \quad (167)$$

With  $\nu = \frac{1}{3}$  and  $|s_g| = 0.2 \text{ nm}^{-1}$ , one finds that  $\Delta x \simeq 2.5 \text{ nm}$  for an edge dislocation.

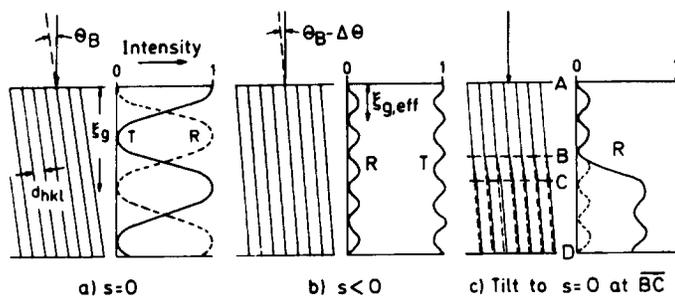
With increasing value of  $\mathbf{g} \cdot \mathbf{b}$  the image peak moves away from the core position. The larger the value of  $\mathbf{g} \cdot \mathbf{b}$  the larger the values of  $s_g$  needed in order to achieve the same precision in the image position. In practice, this limits the values of  $\mathbf{g} \cdot \mathbf{b}$  to  $\leq 2$ .

An example of a weak-beam image in the layered crystal  $\text{RuSe}_2$  is reproduced in Fig. 55.



**Figure 55.** Weak-beam image of four-fold dislocation ribbons in  $\text{RuSe}_2$ . (Courtesy of J. Van Landuyt.)

The dynamical theory, neglecting anomalous absorption for simplicity, leads to essentially the same qualitative results. In terms of this theory the scattered beam, in the first part along a column close to the dislocation on the image side, oscillates with a small amplitude and with a depth period given to a good approximation by the kinematical value  $1/s$  (Fig. 56). In the second part of this column, where  $s_{\text{eff}} \simeq 0$ , the Pendellösung oscillations acquire a large amplitude and a depth period approximated by  $t_g$ . As this second part is thin, only a fraction of an oscillation can develop, and in part three the amplitude of the oscillation and its depth period again become the same as in part one. However, the average intensity level has now become larger in part three, in particular for the column along which  $s_{\text{eff}}$  becomes zero at the level of the dislocation core. Hence the observed intensity at the exit face of that column will be larger than that for columns which are farther away from the dislocation, and will thus show up as a bright line.



**Figure 56.** Weak-beam image formation at a dislocation according to the dynamical theory [43].  $\xi_g \equiv t_g$ ;  $T$ , transmitted beam amplitude;  $R$ , scattered beam amplitude. (a)  $s = 0$ ; (b)  $s < 0$ ; (c) tilt to  $s = 0$  at  $\overline{BC}$ .

### 1.1.1.18 Dislocation Contrast: Dynamical Theory

#### Image Simulation

One-dimensional profiles and two-dimensional maps which describe quantitatively the experimentally observed images are only obtained by applying the dynamical theory including anomalous absorption. Equations (116) or Eqs. (114) have to be integrated, with  $\alpha$  being, in general, a function of  $x$  and  $z$ . For a screw dislocation located a distance  $d$  behind the entrance face this function becomes, for example,

$$\alpha = n \arctg[(z - d)/x] \quad (168)$$

with  $n = \mathbf{g} \cdot \mathbf{b}$ ;  $n$  is an integer for perfect dislocations, but it may be a fraction for partial dislocations. If  $d$  is considered to be a constant, a profile along  $x$  is sufficient to describe the image. For inclined dislocations  $d$  becomes an additional parameter and a two-dimensional map is desirable for comparison with experimental images.

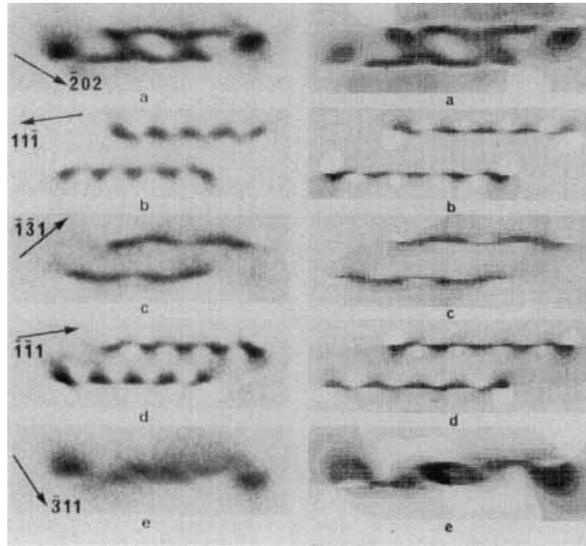
Analytical solutions are difficult, if not impossible, to obtain in most cases. Numerically computed image profiles are available for a number of representative dislocation configurations and are

reviewed below. A semiquantitative analytical discussion of the most striking image properties is possible [15, 45].

In principle, the computation procedure for profiles is a multislice method. It consists in considering a row of columns situated along the  $x$  axis. The integration is performed along a column (i.e., for a fixed  $x$  value) by further dividing this column in thin slices  $dz$ , each slice being considered as perfect with an  $s$  value [ $s_{\text{eff}} = s_g + (\mathbf{g} \cdot d\mathbf{R}/dz)_{x_0}$ ] which depends on  $z$ . The amplitudes of the scattered and transmitted beams can be obtained by the multiplication of a succession of response matrices of the type  $M(dz, s_{\text{eff}})$  (see Sec. 1.1.10 of this Chapter). This procedure is the implementation of the 'column approximation'.

The linear character of the Howie-Whelan system of equations [Eqs. (67) and (69)] and the fact that the displacement field of a dislocation is invariant for a translation along lines parallel to the dislocation line have been exploited by Head [47] and Humble [48] to speed up the computation procedures in order to make it possible to generate rapidly two-dimensional maps which can be compared directly with observed images (Fig. 57).

The different computer programs and subroutines required to generate two-dimensional intensity maps representing



**Figure 57.** Examples of the quantitative agreement that can be achieved between observed and computed dislocation images: (a) observed images for different diffraction vectors; (b) corresponding computer-generated images [48].

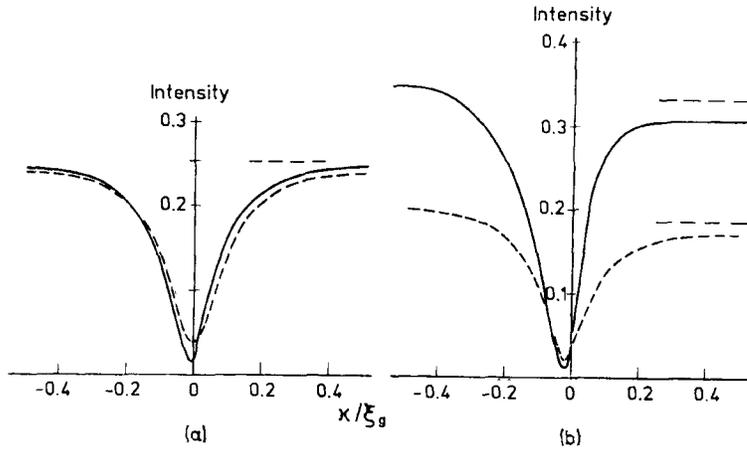
bright and dark field images for a wide variety of single and complex defects are described in full detail by Head et al. [48]. The displacement fields of the defects are computed numerically using anisotropic linear elasticity theory. Subroutines allow one to determine the geometry of the foil, the diffraction conditions, etc., using the Kikuchi line pattern as input data. The defect identification procedure is essentially a trial-and-error method based on the inspired guess of a model based on symmetry considerations and on qualitative geometrical characteristics of the type described in previous sections in this Chapter. The model is then tested and, where necessary, further refined by quantitative comparison of the observed and computer-generated images in which a small number of parameters is varied. As the computing time is short, a comparison exercise does not require an excessive computer effort. Complete listings of the software statements are provided by Head et al. [48].

Remarkable agreement between computed and observed images can be achieved, even for complex defect configurations such as the one illustrated in Fig. 57. One of the important conclusions that has emerged from such simulations is that the extinction criterion for dislocations  $\mathbf{g} \cdot \mathbf{b} = 0$  is only a first approximation and can lead to incorrect conclusions, especially in strongly anisotropic materials where the displacements around a dislocation are, in general, not parallel to  $\mathbf{b}$ , as implied in the analytical expressions based on isotropic linear elasticity.

### Survey of Results of the Two-Beam Dynamical Theory [12, 13, 38]

#### *Images of Screw Dislocations*

The images for  $n = 1$  and  $s = 0$  for a screw dislocation parallel to the foil surfaces and located in the central plane of the foil, exhibit a single dark peak very close to



**Figure 58.** Image for screw dislocations in the central plane of the foil with thickness  $8t_g$ : (a)  $n = 1$ ,  $s = 0$ ; (b)  $s = 0.3$ . (—) Bright field image; (---), dark field image [13, 14].

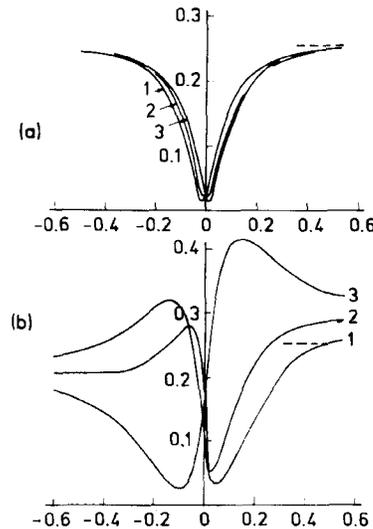
the position of the dislocation core, in both the bright field image and the dark field image (Fig. 58a). This is clearly in contradiction with the results of kinematical theory which predict complementary images; it is a result of anomalous absorption in thick crystals ( $5t_g$  to  $10t_g$ ). The peak width is of the order of  $0.3t_g$  to  $0.4t_g$ .

For screw dislocations close to the surface (Fig. 59) the image becomes clearly one-sided, the image side changing periodically with depth in the crystal.

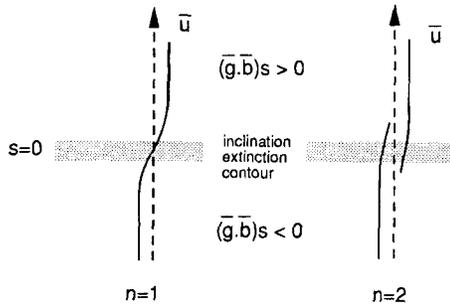
For  $s \neq 0$  and sufficiently large, the dark line is displaced away from the core position in the sense predicted by the intuitive reasoning given in Sec. 1.1.1.17 of this Chapter for the bright field image. The sense of the image shift does not depend on the depth position of the dislocation, but changes with the sign of  $s$ . As a result, the image will shift continuously sideways on crossing an inclination extinction contour in the manner represented in Fig. 60.

For  $s = 0$  as well as for small values of  $s$ , the image shift does depend on the depth positions  $z_0$ . Close to the surfaces the sense of the image shift alternates with a depth

period  $t_g$ , but in the central part of a thick foil the image shift is small. Inclined dislocations in thick foils will thus exhibit oscillating or alternating contrast close to the foil surfaces, but not in the central part (Fig. 61). The oscillations will be in 'phase'

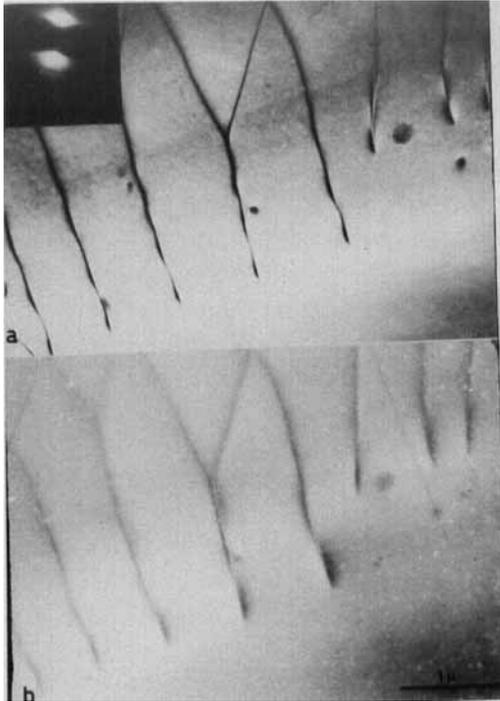


**Figure 59.** Bright field images for screw dislocations in foil with thickness  $8t_g$  at a distance  $d$  from entrance face ( $n = 1$ ;  $z_0 = 8t_g$ ;  $s = 0$ ) [13, 14]. (a) 1,  $d = 4t_g$ ; 2,  $d = 4.25t_g$ ; 3,  $d = 4.50t_g$ . (b) 1,  $d = 7.25t_g$ ; 2,  $d = 7.50t_g$ ; 3,  $d = 7.75t_g$ .



**Figure 60.** Image behavior on crossing an inclination extinction contour [13, 14]: (a)  $n = 1$ ; (b)  $n = 2$ .

close to the entrance face in bright and dark field, but in ‘antiphase’ close to the exit face. The effect is a consequence of anomalous absorption and it applies to all defect images. In thick foils the bright and



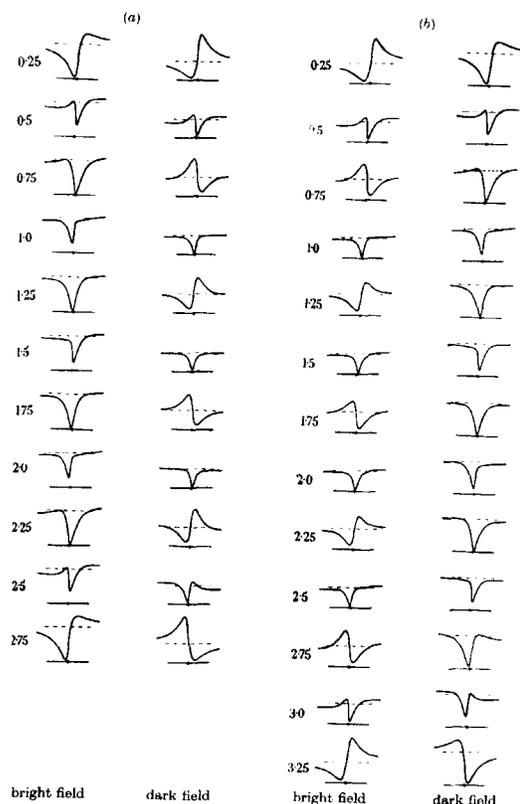
**Figure 61.** Dislocations in  $\text{SnS}_2$  exhibiting oscillating contrast on approaching the surface: (a) bright field image; (b) dark field image. (Courtesy of P. Delavignette.)

dark field images of defects are similar when situated close to the entrance face, but quasi-complementary close to the exit face. This has, in fact, already been pointed out for planar interfaces.

This oscillating contrast can be exploited to provide a depth scale, since for  $s = 0$  the oscillation period is exactly  $t_g$ . In particular, it allows a determination of the foil thickness in units of  $t_g$ , and it makes it possible to determine which end of the dislocation image corresponds to the vicinity of which foil surface (front or rear).

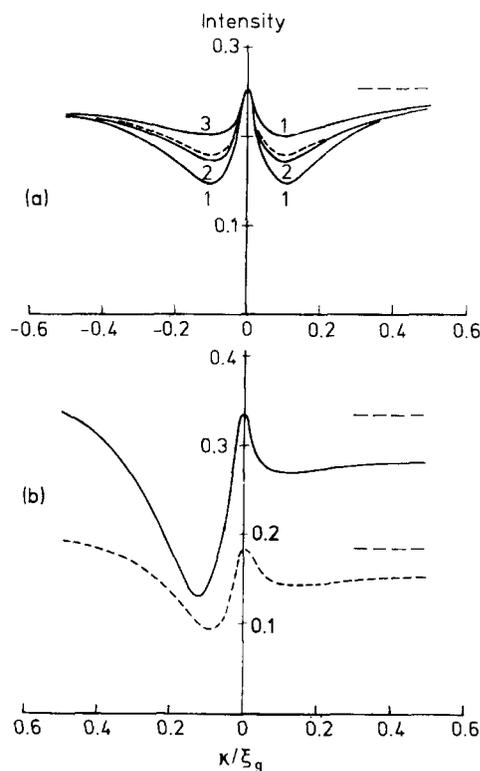
The occurrence of ‘dotted’ images at inclined dislocations can be understood intuitively, as pointed out above (see Sec. 1.1.1.17 of this Chapter) by noting that top and bottom parts of a column passing through the dislocation core are related by a phase jump of  $\pi$  which occurs at the level of the core. Along such columns the intensity profile for  $n = 1$  will be the same as that for a stacking fault with  $\alpha = \pi$ . Whether predominantly ‘dotted’ or ‘alternating’ contrast occurs depends on the thickness of the foil, as does the contrast for a fault with  $\alpha = \pi$ . This is illustrated in Fig. 62 which shows computed bright and dark field profiles for a screw dislocation with  $n = 1$  and  $s = 0$  in foils with a thickness of  $3t_g$  and  $3.5t_g$ , respectively. In the foil with a thickness of  $3t_g$  the bright field image is dotted and the dark field image is alternating, whereas for a foil with a thickness of  $3.5t_g$  the reverse is true.

When  $n = 2$  and  $s = 0$ , the image exhibits two dark peaks, one on each side of the dislocation core. The two peaks are different in strength, their relative strength alternates with a period  $t_g$  with the depth in the foil. These features are illustrated in Fig. 63a; they again lead to ‘oscillating’



**Figure 62.** Image profiles of a screw dislocation ( $s = 0$ ,  $n = 1$ ) at different depths below the entrance face of the foil in two foils of different total thickness: (a)  $3t_g$ ; (b)  $3.5t_g$ . The traces on the left-hand side are bright field images and those on the right are dark field images [13, 14].

contrast at inclined dislocations for  $s \approx 0$ . If  $s \neq 0$  the two peaks become strongly asymmetrical, as shown in Fig. 63b, the sense of the asymmetry depending on the sign of  $s$ . Except for  $s = 0$ , usually only one dark peak is observed as a consequence of the asymmetry. On intersecting an equi-inclination contour with  $s = 0$  the dislocation image will therefore behave as represented in Fig. 60. It is thus possible to deduce the value of  $n$  from the behavior of the image on intersecting an inclination contour. The value of  $n$  gives

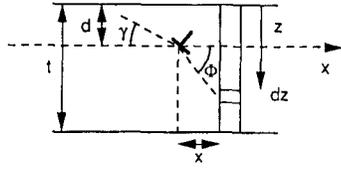


**Figure 63.** Image profile of a screw dislocation; foil thickness  $8t_g$  [13, 14]. (a)  $s = 0$ ,  $n = 2$ ; curves 1–3 correspond to depths of  $4t_g$ ,  $4.25t_g$  and  $4.5t_g$  behind the entrance face, respectively. (b)  $s \neq 0$ ,  $n = 2$  ( $st_g = 0.3$ ). (—) bright field; (---) dark field.

the projection of  $\mathbf{b}$  on the active diffraction vector  $\mathbf{g}$ , and hence allows one to determine the length of  $\mathbf{b}$  once its direction is known. For columns passing through the dislocation core the phase shift at the level of the core when  $n = 2$  is now  $2\pi$ , that is such columns will exhibit the same intensities as the perfect crystal.

#### Images of Edge and Mixed Dislocations

The displacement field of a mixed dislocation with a direction defined by its unit vector  $\mathbf{u}$ , parallel to the foil plane (Fig. 64) is given, according to isotropic



**Figure 64.** Reference system used to describe the displacement field due to mixed dislocations.

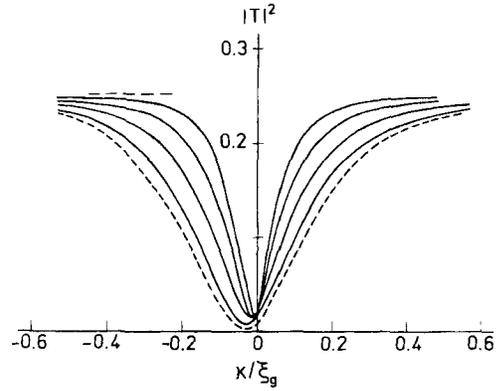
linear elasticity theory, by the expression

$$\mathbf{R} = \frac{1}{2\pi} \left\{ \mathbf{b}\varphi + \frac{\mathbf{b}_e}{4(1-\nu)} \sin(2\varphi) + \left[ \frac{1-2\nu}{2(1-\nu)} \ln|r| + \frac{\cos(2\varphi)}{4(1-\nu)} \right] (\mathbf{b} \times \mathbf{u}) \right\} \quad (169)$$

where  $\nu$  is Poisson's ratio ( $\nu = \frac{1}{3}$ ) and  $\mathbf{b}$  is the Burgers vector, of which  $\mathbf{b}_e$  is the edge component;  $\varphi = \alpha - \gamma$ ;  $r = (x^2 + z^2)^{1/2}$ .

For a pure screw  $\mathbf{b} \times \mathbf{u} = 0$  and  $\mathbf{b}_e = 0$ , and the expression reduces to  $\mathbf{R} = (\mathbf{b}/2\pi)\varphi$ . The term in  $\mathbf{b} \times \mathbf{u}$  describes a displacement perpendicular to the slip plane towards the supplementary half-plane. The slip plane is determined by  $\mathbf{b}$  and  $\mathbf{u}$ ; it forms an angle  $\gamma$  with the foil plane. The character of the dislocation can be quantified by the parameter  $p = (\mathbf{g} \cdot \mathbf{b}_e)/(\mathbf{g} \cdot \mathbf{b})$ , which is 0 for a pure screw and 1 for a pure edge.

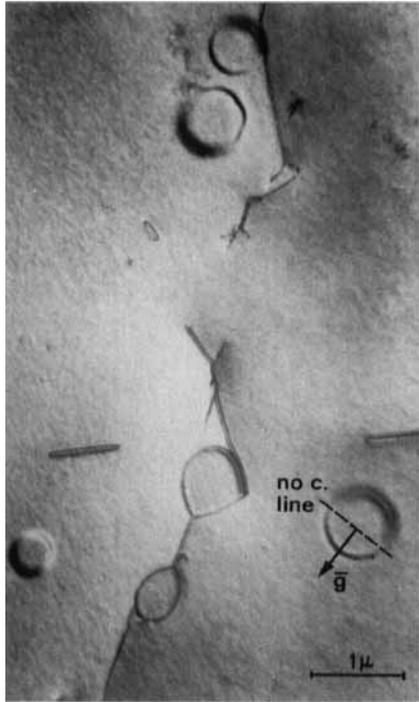
Computed images (Fig. 65) for a mixed dislocation with its slip plane parallel to the foil plane ( $\gamma = 0$ ), for the following values of the parameters  $n = 1$ ,  $s = 0$ ,  $t = 8t_g$ ,  $z_0 = 4t_g$ , and for a number of values for  $p$ , show that the image of a pure edge ( $p = 1$ ) is wider than that of a pure screw ( $p = 0$ ). The narrowest image is obtained for  $p = -\frac{1}{2}$  and the widest for  $p = \frac{1}{2}$ , that is, for  $45^\circ$  mixed dislocations. The full width varies between  $0.3t_g$  and  $0.9t_g$ .



**Figure 65.** Computed images for mixed dislocations with slip plane parallel to the foil plane ( $\gamma = 0$ ) and for various values of the parameter  $p = (\mathbf{g} \cdot \mathbf{b}_e)/(\mathbf{g} \cdot \mathbf{b})$  [13, 14].

Even for  $\mathbf{g} \cdot \mathbf{b} = 0$  a pure edge dislocation may produce contrast because of the term  $\mathbf{b} \times \mathbf{u}$  in the displacement function. For a closed prismatic Frank loop parallel to the foil plane and for the imaging  $\mathbf{g}$  vectors parallel to the loop plane  $\mathbf{g} \cdot \mathbf{b} = 0$ , but nevertheless complete extinction only occurs if moreover  $(\mathbf{b} \times \mathbf{u}) \cdot \mathbf{g} = 0$ , which is only the case if  $\mathbf{u}$  is parallel to the  $\mathbf{g}$  vector. As a result, only those dislocation segments which are parallel to the acting  $\mathbf{g}$  vector will be out of contrast. The line connecting these two segments, called the *line of no contrast*, is perpendicular to the  $\mathbf{g}$  vector (Fig. 66). In a pure Frank loop (i.e., with its Burgers vector perpendicular to the loop plane) a line of no contrast will thus form for all  $\mathbf{g}$  vectors parallel to the loop plane.

If the Burgers vector  $\mathbf{b}$  is inclined with respect to the loop plane there will only be one vector  $\mathbf{g}$  (as well as  $-\mathbf{g}$ ) parallel to the loop plane, for which a line of no contrast occurs; this is the  $\mathbf{g}$  vector perpendicular to the projection of  $\mathbf{b}$  on the loop plane. The argument can be reversed; if among all  $\mathbf{g}$  vectors parallel to the loop plane only one



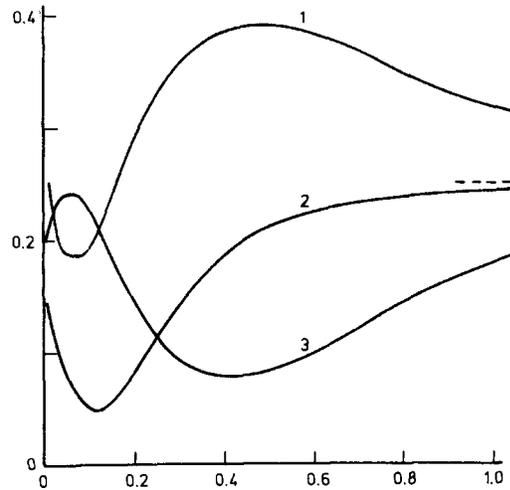
**Figure 66.** Images of prismatic dislocation loops with their Burgers vector parallel to the incident beam in zinc. Note the lines of no contrast perpendicular to the  $g$  vector.

produces a line of no contrast, the loop cannot be a pure Frank loop.

Changing the sign of  $x$  in Eq. (169) for the displacement field changes the sign of  $\varphi$ , but since all terms in Eq. (169) are even functions of either  $\varphi$  or  $x$ , we conclude that the image profile must be symmetrical in  $x$ . In the computed profiles shown in Fig. 67, only the half corresponding to  $x > 0$  is represented; they show that for certain depth positions the image may exhibit two broad dark lines, as for instance in Fig. 66.

#### *Images of Partial Dislocations*

Since the Burgers vectors of partial dislocations are not lattice vectors, the image



**Figure 67.** Computed image profiles (B.F.) for edge dislocations with their Burgers vector parallel to the incident beam. Only half of the profile is shown; it is symmetrical in  $x$  [13, 14].

order  $n \equiv g \cdot b$  may become fractional. For instance, for Shockley partials in face-centered-cubic crystals the Burgers vector is  $\frac{1}{6} [11\bar{2}]$  and the value of  $n$  becomes a multiple of  $\frac{1}{3}$ .

Partial dislocations form the border of stacking faults. The image profile is therefore complicated by the fact that it separates two areas, one of which has the brightness of a perfect region and the other has the contrast of a faulted area at the depth level of the partial dislocation; these brightnesses are, in general, different.

Image profiles have been computed for  $n = \pm\frac{1}{3}$ ,  $\pm\frac{2}{3}$ , and  $\pm\frac{4}{3}$ . For  $n = \pm\frac{1}{3}$  no visible line image is formed since the profile constitutes a continuous transition between the two brightness levels. However, for  $n = \pm\frac{2}{3}$  the image consists of a dark line with a small visibility. Images with  $n = \pm\frac{4}{3}$  are expected to consist of a dark line comparable to that of an ordinary dislocation (Fig. 68). Partial dislocations of the Frank type have pure edge

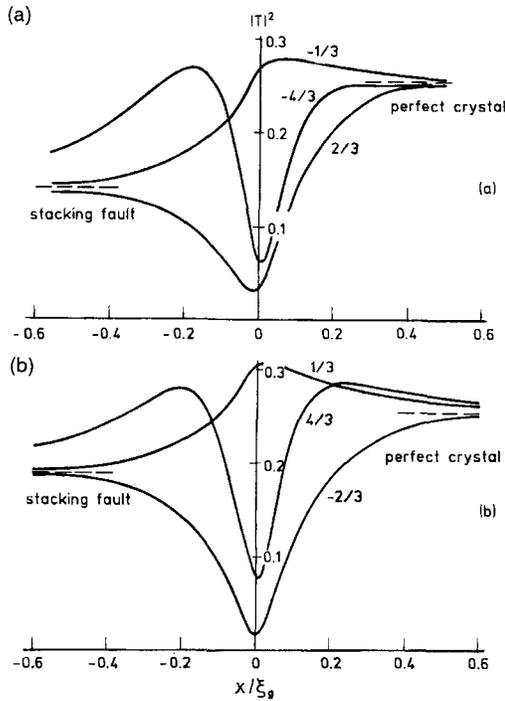


Figure 68. Image profiles (B.F.) for partial dislocations. The  $n$  values are indicated [13, 14].

character; their behavior was discussed in Sec. 1.1.1.18 of this Chapter.

*Images of Dislocation Ribbons*

The images of ribbons of partial dislocations [49] deserve some special consideration. Since the ribbon width is a measure of the energy of the enclosed stacking fault, the exact separation of the partial dislocations needs to be known accurately in order to make possible precise measurements of the stacking fault energy.

The image of a ribbon is *not* the superposition of the images of the separate partial dislocations; it is caused by the strain field of the ribbon, which is obtained in the framework of linear elasticity as the superposition of the strain fields of the two

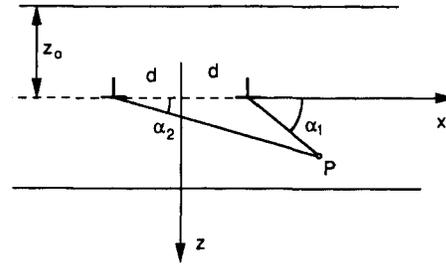


Figure 69. Reference system used in describing the displacement field of a dislocation ribbon [39].

partial dislocations. For a screw ribbon consisting of two Shockley partial dislocations, enclosing an angle of  $60^\circ$ , and for an active diffraction vector oriented along the bisector of the acute angle between Burgers vectors  $\mathbf{b}_1$  and  $\mathbf{b}_2$ , the ribbon behaves to a good approximation, as far as the contrast is concerned, as if it consisted of two screws with  $n$  values ( $n = \mathbf{b} \cdot \mathbf{g}$ ) which are either both  $+1$  or both  $-1$ . This is due to the fact that the edge components are perpendicular to  $\mathbf{g}$  and, therefore, produce residual contrast only.

For the geometry shown in Fig. 69 the phase shift  $\alpha$  caused by the ribbon can be formulated as

$$\alpha = n_1 \alpha_1 + n_2 \alpha_2 \tag{170}$$

with

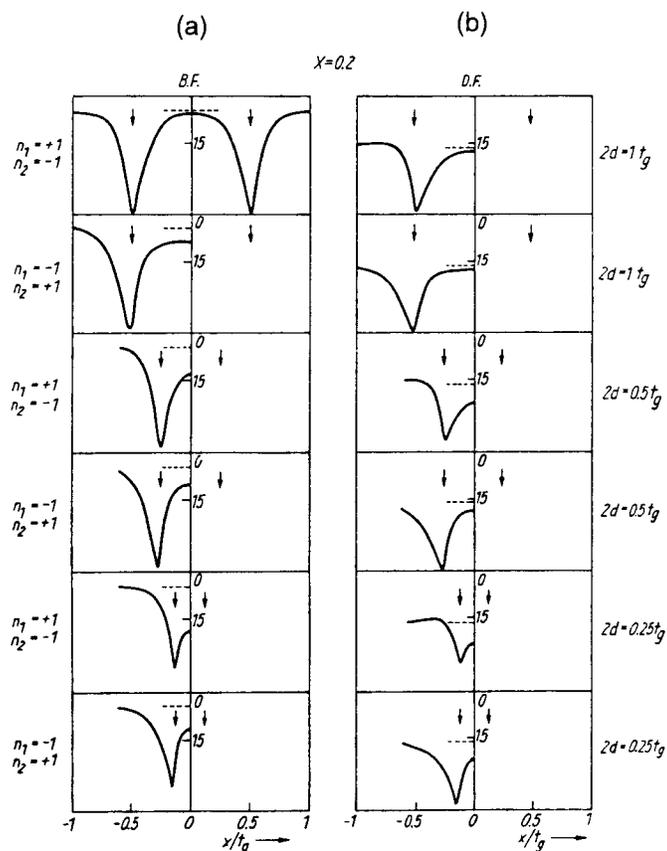
$$\alpha_1 = \text{arctg} \left( \frac{z - z_0}{x - d} \right) \tag{171a}$$

$$\alpha_2 = \text{arctg} \left( \frac{z - z_0}{x + d} \right) \tag{171b}$$

and

$$n_1 = \mathbf{g} \cdot \mathbf{b}_1 \quad \text{and} \quad n_2 = \mathbf{g} \cdot \mathbf{b}_2 \tag{171c}$$

Such a ribbon will produce a symmetrical image when  $n_1 = \pm 1$  and  $n_2 = \mp 1$ , since changing  $x$  into  $-x$  leads to changing  $\alpha_1$



**Figure 70.** (a) Bright field and (b) dark field image profiles for ribbons of varying widths ( $s_g t_g = 0.2$ ) and varying combinations of  $n_1 = \pm 1$ ,  $n_2 = \mp 1$ . The profiles are symmetrical; only one half is shown. BF and DF profiles are represented. The positions of the partial dislocations are indicated by arrows [39].

into  $-\alpha_2$ , and vice versa, and thus  $\alpha(x)$  into  $\alpha(-x)$ . Integrating along columns at  $-x$  and at  $+x$  leads to the same result, and the image is thus symmetrical.

If we change simultaneously  $n_1 = n_2 = +1$  into  $n_1 = n_2 = -1$  and  $x$  into  $-x$  the expression for  $\alpha$  remains unchanged. We conclude from this that the profile for  $n_1 = n_2 = -1$  is the mirror image of that for  $n_1 = n_2 = +1$ . It is thus sufficient to discuss one of these two cases. Bright and dark field image profiles for screw ribbons are reproduced in Fig. 70 for various sign combinations of  $n_1 = \mathbf{g} \cdot \mathbf{b}_1 = \pm 1$  and  $n_2 = \mathbf{g} \cdot \mathbf{b}_2 = \pm 1$ . In all cases the foil had a thickness of  $8t_g$ ; the ribbon was at a

depth  $3.5t_g$  behind the entrance foil and  $st_g = 0.2$ .

When  $n_1$  and  $n_2$  are opposite in sign, as in Fig. 70a, the profiles are symmetrical; only one-half is therefore shown; the complete profile is obtained by a mirror operation. The positions of the dislocations are indicated by arrows. The ribbon widths are clearly different from the peak separations, due to the fact that the image displacements are in the opposite sense for the two dislocations as a result of the sign difference of  $n$ . For the sign combination  $n_1 = +1$  and  $n_2 = -1$  the apparent (observed) width of the ribbon is smaller than the real width; the opposite is true for

$n_1 = -1$  and  $n_2 = +1$ . This width difference is also found to increase with increasing  $s_g$ , since the image displacements increase with  $s_g$ . The effect of a decreasing separation,  $2d$ , on the contrast in the central part of the ribbon is clearly visible; the background intensity being represented by a horizontal dotted line. The center of the image corresponds to a column for which the phase shift is given by

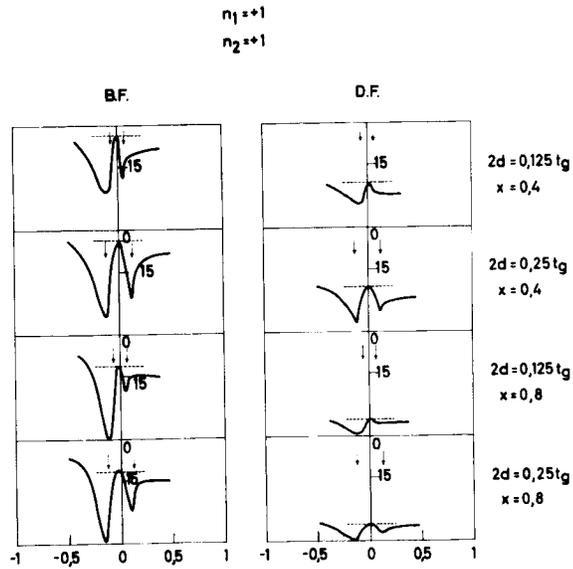
$$\alpha_0 = (n_1 + n_2)\phi - k\pi \quad (172)$$

where  $\phi = \arctg[(z - z_0)/d]$ . The brightness in the center is thus the same as for a stacking fault at  $z = z_0$ , with  $\alpha_0$  as a phase shift.

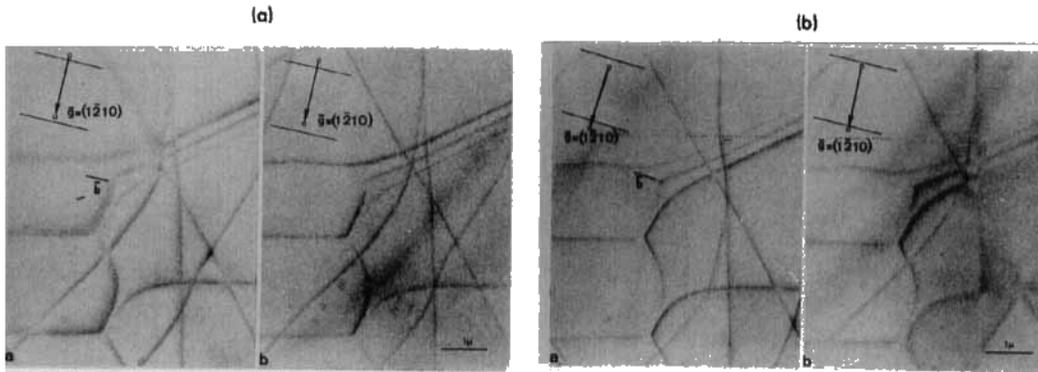
If, on the other hand,  $n_1$  and  $n_2$  have the same sign the image sides are the same for the two partial dislocations and the peak separation in the image is more representative of the real width of the ribbon (Fig. 71b). However, the two partial dislocations are now imaged as lines of different width and brightness for  $s \neq 0$ . This can be understood by noting that the

displacements associated with the two partial dislocations are additive outside of the ribbon, but subtractive in the region between the two partials (i.e., inside the ribbon). The strongest line image is formed outside the ribbon on the image side of the first partial dislocation. For the second partial dislocation the image side is the same as for the first one, but this is now inside the ribbon where the displacements are subtractive and hence the peak is smaller. Changing the sign of  $s$  changes the image side for both partial dislocations. The strongest image will again be outside the ribbon but on the other side, since this is now where the displacements are again additive, albeit in the opposite sense. The brightness in the central column is now the same as the background, since  $\alpha(x = 0) = k\pi$ . Some of these features can be observed in the computed profiles shown in Fig. 71b.

The symmetrical triple ribbons in graphite are formed by three partial dislocations with the same Burgers vector,



**Figure 71.** (a) Bright and (b) dark field image profiles for dislocation ribbons. Foil thickness  $t_g$ ;  $x \equiv st_g$ ;  $z_0 = 3.5t_g$ ;  $n_1 = n_2 = \pm 1$  [39].



**Figure 72.** Image of triple ribbons in graphite under different diffraction conditions (i.e. for different  $s$  values): (a)  $s > 0$ ; (b)  $s < 0$  [39].

separating two fault ribbons (Fig. 72). As a result, the  $n$  values are always the same and hence the image sides for the three partial dislocations, which are determined by the sign of  $ns$ , are also the same. If the displacements are, for instance, additive outside the three-fold ribbon and to the left of it, they are partly additive inside the left fault ribbon, partly subtractive inside the right fault ribbon, and completely subtractive (i.e., additive in the opposite sense) outside the triple ribbon and on the right of it. As a result the peaks marking the three partial dislocations will decrease in magnitude from left to right if the image side is on the left. Changing the sign of  $s$  or changing the image side for the three partial dislocations will invert the sense in which the magnitudes of the peaks decrease. This feature can be observed in the sequence of images shown in Fig. 72.

#### *Dislocation Dipoles*

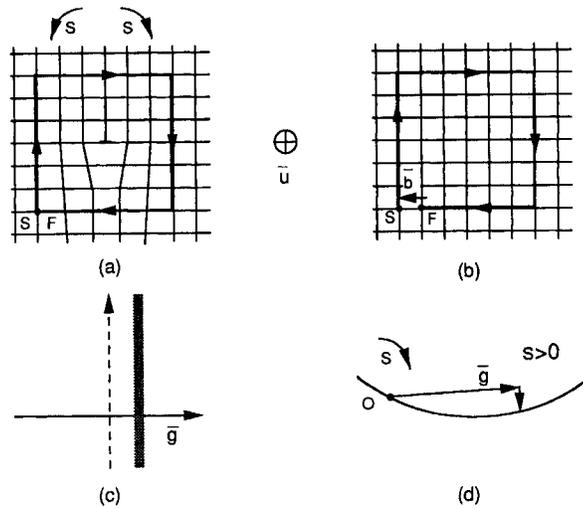
Dislocation dipoles [48] consist of two parallel dislocations with opposite Burgers vector. If the dislocations are restrained to remain in their respective glide planes they take up a stable configuration which

minimizes the elastic energy. For two edge dislocations the regions of expansion in one dislocation tend to overlap the regions of compression in the other one. The configuration is such that the plane formed by the two parallel dislocations encloses an angle  $\phi$  with their slip planes. In the case of two pure edges  $\phi = 45^\circ$ .

As in the case of dislocation ribbons, the image of a dipole is not the superposition of the images of the two separate dislocations. The superposition must be carried out at the level of the strain fields in the framework of the linear elasticity theory. It has been shown by Head et al. [48] that the bright field image of an inclined dipole has a center of symmetry. This symmetry property allows one to distinguish between an image of a narrow dipole and an image of a single dislocation.

#### *The Image Side of Dislocations*

From the dynamical image simulations discussed in Sec. 1.1.1.18 of this Chapter, we concluded that the black line image of a dislocation in the bright field image is systematically one-sided provided that  $s$  is large enough. This is true for  $\mathbf{g} \cdot \mathbf{b} = 1$



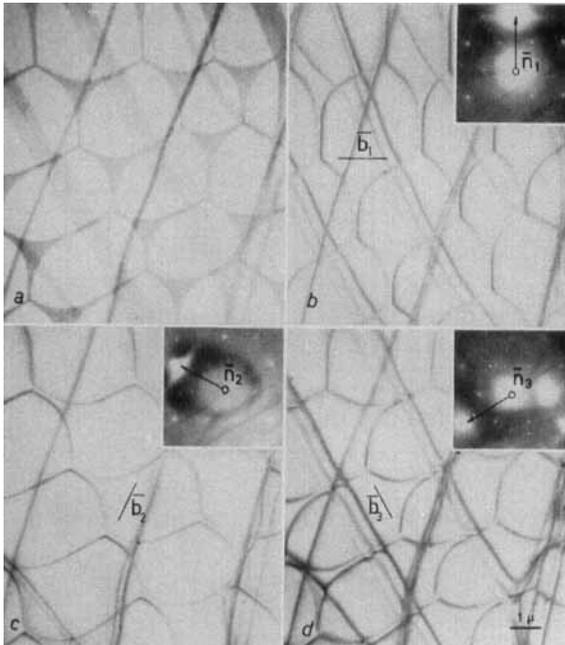
**Figure 73.** Illustration of the FS/RH convention for defining the sense of the Burgers vector of a dislocation: (a) real crystal; (b) perfect reference crystal; (c) relative position of image and dislocation; (d) diffraction conditions.

as well as for  $\mathbf{g} \cdot \mathbf{b} = 2$ . This behavior is different for  $n = 1$  and  $n = 2$  if  $s \approx 0$ , as discussed in Sec. 1.1.1.18 of this Chapter.

The image side (i.e., the position of the black line in the bright field image on a positive print with respect to the dislocation position) is correctly given by the intuitive kinematical considerations discussed in Sec. 1.1.1.17 of this Chapter. According to this theory the image side is on that side of the dislocation core where locally the lattice planes normal to  $\mathbf{g}$  are rotated towards the exact Bragg orientation; in Fig. 73 this is in the sense  $S$ . Finding  $S$  requires a knowledge of the sign of  $s$ ; this can be determined by means of the Kikuchi line pattern, as shown in Sec. 1.1.1.3 of this Chapter. For the edge dislocation shown in Fig. 73 the positive sense (i.e., the unit vector  $\mathbf{u}$ ) was chosen as entering the plane of the drawing. The Burgers vector  $\mathbf{b}$  is determined according to the FS/RH convention as follows. A right-handed closed Burgers circuit when looking along  $\mathbf{u}$  is constructed in the real crystal. In the perfect reference

crystal the corresponding circuit is constructed and the Burgers vector  $\mathbf{b}$  is found as the closure failure of this circuit, joining the final point to the starting point  $\mathbf{b} = \text{FS}$ . For the concrete situation shown in Fig. 73,  $s > 0$  and  $(\mathbf{g} \cdot \mathbf{b})s < 0$ ; the image is indicated by a solid line and the dislocation line by a dotted line. The rule can be formulated as follows: the image side is to the right looking along the positive sense if  $(\mathbf{g} \cdot \mathbf{b})s < 0$ . Changing the sign of one of the three parameters  $\mathbf{g}$ ,  $\mathbf{b}$ , or  $s$  changes the image side.

It should be noted that the descriptions given in different reviews are sometimes confusing and do not always agree because some authors refer to the image as seen along the incident electron beam, whereas other formulations refer to the image as seen from below. The sense of  $\mathbf{u}$  depends on whether the first or the second viewpoint is adopted, but this changes the sign of  $p = (\mathbf{g} \cdot \mathbf{b})s$ . The most direct way is to apply intuitive reasoning, correctly taking into account possible electron optical image rotations.



**Figure 74.** Dislocation network in graphite imaged under four different diffraction conditions leading to stacking fault contrast in (a) and to extinctions in (b)–(d). Note that the triple ribbon loses contrast completely in (c) [56].

### Characterizing Dislocations

A full description of a dislocation line requires a determination of its core geometry and its Burgers vector, that is of the direction, magnitude and sense of  $\mathbf{b}$  [13, 40]. Methods are available to obtain all these elements. The precise *position* of the dislocation can be found by making two images leading to opposite image sides either for active diffraction vectors  $+\mathbf{g}$  and  $-\mathbf{g}$  for the same sign of  $s$ , or for  $+\mathbf{s}$  and  $-\mathbf{s}$  for the same  $\mathbf{g}$  vector. The true position of the dislocation is then between the two images.

The *direction* of the Burgers vector is determined by looking for two diffraction vectors  $\mathbf{g}_1$  and  $\mathbf{g}_2$  for which the dislocation is out of contrast or for which a residual contrast characteristic of  $\mathbf{g} \cdot \mathbf{b} = 0$  is produced. The Burgers vector then has a direction parallel to  $\mathbf{g}_1 \times \mathbf{g}_2$ . An example

of the application of this method to a hexagonal network of dislocations in graphite is shown in Fig. 74. In this particular case a single ‘extinction’ is in fact sufficient since the dislocations were known to be glide dislocations and thus have their Burgers vector in the  $c$  plane. As the foil is prepared by cleavage it is also limited by  $c$  planes. The three families of partial dislocations are seen to be successively brought to extinction using the  $\mathbf{g}$  vectors indicated. Note also the simultaneous extinction of the three partial dislocations in the triple ribbon, showing that they have the same Burgers vector. Their contrast is nevertheless different for the different partials, as discussed in Sec. 1.1.1.18 of this Chapter.

It should be remembered that in highly anisotropic materials the simple extinction criterion  $\mathbf{g} \cdot \mathbf{b} = 0$  is no longer strictly valid, as discussed in Sec. 1.1.1.18 of this

Chapter. In the case of graphite, just mentioned, the extinctions can be observed unambiguously, even though graphite is highly anisotropic. However, due to the presence of the six-fold symmetry axis along  $c$ , the  $c$  plane behaves effectively as elastically isotropic. This is also the case for dislocations in the (111) planes of face-centered-cubic crystals, due to the three-fold symmetry.

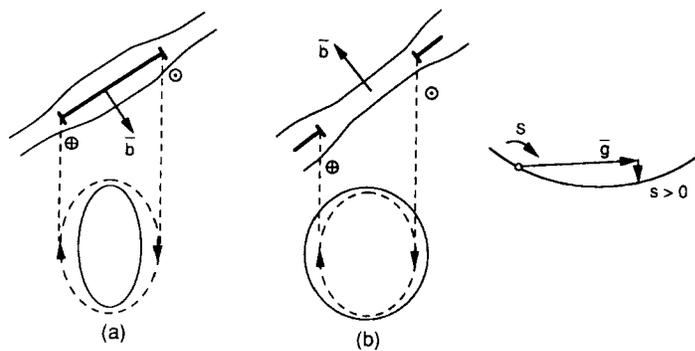
If complete extinction cannot be achieved, one should look for the weakest contrast conditions and deduce from these a plausible Burgers vector, taking the crystal structure into account. Image simulations for various  $\mathbf{g}$  vectors, based on this Burgers vector, can then be compared with the observed images.

The *magnitude* of the Burgers vector for a perfect dislocation can be determined once its direction is known, by looking for diffraction vectors for which  $\mathbf{g} \cdot \mathbf{b} = 2$ . Use is made of the typical contrast effect that occurs where the dislocation crosses a bent extinction contour (see Sec. 1.1.1.18 of this Chapter). If such a diffraction vector is identified we know the length of the projection of  $\mathbf{b}$  on  $\mathbf{g}$ . With a knowledge of the direction of  $\mathbf{b}$  and of its projected length on  $\mathbf{g}$ , the length of  $\mathbf{b}$  can be found.

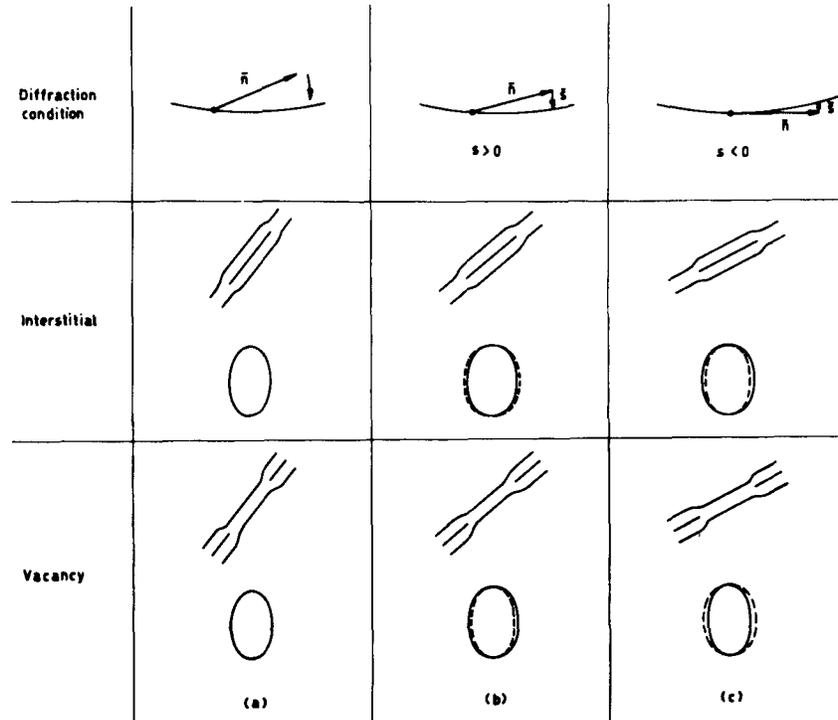
Finally, the sense of  $\mathbf{b}$  is found from the image side which defines the sign of  $(\mathbf{g} \cdot \mathbf{b})_s$ . Knowing the sign of  $s$  from the Kikuchi pattern, the image side allows one to find the sign of  $\mathbf{g} \cdot \mathbf{b}$ . Knowledge of  $\mathbf{g}$  then leads to the sense of  $\mathbf{b}$ .

An important application of the sign determination of the Burgers vector consists in determining whether a Frank loop is due to the precipitation of vacancy or interstitial loops, that is whether  $\mathbf{b}$  is either  $+\frac{1}{3}[111]$  or  $-\frac{1}{3}[111]$ . Applying the relation determining the image side to the loop represented in Fig. 75 it follows that for a loop the image is either completely inside or completely outside the dislocation ring, depending on the sign of  $(\mathbf{g} \cdot \mathbf{b})_s$ , and since  $\mathbf{b}$  is different for a vacancy loop and an interstitial loop, so will be the image side for the same  $\mathbf{g}$  and  $s$ . The type of contrast experiment required for an analysis of the nature of loops is illustrated in Fig. 76. A difficulty arises because of the need to know the sense of inclination of the loop plane. If the loops are known to be parallel to the foil surface a known slope can be imposed by mounting the sample on a wedge. However, this method is not always possible.

Assuming the sense of the slope to be as represented in Fig. 75 and  $\mathbf{g}$  and  $s$  to be as



**Figure 75.** Image contrast of dislocation loops: (a) image inside the loop; (b) image outside the loop.



**Figure 76.** Contrast experiment for determining the nature of dislocation loops. Diffraction vector  $g \equiv n$ . The foil is tilted in the sense indicated.

there shown, it is evident that the image is inside for interstitial loops, whereas it is outside for vacancy loops. Changing the sign of  $s$  by tilting allows one to find the image side and hence to distinguish between the two cases.

An alternative application of the same principle consists in rotating the specimen through the exact Bragg orientation from  $s > 0$  to  $s < 0$  for a given  $g$  (Fig. 76). It is then found that an *interstitial* loop will grow in size because of two effects: (i) the projected size increases and (ii) the image goes from inside for  $s > 0$  to outside for  $s < 0$ . A *vacancy* loop will grow as long as  $s > 0$  because of the geometrical effect; but beyond  $s = 0$  the image side changes and the image size shrinks. The experiment

must clearly be performed starting with loops which are steeply inclined. One can also make use of the asymmetrical image contrast, consisting of a line of no contrast; separating a bright and a dark lobe (or crescent), characteristic of Frank loops seen end-on, which moreover are close to the surface. In the dark field image the asymmetry is the same at the top and bottom of the foil, due to anomalous absorption. If the diffraction vector  $g$  is parallel to  $b$  and points from the bright to the dark lobe in the image, the loop has interstitial character. If  $g$  points from the dark to the bright lobe, the loop is a vacancy loop. A restriction is that the loop must be close to the surface (i.e., within  $\frac{1}{2}t_g$ ). To demonstrate that the latter

condition is satisfied, stereo images are required.

### 1.1.1.19 Moiré Patterns [50–52]

#### Intuitive Considerations

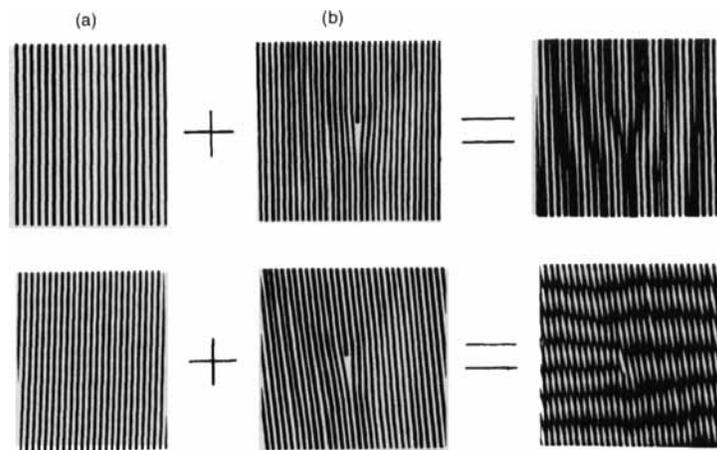
Electron microscopic specimens consisting of two similar thin films superimposed with a small orientation difference produce interference patterns consisting of parallel fringes when a  $g$  vector parallel to the film plane is excited in the two components of the sandwich. In the bright field image this fringe pattern results from the interference between the doubly transmitted and the doubly scattered beams, which enclose a small angle. This angle is usually revealed in the diffraction pattern by a doubling of the spots. A ‘doubly transmitted beam’ is a beam which is transmitted through the first thin film and subsequently through the second; the doubly scattered beam is formed in a similar way.

In a number of cases the geometrical features of such fringes provide useful

information. A geometrical analog, consisting of the superposition of two line patterns (the lines representing lattice planes), one of them containing a dislocation, is shown in Fig. 77. In (a) the directions of the lines are the same, but their spacing is slightly different. In (b) the two patterns have the same line spacing, but the directions of the lines enclose a small angle. The moiré pattern or the superposition pattern shows a magnified representation of a dislocation [51]. Moiré patterns can thus provide ‘geometrical’ magnification, which was especially useful at a time when atomic resolution was not possible. With the development of atomic resolution microscopy, moiré imaging lost most of its importance; however, the geometrical features are still useful [52].

#### Theoretical Considerations

Consider a composite crystal consisting of two plan parallel slabs I and II [50]. Let part II of a column be derived from part I by the displacement field  $u(r)$ . The phase shift between the waves diffracted locally



**Figure 77.** Geometrical analog illustrating the formation of moiré patterns: (a) parallel moiré; (b) rotation moiré. One of the two superimposed foils contains a dislocation.

by the two parts of the crystal is then  $\alpha = 2\pi(\mathbf{g} + \mathbf{s}) \cdot \mathbf{u} \simeq 2\pi(\mathbf{g} \cdot \mathbf{u})$ . This expression is of the same form as the phase shift introduced by a stacking fault, the main difference being that  $\mathbf{u}$  is *not* a constant vector  $\mathbf{R}$  but now depends on  $\mathbf{r}$ , and hence  $\alpha$  is also a function of  $\mathbf{r}$ . The transmitted and scattered amplitudes are then given by Eqs. (113a) and (113b), in which  $\alpha$  enters through the periodic factor  $\exp(i\alpha)$ . Without solving the system of equations it is clear that the loci of the points of equal intensity (i.e., the fringes) are given by  $\exp(i\alpha) = \text{Constant}$ , that is by  $\alpha = \text{Constant} + k2\pi$  (where  $k$  is an integer).

Assuming  $\mathbf{r}$  to be a lattice vector  $\mathbf{g} \cdot \mathbf{r} = \text{Integer}$ , for small difference vectors

$$\Delta \mathbf{g} \cdot \mathbf{r} + \mathbf{g} \cdot \Delta \mathbf{r} = 0 \quad (173)$$

with  $\Delta \mathbf{r} = \mathbf{u}(\mathbf{r})$

and thus

$$\alpha \equiv 2\pi \mathbf{g} \cdot \mathbf{u} = -2\pi \Delta \mathbf{g} \cdot \mathbf{r} \quad (174)$$

Provided  $\mathbf{u}(\mathbf{r})$  is such that  $\Delta \mathbf{g}$  does not depend on  $\mathbf{r}$ , which is true for moiré patterns, the lines of equal intensity are given by  $\Delta \mathbf{g} \cdot \mathbf{r} = \text{Constant} + k$ , where  $k$  is an integer. This equation represents a set of parallel straight lines perpendicular to  $\mathbf{K} = -\Delta \mathbf{g}$ , where  $\mathbf{K}$  can be considered as the wavevector of the moiré fringe system, with wavelength  $\Lambda = 1/K$ .

In the case of a *rotation moiré*,  $K = 2g \sin(\theta/2) \simeq g\theta$  (for small  $\theta$ ); or, expressed in terms of the interplanar spacing  $d_g$  of the active reflection

$$\Lambda_r = d_g / \theta \quad (175)$$

The fringes are parallel to  $\mathbf{g}$  for small  $\theta$ .

For *parallel moiré patterns*,  $\Delta \mathbf{g} = \mathbf{g}_2 - \mathbf{g}_1$  with  $\mathbf{g}_2$  parallel to  $\mathbf{g}_1$  or, in terms

of interplanar spacings,

$$\Delta g = \frac{1}{d_2} - \frac{1}{d_1} = \frac{d_1 - d_2}{d_1 d_2} \quad (176)$$

and

$$\Lambda_p = \frac{d_1 d_2}{d_1 - d_2} \quad (177)$$

The fringes are again perpendicular to  $\Delta \mathbf{g}$ , that is they are also perpendicular to  $\mathbf{g}_1$  and  $\mathbf{g}_2$ .

If an orientation difference as well as a spacing difference is present, mixed moiré patterns are formed. One can always decompose  $\Delta \mathbf{g}$  into components perpendicular and parallel to  $\mathbf{g}$ :

$$\Delta \mathbf{g} = \Delta \mathbf{g}_{\parallel} + \Delta \mathbf{g}_{\perp} \quad (178)$$

As the fringes are still perpendicular to  $\Delta \mathbf{g}$ , they enclose an angle  $\beta$  with the direction of  $\mathbf{g}$  given by  $\tan \beta = \Delta g_{\perp} / \Delta g_{\parallel}$  and

$$|\Delta \mathbf{g}|^2 = |\Delta \mathbf{g}_{\perp}|^2 + |\Delta \mathbf{g}_{\parallel}|^2 \quad (179)$$

and hence

$$1/\Lambda^2 = 1/\Lambda_{\parallel}^2 + 1/\Lambda_{\perp}^2 \quad (180)$$

The intensity variation of the fringe system can be found in a similar way as for stacking faults. For a quantitative theory of the intensity profiles, see [50].

From the theory one can conclude that the positions of the moiré fringes depend on the total thickness of the sandwich, and hence the fringe positions are influenced by surface steps. Furthermore, the fringe positions depend on the specimen tilt.

We have seen that for coherent domain boundaries  $\Delta \mathbf{g}$  is perpendicular to the interface. When such an interface is perpendicular to the incident beam, which is the usual geometry for moiré patterns, the projection of  $\Delta \mathbf{g}$  onto the interface thus vanishes and no moiré fringes are formed.

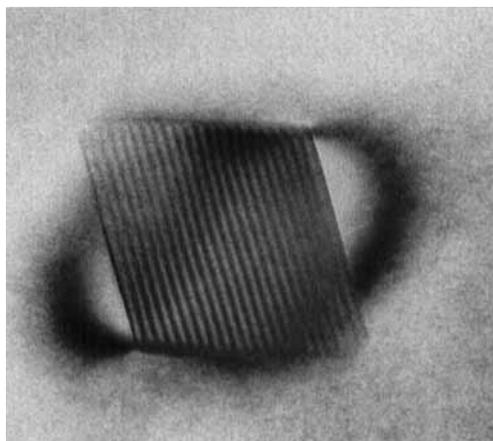
Fringe patterns imaging this type of interface therefore have a different origin. The image for an inclined domain boundary consists of the  $\delta$  fringes described above, which are perpendicular to the projection of  $\Delta g$  onto the foil plane (see Sec. 1.1.1.13 of this Chapter), that is onto the intersection lines of the interface with the foil surfaces.

If  $\Delta g$  has an arbitrary orientation with respect to the contact plane between the two crystal parts, it has a perpendicular component as well as a parallel component with respect to the interface and the image can be a complicated mixture of both types of image. The parallel component gives rise to moiré type fringes, and the perpendicular component to  $\delta$ -type fringes.

#### Applications of Moiré Patterns

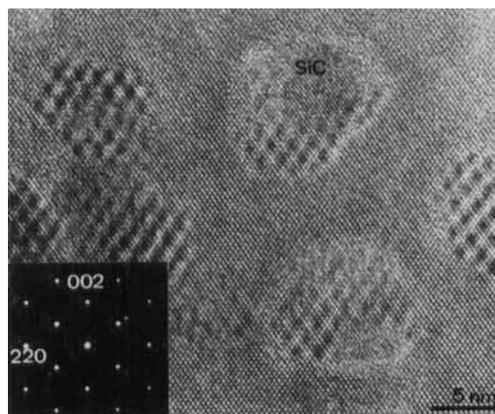
An important application of parallel moiré fringes is the determination of the lattice parameter of one of the two components in a sandwich, the lattice parameter of the other being known. This can be of interest for the identification of plate-like coherent precipitates in a matrix with a known lattice parameter. Moiré fringes formed at the interface between voidite and diamond are shown in Fig. 78, and Fig. 79 shows the moiré fringes at the interface between silicon and silicon carbide (SiC) precipitate particles.

Moiré fringes have also been used as a tool in the study of dislocations. Ending moiré fringes reveal the emergence points of dislocations in one of the two components of the sandwich. The number,  $N$ , of supplementary half-fringes depends on the reflection used to produce the dislocation



**Figure 78.** Moiré pattern formed at the interface between voidite and the diamond matrix. Note the extinction contours revealing the strain field. (Courtesy of G. Van Tendeloo.)

image; it is given by  $N = g \cdot b$ . This number is independent of the character of the dislocation. Supplementary half-fringes cannot therefore be interpreted as necessarily meaning that the corresponding dislocation has edge character. Partial dislocations bordering stacking faults are revealed by a 'fractional' number of



**Figure 79.** Moiré pattern at the interface between silicon and a silicon carbide precipitate. (Courtesy of A. De Veirman.)

supplementary half-fringes; that is, along the trace of the stacking fault the moiré fringes are shifted discontinuously over a fraction  $\mathbf{g} \cdot \mathbf{b}$  such as  $\frac{1}{3}$  or  $\frac{2}{3}$  of the interfringe distance.

The moiré fringes are also shifted by a surface step in one of the components. The fringe shift is not only a function of the step height but also of the deviation parameter and hence of the specimen orientation.

If two or more diffraction vectors are active in both components of the sandwich a crossed grid of moiré fringes is formed, which has the rotation symmetry of the two films.

### 1.1.1.20 Applications of Diffraction Contrast

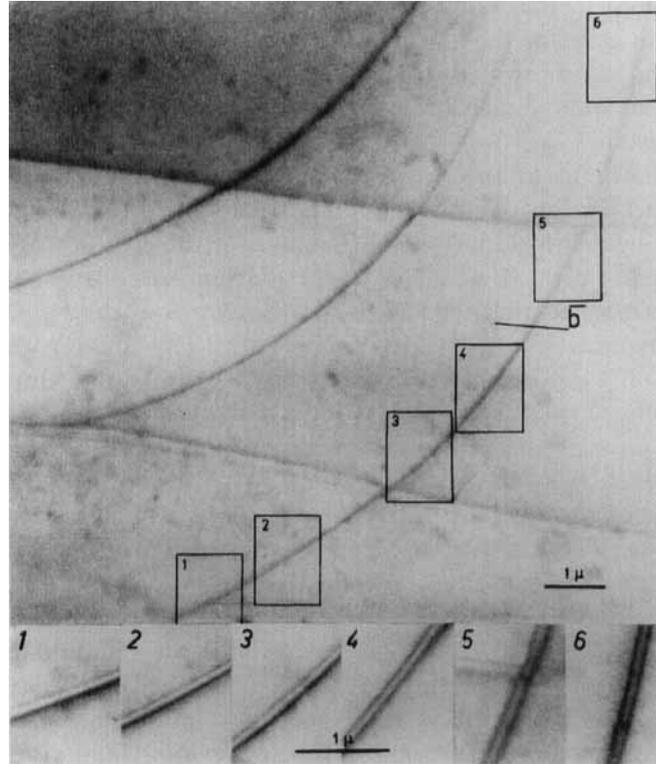
Diffraction contrast images do not reveal the crystal structure but they do image defects such as dislocations, planar interfaces (stacking faults, coherent domain boundaries, and out-of-phase boundaries), discommensurations, point defect clusters. Applications of diffraction contrast to specific classes of materials are described in Vol. II of this handbook. Here we discuss only a few characteristic examples illustrating different types of defect study. The interpretation of the images is only meaningful in terms of the underlying materials problem. Therefore, for each application we sketch as briefly as possible the framework in which the images acquire their significance and interpretation. The choice of the examples, which is admittedly subjective, is mainly motivated by the availability of suitable photographs.

## The Fine Structure of Dislocations

### *Measuring the Stacking Fault Energy*

In most materials the dislocations are not simple line defects but consist in fact of two or more partial dislocations connected by strips of stacking fault or of out-of-phase boundary. The simplest situation arises when glide takes place between two close-packed layers of identical 'spherical' atoms in an elemental face-centered-cubic crystal. The glide motion along the (111) plane in the  $[\bar{1}10]$  direction then follows the valleys, that is it takes place in two steps, each performed by the motion of a partial dislocation, the first with a Burgers vector  $\mathbf{b}_1 = \frac{1}{6} [\bar{2}11]$  and the second with a burgers vector  $\mathbf{b}_2 = \frac{1}{6} [\bar{1}2\bar{1}]$ , enclosing an angle of  $60^\circ$  and leading to a symmetry translation  $\frac{1}{2} [\bar{1}10]$  along the (111) glide plane. Between the two partial dislocations a stacking fault ribbon with a displacement vector, equal to one of the Burgers vectors of the partial dislocations, is present [53].

The two partial dislocations repel one another, since their Burgers vectors enclose an acute angle. In an infinite solid this repulsion is proportional with  $1/d$  (where  $d$  is the partial dislocation separation) and its magnitude is a function of the orientation of the partial dislocations that is a function of their character (screw or edge). The presence of the stacking fault ribbon causes an effective attractive force per unit length between the two dislocations, which is independent of their separation and numerically equal to the stacking fault energy,  $\gamma$ . An equilibrium separation is thus established. Assuming the repulsive force law to be known, it is then possible to deduce the stacking fault energy from the measured equilibrium separation of the



**Figure 80.** Curved dislocation ribbon in the (0001) plane of graphite. Several segments are reproduced as magnified insets. The direction  $b$  of the total Burgers vector, as determined by extinction experiments, is indicated. Note the systematic change in width with orientation [56].

partial dislocations. Dislocation ribbons are thus sensitive probes for measuring the stacking fault energy, a quantity which is difficult to access in any other direct way. The following relations apply in an infinite isotropic solid:

$$d = d_0 \left[ 1 - \left( \frac{2\nu}{2 - \nu} \right) \cos(2\phi) \right] \quad (181)$$

with

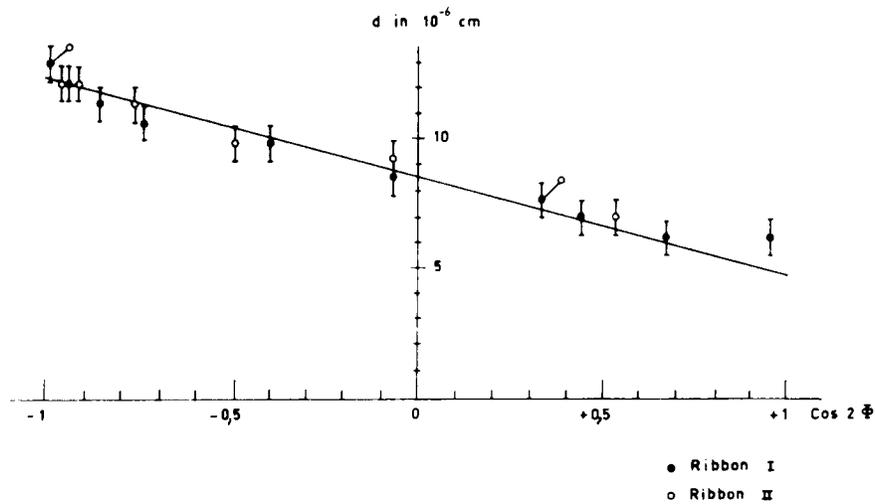
$$d_0 = \frac{\mu b^2 (2 - \nu)}{8\pi\gamma(1 - \nu)} \quad (182)$$

where  $\phi$  is the angle between the total Burgers vector and the ribbon direction;  $\mu$  is the shear modulus and  $\nu$  is Poisson's ratio.

The orientation dependence of the ribbon width can be verified on an image such

as the one shown in Fig. 80, which represents a curved dislocation in a graphite foil. The Burgers vectors were determined using the method described in Sec. 1.1.1.18 of this Chapter. Plotting  $d$  as a function of  $\cos(2\phi)$  the slope of the straight line so obtained gives the effective value of the Poisson ratio as well as the intercept  $d_0$  to be used in the second relation, which then yields a value for the stacking fault energy (Fig. 81).

Using this method it is implicitly assumed that the repulsive force between dislocations is proportional to  $1/d$ , which is only the case in an infinite solid. In a thin foil the repulsive force between dislocations parallel to the foil surfaces decreases with decreasing distance to the specimen surfaces. This behavior can be observed, as

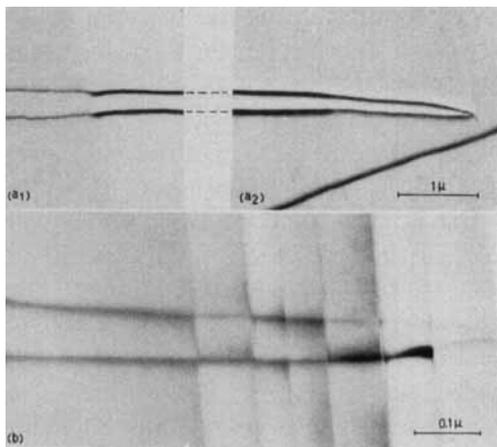


**Figure 81.** Plot of ribbon width  $d$  as a function of  $\cos 2\phi$  [56].

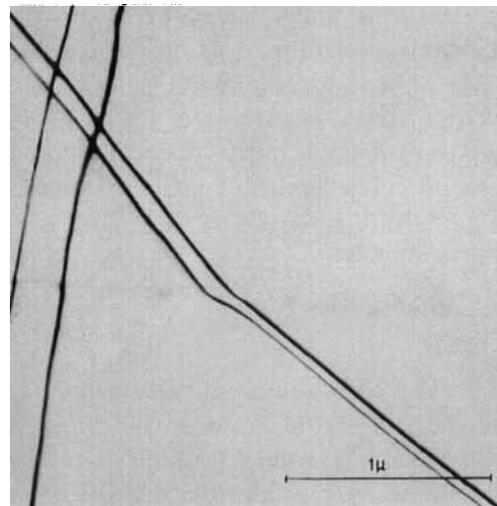
shown in Fig. 82, where a ribbon gradually approaching the surface in a wedge-shaped lamella of tin disulphide, closes as it emerges at the surface [53, 54].

The energy of a dislocation ribbon depends on its distance to the surface. As a result, the shape of minimum energy of a

dislocation ribbon crossing a surface step is not a straight line; 'refraction' of the ribbon as well as a change in width occur on passing underneath the surface step (Fig. 83). The *index of refraction* is the ratio of the total energies of the ribbon in

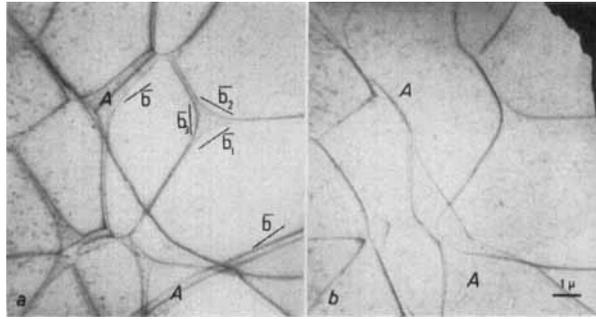


**Figure 82.** A wide ribbon in  $\text{SnS}_2$  gradually approaching the surface. As the ribbon crosses surface steps it becomes discontinuously narrower. The ribbon closes where it emerges in the surface [54].



**Figure 83.** Dislocation ribbon in  $\text{SnS}_2$ . Refraction, accompanied by a width change, occurs on passing underneath a surface step [54].

**Figure 84.** (a) Widely extended dislocation node of partial dislocations in graphite. At A a triple ribbon is present; the three partials have the same Burgers vector as follows from the contrast experiment in (b) where the three partials are simultaneously out of contrast. Nevertheless, the contrast at the three partial dislocations is different in (a) [56].



the two parts of the foil on either side of the surface step.

These images prove that such surface effects are not negligible. Therefore, when measuring stacking fault energies care should be taken to use foils of maximum thickness compatible with the visibility of the dislocations and, moreover, take the widest ribbon as the most representative one. The width of narrow stacking fault ribbons can best be determined by imaging in the weak-beam mode (see Sec. 1.1.1.17 of this Chapter).

Other geometrical configurations involving stacking faults can be used, such as the separation of partial dislocations in triple ribbons in graphite (Fig. 84a) and in close-packed structures, or the radius of curvature of single partial dislocations in a network of extended–contracted nodes (Fig. 84b). In the latter case one has, approximately,

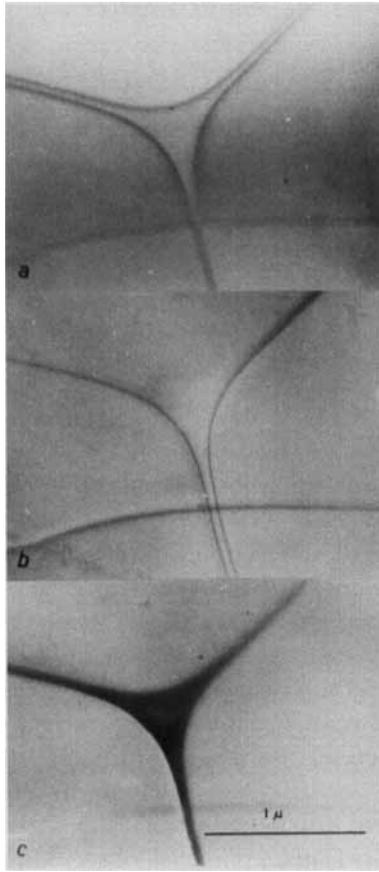
$$\gamma = \frac{\mu b^2}{2R} \quad (183)$$

where  $R$  is the radius of curvature and  $b$  is the Burgers vector of the partial dislocation. Isolated extended nodes such as the one shown in Fig. 85 (observed in AlN) are particularly suitable. More accurate relations are discussed by Nabarro [53] and Siems et al. [55].

### *Multiribbons*

Ordering in alloys based on close-packed structures leads to long symmetry translations along the glide directions in the close-packed glide planes. This results in ribbons consisting of several partial dislocations, separated either by stacking faults or by out-of-phase boundaries. The equilibrium separation of superdislocations (i.e., perfect dislocations with respect to the basic lattice, but partial dislocations with respect to the ordered structure) can be used to derive values of the antiphase boundary energy in the same way as described above for stacking faults. In  $\text{Ni}_4\text{Mo}$ , as many as ten partial dislocations are connected by faults and antiphase boundaries [53].

The dislocations involved in glide between the close-packed layers of anions (X) in layered ionic sandwich crystals of  $\text{CdI}_2$  of the type  $\text{AX}_2$  (XAXXAX...) or  $\text{AX}_3$  are of particular interest. The glide motion takes place between the two weakly van der Waals bonded close-packed anion layers. Dislocations can thus dissociate into two or more Shockley partial dislocations. Although in the close-packed layers, between which glide takes place, all X atoms seem equivalent, the A cations in the adjacent central layers of the



**Figure 85.** Isolated extended node in the (0001) plane of AlN. (Courtesy of P. Delavignette.)

sandwiches may form configurations which impose a large unit mesh in the glide plane, either because not all octahedral cations sites are occupied (e.g., in  $\text{CrCl}_3$  and  $\text{CrBr}_3$ ) or because the cations form metal-metal bonded clusters leading to a slight deformation of the close-packed layers (e.g., in  $\text{NbTe}_2$  and  $\text{TaTe}_2$ ) and the creation of a superperiod.

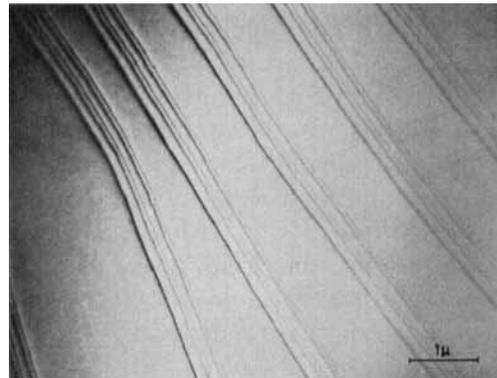
In the chromium trihalides [53, 56] multiribbons containing either four or six partial dislocations are observed. Assuming glide along the close-packed anion layers to take place by the propagation of

Shockley partial dislocations, two types of stacking fault ribbons can be distinguished:

- (i) faults violating only the chromium stacking (i.e., involving only third neighbors), and
- (ii) faults violating the stacking of the chromium ions as well as that of the anions (i.e., involving next-nearest neighbors).

Intuitively, it is clear that the type (ii) faults will have a larger energy than those of type (i). The six-fold ribbons correspond to a 'straight' zigzag glide path along the close-packed directions in the (0001) glide plane of the anion sublattice; they contain the two types of fault in an alternating fashion, the outer ribbons corresponding to high energy faults. Diffraction contrast images of such ribbons are shown in Fig. 86. The outer ribbons are clearly the narrowest ones. The structure of the four-fold ribbons can similarly be related to the structure.

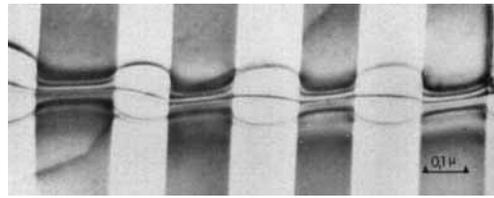
Also in  $\text{NbTe}_2$  [57, 58], which has a deformed  $\text{CdI}_2$  structure, six-fold ribbons



**Figure 86.** Six-fold ribbons of partial dislocations in the (0001) plane of  $\text{CrCl}_3$ . The Burgers vectors of the partial dislocations form a zigzag glide path [53, 56].

occur. In this structure the niobium ions form clusters of three parallel close-packed rows, having a somewhat smaller separation than in the ideal hexagonal structure which probably occurs only in the temperature range in which the crystal is grown. The resulting structure then becomes monoclinic on cooling. The unit mesh in the glide plane is a centered rectangle, which can adopt three different but equally probable orientations differing by  $60^\circ$ . As a consequence, the room temperature structure is fragmented into domains corresponding to the three possible orientations of the clustered niobium rows. The monoclinic symmetry causes the glide paths along the three close-packed directions within the same domain to become non-equivalent.

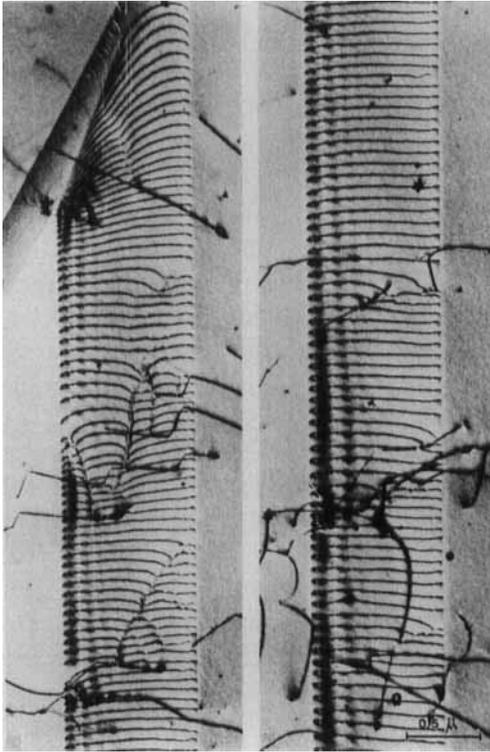
The zigzag glide paths in the direction enclosing an angle of  $30^\circ$  with the long side of the rectangular mesh consist of six partial dislocations, whereas the glide path along the other close-packed direction (i.e., along the short side of the rectangle) repeats after two partial dislocations. The Burgers vector is conserved all the way along the dislocation lines. Hence when a six-fold ribbon passes through a domain wall the glide path changes its orientation relative to the underlying structure. A six-fold ribbon in one domain is thus transformed into three separate two-fold ribbons in the adjacent domain. Whereas in the six-fold ribbon the six partial dislocations are held together by stacking faults, this is no longer the case with the three two-fold ribbons which repel one another and hence develop 'bulges'. The image shown in Fig. 87 illustrates the behavior of a six-fold ribbon intersecting a set of parallel domain boundaries in  $\text{NbTe}_2$  [58].



**Figure 87.** Six-fold ribbon of partial dislocations in  $\text{NbTe}_2$  intersecting domain boundaries along which the underlying structure changes by  $60^\circ$  in orientation. In half of the domains the six-fold ribbons separate into three two-fold ribbons which form bulges as a result of repulsive forces [58].

#### *Plastic Deformation: Glide Dislocations*

Plastic deformation is a subject which was studied intensely in its early stages by means of diffraction contrast. High voltage (about 1000 kV) electron microscopy has been of considerable interest in this respect because it can be used to study thicker foils, which are more representative of a bulk material than the thin foils required at 100 kV. Figure 88 shows a procession of glide dislocations in face-centered cubic stainless steel, confined to their (111) glide plane, as observed in high voltage electron microscopy. The strictly planar arrangement implies that the dislocations are dissociated and that, for this reason, the cross-glide is a difficult process. The dissociation is too small to be directly observable at this resolution, but it has been found from other images that the stacking fault energy is rather low in stainless steel. Note the periodic contrast of the dislocations in the vicinity of their emergence points in the foil surfaces and the absence of such contrast in the central part of the foil. Figure 89 is an image of a low stacking fault energy alloy (Cu-Ga); the dissociation is clearly visible and stacking fault fringes can be observed between the partial dislocations.



**Figure 88.** High voltage electron micrograph showing a procession of dislocations confined to a glide plane in stainless steel. Note the wavy contrast close to the surfaces and its absence in the central part of the foil.

Figure 90 shows a network of intersecting glide dislocations confined to the (111) glide plane in a face-centered cubic copper alloy (Cu-Ga) with a low stacking fault energy. One set of dislocation nodes is dissociated and gives rise to the dark triangular areas; the other set is contracted. Such nodes allow one to deduce the stacking fault energy from the curvature of the partial dislocations forming the extended nodes.

Figure 91 shows glide dislocations in the layer plane (001) of  $\text{NbTe}_2$ , which is parallel to the foil plane, the specimen having been obtained by cleavage [58]. In every other domain the dislocation

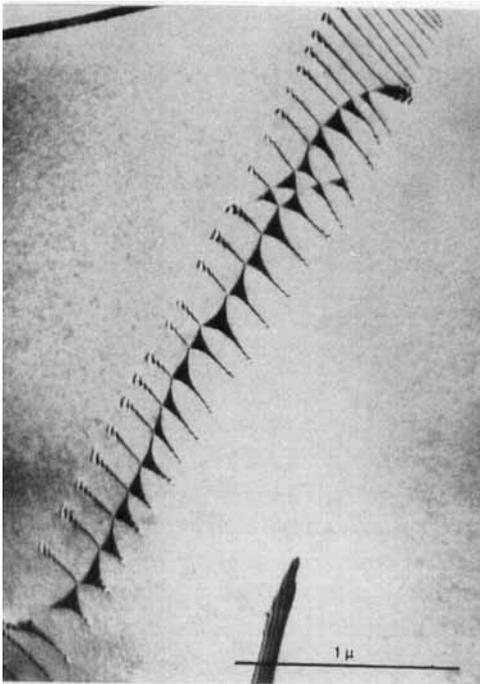


**Figure 89.** Ribbons dissociated in Shockley partials observed in a Cu-Ga alloy. (Courtesy of A. Art.)

multiribbons consist of six partial dislocations; in the remaining domains the dislocations are simple ribbons, as described above. The image illustrates the strong interactions of the glide dislocations with the domain walls. On entering a domain in which the six-fold ribbons would have to be formed, the single ribbons line up with the domain wall, in this way minimizing the generation of stacking faults. This leads to an effective interaction between the dislocation ribbons and the domain walls.

### The Structure of Subgrain Boundaries

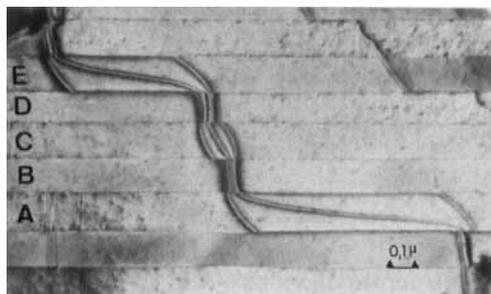
Small-angle grain boundaries can be described in terms of arrays of dislocation



**Figure 90.** Network of dissociated dislocations in a Cu-Ga alloy with a low stacking fault energy. (Courtesy of A. Art.)

lines. Diffraction contrast electron microscopy has contributed significantly to firmly establishing dislocation models for such boundaries.

A general subgrain boundary is characterized by five parameters describing its



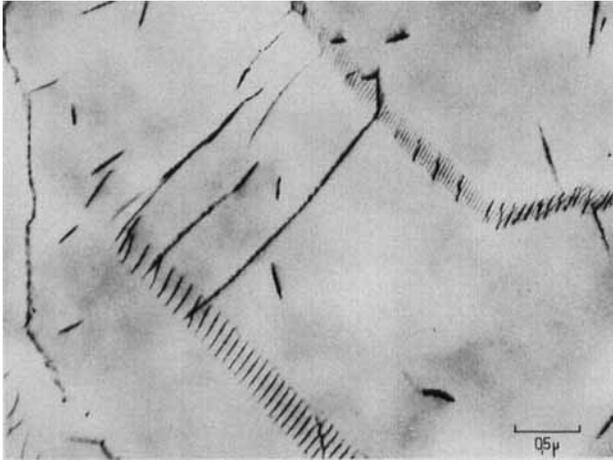
**Figure 91.** Glide dislocations in the layer plane of  $\text{NbTe}_2$ . Note the interaction between dislocations and twin domain walls [58].

geometry: the rotation axis, the rotation angle, and the normal to the contact plane. These parameters can be determined by the combined use of the spot diffraction pattern and of the Kikuchi line pattern. The diffraction contrast image then allows one to visualize the geometry of the dislocation lines and, using the extinction criterion, to determine their Burgers vectors. If the rotation axis is parallel to the contact plane the boundary is a tilt boundary and the dislocation configuration consists of parallel lines. If, on the other hand, the rotation axis is perpendicular to the contact plane the boundary consists of a network of intersecting dislocations. Depending on the symmetry of the contact plane this network may ideally consist of square meshes in (100) planes or of hexagonal meshes in (111) planes.

Figure 92 shows two tilt boundaries in body-centered niobium consisting of parallel dislocation lines. Some of the dislocations are decorated by small precipitate particles.

The image of the twist boundaries in platinum [59] shown in Fig. 93a illustrates a hexagonal network containing three intersecting families of dislocations with Burgers vectors enclosing angles of  $120^\circ$ . Figures 93b and 93c represent square networks, consisting of dislocations with mutually perpendicular Burgers vectors.

Figure 94 shows a well-developed hexagonal network of undissociated dislocations in the (0001) plane of hexagonal zinc [60]. All dislocations are mobile in the plane of the boundary. The right part of the boundary moved along the (0001) glide plane during the exposure, leading to blurring of the image. In the



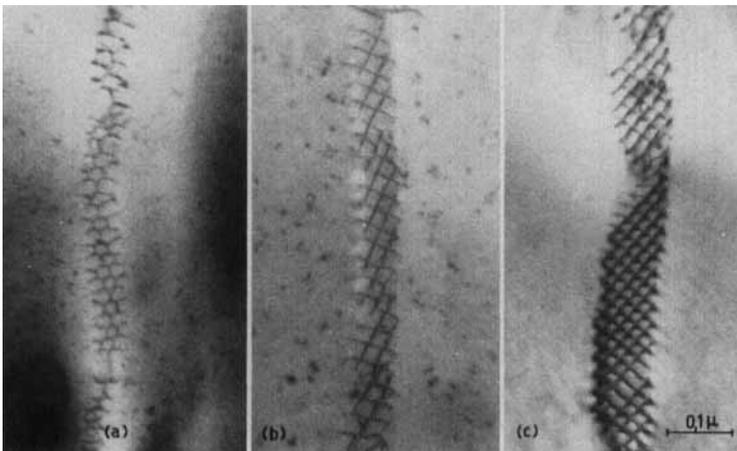
**Figure 92.** Tilt boundary consisting of sets of parallel dislocations in niobium. Some of the dislocations are decorated by small particles. (Courtesy of A. Fourdeux and A. Berghezan.)

left part of the image some of the dislocations, leaving the network and terminating in the foil surfaces, have become sessile, and they thus pinned the network in that part.

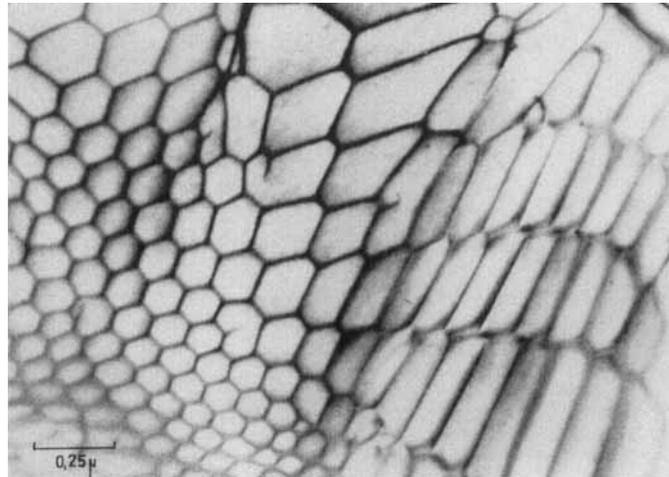
Figure 21 shows a hexagonal network of widely extended dislocations in the basal plane of graphite [56, 61]. The network is, in fact, a glissile twist boundary. From the curvature of the partial dislocations in the extended nodes one can deduce the stacking fault energy.

### Point Defect Clusters

Vacancies in quenched metals form disc-shaped agglomerates in (111) face-centered-cubic or (0001) hexagonal-close-packed layers, limited by Frank-type dislocation loops. If the stacking fault energy is large enough the loop is 'unfaulted', since energy is gained by nucleating a Shockley partial dislocation and sweeping the loop, transforming the sessile Frank loop into a perfect glissile loop. Such unfaulted loops



**Figure 93.** Twist boundaries in platinum [59]: (a) hexagonal network; (b, c) square networks. (Courtesy of E. Ruedl.)



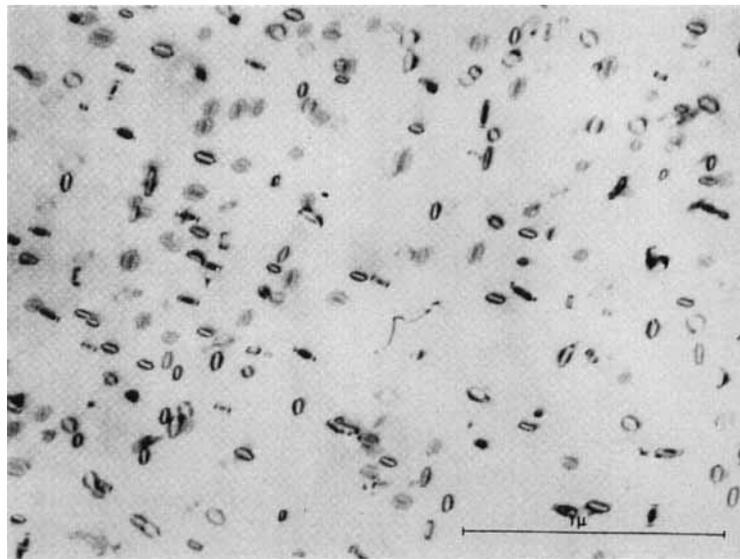
**Figure 94.** Hexagonal network of undissociated dislocations in the (0001) plane of zinc [60].

in quenched aluminum are shown in Fig. 95 [62, 63].

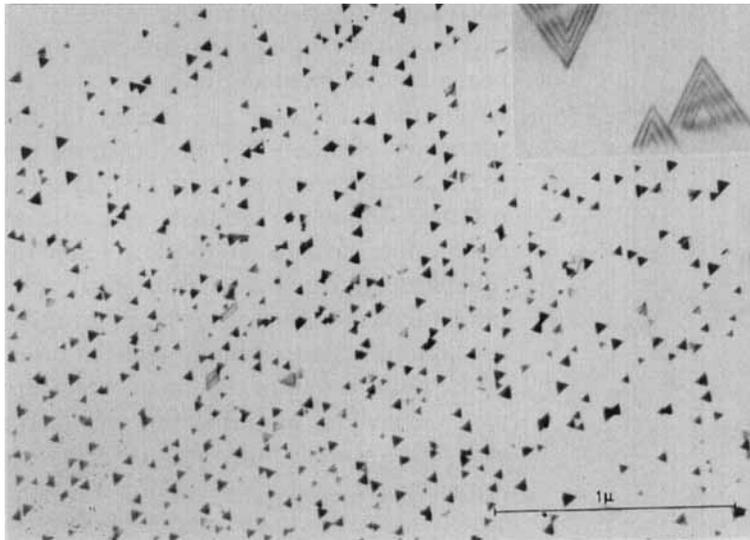
If the stacking fault energy is small enough, which is true in, for instance, gold and Ni-Co alloys, the Frank loop is transformed into a stacking fault tetrahedron consisting of four intersecting triangular stacking faults in  $\{111\}$  planes, limited along their intersection lines by

edge-type stair-rod dislocations with a Burgers vector of the type  $\frac{1}{6}[110]$ . For intermediate values of the stacking fault energy the Frank loops may remain faulted. Figure 96 shows stacking fault tetrahedra in gold imaged in diffraction contrast [62, 64].

Faulted Frank loops in silicon are shown imaged in diffraction contrast in

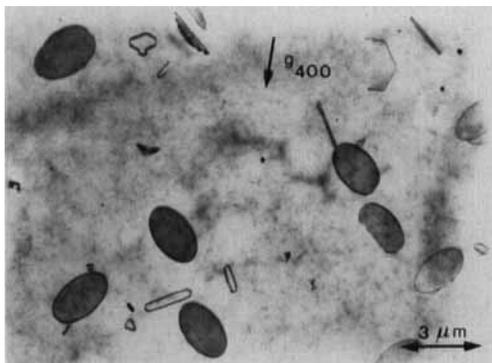


**Figure 95.** Unfaulted dislocation loops in quenched aluminum [62, 63].



**Figure 96.** Diffraction contrast image of stacking fault tetrahedra in quenched gold. The inset shows a magnified image [62, 64].

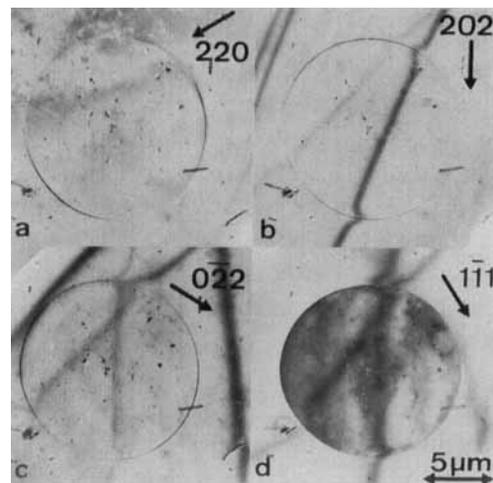
Fig. 97. The presence of the stacking fault causes contrast inside the loop. Figure 98 shows a contrast experiment on an extrinsic Frank-type dislocation loop in silicon. Note the presence of a line of no contrast perpendicular to the active  $g$  vector and the deformation of the extinction contours where they cross the dislocation loop. For  $g = [1\bar{1}1]$  the loop exhibits stacking fault contrast, as do the loops in Fig. 97.



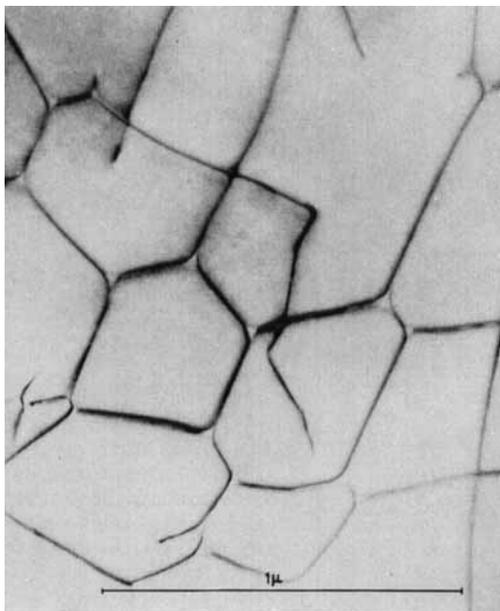
**Figure 97.** Faulted Frank loops in silicon due to interstitials. The loops exhibit stacking fault contrast. (Courtesy of H. Bender.)

### Planar Interfaces

It is well known that two simple types of stacking fault can occur in the face-centered-cubic structure. The *intrinsic fault*, formed either by the extraction of a layer or by glide of a Shockley partial, is represented by the stacking symbol



**Figure 98.** Contrast experiment on a faulted loop in silicon.



**Figure 99.** Network of extended dislocations in silicon; all nodes are dissociated. (Courtesy of E. Aerts.)

abcabcabc... The *extrinsic fault*, formed for instance by the precipitation of interstitials in a Frank loop, corresponds to the stacking symbol abcababc... The two types of fault have comparable energies in certain materials. In a network of dissociated dislocations all nodes are then dissociated; this is, for instance, the case in silicon (Fig. 99) and in certain alloys (e.g., Ag-Sn). The two kinds of fault have opposite displacement vectors of the type  $a/3 [111]$  and can thus be distinguished by the characteristic fringe pattern which they produce when situated in inclined planes (see Sec. 1.1.1.13 and Sec. 1.1.1.20 of this Chapter) [65–67].

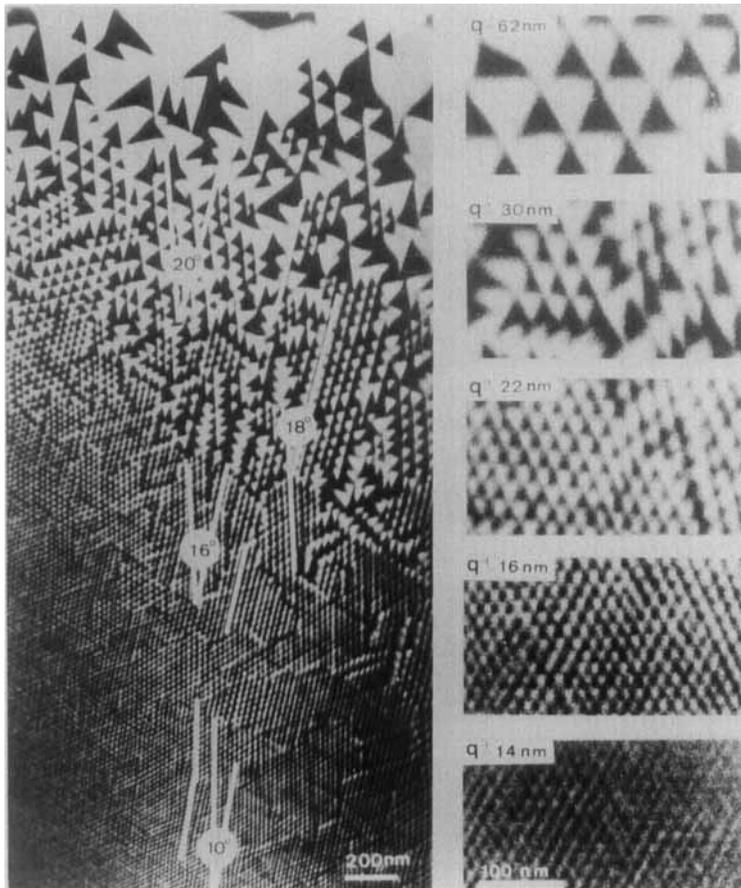
### Domain Structures

Phase transformations are usually accompanied by a decrease in symmetry with

decreasing temperature. As a result, a single crystal of a higher symmetric phase becomes fragmented into domains of which the structures are related by the symmetry elements lost in the transition to the lower symmetry phase [68]. The lost rotation symmetry elements give rise to orientation variants of the low temperature phase of which the number is given by the ratio of the order of the point group of the high temperature phase and the order of the point group of the low temperature phase. The loss of translation symmetry gives rise to translation variants related by displacement vectors given by the lost lattice translations. Their number is determined by the ratio of the volumes of the primitive unit cells of the low and high temperature phases.

Orientation variants are separated by domain boundaries, whereas translation variants are separated by out-of-phase boundaries. The orientation of the domain boundaries is determined by the requirement that the strain energy should be a minimum. This will be the case for strain-free interfaces. As a result, the orientation of certain interfaces ( $W$ ) follows entirely from symmetry, whereas others ( $W'$ ) have orientations which depend on the lattice parameters of the two phases involved, at the transition temperature [69–71].

For example, in the  $\alpha$ - $\beta$  transition of quartz, referred to above, the  $\alpha$  phase has the point group 32 (order 6) and the  $\beta$  phase has the point group 622 (order 12). The number of orientation variants in the  $\alpha$  phase is thus  $12/6 = 2$  ( $\alpha_1$  and  $\alpha_2$ ; Dauphiné twins) and they are related by the lost  $180^\circ$  rotation about the three-fold axis. There is no change in translation symmetry. Images of domain fragmented  $\alpha$  phase are shown in Fig. 100. In the case



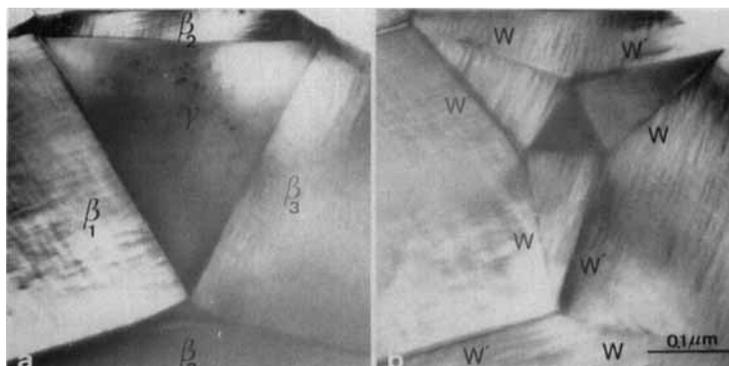
**Figure 100.** Domain fragmentation in quartz as a function of temperature. A temperature gradient is present across the specimen. At the highest temperature the incommensurately modulated phase is observed [28].

of quartz the situation is actually somewhat more complicated by the occurrence of an intermediate incommensurate phase between the  $\alpha$  and  $\beta$  phases and which is only stable within a narrow temperature range ( $\approx 1.5$  K). This phase was discovered using diffraction contrast electron microscopy [30]. It consists of a regular texture of triangular prisms parallel to the  $c$  axis, of  $\alpha_1$  and  $\alpha_2$  structure. The size of the triangular prisms decreases with increasing temperature in the vicinity of the transition temperature (Fig. 100).

Quite striking domain structures were studied using diffraction contrast in the

monoclinic room temperature phase of ferroelastic lead orthovanadate  $[\text{Pb}_3(\text{VO}_4)_2]$  [70, 71]. The structure is rhombohedral at high temperature ( $\gamma$  phase), but on cooling it transforms at  $120^\circ\text{C}$  into a monoclinic structure ( $\beta$  phase) which is stable at room temperature. The rhombohedral parent phase is fragmented into domain patterns which minimize the strain energy. They consist of combinations of completely symmetry determined walls ( $W$ ) and walls ( $W'$ ), the strain-free orientation of which depends on the lattice parameters below and above the  $\gamma \leftrightarrow \beta$  transition temperature

**Figure 101.** Domain pattern in lead orthovanadate  $[\text{Pb}_3(\text{VO}_4)_2]$  resulting from the  $\gamma \rightarrow \beta$  phase transition. The central triangle of the star pattern is still in the  $\gamma$  phase. The two images refer to the same area; in (b) the temperature was somewhat lower than that in (a) [71].

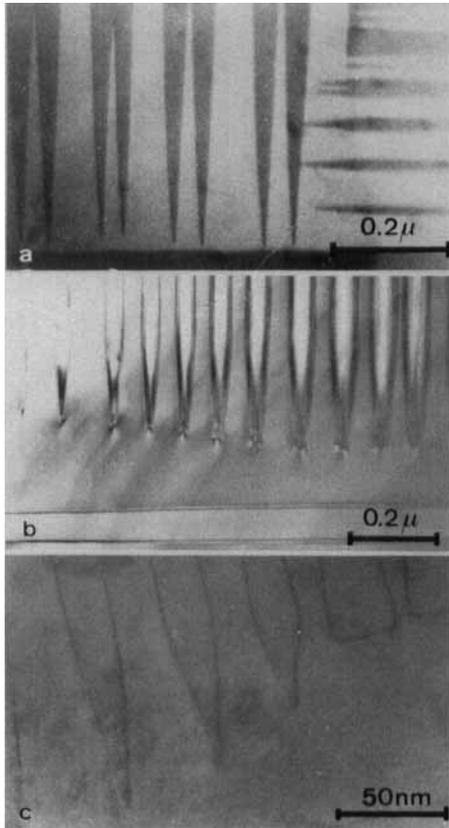


(i.e., on the spontaneous strain tensor). The most striking configuration is the pattern shown in Fig. 101; it contains three concentric 'stars' of decreasing size. The pattern in Fig. 101a shows a central triangle of metastable  $\gamma$  phase surrounded by areas consisting of three different variants of the  $\beta$  phase. On cooling this triangle transforms further 'in situ' into the configuration shown in Fig. 101b; the  $\gamma$  triangle becomes smaller and is rotated by  $180^\circ$ . Similar patterns occur in other domain textures resulting from a phase transformation between the parent and product phases belonging to the same point groups as  $\gamma$  and  $\beta$  lead orthovanadate, respectively.

The compound  $\text{YBa}_2\text{Cu}_3\text{O}_7$  is tetragonal at high temperature where the  $\cdots\text{O}-\text{Cu}-\text{O}-\text{Cu}\cdots$  chains in the  $\text{CuO}$  layers are disordered. Below the transition temperature, which depends on the oxygen content, the chains order in any given area along one out of two mutually perpendicular, equally probable, orientations, which then becomes the  $b_0$  direction of the orthorhombic structure. The disorder-order transition thus produces two structural variants with their  $b_0$  axes roughly perpendicular and which are twin related by a mirror operation with

respect to  $(110)$  or  $(1\bar{1}0)$ . These two orientation variants are revealed, using different imaging modes, in Fig. 102 [72].

The ordering of magnetic moments below the Néel temperature in antiferromagnetic materials is usually accompanied by a structural phase transition. This leads to the formation of an antiferromagnetic domain structure of which the domain walls coincide with those due to the structural phase transition [73]. Such a combined transition occurs, for instance, in  $\text{NiO}$ , which has the sodium chloride structure above the Néel point ( $525\text{ K}$ ). Below this temperature the nickel spins order in such a way that the spins in one of the families of  $(111)$  planes order ferromagnetically, the spin direction being parallel to these  $(111)$  planes, whereas successive  $(111)$  sheets contain oppositely oriented spins. As a consequence of magnetostriction, the structure contracts along the  $\langle 111 \rangle$  direction perpendicular to these sheets, and the lattice becomes rhombohedral ( $\alpha = 90^\circ 4'$ ). The rhombohedral structures in adjacent antiferromagnetic domains contract along different  $\langle 111 \rangle$  directions, and as a result such domains are separated by coherent twin boundaries with a very small twinning vector, which are imaged as  $\delta$ -fringe patterns. Two such



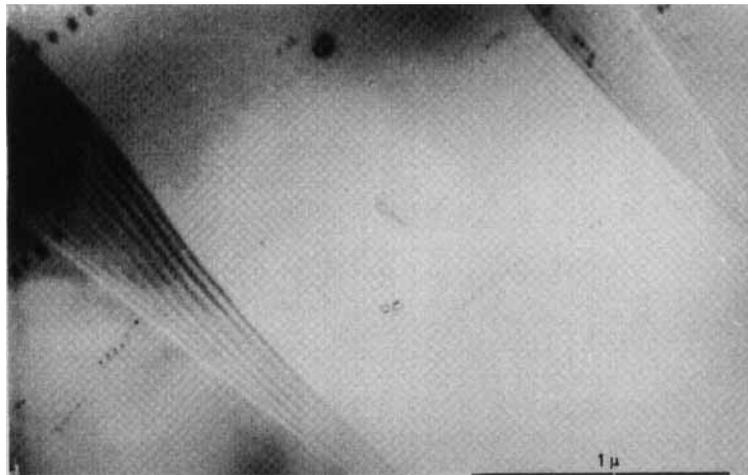
**Figure 102.** Orthorhombic twins in  $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$  revealed using three different contrast modes: (a) domain contrast; (b) interface contrast; (c) high resolution imaging. Note the strain at the twin tips in (b) [72].

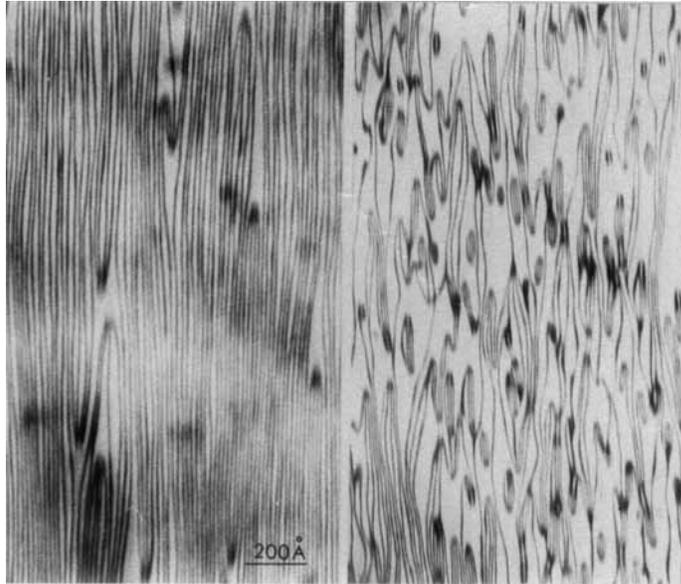
domain walls are shown in Fig. 103 which is a bright-field image of two parallel domain walls for which the  $\delta$  values are opposite in sign. This is reflected in the opposite nature of the edge fringes for the two boundaries.

### The Structure of Ordered Alloys

Observations of 'discommensurations' and of 'discommensuration nodes' [74] were performed on the alloy  $\text{Ni}_{3+}\text{Mo}$  [75], using diffraction contrast, at a time when the term 'discommensuration' had not yet been introduced. The interfaces shown in Fig. 104 were described as 'out-of-phase boundaries' with a displacement vector equal to one-quarter of a lattice vector. Although in alloys there is no essential difference between out-of-phase boundaries and 'discommensuration walls', the

**Figure 103.** Two domain walls with opposite  $\delta$  in antiferromagnetic nickel oxide; they are imaged as  $\delta$ -fringe patterns [73].





**Figure 104.** Four-fold discommensuration nodes in  $\text{Ni}_3\text{Mo}$ , revealed by diffraction contrast [75].

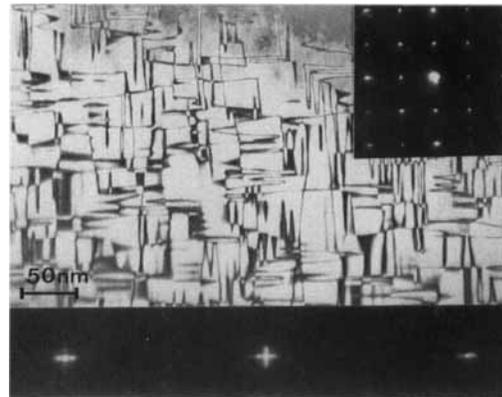
defects shown in Fig. 104 would, at present, presumably be termed 'discommensurations' by most authors.

Conservative antiphase boundaries in the alloy  $\text{Cu}_3\text{Pd}$  with  $L1_2$  structure revealed by diffraction contrast are shown in Fig. 105. These boundaries represent the first stage in the formation of a one-dimensional long period antiphase boundary structure from the disordered phase. A number of nonconservative antiphase boundaries become unstable and start 'meandering', forming parallel sets of conservative antiphase boundaries [76, 77].

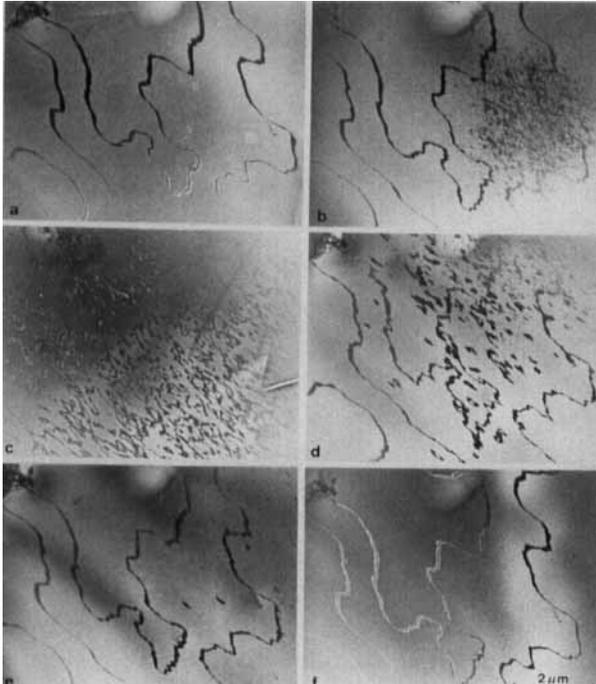
### Minerals

Anorthite ( $\text{CaAl}_2\text{Si}_2\text{O}_8$ ) is a complicated silicate which has a primitive triclinic Bravais lattice (space group  $P\bar{1}$ ) at room temperature. Above  $T_c = 514\text{ K}$  the same unit cell becomes body centered ( $I\bar{1}$ ). This can be concluded from the diffraction pattern since the spots of the type

$h + k + l = \text{Odd}$  gradually disappear above  $T_c$ . On cooling the crystal from the high temperature phase to room temperature, it breaks up into two translation variants separated by very 'ragged' antiphase boundaries with a  $\frac{1}{2}[111]$  displacement vector. No orientation variants are formed. The domain boundaries are



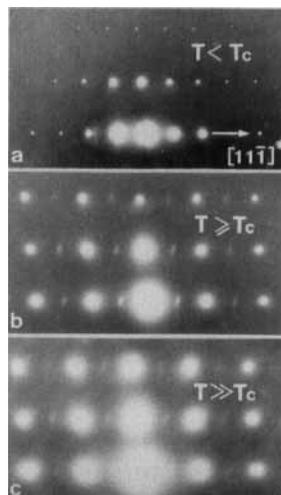
**Figure 105.** Diffraction contrast image of the first stage in the formation of a one-dimensional long period structure in  $\text{Cu}_3\text{Pd}$ . Note the 'meandering' of the antiphase boundaries [77].



**Figure 106.** Evolution of antiphase boundaries in anorthite ( $\text{CaAl}_2\text{Si}_2\text{O}_8$ ) during a heating-cooling cycle from room temperature up to above 514 K. All images refer to the same area. Note the memory effect [78].

revealed by diffraction contrast dark field imaging in reflections for which  $h + k + l = \text{Odd}$ . On heating above 514 K

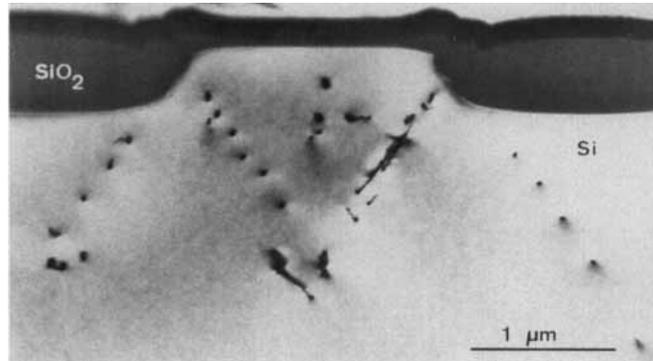
the boundaries disappear, but on cooling they reappear at exactly the same place and with the same shape as before (i.e., there is a pronounced memory effect, presumably due to impurity pinning). This is illustrated by the heating-cooling cycle shown in Fig. 106; the corresponding diffraction patterns along [101] are shown in Fig. 107 [78].



**Figure 107.** Evolution of the diffraction pattern during the same heating-cooling cycle as in Fig. 106: (a)  $T < T_c$ ; (b)  $T \geq T_c$ ; (c)  $T \gg T_c$ .

### Fabrication-induced Defects in Semiconductors

Semiconductor single-crystal ‘chips’ often undergo a long sequence of fabrication steps (thermal treatment, oxidation, etching, etc.) some of which can be accompanied by a deterioration of the crystal’s physical properties and hence affect the performance of the final device. The



**Figure 108.** TEM image of a cross-section of a field-effect device. Dislocations are emitted from the edges of the constriction in the silicon oxide layer; the dislocations are seen end-on [79].

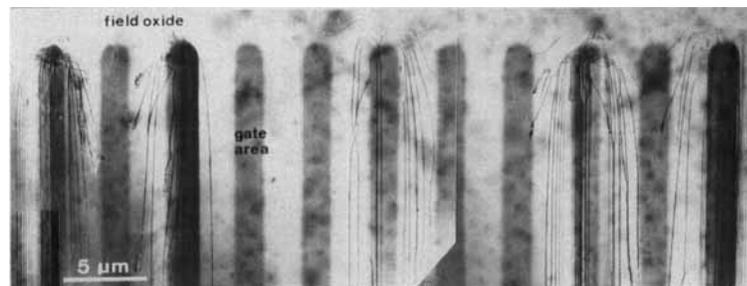
micro-miniaturization of electron devices makes detailed control of the crystal perfection strongly dependent on electron microscope techniques; both on high resolution images of cross-section specimens of devices, and on high voltage electron microscopy for the study of 'thick' specimens at low resolution and small magnification.

Figure 108 shows a transmission electron microscopy (TEM) image of processions of dislocations observed end-on in a cross-sectional view of a field-effect device. At the edge of the constriction in the silicon oxide layer sources have generated dislocations along the glide planes of maximum resolved shear stress, in order to relieve the stresses generated by the oxidation process. The dislocations apparently form 'inverse' pile-ups, their spacing being smallest close to the source [79].

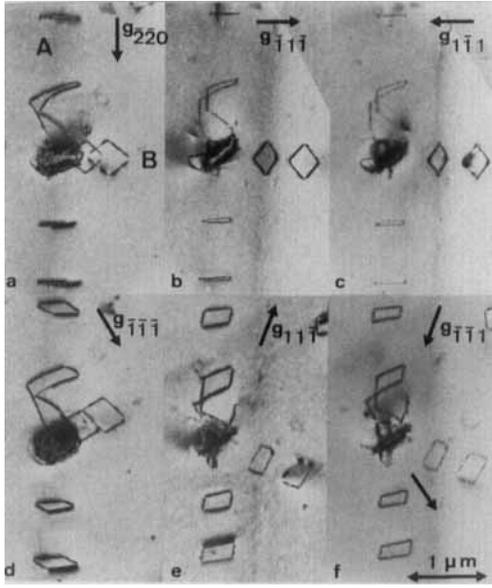
Figure 109 shows 'finger'-shaped gate areas formed in a field oxide layer on a silicon chip. They have similarly generated stresses (see above) which are relieved by dislocation generation. In this case the dislocations are imaged in a plane view.

Oxide or other precipitate particles may put the surrounding silicon matrix under a compressive stress. This stress is often large enough to give rise to 'prismatic punching' whereby discs of self-interstitials surrounded by a loop of perfect dislocation are emitted. Such loops are glissite on a cylindrical surface of which the cross-section is determined by the precipitate's shape and the direction of the generators by the Burgers vector of the dislocations (i.e.,  $\frac{1}{2}\langle 110 \rangle$ ) (Fig. 110).

Interfacial dislocations are often formed at the interface between the silicon substrate and metallic layers used as

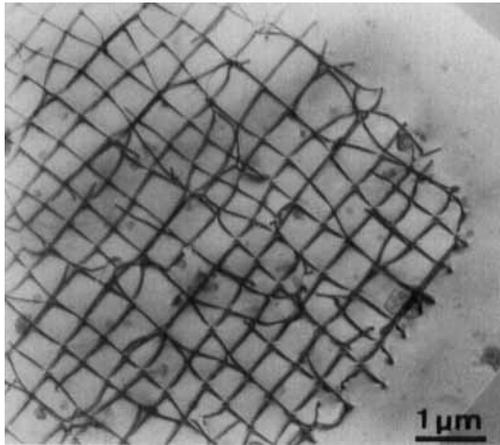


**Figure 109.** Finger-shaped gate areas in a field oxide. Dislocations are generated along the edges; they are observed in a plane view. (Courtesy of Vanhellemont.)



**Figure 110.** Prismatic punching around a precipitate particle in a silicon matrix imaged by six different diffraction vectors. (Courtesy of H. Bender.)

electrical contacts. Figure 111 shows the networks of misfit dislocations between silicon and aluminum; the image was obtained in a plane view using diffraction contrast.



**Figure 111.** Network of interfacial dislocations in the contact plane between silicon and aluminum.

## Various Applications

### *In Situ Studies*

The availability of cooling and heating specimen holders allows in situ study of the phenomena accompanying phase transitions. When going through a disorder–order transition, different phases of the domain fragmentation can be followed. The creation and elimination of discommensuration walls is directly observable in dark field images made in clusters of incommensurate reflections [80, 81].

When performing such observations one should be aware of the effect of the electron beam on the specimen, which results in an increase in temperature (depending on the thermal conductivity of the foil) and may also cause some radiation effects (which may interfere with the transition).

### *Radiation Damage*

Electron microscopy, in particular high voltage electron microscopy, has been used extensively to study in situ radiation effects as well as postirradiation defect configurations. The point defects, precipitates, and small dislocation loops can be characterized using the methods described above (see Sec. 1.1.1.18 of this Chapter).

### *Radiation Ordering*

Some surprising results were obtained from in situ studies of ordering alloys that exhibit a short-range order state (e.g.,  $\text{Ni}_4\text{Mo}$ ). When irradiated with 1 MV electrons at low temperature, ordered  $\text{Ni}_4\text{Mo}$  becomes completely disordered. When irradiating in a temperature range below, but close to, the

order–disorder transition temperature, the irradiation causes the alloy to assume order up to a certain degree. The order parameters can be determined by following the evolution of the intensity of the order diffraction spots. These phenomena result from the competition between the ordering effect due to radiation-enhanced diffusion at the irradiation temperature and the disordering effect of the irradiation as a result of atomic collisions. Over a certain temperature range the short-range order state is produced by irradiation. Certain alloy phases which could not be ordered by thermal treatment (e.g., Ni<sub>4</sub>W), were found to assume order under electron irradiation.

#### *Determination of the Type of Stacking Fault*

Close-packed layers of atoms can be stacked in an infinite number of ways, all of which have nearly the same free energy. Two essentially different types of stacking mode are usually distinguished: the face-centered-cubic stacking mode (...ABCABC...) and the hexagonal-close-packed mode (...ABAB..., or ...ACAC..., or ...BCBC...).

In face-centered stacking two different types of stacking fault are often distinguished. If a single atomic layer is extracted and the gap so created is closed by a displacement over a vector  $\mathbf{R}_0 = \frac{1}{3}[111]$ , the resulting sequence is ...BCA<sub>↓</sub>CABCABC...

This is called an *intrinsic stacking fault*. It is formed, for instance, by the precipitation of a layer of vacancies, but it is also generated in the wake of a glissile Shockley partial dislocation with Burgers vector  $\frac{1}{6}[11\bar{2}]$  on a (111) glide plane.

If a single atomic layer is inserted, when for instance a layer of interstitials precipitates, the resulting sequence is

...ABC<sub>↑</sub>BABCABC...

This is called an *extrinsic stacking fault*. The displacement vector is now  $\mathbf{R}_0 = -\frac{1}{3}[111]$  (i.e., the opposite of the previous one). A single glide dislocation cannot generate such a fault. In both faults two triplets in the hexagonal configuration occur, but in a different configuration.

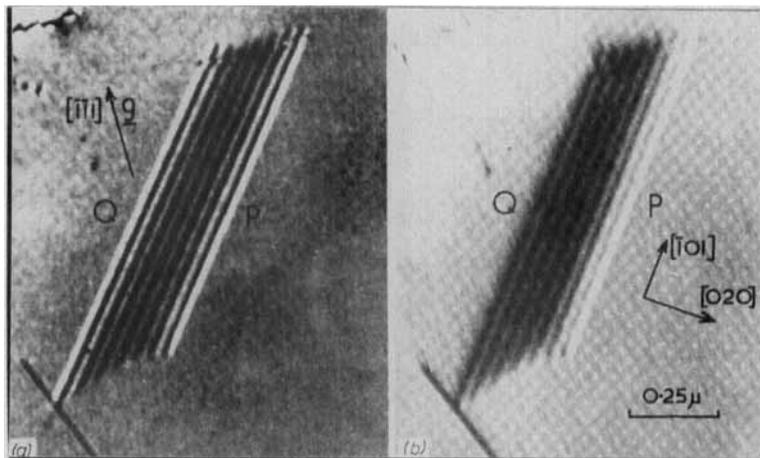
To make a detailed interpretation of partial dislocation–fault configurations in face-centered-cubic metals, it is important to distinguish between intrinsic and extrinsic stacking faults. It has been shown that this information can be obtained from the nature of the edge fringes in stacking fault images [82–84]. In particular [85], it has been demonstrated how this information can be obtained from a single dark field image made in a well-defined reflection.

One can distinguish three classes of reflections in the face-centered-cubic structure, depending on whether or not  $h+k+l = \text{Three-fold}$ ,  $\text{Three-fold} + 1$ , or  $\text{Three-fold} - 1$ . The reflections for which  $h+k+l = \text{Three-fold}$  lead to  $\alpha = k \cdot 2\pi$  and, therefore, do not give rise to a fringe pattern. Reflections such as  $\{200\}$ ,  $\{2\bar{2}0\}$ , and  $\{440\}$ , for which  $h+k+l = \text{Three-fold} - 1$ , will be called type A, whereas reflections such as  $\{1\bar{1}1\}$ ,  $\{220\}$ , and  $\{400\}$ , for which  $h+k+l = \text{Three-fold} + 1$ , will be called type-B reflections.

The edge fringes in bright and dark field images for all possible combinations of the type of active reflection (A or B), the sense of inclination of the fault planes, and the nature of the fault (E or I) are represented schematically in Table 2 where the diffraction vector is assumed to point to the right.

**Table 2.** Determination of the type of stacking fault in the face-centered-cubic structure.

Bright field				Dark field			
A		B		A		B	
⋮	D	⋮	B	⋮	B	⋮	D
⋮		⋮		⋮		⋮	
⋮	B	⋮	D	⋮	B	⋮	D
⋮		⋮		⋮		⋮	
⋮	B	⋮	D	⋮	D	⋮	B
⋮		⋮		⋮		⋮	
⋮	B	⋮	D	⋮	B	⋮	D
⋮		⋮		⋮		⋮	
⋮	D	⋮	B	⋮	D	⋮	B
⋮		⋮		⋮		⋮	



**Figure 112.** Bright (a) and dark field (b) image pair of a stacking fault in a small stacking fault energy alloy. From Table 2 it can be concluded that the fault is intrinsic.

Note that for a given type of fault and a given type of vector the nature of the edge fringes in the dark field image is independent of the sense of inclination of the fault plane. A simple rule can thus be formulated: if in the dark field image the  $g$  vector, its origin being put in the center of the

fringe pattern, points towards a bright fringe and the operating reflection  $g$  is of type A, then the fault is intrinsic. If one of the parameters (the nature of the edge fringes or the class of the operating reflection) changes, the conclusion also changes. The nature of the edge fringes in the bright

field image also allows one to determine the sense of inclination of the fault plane. In applying the present method one must be aware of the fact that the nature of the edge fringes is only well defined in sufficiently thick foils, where anomalous absorption is important.

Figure 112 shows a bright and dark field image pair which allows us to conclude that the fault being imaged is an intrinsic fault.

### Acknowledgements

Thanks are due to the numerous colleagues for providing photographs for illustrations.

Reference is made in the captions to the original publications or in the case of unpublished photographs to the original author. In particular Prof. G. Van Tendeloo provided several unpublished photographs.

### 1.1.1.21 References

- [1] N. F. Mott, H. S. W. Massey, *The Theory of Atomic Collisions*, Clarendon, Oxford, **1949**.
- [2] J. A. Ibers, B. K. Vainshtein, in *International Crystallographic Tables*, Volume III, Tables 3.3.3 A(1) and A(2), Kynock Press, Birmingham, **1962**.
- [3] W. L. Bragg, *Nature* **1929**, *124*, 125.
- [4] P. P. Ewald, *Ann. Phys.* **1917**, *54*, 519.
- [5] S. Takagi, *Acta Crystallogr.* **1962**, *15*, 1311.
- [6] S. Kiguchi, *Jpn J. Phys.* **1928**, *5*, 23.
- [7] C. G. Darwin, *Phil. Mag.* **1914**, *27*, 315, 675.
- [8] G. Thomas, in *Diffraction and Imaging Techniques in Material Science* (Eds.: S. Amelinckx, R. Gevers, J. Van Landuyt), North-Holland, Amsterdam, **1978**, p. 217.
- [9] R. Gevers, in *Diffraction and Imaging Techniques in Material Science* (Eds.: S. Amelinckx, R. Gevers, J. Van Landuyt), North-Holland, Amsterdam, **1978**, p. 9.
- [10] P. B. Hirsch, A. Howie, M. J. Whelan, *Phil. Mag.* **1962**, *7*, 2095.
- [11] P. B. Hirsch, A. Howie, M. J. Whelan, *Phil. Trans. R. Soc., Ser. A* **1960**, *252*, 499.
- [12] R. Gevers, *Phil Mag.* **1962**, *7*, 59, 651, 769, 1681.
- [13] A. Howie, M. J. Whelan, in *Proc. European Reg. Conf. on Electron Microscopy, Delft*, Vol. 1, **1960**, p. 194.
- [14] A. Howie, M. J. Whelan, *Proc. R. Soc., Ser. A* **1961**, *263*, 206, 217.
- [15] A. Amelinckx, D. Van Dyck, *Electron Diffraction Techniques*, Vol. 2, IUCr Monographs on Crystallography 4, Oxford Science Publications, Oxford, **1993**, p. 000.
- [16] C. G. Darwin, *Phil. Mag.* **1914**, *27*, 315, 675.
- [17] S. Amelinckx, in *Ullmann's Encyclopedia of Industrial Chemistry*, Vol. B6, VCH, Weinheim, **1994**, p. 229.
- [18] S. Amelinckx, in *Solid State Physics* (Eds.: F. Seitz, D. Turnbull), Academic Press, London, **1964**, p. 405.
- [19] S. Amelinckx, in *Solid State Physics* (Eds.: F. Seitz, D. Turnbull), Academic Press, London, **1964**, p. 409.
- [20] J. W. Menter, *Proc. R. Soc., Ser. A* **1956**, *236*, 119.
- [21] S. Amelinckx, *J. Electron. Microscop. Technol.* **1986**, *3*, 131.
- [22] H. Yoshioka, *J. Phys. Soc. Jpn* **1957**, *12*, 628.
- [23] G. Borrmann, *Phys. Z.* **1941**, *42*, 157; **1950**, *127*, 297.
- [24] S. Amelinckx, J. Van Landuyt, in *Diffraction and Imaging Techniques in Material Science* (Eds.: S. Amelinckx, R. Gevers, J. Van Landuyt), North-Holland, Amsterdam, **1978**, p. 107.
- [25] R. Gevers, J. Van Landuyt, S. Amelinckx, *Phys. Stat. Solidi* **1965**, *11*, 689.
- [26] G. Van Tendeloo, S. Amelinckx, *Phys. Stat. Solidi A* **1974**, *22*, 621.
- [27] G. Van Tendeloo, S. Amelinckx, *Acta Crystallogr., Sect. A* **1974**, *30*, 431.
- [28] J. Van Landuyt, G. Van Tendeloo, S. Amelinckx, M. B. Walker, *Phys. Rev. B* **1985**, *31*, 2986.
- [29] G. Dolino, P. Bachheimer, B. Berge, C. M. Zeyen, G. Van Tendeloo, J. Van Landuyt, S. Amelinckx, *J. Phys.* **1984**, *45*, 901.
- [30] G. Van Tendeloo, J. Van Landuyt, S. Amelinckx, *Phys. Stat. Solidi A* **1976**, *33*, 723.
- [31] M. Meulemans, P. Delavignette, F. Garcia-Gonzales, S. Amelinckx, *Mat. Res. Bull.* **1970**, *5*, 1025.
- [32] C. Boulesteix, J. Van Landuyt, S. Amelinckx, *Phys. Stat. Solidi A* **1976**, *33*, 595.
- [33] C. Manolikas, J. Van Landuyt, S. Amelinckx, *Phys. Stat. Solidi A* **1979**, *53*, 327.
- [34] R. Serneels, M. Snijkers, P. Delavignette, R. Gevers, S. Amelinckx, *Phys. Stat. Solidi B* **1973**, *58*, 277.
- [35] J. D. Eshelby, A. N. Stroh, *Phil. Mag.* **1951**, *42*, 1401.

- [36] M. Snijkers, R. Serneels, P. Delavignette, R. Gevers, S. Amelinckx, *Cryst. Latt. Def.* **1972**, *3*, 99.
- [37] M. Mannami, *J. Phys. Soc. Jpn* **1962**, *17*, 1160.
- [38] M. Mannami, *Acta Crystallogr.* **1960**, *13*, 363.
- [39] P. Delavignette, R. Trivedi, R. Gevers, S. Amelinckx, *Phys. Stat. Sol.* **1966**, *17*, 221.
- [40] M. F. Ashby, L. M. Brown, *Phil. Mag.* **1963**, *8*, 1083, 1649.
- [41] M. Wilkens, in *Diffraction and Imaging Techniques in Material Science* (Eds.: S. Amelinckx, R. Gevers, J. Van Landuyt), North-Holland, Amsterdam, **1978**, p. 185.
- [42] U. Essmann, M. Wilkins, *Phys. Stat. Solidi* **1964**, *4*, K53.
- [43] D. J. H. Cockayne, M. J. Jenkins, I. L. E. Ray, *Phil. Mag.* **1971**, *24*, 1383.
- [44] D. J. H. Cockayne, I. L. E. Ray, M. J. Whelan, *Phil. Mag.* **1969**, *20*, 1265.
- [45] R. De Ridder, S. Amelinckx, *Phys. Stat. Solidi B* **1971**, *43*, 541.
- [46] R. Gevers, *Phys. Stat. Solidi* **1963**, *3*, 415.
- [47] P. Humble, in *Diffraction and Imaging Techniques in Material Science* (Eds.: S. Amelinckx, R. Gevers, J. Van Landuyt), North-Holland, Amsterdam, **1978**, p. 315.
- [48] A. K. Head, P. Humble, L. M. Clarebrough, A. T. Morton, C. T. Forwood, 'Computed Electron Micrographs and Defect Identification', in *Defects in Crystalline Solids*, Vol. 7 (Eds.: S. Amelinckx, R. Gevers, J. Nihoul).
- [49] P. Delavignette, R. Trivedi, R. Gevers, S. Amelinckx, *Phys. Stat. Sol.* **1966**, *17*, 221.
- [50] R. Gevers, *Phys. Stat. Solidi* **1963**, *3*, 2289.
- [51] D. W. Pashley, J. W. Menter, G. A. Bassett, *Nature* **1957**, *179*, 752.
- [52] G. A. Bassett, J. W. Menter, D. W. Pashley, *Proc. R. Soc., Ser. A* **1958**, *246*, 345.
- [53] F. R. N. Nabarro (Ed.), *Dislocation in Solids*, North-Holland, Amsterdam, **1979**.
- [54] R. Siems, P. Delavignette, S. Amelinckx, *Phys. Stat. Solidi* **1962**, *2*, 421, 636.
- [55] R. Siems, P. Delavignette, S. Amelinckx, *Z. Phys.* **1961**, *165*, 502.
- [56] S. Amelinckx, P. Delavignette, *J. Appl. Phys.* **1962**, *33*, 1458.
- [57] S. Amelinckx, in *Dislocation in Solids* (Ed.: F. R. N. Nabarro), North-Holland, Amsterdam, **1979**, Vol. 2, p. 345.
- [58] J. Van Landuyt, G. Remaut, S. Amelinckx, *Phys. Stat. Solidi* **1970**, *41*, 271.
- [59] E. Ruedl, P. Delavignette, S. Amelinckx, *J. Nucl. Mater.* **1962**, *6*, 46.
- [60] A. Berghezan, A. Fourdeux, S. Amelinckx, *Acta Metall.* **1960**, *9*, 464.
- [61] S. Amelinckx, in *Chemistry and Physics of Carbon*, Vol. 1 (Ed.: P. L. Walker), Marcel Dekker, New York, **1965**, p. 1.
- [62] P. B. Hirsch, J. Silcox, R. Smallmann, K. Westmacott, *Phil. Mag.* **1958**, *3*, 897.
- [63] P. B. Hirsch, R. W. Horne, M. J. Whelan, *Phil. Mag.* **1956**, *1*, 667.
- [64] P. B. Hirsch, J. Silcox, in *Growth and Perfection of Crystals* (Eds.: R. H. Doremus et al.), Wiley, New York, **1958**, p. 262.
- [65] H. Hashimoto, A. Howie, M. J. Whelan, *Proc. R. Soc., Ser. A* **1962**, *269*, 80.
- [66] A. Art, R. Gevers, S. Amelinckx, *Phys. Stat. Solidi* **1963**, *3*, 967.
- [67] M. J. Whelan, P. B. Hirsch, *Phil. Mag.* **1957**, *2*, 1121; *ibid.*, 1303.
- [68] G. Van Tendeloo, S. Amelinckx, *Acta Crystallogr., Sect. A* **1974**, *30*, 431. R. Portier, D. Gracia, *Journal de Physique* **1982**, *43*, C4-17.
- [69] C. Manolikas, J. Van Landuyt, S. Amelinckx, *Phys. Stat. Solidi A* **1979**, *53*, 327.
- [70] C. Manolikas, J. Van Landuyt, S. Amelinckx, *Phys. Stat. Solidi A* **1980**, *60*, 607.
- [71] C. Manolikas, S. Amelinckx, *Phys. Stat. Solidi* **1980**, *61*, 179.
- [72] H. W. Zandbergen, G. Van Tendeloo, T. Okabe, S. Amelinckx, *Phys. Stat. Solidi A* **1987**, *103*, 45.
- [73] P. Delavignette, S. Amelinckx, *Appl. Phys. Lett.* **1963**, *2*, 236.
- [74] S. Amelinckx, *Chimica Scripta* **1978**, **1979**, *14*, 197.
- [75] G. Van Tendeloo, S. Amelinckx, *Phys. Stat. Solidi A* **1974**, *22*, 621.
- [76] D. Broddin, G. Van Tendeloo, J. Van Landuyt, S. Amelinckx, *Phil. Mag.* **1989**, *59*, 47.
- [77] D. Broddin, G. Van Tendeloo, S. Amelinckx, *J. Phys. Condens. Matter* **1990**, *2*, 3459.
- [78] G. Van Tendeloo, S. Ghose, S. Amelinckx, *Phys. Chem. Minerals* **1989**, *16*, 311.
- [79] J. Vanhellefont, S. Amelinckx, *J. Appl. Phys. B* **1987**, *61*, 2176.
- [80] J. Mahy, J. Van Landuyt, S. Amelinckx, Y. Uchida, K. D. Bronsema, S. Smaalen, *Phys. Rev. Lett.* **1985**, *55*, 1181.
- [81] J. Mahy, J. Van Landuyt, S. Amelinckx, K. D. Bronsema, S. Smaalen, *J. Phys. C: Solid State Physics* **1986**, *C19*, 5049.
- [82] H. Hashimoto, A. Howie, M. J. Whelan, *Proc. R. Soc.* **1962**, *A267*, 80.
- [83] H. Hashimoto, M. J. Whelan, *J. Phys. Soc. Jpn* **1963**, *18*, 1706.
- [84] R. Gevers, A. Art, S. Amelinckx, *Phys. Stat. Solidi* **1963**, *3*, 1563.
- [85] A. Art, R. Gevers, S. Amelinckx, *Phys. Stat. Solidi* **1963**, *3*, 967.

## 1.1.2 High-Resolution Electron Microscopy

### 1.1.2.1 Introduction

Among all the techniques used to obtain structural information of materials, high-resolution electron microscopy (HREM) has the great advantage that it yields information about the bulk structure, projected along the direction of electron incidence at a resolution comparable to the interatomic distances. This enables the study of complicated and artificial structures, crystal defects, precipitates and so forth, down to the atomic level. It is our belief that in the future to come, when materials research will gradually evolve into materials design and microstructures into nanostructures, the disposal of a high-resolution electron microscope yielding quantitative structural information at the subangstrom level will be indispensable.

By combining the information from different projections one can in principle obtain three-dimensional structural information. For instance, in the case of silicon, a resolution of 0.1 nm allows observation of the crystal along seven different zone axes.

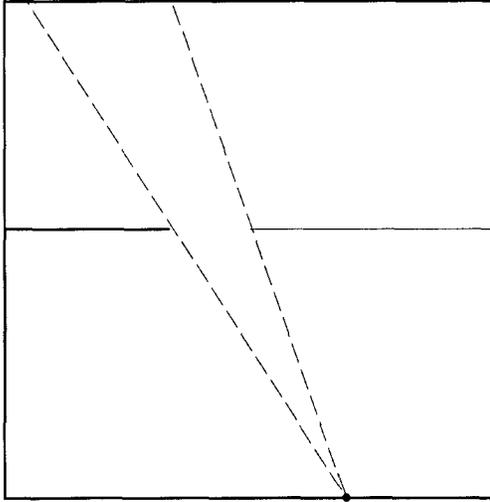
Recent technological improvements allow one to obtain a resolution of about 0.1 nm, so that it will become possible to 'see' the individual atomic columns of which matter is constituted. However, the

potential power of the technique is still severely limited by the problem of quantitative interpretation of the images. For instance, the use of computer simulation images requires much a priori knowledge, which makes HREM very dependent on other techniques. The situation can be compared with the early days of X-ray diffraction when the power of the technique became obvious but the applications were limited by lack of interpretation. Recent developments make it possible to retrieve the object structure directly from the electron micrographs. In this way HREM becomes much more powerful and independent. We will discuss future prospects in this direction and we will also show that the ultimate resolution is finally limited by the object itself.

### 1.1.2.2 Principles of Image Formation

#### Basic Principles

Let us first consider, as an illustrative example, the simplest imaging device: the camera obscura. This is a black box with a pinhole (Fig. 1). The object, represented by the function  $f(x)$ , is projected through the



**Figure 1.** The simplest imaging device: the projection box.

pinhole (aperture) to the image (for simplicity we take the function and the camera to be one-dimensional). A point  $x$  in the image gets contributions from an area of the object as seen through the window. This can be mathematically expressed as

$$f_{im}(x) = \int a(x' - x)f(x') dx' \quad (1)$$

where  $a(x)$  is the aperture function, which is equal to one in the aperture and zero elsewhere. Equation (1) is the definition of a convolution product

$$f_{im} = a(x) * f(x) \quad (2)$$

If we take the Fourier transform of Eq. (2), we obtain

$$\tilde{f}_{im}(g) = \tilde{a}(g)\tilde{f}(g) \quad (3)$$

i.e. the Fourier transform of a convolution product is the product of the Fourier transforms.  $g$  is the spatial frequency. This is illustrated on the right-hand side of Fig. 2.  $\tilde{a}(g)$  is usually called the (modulation) transfer function or MFT of the imaging device. It is shown schematically in Fig. 3.

Every imaging device can be characterized by its transfer function (band filter), which describes the magnitude with which a spatial frequency  $g$  is transferred through the device. The noise,  $N$ , is also indicated.

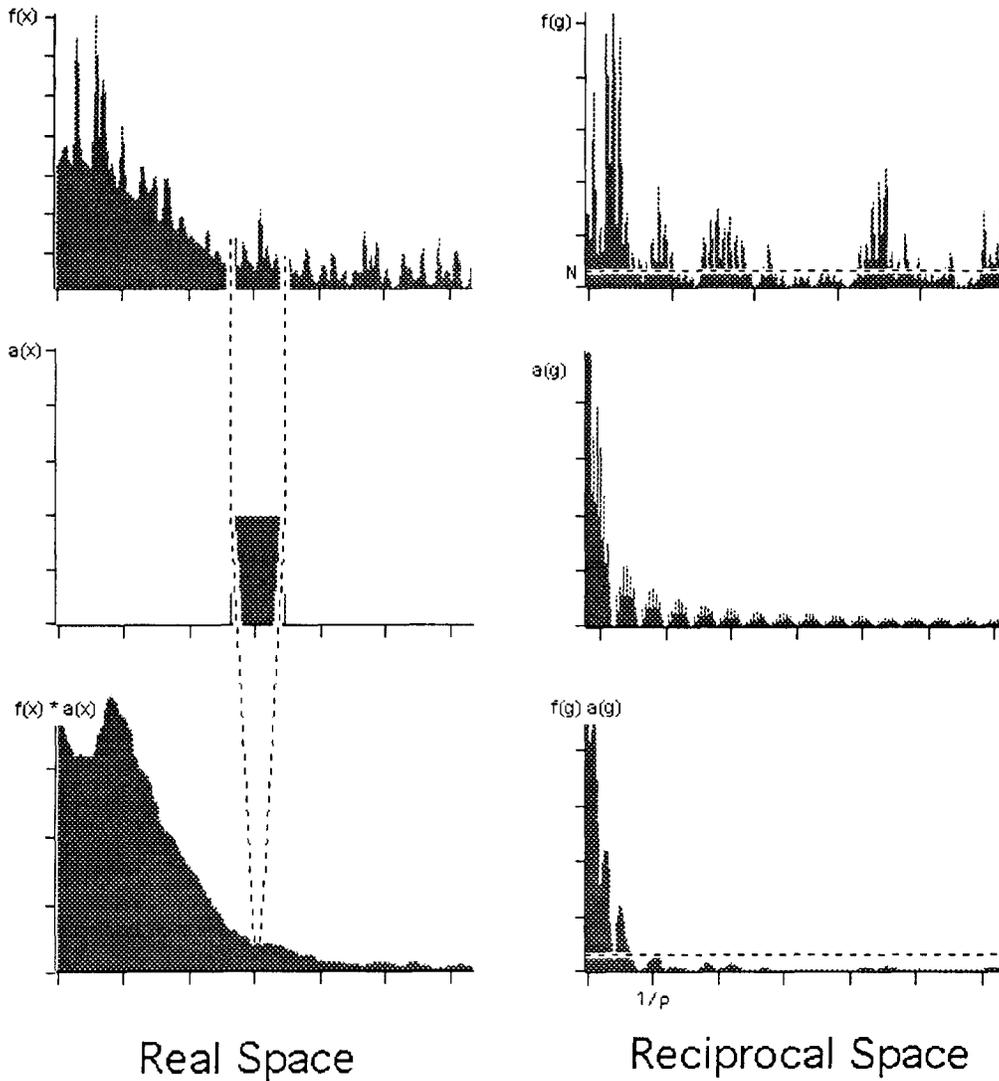
## Resolution

Usually, the resolution of the instrument  $\rho$  is defined from the cut-off  $1/\rho$  between the signal and noise beyond which no spatial information is transferred. This is the type of resolution in the sense as defined by Rayleigh. The Fourier transform of the transfer function to real space is usually called the impulse response function (IRF). It is the generalization of the aperture function of the camera obscura. It is a sharply peaked function which represents the image of a point object. The width of the IRF is also related to the Rayleigh resolution. The sharper the IRF, the better the resolution. This is demonstrated in Fig. 4. If the transfer function is known, the original image can be restored up to the resolution  $\rho$  by dividing by  $\tilde{a}(g)$ . This is called image restoration or deblurring. If an imaging device consists of a series of subdevices, the total transfer function is the product of the transfer functions of the subdevices.

### 1.1.2.3 The Electron Microscope [1]

#### Transfer Function

The image formation in an electron microscope is a coherent process, so that the object function as well as the transfer



**Figure 2.** Mathematical formulation of the image formation in a projection box in real space (left) and reciprocal space (right).

functions are complex functions with an amplitude and a phase component. The transfer function now has a particular physical meaning.

The wavefunction  $\psi(\mathbf{R})$  at the exit face of the object can be considered as a planar source of spherical waves (Huyghens

principle) ( $\mathbf{R}$  is taken in the plane of the exit face). It thus acts as a diffraction grating. According to Fraunhofer's diffraction theory the complex amplitude of the diffracted wave in the direction given by the reciprocal vector  $\mathbf{g}$  (or spatial frequency) is given by the Fourier transform

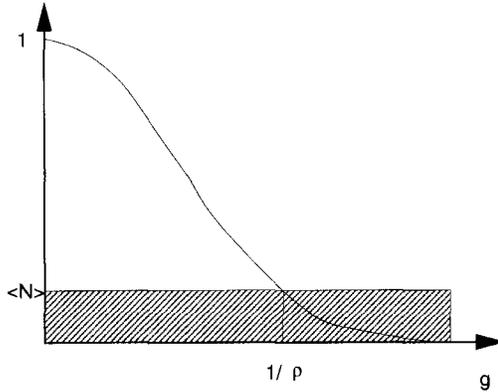


Figure 3. Transfer function.

of the object function, i.e.

$$\psi(\mathbf{g}) = F_{\mathbf{g}}\psi(\mathbf{R}) \quad (4)$$

Now the objective lens is placed behind the object. Hence it focuses each parallel diffracted beam into another point of the focal plane, whose position is given by the reciprocal vector  $\mathbf{g}$  characterizing the diffracted beam. Thus the back focal plane of the objective lens contains the Fourier transform of the object. In an electron microscope it is possible to change the lens settings so as to image the focal plane directly. In this way one can see the diffraction pattern, given by  $|\psi(\mathbf{g})|^2$ . If the object is periodic, such as a crystal, the diffraction pattern will consist of sharp spots. A continuous object will give rise to a continuous diffraction pattern. In the second stage of the imaging process, the back focal plane acts, in its turn, as a set of Huyghens sources of spherical waves which interfere, through a system of lenses, in the image plane. This stage in the imaging process is described by an inverse Fourier transform which reconstructs the object function  $\psi(\mathbf{R})$  (usually enlarged) in the image plane (Fig. 5). The intensity in the image plane is then given by  $|\psi(\mathbf{R})|^2$ .

In practice, by inserting an aperture in the focal plane of the objective lens, it is possible to obtain an image in which only selected beams contribute (see Sect. 1.1.2.5). By selecting only one beam, all interference disappears and the image shows only contrast proportional to the local intensity of that particular diffraction beam. This mode of imaging is called diffraction contrast (see Sect. 1.1.1 of this Chapter).

The electron microscope however is not a perfect imaging device. On passing through the objective lens, each electron beam  $\mathbf{g}$  (i.e. the spatial frequency  $g$ ) undergoes a phase shift and an amplitude reduction (damping). Hence the transfer function takes the form

$$T(\mathbf{g}) = A(\mathbf{g}) \exp[-i\chi(\mathbf{g})]D(\alpha, \Delta, \mathbf{g}) \quad (5)$$

$\chi(\mathbf{g})$  is the phase shift.  $A(\mathbf{g})$  describes the effect of the beam selecting aperture and the damping. For a derivation of the exact expression, see Appendix A. The wavefunction at the image plane is now given by

$$\Phi(\mathbf{R}) = F_{\mathbf{R}}^{-1}T(\mathbf{g})\psi(\mathbf{g}) \quad (6)$$

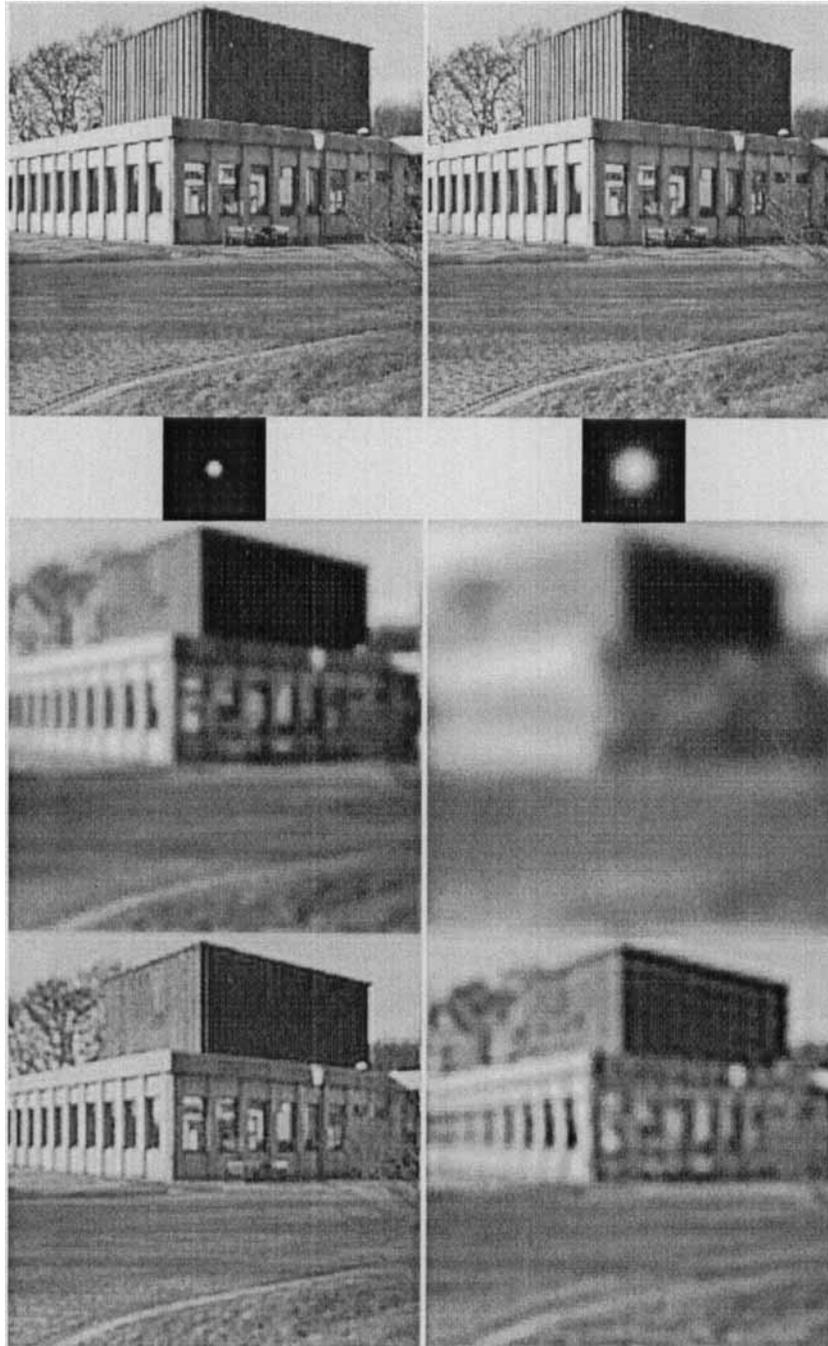
and the image intensity by

$$I(\mathbf{R}) = |\Phi(\mathbf{R})|^2 \quad (7)$$

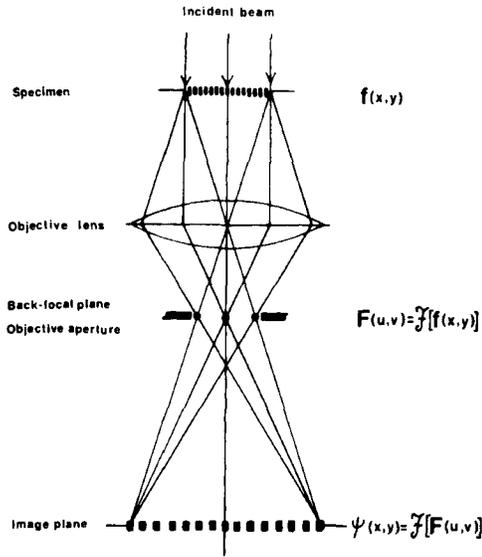
Equation (5) is called the coherent approximation; it is valid for thin objects. For thicker objects one uses the concept of transmission cross-coefficient (TCC). Here, the Fourier components of the image intensity are given by

$$\begin{aligned} I(\mathbf{g}) &= F_{\mathbf{g}}(I(\mathbf{R})) \\ &= \int \psi(\mathbf{g} + \mathbf{g}')\tau(\mathbf{g} + \mathbf{g}', \mathbf{g}')\psi^*(\mathbf{g}') d\mathbf{g}' \end{aligned} \quad (8)$$

with  $\tau$  the TCC which describes how the beams  $\mathbf{g}'$  and  $\mathbf{g} + \mathbf{g}'$  are coupled in the



**Figure 4.** Image formation through a pinhole (impulse response function). From top to bottom: original image (Centre for Electron Microscopy, Antwerp); Gaussian pinholes with different sizes; blurred image, deblurred image.



**Figure 5.** Schematic representation of the image formation by the objective lens in a transmission electron microscope. The corresponding mathematical operations are indicated (see text with  $\mathbf{R} = (x, y)$  and  $\mathbf{G} = (u, v)$ ).

Fourier component  $I(\mathbf{g})$ . For a derivation of Eq. (8) see Appendix A.

The total image formation process is shown schematically in Fig. 6.

### Impulse Response Function

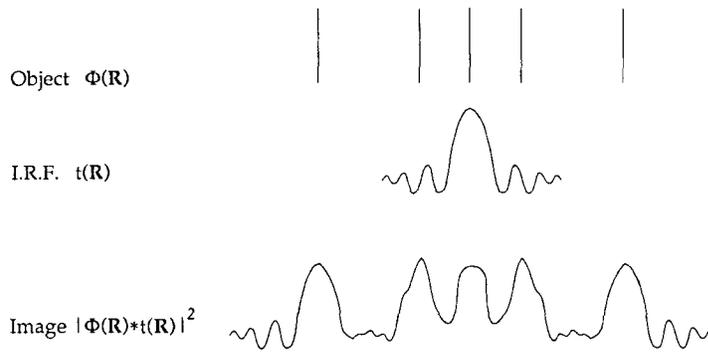
As explained in Sec. 1.1.2.2 the image transfer can also be described in real

space as a convolution product

$$\Phi(\mathbf{R}) = \psi(\mathbf{R}) * t(\mathbf{R}) \tag{9}$$

where  $\psi(\mathbf{R})$  is the object wave in real space and  $t(\mathbf{R})$  is the Fourier transform of the transfer function. For a hypothetical ideal pointlike object,  $\psi(\mathbf{R})$  would be a delta function or 'impulse' so that  $\Phi(\mathbf{R}) = t(\mathbf{R})$ , that is, the microscope would reveal  $t(\mathbf{R})$ , which is therefore called the impulse response function. This is the equivalent of the window function in Fig. 2. If the transfer function was constant (i.e., perfectly flat) over the whole spatial frequency range, the impulse response would be a delta function so that  $\Phi(\mathbf{R}) = \psi(\mathbf{R})$ , that is, the wavefunction in the image plane represents exactly the wavefunction of the object. In a sense the microscope is perfect. However, in practice the transfer function cannot be made arbitrarily flat as is shown in Fig. 7. The impulse response function is still peaked as shown in Fig. 8.

Hence, as follows from Eq. (9), the object wavefunction  $\Phi(\mathbf{R})$  is then smeared out (blurred) over the width of the peak. This width can then be considered as a measure for the resolution in the sense as originally defined by Rayleigh. The width of this peak is the inverse of the width of the constant plateau of the transfer



**Figure 6.** Schematic representation of the whole image formation process in the electron microscope.

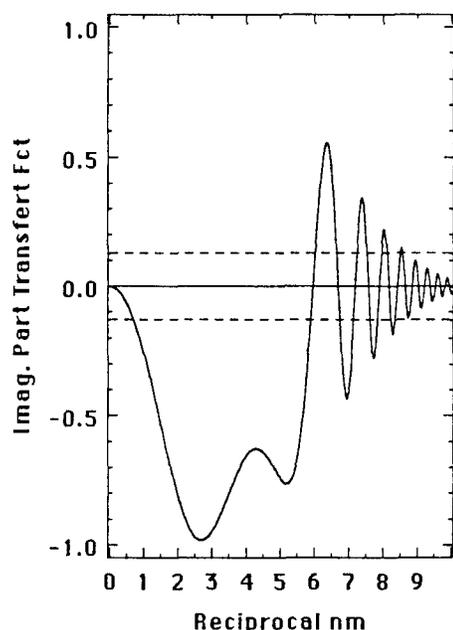


Figure 7. Transfer function (imaginary part) of a 300 keV electron microscope at optimum defocus.

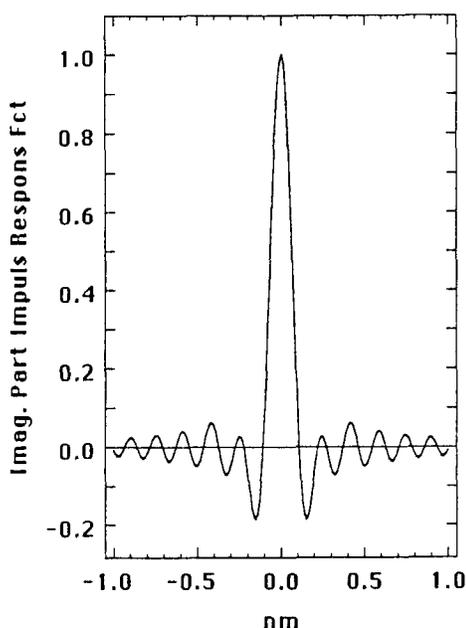


Figure 8. Impulse response function (imaginary part) corresponding to the transfer function of Fig. 6.

function in Fig. 7. From another point of view one can argue that if all the spatial frequencies have the same phase shift, the information is transferred forward and keeps a point-to-point relation to the object.

However, the information beyond this plateau is still contributing to the image but with a wrong phase. It is scattered outside the peak of the impulse response function and thus redistributed over a larger area in the image plane.

### Imaging at Optimum Focus: Phase Contrast Microscopy

In an ideal microscope, the image wavefunction would exactly represent the object wavefunction, and the image intensity for a pure phase object function would be

$$|\Phi(\mathbf{R})|^2 = |\psi(\mathbf{R})|^2 = |\exp[i\varphi(\mathbf{R})]|^2 = 1 \quad (10)$$

that is, the image would show no contrast. This can be compared with imaging a glass plate with variable thickness in an ideal light microscope.

Assuming a weak phase object (WPO) one has  $\varphi(\mathbf{R}) \ll 1$  so that

$$\psi(\mathbf{R}) \approx 1 + i\varphi(\mathbf{R}) \quad (11)$$

The constant term 1 contributes to the central beam (zeroth Fourier component) whereas the term  $i\varphi$  mainly contributes to the diffracted beams. If the phases of the diffracted beams can be shifted over  $\pi/2$  with respect to the central beam, the amplitudes of the diffracted beams are multiplied by  $\exp(i\pi/2) = i$ . Hence the image term  $i\varphi(\mathbf{R})$  becomes  $-\varphi(\mathbf{R})$ . It is as if the

object function had the form

$$\Phi(\mathbf{R}) = 1 - \varphi(\mathbf{R}) \approx \exp[-\varphi(\mathbf{R})]$$

that is, the phase object now acts as an amplitude object. The image intensity is then

$$|\Phi(\mathbf{R})|^2 \approx 1 - 2\varphi(\mathbf{R}) \quad (12)$$

which is a direct representation of the phase of the object. In light microscopy, this has been achieved by F. Zernike using a quarter-wavelength plate.

In electron microscopy, phase contrast imaging can be achieved by making the transfer function as constant as possible. From Eq. (37) it is clear that oscillations occur due to spherical aberration and defocus. However, the effect of spherical aberration which, in a sense, makes the objective lens too strong for the most inclined beams, can be compensated somewhat by slightly underfocusing the lens. The optimum defocus value (also called Scherzer\* defocus) for which the plateau width is maximal, is given by

$$\epsilon_s = -1.2(\lambda C_s)^{1/2} = -1.2 \text{ Sch} \quad (13)$$

with  $1 \text{ Sch} = (\lambda C_s)^{1/2}$  the Scherzer unit, with  $\lambda$  the electron wavelength and  $C_s$  the spherical aberration. Typical values are  $\lambda = 2 \text{ pm}$  (300 keV) and  $C_s = 1 \text{ mm}$ , yielding  $\epsilon_s = -54 \text{ nm}$ .

The transfer function for this situation is depicted in Fig. 7. The phase shift  $\chi(g)$  is nearly equal to  $-\pi/2$  for a large range of spatial coordinates  $g$ . The Scherzer plateau extends nearly to the first zero at a spatial frequency

$$g \approx 1.5 C_s^{-1/4} \lambda^{-3/4} \quad (14)$$

\* Named after Otto Scherzer who was the first to describe the image formation in the electron microscope in this way [1].

In this mode one can image directly the phase of the object. Now a thin material object acts as a phase object (see Appendix B) in which the phase is proportional to the electrostatic potential of the atoms projected along the viewing direction. Hence, if the object was very thin, optimum focus imaging would directly reveal atoms as dark areas, and empty spaces as light. The details are explained in Sec. 1.1.2.6. However, this argument only holds for distances which are well above the point resolution of the microscope. Furthermore, the thickness up to which an object can be considered as WPO is unrealistically small (e.g., 1 nm) and is rarely met in practice.

## Resolution of the Electron Microscope

In principle the characteristics of the electron microscope can be completely defined by its transfer function, that is, by the instrumental aberrations. However, a clear definition of resolution is not easily given for an electron microscope. For instance, for thick specimens, there is not necessarily a one-to-one correspondence between the projected structure of the object and the wavefunction at the exit face of the object, so that the image does not show a simple relationship.

If one wants to determine a 'resolution' number, this can only be meaningful for thin objects. Furthermore one has to distinguish between *point resolution* (or *structural resolution*), as the finest detail that can be interpreted in terms of the structure, and the *information limit*, which is the

finest detail that can be resolved by the instrument, irrespective of a possible interpretation.

As explained in Sec. 1.1.2.6, the electron microscope in the phase contrast mode at optimum focus directly reveals the projected potential, that is, the structure, of the object, provided the object is very thin. All spatial frequencies  $g$  with a nearly constant phase shift are transferred forward from object to image. Hence the point resolution can be obtained from the first zero of the transfer function (14) as

$$\rho_s = \frac{1}{g} \approx 0.65 C_s^{1/4} \lambda^{3/4} = 0.65 \text{ Gl} \quad (15)$$

$\text{Gl} = C_s^{1/4} \lambda^{3/4}$  is called the Glaser unit. The point resolution is also equal to the 'width' of the impulse response function. The information beyond  $\rho_s$  is transferred with a non-constant phase and, as a consequence, is redistributed over a larger image area.

The *information limit* can be defined as the finest detail that can be resolved by the instrument. It corresponds to the maximal diffracted beam angle that is still transmitted with appreciable intensity, that is, the transfer function of the microscope is a spatial band filter which cuts all information beyond the information limit. For a thin specimen, this limit is mainly determined by the envelope of chromatic aberration (temporal incoherence) and beam convergence (spatial incoherence) (see Appendix A). In principle, beam convergence can be reduced using a smaller illuminating aperture and a larger exposure time. If chromatic aberration is predominant, the damping envelope function is given by Eq. (40), from which the information limit can be

estimated as

$$\rho_l = \frac{1}{g} = \left( \frac{\pi \lambda \Delta}{2} \right)^{1/2} \quad (16)$$

The information limit is usually smaller than the point resolution. Typical values are  $\lambda = 2 \text{ pm}$  (300 keV),  $C_s = 1 \text{ mm}$ ,  $\Delta = 5 \text{ nm}$ ,  $\rho_s = 0.2 \text{ nm}$ , and  $\rho_l = 0.13 \text{ nm}$ . The point resolution can be improved by reducing  $C_s$  and reducing  $\lambda$  (i.e., increasing the voltage). The information limit can be improved by improving the coherence and reducing  $\lambda$ .

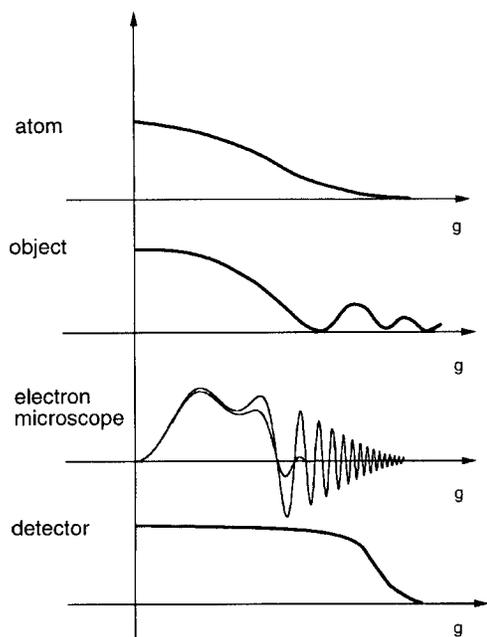
### 1.1.2.4 Resolution Limits [2]

The electron microscope can be considered as an information channel that transfers information from the object to the observer. The channel can be considered as consisting of four subchannels, each of which has a transfer function that limits the resolution.

These are schematically represented in Fig. 9. We will discuss each of the four different subchannels in more detail.

### The Atom

The electron interacts with the electrostatic potential of the atom. A typical electrostatic potential of an atom has a Gaussian-like shape. The average width is of the order of 0.05–0.1 nm. The Fourier transform of the potential is also Gaussian shaped, and is called the scattering factor. The atom itself can be considered as the ultimate probe with which the object can be sampled. So the final resolution in the

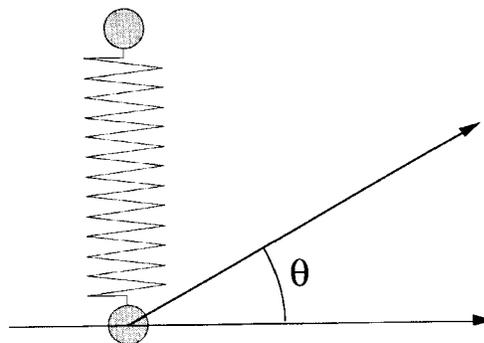


**Figure 9.** Transfer functions of the subchannels in the electron microscope.

sense of Rayleigh will always be of the order of 0.1 nm. Furthermore, the thermal atom motion causes the scattering factor to decrease by the so-called Debye–Waller factor, which even reduces the resolution.

### The Object

In practice the imaging in the electron microscope is two-dimensional whereas the object itself is three-dimensional. In a sense the three-dimensional structure is forced into a two-dimensional imaging process. This also has an effect on the resolution. The conservation of energy requires that the wave vector of the electrons should be on the Ewald sphere. Hence the scattered intensity is reduced with increasing distance from the origin.



**Figure 10.** Schematic representation of electron scattering at a configuration of two atoms connected by a spring.

The ultimate resolution limit imposed by the object can be estimated as follows. By studying an object one is in fact interested in knowing the configuration of atoms. However, in order to get information with sufficient detail, the incident electron has to approach the atoms closely so that it can transmit energy, change the configuration, and destroy the coherence. This poses ultimate intrinsic limits to the resolution. We will demonstrate it in a very simplified way.

Consider two neighboring atoms connected by a spring (Fig. 10). The energy transferred by the incident electron is given by the formula

$$\frac{\Delta E}{E} = \frac{4m}{M} \sin^2\left(\frac{\theta}{2}\right) \quad (17)$$

with  $M$  the mass of the target atom. The diffraction angle is related to the resolution by the Bragg formula

$$\rho = \frac{\lambda}{\sin \theta} \quad (18)$$

The energy of the spring considered as a harmonic oscillator is given by

$$\Delta E = \frac{h}{2\pi} \sqrt{\frac{k}{M}} \quad (19)$$

with

$$k = \frac{dK}{dr} \quad (20)$$

$K$  is the Coulomb force and

$$E = \frac{h^2 k^2}{2m} \quad (21)$$

Finally one obtains, using Eqs. (17)–(21), for the ultimate resolution

$$\rho = \frac{\epsilon R_B}{Z^{1/4}} \quad (22)$$

with  $R_B$  the Bohr radius,  $Z$  the atomic number, and  $\epsilon$  a configuration constant approximately equal to 1.5. Hence, the electrons that carry out information beyond  $\rho$  will transfer energy to the system and are lost for coherent imaging. The ultimate resolution for coherent imaging is thus of the order of 0.1 nm for light atoms and 0.05 nm for heavy atoms. Note also that this limit is independent of the type of the imaging particles.

## The Electron Microscope

The ultimate resolution is determined by the subchannel with the worst resolution. Thus far, the weakest part has been the electron microscope itself.

As shown in Sec. 1.1.2.3 the point resolution  $\rho_s$  of the electron microscope can be improved by reducing the spherical aberration constant  $C_s$  and/or by increasing the voltage. However, since  $C_s$  depends mainly on the pole piece dimension and the magnetic materials used, not much improvement can be expected. Hence, at present, all high-resolution electron microscopes yield comparable values for  $C_s$  for comparable designs. Furthermore, the effect of

$C_s$  on the resolution is rather limited. In the far future, a major improvement can be expected by using superconducting lenses or microlenses.

Another way of increasing the resolution is by correcting the third-order spherical aberration by means of a system of quadrupole, hexapole, and/or octupole lenses [3].

Increasing the voltage is another way of increasing the resolution. However, apart from the increased cost, increasing the voltage might increase the displacive radiation damage of the object, especially for light atoms.

A promising way of increasing the resolution is by restoring the information that is present between  $\rho_s$  and  $\rho_l$  and that is still present in the image, albeit with the wrong phase. For this purpose, holography combined with image processing will be indispensable. In that case, the resolution will be determined by  $\rho_l$ .  $\rho_l$  can be improved drastically by using a field emission gun (FEG) which reduces the spatial as well as the temporal incoherence. Taking all these considerations into account, an ultimate resolution of the electron microscope of 0.1 nm is within reach even at intermediate voltages.

Since resolution is a trade-off between signal and noise, some improvement can still be expected by reducing the noise. Specimen noise (inelastic scattering) can be reduced by energy filtering and the recording noise can be improved by using slowscan CCD cameras. However, if we assume that the total transfer function is Gaussian, an improvement in the signal-to-noise ratio from 20 to 100 only results in an improvement of the resolution with 25%. Ultimately, one cannot avoid the Poisson noise due to the counting statistics (quantum noise or shot noise) which is

limited by the radiation dose. One cannot arbitrarily increase the dose without increasing the recording time and/or damaging the object. Hence, it can be expected that the ultimate resolution attainable with this technique will not exceed 0.05 nm. At that resolution level it will not be the microscope that will limit resolution but rather the object itself. It is worth noticing that when testing the resolution capabilities of an electron microscope close to 0.1 nm, it is hard to find objects with which this resolution can be demonstrated. Conversely, only suitable objects can benefit from a further resolution improvement. Hence we believe that more effort has to be done to optimize the specimen preparation technique.

### The Recording Device

If one wants to record information up to the information limit, special requirements have to be met by the recording device. First in real space one needs at least three pixels per resolution unit. In reciprocal space the finest oscillations of the transfer functions (Fig. 7) have to be sampled sufficiently, which means about eight pixels per oscillation. The oscillations can be reduced over the whole spatial frequency range by using an optimum focus called the Lichte focus (see Sec. 1.9 of this Chapter). Totally this gives a minimum number of pixels given by the formula

$$N > 30 \left( \frac{\rho_S}{\rho_I} \right)^4 \quad (23)$$

One now has  $N > 500$  for  $\rho_S = 0.2$  nm and  $\rho_I = 0.1$  nm. This requires a very good recording system. If one uses off-axis

holography (see Sec. 1.9 of this Chapter) then this number has to be multiplied by three. The newest generation of slow-scan CCD cameras with fiber-coupled scintillators might be a solution to this problem. Furthermore, when cooled, these cameras are able to detect nearly all single electrons.

### 'Seeing' Atoms: A New Situation

It is clear that with the recent technological improvements the resolution approaches the physical limits, which is comparable to the size of the atoms (strictly speaking one does not 'see' individual atoms but rather individual columns in projection).

If atoms, the building blocks of matter, can be seen individually, one can introduce a new concept of resolution, based on the information theory of Shannon. In 1965, Gabor argued that one could not use information theory for microscopy since this theory was devised to establish communication between a sender and a receiver which communicated with a common vocabulary of messages. However, in the microscope one communicates between an unknown object and an observer. Since the object is unknown one cannot define a set of messages. However, when it is possible to see the atoms, they can be considered as messages in the original sense of Shannon and hence we can use communication theory. Consider for example the simplified case depicted in Fig. 11.  $d$  is a one-dimensional periodicity containing atoms. For each atom the position has to be determined. The right-hand side shows the transfer function. It is sampled in distances  $1/d$ . The total number of degrees of freedom is equal to the total number of

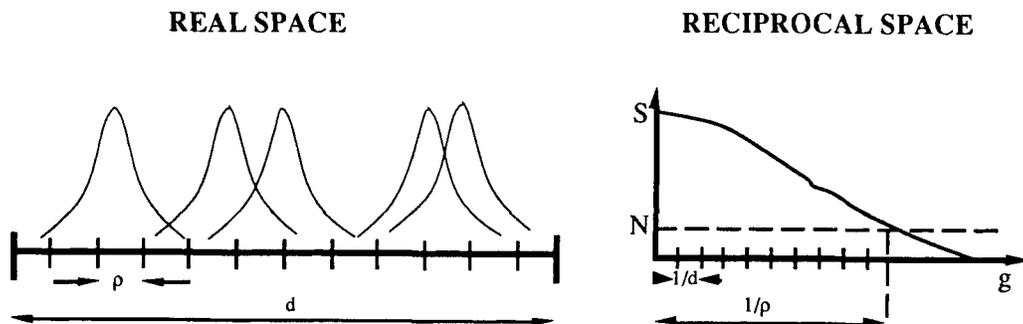


Figure 11. Schematic representation of the information content.

spatial frequencies in the transfer function, given by

$$n = \frac{1/\rho}{1/d} = \frac{d}{\rho} \quad (24)$$

from which

$$\frac{n}{d} = \frac{1}{\rho}$$

Hence each unit of length  $\rho$  carries one degree of freedom. The average information of each degree expressed in bits can be calculated as follows: calling  $S$  the signal and  $N$  the noise (Fig. 11), if we consider the noise as the smallest piece of information. The number of bits needed to describe the signal is then

$$\log_2 \left( \frac{N + S}{N} \right) = \log_2 \left( 1 + \frac{S}{N} \right)$$

Strictly speaking this means that a microscope has an information capacity of one number per resolution unit. Each number has on the average  $\log_2(1 + S/N)$  bits of information (precision). In two dimensions one can use the same arguments (Fig. 12). For a resolution  $\rho$  the number of degrees of freedom per unit area is equal to about three per  $\rho^2$ , which in fact means for an electron microscope with a resolution of 0.1 nm that one can determine approximately the coordinates of one atom per  $0.01 \text{ nm}^2$ . If one investigates a crystal viewed along a simple zone axis and the number of columns in projection is less than one per  $\rho^2$ , the complete structure in projection can be solved. If, however, the number of atom columns in projection exceeds the capacity of the microscope, as when one investigates a complex zone

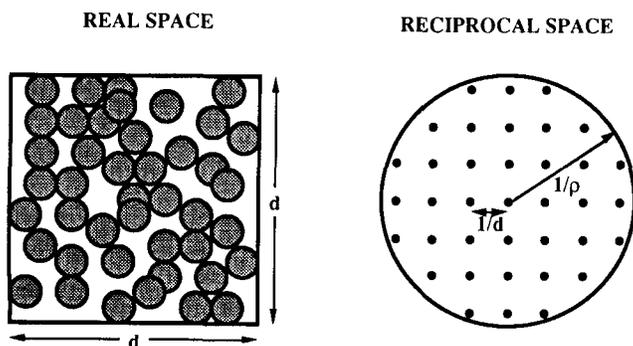


Figure 12. Information content in two-dimensional images.

axis or when the object is disordered, then the capacity of the microscope is exceeded and the information channel is blocked. In such a situation a structure cannot be solved completely and only limited information can be obtained from the images. It is thus more appropriate to regard the resolution capacity of the electron microscope as the number of data points per unit area. The situation is comparable to the case of X-ray diffraction. If the number of beams exceeds the number of atom coordinates to be determined (usually by a factor of two to three) then the structure can in principle be solved completely.

### 1.1.2.5 High-Resolution Electron Microscopy in Practice

#### The Electron Microscope

The resolution capability of an electron microscope is related to the size of the objective pole piece gap and thus also to the tilting possibilities. In cases where tilting is not necessary, as often is the case for the study of biological specimens, one can use a high-resolution stage with a very small bore, but without tilting possibilities. Generally, for materials research, a goniometer stage is indispensable. This can either be a double tilt or a rotation-tilt stage, although the former is easier to use. Moreover, one has to make a compromise between the maximal tilt angle and the resolution. The higher the resolution, the smaller the tilt angle and the more patience is required from the operator to find a suitable orientation. However, at

high voltages ( $\geq 300$  keV) it is possible to combine high resolution with large tilt possibilities. One can usually choose between a top entry and a side entry goniometer stage. The top entry stage normally yields a higher resolution and better mechanical stability (less thermal specimen drift). The side entry stage has much better possibilities for tilting, heating, strain experiments, and so on, usually at the expense of resolution. However, recently designed side entry stages seem to compete with top entry stages with respect to resolution and stability.

For high-resolution work many advantages are offered if the electron microscope is equipped with an image-intensifying camera mounted below the photographic plate holder. This facilitates the observation of the high-resolution images and the adjustment of the microscope. A suitable rule of thumb for making good use of the camera is that the resolution of the camera, divided by the primary magnification of the electron microscope, should be much smaller than the resolution of the electron microscope. For instance for 300 keV electrons the resolution of the best CCD camera is of the order of  $50\ \mu\text{m}$  (mainly determined by the resolution of the transmission fluorescent screen). Therefore, a primary magnification of  $10^6$  is required in order that  $50\ \mu\text{m} \times 10^{-6} = 0.05\ \text{nm}$  should be much less than the resolution of the electron microscope.

Before starting high-resolution work, it is necessary to determine the most important optical parameters of the instrument for later use in image simulation and reconstruction. For very high resolution (below  $0.2\ \text{nm}$ ) the standard correction procedure for the aberrations is not sufficient, and methods have been

developed for automatic alignment. Furthermore, the spherical aberration constant depends on the position of the specimen in the objective lens. For this purpose the microscope should be equipped with a CCD camera and on-line image-processing capabilities. A series of images is taken from a thin amorphous film at different focus values. Each of the images is then digitized and numerically Fourier transformed so as to obtain the so-called diffractogram. Originally, diffractograms were obtained by diffraction with a laser beam in an optical bench, but due to the disposal of good CCD cameras and fast computers this method is not much in use anymore. The diffractogram represents the contrast transfer function corresponding with the particular focus. Accurate values for the spherical aberration constant  $C_s$  and for the defocus values can be derived by curve fitting. In this way the defocus steps can be calibrated. The point resolution can be measured from the first zero of the transfer function and from the damping function. The instrumental resolution, the chromatic aberration, and the beam convergence can also be determined.

A review of aberrations and their determination is given by Saxton [4]. Recently, it has been discovered [5] that third-order astigmatism, which is usually believed to be unimportant, may become very important for very high resolution. The effect of this astigmatism is not visible in the diffractogram. Recently it has been shown how it can be measured and corrected [4]. It is also interesting to note that when holographic reconstruction methods are used (see Sec. 1.1.2.6) the microscope has not to be tuned perfectly since the residual aberrations can be corrected for afterwards. One of the major problems encountered

in obtaining high-resolution images is the mechanical stability. During the exposure (of the order of seconds) the drift of the specimen has to be much less than the resolution of the instrument. This can be overcome somewhat by decreasing the primary magnification so that the intensity is increased and the recording time reduced. The resolution and sensitivity of the photographic plate or the camera then are the limiting factors.

## The Specimen

Specimens for HREM are prepared using the same techniques as for usual transmission electron microscopy, that is, ion beam milling, chemical and electrolytical thinning, cleavage, crushing, and so forth. The only requirements are that the specimen should be sufficiently thin, that is, less than about 10 nm. Furthermore, the specimen should be sufficiently clean and free of contamination. For instance, when a specimen such as an alloy is thinned using ion milling, it can be recommended to finish with a chemical polishing or plasma cleaning, in order to remove the amorphous layer.

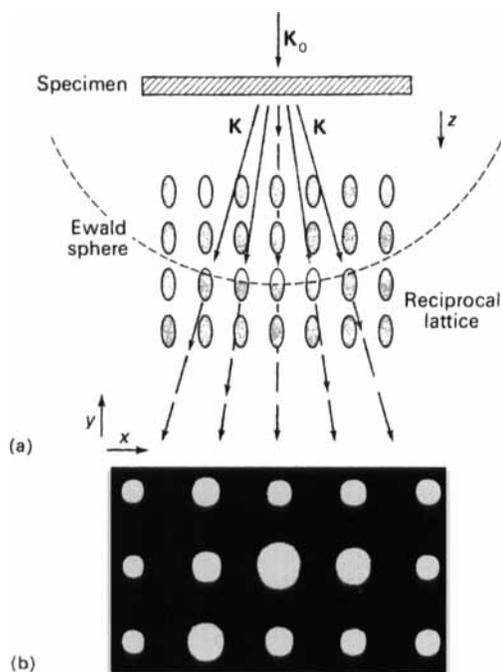
For brittle specimens, the crushing method is usually applied. Here the specimens are ground in an agate mortar after which the powder is suspended in methanol and some drops of this suspension are deposited on a thin perforated supporting film.

In cases where cleavage along a required plane does not occur easily, it might be worthwhile to grind the specimens at liquid nitrogen temperature where the material is more brittle and the cleavage more isotropic. In cases where the required cleavage is almost impossible, for example

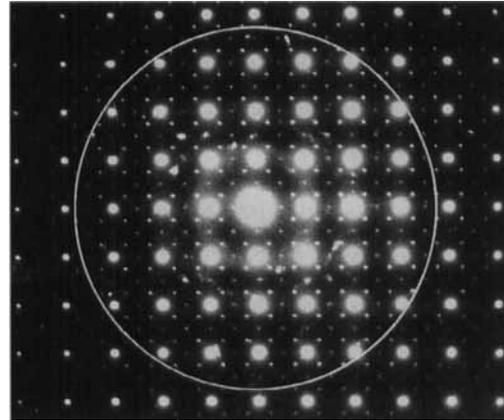
when hair-grown crystals have to be cleaved perpendicular to the hair filaments, one can try to embed the crystals in an epoxy resin and to cut them by microtomy.

For details of specimen preparation we refer the reader to Robards and Wilson [6].

Crystalline specimens with a unit cell with two large and one small lattice parameter are most ideal for HREM. In that case the reciprocal lattice consists of dense planes (Laue zones) which are largely separated. Such crystals can be oriented with their short axis parallel to the incident beam so that the nearly flat Ewald sphere touches the Laue zone through the origin (Fig. 13). Hence a large number of diffracted beams are excited simultaneously and maximal information is present in the image. This is shown in Fig. 14. In this



**Figure 13.** Formation of the diffraction pattern. The simultaneously excited electron beams can be used for image formation.

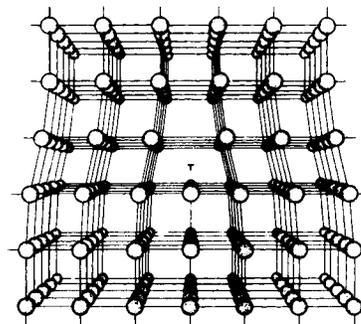


**Figure 14.** Typical diffraction pattern as used for structure imaging. The aperture selecting the contributing beams is also indicated.

situation, the electrons propagate parallel to a zone axis, that is, parallel to the atom rows. Only in this way can a possible interpretation of the images in terms of the projected structure be meaningful. The same argument holds also for crystals with defects.

As an example, Fig. 15 shows a schematized model of a dislocation viewed along the atomic columns.

It is also possible, using an aperture placed in the focal plane of the objective lens, to select a particular set of beams so that the images contain specific



**Figure 15.** Structure model of a crystal containing a dislocation viewed along the atomic columns.

information. If the central beam is not included, these images are called dark field images. Examples are given in Sec. 1.1.2.7.

When the periodicity along the incident beam direction is not small, the Laue zones are not distant from each other so that more Laue zones are excited at the same time. In this case, the diffraction pattern also shows the intersection of the Ewald sphere with the higher-order Laue zones (HOLZs), which consist of concentric circles. In that case the three lattice parameters can be deduced from one diffraction pattern.

When dealing with a crushed specimen suspended on holey support film, one has to search for thin, wedge-shaped flakes, partly covering a hole of the film. Then the specimen has to be oriented with the required zone axis parallel to the incident beam. However, when the tilt possibilities are not large (e.g.,  $10^\circ$ ) the majority of the specimens cannot be properly oriented, except when the cleavage is preferentially perpendicular to the zone axis. It can then be tried to search for a suitable crystal particle in diffraction mode. A crystal with a nearly good orientation can easily be recognized since in that case the intersection of the Ewald sphere with the zeroth Laue zone is an arc of a circle through the origin, which, with some experience, can easily be recognized. The smaller the radius of curvature of the arc, the closer to the exact zone orientation. After finding a suitably thin part with the proper orientation, one has to adjust the focus. When the specimen is very thin, the zero focus corresponds to minimal contrast (see Sec. 1.1.2.6). By underfocusing, a bright fringe appears at the border of the specimen. Maximal contrast appears close to the optimum defocus. On going through focus, the contrast reverses.

In practice, since the focus is not exactly known, especially in the case of thicker specimens, one has to take a series of images at gradually different focus settings, recorded approximately around the optimum defocus. This is called a through-focus series (Fig. 16). When a CCD camera is available, the number of recorded through-focus images can usually be reduced. When dealing with a specimen that is unstable in the electron beam, the specimen can be completely destroyed within a period of seconds. Here a minimal exposure technique has to be used. In some cases, only one picture can be taken. Here also the availability of an image intensifier allows use of a much smaller beam intensity and an increase in the time of observation.

A parameter that is difficult to determine is the specimen thickness. In the case of a crystalline specimen with a small unit cell, an estimate of the thickness in thicker areas can sometimes be obtained from thickness contours in a two-beam case. Tilt experiments can also provide some information. However, in general the thickness of the foil in thin areas cannot be measured accurately.

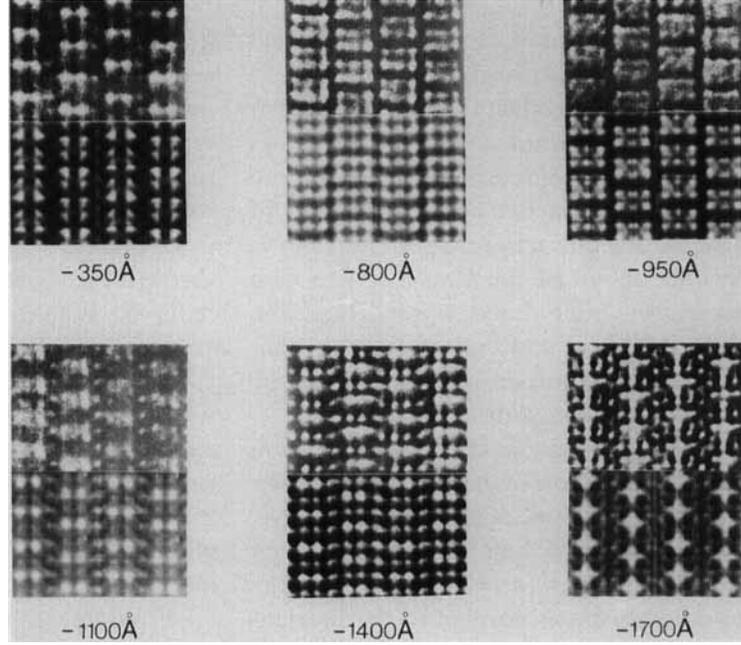
### 1.1.2.6 Interpretation of the Images

#### Intuitively Interpretable Images

##### Images of Thin Objects

###### (i) *Optimum focus*

When the phase change in the phase object is very small (weak phase object) we can use expression (56) derived in Appendix B,



**Figure 16.** Comparison of experimental images (top row) (courtesy of S. Iijima) and computer-simulated images (bottom row) for  $\text{Ti}_2\text{Nb}_{10}\text{O}_{29}$  as a function of defocus.

which can be expanded up to the first power as

$$\psi(\mathbf{R}) \simeq 1 + i\sigma V_p(\mathbf{R}) - \mu(\mathbf{R}) \quad (25)$$

$V_p(\mathbf{R})$  is the projected potential of the object,  $\mu(\mathbf{R})$  is the absorption function. The Fourier transform, yielding the amplitude in the back focal plane now becomes

$$\psi(\mathbf{g}) = \delta(\mathbf{g}) + i\sigma V_p(\mathbf{g}) - M(\mathbf{g}) \quad (26)$$

with the Dirac function  $\delta(\mathbf{g})$  representing the transmitted beam.

From Equation (5) the image amplitude (without aperture) is now

$$\begin{aligned} \phi(\mathbf{R}) &= \mathcal{F}_R \psi(\mathbf{g}) e^{-i\chi(\mathbf{g})} \\ &= \mathcal{F}_R [\delta(\mathbf{g}) + \sigma V_p(\mathbf{g}) \sin \chi(\mathbf{g}) \\ &\quad - M(\mathbf{g}) \cos \chi(\mathbf{g}) \\ &\quad + i\sigma V_p(\mathbf{g}) \cos \chi(\mathbf{g}) \\ &\quad + iM(\mathbf{g}) \sin \chi(\mathbf{g})] \end{aligned} \quad (27)$$

At the optimum defocus the transfer function shows a nearly flat region for which  $\sin \chi(\mathbf{g}) \approx -1$  and  $\cos \chi(\mathbf{g}) \approx 0$  for all contributing beams.

Now Eq. (27) becomes

$$\begin{aligned} \phi(\mathbf{R}) &\approx \mathcal{F}_R [\delta(\mathbf{g}) - \sigma V_p(\mathbf{g}) - iM(\mathbf{g})] \\ &= 1 - \sigma V_p(\mathbf{R}) - i\mu(\mathbf{R}) \end{aligned}$$

and the image intensity to the first order is

$$I(\mathbf{R}) \approx 1 - 2\sigma V_p(\mathbf{R}) \quad (28)$$

At the optimum focus, the electron microscope acts as a phase contrast microscope so that the image contrast of a thin object is proportional to its electrostatic potential  $V_p(\mathbf{R})$  projected along the direction of incidence. This theory can be generalized for larger phase changes [7].

#### (ii) *Small Defocus*

In the absence of apertures and spherical aberration, the image amplitude is

given by

$$\phi(\mathbf{R}) = \mathcal{F}_{\mathbf{R}} \exp[-i\pi\epsilon\lambda(\mathbf{g}^2)]\psi(\mathbf{g}) \quad (29)$$

which can be expanded for small defocus values as

$$\begin{aligned} \phi(\mathbf{R}) &= \mathcal{F}_{\mathbf{R}}[1 - i\pi\epsilon\lambda(\mathbf{g}^2)]\psi(\mathbf{g}) \\ &= \psi(\mathbf{R}) - i\pi\epsilon\lambda \mathcal{F}_{\mathbf{R}}(\mathbf{g}^2)\psi(\mathbf{g}) \end{aligned}$$

or explicitly

$$\begin{aligned} \phi(\mathbf{R}) &= \psi(\mathbf{R}) - i\pi\epsilon\lambda \\ &\quad \times \int \exp[2\pi i(\mathbf{g} \cdot \mathbf{R})](\mathbf{g}^2)\psi(\mathbf{g}) \, d\mathbf{g} \end{aligned}$$

This expression can be transformed elegantly as

$$\begin{aligned} \phi(\mathbf{R}) &= \psi(\mathbf{R}) + \frac{i\pi\epsilon\lambda}{4\pi^2} (\nabla^2) \\ &\quad \times \int \exp[2\pi i(\mathbf{g} \cdot \mathbf{R})]\psi(\mathbf{g}) \, d\mathbf{g} \\ &= \left(1 + \frac{i\epsilon\lambda}{4\pi} (\nabla^2)\right) \psi(\mathbf{R}) \end{aligned}$$

which for a phase object, where  $\psi(\mathbf{R}) = \exp[i\sigma V(\mathbf{R})]$ , becomes

$$\begin{aligned} \phi(\mathbf{R}) &= \left(1 - \frac{\epsilon\lambda\sigma\Delta V_p(x,y)}{4\pi} \right. \\ &\quad \left. - \frac{i\epsilon\lambda\sigma^2}{4\pi} (\nabla V(\mathbf{R}))^2\right) \\ &\quad \times \exp[i\sigma V_p(\mathbf{R})] \end{aligned}$$

So that the image density  $|\phi(\mathbf{R})|^2$  is given to the first order by

$$I(\mathbf{R}) = 1 - \frac{\epsilon\lambda\sigma\Delta V_p(\mathbf{R})}{2\pi} \quad (30)$$

with  $\Delta$  the Laplacean operator. Since the projected  $V_p(\mathbf{R})$  is of electrostatic origin, it obeys the Poisson equation,

$$\Delta V_p(\mathbf{R}) = -4\pi\rho_p(\mathbf{R})$$

so that finally

$$I(\mathbf{R}) = 1 + 2\epsilon\lambda\sigma\rho_p(\mathbf{R}) \quad (31)$$

The image contrast is thus proportional to the projected charge density; it disappears in the Gaussian image plane ( $\epsilon = 0$ ) and reverses with the sign of the defocus. This effect can be used to find approximately the zero focus. Expression (31) is the projected charge density (PCD) approximation [8]. Its validity is restricted to resolutions for which the spherical aberration is not too important. It is interesting to note that exactly at zero focus the image of the phase object shows no contrast. The only contrast observed stems from absorption, by which electrons are removed from the image formation. The use of a small objective aperture which eliminates electrons from the imaging process also gives rise to a kind of absorption contrast. This is the case for biological samples, 'colored' with heavy atoms.

For phase objects at small underfocus, neglecting the spherical aberration, the image contrast shows the projected charge density. For WPOs at optimum defocus, including the spherical aberration, the image contrast represents the projected potential. Although both results are in fact approximations, it can be concluded that between zero focus and optimum defocus, the details of the high-resolution images of very thin objects which are larger than the point resolution can be interpreted intuitively in terms of the projected specimen structure. For instance, groups of atoms are imaged with black contrast, and holes are imaged with white contrast. Examples are given in Sec. 1.1.2.7. However, these cases are rarely met in practice.

### Interpretation Using Channeling

The wavefunction in the image plane can be written as the convolution product of the wavefunction at the exit face of the crystal with the impulse response function  $t(\mathbf{R})$  of the electron microscope, Eq. (9). When we now use the expression derived for the wavefunction using the channeling theory (Eq. (89)), we obtain

$$\psi(\mathbf{R}) = 1 + \sum_i \left[ \exp\left(\frac{-iE_i z}{E \lambda}\right) - 1 \right] \times C_i \phi_i(\mathbf{R} - \mathbf{R}_i) * t(\mathbf{R}) \quad (32)$$

The summation runs over the columns  $i$  of the object, parallel to the electron beam with position  $\mathbf{R}_i$ . Each column contributes with a localized wavefunction  $\phi_i$  multiplied with a factor which varies with depth. The periodicity is inversely proportional to  $E_i$  which is related to the mass density of the column.

If the microscope is operated close to optimum focus and in axial mode, the impulse response function is sharply peaked (see Fig. 8). If also the localized wavefunction  $\phi_i$  is highly peaked and if the distance between the columns is larger than the width of the impulse response function  $t(\mathbf{R})$ , the overlap between convolution products  $\phi_i * t(\mathbf{R})$  of adjacent sites can be assumed to be small so that the image intensity is

$$I(\mathbf{R}) = |\psi(\mathbf{R})|^2 \approx \sum_i 4C_i^2 \sin^2\left(\frac{E_i z}{2E\lambda}\right) \times |\phi_i(\mathbf{R} - \mathbf{R}_i) * t(\mathbf{R})|^2 \quad (33)$$

Each column is thus imaged separately. The contrast of a particular column varies periodically with thickness. The

periodicity can be different for different types of columns. It is interesting to note that the functions  $\phi_i$  as well as  $t(\mathbf{R})$  are symmetrical around the origin, provided the objective aperture is centered around the optical axis. Hence, the image of a column is rotationally symmetric around the position  $\mathbf{R}_i$  of the columns. The intensity at  $\mathbf{R}_i$  is a maximum or a minimum. The positions of the columns can thus be determined from the positions of the intensity extrema.

When the resolution of the microscope is insufficient to discriminate the individual columns, or the focus is not close to optimum, the overlap between the convolution products of adjacent columns cannot be avoided and the interpretation of the contrast is not straightforward. In that case image simulation or object wave reconstruction is required.

### Building Block Structures

It often happens that a family of crystal structures exists, all members of which consist of a stacking of simple building blocks but with a different stacking sequence. This is for instance the case in mixed-layer compounds, polytypes, periodical twins, and substitutional columnar binary alloys. In a broader sense also periodic interfaces such as antiphase boundaries and crystallographic shear planes can be considered as building block systems.

In order to characterize the stacking sequence of unknown members of the family, one can search for a member with a known stacking sequence. Often, this will be a simple member, the structure of which has been determined by X-ray diffraction.

From this system, high-resolution micrographs are taken. The image characteristics of each type of building block in these images are then called the imaging code: since the high-energy electrons are scattered forward, it can be expected that the image of a building block will not be severely affected by its surroundings, provided the crystal is not too thick and the focus is close to the optimum focus. Now the building blocks of the unknown members are imaged with the same code, provided the experimental conditions (thickness, focus) are the same, so that the stacking sequence can be read directly from the high-resolution micrographs. In general, the technique of the imaging code is applicable up to crystal thicknesses of a few tens of nanometers, and also to other imaging modes.

Examples are given in Sec. 1.1.2.7.

A particular type of building block structure is the substitutional binary alloys with a column structure. In a substitutional binary alloy, the two types of atoms occupy positions on a regular lattice, usually face-centered cubic (f.c.c.). Since the lattice, as well as the types of the atoms and the average composition, is known, the problem of structure determination is then reduced to a binary problem of determining which atom is located at which lattice site.

Particularly interesting are the alloys in which columns are found parallel to a given direction and which consist of atoms of the same type. Examples are the gold–manganese system and other f.c.c. alloys [9]. If viewed along the column direction, which is usually  $[001]_{\text{f.c.c.}}$ , the high-resolution images contain sufficient information to determine unambiguously the type and position of the individual

columns. Even if the resolution of the microscope is insufficient to resolve the individual lattice positions, which have a separation of about 0.2 nm, it is possible to reveal the minority columns only, which is sufficient to resolve the complete structure. This can be done using dark field superlattice imaging, which is explained in the next section.

## Selective Beam Imaging

In some cases it is possible to reveal only specific information by selecting the appropriate beams for contributing to the image function. For this one uses an aperture in the focal plane of the objective lens. An illustrative example is the imaging of binary f.c.c. alloys with a column structure as described above.

Let us consider a binary system consisting of columns of A atoms, which are the majority atoms, and columns of B atoms. We call the fraction of columns respectively  $m_A$  and  $m_B$ ;  $m_A + m_B = 1$  with  $m_A > m_B$ .

Using the channeling theory (see Appendix D), it can then be shown that, when the structure is imaged without the f.c.c., reflections and overlap between images of adjacent columns can be neglected, the image intensity is given by

$$|\psi(\mathbf{R})|^2 \approx \sum_i \sigma_i^2 |\Delta\phi(\mathbf{R} - \mathbf{R}_i) * t(\mathbf{R})|^2 \quad (34)$$

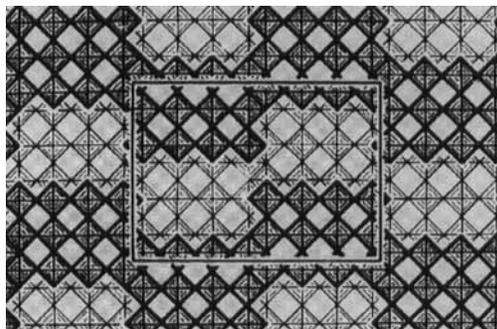
where  $\Delta\phi = \phi_B - \phi_A$  with  $\phi_A$  (respectively  $\phi_B$ ) the wavefunctions of the columns A (respectively B), and  $\sigma_i$  the Flynn occupation parameters defined as  $\sigma_i = m_A$  for a B column and  $\sigma_i = -m_B$  for an A column. The interpretation of Eq. (34) is now as

follows. The atom columns are imaged as bright peaks, the shape of the peaks being given by  $|\Delta\phi(\mathbf{R}) * t(\mathbf{R})|^2$  and the height of the peaks being proportional to  $m_A^2$  for the B-atom columns and  $m_B^2$  for the A-atom columns.

As a consequence, the minority atoms B are imaged as brighter dots than the majority atoms A, the ratio of brightness being equal to  $(m_A/m_B)^2$ . When this ratio is large (e.g.,  $>10$ ), the minority atoms will be visible as bright dots on a dark background. An example of this high-resolution dark field imaging applied to the  $\text{Au}_4\text{Mn}$  alloy is given in Sec. 1.1.2.7. A nice example of selective beam imaging is the quantitem technique developed by Ourmadz and co-workers in which they use reflections that are sensitive to the composition of atom columns so as to obtain images from which the composition can be deduced by pattern recognition techniques [10]. A survey of selective imaging methods and applications is given by Amelinckx et al. [11].

## Image Simulation

When no obvious imaging code is available, the interpretation of high-resolution images often becomes a precarious problem since, especially at very high resolution, the image contrast can vary drastically with focus. As a typical and historical example, high-resolution images obtained by Iijima for the complex oxide  $\text{Ti}_2\text{Nb}_{10}\text{O}_{25}$  with a point resolution of approximately 0.35 nm (100 keV) are shown in Fig. 16 (top row). The structure as reproduced in Fig. 17 consists of a



**Figure 17.** Schematic representation of the unit cell of  $\text{Ti}_2\text{Nb}_{10}\text{O}_{29}$  consisting of corner-sharing  $\text{NbO}_6$  octahedra with the titanium atoms in tetrahedral sites.

stacking of corner- or face-sharing  $\text{NbO}_6$  octahedrons with the titanium atoms in tetrahedral positions. High-resolution images are taken at different focus values, causing the contrast to change drastically. The best resemblance to the structure can be obtained near the optimum Scherzer defocus, which is  $-90$  nm in this particular case. However, the interpretation of such high-resolution images never appears to be trivial. The only way out remains in the comparison of the experimental images with those calculated for various trial structures. During the imaging process, the electrons undergo three distinct interactions, as shown schematically in Fig. 6. Each of these interactions is known and can be calculated by the computer. First, the electron scatters dynamically in the crystal. This interaction can be simulated using the multislice methods explained in Appendix C. However, as an input to the program one has to specify all the object parameters such as the unit cell, position and type of cell atoms, thermal atom factors (Debye–Waller factors), object orientation, and thickness. The result of this calculation yields the wavefunction at

the exit face of the crystal. In a second step, the formation of the image in the electron microscope is simulated using the expressions of Appendix A. Here all the instrumental parameters have to be specified. Finally, the electron intensity in the image plane is calculated by squaring the wavefunction, and is displayed as a half-tone image on a high-resolution screen or printer. Different commercial software packages exist for high-resolution image simulations. References are given in Ref. [12].

In practice, the image simulation requires a number of input parameters that are not accurately known such as specimen thickness and focus value. For this reason one often compares whole series of images simulated for various focus values and/or specimen thicknesses which are compared with experimental through-focus and/or through-thickness series, which makes the method more sensitive. As an example, the series of images simulated using the model of Fig. 17 for different focus values are shown in Fig. 16 (bottom row) and reveal a close resemblance to the experimental images. Other examples are given in Sec. 1.1.2.7.

A drawback of the present use of image simulation is that the image comparison is usually done visually and not quantitatively or in a recursive refinement. As a consequence, the technique can only be used if the number of plausible models is very limited. This makes HREM, despite its potential power for structure investigation, very dependent on other techniques. Direct methods, which extract the information from the images in a direct way, are much more promising.

## Quantitative Interpretation

In principle one is usually not interested in high-resolution images as such but rather in the structure of object under study. High-resolution images are then to be considered as data planes from which the structural information has to be extracted in a quantitative way. Ideally this should be done as follows. One has a model for the object and for the imaging process, including electron object interaction, microscope transfer, and image detection (see Fig. 6). The model contains parameters that have to be determined by the experiment. The parameters can be estimated from the fit between the theoretical images and the experimental images. The goodness of the fit is evaluated using a matching criterion (fitness function) such as likelihood, least squares, or the  $R$  factor (cf. X-ray crystallography). This fitness function can be calculated for each set of parameters. The optimal fit then yields the best estimates for the parameters of the model that can be derived from the experiment. In a sense one is searching for a maximum (or minimum) depending on the criterion) of the fitness function in the parameter space, the dimension of which is equal to the number of parameters. The object model that describes the interaction with the electrons consists of the assembly of the electrostatic potentials of the constituent atoms. Since for each atom type the electrostatic potential is known, the model parameters then reduce to atom numbers and coordinates, thermal atom factors, object thickness, and orientation (if inelastic scattering is neglected).

The imaging process is characterized by a small number of parameters such as

defocus, spherical aberration etc, that are not accurately known.

A major problem is now that the structural information of the object can be strongly delocalized by the image transfer in the electron microscope (see Figs. 6 and 8), so that the influence of the model parameters is completely scrambled in the high-resolution images. Due to this coupling, one has to refine all parameters simultaneously. As a consequence, the dimension of the parameter space is so high that one cannot use advanced optimization techniques such as genetic algorithms, simulating annealing, tabu search, and so forth, without the risk of ending in local maxima. Furthermore, for each new model trial one has to perform a tedious image calculation so that the procedure is very cumbersome. The problem is only manageable if the object is a crystal with a very small unit cell and hence a small number of object parameters [13], or if sufficient prior information is available to reduce the number of parameters drastically.

In X-ray crystallography, this problem can be solved by using direct methods which provide a pathway towards the global maximum. In HREM, this problem can be solved by deblurring the information, so as to unscramble the influence of the different object parameters of the image, and thus reduce the dimension of the parameter space. As described in Sec. 1.1.2.4 this can be achieved by high-voltage microscopy, by correcting the microscopic aberrations, or by holographic techniques.

Holographic techniques have the particular advantage that they first retrieve the whole wavefunction in the image plane, that is, amplitude and phase. In this way, they use all possible information. In the other two methods, one starts from the image intensity only, and the information

that is predominantly present in the phase is inevitably missed. Ideally, high-voltage microscopy or aberration correction is combined with holography, which leads to the advantage of holography but with a broader field of view. However, this has not yet been done in practice.

As explained above, the whole purpose is to unscramble the object information in the images, that is, to undo the image formation process, so as to uncouple the object parameters and to reduce the size of the parameter space. In this way it is possible to reach the global maximum (i.e., best fit) which leads to an approximate structure model.

This structure model then provides a starting point for a final refinement by fitting with the original images (i.e., in the high-dimensional parameter space) that is sufficiently close to the global maximum so as to guarantee fast convergence.

It has to be noticed that, in the case of perfect crystals, one can combine the information in the high-resolution images with that of the electron diffraction pattern, which in principle can also be recorded by the CCD camera. Since the diffraction patterns usually yield information up to higher spatial frequencies than the images, one can in this way extend the resolution to beyond 0.1 nm. Recently [14] very accurate structure refinements for unknown structures have been achieved with  $R$  factors below 5% (which is comparable to X-ray results). Here one starts with a first estimate of the structure as obtained from exit wave reconstruction (see further) which is then refined iteratively by using the electron diffraction data.

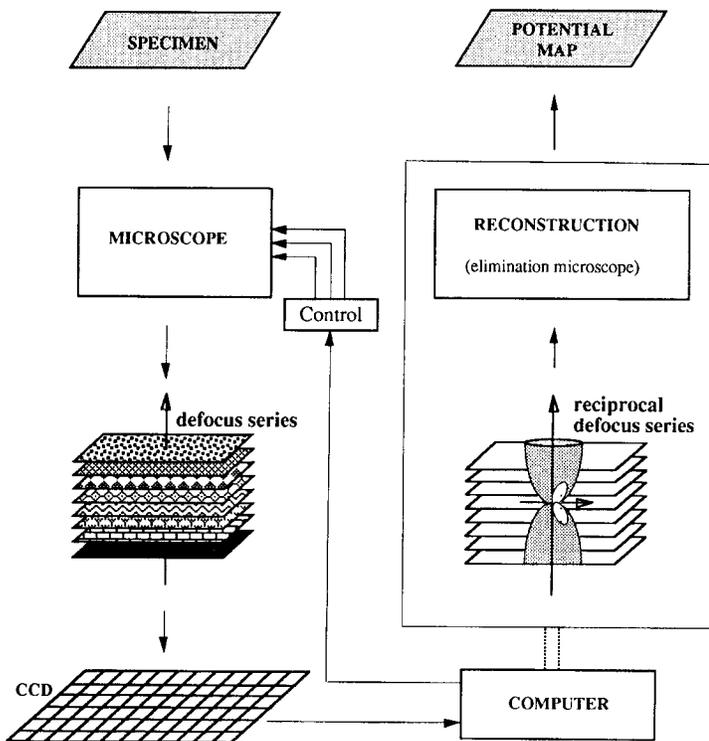
We will now focus attention mainly on the holographic reconstruction methods. Undoing the scrambling from object to

image consists of three stages. First, one has to reconstruct the wavefunction in the image plane (phase retrieval). Then one has to reconstruct the exit wave of the object. Finally, one has to 'invert' the scattering in the object so as to retrieve the object structure.

**Phase Retrieval**

The phase problem can be solved by holographic methods. Two methods exist for this purpose: off-axis holography and focus variation, which is a kind of in-line holography. In off-axis holography, the beam is split by an electrostatic biprism into a reference beam and a beam that traverses the object. Interference of both beams in the image plane then yields fringes, the positions of which yield the phase information.

In order to retrieve this information, a very high-resolution camera (CCD), a powerful image processor, and a field emission source to provide the necessary spatial coherence are needed. Details are given in Sec. 1.9 of this Chapter. In the focus variation method, the focus is used as a controllable parameter so as to yield focus values from which both amplitude and phase information can be extracted [15, 16]. Images are captured at very close focus values so as to collect all the information in the three-dimensional image space. Each image contains linear information and non-linear information. By Fourier transforming the whole three-dimensional image space, the linear information of all images is superimposed onto a sphere in reciprocal space, which can be considered as an Ewald sphere (Fig. 18). By filtering out this linear



**Figure 18.** Schematic representation of the phase retrieval procedure. The paraboloid which contains the linear information in reciprocal space is also shown.

information the phase can be retrieved (for details, see Appendix E).

The results indicate that focus variation is more accurate for high spatial frequencies whereas off-axis holography is more accurate for lower spatial frequencies but puts higher demands on the number of pixels in order to detect the high spatial frequencies.

The choice of focal values can also be optimized using a criterion that is currently used for experiment design [17]. It turns out that the choice of equidistant focus values is close to optimal.

### Exit Wave Reconstruction

The wavefunction at the exit face of the object can be calculated from the wavefunction in the image plane by applying the inverse phase transfer function of the microscope. This procedure is straightforward, provided the proper parameters describing the transfer function (such as the spherical aberration constant  $C_s$ ). As is clear from Fig. 7, the retrieval of information up to the information limit requires that the transfer function should be known with high accuracy. This requires an accuracy of less than 0.01 nm for  $C_s$  and 5 nm for  $\epsilon$ . Two remarks have to be made:

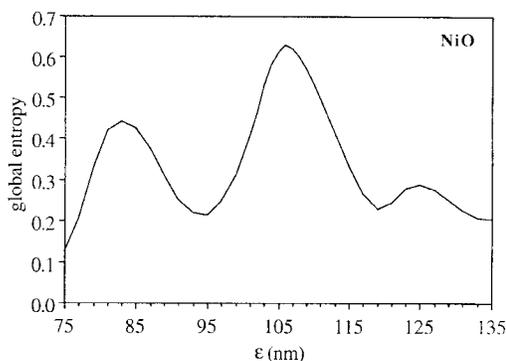
- (i) In principle the alignment of the microscope has not to be perfect provided the amount of misalignment is known so that it can be corrected for in the reconstruction procedure.
- (ii) An accurate measurement of  $C_s$  and  $\epsilon$  can only be performed if sufficient information is known about the object (e.g., a thin amorphous object can be

considered as a white noise object) from which the transfer function can be derived from the diffractogram.

We are thus faced with an intrinsic problem. An accurate determination of the instrumental parameters requires knowledge of the object. On the other hand, the most interesting objects under investigation are not fully known. Hence, the fine tuning of the residual aberrations has to be done on the object under study, based on some general assumptions that do not require a knowledge of the specimen structure such as the crystal potential is real, the structure is atomic, and so forth.

For instance, if the object is thin, the phase of the exit wave would show the projected potential which is sharply peaked at the atom columns. If the exit face is reconstructed with a slight residual defocus, these peaks would be blurred. Hence it can be expected that the peakiness of the phase is maximal at the proper defocus. The peakiness can be evaluated by means of an entropy using the Shannon formula. If the object is thicker, it can be expected from the channeling theory (see Eq. (89)) that the amplitude of  $\psi - 1$  is peaked, and thus also its entropy. Hence, a weighted entropy criterion may be used for fine tuning the residual defocus. This is shown in Fig. 19. Details are given by Tang et al. [18].

Figure 20 shows the exit wave of an object of  $\text{YBa}_2\text{Cu}_4\text{O}_8$  (high  $T_c$  superconductor), which was historically the first experimental result obtained with the focus variation method. The microscope used was a Philips CM20 ST equipped with field emission source and  $(1024)^2$  slow-scan CCD camera developed in the

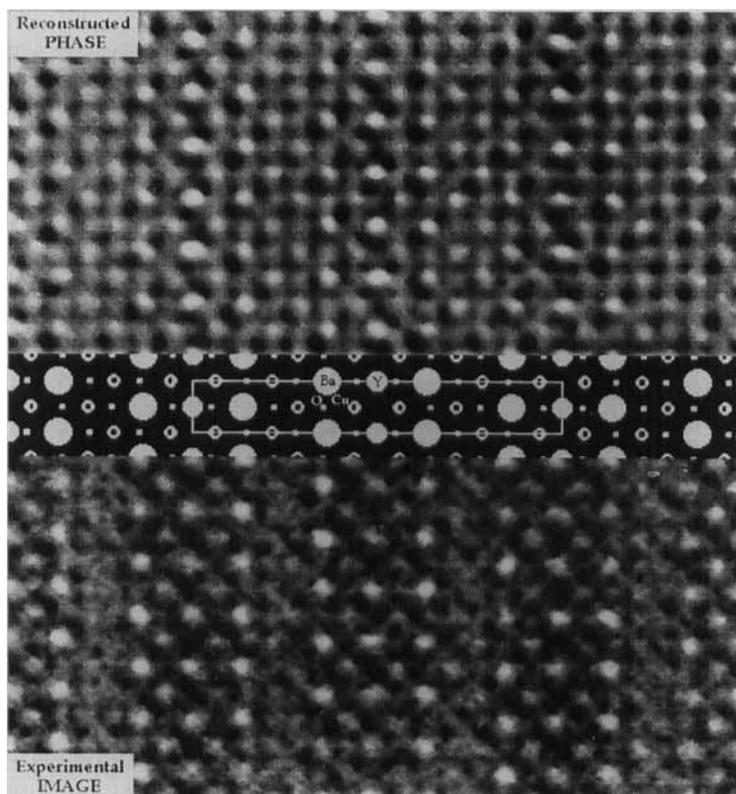


**Figure 19.** Global exit wave entropy as a function of residual defocus for  $\text{TiO}_2$ .

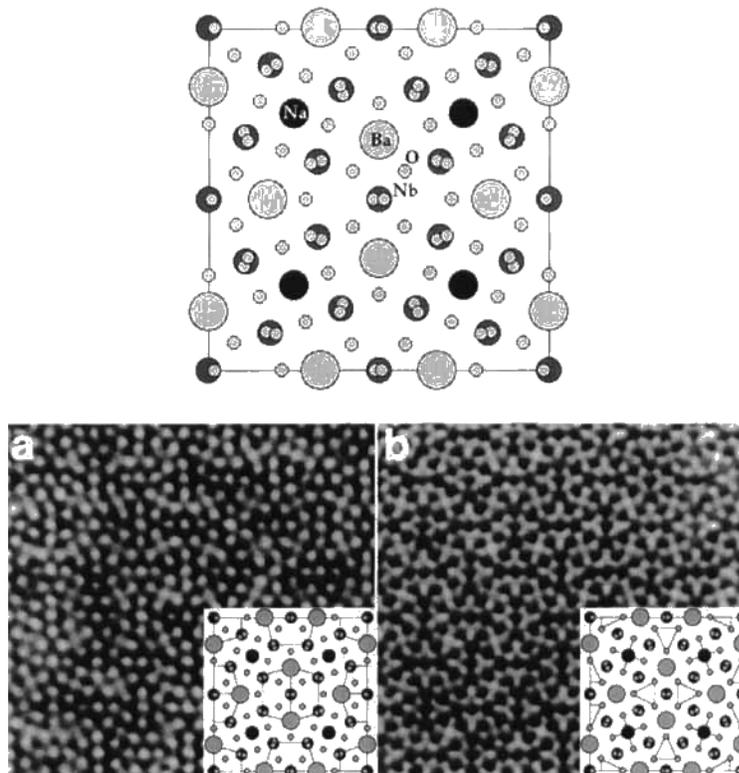
framework of a Brite-Euram project. In this case, the object is very thin so that the phase of the wavefunction directly reveals the projected potential of the

atom columns. The oxygen columns adjacent to the yttrium columns could just be observed, proving a resolution of 0.13 nm. However, when the object is thicker, the one-to-one correspondence between the wavefunction and the projected structure is not so straightforward due to the dynamic diffraction. This is shown in Fig. 21 for  $\text{Ba}_2\text{NaNb}_5\text{O}_{15}$  where the heavy columns (barium and niobium) are revealed in the amplitude and the bright columns (sodium and oxygen) in the phase. In this case it is necessary to invert in a sense the electron scattering in the object so as to retrieve the projected structure.

It should be noted that, once the exit wave is reconstructed, it is in principle possible to recalculate all the images of



**Figure 20.** Experimentally reconstructed exit wave for  $\text{YBa}_2\text{CuO}_8$ . Top, reconstructed phase; center, structure model; bottom, experimental image.



**Figure 21.** Experimentally reconstructed exit wave for  $\text{Ba}_2\text{NaNb}_5\text{O}_{15}$ . (a) Amplitude. (b) Phase. The structure model is shown at the top.

the Fourier series which perfectly fit in the experimental images within the noise level. Hence, the reconstructed exit wave contains all experimentally attainable object information. In practice, one thus will not have to store the original images but only the reconstructed wave.

### Structure Retrieval

The final step consists of retrieving the projected structure of the object from the wavefunction at the exit face. If the object is thin enough to act as a phase object, the phase is proportional to the electrostatic potential of the structure, projected along the beam direction so that the retrieval is

straightforward. If the object is thicker, the problem is much more complicated. In principle one can retrieve the projected structure of the object by an iterative refinement based on fitting the calculated and the experimental exit wave. As explained earlier, this is basically a search procedure in a parameter space. However, since the exit wave is much more locally related to the structure of the object than the original images the dimension of the parameter space is much smaller. Nevertheless it is possible to end in a local maximum [19]. However, an approximate structure can be obtained in a more direct way. If the object is a crystal viewed along a zone axis, the incident beam is parallel to the atom columns. It can be shown that in

such a case the electrons are trapped in the positive electrostatic potential of the atom columns, which then act as channels. This effect is known as electron channeling, and is explained in detail in Appendix D.

If the distance between the columns is not too small, a one-to-one correspondence between the wavefunction at the exit face and the column structure of the crystal is maintained. Within the columns, the electrons oscillate as a function of depth without, however, leaving the column. Hence the classical picture of electrons traversing the crystal as plane-like waves in the direction of the Bragg beams, which historically stems from X-ray diffraction, is in fact misleading. It is important to note that channeling is not a property of a crystal, but occurs even in an isolated column and is not much affected by the neighboring columns, provided the columns do not overlap. Hence the one-to-one relationship is still present in the case of defects such as translation interfaces or dislocations provided they are oriented with the atom columns parallel to the incident beam.

The basic result is that the wavefunction at the exit face of a column is expressed as Eq. (88):

$$\psi(\mathbf{R}, z) = 1 + \left[ \exp\left(-i\pi \frac{E}{E_0} kz\right) - 1 \right] \phi(\mathbf{R}) \quad (35)$$

This result holds for each isolated column. In a sense, the whole wavefunction is uniquely determined by the eigenstate  $\phi(\mathbf{R})$  of the Hamiltonian of the projected columns and its energy  $E$ , which are both functions of the 'density' of the column and the crystal thickness  $z$ . It is clear from Eq. (35) that the exit wave is peaked at the

centre of the column and varies periodically with depth. The periodicity is inversely related to the 'density' of the column. In this way the exit wave still retains a one-to-one correspondence with the projected structure. Furthermore, it is possible (see Eq. (82)) to parameterize the exit wave in terms of the atomic number  $Z$  and the interatomic distance  $d$  of the atoms constituting the column. This enables the projected structure of the object to be retrieved from matching with the exit wave.

In practice it is possible to retrieve the positions of the columns with high accuracy (0.01 nm) and to obtain a rough estimate of the density of the columns. Figure 22 shows a map of the projected potential  $\text{Ba}_2\text{NaNb}_5\text{O}_{15}$  retrieved from the exit wave of Fig. 21. Here all atoms are imaged as white dots with an intensity roughly proportional to the weight of the columns.

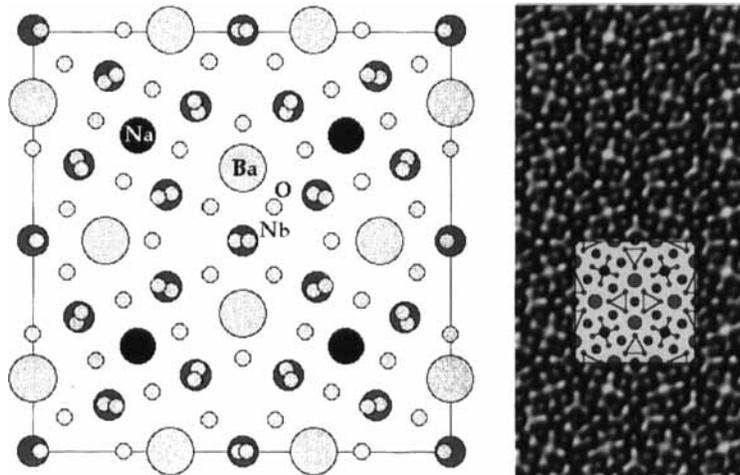
In principle, the three-dimensional structure can be retrieved by combining the information from different zone orientations. However, the number of 'visible' zone orientations is limited by the resolution of the electron microscope.

### Intrinsic Limitations

It should be noticed that HREM, even combined with quantitative reconstruction methods, has its intrinsic limitations.

Although the positions of the projected atom columns can be determined with high accuracy (0.01 nm), the technique is less sensitive for determining the mass density of the columns and to get information about the bonds between atoms. Besides,

**Figure 22.** Experimentally retrieved structure for  $\text{Ba}_2\text{NaNb}_5\text{O}_{15}$ .



due to the high speed of the electrons, they only sense a projected potential so that no information can be obtained about the distribution of this potential along the columns. Three-dimensional information can be obtained though, by investigating the same object along different zone axes. Furthermore, as shown above, for some object thicknesses, atom columns can become extinct so that they cannot be retrieved from the exit wave.

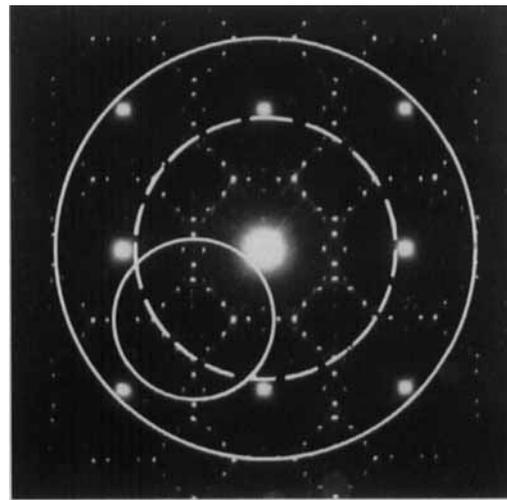
### 1.1.2.7 Case Studies

#### Alloys

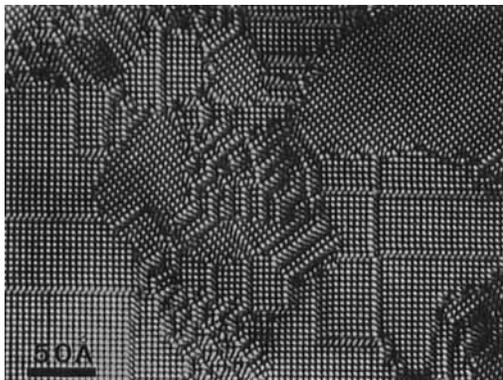
As shown in Sec. 1.1.2.6, substitutional alloys with a column structure are particularly interesting subjects for HREM studies. If the alloys are oriented with the columns parallel to the electron beam, one can directly visualize the minority columns by imaging without the reflections of the basic f.c.c. structure. The diffraction

pattern and the beam selecting aperture are shown in Fig. 23.

Figure 24 shows a historical experimental high-resolution micrograph of  $\text{Au}_4\text{Mn}$  viewed along the (100) zone using this dark field superlattice imaging mode



**Figure 23.** Diffraction pattern and beam selecting apertures for f.c.c.-based substitutional alloys. For high-resolution dark field imaging, the f.c.c. reflections are excluded and only the superstructure reflections of one reciprocal unit cell are included.



**Figure 24.** Dark field superlattice image of the alloy  $\text{Au}_4\text{Mn}$  viewed along the column direction  $[001]$ . The manganese columns are selectively imaged as bright dots [9].

using a 200 keV electron microscope with only 0.25 nm point resolution. The bright dots correspond to the configuration of manganese columns. Two orientation variants are present, as well as a number of antiphase boundaries.

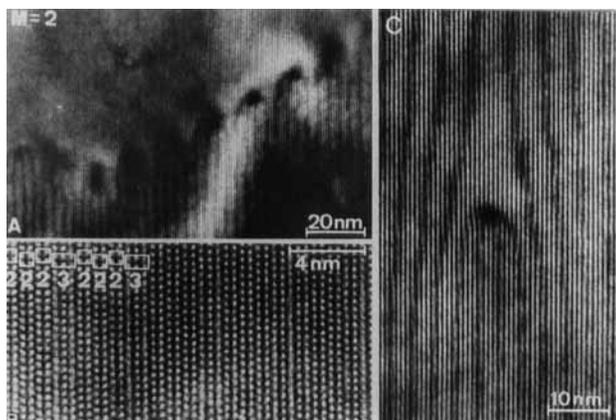
The imaging characteristics in which the manganese atoms are revealed as bright dots are preserved even close to the interfaces. This allows deduction of the displacement vectors and the orientations directly from the images.

Figure 25 shows a large period antiphase boundary superstructure 2223 (Fig. 25b). Discommensurations can be seen in a dark field line resolution image (Fig. 25c). Figure 26 shows different models of superstructures in the alloy  $\text{Nb}_5\text{Ga}_{13}$ . Figure 27 shows high-resolution images of  $\text{Nb}_5\text{Ga}_{13}$  viewed along  $[110]$ . From the high-resolution images it can be concluded that the actual structure is that corresponding to the model in Fig. 26a.

## Crystal Defects

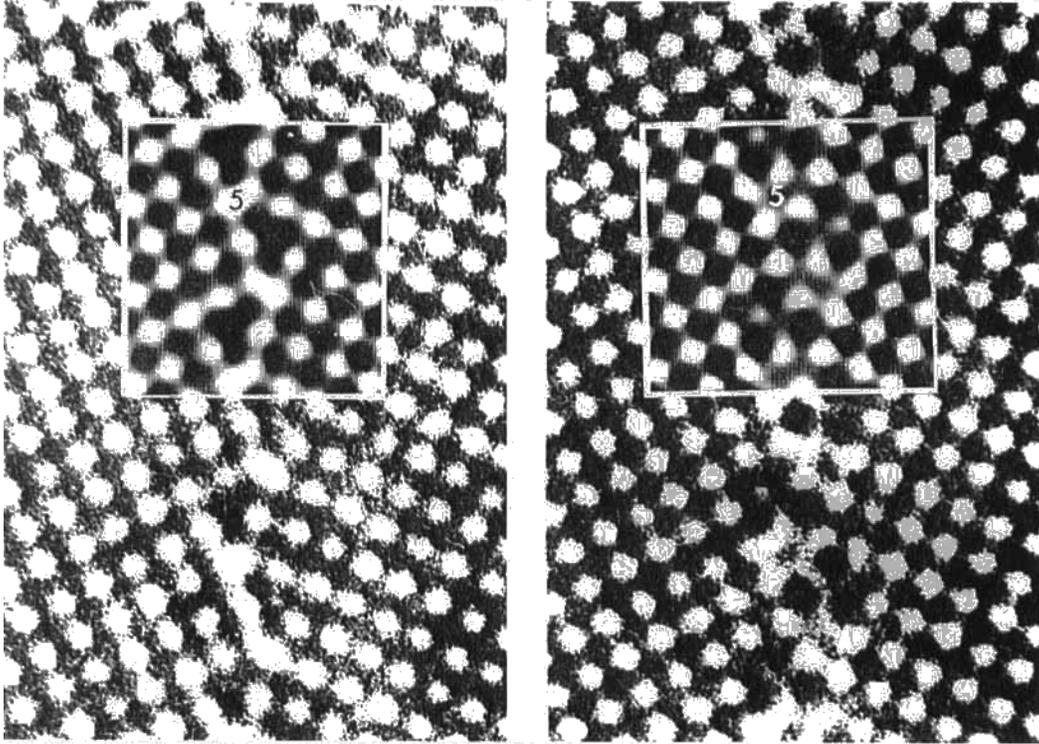
### Grain Boundaries

In non-coherent interfaces the atom columns can still be imaged by white dots but the exact positions of the dots do not necessarily correspond with the exact positions of the columns so that comparison with simulated images is necessary. An example is given in Fig. 28 showing a  $(310)\Sigma = 5$  grain boundary in germanium viewed along  $[001]$  which contains many dislocations. In agreement with the



**Figure 25.**  $[100]$  zone of the f.c.c.-based ordered alloy  $\text{Au}_4\text{Mn}$  [20]. The long period antiphase boundary superstructure 2223 is revealed in the high-resolution image (b). A dark field line resolution image is used to reveal the discommensurations (i.e., the 3-strips).





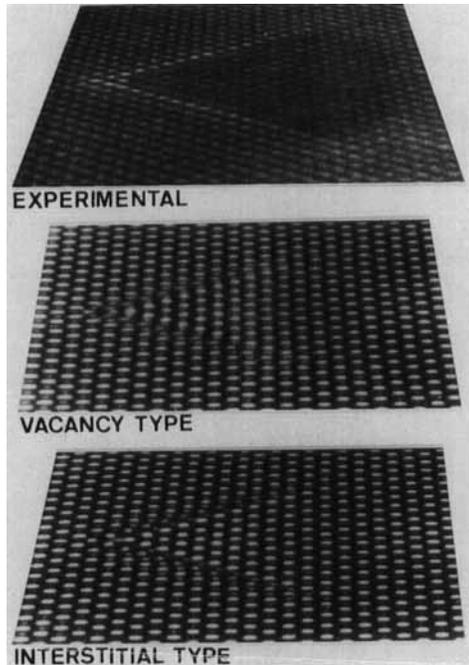
**Figure 28.**  $\Sigma = 5$  grain boundary in germanium. (Courtesy of J. L. Rouvière.)

tetrahedron as well as for the interstitial type. Approximately  $10^5$  atoms are involved in the calculations. By carefully looking along a glancing incidence (Fig. 29), it is clear that the displacement of the bright dots in the interstitial type is directed towards the point, whereas in the vacancy type the displacement is in the opposite sense, in agreement with experiment. Hence it can be concluded that the stacking fault tetrahedra in silicon are of the vacancy type.

This is a very fortunate case where only two different atomic models are possible which are so different that discrimination can be made on the basis of only a few image simulations.

### Dislocations

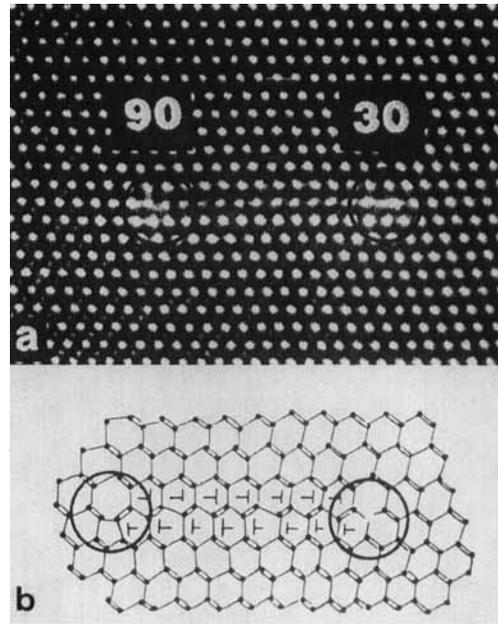
When edge dislocations are parallel to the zone axis of observation and when the resolution of the microscope is sufficient to discriminate the individual atoms the dislocation structure can be unravelled. Figure 30a represents the high-resolution image of a  $60^\circ$  dissociated dislocation in silicon. The dislocation is dissociated in two Schockley partials, one  $90^\circ\text{D}$  and one  $30^\circ\text{D}$  enclosing an intrinsic stacking fault. From the high-resolution image one can estimate the dissociation energy. The main feature is that the partials consist of well-defined structural units (Fig. 30b) that also occur in the dislocation cores. Figure



**Figure 29.** SFT in silicon, viewed under glancing angle. Top: experimental image. Middle: simulated image for a vacancy-type SFT. Bottom: simulated image for an interstitial-type SFT. From this it can be concluded that the SFT is of the vacancy type. (Courtesy of W. Coene and H. Bender.)

30c shows the dislocation structures occurring at a  $\Sigma = 9$  grain boundary in silicon. Here also the same structural units can be observed.

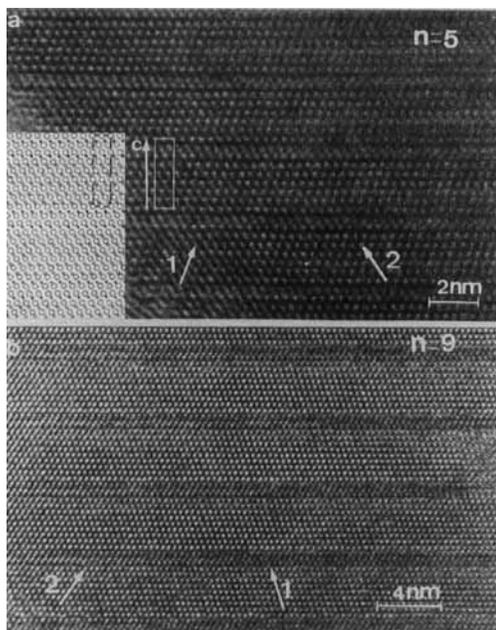
Figure 31 shows an image of a  $60^\circ$  dislocation in silicon dissociated into a  $30^\circ$  and a  $90^\circ$  partial. In Fig. 31 one can observe, in between the partials, different sets of 0.33 nm fringes. What is particularly interesting is that these fringes are caused by  $(\bar{4}23)/3$  reflections that are forbidden in the bulk crystal. Hence they reveal detailed information about the dislocation cores at the subnanometer level. For details the reader is referred to Alexander et al. [22].



**Figure 30.** A  $60^\circ$  dissociated dislocation in silicon. (Courtesy of J. Thibault-Desseaux.)



**Figure 31.** Image of a dissociated  $60^\circ$  dislocation in silicon. In between the  $30^\circ$  and  $90^\circ$  partials, different sets of 0.33 nm fringes are visible. These fringes are caused by forbidden  $(\bar{4}22)/3$  reflections and contain high-resolution information about the dislocation cores. (Courtesy of J. C. H. Spence.)



**Figure 32.** Mixed-layer compounds  $\text{As}_2\text{Te}_3(\text{GeTe})_n$  with (a)  $n = 5$  and (b)  $n = 9$  [23]. The image for  $n = 5$  is compared with a structure model in the inset. The  $\text{As}_2\text{Te}_3$  layer can clearly be distinguished from the (GeTe) block.

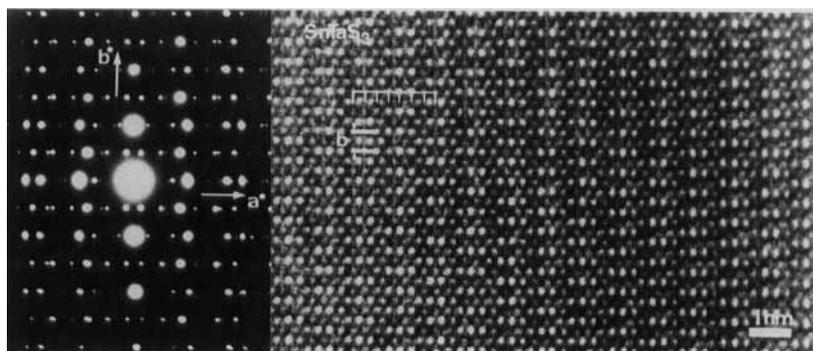
### Mixed-Layer Compounds

Mixed-layer compounds can generally be defined as an alternation of stackings of

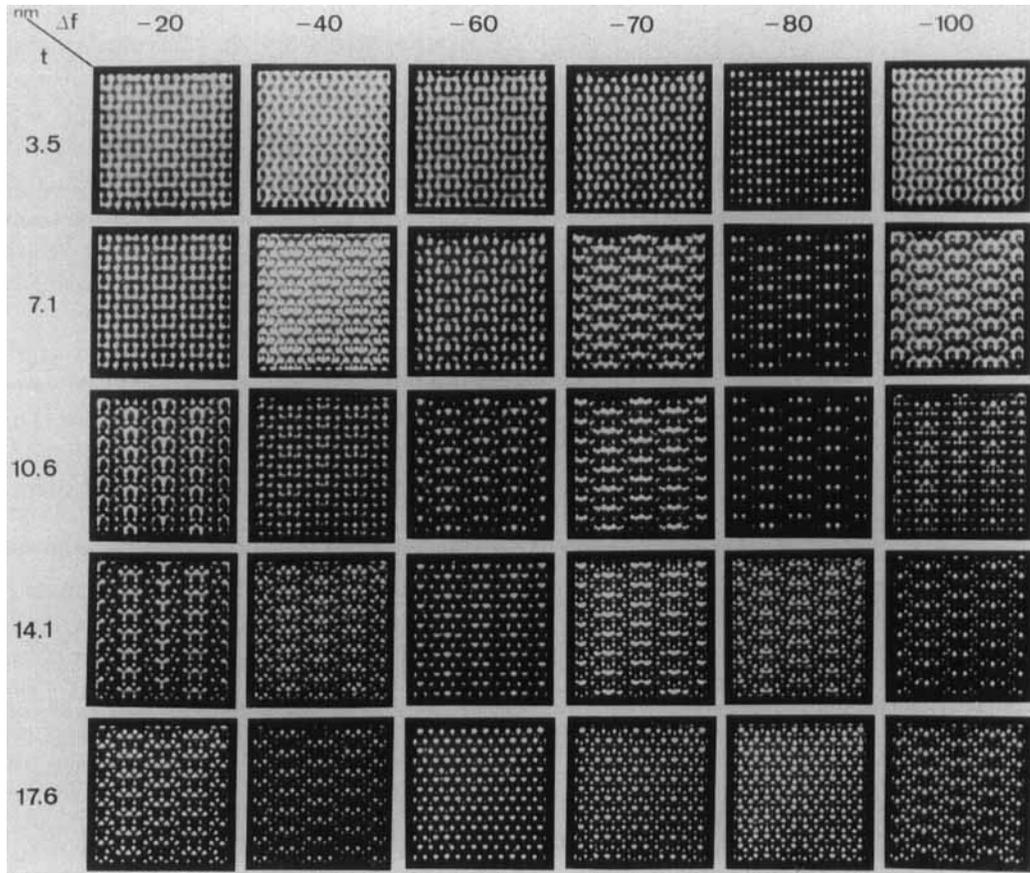
blocks of different types in which the number of blocks and the stacking sequence can be altered by changes in composition and so forth. In a sense, a system with periodical translation interfaces or twin planes can also be considered as a mixed-layer compound. In the literature a more restricted definition can be found in which each type of the basic block has to occur as an existing member of the series.

Figure 32 reveals a mixed layer compound of the series  $\text{As}_2\text{Te}_3(\text{GeTe})_n$  with  $n = 5$  (Fig. 32a) and  $n = 9$  (Fig. 32b). The systems consist of a regular stacking of  $\text{As}_2\text{Te}_3$  layers with (GeTe) blocks. A structure model is shown in the inset.

Figure 33 shows an example of an incommensurate stacking in the mixed-layer compound  $\text{SnTaS}_3$ . This is a compound of the type  $\text{MTS}_3$ , which in this case consists of a stacking of hexagonal  $\text{TaS}_2$  and tetragonal  $\text{SnS}$  layers. The image simulations which confirm the model are shown in Fig. 34. Figure 35 shows the images of a homologous series of polytypoids of the type  $\text{La}_n\text{Ti}_{n-\delta}\text{O}_{3n}$ . The layer sequences can be deduced directly from the images.



**Figure 33.** Incommensurate mixed layer compound of the type  $\text{MTS}_3(\text{SnTaS}_3)$  consisting of the juxtaposition of hexagonal  $\text{TaS}_2$  and tetragonal  $\text{SnS}$  layers [24]. (a) The diffraction pattern is the superposition of the diffraction patterns of the separate layers, supplemented with double diffraction spots. (b) High-resolution image along the normal to the layer plane.



**Figure 34.** Image simulation corresponding with the experimental situation of Fig. 33 [24]. Images are simulated for various focus values  $\Delta f$  and crystal thickness  $H$  (in nm). The best match is obtained for  $\Delta f = -60$  nm and  $t = 10.6$  nm.

### High- $T_c$ Superconductors

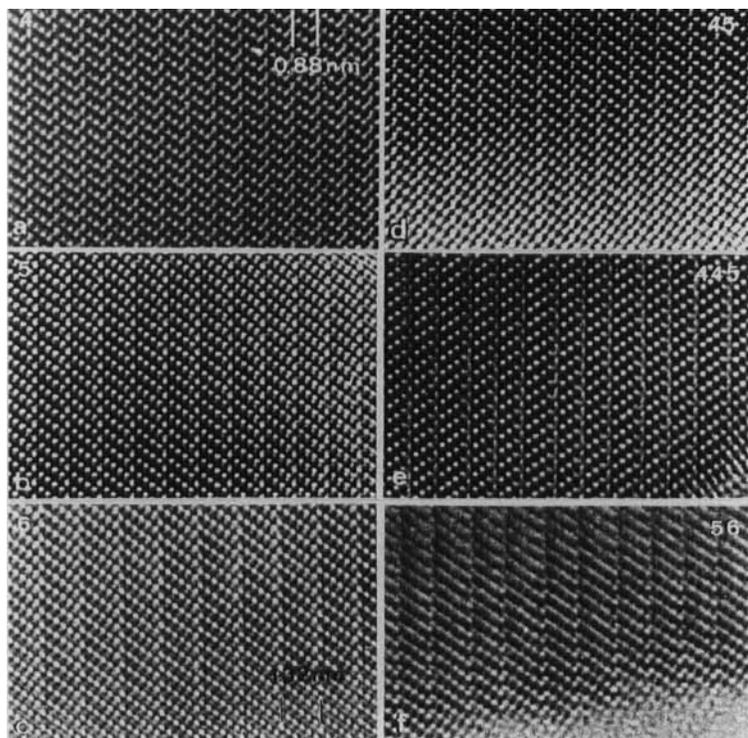
Figure 36 shows a high-resolution image of superconducting  $\text{BiSrCaO}$  viewed along  $[100]_p$ . From this it can be seen that the cleavage of the crystal has taken place in between two adjacent  $\text{BiO}$  layers.

Figure 37 shows a high-resolution image of superconducting  $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ .  $(\text{CuO})_2$  double layers as well as  $\text{CuO}$  single layers can be identified. Cleavage has taken place at the  $\text{CuO}$  layer.

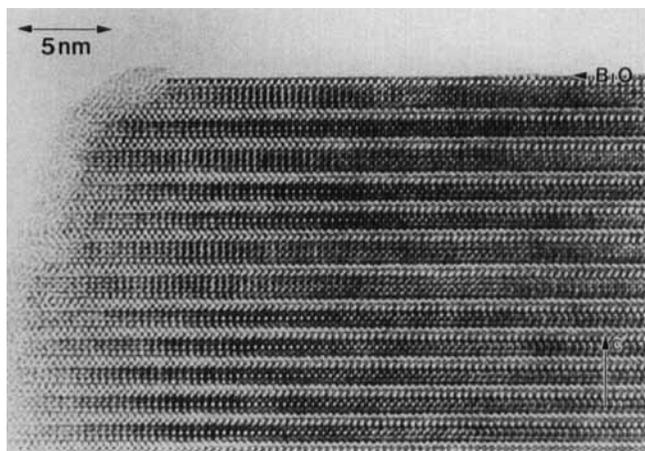
From these images it is clear that if the electron beam is nearly parallel to the surfaces of the crystal, the profile of the surface can be visualized directly. This enables one to characterize the surface layers and to detect surface relaxation or modification effects.

### Minerals

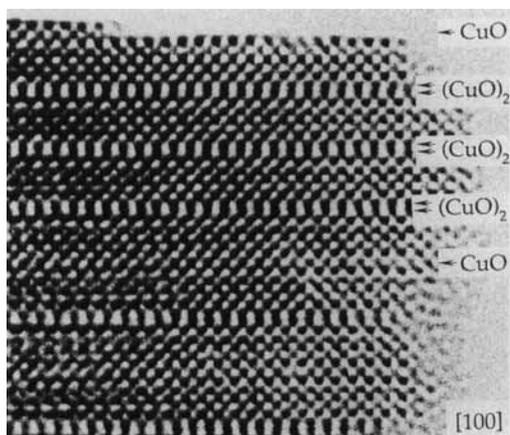
Figure 38 shows the image of the mineral hollandite ( $\text{Ba}_x\text{Mn}_8\text{O}_{16}$ ) viewed along



**Figure 35.** High-resolution image along the  $[1\ 1\ \bar{2}0]^*$  zone of a homologous series of polytypoids with composition  $\text{La}_n\text{Ti}_{n-\delta}\text{O}_{3n}$  [25]. The layer sequences in the different polytypoids, as deduced from the image and from the diffraction pattern, are indicated.

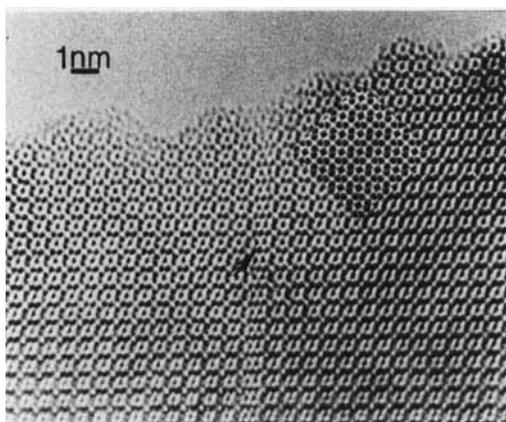


**Figure 36.** Superconducting  $\text{BiSrCaO}$  compound exhibiting a cleaved surface [26]. The high-resolution image along  $[1\ 0\ 0]_p$  proves that cleavage took place between the two adjacent  $\text{BiO}$  layers in the structure.



**Figure 37.** Superconducting  $\text{YBa}_2\text{Cu}_3\text{O}_4$ . The  $(\text{CuO})_2$  double layer as well as  $(\text{CuO})$  single layers can be identified. Cleavage has taken place at the  $\text{CuO}$  layer [27].

[0 1 0]. In this case the images can be interpreted intuitively in terms of the projected potential (see Sec. 1.1.2.6). As can be judged by comparing with the structure model in Fig. 39, a planar defect can be identified. The image matches with the computer simulation (inset).



**Figure 38.** [0 1 0] zone image of the mineral hollandite ( $\text{Ba}_x\text{Mn}_8\text{O}_{16}$ ). The barium ions in the channels of the  $\text{MnO}_2$  framework are revealed as dark dots. A planar defect consisting of double channels is present. In the top right corner a computer simulated image is superposed on the actual image [28].

## Carbon Structures

Figure 40 shows the image of a  $\text{C}_{60}$  crystal ('bucky balls') containing intersecting intrinsic stacking faults. A model for the intersection is shown in the inset.

Figure 41 shows an image of a rubidium intercalated  $\text{C}_{60}$  crystal ( $\text{Rb}_6\text{C}_{60}$ ). Also in this case the contrast can be explained intuitively in terms of the projected structure (Fig. 42). Figure 43 shows a high-resolution image of straight and helix-shaped nanotubes in which the (000 1) graphite planes are resolved.

## Cage Structures

Figures 44 and 45 show images of the cage structure of  $\text{Na}_8\text{Si}_{46}$  and  $\text{Na}_x\text{Si}_{136}$  viewed along (00 1), respectively (0 1 1). Since these structures are very open, the images can be directly interpreted in terms of the projected structure viewed along these directions (Fig. 46, respectively Fig. 47).

### 1.1.2.8 Appendices

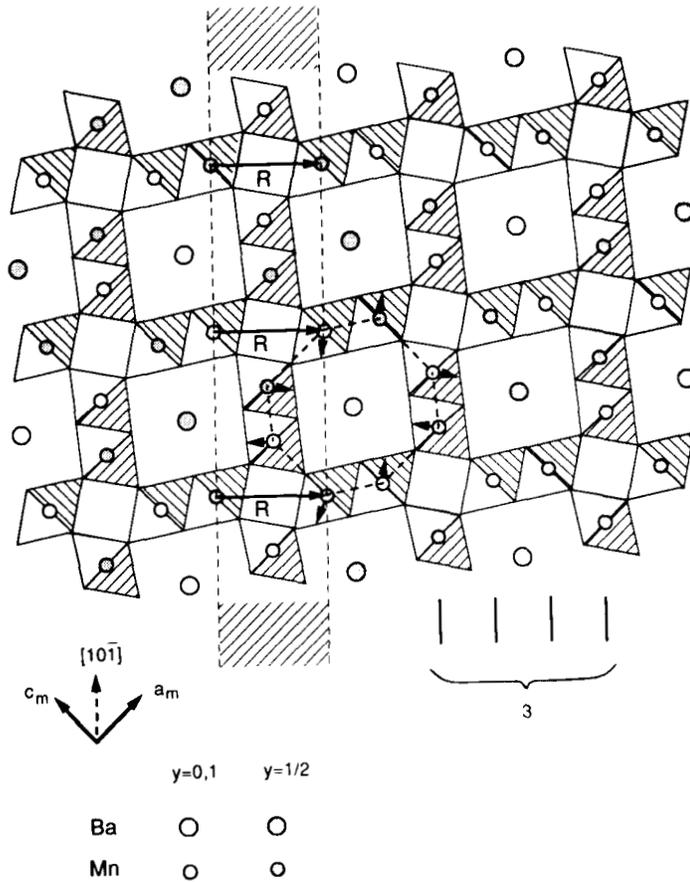
#### Appendix A: Image Formation [1]

The wavefunction in the image plane is given by

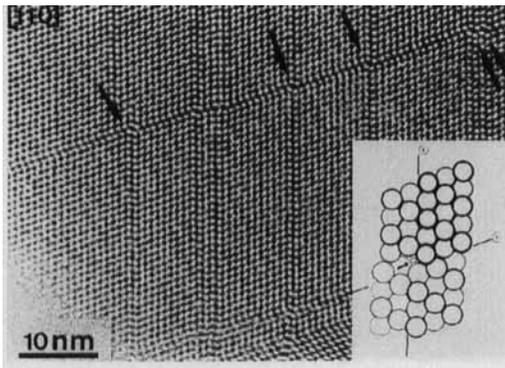
$$\phi(\mathbf{R}) = F_{\mathbf{R}}^{-1} A(\mathbf{g}) \exp[-i\chi(\mathbf{g})] \times D(\alpha, \Delta, \mathbf{g}) F_{\mathbf{g}} \psi(\mathbf{R}) \quad (36)$$

$A(\mathbf{g})$  represents the physical aperture with radius  $g_A$  selecting the imaging beams, thus

$$A(\mathbf{g}) = \begin{cases} 1 & \text{for } |g| \leq g_A \\ 0 & \text{for } |g| > g_A \end{cases}$$



**Figure 39.** Schematic representation of the hollandite structure along the [0 1 0] zone. A double row planar defect as seen in Fig. 38 is generated by removing the slice indicated by dotted lines and closing the gap [28].

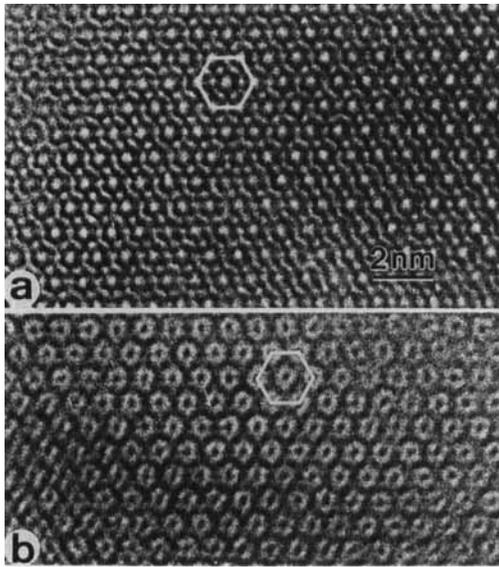


**Figure 40.** [1 1 0] zone image of a  $C_{60}$  crystal containing intersecting intrinsic stacking faults. Along the intersection line a row of partial vacancies is formed as represented schematically in the inset [29].

The total phase shift due to spherical aberration and defocus is

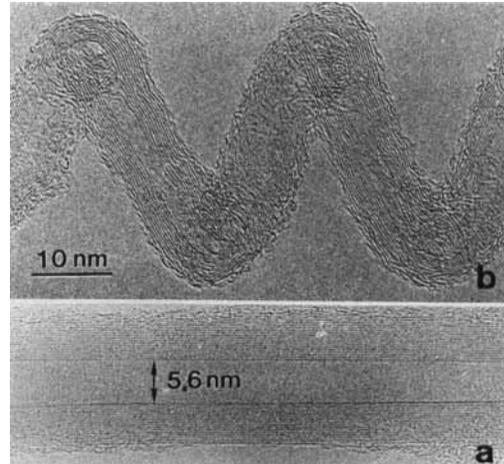
$$\chi(g) = \frac{1}{2}\pi C_s \lambda^3 g^4 + \pi \epsilon \lambda g^2 \quad (37)$$

with  $C_s$  the spherical aberration coefficient,  $\epsilon$  the defocus, and  $\lambda$  the wavelength. The imaging process is also influenced by spatial and temporal incoherence effects. Spatial incoherence is caused by the fact that the illuminating beam is not parallel but can be considered as a cone of incoherent plane waves (beam convergence). The image then results from a superposition of the respective image intensities. Temporal incoherence results



**Figure 41.** (a) Rubidium intercalated  $C_{60}$  crystal ( $Rb_6C_{60}$ ). The circular bright areas image the  $C_{60}$  molecules: the dark areas represent the Rb ions. The structure is imaged along the  $[1\ 1\ 1]$  zone; (b) different defocus [30].

from fluctuations in the energy of the thermally emitted electrons, in the fluctuation of the lens currents, and of the accelerating voltage. All these effects cause the focus  $\epsilon$  to fluctuate. The final image is then



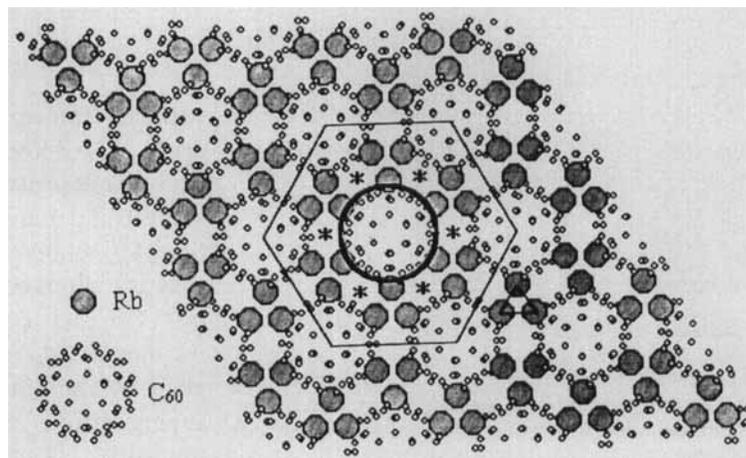
**Figure 43.** High-resolution images of carbon nanotubes [31]. (a) Straight tube; only the  $(000\ 1)$  planes are resolved. (b) Helix-shaped nanotubes;  $(000\ 1)$  planes are resolved.

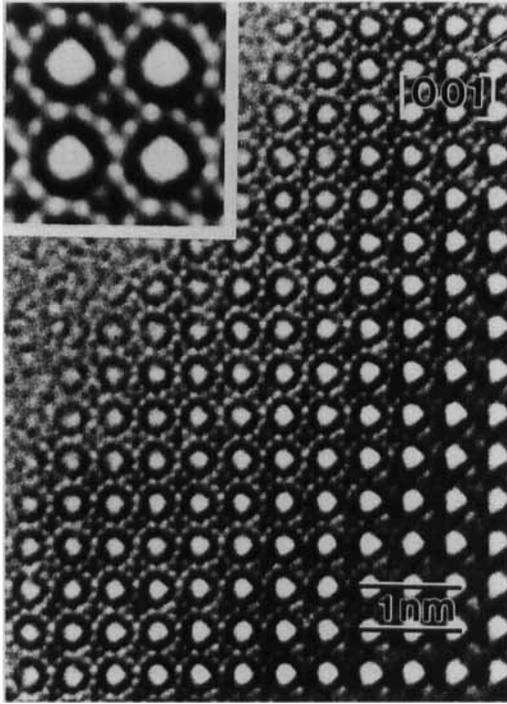
the superposition (integration) of the images corresponding with the different incident beam directions  $\mathbf{K}$  and focus values  $\epsilon$ , that is,

$$I(\mathbf{R}) = \int_{\epsilon} \int_{\mathbf{K}} |\phi(\mathbf{R}, \mathbf{K}, \epsilon)|^2 f_s(\mathbf{K}) f_T(\epsilon) d\mathbf{K} d\epsilon \quad (38)$$

where  $\phi(\mathbf{R}, \mathbf{K}, \epsilon)$  denotes that the

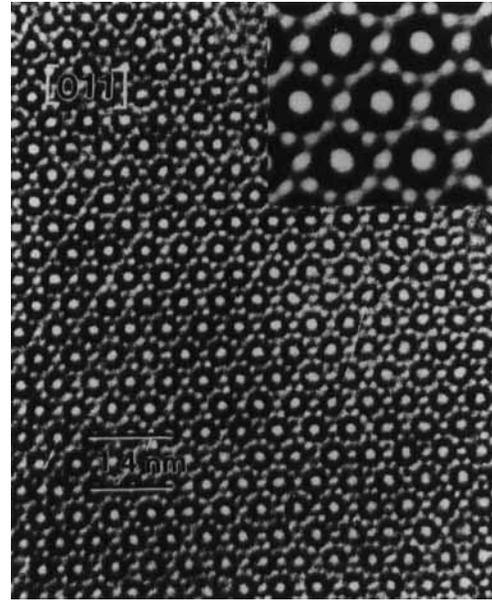
**Figure 42.** Structure model, projected along the  $[1\ 1\ 1]$  of the intercalate  $Rb_6C_{60}$ . This model can be compared with the image of Fig. 41b [30].





**Figure 44.** High-resolution image of the cage structure  $\text{Na}_8\text{Si}_{46}$  [32]. The bright dots correspond to open channels in the structure. The inset is an averaged image.

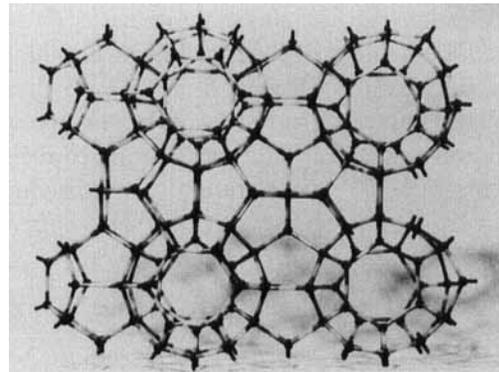
wavefunction in the image plane also depends on the incident wavevector  $\mathbf{K}$  and on the defocus  $\epsilon$ .  $f_s(\mathbf{K})$  and  $f_c(\epsilon)$  are the probability distribution functions of  $\mathbf{K}$ , respectively  $\epsilon$ . Expressions (36), (37), and (38) are the basic expressions describing the whole real imaging process. They are also used for the computer simulation of high-resolution images. However, the computation of Eq. (38) requires the computation of  $\psi(\mathbf{R})$  for a large number of defocus values and beam directions, which in practice is a horrible task. For this reason, Eq. (38) has often been approximated. In order to study the effect of chromatic aberration and beam convergence (on a more intuitive basis)



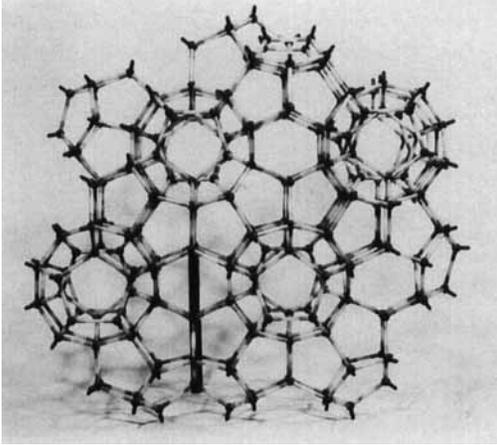
**Figure 45.** High-resolution image of the cage structure  $\text{Na}_v\text{Si}_{136}$  [32]. The bright dots correspond to open channels in the structure. The inset is an averaged image.

we assume a disk-like effective source function

$$f_s(\mathbf{K}) = \begin{cases} 1 & \text{for } |\mathbf{K}| < \alpha/\lambda \\ 0 & \text{for } |\mathbf{K}| > \alpha/\lambda \end{cases}$$



**Figure 46.** Model for the structure of  $\text{Na}_8\text{Si}_{46}$  as viewed along the same zone as the image of Fig. 44 [32].



**Figure 47.** Model for the cage structure  $\text{Na}_x\text{Si}_{136}$  which can be compared with the image of Fig. 45 [32].

with  $\alpha$  the apex angle of the illumination cone. We assume further that the integrations over defocus and beam convergence can be performed coherently, that is, over the amplitudes rather than the intensities. This latter assumption is justified when the intensity of the central beam is much larger than the intensities of the diffracted beams so that cross-products between diffracted beam amplitudes can be neglected. We assume that the defocus spread  $f_T(\epsilon)$  is a Gaussian centered on  $\epsilon$  with a half-width  $\Delta$ . Assuming the object function  $\psi(\mathbf{R})$  to be independent of the inclination  $\mathbf{K}$ , which is only valid for thin objects, one then finally finds that the effect of the chromatic aberration, combined with beam convergence, can be incorporated by multiplying the transfer function with an effective aperture function

$$D(\alpha, \Delta, \mathbf{g}) = B(\Delta, \mathbf{g})C(\alpha, \Delta, \mathbf{g}) \quad (39)$$

where

$$B(\Delta, \mathbf{g}) = \exp\left(-\frac{1}{2}\pi^2\lambda^2\Delta^2g^4\right) \quad (40)$$

representing the effect of the defocus

spread, and

$$C(\alpha, \Delta, \mathbf{g}) = 2J_1(|\mathbf{q}|)/|\mathbf{q}| \quad (41)$$

with  $J_1$  the Bessel function and  $|\mathbf{q}| = (\mathbf{q} \cdot \mathbf{q})^{1/2}$ , which may be a complex function for a complex  $\mathbf{q}$

$$\mathbf{q} = 2\pi\alpha\mathbf{g}[\epsilon + \lambda g^2(\lambda C_s - i\pi\Delta^2)] \quad (42)$$

$C(\alpha, \Delta, \mathbf{g})$  represents the combined effect of beam convergence and defocus spread.

The total image transfer can now be described from Eqs. (5) and (6) as

$$\begin{aligned} \phi(\mathbf{R}) = & F_{\mathbf{R}}^{-1}A(\mathbf{g})\exp[-i\chi(\mathbf{g})] \\ & \times D(\alpha, \Delta, \mathbf{g})F_{\mathbf{g}}\psi(\mathbf{R}) \end{aligned} \quad (43)$$

that is, the effective aperture yields a damping envelope function for the phase transfer function. Other approximations for including the effects of beam convergence and chromatic aberrations [33] using a Gaussian effective source lead to a similar damping envelope function. Experimentally obtained transfer functions confirm this behavior.

In Eq. (43) the incoherent effects are approximated by a coherent envelope function. Hence it is called the coherent approximation. It is usually valid for thin objects. A full treatment of incoherent effects requires the calculation of the double integral in Eq. (38). Another approximation which is valid for thicker objects is based on the concept of the TCC [34]. Here it is assumed that beam convergence and defocus spread do not influence the diffraction in the object. Hence in Eq. (5) they do not appear in the object wavefunction but only in the phase transfer function. Now the wavefunction in the image plane (Eq. (5)) can be written as

$$\Phi(\mathbf{R}, \mathbf{K}, \epsilon) = F_{\mathbf{R}}^{-1}T(\mathbf{g}, \mathbf{K}, \epsilon)\psi(\mathbf{g}) \quad (45)$$

with

$$T(\mathbf{g}, \mathbf{K}, \epsilon) = A(\mathbf{g}) \exp[-i\chi(\mathbf{g}, \mathbf{K}, \epsilon)] \quad (46)$$

Substituting Eq. (8) into Eq. (38) then yields, after Fourier transforming,

$$\begin{aligned} I(\mathbf{g}) &= F_{\mathbf{g}}[I(\mathbf{R})] \\ &= \int \psi(\mathbf{g} + \mathbf{g}') \tau(\mathbf{g} + \mathbf{g}', \mathbf{g}') \psi^*(\mathbf{g}') d\mathbf{g}' \end{aligned} \quad (47)$$

with

$$\begin{aligned} \tau(\mathbf{g} + \mathbf{g}', \mathbf{g}') &= \int_{\epsilon} \int T^*(\mathbf{g} + \mathbf{g}', \mathbf{K}, \epsilon) \\ &\quad \times T(\mathbf{g}', \mathbf{K}, \epsilon) d\mathbf{K} d\epsilon \end{aligned} \quad (48)$$

$\tau$  is the transmission cross-coefficient. It describes how the beams  $\mathbf{g}'$  and  $\mathbf{g} + \mathbf{g}'$  are coupled to yield the Fourier component  $\mathbf{g}$  of the image intensity.

## Appendix B: Scattering of an Electron in a Thin Object

We will now follow a classical approach.

The non-relativistic expression for the wavelength of an electron accelerated by an electrostatic potential  $E$  is given by

$$\lambda = \frac{h}{\sqrt{2meE}} \quad (49)$$

with  $h$  the Planck constant,  $m$  the electron mass, and  $e$  the electron charge.

During the motion through an object with local potential  $V(x, y, z)$  the wavelength will vary with the position of the electron as

$$\lambda'(x, y, z) = \frac{h}{\sqrt{2me[E + V(x, y, z)]}} \quad (50)$$

For thin phase objects and large accelerating potentials the assumption can be made

that the electron keeps traveling along the  $z$  direction so that by propagation through a slice  $dz$  the electron suffers a phase shift:

$$\begin{aligned} d\chi(x, y, z) &= 2\pi \frac{dz}{\lambda'} - 2\pi \frac{dz}{\lambda} \\ &= 2\pi \frac{dz}{\lambda} \left( \frac{\sqrt{E + V(x, y, z)}}{\sqrt{E}} - 1 \right) \\ &\simeq \sigma V(x, y, z) dz \end{aligned} \quad (51)$$

with

$$\sigma = \pi/\lambda E$$

so that the total phase shift is given by

$$\chi(x, y) = \sigma \int V(x, y, z) dz = \sigma V_p(x, y) \quad (52)$$

where  $V_p(x, y)$  represents the potential of the specimen projected along the  $z$  direction.

Under this assumption the specimen acts as a pure phase object with transmission function

$$\psi(x, y) = \exp[i\sigma V_p(x, y)] \quad (53)$$

When the object is very thin, one has

$$\psi(x, y) \approx 1 + i\sigma V_p(x, y) \quad (54)$$

This is the weak phase approximation.

The effect of all processes, prohibiting the electrons from contributing to the image contrast, including the use of a finite aperture can in a first approximation be represented by a projected absorption function in the exponent of Eq. (53), so that

$$\psi(x, y) = \exp[i\sigma V_p(x, y) - \mu(x, y)] \quad (55)$$

or

$$\psi(\mathbf{R}) = \exp[i\sigma V_p(\mathbf{R}) - \mu(\mathbf{R})] \quad (56)$$

with  $\mathbf{R} = (x, y)$  the vector in the plane perpendicular to  $z$ .

## Appendix C: Scattering of an Electron in a Thick Object [35]

### Optical Approach: The Multislice Method [36]

Although the multislice formula can be derived from quantum mechanical principles, we follow a simplified version of the more intuitive original optical approach. A more rigorous treatment is given in the next section.

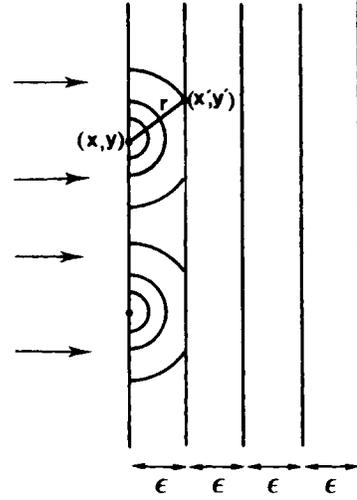
Consider a plane wave, incident on a thin specimen foil and nearly perpendicular to the incident beam direction  $z$ . If the specimen is sufficiently thin, we can assume the electron to move approximately parallel to  $z$  so that the specimen acts as a pure phase object with transmission function Eq. (53):

$$\psi(x, y) = \exp[i\sigma V_p(x, y)]$$

A thick specimen can now be subdivided into thin slices, perpendicular to the incident beam direction. The potential of each slice is projected into a plane which acts as a two-dimensional phase object. Each point  $(x, y)$  of the exit plane of the first slice can be considered as a Huyghens source for a secondary spherical wave with amplitude  $\psi(x, y)$  (Fig. 48).

Now the amplitude  $\psi(x', y')$  at the point  $(x', y')$  of the next slice can be found by the superposition of all spherical waves of the first slice, that is, by integration over  $x$  and  $y$ , yielding

$$\begin{aligned} \psi(x', y') = & \int \exp[i\sigma V_p(x, y)] \\ & \times \frac{\exp(2\pi i k r)}{r} dx dy \end{aligned}$$



**Figure 48.** Schematic representation of the propagation effect of electrons between successive slices of thickness  $\epsilon$ .

When  $|x - x'| \ll \epsilon$   $|y - y'| \ll \epsilon$ , with  $\epsilon$  the slice thickness, the Fresnel approximation can be used, that is,

$$\begin{aligned} r &= \sqrt{(x - x')^2 + (y - y')^2 + \epsilon^2} \\ &\approx \epsilon \left( 1 + \frac{(x - x')^2}{2\epsilon^2} + \frac{(y - y')^2}{2\epsilon^2} \right) \end{aligned}$$

so that

$$\begin{aligned} \psi(x', y') &\approx \frac{\exp(2\pi i k \epsilon)}{\epsilon} \\ &\times \int \exp[i\sigma V_p(x, y)] \\ &\times \exp\left[i \frac{\pi k}{\epsilon} [(x - x')^2 + (y - y')^2]\right] dx dy \end{aligned}$$

which, apart from constant factors, can be written as a convolution product:

$$\begin{aligned} \psi(x, y) = & \epsilon [i\sigma V_p(x, y)] \\ & * \exp[i\pi k(x^2 + y^2)/\epsilon] \end{aligned} \quad (57)$$

where the convolution product of two functions is defined as (in one dimension)

$$f(x) * g(x) = \int f(x')g(x - x') dx'$$

If the wavefunction at the entrance face is  $\psi(x, y, 0)$ , instead of a plane wave one has for the wavefunction at the exit face

$$\psi(x, y, \epsilon) = \{\psi(x, y, 0) \exp[i\sigma V_p(x, y)]\} \\ * \exp[i\pi k(x^2 + y^2)/\epsilon] \quad (58)$$

This is the Fresnel approximation in which the emerging spherical wavefront is approximated by a paraboloidal wavefront.

The propagation through the vacuum gap from one slice to the next is thus described by a convolution product in which each point source of the previous slice contributes to the wavefunction in each point of the next slice. The motion of an electron through the whole specimen can now be described by an alternating of phase object transmissions (multiplications) and vacuum propagations (convolutions). In the limit of the slice thickness  $\epsilon$  tending to zero, this multislice expression converges to the exact solution of the non-relativistic Schrödinger equation in the forward-scattering approximation.

In the original multislice method one used the Fourier transform of Eq. (58) where the real space points  $(x, y)$  are transformed into diffracted beams  $\mathbf{g}$  and where convolution and normal products are interchanged, that is,

$$\psi(\mathbf{g}, \epsilon) = \{\psi(\mathbf{g}, 0) * \exp[i\sigma V_g]\} \\ \times \exp[i\pi g^2 \epsilon/k] \quad (59)$$

where  $V_g$  are the structure factors (Fourier transforms of the unit cell potential).

The wavefunction at the exit face of the crystal can now be obtained by successive

application of Eq. (58) or (59). This can either be done in real space (Eq. (58)) or in reciprocal space (Eq. (59)). The major part of the computing time is required for the calculation of the convolution product, which is proportional to  $N^2$  ( $N$  is the number of sampling points (real space) or beams (reciprocal space)).

Since the Fourier transform of a convolution product yields a normal product (with calculation time proportional to  $N$ ), a large gain in speed can be obtained by alternatively performing the propagation in reciprocal space and the phase object transmission in real space [37]. In this way the computing time is devoted to the Fourier transforms and is proportional to  $N \log_2 N$ .

Another way of increasing the speed is in the so-called real space method [38]. Here the whole calculation is done in real space using Eq. (58) but the forward scattering of the electrons is exploited so as to calculate the convolution effect of the propagation only in a limited number of adjacent sampling points. In this way, the calculation time is proportional to  $N$ . This method does not require a periodic crystal and is thus suitable for calculation of crystal defects.

### Quantum Mechanical Approach [35]

As is clear from Eq. (5) the calculation of the image wavefunction  $\Phi(\mathbf{R})$  requires the knowledge of  $\psi(\mathbf{R})$ , that is, the wavefunction at the exit face of the object. This can be obtained by numerically solving the Schrödinger equation in the object. For convenience we will now follow a simplified more intuitive approach, which leads to the correct results.

If we assume that the fast electron, in the direction of propagation ( $z$  axis) behaves as a classical particle with velocity  $v = \hbar k/m$  we can consider the  $z$  axis as a time axis with

$$t = \frac{mz}{\hbar k} \quad (60)$$

Hence we can start from the time-dependent Schrödinger equation

$$-\frac{\hbar}{i} \frac{\partial \psi}{\partial t}(\bar{\mathbf{R}}, t) = H\psi(\bar{\mathbf{R}}, t) \quad (61)$$

with

$$H = -\frac{\hbar^2}{2m} \Delta_{\mathbf{R}} - eU(\mathbf{R}, t) \quad (62)$$

with  $U(\mathbf{R}, t)$  the electrostatic crystal potential,  $m$  and  $k$  the relativistic electron mass and wavelength, and  $\Delta_{\mathbf{R}}$  the Laplacian operator acting in the plane ( $\mathbf{R}$ ) perpendicular to  $z$ .

Using Eq. (60) we then have

$$\frac{\partial \psi(\mathbf{R}, z)}{\partial z} = \frac{i}{4\pi k} (\Delta_{\mathbf{R}} + V(\mathbf{R}, z))\psi(\mathbf{R}, z) \quad (63)$$

with

$$V(\mathbf{R}, z) = \frac{2me}{\hbar^2} U(\mathbf{R}, z) \quad (64)$$

This is the well-known high-energy equation in real space which can also be derived from the stationary Schrödinger equation in the forward-scattering approximation [35].

In HREM of crystalline objects, the object is usually oriented along a zone axis, so that the electrons are traveling parallel to the atom columns. If the periodicity along the column direction is not too large (less than 1–2 nm), the fast electron does not feel this variation. In fact

it sees the potential as constant along  $z$ . In other words, the effect of higher-order Laue zones or upper layer lines is negligible. This is the projection approximation, which is usually valid for most high-resolution conditions, Now Eq. (63) becomes

$$\frac{\partial \psi(\mathbf{R}, z)}{\partial z} = \frac{i}{4\pi k} (\Delta_{\mathbf{R}} + V(\mathbf{R}))\psi(\mathbf{R}, z) \quad (65)$$

with

$$V(\mathbf{R}) = \frac{2me}{\hbar} \frac{1}{z} \int_0^z U(\mathbf{R}, z) dz \quad (66)$$

the potential, averaged (projected) along  $z$ . In the time-dependent Schrödinger picture (Eq. (61)) the electron walks as a function of time in a two-dimensional potential of projected atom columns.

Equation (65) can also be transformed to reciprocal space.

Assuming  $V(\mathbf{R})$  to be periodic in two dimensions, we can expand it in Fourier series,

$$V(\mathbf{R}) = \sum_{\mathbf{g}} V_{\mathbf{g}} \exp 2\pi i \mathbf{g} \cdot \mathbf{R} \quad (67)$$

with  $\mathbf{g}$  in the zone plane.  $V_{\mathbf{g}}$  are structure factors. Similarly we have

$$\psi(\mathbf{R}) = \sum_{\mathbf{g}} \psi_{\mathbf{g}}(z) \exp 2\pi i \mathbf{g} \cdot \mathbf{R} \quad (68)$$

$\psi_{\mathbf{g}}(z)$  represents the amplitude of the beam  $\mathbf{g}$  at a depth  $z$ . Substitution in Eq. (65) then yields

$$\frac{d\psi_{\mathbf{g}}}{dz} = i\pi \left[ 2s_{\mathbf{g}}\psi_{\mathbf{g}}(z) + \sum_{\mathbf{g}'} V_{\mathbf{g}-\mathbf{g}'}\psi_{\mathbf{g}'}(z) \right] \quad (69)$$

with

$$s_{\mathbf{g}} = g^2/2k \quad (70)$$

the excitation error, which is approximately equal to the distance between the

reciprocal node  $\mathbf{g}$  and the Ewald sphere, measured along  $z$ . This system of coupled first-order differential equations was derived in the early 1960s (for references, see Van Dyck [35]). Most of the image simulation programs are based on a numerical solution of the dynamical equation in real space (Eq. (65)) or reciprocal space (Eq. (69)), or a combination of both.

The dynamic relation (63) or (65) is a mixture of two equations, each representing a different physical process.

$$\frac{\partial\psi(\mathbf{R}, z)}{\partial z} = \frac{i}{4\pi k} \Delta_{\mathbf{R}}\psi(\mathbf{R}, z) \quad (71)$$

is a complex diffusion-type of equation, which represent the free electron propagation and whose solution can be represented formally as

$$\psi(\mathbf{R}, z) = \exp\left(\frac{i\Delta_{\mathbf{R}}z}{4\pi k}\right)\psi(\mathbf{R}, 0) \quad (72)$$

It can also be written as a convolution product:

$$\psi(\mathbf{R}, z) = \exp\frac{i\pi k R^2}{z} * \psi(\mathbf{R}, 0) \quad (73)$$

The other part of Eq. (65) is a differential equation,

$$\frac{\partial\psi(\mathbf{R}, z)}{\partial z} = \frac{i}{4\pi k} V(\mathbf{R}, z)\psi(\mathbf{R}, z) \quad (74)$$

which represents the scattering of the electron by the crystal potential. It can be readily integrated in real space, yielding

$$\psi(\mathbf{R}, z) = \exp\left(\frac{i}{4\pi k} V(\mathbf{R})z\right)\psi(\mathbf{R}, 0) \quad (75)$$

with  $V(\mathbf{R})$  the projected potential as defined in Eq. (66).

The wavefunction in real space is multiplied with a phase factor which is

proportional to the electrostatic potential of the object projected along  $z$  and which is called the phase object function. The solution of the complete dynamic equation (65) can be written formally as

$$\psi(\mathbf{R}, z) \exp\left(\frac{i}{2\pi k} [\Delta_{\mathbf{R}} + V(\mathbf{R})]z\right)\psi(\mathbf{R}, 0) \quad (76)$$

For the explicit calculation, slice methods are the most appropriate. Here the crystal is cut into thin slices with thickness  $\epsilon$  perpendicular to the incident beam.

If the slice thickness is sufficiently small, the solution within one slice is approximated by

$$\begin{aligned} \psi(\mathbf{R}, z + \epsilon) &= \exp\left(\frac{i}{2\pi k} \Delta_{\mathbf{R}}\epsilon\right) \\ &\times \exp\left(\frac{i}{4\pi k} V(\mathbf{R})\epsilon\right)\psi(\mathbf{R}, z) \end{aligned}$$

or explicitly

$$\begin{aligned} \psi(\mathbf{R}, z + \epsilon) &= \exp\frac{i\pi k R^2}{\epsilon} \\ &* \left[ \exp\left(\frac{i}{4\pi k} V(\mathbf{R})\epsilon\right)\psi(\mathbf{R}, 0) \right] \quad (77) \end{aligned}$$

This expression is essentially the same as Eq. (58), which was derived from an optical approach. In practice the wavefunction is sampled in a network of closely spaced points. At each point the wavefunction is multiplied with the phase object function. Then the wavefunction is propagated to the next slice, and so on. Calling  $N$  the number of sampling points, the phase object requires a calculation time proportional to  $N^2$ . In reciprocal space, direct and convolution products are

interchanged, yielding

$$\begin{aligned} \psi_g(z + \epsilon) &= \exp \frac{i\pi g^2 \epsilon}{k} \\ &\times \left[ \mathcal{F}_g \left( \exp \frac{i}{4\pi k} V(\mathbf{R}) \epsilon \right) * \psi_g(z) \right] \end{aligned} \quad (78)$$

This expression is the same as Eq. (59). Now the calculation time of the propagation is proportional to  $N$ , the number of beams, whereas the scattering in the phase object gives a calculation time proportional to  $N^2$ .

In order to speed up the calculation, the phase object is calculated in real space, and the propagator in reciprocal space [37]. Between each a fast Fourier transform is performed, the calculation time of which is only proportional to  $N \log_2 N$ . In the standard slice programs the object is assumed to be a perfect crystal. Defects are treated by the periodic continuation method in which the defect is artificially repeated so as to create an artificial super-crystal.

In the real space method, proposed in [38], the whole calculation is performed in real space, but due to the forward scattering of the electrons the propagation effect is limited to a local area so that the calculation time remains proportional to  $N$ . This is particularly interesting for treating extended or aperiodic structures.

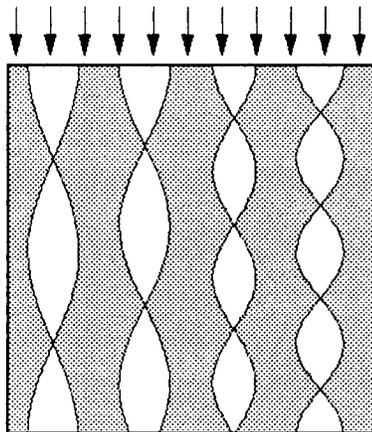
## Appendix D: Electron Channeling

### Principle

Although the slice methods are valuable for numerical purposes, they do not

provide much physical insight into the diffraction process. There is need for a simple intuitive theory that is valid for larger crystal thicknesses. In our view, a channeling theory fulfills this need. Indeed, it is well known that, when a crystal is viewed along a zone axis, that is, parallel to the atom columns, the high-resolution images often show a one-to-one correspondence with the configuration of columns provided the distance between the columns is large enough and the resolution of the instrument is sufficient. This is for instance the case in ordered alloys with a column structure (see Secs. 1.1.2.6 and 1.1.2.7). From this, it can be suggested that, for a crystal viewed along a zone axis with sufficient separation between the columns, the wavefunction at the exit face mainly depends on the projected structure, that is, on the type of atom columns. Hence, the classical picture of electrons traversing the crystal as plane-like waves in the directions of the Bragg beams which stems from the X-ray diffraction picture and upon which most of the simulation programs are based is, in fact, misleading. The physical reason for this 'local' dynamic diffraction is the channeling of the electrons along the atom columns parallel to the beam direction. Due to the positive electrostatic potential of the atoms, a column acts as a guide or channel for the electron [39, 40] within which the electron can scatter dynamically without leaving the column (Fig. 49).

The importance of channeling for interpreting high-resolution images has often been ignored or underestimated, probably because of the fact that, for historical reasons, dynamic electron diffraction is often described in reciprocal



**Figure 49.** Schematic representation of electron channeling.

space. However, since most of the high-resolution images of crystals are taken in a zone axis orientation, in which the projected structure is the simplest, but in which the number of diffracted beams are the largest, a simple real-space channeling theory yields a much more useful and intuitive, albeit approximate, description of the dynamic diffraction, which allows an intuitive interpretation of high-resolution images, even for thicker objects.

We will now give a simplified version of the basic principles of the theory. For more details the reader is referred to the literature [40].

### Isolated Columns

Consider an isolated column of atoms, parallel to the electron beam. If we now consider the depth proportional to the time, the dynamic equation (65) represents the walk of an electron in the two-dimensional projected potential of the columns.

The solution can be expanded in eigenfunctions (eigenstates) of the Hamiltonian

$$\psi(\mathbf{R}, z) = \sum_n C_n \phi_n(\mathbf{R}) \exp\left(-i\pi \frac{E_n}{E} \frac{z}{\lambda}\right) \quad (79)$$

where

$$H\phi_n(\mathbf{R}) = E_n\phi_n(\mathbf{R}) \quad (80)$$

with the Hamiltonian

$$H = -\frac{\hbar^2}{2m} \Delta_{\mathbf{R}} - eU(\mathbf{R}) \quad (81)$$

$U(\mathbf{R})$  is the projected potential of the column,

$$E = \frac{\hbar^2 k^2}{2m} \quad (82)$$

the incident electron energy, and  $\lambda$  is the electron wavelength. For  $E_n < 0$  the eigenstates are bound to the column. We now rewrite Eq. (79) as

$$\psi(\mathbf{R}, z) = \sum_n C_n \phi_n(\mathbf{R}) + \sum_n C_n \phi_n(\mathbf{R}) \times \left[ \exp\left(-i\pi \frac{E_n}{E} \frac{z}{\lambda}\right) - 1 \right] \quad (83)$$

The coefficients  $C_n$  are determined from the boundary condition

$$\sum_n C_n \phi_n(\mathbf{R}) = \psi(\mathbf{R}, 0) \quad (84)$$

In the case of plane wave incidence one thus has

$$\sum_n C_n \phi_n(\mathbf{R}) = 1 \quad (85)$$

so that

$$\psi(\mathbf{R}, z) = 1 + \sum_n C_n \phi_n(\mathbf{R}) \times \left[ \exp\left(-i\pi \frac{E_n}{E} \frac{z}{\lambda}\right) - 1 \right] \quad (86)$$

Only states will appear in the summation, for which

$$|E_n| \geq \frac{E\lambda}{z} \quad (87)$$

These are bound states with deep energy levels, which are localized near the column cores. In practice if the atom column is not too heavy and the distance between columns not too close (e.g., larger than 0.1 nm) only one eigenstate will appear, which can be compared to the 1s state of an atom.

We then have

$$\begin{aligned} \psi(\mathbf{R}, z) = & 1 + C\phi(\mathbf{R}) \\ & \times \left[ \exp\left(-i\pi \frac{E}{E_0} \frac{z}{\lambda}\right) - 1 \right] \quad (88) \end{aligned}$$

A very interesting consequence of this description is that, since the state  $\phi$  is very localized at the atom core, the wavefunction for the total object can be expressed as a superposition of the individual column functions  $\phi_i$  so that Eq. (88) in that case becomes

$$\begin{aligned} \psi(\mathbf{R}, z) = & 1 + \sum_i C_i \phi_i(\mathbf{R} - \mathbf{R}_i) \\ & \times \left[ \exp\left(-i\pi \frac{E}{E_0} \frac{z}{\lambda}\right) - 1 \right] \quad (89) \end{aligned}$$

where the summation runs over all the atomic columns of the object, parallel to the electron beam.

The interpretation of Eq. (89) is simple. Each column  $i$  acts as a channel in which the wavefunction oscillates periodically with depth. The periodicity is related to the 'weight' of the column, that is, proportional to the atomic number of the atoms in the column and inversely proportional to their distance along the column. The importance of these results lies in the fact

that they describe the dynamic diffraction for larger thicknesses than the usual phase grating approximation and that they require only the knowledge of one function  $\phi_i$  per column (which can be tabulated similar to atom scattering factors or potentials). Furthermore, even in the presence of dynamical scattering, the wavefunction at the exit face still retains a one-to-one relation with the configuration of columns for perfect crystals as well as for defective crystals as far as they consist of columns parallel to the electron beam. Hence this description is very useful for interpreting high resolution images. Equation (89) is valid for light columns, such as Si(1 1 1) or Cu(1 0 0) with an accelerating voltage up to about 300 keV. When the atom columns are 'heavier' and/or the accelerating voltage higher or when the atom columns are very close (0.1 nm), the basic concept of channeling remains valid but the expression (89) has to be modified.

Figure 50 shows the electron density  $|\psi(\mathbf{R}, t)|^2$  as a function of depth in an Au<sub>4</sub>Mn alloy crystal for 200 keV incident electrons. The corners represent the projection of the manganese column. The square in the center represents the four gold columns. The distance between adjacent columns is 0.2 nm. The periodicity along the direction of the column is 0.4 nm. From these results it is clear that the electron density in each column fluctuates nearly periodically with depth. For gold this periodicity is about 4 nm, and for manganese 13 nm. These periodicities are nearly the same as for isolated columns so that the influence of neighboring columns in this case is still small. The energies of the respective s states are respectively about 250 and 80 eV.

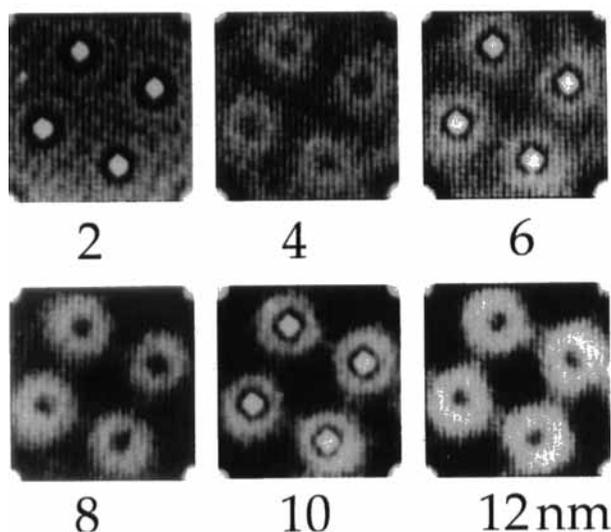


Figure 50. Electron density as a function of depth in  $\text{Au}_4\text{Mn}$  (see text).

It has to be stressed that the derived results are only valid in a perfect zone axis orientation. A slight tilt can destroy the symmetry and excite other, non-symmetric states, so that the results become much more complicated. It is interesting to note that channeling has usually been described in terms of Bloch waves. However, as follows from the foregoing, channeling is not a mere consequence of the periodicity of the crystal but occurs even in an isolated column parallel to the beam direction.

In this view, the Bloch wave approach is only of mathematical importance.

### Diffraction Pattern

Fourier transforming the wavefunction (89) at the exit face of the object yields the wavefunction in the diffraction plane, which can be written as

$$\psi(\mathbf{g}, z) = \delta(\mathbf{g}) + \sum_i \exp(-2\pi i \mathbf{g} \cdot \mathbf{R}_i) F_i(\mathbf{g}, z) \quad (90)$$

In a sense the simple kinematical expression for the diffraction amplitude holds, provided the scattering factor for the atoms is replaced by a dynamical scattering factor for the columns, in a sense as obtained in [42] and which is defined by

$$F_i(\mathbf{g}, z) = \left[ \exp\left(\frac{-i\pi E_i z}{E \lambda}\right) - 1 \right] C_i f_i(\mathbf{g}) \quad (91)$$

with  $f_i(\mathbf{g})$  the Fourier transform of  $\phi_i(\mathbf{R})$ . It is clear that the dynamical scattering factor varies periodically with depth. This periodicity may be different for different columns.

In the case of a monoatomic crystal, all  $F_i$  are identical. Hence  $\psi(\mathbf{g}, z)$  varies perfectly periodically with depth. In a sense the electrons are periodically transferred from the central beam to the diffracted beams and back. The periodicity of this dynamic oscillation (which can be compared with the Pendelösung effect) is called the dynamic extinction distance. It has for instance been observed in  $\text{Si}(111)$ . An

important consequence of Eq. (90) is the fact that the diffraction pattern can still be described by a kinematical type of expression so that existing results and techniques (e.g., extinction rules) that have been based on the kinematical theory remain valid to some extent for thicker crystals in zone orientation.

### Parameterization

Suppose all atom potentials would have the same form multiplied with a constant proportional to the atomic number  $Z$  one could expect that the eigenstate  $\phi(\mathbf{r})$  and its energy  $E$  could scale with  $Z/d$  with  $d$  the repeat distance of atoms in the column. In this way, the wavefunction could be completely parameterized in terms of  $Z/d$  and  $Z$ . Even if Eq. (88) is only qualitatively correct, this would allow the position of each column to be determined very accurately, and the ‘weight’ of the column and the crystal thickness approximately. In practice it has been found that the bound state scales with  $Z^{1/2}$  for all atoms of the periodic table (calculated using Smith and Burge parameters) that  $E$  roughly scales with  $Z/d$ . In this way, a robust parameterization is possible which enables the retrieval of the projected structure from matching with the exit wavefunction [43].

### Appendix E: Phase Retrieval

Consider an image plane at a particular focus value for which we want to retrieve the phase. For convenience we choose the origin of focus in that plane. Writing the wavefunction as a Fourier integral,

we have

$$\psi(\mathbf{R}, 0) = C + \int_{\mathbf{g} \neq 0} \phi(\mathbf{g}) \exp(2\pi i \mathbf{g} \cdot \mathbf{R}) d\mathbf{g} \quad (92)$$

where  $\phi(\mathbf{g})$  are the Fourier components. We have separated the zeroth-order component (constant term).

The wavefunction at a particular focus value  $z$  can be obtained from Eq. (92) by propagation, that is,

$$\psi(\mathbf{R}, z) = C + \int_{\mathbf{g} \neq 0} \phi(\mathbf{g}) \exp(2\pi i \mathbf{g} \cdot \mathbf{R}) \times \exp(-i\pi \lambda g^2 z) d\mathbf{g} \quad (93)$$

Three-dimensional Fourier transforming the image intensity  $|\psi(\mathbf{R}, z)|^2$  now yields, using Eq. (93),

$$\begin{aligned} I(\mathbf{g}, \xi) = & |C|^2 \delta(\mathbf{g}) + C^* \phi(\mathbf{g}) \delta\left(\xi - \frac{\lambda g^2}{2}\right) \\ & + C \phi^*(-\mathbf{g}) \delta\left(\xi + \frac{\lambda g^2}{2}\right) \\ & \times \int_{\substack{\mathbf{g}' \neq 0 \\ \mathbf{g} + \mathbf{g}' \neq 0}} d' \mathbf{g} \phi^*(\mathbf{g}') \phi(\mathbf{g} + \mathbf{g}') \\ & \times \delta\left(\xi - \frac{\lambda}{2} [(\mathbf{g} + \mathbf{g}')^2 - g'^2]\right) \end{aligned} \quad (94)$$

where  $\delta$  are Dirac functionals and  $\mathbf{g}$  and  $\xi$  are the conjugates of  $\mathbf{R}$  and  $z$ . The first term on the right-hand side of Eq. (94) only yields a contribution in the origin. The second and third linear terms give a sharply peaked contribution which is located on a paraboloid in reciprocal space which can be considered as the Ewald sphere in a vacuum (Fig. 18). The last term gives a contribution which is more continuously spread through reciprocal space. It is immediately clear that by selecting the information concentrated on the

paraboloid one directly obtains  $\phi(\mathbf{g})$  and  $\phi^*(\mathbf{g})$ , so that from Eq. (92) the total wavefunction at focus 0 is retrieved. In principle, this can be done by taking a nearly continuous series of images at very small focus intervals, three-dimensional Fourier transforming, and selecting the sphere. However, this procedure is rather impractical. Hence we proceed as follows. We take a series of images at focus values  $z_1, z_2, z_3, \dots$ . The focus interval is of the order of 10 nm. Each of the images  $I(\mathbf{R}, z_n)$  is then transformed into  $I(\mathbf{g}, z_n)$ . Finally we calculate the series

$$\sum_n I(\mathbf{g}, z_n) \exp(-i\pi\lambda g^2 z_n)$$

In this way, the sphere for  $\xi > 0$  is in a sense projected in the plane, apart from a known weighting factor. We can do the same for  $\xi < 0$ . From these data it is easy to calculate  $\phi(\mathbf{g})$ . In a sense, all the images are back propagated to zero focus, where the linear part of each image superimposes and increases with respect to the nonlinear part. However, as seen from Eq. (94), the integral also gives a contribution to the sphere which may influence the results. This contribution can be corrected for by using Eq. (94) in an iterative way. Another advantage of this method is that it is relatively easy to compensate for the effect of chromatic aberration. It is well known that chromatic aberration results from a spread in the focus due to instabilities in voltage and lens current. Hence the image intensity is convoluted with a focal spread function  $I(\mathbf{R}, z) = I_0(\mathbf{R}, z) * f(z)$ . In reciprocal space, the convolution product is a direct product with the Fourier transform of  $f(z)I(\mathbf{g}, \xi) = I_0(\mathbf{g}, \xi)f(\xi)$ . If the spread function  $f$  is known, it is easy to compensate for this effect by dividing  $f(\xi)$ . Since

this has to be done only at the sphere, blowing up effects are largely reduced. Another advantage of the method is that, since the contribution of the noise is more homogeneously distributed in space, the selection of the sphere automatically increases the signal-to-noise ratio. In practice the whole procedure is performed in two steps. A first approximation of the exit wave is obtained by the paraboloid method described above; then this result is refined using a maximum likelihood method as proposed by Coene et al. [16].

## Acknowledgments

The author wishes to thank J. Janssen, M. Op de Beeck, D. Tang, A. Thust, J. Van Landuyt and G. Van Tendeloo for the use of material and/or stimulating discussion.

## 1.1.2.9 References

- [1] O. Scherzer, *J. Appl. Phys.* **1949**, 20, 20; J. C. H. Spence, *Experimental High Resolution Electron Microscopy*, Clarendon Press, Oxford **1988**.
- [2] D. Van Dyck, A. F. De Jong, *Ultramicroscopy* **1992**, 47, 266; A. F. De Jong, D. Van Dyck, *Ultramicroscopy* **1993**, 49, 66; D. Van Dyck, M. Op de Beeck, W. M. J. Coene, *MSA Bull.* **1994**, 24(2), 427.
- [3] H. Rose, *Proc. XIIth International Congress on Electron Microscopy (Seattle)*, San Francisco Press, San Francisco **1990**, pp. 202–203; H. Rose, *Optik* **1990**, 85, 19.
- [4] W. O. Saxton, *J. Microsc.* **1995**, 179(2), 210.
- [5] O. Krivanek, *Ultramicroscopy* **1994**, 55, 419.
- [6] D. W. Robards, A. J. Wilson, *Procedures in Electron Microscopy*, Wiley, Chichester **1993**.
- [7] J. M. Cowley, S. Iijima, *Z. Naturforsch.* **1972**, 27a(3), 445.
- [8] D. F. Lynch, A. F. Moodie, M. A. O'Keefe, *Acta Crystallogr.* **1975**, A31, 300.
- [9] D. Van Dyck, G. Van Tendeloo, S. Amelinckx, *Ultramicroscopy* **1982**, 10, 263; S. Amelinckx, G. Van Tendeloo, J. Van Landuyt, *Bull. Mater. Sci.* **1984**, 6(3), 417; S. Amelinckx, G. Van Tendeloo, J. Van Landuyt, *Ultramicroscopy* **1985**, 18, 395.

- [10] C. Kisielowski, P. Schwander, F. H. Baumann, M. Seibt, Y. Kim, A. Ourmadz, *Ultramicroscopy* **1994**, 58, 131.
- [11] S. Amelinckx, *Acta Crystallogr.* **1995**, B51, 486.
- [12] P. A. Stadelman, *Ultramicroscopy* **1987**, 21, 131; Anon, *Multislice Computer Program*, High Resolution Electron Microscope Facility, Arizona State University, Tempe, AZ; R. Kilaas, R. Gronsky, *Ultramicroscopy* **1982**, 11, 289; D. Van Dyck, W. Coene, *Ultramicroscopy* **1984**, 15, 29; D. Van Dyck, W. Coene, *Ultramicroscopy* **1984**, 15, 41; D. Van Dyck, W. Coene, *Ultramicroscopy* **1984**, 15, 287.
- [13] R. Bierwolf, M. Hohenstein, *Ultramicroscopy* **1994**, 56, 32.
- [14] J. Jansen, D. Tang, H. W. Zandbergen, M. Schenk, *Acta Cryst. A*, **1996**, in press.
- [15] P. Schishe in *Image Processing of Computer-aided Design in Electron Optics* (Ed.: P. Hawkes) **1973**; W. O. Saxton in *Proceedings of the XIth International Congress on Electron Microscopy, Kyoto*.
- [16] D. Van Dyck in *Proceedings of the XIIth International Congress on Electron Microscopy (Seattle)*, San Francisco Press, San Francisco **1990**, pp. 26–27; W. Coene, G. Janssen, M. Op de Beeck, D. Van Dyck, *Phys. Rev. Lett.* **1992**, 29, 37; M. Op de Beeck, D. Van Dyck, W. Coene in *Electron Holography* (Eds.: A. Tonomura et al.), North-Holland/Elsevier, Amsterdam **1995**, pp. 307–316.
- [17] M. A. O. Miedema, A. H. Buist, A. Vandenbos, *IEEE Trans. Instrum. Measur.* **1994**, 43(2), 181.
- [18] D. Tang, H. Zandbergen, J. Jansen, M. Op de Beeck, D. Van Dyck, *Ultramicroscopy* (in press).
- [19] A. Thust, K. Urban, *Ultramicroscopy* **1992**, 45, 23.
- [20] D. Broddin, G. Van Tendeloo, S. Amelinckx, *J. Phys. C: Condens. Matter* **1990**, 2, 3459.
- [21] M. Takeda, G. Van Tendeloo, S. Amelinckx, *Acta Crystallogr.* **1988**, A44, 938.
- [22] H. Alexander, J. C. H. Spence, D. Shindo, H. Gottschalk, N. Long, *Phil. Mag.* **1986**, 53(5), 617.
- [23] S. Kuypers, G. Van Tendeloo, S. Amelinckx, H. W. Shu, S. Faulmes, J. Flahaut, P. Laruelle, *J. Solid State Chem.* **1988**, 73, 192.
- [24] S. Kuypers, J. Van Landuyt, S. Amelinckx, *J. Solid State Chem.* **1990**, 86, 212.
- [25] R. Bontchev, B. Darrut, J. Parriat, F. Weill, G. Van Tendeloo, S. Amelinckx, *Eur. J. Solid State Inorg. Chem.* **1993**, 30, 521.
- [26] H. W. Zandbergen, W. A. Groen, F. C. Mijl-hoff, G. Van Tendeloo, S. Amelinckx, *Physica* **1988**, C156, 325.
- [27] T. Krekels, G. Van Tendeloo, S. Amelinckx, *Solid State Commun.* **1991**, 79, 607.
- [28] L. C. Nistor, G. Van Tendeloo, S. Amelinckx, *J. Solid State Chem.* **1994**, 109, 152.
- [29] S. Muto, G. Van Tendeloo, S. Amelinckx, *Phil. Mag.* **1993**, B67, 443.
- [30] X. F. Zhang, X. B. Zhang, D. Bernaerts, G. Van Tendeloo, S. Amelinckx, J. Van Landuyt, H. Werner, *Ultramicroscopy* **1994**, 55, 25.
- [31] S. Amelinckx, D. Bernaerts, X. B. Zhang, G. Van Tendeloo, J. Van Landuyt, *Science* **1994**, 267, 1334; S. Amelinckx, X. B. Zhang, D. Bernaerts, X. F. Zhang, V. Ivanov, J. B. Nagy, *Science* **1994**, 265, 635.
- [32] L. Nistor, G. Van Tendeloo, S. Amelinckx, C. Cros, *Phys. Status Solidi (a)* **1994**, 146, 119.
- [33] J. Frank, *Optik* **1973**, 38, 519; P. L. Fejes, *Acta Crystallogr.* **1977**, A33, 109.
- [34] M. Born, E. Wolf, *Principles of Optics*, Pergamon Press, London **1975**, Chap. X.
- [35] D. Van Dyck, *Adv. Electron. Electron Phys.* **1985**, 65, 295.
- [36] J. M. Cowley, A. F. Moodie, *Acta Crystallogr.* **1957**, 10, 609.
- [37] K. Ishizuka, N. Uyeda, *Acta Crystallogr.* **1977**, A33, 740.
- [38] D. Van Dyck, W. Coene, *Ultramicroscopy* **1984**, 15, 29.
- [39] J. Lindhard, *Mat. Fys. Medd. Dan. Vid. Selsk* **1965**, 34, 1; A. Tamura, Y. K. Ohtsuki, *Phys. Status Solidi (b)* **1974**, 73, 477; A. Tamura, F. Kawamura, *Phys. Status Solidi (b)* **1976**, 77, 391; B. Buxton, J. E. Loveluck, J. W. Steeds, *Phil. Mag.* **1978**, A3, 259.
- [40] K. Kambe, G. Lehmppfuhl, F. Fujimoto, *Z. Naturforsch.* **1974**, 29A, 1034; D. Van Dyck, J. Danckaert, W. Coene, E. Selderslaghs, D. Broddin, J. Van Landuyt, S. Amelinckx in *Computer Simulation of Electron Microscope Diffraction and Images* (Eds.: M. Krakow, M. O'Keefe), TMS Publications, MMM Society **1989**, pp. 107–134; D. Van Dyck, M. Op de Beeck in *Electron Holography* (Eds.: A. Tonomura et al.), North-Holland, Amsterdam **1994**, pp. 297–307; M. V. Berry, K. E. Mount, *Rep. Progr. Phys.* **1972**, 35, 315; B. F. Buxton, J. E. Loveluck, J. W. Steeds, *Phil. Mag.* **1978**, 38(3), 259.
- [41] K. Kambe, G. Lempfuhl, F. Fujimoto, *Z. Naturforsch.* **1974**, 29a, 1034.
- [42] D. Shindo, M. Hirabayashi, *Acta Crystallogr.* **1988**, A44, 954.
- [43] D. Van Dyck, M. Op de Beeck, *Electron Holography* (Eds.: A. Tonomura, L. F. Allard, G. Pozzi, D. C. Joy, Y. A. Ono), Elsevier Science B.V. **1995**, p. 297.
- [44] S. Amelinckx, D. Van Dyck in *Electron Diffraction Techniques* (Ed.: J. M. Cowley), Vol. 2, *IUCr Monographs on Crystallography*, International Union of Crystallography, Oxford University Press, Oxford **1992**, pp. 1–222.
- [45] H. Alexander, J. C. H. Spence, D. Shindo, H. Gottschalk, N. Long, *Phil. Mag.* **1986**, A53(5), 627.

## 1.2 Reflection Electron Microscopy

### 1.2.1 Introduction

In the early days of electron microscopy, Ruska [1] attempted to obtain images by reflecting electrons from the surfaces of solids, with limited success. The idea was taken up by several groups in the 1950s [2, 3] and images of surfaces showing moderately good resolution were obtained. However, in order to get useful image intensity, it was necessary to use electrons scattered at small angles, 5 to 10 degrees, from the surface so that the images suffered from severe foreshortening. When SEI (secondary electron imaging) was developed and gave comparable resolution with normal beam incidence and no foreshortening, the reflection mode fell out of favor. The reason why it was later revived was that it was realized that, if the images were obtained using strong diffracted beams from crystal surfaces, the image contrast could be highly sensitive to small changes in crystal structure, orientation or composition of the top few atomic layers of the surface. Thus it could give information, not available from SEI, of great value in surface science.

The first images obtained in 100 keV electron microscopes using strong diffracted beams confirmed the sensitivity of the method to surface structure [4, 5]. Bulk

crystals were mounted in the usual specimen stage of the microscope so that the electron beam was incident on a flat crystal surface at a grazing angle of a few degrees. The reflection high energy electron diffraction (RHEED) pattern could be observed in the selected area electron diffraction (SAED) mode and a strong diffraction spot (usually the specular beam) could be chosen to pass through the objective aperture and form the image. The reflection electron microscopy (REM) imaging mode could then be likened to that for the so-called 'high-resolution' mode of dark-field imaging in transmission. For a diffraction angle of  $\theta$ , the incident beam is tilted by an angle of  $2\theta$  to the axis of the objective lens and the crystal face is tilted at an angle  $\theta$ , so that the diffracted beam passes down the axis through a centered objective aperture.

Also it was realized that, as in the case of transmission microscopy, an equivalent scanning form of imaging (SREM) was possible. Initially, it was demonstrated that useful images could be obtained with medium-energy (5–15 keV) electrons in the scanning mode when the small electron probe from a field-emission gun was scanned over a surface and a strong diffracted beam was detected to form the image [6, 7].

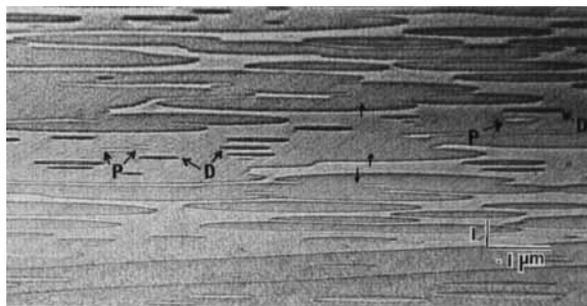
For meaningful research in surface science it is necessary, for most materials,

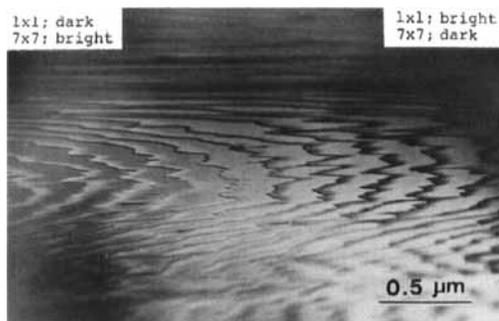
to study surfaces prepared and maintained under ultrahigh vacuum (UHV) conditions. For imaging with moderate resolution and medium-energy electrons this may be achieved by adding a field-emission gun to an ultrahigh vacuum chamber with only a simple lens system to focus the electron probe on the specimen. Standard specimen manipulators could be used and the reflection imaging could be combined with an imaging of the surface using secondary or Auger electrons or analysis of the surface composition using AES or other methods. A number of useful systems have been developed along these lines and applied to surface structure problems [8–10]. However, the spatial resolution achievable in such systems is limited to 5–10 nm. To achieve normal electron microscope resolutions under clean surface conditions, it is necessary to build a UHV specimen environment into a TEM instrument. By doing this, and providing a variety of devices for in situ specimen treatment, Yagi and coworkers [11] showed that images of clean surfaces could be obtained with resolutions approaching 1 nm and other groups have done likewise [12]. Later it was demonstrated that a resolution of almost 0.3 nm could be achieved in a 1 MeV electron microscope [13].

The reflection electron microscopy images obtained in electron microscopes at 100 keV or more make use of diffracted beams making angles of only a few degrees with the surface and so are foreshortened by a factor of 20 to 50 times. The magnification and resolution of the images in the direction of the incident beam are then severely limited and the high resolution of the images is possible only in the direction across the surface, perpendicular to the beam. However, the value of the technique is immediately apparent, in spite of this disadvantage. Steps on a crystal surface only one atom high are clearly visible, as in Fig. 1. Strong contrast is shown between surface regions that differ only in the structure of the first one or two monolayers of atoms, as seen in Fig. 2. Small bumps, dips, impurities or projections on the surfaces become obvious (see Figs. 8 and 10). The changes in surface structure or composition resulting from in situ treatments can be followed in detail. Also the images may be correlated with RHEED patterns revealing the crystal structure and with microanalysis using EELS [14] or EDS to determine the chemical composition of surfaces with high spatial resolution.

REM images can be obtained using any TEM instrument if a bulk specimen is

**Figure 1.** REM image of one-atom high surface steps on a Pt (111) surface. Arrows indicate the step-down direction. Circular features, foreshortened to thin ellipses, are atom-high projections, P, or depressions, D. Note the intensity changes across the steps. (Courtesy of T. Hsu.)





**Figure 2.** REM image of Si (111) face showing the formation of the  $7 \times 7$  superstructure on the top side of surface steps. Differences in diffraction conditions give differences in contrast for the two sides of the image. (Reprinted with permission of K. Yagi [60].)

mounted so that the flat face of a crystal is almost parallel to the incident electron beam [15]. For many specimens such as noble metals and some refractory oxides the vacuum of about  $10^{-7}$  torr of standard microscopes is not a serious hindrance to much research concerned with the structures of surfaces although, when chemical or physical surface reactions are to be studied, extreme precautions must be taken to avoid complications due to contamination of the surface. For semiconductors, most metals and other inorganic compounds, UHV is essential in most cases and access to one of the few microscopes modified for UHV operation is required. High voltage microscopes, operating at up to 1 MeV, can give excellent results [13] even though the foreshortening of the images is increased because of the smaller diffraction angles. Scanning transmission electron microscopy (STEM) instruments, which normally have a vacuum better than  $10^{-8}$  torr, may also be used with the incident beam scanned over the surface at grazing incidence to give SREM images [16].

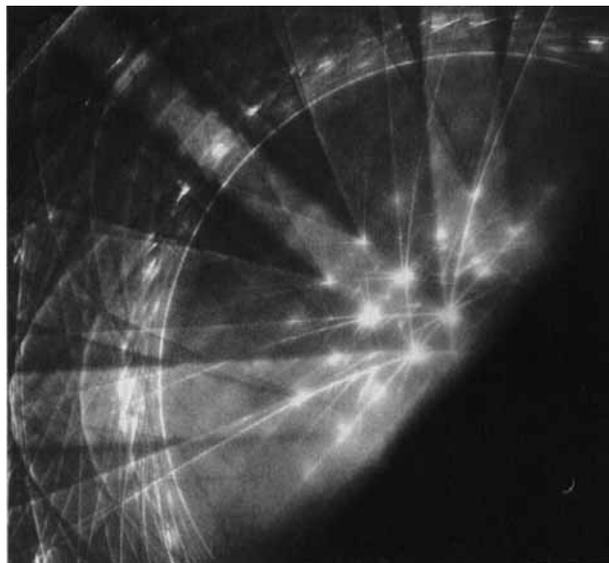
The main limitation on the resolution of REM images comes from the large energy spread of the diffracted electrons [17]. For the usual RHEED geometry, the average distance traveled in the crystal by an emerging 100 keV electron may be 100 nm or more; greater than the path-length for inelastic scattering processes such as the excitation of plasmons or the excitation of outer-shell or valence-band electrons of the solid, with energy losses of 10 to 30 eV. Also it may be noted that the electron beam traveling almost parallel to a specimen surface can excite surface plasmons while it is within 1 or 2 nm of the surface, on either side, and the energy losses from this source are in the range of 5 to 20 eV. The strong RHEED diffraction spots may contain 50 to 80% of electrons which have lost more than 3 eV of energy [18]. The average energy loss is of the order of 30 eV and the energy spread is about the same. The chromatic aberration of the objective lens in a 100 keV microscope then limits the resolution to about 1 nm. Better resolution can be obtained by use of higher voltage microscopes (up to 1 MeV) since the loss of resolution from this cause depends on the ratio of the energy loss to the initial electron energy. Otherwise, both the resolution and the contrast of REM images may be improved by use of an energy filter, such as is standard equipment for STEM instruments and is now available for some TEM instruments.

## 1.2.2 Reflection High Energy Electron Diffraction Patterns

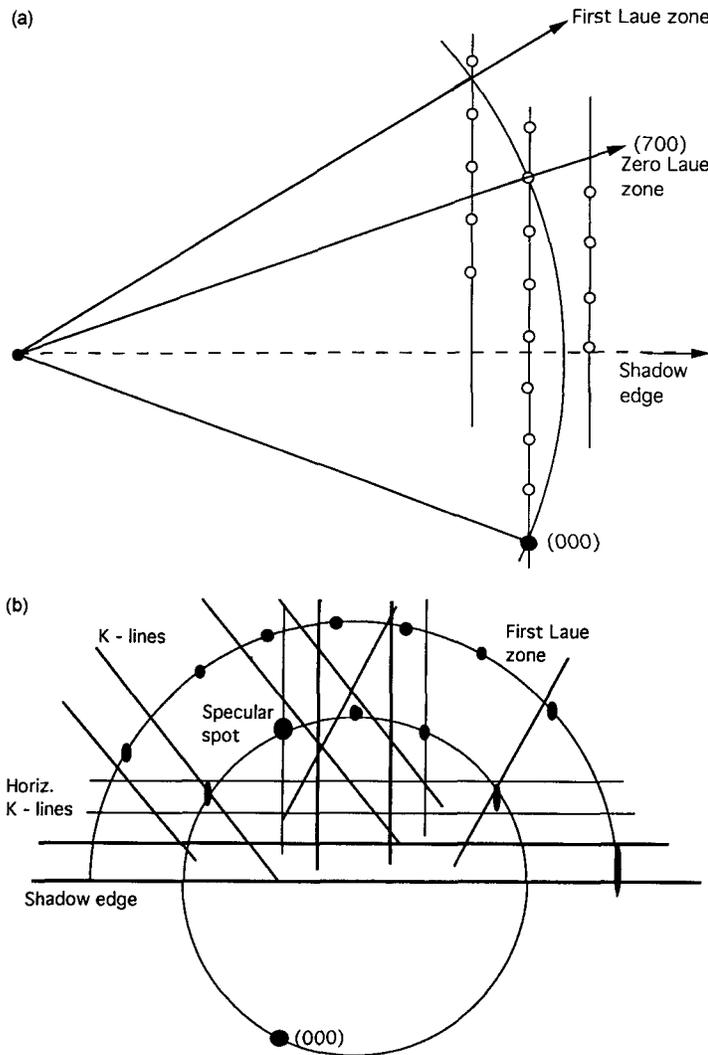
The RHEED pattern produced when a high-energy electron beam strikes a crystal

surface at grazing incidence includes sharp strong diffraction spots lying on a set of concentric circles known as Laue circles, corresponding to the Laue zones in which the Ewald sphere cuts the set of parallel planes of reciprocal lattice points in reciprocal space. In practice, the strong spots do not appear sharp because they are normally grossly overexposed when recorded on photographic film and greatly broadened by halation. Also, with the overexposure, the diffuse background around the spots, arising from thermal diffuse scattering and from the small-angle energy-loss electron distribution, blackens the film and makes the spots seem large. In the background, there is a pattern of Kikuchi lines (sometimes referred to simply as K-lines) resulting from the diffraction of diffusely scattered electrons by the crystal lattice planes. The K-line patterns are accompanied in many cases by sets of parabolas and circles (Fig. 3).

The configuration of the sharp Bragg spots can be derived from a reciprocal space diagram such as Fig. 4a. Because the crystal is terminated at a surface, and most of the diffraction occurs in a thin surface layer, continuous lines of scattering power, perpendicular to the surface, pass through the sharp maxima around the reciprocal lattice spots. Hence a sharp spot is generated whenever the Ewald sphere cuts one of these lines, giving rise to sets of spots lying on the Laue circles, as suggested in Fig. 4b. The spots are very bright when the Ewald sphere cuts close to a reciprocal lattice point. The spot most commonly used for REM imaging is the specular reflection, given by the mirror image of the incident beam in the plane of the crystal surface. This spot is most intense when the incident beam is at the Bragg angle for lattice planes parallel to the surface. No spots appear below the 'shadow edge' which marks the intersection of the surface plane with the recording



**Figure 3.** RHEED pattern from the (110) surface of GaAs showing Bragg spots, K-lines, parabolas and circles. (Reprinted with permission of K. Yagi [61].)



**Figure 4.** The geometry of RHEED patterns. (a) Section of reciprocal space in the plane of the incident and diffracted beams showing the intersection of the Ewald sphere with the reciprocal lattice. (b) The resulting pattern of spots and K-lines.

plate unless, as sometimes happens, there is transmission diffraction through small projections or edges of the crystal close to the termination of the crystal face.

Energy-loss processes involving the excitation of plasmons and single-electron excitations give rise to small-angle scattering ( $10^{-3}$  radians or less). The combination of such inelastic scattering with the elastic scattering gives the peaks of diffuse

scattering around each Bragg reflection spot. The thermal diffuse scattering, which involves energy losses of the order of only  $10^{-2}$  eV, gives high-angle scattering to  $10^{-1}$  radians or more but is also peaked around the Bragg spots. Multiple scattering occurs for each process separately and between processes, so that the RHEED pattern has a broad diffuse background of, mainly, inelastically scattered electrons.

These electrons are diffracted by the crystal lattice, giving rise to the K-line patterns. These are analogous to the Kossel patterns produced when X-rays, generated inside a crystal, are diffracted by the lattice planes. However, in the case of high energy electrons, the wavelengths and diffraction angles are so small that, instead of the strongly curved Kossel lines, the K-lines produced appear as parallel sets of straight lines. The separation of a pair of parallel lines is twice the Bragg angle for diffraction from a set of crystal lattice planes, since a line is generated when the electrons are incident on the lattice planes at the Bragg angle from one side or the other. Multiple sets of parallel lines are generated, by several orders of diffraction from the same planes (see Figs. 3 and 4b). Usually a set of sharp horizontal K-lines is seen, parallel to the shadow-edge, corresponding to the diffraction from planes parallel to the crystal surface. The relative intensities and detailed intensity distributions of K-lines have been described in terms of dynamical diffraction theory for high-energy electrons by various authors [19].

Sets of K-lines, corresponding to sets of planes such as those associated with the reciprocal lattice points on a common reciprocal lattice plane, may be tangential to a common asymptotic circle or parabola. Strong parabolas and circles are seen in K-line patterns, close to, but slightly displaced towards the shadow edge from, these asymptotic curves. These parabolas and circles may be ascribed to a resonance phenomenon in which the diffusely scattered electrons are strongly confined to sets of planes or rows of atoms parallel to the crystal surface [20, 21]. When a convergent incident

beam is used, sets of parabolas can also appear corresponding to channelling of the incident-beam electrons along planes perpendicular to the surface [22].

The resonance condition for planes parallel to the surface corresponds to the geometry for which a strong specular beam is produced because the incident beam is at the Bragg angle for planes parallel to the surface and also a strong diffracted beam is produced so that it runs almost parallel to the surface. This condition is indicated when the specular beam sits at the intersection of a horizontal K-line and either a strong inclined K-line, or, for diffusely scattered electrons, a strong parabola. Then the intensity of the specular reflection, and of the whole diffraction pattern, is greatly enhanced, so that this condition is often sought to produce high-intensity REM images. The resonance condition sometimes involves a channeling phenomenon in which the electron wave is channeled along the top one or two top surface layers of atoms and penetrates very little into the bulk of the crystal [23, 24]. Then the REM image becomes highly sensitive to the structure and composition of the top surface layers and the image contrast for surface defects is modified in ways which will be described later.

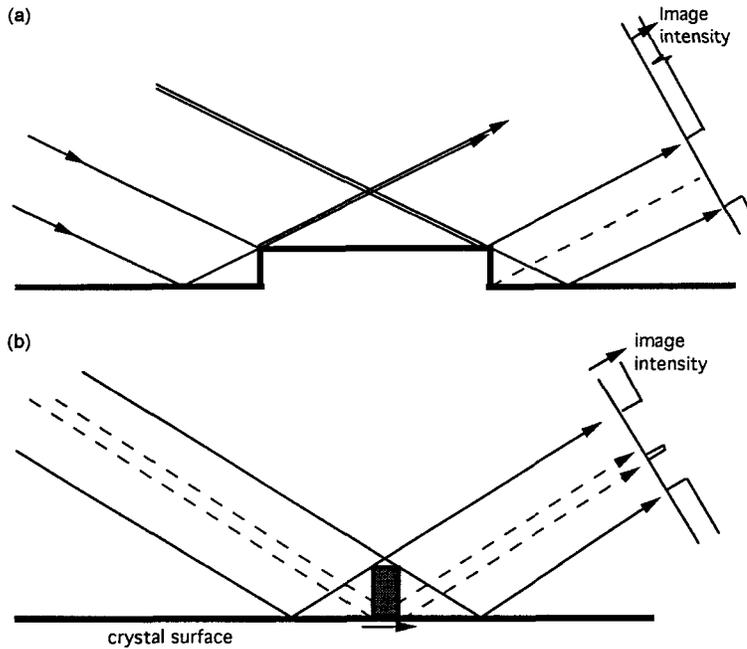
The geometry of the pattern of K-lines and circles and parabolas depends only on the orientation of the crystal relative to the recording film and is independent of the orientation of the incident beam since the pattern is derived from diffusely scattered electrons which may be considered as generated at points within the crystal. As a crystal face is rotated in azimuth, the K-line pattern rotates with it and hence gives a very sensitive indication of crystal

orientation. However, the incident beam direction determines the overall pattern of intensities of the scattered electrons and so governs the relative intensities and visibilities of the various lines and curves.

Because the refractive index of solids for high-energy electrons is slightly greater than unity, there is a refractive bending of all diffracted beams leaving a flat surface at a small angle, resulting in a displacement of all RHEED spots and lines toward the shadow edge. For marginal cases, the displacement may be as great as  $10^{-2}$  radians but the displacement decreases rapidly for higher angles of diffraction. This introduces a distortion of the geometry of the RHEED pattern. The refraction effect also leads to a total internal reflection of some diffracted beams directed almost parallel to the surface, often contributing to the surface resonance effect.

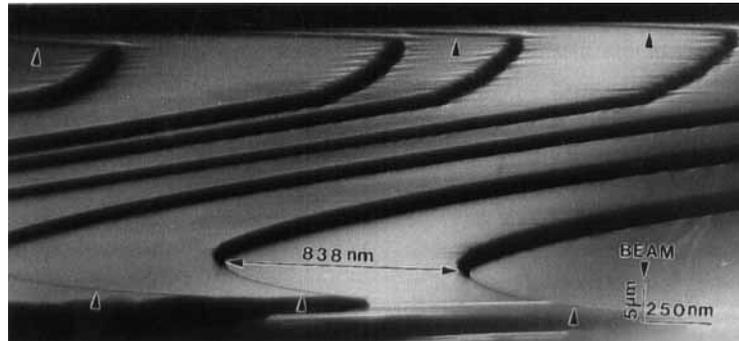
### 1.2.3 Image Contrast for Surface Steps and Projections

For steps or projecting particles on an otherwise perfectly flat crystal face which are large compared with the resolution limit of the microscope, the REM image contrast can be described, to a good approximation, in terms of simple geometric optics. Fig. 5a, for example, illustrates the case for large steps for an exaggerated angle of incidence. For a down-step (as seen when looking towards the electron source) there is no discontinuity in the intensity diffracted from the crystal surface before the step and from the top of the step, so that the step is invisible apart from minor perturbations due to some transmission through the step edge. For an up-step, however, no electrons reach the image from the exit face of the



**Figure 5.** Geometric-optics diagrams of the diffracted beams from a large projection (a) and a thin projection or small particle, on a crystal surface and the resulting image intensities (b).

**Figure 6.** REM image of large curved steps and small intersecting steps (arrowed) on an oxygen-annealed rutile (1 0 0) surface [62].

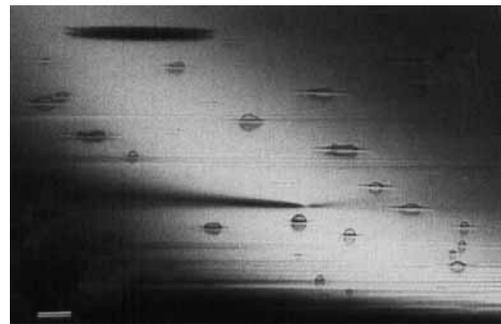


step or from a region of the crystal surface after the step which is shaded from the incident beam. A black region therefore appears in the image of width equal to roughly twice the step height. The appearance of the step varies with the angle that the step makes with the incident beam direction. When the step becomes parallel to the incident beam direction, the width of the black shadow decreases to zero, and the form of the step is seen in profile. The consequent appearance of large curved steps in the image is as illustrated in Fig. 6.

For a projection from the surface which is thin in the beam direction, or for a small foreign particle sitting on the surface, the corresponding geometric-optics construction is as shown in Fig. 5b. The projection is illuminated by the electrons diffracted from the surface behind it (as seen when looking towards the electron source) and so forms an image as in transmission microscopy. The projection also prevents the incident beam from being diffracted from the surface in front of it. Hence the image of the projection is accompanied by an inverted mirror-image, as is evident in Fig. 7.

A common feature of such double images for particles on crystal surfaces is that the particle image and the mirror

image are separated by a bright line. This effect has been attributed to a channeling phenomenon. Under the surface resonance conditions, electrons may be channeled to flow along the top one or two layers of atoms on the surface. Under this condition they may travel for unusually large distances along the surface before being diffracted out of the surface. The presence of a particle on the surface may actually enhance this effect since the presence of foreign atom layers on the surface may decrease the probability of the electrons escaping from the surface layers. Hence the electrons may pass underneath the



**Figure 7.** REM image of small transparent foreign particles on an Si (1 1 1) surface showing the direct image and mirror image and the bright line between them in each case (plus a dislocation strain field). Scale bar: approx. 50 nm. (Courtesy of T. Hsu.)

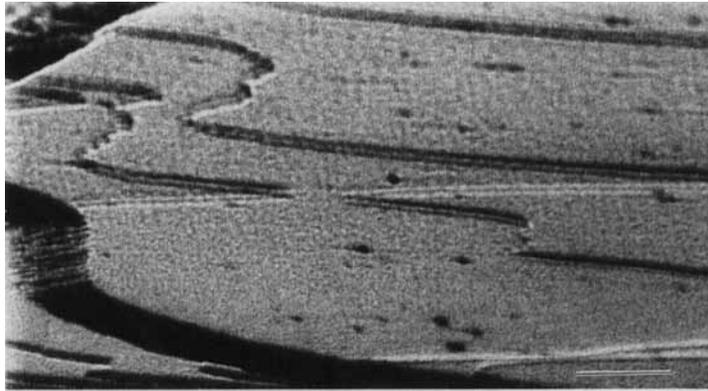
particle with little decrease of intensity and give a bright line in the image, as suggested in Fig. 5b.

For very low steps on a surface, the step height is often equal to the thickness of one or two monolayers of atoms, which is less than the resolution limit for most REM imaging. Then the geometric-optics pictures used for high steps can no longer apply and the strong contrast normally seen must be attributed to a phase-contrast mechanism. The two beams labeled 1 and 2 in Fig. 5a, coming from before the step and on top of the step, overlap and give rise to interference effects depending on their relative phase. For a step height  $h$  and with equal angles of incidence and reflection of the beams with the surface equal to  $\theta_B + \Delta$ , where  $\theta_B$  is the Bragg angle (and ignoring refraction effects) the beams have a phase difference of  $(2\pi/\lambda) \cdot 2\theta_B h(1 + \Delta/\theta_B)$  [25]. This is equal to  $2\pi n(1 + \Delta/\theta_B)$  if the step height is  $n$  times the crystal plane spacing. If this phase difference is small, or close to a multiple of  $2\pi$ , the contrast given by the step is the same as that at the edge of a weak phase object in a TEM image, that is, the contrast is zero for the objective lens exactly focused on the step (if the effect of the spherical aberration of the lens is ignored) but the out-of-focus images show a black-white contrast which reverses from overfocus to underfocus and gets stronger and broader as the amount of defocus increases. As for Fresnel fringes, the width of the dark and bright lines is proportional to the square root of the defocus.

Because of the glancing angle of incidence, the amount of defocus occurring in a normal REM image of an extended flat crystal face may be very large (many

micrometers, positive and negative) so that strong contrast is seen over most of the field of view. If the step is an up-step rather than a down-step the same argument applies except that the sign of the phase difference is reversed and hence the dark-light contrast of the step is reversed. Hence an analysis of the image contrast for a through-focus series of images gives an immediate indication of whether the steps go up or down [11].

It is readily shown that if the incident beam is at the Bragg angle for the lattice planes parallel to the surface and the step height is equal to the lattice plane spacing, the phase difference of the beams scattered before and after the step is a multiple of  $2\pi$ , giving no contrast. However, this can only occur if it can be assumed that the refraction effect is negligible, that the incident beam is exactly at the Bragg angle and that there is no variation of the lattice plane spacing at the step due to elastic strain effects. In most cases, if it is assumed that the potential function drops sharply at the crystal surface from the value in vacuum to the inner-potential value inside the crystal, the refraction effect at the surface ensures that the angle of incidence of the beam in vacuum is less than the Bragg angle for strong reflection within the crystal. Then the deviation,  $\Delta$ , from the Bragg angle in the above equation may be quite large. Also, for many materials, it is thought that there is a relaxation effect around a step so that the step height deviates from the bulk interplanar spacing by a few percent. In either of these cases, or if the image is obtained at an incident angle which is not exactly the Bragg angle, the step is no longer equivalent to a weak phase object. A second order approximation then suggests that a sharp black line



**Figure 8.** REM image of a Pt (111) surface, plus some small evaporated Au particles, showing the doubling of the images of small steps [63]. Scale bar: 40 nm.

may appear at the in-focus position, and for the images of the step with moderate amounts of defocus the antisymmetric dark-light contrast may be strongly perturbed [25].

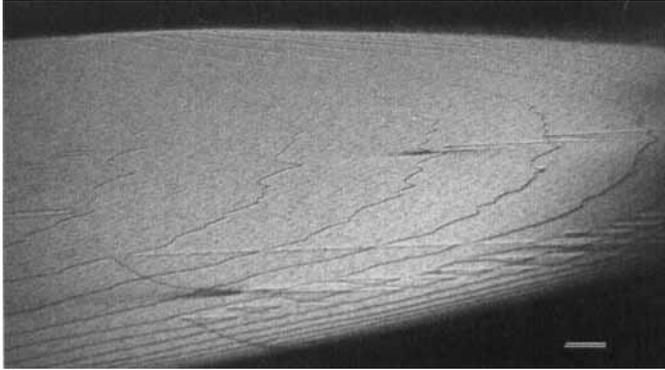
In some cases the in-focus images of small steps appear as pairs of dark or light lines with a separation, in the image plane, of about 3 nm (Fig. 8). The step image doubling persists for out-of-focus images until it is swamped by the usual dark-light defocus contrast. This effect has been variously attributed to dynamical diffraction effects [26] or to strain fields due to lattice relaxation around the step [27]. However, a systematic investigation has shown that the effect is associated with the occurrence of surface resonance [28]. The doubling has been simulated by Anstis [29] using many-beam dynamical diffraction calculations.

### 1.2.4 Dislocations and Stacking Faults

Because the contrast in an REM image derives from the variation of intensity of a

diffracted beam, the small changes of lattice orientation associated with local strain fields may be clearly visible. When a dislocation line emerges from the bulk and intersects a crystal surface, the associated strain fields distort the surface and so give rise to characteristic patterns of dark and light lobes which, when foreshortened, appear as streaks, extending across the surface, nearly perpendicular to the incident beam [30]. The sensitivity to lattice strain is such that the streaks may extend for distances of several micrometers (Fig. 9).

The nature of the dislocation may be derived from the configuration of the streaks and their variation with the diffraction conditions, in much the same way as in TEM [31]. For a pure screw dislocation normal to the surface, the surface distortion is similar to that of the lattice planes in bulk. When viewed from the incident beam direction, the surface lattice planes are tilted up on one side and down on the other side of the dislocation core. If the bulk crystal is exactly at the Bragg angle, a tilt in either direction decreases the intensity and gives a dark streak on each side. If the incident beam is slightly off the Bragg angle, the intensity is increased on one



**Figure 9.** REM image of a Pt (111) surface showing slip trace steps from moving dislocations interacting with growth steps. Contrast from the dislocation strain fields appears at the terminations of the slip traces. Scale bar: approx. 50 nm. (Courtesy of T. Hsu.)

side and decreased on the other so that the dislocation gives an antisymmetrical, black–white pair of streaks.

An emerging edge dislocation gives a more complicated pattern of streaks. The well-known rule that the contrast vanishes for  $\mathbf{g} \cdot \mathbf{b} = 0$ , where  $\mathbf{g}$  is the diffraction vector and  $\mathbf{b}$  is the Burgers vector, holds for two-beam conditions, as in TEM, so that no contrast should be given for an edge dislocation perpendicular to the surface when imaged with a specular reflection. Even for this case, contrast may appear because the strain field varies the azimuthal rotation of the lattice relative to the beam and an azimuthal rotation can strongly influence the amplitudes of non-specular beams which may, in turn, affect the specular beam intensities through many-beam dynamical diffraction effects. Shuman [30] showed that edge dislocations can give images with four dark and light streaks.

When dislocations run parallel to a surface but below it, as in the case of dislocation loops which are wholly or partially buried, their strain fields may extend to the surface and perturb the orientations of the surface layers. Then diffuse, curved lines are seen in the REM image [32].

Frequently the contrast pattern indicating the emergence of a dislocation is seen at the termination of a step line. The step may be a growth step terminating in a screw dislocation or it may be a slip-trace, marking the passage of the dislocation as it has migrated through the crystal or glided along a fault plane (Fig. 9).

Planar stacking faults in the bulk crystal intersect the surface in straight lines which may or may not involve a surface step, depending on whether or not the displacement vector,  $\mathbf{R}$ , for the fault has a component perpendicular to the surface. In either case, the lattice displacement across the fault gives dark or light contrast, depending on the diffraction vector. Shuman [30] predicted that for a fault intercept with the surface, running perpendicular to the incident beam, there would be an oscillatory contrast similar to that seen in TEM images of faults but with contrast decreasing rapidly with depth of the fault below the crystal surface. However, no clear evidence of such contrast has been seen. If there is a surface step, the phase-contrast image, as described above, is added to the fault contrast and can be the predominant effect. For f.c.c. metals, fault lines have been seen joining the characteristic

contrast features corresponding to the two partial dislocations [33].

When f.c.c. metals such as Au and Pt are crystallized by rapid cooling from high temperatures, deformation fault lines or slip traces often appear on the surfaces, cutting across the curved growth steps as in Fig. 9. The points of intersection of the growth and fault steps are highly unstable and, if the sample has been at sufficiently high temperatures so that surface diffusion is possible, the sharp intersection points are rounded off, giving characteristic rounded cusps, lying along the straight line of the fault step, as seen in the image [34].

### 1.2.5 Surface Layers, Superlattices and Lattice Fringes

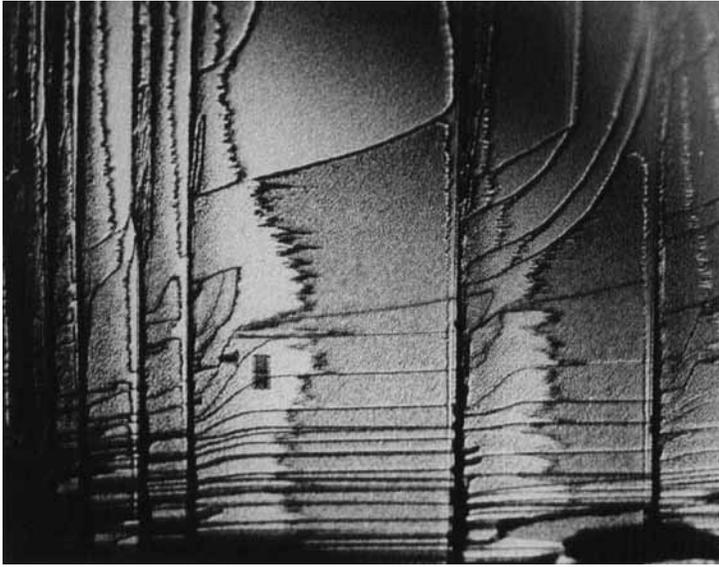
The high sensitivity of REM contrast to surface structure was first illustrated by the images obtained by Osakabe et al. [35], showing the growth of regions of the high-temperature Si (1 1 1)  $7 \times 7$  structure on an Si (1 1 1),  $1 \times 1$  face. The  $7 \times 7$  structure appeared as dark or light bands along the low sides of growth steps (see Fig. 2), showing, for the first time, the importance of the steps as growth sites for this surface superlattice. Even more striking contrast was given when gold was evaporated on the Si (1 1 1) surface, giving an Au  $5 \times 1$  superlattice structure with one third of a monolayer of Au. The Au superlattice was formed in domains in each of the three possible equivalent orientations and the different domain orientations gave very different image intensities [36]. With

medium-energy SREM imaging, a similar contrast for different domain orientations was shown for Si (1 1 1)  $2 \times 1$  superlattices [37].

Since then many such observations of surface structures and their domain configurations have been made. The notable advantage of the REM–RHEED method, as compared with the diffraction techniques such as LEED or other imaging methods such as LEEM or SEM is that the domain structures and their interactions with surface steps or other surface features may be examined with a spatial resolution of 1 nm or better.

For some oxide crystal faces, such as the  $\text{Al}_2\text{O}_3$  (0 1 1) surface imaged in Fig. 10, areas of very different intensity are separated by steps, whereas other, darker steps separate regions of the same intensity [38]. It is believed that these observations arise because the surface plane may cut the unit cell of the crystal at various levels giving different terminations of the lattice. The differences in contrast across the steps depend on whether the height of the steps is equal to a fraction, or the whole, of the unit cell repeat distance. In the case of Fig. 10, the darker areas appear rough, whereas the lighter areas are smooth. It is suggested that the difference of intensity is due not just to the difference in lattice termination, but to a difference in chemical reactivity of the exposed surface atom layers which leads to a reaction with the remanent gases in the relatively poor vacuum of a normal TEM for some terminations.

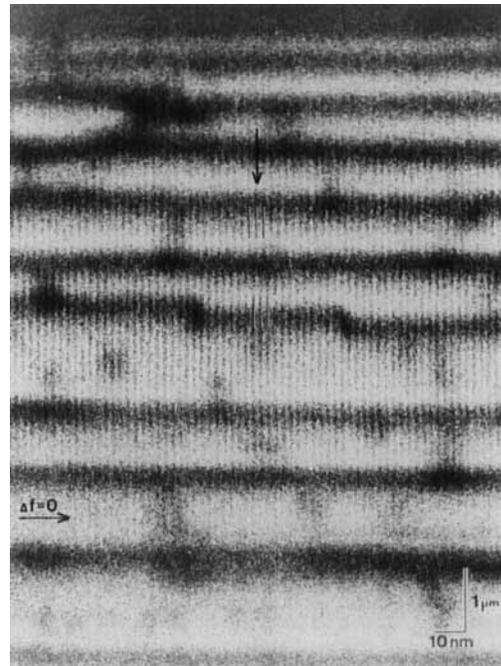
The case is somewhat different for the (1 1 1) faces of noble metals where, again, there are intensity differences across surface steps [39] (see Fig. 1). Here surfaces terminated at different levels within a unit



**Figure 10.** REM image from the (0 1 1) surface of an  $\alpha$ -alumina crystal showing regions with differences of intensity corresponding to different surface terminations and step images which are different for step-heights of whole or partial unit cells [38].

cell are exactly equivalent. The contrast is explained if there are stacking faults parallel to the surface in the near-surface planes; for example, if the top three planes have an h.c.p. sequence rather than an f.c.c. sequence. Then, when an atom-high step occurs, there can be a change from one sequence to the other and hence a change of diffraction intensities.

Many cases of surface reconstruction involve the formation of surface superlattices. When the superlattice periodicities exceed the resolution limit for the REM imaging, lattice fringes of the corresponding periodicity become visible. For the Si (1 1 1)  $7 \times 7$  structure the fringes corresponding to the 2.3 nm periodicity are clearly seen in such images as Fig. 11, due to Tanishiro et al. [40]. In later work with the high voltage UHV microscope, the basic lattice spacing of 0.34 nm was resolved by Koike et al. [13]. In images like these it is possible to observe the interactions of the periodic structures with the



**Figure 11.** Lattice fringes of the  $7 \times 7$  structure on an Si (1 1 1) face crossing surface steps which may, or may not, be out-of-phase boundaries of the  $7 \times 7$  structure. (Courtesy of K. Yagi [40].)

surface defects such as steps, emerging dislocations and impurities. Superlattice fringes with periodicities in the range of 1–5 nm have also been observed for metals by Lehmpfuhl and Uchida [41] and for various oxides by Liu et al. [42].

The observation of surface lattice fringes in REM, however, is not quite as straightforward as for the TEM case. The fringes of Fig. 11 are obtained when the objective aperture transmits a  $1 \times 1$  reflection from the underlying crystal plus several closely-spaced spots due to the surface superlattice, lying in a row parallel to the shadow-edge of the RHEED pattern. The superlattice spots lie on the Laue circle and so are often not in a line parallel to the shadow edge. If a group of superlattice spots which are on a line inclined to the shadow-edge are used to form the image, complications arise. The foreshortening factor and therefore the magnification in the beam direction, is different for each spot.

## 1.2.6 RHEED and REM Theory: Quantitative Interpretations

To allow quantitative interpretations of RHEED or REM intensities, two conditions must be met. Experimentally, the diffraction conditions must be determined with high precision. Then calculations must be made using an adequate many-beam dynamical diffraction formulation for the likely models of the structure. The dynamical diffraction theory is more complicated for the reflection case than for transmission of high-energy electrons,

but a number of adequate approaches are now available.

The original dynamical theory of electron diffraction was formulated by Bethe [43] for the LEED case and was adapted in its simplest two-beam form by Collela [44] for RHEED and by Shuman [30] to describe the contrast given by crystal defects in REM images. However, the Bethe theory assumes that a perfect crystal lattice is cut off discontinuously at a planar boundary. It must be considerably modified to include the gradual decay of the crystal potential into the vacuum and the perturbation of the surface structure by relaxation, reconstructions or chemically modified layers. Also a many-beam form is usually essential.

Current theoretical treatments may be divided into two types. In one, the crystal potential distribution is divided into thin layers, parallel to the surface. Then the propagation of the electron wave into and out of the crystal is considered in terms of its modification by each layer and its transfer between layers [45–47]. This approach works very well for crystals which are perfect in the two dimensions of the flat surface, with relatively small periodicities. As the surface unit cell dimensions become larger, the number of beams involved increases rapidly and the size of the computation grows even more rapidly.

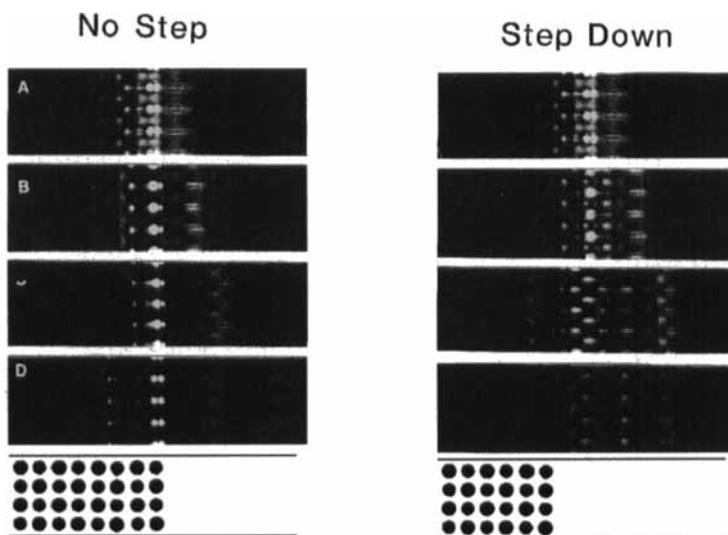
For nonperiodic objects, such as surface steps or boundaries of surface reconstructions, it is necessary to assume an artificial periodicity, using the assumption of periodic continuation. The assumption of large periodicity in real space corresponds to the sampling at small intervals in reciprocal space of the diffuse scattering arising from the nonperiodic object. The

repeat distance along the surface which must be considered is very large because of the small angles of incidence in REM, so that the number of sampling points in reciprocal space, and so the effective number of 'beams' in the calculation, is very large. However, even for this case, some successful calculations have been made by McCoy and Maksym [48].

In the other type of treatment, thin slices of the crystal perpendicular to the surface and almost normal to the incident beam are considered and transmission through these slices is simulated using the same sort of multislice, forward-scattering theory, as introduced by Goodman and Moodie [49], and used very successfully for transmission through thin crystals. The effect of the potential distribution in each slice of the specimen is given by a simple phase-object approximation since for forward scattering a small-angle approximation may be made. Then the propagation of the electron wave between slices is described in terms of Fresnel diffraction

theory by convolution with a propagation function. The structure within each slice of the crystal in this case, however, is essentially non-periodic, since the slice cuts across the surface, and so an artificial periodicity must be assumed, as if the diffraction were from the faces of a set of parallel crystal slabs [50, 51]. A large number of beams must be considered even for a sharply terminated perfect crystal, but then only a small further complication is introduced by modifications of the surface structure. Also it is easy to introduce variations of structure in the incident beam direction, such as surface steps or projections or crystal faults, by varying the content of successive slices [24] (Fig. 12). This approach has been further modified and extended and applied to the simulation of REM images of surface steps and faults by Ma and Marks [52] and Anstis [29].

The fact that RHEED spots and REM images normally contain a high proportion of inelastically scattered electrons



**Figure 12.** Calculations of the electron intensity distributions in planes perpendicular to the (100) surface of an MgO crystal at a distance, along the crystal face, of (A) 180, (B) 240, (C) 300 and (D) 480 slices, for a slice thickness of 0.2105 nm, after the entry of a narrow 100 kV electron beam into the crystal to give the (400) specular reflection. For the right-hand figures, there is a step-down after slice number 220. The bottom figures show the atom positions. The channeling of electrons along the surface planes of atoms is evident. (Courtesy of Z. L. Wang.)

introduces a complication which can be treated theoretically by including the inelastic scattering processes in the simulations; a difficult process [53]. It can be treated experimentally by use of an energy filter to remove the inelastically scattered electrons, implying that the calculations can be carried out for elastic scattering only but modified by the use of appropriate absorption functions.

### **1.2.7 Applications of Reflection Electron Microscopy**

Of the many successful applications of REM imaging, a few may be mentioned as illustrating the scope and potential of the method. The capability for imaging surface steps has led to studies of step movements in the processes of crystal growth and evaporation, and of the movements of steps under the influence of applied electric fields by Yagi et al. [54]. When a current flows through a specimen of Si, the surface steps may be seen to bunch together or spread apart, depending on the direction of the current, indicating that there is an essential asymmetry in the surface structure. Likewise, the migration of metal atoms on the Si surfaces may depend on the current direction. Such surprising results have important inferences for the understanding of surface migration processes.

The observation of steps and their decoration has provided evidence regarding the importance of steps for crystal growth processes. Steps frequently form the nuclei for the condensation of impurity

atoms arriving from the ambient atmospheres or by migration from the bulk of the specimen, as seen by Crozier et al. [55]. Small crystals growing on crystal surfaces as a result of deposition or chemical reaction are frequently seen to nucleate at steps or other defects of the surface [12].

The morphology of surfaces, described in terms of roughness or smoothness on an atomic scale, the occurrence of large steps and facets, or the frequency and regularity of small steps (Figs. 6, 8, and 10), provides information of importance in relation to the use of the surfaces as bases for epitaxial growth of thin films, for example of magnetic films or superconductors, for technical purposes. The REM technique has the advantage in such assessments, of allowing rapid surveys of large areas, plus studies of selected areas with high spatial resolution. Studies have been made by Wang and Bentley [56] at quite high temperatures, in situ, to show the changes of surface morphology taking place during the process of annealing of ceramic surfaces.

It has recently been shown that domain boundaries intersecting the surfaces of ferroelectric crystals can be seen and their movements under applied electric fields can be observed [57]. In this way it may be possible to throw light on a number of questions relating to the relevance of surface domain structures in relation to ferroelectric device properties.

A development of great potential significance has been the application of electron holography in the REM mode by Osakabe et al. [58] and Banzhof et al. [59]. As in transmission off-axis holography (Sec. 1.8 of this Chapter), an electrostatic biprism is inserted near to the image plane of the objective lens of a TEM

instrument. The electron waves reflected from the area of interest, the region of a flat crystal surface containing a defect, are made to interfere with the reference wave which is the wave reflected from an adjacent area of perfect crystal face. Perturbations of the pattern of interference fringes indicate the differences in phase of the two waves due to differences of height of the local surface relative to the ideally perfect crystal surface. Hence it is possible to map the local distortions of the surface. It has been shown that, using this technique, measurements of step heights or the strain fields around emerging dislocations may be made with an accuracy of 0.01 nm or better.

## 1.2.8 References

- [1] E. Ruska, *Z. Phys.* **1933**, 83, 492.
- [2] J. W. Menter, *J. Photogr. Sci.* **1953**, 1, 12.
- [3] K. Ito, T. Ito, M. Watanabe, *J. Electron. Microsc.* **1954**, 2, 10.
- [4] J. S. Halliday, R. C. Newman, *Br. J. Appl. Phys.* **1960**, 11, 158.
- [5] P. E. H. Nielsen, J. M. Cowley, *Surf. Sci.* **1976**, 54, 340.
- [6] J. M. Cowley, J. L. Albain, G. G. Hembree, P. E. H. Nielsen, F. A. Koch, J. D. Landry, H. Shuman, *Rev. Sci. Instrum.* **1975**, 46, 826.
- [7] E. S. Elibol, H.-J. Ou, G. G. Hembree, J. M. Cowley, *Rev. Sci. Instrum.* **1985**, 56, 1215.
- [8] T. Ichinokawa, *Ultramicroscopy* **1984**, 15, 193.
- [9] M. Ichikawa, T. Doi, M. Ichihashi, K. Hayakawa, *Jpn. J. Appl. Phys.* **1984**, 23, 913.
- [10] P. A. Bennett, A. P. Johnson in *Reflection High Energy Electron Diffraction and Reflection Imaging of Surfaces* (Eds.: P. K. Larsen, P. T. Dobson), Plenum Press, New York **1988**, p. 371.
- [11] N. Osakabe, Y. Tanishiro, K. Yagi, G. Honjo, *Surf. Sci.* **1981**, 102, 424.
- [12] A. V. Latyshev, A. B. Krasilnikov, A. L. Aseev, *Microsc. Res. Tech.* **1992**, 20, 341.
- [13] H. Koike, K. Kobayashi, S. Ozawa, K. Yagi, *Jpn. J. Appl. Phys.*, **1989**, 28, 861.
- [14] Z. L. Wang, J. Bentley, *Microsc. Microanal. Microstruct.* **1991**, 2, 301.
- [15] T. Hsu, *Microsc. Res. Tech.* **1992**, 20, 318.
- [16] J. Liu, J. M. Cowley, *Ultramicroscopy* **1993**, 48, 381.
- [17] J. M. Cowley, *Ultramicroscopy* **1992**, 47, 187.
- [18] Z. L. Wang, J. M. Cowley, *Surf. Sci.* **1988**, 193, 501.
- [19] K. Okamoto, T. Ichinokawa, Y.-H. Ohtsuki, *J. Phys. Soc. Jpn.* **1971**, 30, 1690.
- [20] A. G. Emslie, *Phys. Rev.* **1934**, 45, 43.
- [21] M. Gajdardziska-Josifovska, J. M. Cowley, *Acta Crystallogr.* **1991**, A47, 74.
- [22] L. Wang, J. M. Cowley, *Ultramicroscopy* **1994**, 55, 228.
- [23] H. Martin, G. Meyer-Ehmsen, *Surf. Sci.* **1985**, 151, 570.
- [24] Z. L. Wang, P. Lu, J. M. Cowley, *Ultramicroscopy* **1987**, 23, 205.
- [25] J. M. Cowley, L.-M. Peng, *Ultramicroscopy* **1985**, 16, 59.
- [26] Y. Uchida, G. Lehmpfuhl, *Ultramicroscopy* **1987**, 23, 53.
- [27] L.-M. Peng, J. M. Cowley, T. Hsu in *Proceedings of the 45th Annual Meeting of the Electron Microscopy Society of America* (Ed.: G. W. Bailey), San Francisco Press, San Francisco **1987**, p. 34.
- [28] N. Yao, J. M. Cowley, *Microsc. Res. Tech.* **1992**, 20, 413.
- [29] G. R. Anstis, in *Proc. Internat. Conf. Electron Microsc.-13*, Paris **1994**, p. 1027.
- [30] H. Shuman, *Ultramicroscopy* **1977**, 2, 361.
- [31] L.-M. Peng, J. M. Cowley, *Micron Microscopia Acta.* **1987**, 18, 171.
- [32] L.-M. Peng, J. M. Cowley, T. Hsu, *Micron Microscopia Acta.* **1987**, 18, 179.
- [33] F. Tsai, J. M. Cowley, *Ultramicroscopy* **1992**, 45, 43.
- [34] T. Hsu, J. M. Cowley, *Ultramicroscopy* **1983**, 11, 239.
- [35] N. Osakabe, Y. Tanishiro, K. Yagi, G. Honjo, *Surf. Sci.* **1981**, 109, 353.
- [36] N. Osakabe, Y. Tanishiro, K. Yagi, G. Honjo, *Surf. Sci.* **1980**, 97, 393.
- [37] J. M. Cowley in *Surface and Interface Characterization by Electron Optical Methods* (Eds.: A. Howie, U. Valdre), Plenum Press, New York and London **1988**, p. 127.
- [38] N. Yao, Z. L. Wang, J. M. Cowley, *Surf. Sci.* **1989**, 208, 533.
- [39] T. Hsu, J. M. Cowley in *The Structure of Surfaces* (Eds.: M. A. van Hove, S. Y. Tong), Springer-Verlag, Berlin **1984**, p. 55.
- [40] Y. Tanishiro, K. Takayanagi, K. Yagi, *J. Microsc.* **1986**, 142, 211.
- [41] G. Lehmpfuhl, Y. Uchida, *Surf. Sci.* **1990**, 235, 295.
- [42] J. Liu, L. Wang, J. M. Cowley, *Surf. Sci.* **1992**, 268, L293.

- [43] H. A. Bethe, *Ann. Phys.* **1928**, 87, 55.
- [44] R. Colella, *Acta Crystallogr.* **1972**, A28, 11.
- [45] P. A. Maksym, J. L. Beeby, *Surf. Sci.* **1984**, 140, 77.
- [46] A. Ichimiya, *Jpn. J. Appl. Phys.* **1983**, 22, 176.
- [47] S. Y. Tong, T. C. Zhao, H. C. Poon in *Reflection High Energy Electron Diffraction and Reflection Electron Imaging of Surfaces* (Eds.: P. K. Larsen, P. J. Dobson), Plenum Press, New York and London **1988**, p. 63.
- [48] J. M. McCoy, P. A. Maksym, *Surf. Sci.* **1994**, 310, 217.
- [49] P. Goodman, A. F. Moodie, *Acta Crystallogr.* **1974**, A30, 280.
- [50] J. M. Cowley, P. M. Warburton in *The Structure and Chemistry of Solid Surfaces* (Ed.: G. A. Somorjai), J. Wiley and Sons, New York **1967**, p. 6-1.
- [51] L.-M. Peng, J. M. Cowley, *Acta Crystallogr.* **1986**, A42, 552.
- [52] Y. Ma, L. D. Marks, *Microsc. Res. Tech.* **1992**, 20, 371.
- [53] Z. L. Wang, P. Lu, *Ultramicroscopy* **1988**, 26, 217.
- [54] K. Yagi, Y. Akira, I. Homma, *Microsc. Res. Tech.* **1992**, 20, 333.
- [55] P. A. Crozier, M. Gajdardziska-Josifovska, J. M. Cowley, *Microsc. Res. Tech.* **1992**, 20, 426.
- [56] Z. L. Wang, J. Bentley, *Ultramicroscopy* **1993**, 51, 64.
- [57] F. Tsai, J. M. Cowley, *Ferroelectrics* **1993**, 140, 203.
- [58] N. Osakabe, N. Endo, J. Matsuda, A. Fukuhara, A. Tonomura, *Phys. Rev. Letts.* **1989**, 62, 2969.
- [59] H. Banzhof, K. H. Herrmann, H. Lichte, *Microsc. Res. Tech.* **1992**, 20, 450.
- [60] K. Yagi, *J. Appl. Crystallogr.* **1987**, 20, 147.
- [61] N. Yao, J. M. Cowley, *Ultramicroscopy* **1989**, 31, 149.
- [62] L. Wang, J. Liu, J. M. Cowley, *Surf. Sci.* **1994**, 302, 141.
- [63] N. Yao, J. M. Cowley, *Ultramicroscopy* **1990**, 33, 237.

# 1.3 Electron Energy-Loss Spectroscopy Imaging

## 1.3.1 Introduction

Electron energy-loss spectroscopy (EELS) measures the spectral distribution of energy transferred from a monochromatic incident electron beam into a specimen. Two main types of geometry are concerned: low energy primary beams reflected by solid surfaces, and high energy primary beams transmitted through thin foils or at glancing incidence along surfaces. In the first case, using monochromators, the investigated excitation spectrum covers the millielectronvolt to electronvolt range (which is equivalent to the infra red domain of the electromagnetic spectrum) and is largely a vibrational spectroscopy. In the second case, corresponding to the electron microscopy situation, the inelastic scattering events involved encompass a whole range of excitations from the electronvolt to the kiloelectronvolt range, which is equivalent to the visible to soft X-ray domain. It is, therefore, a spectroscopy of electron states, related to the excitation of the valence and conduction electrons in the low energy-loss domain, i.e. from 1 to 50 eV, and of the atomic-core electrons, in the high

energy-loss domain from about 50 eV up to one- or several-thousand of electronvolts.

When performed in the context of an electron microscope column, the EELS technique obviously introduces a new dimension to electron microscopy. As a consequence of the elemental specificity of the signal delivered, EELS is now well established as an essential component of analytical microscopy, extending the accessible spatial resolution to the sub-nanometer level. Furthermore, it is largely accepted that the access to the information stemming from inelastic processes, in parallel to that contained in the elastic ones, constitutes an essential step forward in promoting electron imaging and diffraction as real quantitative tools. This chapter describes the present state of instrumentation and methodology accessible in the field of EELS imaging and emphasizes the great diversity of potential fields of application through a selection of recent investigations. For a more complete knowledge of the impact of EELS as a spectroscopy of electron states (compared to photon beam techniques), the reader is advised to consult other reviews [1–5].

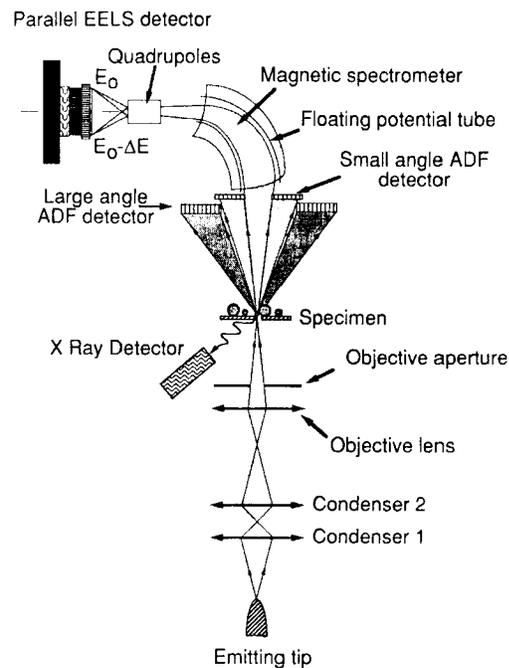
## 1.3.2 Instrumentation and Methods

### 1.3.2.1 An Historical Survey

One can trace back the origin of the use of electron energy losses as a micro-analytical technique to the early work of Hillier and Baker [6], five decades ago. However, it only became used practically in the 1960's, as a consequence of the progress achieved in realizing and coupling well-adapted analyzers and filters to an electron microscope column. Two major approaches (scanning or fixed beam) have been explored, leading to energy-analyzing and energy-selecting microscopies.

In the energy-analyzing mode, a spectrometer transforms a point on the object into an EELS spectrum. From the prototypes designed by Wittry [7] to be introduced at the bottom of a conventional transmission electron microscope (CTEM) column, and by Crewe et al. [8] as an essential part of their first scanning transmission electron microscope (STEM) microscope, it has evolved into the present Gatan commercially available attachment realized by Krivanek and Swann [9]. It is basically made of a magnetic sector with tilted and curved entrance and exit surfaces of the pole pieces in order to deliver a second-order aberration-corrected image of the spectrometer object function. This combination of a STEM column with such an EELS spectrometer is particularly well suited to recording EELS spectra from local areas. When governed by the intensity distribution within the probe issued from a field emission source, the accessible spatial resolution can be as small as a

fraction of a nanometer. Furthermore, the upgrading of the initial serial detection system (which comprised a slit and a scintillator-photomultiplier device) by a parallel array of typically 1000 diodes which measure in parallel all the spectral channels, has been responsible for a major breakthrough in terms of detection efficiency [10]. Such a combination has demonstrated a routinely accessible energy resolution of about 0.5 eV on core edges and identification capabilities approaching the single atom level, when coupled to a FEG STEM [11, 12]. Figure 1 shows the basic configuration and major components



**Figure 1.** Typical example of the STEM + PEELS configuration for point analysis. It offers the simultaneous capability of visualizing the object topography through the annular detectors (small-angle for diffraction contrast, large-angle for Z-contrast) and of analyzing with the magnetic spectrometer the energy-loss distribution of the forwardly transmitted beam. (ADF = annular dark field.).

of the dedicated STEM equipped with an EELS spectrometer and a parallel detector array, in operation in Orsay. Although the performance of the diode array has not yet been optimized, it constitutes the best approach presently available for the acquisition of single EELS spectra originating from well-defined areas on the specimen, and a variety of processing techniques have been developed to extract the useful information from these spectra [13, 14]. An alternative solution to the magnetic sector is the Wien filter. It has been implemented on a dedicated VG STEM by Batson [15], and this unique system on an upgraded microscope provides the best presently attainable performance in terms of both energy resolution (of the order of 0.15 eV) and spatial resolution (down to 0.2 nm). However, it is not suitable for routine use and has not been duplicated by any other group or manufacturer.

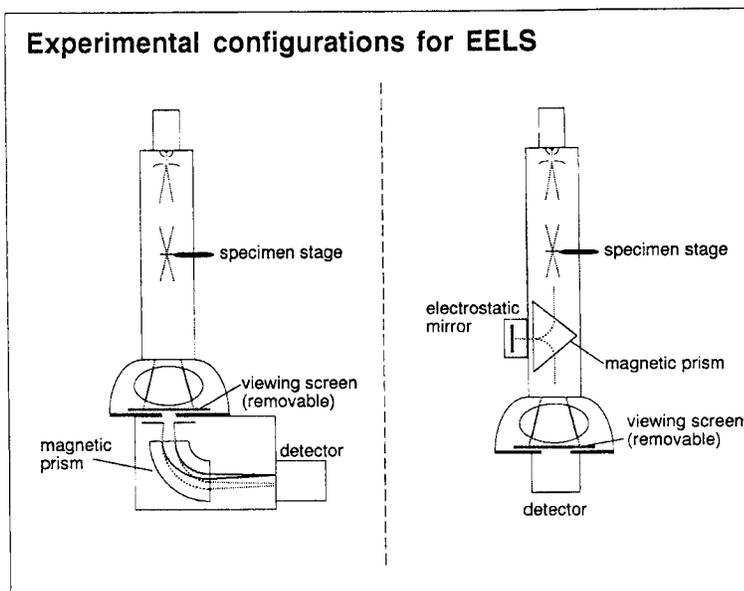
In the energy-selecting mode, an imaging stage is added to the spectrometer, transforming the spectrum behind the energy-selecting slit into an image corresponding to a specific energy loss. This approach has been pioneered by Castaing and Henry [16], who produced the first high-quality energy-filtered images using a magnetic prism–electrostatic mirror device incorporated in the imaging part of a microscope column. Obviously, their realization has offered quite new possibilities to conventional transmission electron microscopy (TEM), such as a type of ‘colour’ microscopy, as different images corresponding to different types of inelastic interaction can then be obtained from a selected area. The first micro-analytical applications followed rather quickly: El Hili [17] used the change of

plasmon energy in different metals and alloys and Colliex and Jouffrey [18] showed energy-filtered images with electrons corresponding to a specific core loss.

In spite of these promising results, the first commercial attempt at producing an energy-filtering microscope using the Castaing–Henry design, by the French company Sopenem, has not been a clear success. It took nearly 20 years before Zeiss introduced to the market its 902 microscope, using the same concept, then its fully magnetic substitute, known as the ‘omega filter’, in the 912 machine. As a matter of fact this second version also had its origin in Castaing’s group [19], before being developed by Zanchi et al. [20] and optimized by Rose and coworkers [21, 22].

However, these filters are only available on dedicated instruments and cannot be retrofitted to existing microscopes. An alternative possibility, first explored by Shuman and Somlyo [23], consists of an imaging spectrometer where a simple magnetic sector acting as an analyzer is followed by an electron optics transfer system which displays the energy filtered image on a two-dimensional (2D) recording medium. The very sophisticated system designed, produced and tested by Krivanek and coworkers [24, 25], known as the Gatan imaging filter (GIF), can be adapted to any TEM column (even on megavolt instruments) and provides in a quite user-friendly configuration energy-filtered images with atomic resolution. Figure 2 compares the two approaches to the fixed-beam energy filtering presently commercially available, the Gatan 1-6-6 filter and the Zeiss  $\Omega$  filter.

**Figure 2.** Typical examples of the incorporation of filtering devices in a TEM column for energy selected imaging (ESI). A change in the excitation of the first intermediate lens replaces the diffraction pattern in the entrance object plane of either filter by an image of the specimen, and replaces the final energy-filtered image on the detector by an energy-filtered diffraction pattern.

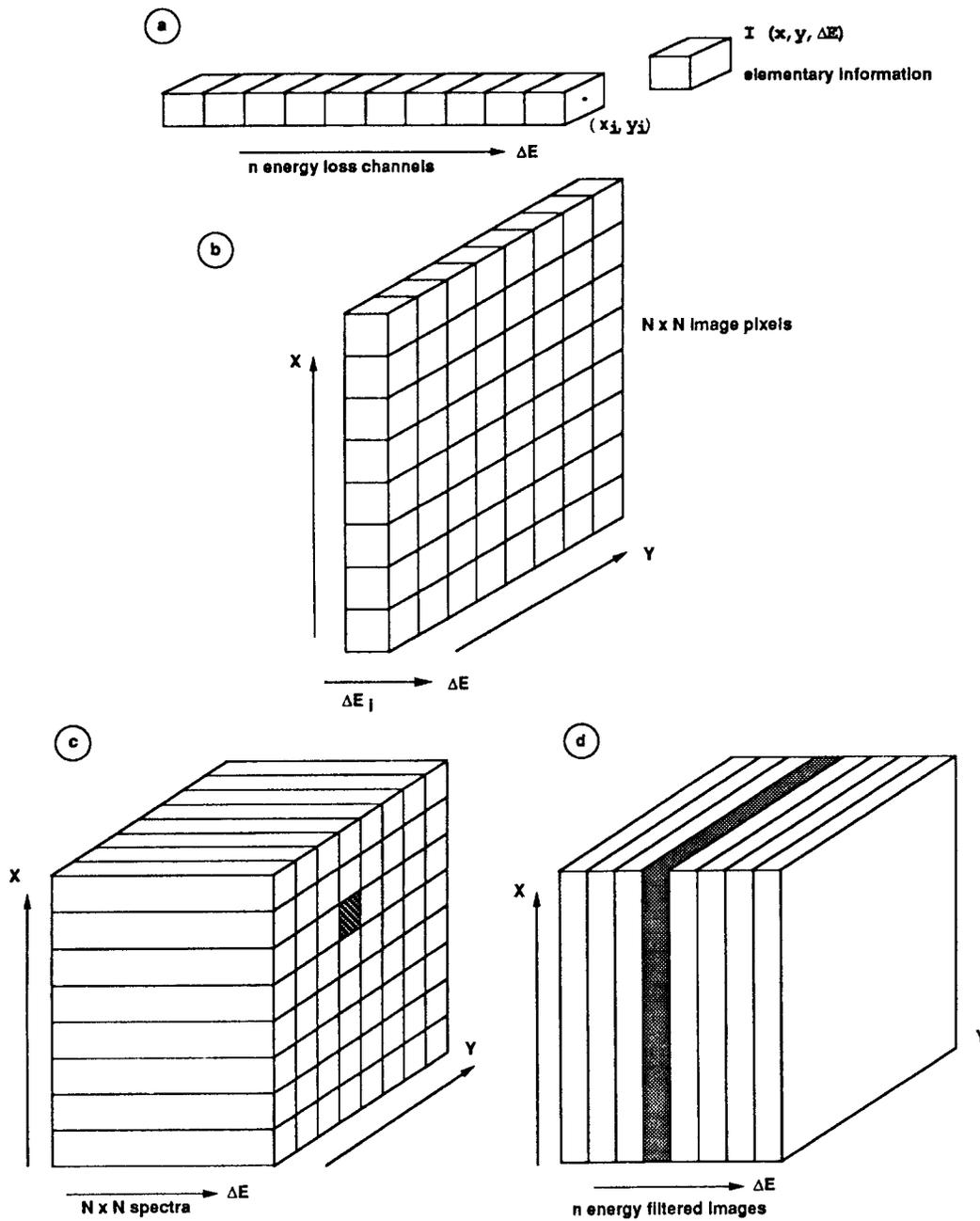


### 1.3.2.2 A New Dimension in EELS Imaging: From Point Analysis and Energy-Filtered Images to Spectrum Imaging

The introduction of the energy-loss parameter ( $\Delta E$ ) adds a new dimension of information to the current 2D spatial information provided for a given pixel position ( $x, y$ ) by any type of scattering responsible for the contrast observed in a conventional electron micrograph. One can then measure the intensity  $I(x, y, \Delta E)$  within an elementary volume defined within a three-dimensional (3D) space with two axes related to the position and one to the spectral information. For any of these elementary volumes, a given dose of electrons ( $J \times \tau$ ) incident on the entrance surface of the specimen, is required in order to produce a measurable signal with a given signal-to-noise ratio, which

depends practically on the physical parameters (thickness and composition) of the specimen and of the investigated spectral channel.

Basically, a spectrometer in a STEM mode with parallel acquisition provides a full spectrum  $I(\Delta E)$  covering  $n$  energy-loss channels for each probe position, i.e. for each pixel defined by its coordinates ( $x, y$ ) on the specimen within the time interval  $\Delta t_1$  set by the dose requirements (see Fig. 3a). On the other hand an energy filtering microscope (EFTEM) provides a complete 2D image made up of  $N \times N$  pixels, using only the electrons contained within an energy band defined by the selection slit (see Fig. 3b). Similar criteria of signal-to-dose ratio apply for the evaluation of the image recording time  $\Delta t_2$ . To obtain the same intensity  $I(x, y, \Delta E)$  on a given area element within a given energy-loss channel, it implies, as stated above, that  $J_1 \times \Delta t_1 = J_2 \times \Delta t_2$ , all



**Figure 3.** Definition of: (a) the parallel acquired EELS spectrum in a STEM configuration; (b) the energy filtered image in an EFTEM configuration; the image-spectrum acquired either as a 2D collection of parallel EELS spectra (c) or as a sequence of energy filtered images (d).

other parameters governing the measured signal being equal. The ratio between the recording times  $\Delta t_1$  and  $\Delta t_2$  depends on the ratio of the primary flux of electrons available to the specimen. Practically, this is about  $10^2$ – $10^3$  times higher in a nanoprobe delivered from a field emission source in a STEM than in a submicroprobe delivered from a LaB6 filament in an EFTEM, and  $\Delta t_1/\Delta t_2 = a$  is of the order of  $10^{-2}$ – $10^{-3}$ .

A few years ago, Jeanguillaume and Colliex [26] introduced the concept of the spectrum-image in EELS digital acquisition and processing. It was originally stimulated by the access to parallel EELS recording devices in the STEM and by the availability of computers with rapidly increasing capacities for data storage and handling. As a matter of fact, this concept can be applied to any case where spectra can be acquired as a function of a probe location, such as energy-dispersive X-ray (EDX) or Auger analysis under a primary electron beam, inelastic tunnelling spectroscopy under the scanning tip, and secondary ion mass spectrometry (SIMS) by sputtering under an ion probe. A spectrum image is a 3D ensemble of numbers  $I(x, y, \Delta E)$ : the first two axes correspond to the  $(x, y)$  position on the specimen, similarly to any image and the third axis is associated with the energy-loss dimension ( $\Delta E$ ). As shown in Figs. 3c and d, the spectrum image can be built with the basic entities provided by the STEM + PEELS analyzing microscope (Fig. 3a) and by the EFTEM filtering microscope (Fig. 3b). It can be described either as a collection of  $N \times N$  spectra acquired while scanning the STEM probe over the specimen surface, or as a stack of  $n$  filtered images recorded successively

while ramping the energy-loss value of the electrons transmitted through the selection slit. The total time  $T$  needed for recording this amount of  $N \times N \times n$  values is, in the first case,  $N \times N \times \Delta t_1$  and, in the second case  $n \times \Delta t_2$ , but the total dose  $D$  received by the elementary area  $\Delta x \Delta y$  is proportional to  $J_1 \times \Delta t_1$  for the STEM and to  $J_2 \times n \times \Delta t_2$  for the EFTEM. This leads to:

$$T_1/T_2 = N \times N \times a/n \quad D_1/D_2 = 1/n \quad (1)$$

Consequently, one can summarize this comparison between the two approaches, as follows:

(i) The parallel EELS + STEM mode is always better in terms of the dose required, because of the intrinsic superiority of the technique in recording all energy-loss channels in parallel, while the EFTEM technique requires as many images (and irradiations) as the number of energy-loss channels required.

(ii) The comparison is not as obvious when considering the total time required for the acquisition. For small numbers of pixels and high numbers of loss channels, the PEELS + STEM is superior, while for large  $N$  and small  $n$  values the EFTEM provides better results. Of course, for a single filtered image, such as the selection of the elastic image, filtering devices have to be recommended.

In any case, however it has been acquired, the great advantage of the spectrum-image is that it can be processed a posteriori, offering access to the exact relationship between the spatial coordinate and the spectral feature. For instance, elements in unexpected locations may be found without any a priori operator

decision regarding where to locate the beam for data collection. Furthermore, it allows summation of spectra within segmented areas, thus improving the statistical significance of the data. On the other hand, the major limitation is the large data capacity required (32 Mbyte for a  $128 \times 128 \times 1024 \times 16$  bit spectrum-image) and the time needed for access and manipulation of data in any type of off-line processing. But this difficulty will surely gradually vanish as faster and cheaper computers become available.

The hardware and software required for implementing spectrum-imaging techniques have been used over the last few years by a number of groups. In particular, Hunt and Williams [27] have published the first complete description of a system installed on a VG HB501 at NIH, Bethesda, MD, and on a Philips 430 TEM/STEM at Lehigh University, Pennsylvania, USA, illustrated with results obtained in both life and materials sciences [28, 29]. On the other hand, the access to

image-spectra using a fixed-beam filtering microscope has been demonstrated by Lavergne et al. [30]. The number of facilities that have developed their own equipment has increased during the past year, and the examples given in the following sections of this Handbook have been acquired and processed using the system described previously [31]. The system consists of a combination of a FEG VG STEM with a Gatan 666 PEELS in which the array of photodiodes has been modified to reduce the read-out noise. It makes profiling and mapping of EELS spectra achievable with (sub)nanometer spatial resolution. The experimental parameters routinely employed are summarized in Table 1. For reasons already discussed (limited number of pixels), the present installation is particularly well suited to the investigation of line spectra, that is, to one-dimensional (1D) spatial data acquired when scanning the probe along a given line on the specimen.

**Table 1.** Parameters currently used for the acquisition of spectrum lines and spectrum images with a PEELS + STEM configuration

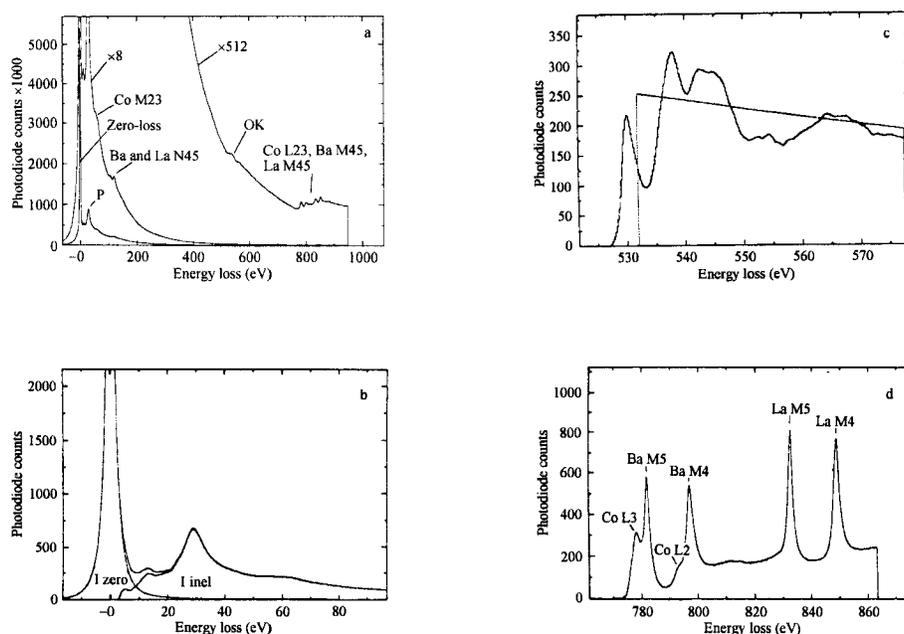
	Probe size (nm)	Typical extent of the used signal (nm)	No. of pixels involved	Pixel step on the specimen (nm)	Average recording time per spectrum
<b>Spectrum-line</b>					
Low losses (plasmons, dielectric coefficients)	1–2	>1–2	64–256	0.5–1	25–250 ms
Elemental mapping from core-losses	0.5	<0.5	32–128	0.3–0.5	1 s
Fine structures on core-losses	0.5	<0.5	32–128	0.3–0.5	1–5 s
<b>Spectrum image</b>					
Elemental mapping from core-losses	1–2	<0.5	32 × 32 or 64 × 64	0.5–1	0.1–1 s

Another interesting possibility is the access to time-resolved spectroscopy ('chronospectroscopy' [32]), in which case the incident beam is maintained fixed on the specimen and sequences of spectra are acquired at fixed time intervals and reflect in particular the change of the specimen as a function of the dose. This possibility has proved to be of great use when investigating beam-sensitive specimens, in which case it is possible to estimate the zero-dose spectrum by techniques of back-extrapolation.

### 1.3.3 Understanding and Extracting the Information Contained in an EELS Spectrum

#### 1.3.3.1 Anatomy of an EELS Spectrum

A typical EELS spectrum [1–5] exhibits the characteristic features shown in Fig. 4: a zero-loss peak; a major contribution (the low-loss region) extending from



**Figure 4.** An EELS spectrum acquired on a thin foil of nominal composition  $\text{LaBaCo}_2\text{O}_{5+\delta}$  (specimen courtesy of F. Studer, Caen). (a) The whole spectrum extending over 1000 eV covers four orders of magnitude in intensity for comparison with the non-saturated zero-loss peak and the details of the overlapping Co, Ba and La edges around 800 eV loss. (b) The tail of the zero-loss peak and the onset of inelastic scattering, which can be used to evaluate the local thickness in terms of the inelastic mean free path ( $t/\lambda_{\text{in}}$  is of the order of 0.63 in this present spectrum). (c) The oxygen K edge after background subtraction and multiple loss deconvolution, compared with an atomic calculation of the relevant cross-section (using the Sigmap program of Egerton). Note how the fine structures modulate the calculated cross-section over the ELNES range. (d) The overlapping white lines and edges for  $\text{CoL}_{2,3}$ ,  $\text{BaM}_{4,5}$  and  $\text{LaM}_{4,5}$ .

about 5 to 50 eV and corresponding to the excitation of electrons in the valence band and low-lying levels; and, superimposed on a continuously decreasing background, a succession of element-specific edges at increasing energy losses. All these features contain useful information, which in certain cases can only be reached after some lengthy procedure. Briefly, the low-loss structures reflect a mixture of collective ('plasmon') and interband processes, i.e. some kind of average electron properties.

For example in a material exhibiting quasi-free-electron gas behavior, the energy of the plasmon peak is proportional to the mean density of electrons.

On the other hand, the core-loss features that appear between 100 and 1000 eV are chemically representative and constitute the major route to nano- and micro-analytical applications. However, it is obvious that all visible edges do not display similar shapes, their general behavior being imposed by atomic considerations. If the

**Table 2.** Summary of the different types of information accessible through a detailed analysis of an EELS spectrum

Spectrum domain	Information accessible	Required processing technique	Field of application
Whole spectrum and zero loss in particular	Thickness, total inelastic scattering	Measurement of unsaturated zero-loss peak	Very general
Low-loss region	Average electron density	Measurement of plasmon line properties	Microanalysis in metallic alloys
Low-loss region	Interband transitions/joint density of states	Kramers-Kronig transformation; calculation of dielectric constants (critical-point modeling)	Optical and transport properties; comparison with VUV spectra; intergranular van der Waals forces
Low-loss region	Interface/surface properties	Study of the interface plasmon modes	Interface and boundary structure and chemistry
Core-loss region	Qualitative and quantitative elemental analysis	Measurement of core-edge weight	Nanoanalysis of any type of material
Core-loss region	Site symmetry; bonding type; bond lengths	Analysis of core-ELNES; comparison with fingerprints; molecular orbital, multiple scattering or band structure calculations	Site-selected valence state; charge transfer; bonding and structural environment
Core-loss region	Radial distribution function	Analysis of core-EXELFS	Site-selected crystal coordination
Core-loss region	Density of holes on local states	Measurement of white-line intensities	Electron configuration in intermetallics, insulators, and superconductors; magnetic properties

ELNES, energy loss near edge structure; EXELFS, extended electron energy loss fine structure; VUV, vacuum ultraviolet.

oxygen K edge exhibits a rather typical sawtooth profile modulated by fine structures, the cobalt L<sub>23</sub> edge and the Ba and La M<sub>45</sub> edges appear as sharp lines ('white lines') followed by a step function. These are due to strong transitions from the initial levels with p- and d-type symmetry to unoccupied bound states of 3d and 4f symmetry, respectively, reflecting (at least indirectly) the population of d or f holes on the excited atom. Table 2 summarizes the different types of information accessible through an analysis of the different spectral regions.

### 1.3.3.2 Methods Available for Data Processing and Specific Information Extraction

In parallel with the continuous improvement in understanding the physical content of EELS spectra, methods have been elaborated for processing spectra in order to extract from them the different types of specific information listed in Table 2 [33–35]. The most important methods are those dealing with quantitative elemental analysis. The first step consists in evaluating the local thickness and the importance of multiple scattering events. Assuming a Poisson distribution for the occurrence of independent scattering events

$$P_n(t) = \frac{1}{n!} \left( \frac{t}{\lambda_i} \right)^n \exp \left( - \frac{t}{\lambda_i} \right) \quad (2)$$

one calculates

$$t/\lambda_i = \log(I/I_0) \quad (3)$$

where the thickness is scaled to the inelastic mean free path ( $\lambda_i$ ) corresponding to all possibilities of an inelastic event,

whatever may be the scattering angle  $\beta$  or the energy loss  $\Delta E$ . The information is then contained in the ratio between the intensity of the zero-loss peak ( $I_0$ ) compared to that of the total energy loss spectrum with the zero-loss peak included ( $I$ ). As soon as the ratio  $t/\lambda_i$  is higher than about 0.5, deconvolution techniques must be employed in order to recover the single scattering signals.

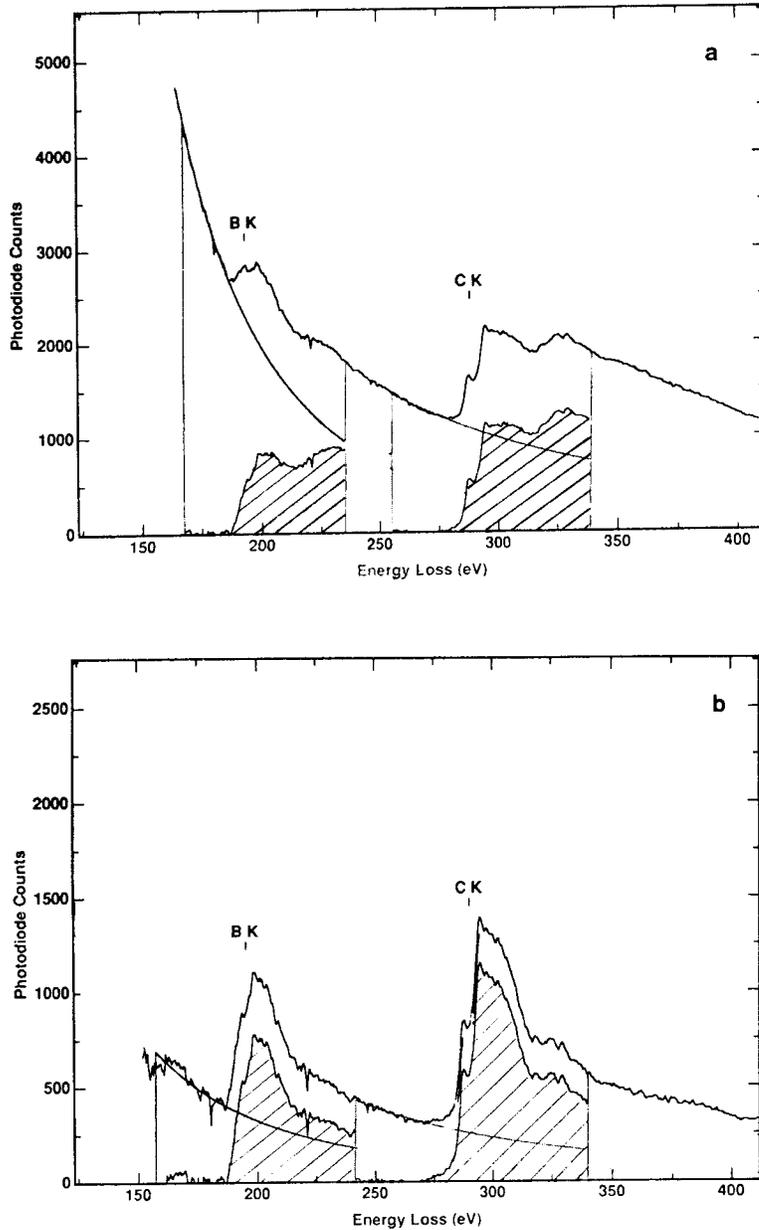
The basic formula for quantitative elemental analysis with a core-loss relates the measured signal ( $S$ ) to the number of atoms generating it:

$$S = I_{\text{inc}} N \sigma \tau \quad (4)$$

where  $I_{\text{inc}}$  is the incident beam current,  $N$  is the number of atoms per unit specimen area giving rise to the measured signal,  $\sigma$  is the relevant cross-section, and  $\tau$  is the counting time.

Egerton [34] has shown that a good approximation, when measuring the signal for a given angle of collection ( $\beta$ ) and integrated over an energy window ( $\Delta$ ) over threshold, is to introduce in the above formula the low-loss spectrum [ $I_0(\beta, \Delta)$ ] and the cross-section [ $\sigma(\beta, \Delta)$ ] measured under the same conditions as the signal [ $S(\beta, \Delta)$ ].

There are two types of problem associated with achieving a good quantitative analysis: the measurement of the characteristic signal  $S$  as discriminated from the background, and accessing reliable cross-section values. The extraction of the characteristic edge implies background modeling over a fitting window below threshold and extrapolation of this model curve under the signal before subtraction (Fig. 5a). The most generally used model is of the  $A \cdot \Delta E^{-R}$  type. It is a simple method of broad applicability and requires no



**Figure 5.** Quantitative analysis of a  $BC_3$  thin foil with  $t/\lambda_i = 0.87$ , using the background extrapolation technique under the BK and CK edges: (a) original spectra, (b) deconvoluted spectra. Using the Sigmak cross-sections, one finds in the first case  $B/C = 0.32$  and in the second case  $B/C = 0.34$ .

other information. However, it may fail because the power-law model is not well adapted in the low-loss regime (i.e. up to 100–150 eV). Furthermore, this method cannot be used where there are

overlapping edges, and the accuracy of the measurement of the signal depends on the quality of the fit [estimated by an  $h$  factor (following Egerton) or by the reduced  $\chi^2$  term].

The cross-sections are satisfactorily estimated using a hydrogenic model for the K and L shells, and the relevant values of  $\sigma(\beta, \Delta, E_0)$  calculated by Egerton [36] can be found in the routinely available quantification procedures such as the ELP software available from Gatan. For the outer shells of M, N or O type, the calculations are much more complex and require an Hartree–Slater description [37]. The values calculated by Rez are now available in the latest version of the ELP software. The reliability of these calculated values has been tested on reference specimens by many authors, in particular by Hofer [38]. Hofer has shown that for the K and L shells the accuracy of the calculations is of the order of  $\pm 5\%$  for most light-element edges up to the transition metal series. However, for heavier elements the edges involved require the use of the more complex calculations and discrepancies up to 40% have been found when comparing calculated values with measured ones.

When one is only interested in measuring relative concentrations of different elements (A and B), the above formula can be simplified as:

$$\frac{N_A}{N_B} = \left[ \frac{S_A(\beta, \Delta)}{S_B(\beta', \Delta')} \right] \left[ \frac{\sigma_B(\beta', \Delta')}{\sigma_A(\beta, \Delta)} \right] \quad (5)$$

where the ratio of cross-sections corresponds to a scaling  $k$  factor (following the terminology used in EDX analysis of thin foils), reduced to the ratio of the probability of occurrence of the edges involved.

For specimen thicknesses ranging from about  $0.5\lambda_i$  to  $1.5\lambda_i$ , the same formula can still be applied, by replacing the experimental signals with the deconvoluted ones. Two alternatives can be

used practically: (i) the Fourier-log method and (ii) the Fourier-ratio method. Both methods require a knowledge of an unsaturated low-loss spectrum from the same area. The first technique applies to the whole spectrum from zero loss to core losses, background included. The second method only requires independently recorded core-loss and low-loss spectra. Although it may introduce more noise, the first technique is generally recommended. Figure 5b shows the quantitative analysis applied to the deconvoluted spectra obtained from the original shown in Fig. 5a.

### 1.3.3.3 Alternative Approaches using Multiple Least-Squares Techniques

In many situations, the above standard technique consisting of background modeling, extrapolation, and subtraction is not adapted. One can quote the case of overlapping edges, the need for evaluating the weight of different components within an edge, and, in general, the detection of very small concentrations. In the latter case, the signal appears only as a weak modulation over an intense background and its measurement can be made more difficult when using a parallel recording device composed of a collection of individual detection units due to the inhomogeneities between the detection properties of the different detectors. For these reasons, new tools for recording and processing EELS data have been introduced. To reduce the influence of the variations of read-out and dark-current noises in the different channels of

a diode array used for parallel recording, the use of difference spectra (first or second difference) is a simple approach which requires a double or triple exposure of the same spectrum shifted by a given energy increment ( $\delta$ ) and covering different detector distributions. This method has proved to be quite efficient [13, 14].

In order to process all these different types of EELS spectra, multiple least-squares (MLS) fitting techniques are being rapidly developed [39–41]. These methods search for the combination of reference signals which best reproduce an experimental curve, encompassing both the background and the characteristic edges. The procedure is quite versatile and can be used for different purposes. It implies the use of reference spectra, either calculated or previously obtained from standards. The thickness effect and its associated multiple scattering events can be accounted for by convolution techniques. The major limitation in using MLS fitting techniques for EELS spectra, as compared to the simpler case of processing EDX data where all characteristic lines can be fitted with similar profiles, is due to the fact that the detailed shape of the edge exhibits variable fine structure, depending on the local environment of the atom. For example, one cannot use cation edges recorded from metals when the cation is inserted in an ionic compound.

These limitations being kept in mind, its application for quantitative analysis technique has been demonstrated successfully by Leapman [13] on different examples. In particular, the detection limits have been improved considerably when the magnitude of the derivatives of the cross-section

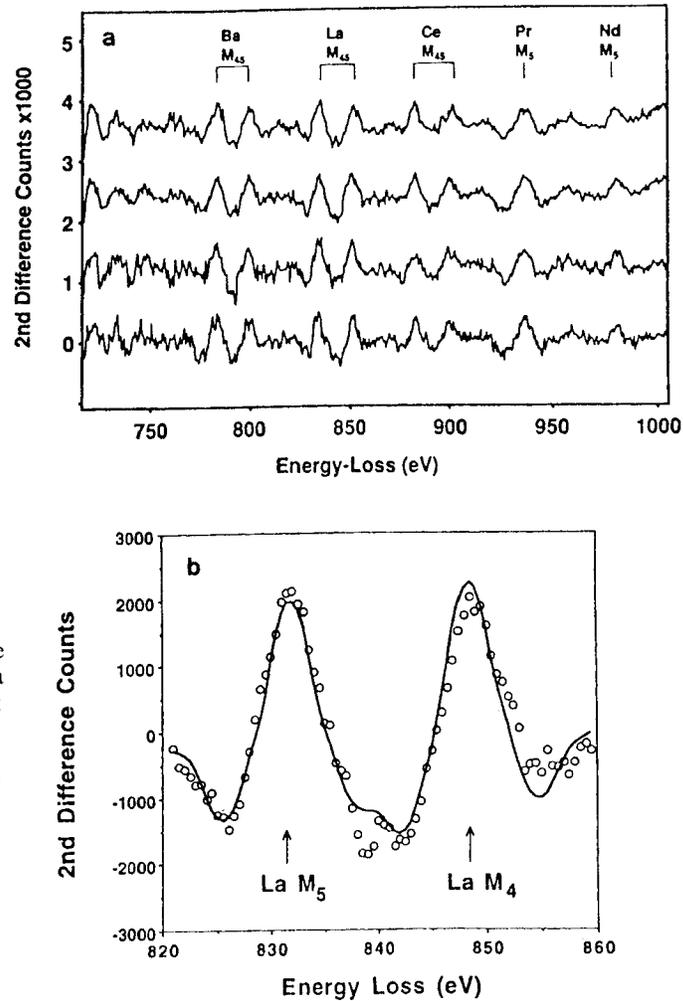
is high (in particular when the edge exhibits intense white lines at threshold). Figure 6 shows an example from Leapman and Newbury [42], extracted from a test experiment performed on a glass specimen containing 66 elements, and in particular transition and rare earth elements at trace concentrations. When considering the quality of the fit for the La  $M_5$ – $M_4$  lines, one can estimate the detection limit for such an element under the experimental conditions used to be of the order of 10 ppm. One further parameter controlling the detectability limit is then the edge shape:

(i) 3d transition elements and preceding alkaline earth  $L_{23}$  edges (from Ca to Ni), as well as 4f lanthanides and preceding alkaline earth  $M_{45}$  edges (from Ba to Yb) can be detected in the 10–100 ppm atomic concentration range;

(ii) light elements identifiable through their K edge (Li to P) or their  $L_{23}$  edge (Mg to Cl) and 4d transition elements identifiable through their  $M_{45}$  edge (Sr to Rh) can usually be detected in the 100–1000 ppm range; and

(iii) for all elements with no clear discontinuity at threshold (Ga, Ge, As, . . . , Sb, Te, W, . . .) it seems reasonable to estimate a normal detection limit around an atomic concentration of 1%.

Another important field of application for these fitting techniques is to obtain access to quantitative measurements of the different contributions within the fine structures of a given edge, i.e. the relative strength of a pre-peak on an O K edge such as the one shown in Fig. 4c, or the weight of the unoccupied d holes as compared to that of transitions to continuum states  $[(I_{L3} + I_{L2})/I_{\text{cont}}]$  on an  $L_{23}$

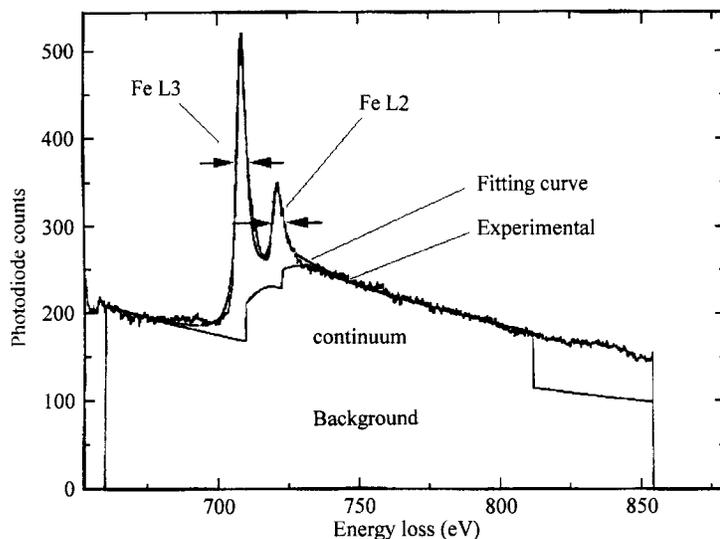


**Figure 6.** Detection of elements in the concentration range 10–100 ppm in a reference glass, using the MLS fitting technique on a second difference spectrum. (a) Four different spectra exhibiting the  $M_{45}$  white lines for the light rare-earth elements from 750 to 1000 eV. (b) Best fit for the detection of the La signal corresponding to a nominal concentration of 76 ppm. (Courtesy of Leapman and Newbury [42].)

edge of a transition metal. Figure 7 demonstrates how the MLS fitting technique described by Manoubi et al. [41] can be used to evaluate separately the weight of the two white lines ( $I_{L3}$  and  $I_{L2}$ ) and of the transitions to continuum states when trying to simulate an experimental Fe  $L_{23}$  edge as a sum of two Lorentzian curves for the white lines and of a Hartree–Slater cross-section in order to account for transitions to continuum states.

### 1.3.4 Applications of EELS Imaging

As described above, EELS spectroscopy and imaging techniques provide a great variety of approaches, adding a new dimension to TEM, that is, access to more refined quantitative or analytical microscopies (see, for example, Reimer et al. [43]). In this Section my intention is to illustrate some recent trends in the use



**Figure 7.** Use of a MLS fitting technique to measure the relative importance of different components in the fine structure of an FeL<sub>23</sub> edge. The specimen is a transverse section of an Fe/Fe<sub>x</sub>Ge<sub>1-x</sub> multilayer. An accurate measurement of the weight of the white lines and of their ratio is required for investigating the distribution of the electron states and the magnetic moment on the excited Fe atom. (Study performed in collaboration with K. Suenaga.)

of an imaging EELS spectrometer (either of the filter type or as a STEM + PEELS combination) in the field of materials science. Applications cover situations where the analysis is performed over many pixels, extending from a line profile of typically 100 points of analysis, to a full image comprising typically 1000 × 1000 pixels. As already explained, in the first case it is easy to record and process many spectral channels for each pixel, while in the latter case one can practically record only a limited number of images.

#### 1.3.4.1 Advances in EELS Elemental Mapping

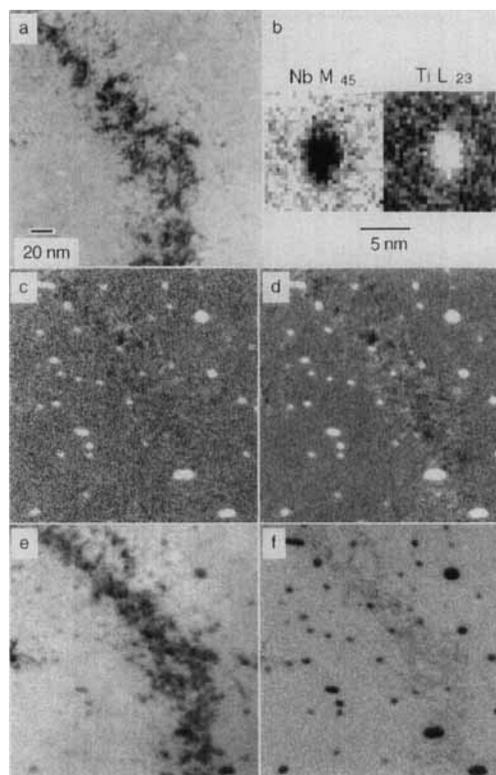
In essence, one makes a map of the spatial origin of chemically significant signals such as characteristic core edges. It is demonstrated here how the results are more quantitative when one has access to a greater number of energy loss channels

for each pixel. The characteristic signal is always superimposed on a non-characteristic background. It is therefore obvious that with a single energy-loss image it is impossible to attribute a varying contrast to a change in chemistry. On the other hand, when processing a single EELS spectrum, it has been pointed out that the highest accuracy can only be obtained when the whole spectrum is known. In the case of moderately thick specimens, this includes the low-loss domain which is needed for making deconvolution corrections. In intermediate cases, the major problem is to discriminate changes in contrast due to compositional changes from those due to variations in thickness or orientation.

The first real progress in this direction was made when Jeanguillaume and co-workers [44, 45] proposed and realized the three-window technique in which two energy-filtered background images are acquired below the edge and one above the edge. An extrapolated background

image is then calculated, independently for each pixel, and subtracted from the image above the ionization threshold, thus providing a net image which can be considered as an elemental map. This technique avoids the risk of artifacts but is not strictly quantitative. In fact, the results should be scaled to the normal unfiltered bright field image corresponding to the  $I_0$  term in Eq. (3). Nevertheless, the technique is a clear improvement on the two-window method in which a post-edge image was subtracted from a 'scaled' pre-edge image. However, another type of two-window method, known as the 'jump-ratio map', has been proposed recently [46]. The elemental map is simply made by dividing a post-edge by a pre-edge image. Hofer et al. [47] have demonstrated (Fig. 8), that this method produces maps with minimum added noise and which are only weakly affected by diffraction contrasts arising in crystalline materials. However, without denying these successes, I would not recommend systematic use of this technique because the interpretation of the measured contrast in terms of elemental composition is far from quantitative.

On the contrary, it is clear that any increase in the amount of information available and used leads to better quantification of the data. Crozier [48] has pointed out that for large-area mapping of specimens with complex morphology and varying thickness, inspection of the  $t/\lambda$  map derived from simultaneously recorded unfiltered and zero-loss images is a good basis on which to define which pixels in an image may be strongly affected by plural scattering. In an effort to assess the qualities of the different software available for elemental mapping from a



**Figure 8.** EELS mapping of nanometer-sized precipitates by processing several energy-filtered images acquired with an imaging filter attached on a 200-keV transmission electron microscope. The specimen was a thin foil of niobium containing TiO precipitates. The TiL<sub>23</sub> edge was used for Ti mapping, and the niobium M<sub>45</sub> edge was used for Nb mapping (b). The results obtained using the three-window technique (c, e), and the jump-ratio map (d, f). (a) Normal unfiltered TEM bright field micrograph showing the weak contrast of the precipitates close to a bend contour. (Data courtesy of F. Hofer et al.)

sequence of energy filtered images, Bonnet et al. [49] have pointed out the advantage of collecting an increased number of energy-filtered images both below and above the threshold in order to improve the quality of the fit and to enhance the signal per pixel. Furthermore, the signal-to-noise ratio (SNR) may benefit from local smoothing of the experimental data

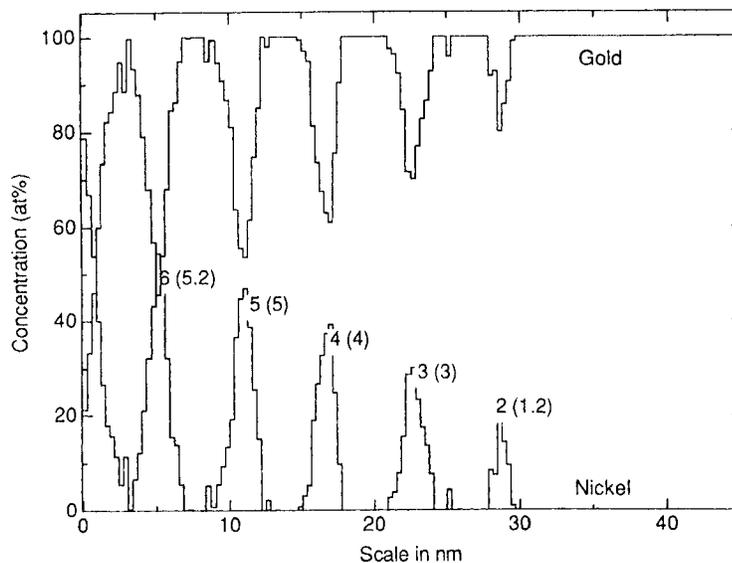
before processing, as long as it does not destroy significantly the spatial resolution. A test specimen consisting of clusters comprising a small number of rare-earth or uranide atoms has been investigated in detail by Mory and Colliex [50] in order to estimate the limits of detection of the chemical mapping techniques. Introducing double exposure during the acquisition sequence (three pre-edge images and one post-edge image), these authors have demonstrated that the elemental identification of numbers of atoms below 10 can be performed with a good SNR. In a further step, it has been demonstrated that the implementation of linear optimal unbiased filters, such as those provided by geostatistical techniques (kriging procedure), could improve the SNR further, i.e. progressing toward single-atom identification [51].

The relationship between the different types of accessible performance (spatial resolution, minimum detectable number of atoms, and minimum detectable concentration) as a function of the physical constraints involved (SNR and cross-section for the used edge, and local specimen thickness) and of the experimental parameters used (primary dose and primary voltage) has been discussed by many authors, in particular by Colliex [52]. The examples cited above (concentrations of approximately  $10^{-5}$  [42], and identification of a single atom [50, 51]) correspond to well-defined conditions in which the total current prevails in the first case and the local dose (need for a field emission source) in the second case. The intermediate case, where a filtering device on a TEM is used for the quantitative analysis of nanostructures, has been discussed theoretically and practically for typical

examples (interfaces and nanotubes) by Berger et al. [53] and by Kurata et al. [54]. In both cases a subnanometer spatial resolution and a detection limit of a few atoms has been demonstrated, the signal-to-noise ratio being noticeably improved when summing the data over several equivalent pixels.

In all these cases, the edge is sufficiently clear so that the technique of extrapolating the background is sufficient to discriminate and measure the characteristic signal. However, as discussed above, in the case of single spectrum processing, there are many situations that deviate from this simple approach, in particular when the SNR is so low that the signal cannot be distinguished from fine structures on an edge at lower energy, or for very low concentration. It is then necessary to use some kind of MLS technique and the only possible means of making the calculations is to have access to a complete energy-loss spectrum for each pixel. The unique solution is the STEM + PEELS approach. Tencé et al. [55] have recently described some applications, using this type of routine for fitting the experimental spectrum acquired for each pixel with a linear combination of reference edges. When these references are recorded during the acquisition of other parts of the specimen, the technique is very powerful, as has been demonstrated by the analysis of an Au–Ni multilayer (Fig. 9). In this example, the situation is made very complex by the superposition of the edges involved (strong overlap between the Au  $O_{23}$  and  $N_{67}$  edges at 54 and 83 eV, respectively and the Ni  $M_{23}$  edge at 68 eV, requiring a MLS fit on the second-difference spectra) and by the spatial resolution required to identify a monolayer of Ni. Another highly

**Figure 9.** Chemical analysis of a transverse section of a Au–Ni multilayer obtained by processing a sequence of 64 second-difference spectra (energy-loss range 40–120 eV) acquired with step increments of 0.6 nm. It is important that the reference Au and Ni profiles are extracted from spectra recorded in the same sequence from areas containing only gold or nickel. Numbers in parentheses are proportional to the summed signal for each Ni layer; the other numbers correspond to the nominal values obtained during the molecular beam epitaxy deposition process. (Specimen courtesy of P. Bayle and J. Thibault.)



impressive study has been performed by Leapman et al. [56] concerning the mapping of Ca at very low concentration (a few  $10^{-5}$  at. conc.) in a cryosection of cerebellar cortex. Here again MLS fitting techniques in the second difference mode were used on spectra summed within segmented areas defined on maps or images obtained at higher doses. These two types of result demonstrate clearly the trends to be followed for extracting the optimum information from EELS elemental analysis in situations at the limit.

### 1.3.4.2 Beyond Elemental Analysis: Mapping of EELS Fine Structures

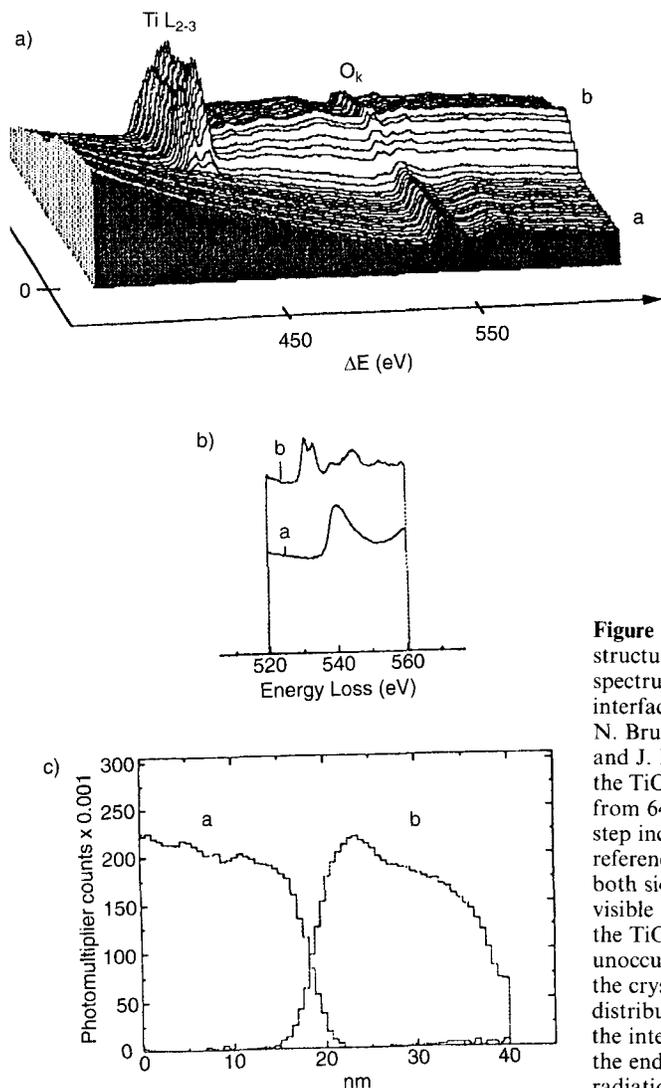
Having recorded a series of spatially resolved EELS spectra, either in the spectrum-line or the spectrum-image mode, it is possible to extend the information

extracted beyond simple elemental mapping or profiling by considering the complete spectral distribution over all relevant channels. Generally, the main interest is to identify which spectral feature is changing and at which probe position it is changing. Practically, the useful changes concern only a reduced number of energy-loss channels and a small number of pixels. Data-processing tools are being developed that will be able to answer these questions in an unbiased fashion.

The first solution is to calculate spatial-difference spectra, defined as the difference between spectra recorded at selected positions on the specimen; for example, on the matrix on each side of an interface and on the boundary itself [57]. This approach has been shown to be quite efficient in determining the bonding state and electronic structure of atoms at the interface itself, and thus opening up new avenues of investigating the atomic bonding

ies across interfaces [58, 59]. How-  
 is approach must be performed  
 ly, spectrum after spectrum, and  
 r tedious and subjective.  
 entioned earlier, the present aim is  
 rove the quantitative aspects of  
 g or mapping EELS fine structures,  
 nsequently any associated infor-  
 such as bonding type, valence

state, d-state occupancy, and site sym-  
 metry. The means of achieving this  
 involves the use of MLS fitting techniques,  
 the preliminary step being to define or  
 select templates of the fine structures  
 which are to be searched along the line  
 scan. This definition can be achieved  
 a priori, relying on the knowledge of the  
 specimen and of the problem (for instance



**Figure 10.** Mapping of the fine structures on a O K edge along a spectrum line across a SiO<sub>2</sub>-TiO<sub>2</sub> interface (in collaboration with N. Brun, K. Yu-Zhang, D. Hurtaud and J. Rivory). (a) Spectrum line from the TiO<sub>2</sub> area into the SiO<sub>2</sub> area, made from 64 spectra (acquisition time 8 s, step increments 0.6 nm). (b) Selected reference O K edge profiles selected on both sides of the interface; the doublet visible on the front side of the edge in the TiO<sub>2</sub> case reflects the existence of unoccupied O p-Ti d-orbitals split by the crystal field effect. (c) MLS fit distribution of both references across the interface; the intensity decrease at the end of the scan is due to increased radiation damage on the TiO<sub>2</sub> side.

by selecting reference spectra in the sequence far from the regions where spectral variations are likely to occur). The definition can also be realized using pattern recognition and classification techniques which sort all the spectra of a sequence into families exhibiting a certain degree of similarity. One of these approaches, the use of artificial neural networks (ANNs) has recently been introduced for the analysis of near-edge structure components across Si–SiO<sub>2</sub> interfaces [60]. The basic idea is to consider any spectrum as a vector in a multidimensional space, the number of dimensions being equal to the number of energy-loss channels in the range of interest. Several criteria (e.g. scalar product and cross-correlation coefficient) can be used to select the basis vectors of this space, on which the projection coefficients of all spectra in the sequence provides a measurement of their weight in terms of the basis spectra. Other methods, such as the factorial analysis of correspondence, would provide other routes to classifying in a hierarchical order the variance of components within a large data set [61]. In all cases the full impact of these classification and analysis tools relies on a satisfactory physical understanding of the basis vectors or principal components.

It is obvious that these novel possibilities will open up quite large fields of, as yet undefined, applications. Interfaces, multilayers, nano-objects, and clusters constitute classes of objects particularly well suited to such subnanometer mapping of electronic structures. Among the many problems under present investigation [62] is the mapping of the p–d electron-state hybridization across a SiO<sub>2</sub>–TiO<sub>2</sub> interface (Fig. 10).

The future of EELS mapping therefore relies on several developments in terms of instrumentation, with access to atomic-size electron probes such as the one demonstrated by Batson [63] and theoretical interpretation, applying for instance to the low-loss domain (maps of dielectric constants, optical coefficients, adhesion forces, etc.).

## Acknowledgements

This review is the fruit of the collaboration of many colleagues over two decades, and I am very pleased to thank all of them who have contributed at different stages: R. Castaing, B. Jouffrey, P. Trebbia, C. Mory, C. Jeanguillaume, O. L. Krivanek, M. Tencé, N. Bonnet, E. Lefèvre, M. G. Walls, D. Bouchet, N. Brun, O. Stephan, and P. Redlich.

## 1.3.5 References

- [1] H. Raether, *Springer Tracts Modern Phys.* **1965**, 38, 85.
- [2] S. E. Schnatterly, *Solid State Phys.* **1979**, 14, 275.
- [3] H. Raether, *Springer Tracts Modern Phys.* **1980**, 88, 1.
- [4] P. Schattschneider, *Fundamentals of Inelastic Electron Scattering*, Springer Verlag, Wien, **1986**.
- [5] J. Fink, *Topics Appl. Phys.* **1992**, 69, 203.
- [6] J. Hillier, R. F. Baker, *J. Appl. Phys.* **1944**, 15, 663.
- [7] D. B. Wittry, *Br. J. Appl. Phys.* **1969**, 3, 1757.
- [8] A. V. Crewe, M. S. Isaacson, D. E. Johnson, *Rev. Sci. Instrum.* **1971**, 42, 411.
- [9] O. L. Krivanek, P. R. Swann, in *Quantitative Microanalysis with High Spatial Resolution* (Eds. G. W. Lorimer, M. H. Jacobs, P. Doig), The Metals Society, London **1981**, p. 136.
- [10] O. L. Krivanek, C. C. Ahn, R. B. Keeney, *Ultramicroscopy* **1987**, 22, 103.
- [11] D. Bouchet, C. Colliex, P. Flora, O. L. Krivanek, C. Mory, M. Tencé, *Microsc. Microanal. Microstruct.* **1990**, 1, 443.
- [12] O. L. Krivanek, C. Mory, M. Tencé, C. Colliex, *Microsc. Microanal. Microstruct.* **1991**, 2, 257.

- [13] R. D. Leapman, in *Transmission Electron Energy Loss Spectrometry in Materials Science* (Eds. M. M. Disko, C. C. Ahn, B. Fultz), TMS, Warrendale **1992**, p. 47.
- [14] M. K. Kundmann, O. L. Krivanek, *Microsc. Microanal. Microstruct.* **1991**, 2, 245.
- [15] P. E. Batson, *Rev. Sci. Instrum.* **1986**, 57, 43.
- [16] R. Castaing, L. Henry, *C. R. Acad. Sci., Ser. B* **1962**, 255, 76.
- [17] A. El Hili, *J. Microsc.* **1966**, 5, 669.
- [18] C. Colliex, B. Jouffrey, *C. R. Acad. Sci., Ser. B* **1970**, 270, 673.
- [19] S. Senoussi, L. Henry, R. Castaing, *J. Microsc.* **1971**, 11, 19.
- [20] G. Zanchi, J. P. Perez, J. Sévely, *Optik* **1975**, 43, 495.
- [21] H. Rose, W. Pejas, *Optik* **1979**, 54, 235.
- [22] S. Lanio, *Optik* **1986**, 73, 99.
- [23] H. Shuman, A. P. Somlyo, in *Analytical Electron Microscopy* (Ed. R. H. Geiss), San Francisco Press, San Francisco, CA **1981**, 202.
- [24] O. L. Krivanek, A. J. Gubbens, N. Dellby, *Microsc. Microanal. Microstruct.* **1991**, 2, 315.
- [25] O. L. Krivanek, A. J. Gubbens, N. Dellby, C. E. Meyer, *Microsc. Microanal. Microstruct.* **1992**, 3, 187.
- [26] C. Jeanguillaume, C. Colliex, *Ultramicroscopy* **1989**, 28, 252.
- [27] J. A. Hunt, D. B. Williams, *Ultramicroscopy* **1991**, 38, 47.
- [28] R. D. Leapman, J. A. Hunt, in *Microscopy: The Key Research Tool*, EMSA **1992**, p. 39.
- [29] R. D. Leapman, J. A. Hunt, R. A. Buchanan, S. B. Andrews, *Ultramicroscopy* **1993**, 49, 225.
- [30] J. L. Lavergne, J. M. Martin, N. Belin, *Microsc. Microanal. Microstruct.* **1992**, 3, 517.
- [31] C. Colliex, M. Tencé, E. Lefèvre, C. Mory, H. Gu, D. Bouchet, C. Jeanguillaume, *Mikrochim. Acta* **1994**, 114/115, 71.
- [32] M. Tencé, M. G. Walls, C. Jeanguillaume, C. Colliex, X. Thomas, O. Jbara, J. Cazaux, *Inst. Phys. Conf. Ser.* **1989**, 98, 311.
- [33] C. Colliex, *Adv. Opt. Electron Microsc.* **1984**, 9, 65.
- [34] R. F. Egerton, *Electron Energy Loss Spectroscopy in the Electron Microscope*, Plenum, New York, **1986**.
- [35] M. M. Disko, C. C. Ahn, B. Fultz, *Transmission Electron Energy Loss Spectrometry in Materials Science*, TMS, Vol. 2, Warrendale, **1992**.
- [36] R. F. Egerton, *Ultramicroscopy* **1979**, 4, 169.
- [37] R. D. Leapman, P. Rez, D. F. Mayers, *J. Chem. Phys.* **1980**, 72, 1232.
- [38] F. Hofer, *Microsc. Microanal. Microstruct.* **1991**, 2, 215.
- [39] H. Shuman, A. P. Somlyo, *Ultramicroscopy* **1987**, 21, 23.
- [40] R. D. Leapman, C. R. Swyt, *Ultramicroscopy* **1988**, 26, 393.
- [41] T. Manoubi, M. Tencé, M. G. Walls, C. Colliex, *Microsc. Microanal. Microstruct.* **1990**, 1, 23.
- [42] R. D. Leapman, D. E. Newbury, *Anal. Chem.* **1993**, 65, 2409.
- [43] L. Reimer, I. Fromm, C. Hülk, R. Rennekamp, *Microsc. Microanal. Microstruct.* **1992**, 3, 141.
- [44] C. Jeanguillaume, P. Trebbia, C. Colliex, *Ultramicroscopy* **1978**, 3, 237.
- [45] C. Jeanguillaume, M. Tencé, P. Trebbia, C. Colliex, *Scanning Electron Microscopy*, Vol. 2, SEM Inc, AMF O'Hare, Chicago, IL **1983**, p. 745.
- [46] O. L. Krivanek, A. J. Gubbens, M. K. Kundmann, G. C. Carpenter, *Proc. 51st EMSA Meeting* (Eds. G. W. Bailey, C. L. Rieder), San Francisco Press, San Francisco, CA **1993**, p. 586.
- [47] F. Hofer, P. Warblicher, W. Grogger, *Ultramicroscopy* **1995**, 59, 15.
- [48] P. A. Crozier, *Ultramicroscopy* **1995**, 58, 157.
- [49] N. Bonnet, C. Colliex, C. Mory, M. Tencé, *Scanning Microscopy Supplement 2*, Scanning Microscopy International, AMF O'Hare, Chicago, IL **1988**, p. 351.
- [50] C. Mory, C. Colliex, *Ultramicroscopy* **1989**, 29, 339.
- [51] B. Bhanu Prasad, D. Jeulin, C. Daly, C. Mory, M. Tencé, C. Colliex, *Microsc. Microanal. Microstruct.* **1991**, 2, 107.
- [52] C. Colliex, *Microsc. Microanal. Microstruct.* **1991**, 2, 403.
- [53] A. Berger, J. Mayer, H. Kohl, *Ultramicroscopy* **1994**, 55, 101.
- [54] H. Kurata, S. Isoda, T. Kobayashi, *Microsc. Microanal. Microstruct.* **1995**, 6, 405.
- [55] M. Tencé, M. Quartuccio, C. Colliex, *Ultramicroscopy* **1995**, 58, 42.
- [56] R. D. Leapman, J. A. Hunt, R. A. Buchanan, S. B. Andrews, *Ultramicroscopy*, **1993**, 49, 225.
- [57] H. Müllejjans, J. Bruley, *Ultramicroscopy* **1994**, 53, 351.
- [58] J. Bruley, M. Tseng, D. B. Williams, *Microsc. Microanal. Microstruct.* **1995**, 6, 1.
- [59] C. Scheu, G. Dehm, H. Müllejjans, R. Brydson, M. Rühle, *Microsc. Microanal. Microstruct.* **1995**, 6, 19.
- [60] C. Gatts, G. Duscher, H. Müllejjans, M. Rühle, *Ultramicroscopy* **1995**, 59, 229.
- [61] P. Trebbia, N. Bonnet, *Ultramicroscopy* **1990**, 34, 165.
- [62] N. Brun, **1995**, private communication.
- [63] P. Batson, *Nature* **1993**, 366, 727.

# 1.4 High Voltage Electron Microscopy

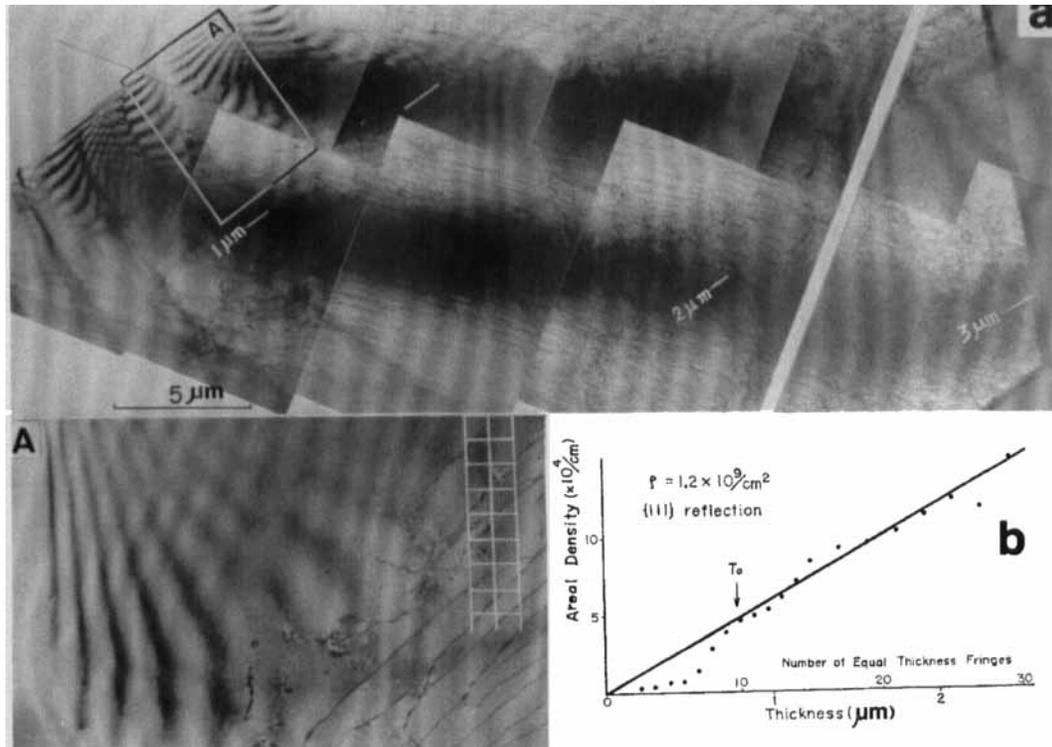
## 1.4.1 Introduction

Because the behavior of materials is determined by lattice defects (point defects, dislocations, etc.), much attention has been paid to their direct observation, particularly to individual dislocations [1, 2]. Lattice defects, however, have strong strain fields associated with them in crystals, and thus their behavior is very sensitive to the thickness of the specimen. Figures 1 and 2 show the thickness effect on the dislocation density [3] and the cell formation [4] in aluminum crystals, respectively. The dislocation density decreases quite remarkably (Fig. 1) when the thickness becomes smaller than a critical value (i.e., 0.4–0.8  $\mu\text{m}$  in cold-worked aluminum). This is a function of both the kind of material and the dislocation density [3]. With no applied stress, the position of each individual dislocation is determined by the local internal stress, which is a function of the dislocation density and the interaction among different families of dislocations in the local region, in addition to the image force. Therefore, to observe the same distribution of dislocations as that in bulk materials, the critical thickness increases more than that required for the dislocation density [5]. The formation of cell structures depends on the mutual interaction among

moving dislocations, and thus the specimen thickness for this process must be larger than the mean free path of dislocations. Thus, the critical thickness increases further to a few micrometers to become representative for the cell structures of bulk specimens [5]. For example, the critical thickness in Fig. 2 is about 3  $\mu\text{m}$ . Additionally, recovery and recrystallization, phase transformations, etc. are also very sensitive to the specimen thickness [3].

The above facts show that not only the behavior of lattice defects but also their density markedly changes when the specimen thickness is smaller than the critical value which is a function of the behavior of related lattice defects.

However, the maximum observable specimen thickness with the 100 kV class electron microscopes (EMs) is smaller than these critical thicknesses in general, especially for the dislocation behavior [5]. For this reason there is a need for high voltage electron microscopes (HVEMs) for which the maximum observable thickness of the specimens is larger than the critical values. For such purposes, a few 0.5 MV HVEMs were constructed despite the technical difficulties (which are considered to be roughly proportional to the third power of the accelerating voltage required to keep the same mechanical and electrical stability at higher accelerating voltages).



**Figure 1.** Thickness effect on the dislocation density in aluminum [3]. (a) Specimen was finished wedge-shaped after deformation at room temperature; micrograph was taken at 0.5 MV. Equal-thickness fringes appear clearly near the specimen edge at the left-hand side. Micrograph A shows an enlargement of the framed part A; (b) shows the relationship between the areal density of dislocations and the specimen thicknesses indicated in micrograph (a).

In 1965 it was found that the same dislocation behavior as that in bulk materials can be observed at 0.5 MV in light metals whose atomic number is smaller than about 20, and in situ experiments were carried out on various phenomena in such metals [3, 5].

Since 1965, the HVEMs and their related accessories have rapidly developed and widely applied to the natural sciences, especially materials science. Two ultra-HVEMs operating at 3 MV have been installed independently in Osaka and Toulouse in 1970, and, to date, about 60 HVEMs, whose accelerating voltages are

higher than 0.5 MV, have been installed around the world [6].

This account deals with the usefulness and the applications of HVEMs, mainly in the field of materials science.

### 1.4.2 Voltage Dependence of Operational Features of Electron Microscopes

It was shown experimentally [7] that with increasing accelerating voltage both



**Figure 2.** Thickness effect on the formation of cell structures in aluminum [4]. The wedge-shaped specimen was deformed by HVEM to form cell structures, and the micrograph was taken at 0.5 MV. Local thicknesses of the specimen are indicated in the micrograph.

spherical and chromatic aberrations decrease remarkably as well as the wavelength of electrons. Furthermore, the following advantages can be obtained at higher accelerating voltages: (a) a decrease in contamination of the specimens during observation [8]; (b) a decrease in irradiation damage in biological and polymer specimens [9]; (c) a decrease in temperature

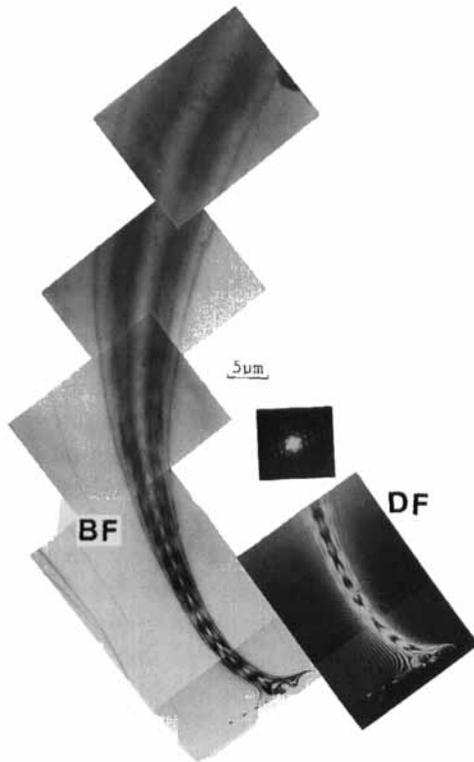
rise of the specimens during observation [6]; and (d) an increase in the simultaneous excitation of many waves. Point (d) is closely related to a phenomenon called electron channeling [10] or the Bloch wave channeling [11, 12], as mentioned below. Increase of the maximum observable thickness and visibility of lattice defects in adjacent different crystals at the same time are also related to this phenomenon [6]. Additional advantages are: (e) an easy determination of the structure factor of crystals [13, 14] and (f) accurate studies of electron irradiation damage in materials [15, 16]. With regard to (f) electron irradiation damage during observation can be effectively suppressed under suitable conditions, as mentioned below [6]. There are a number of other applications [6].

#### 1.4.2.1 Electron Channeling at High Accelerating Voltages

A theoretical estimation was made in 1962 of the voltage dependence on the penetration power of electrons [17]. The result showed that the penetration power is almost saturated at about 500 kV, and that it increases only 3.3 times as compared to that at 100 kV even when the accelerating voltage increases up to infinity. Many electron microscopists concluded from these results that the penetration power is proportional to the maximum observable thickness of the specimens, and this resulted in a pessimistic estimation of the voltage dependence on the maximum observable thickness.

However, many-beam effects become very important for the maximum observable thickness at high voltages. When the

accelerating voltage increases, a bright band appears at the symmetry position of bend contours. This phenomenon is due to the many-beam effect on the anomalous transmission of electrons. In the case of aluminum, the anomalous transmission at the symmetry position of (1 1 1) systematic reflections begins to appear clearly at about 1 MV, and its intensity remarkably increases with increasing accelerating voltage [6]. Figure 3 shows the bright-field image (BF) and a dark-field image (DF) of the (1 1 1) extinction contour at 3 MV.



**Figure 3.** Anomalous transmission of electrons at the symmetry position of (1 1 1) reflections in aluminum at 3 MV [6]. Micrographs at the left- and right-hand sides show the bright (BF) and dark-field (DF) images, respectively. An electron diffraction pattern was taken at the symmetry position.

A bright band appears at the symmetry position of (1 1 1) systematic reflections in the BF, and at both sides of the (1 1 1) contour in the DF.

The axial and planar channeling of electrons can be estimated from the effective width of Kikuchi bands [12]. The axial ( $\alpha_a$ ) and planar ( $\alpha_p$ ) channeling parameters are defined as follows:

$$\alpha_a = \frac{4mZ_1Z_2e^2}{h^2Nd^2} \quad (1)$$

$$\alpha_p = \frac{4mV_h}{h^2b_h^2} \quad (2)$$

where  $Z_1 = 1$ , and  $m, Z_2, e, h, N$  and  $d$  in Eq. (1) are the relativistic mass of an electron, the atomic number of crystal atoms, the charge of an electron, Planck's constant, the atomic density, and the distance along the crystal axis, respectively, and  $V_h$  and  $b_h$  in Eq. (2) are the Fourier coefficient of the crystal potential and the reciprocal vector, respectively. Experimental results show that the classical model is adequate when these two parameters are larger than unity (i.e.,  $\alpha_p > 1.1$ ) and that the degree of anomalous transmission at the symmetry position of reflections is also well correlated with the value of  $\alpha_p$  irrespective of both the kind of crystal and the indices of the reflection [18].

Table 1 shows the calculated  $\alpha_a$  and  $\alpha_p$  by Eqs. (1) and (2). The values of both channeling parameters are proportional to the electron intensity passing through the material. Thus it is concluded, from Eqs. (1) and (2), that the electron intensity increases linearly with increasing accelerating voltage, because the relativistic mass of electrons linearly increases with the accelerating voltage. This means that the maximum observable specimen thickness

**Table 1.** (a) Calculated axial channeling parameter,  $\alpha_a$ .

Voltage (MV)	Beam direction		
	$\langle 110 \rangle$	$\langle 200 \rangle$	$\langle 222 \rangle$
0.5	3.1 <sub>7</sub>	1.5 <sub>8</sub>	0.5 <sub>2</sub>
1.0	4.7 <sub>4</sub>	2.3 <sub>7</sub>	0.7 <sub>9</sub>
1.5	6.3 <sub>1</sub>	3.1 <sub>5</sub>	1.0 <sub>5</sub>
2.0	7.8 <sub>8</sub>	3.9 <sub>4</sub>	1.3 <sub>1</sub>
2.5	9.4 <sub>5</sub>	4.7 <sub>2</sub>	1.5 <sub>7</sub>
3.0	11.0 <sub>2</sub>	5.5 <sub>1</sub>	1.8 <sub>3</sub>

**Table 1.** (b) Calculated planar channeling parameter,  $\alpha_p$ .

Voltage (MV)	Order of reflection		
	$\{111\}$	$\{200\}$	$\{220\}$
0.5	0.8 <sub>9</sub>	0.5 <sub>5</sub>	0.1 <sub>7</sub>
1.0	1.3 <sub>3</sub>	0.8 <sub>2</sub>	0.2 <sub>6</sub>
1.5	1.7 <sub>7</sub>	1.1 <sub>0</sub>	0.3 <sub>5</sub>
2.0	2.2 <sub>1</sub>	1.3 <sub>7</sub>	0.4 <sub>3</sub>
2.5	2.6 <sub>5</sub>	1.6 <sub>4</sub>	0.5 <sub>2</sub>
3.0	3.0 <sub>9</sub>	1.9 <sub>2</sub>	0.6 <sub>1</sub>

is expected to increase linearly with increasing accelerating voltage when only the electron intensity is taken into account.

The electron intensity also increases with increasing the atomic density along the electron path, that is, with decreasing  $d$  in Eq. (1) and  $b_h$  in Eq. (2), respectively.

Two important phenomena, due to electron channeling, are: (a) the orientation dependence of electron irradiation damage and (b) the straight path of electrons along the incident direction in materials. According to (a), the rate of electron irradiation damage is proportional to the electron intensity, and thus the damage rate remarkably decreases when higher order reflections and/or higher order directions are used, as mentioned above [19]. Furthermore, an image-intensifier system is always used for the observation of images with the HVEMs.

Thus, the intensity of the incident beam can be decreased by more than three orders of magnitude. Specimen damage during observation can therefore be suppressed to an extremely small level. According to (b), electron channeling occurs such as the total reflection on inner surfaces of a fine column on an atomic scale along the incident electron beam, as recognized by the electron irradiation induced foreign-atom implantation (below).

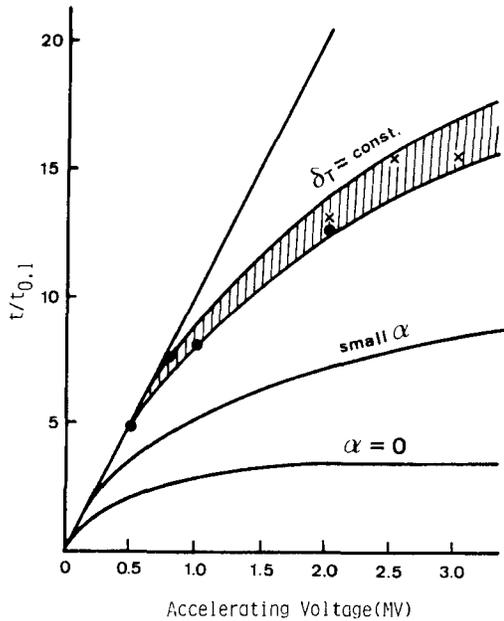
#### 1.4.2.2 The Objective Aperture Effect on the Maximum Observable Specimen Thickness

As already mentioned, electron intensity, transmitted through materials by electron channeling, increases linearly with increasing accelerating voltage. Practically, however, various aberrations of electron microscope images must be taken into account for the determination of the maximum observable specimen thickness. Here, the maximum observable thickness is discussed under a fixed image resolution.

The total aberration ( $\delta_T$ ) of electron microscope images consists of the diffraction ( $\delta_{\text{diff}}$ ), spherical ( $\delta_{\text{sph}}$ ) and chromatic ( $\delta_{\text{chr}}$ ) aberrations. This is expressed as

$$\delta_T = (\delta_{\text{diff}}^2 + \delta_{\text{sph}}^2 + \delta_{\text{chr}}^2)^{1/2} \quad (3)$$

where  $\delta_{\text{diff}} \sim 0.6 \lambda / \alpha$ ,  $\delta_{\text{sph}} = C_s \alpha^3$ , and  $\delta_{\text{chr}} = C_c \alpha \Delta E / E$ . Here,  $\lambda$ ,  $\alpha$ ,  $C_s$  and  $C_c$ , and  $E$  and  $\Delta E$  are the electron wavelength, effective aperture angle, spherical and chromatic aberration coefficients, and energy and energy fluctuation of electrons, respectively. In these three terms,  $\delta_{\text{diff}}$  is divided by  $\alpha$ , but the terms  $\delta_{\text{sph}}$  and  $\delta_{\text{chr}}$  are



**Figure 4.** Voltage dependence of the maximum observable thickness of specimens [6, 20]. The maximum observable thickness at each voltage ( $t$ ) is normalized against that at 100 kV ( $t_{0.1}$ ). The hatched curve corresponds to measurements at a fixed resolution. Marks in the hatched curve show measured values on various metals and alloys.

multiplied by  $\alpha^3$  and  $\alpha$ , respectively. Thus, there is an optimum aperture size that minimizes  $\delta_T$  given by Eq. (3), depending on the accelerating voltage. Under this condition, the diffraction contrast of images clearly results from higher order excited waves, which are cut off by the objective aperture, because their intensity markedly increases at high voltages. Figure 4 shows the aperture effect on the image contrast as a function of the accelerating voltage [20]. The hatched band shown in Fig. 4 was obtained experimentally for the dislocation images in various materials with a fixed image resolution of 0.5 nm [6]. The lowest curve of  $\alpha = 0$  in Fig. 4 corresponds to a case in which only

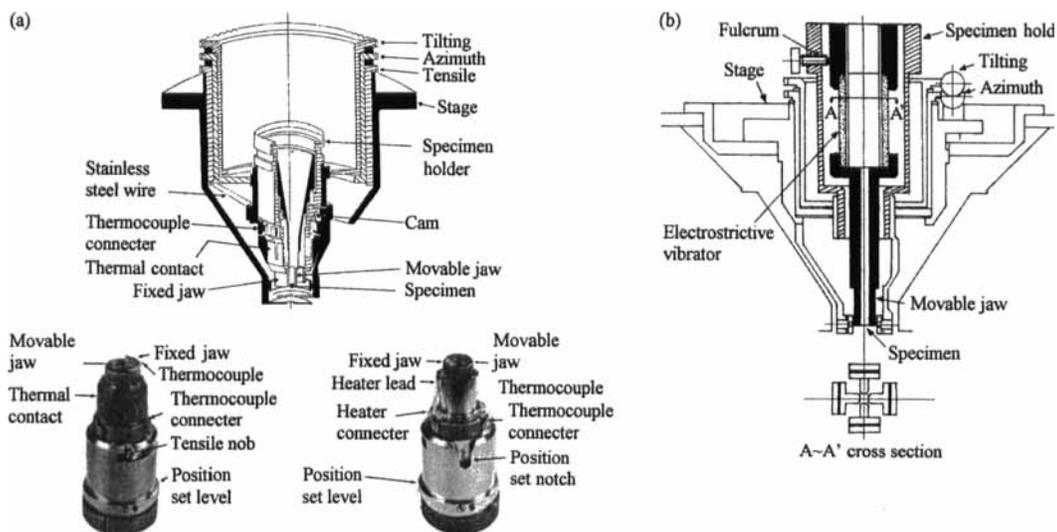
one wave is excited, that is, it represents the voltage dependence of the penetration power of electrons.

It is noted in Fig. 4 that the maximum observable thickness of specimens at 3 MV practically increases more than 15 times compared to that at 100 kV. In other words,  $t/t_{0.1} > 15$ , even when the image resolution is fixed, while it deviates from a linear relation for accelerating voltages higher than about 0.5 MV [6]. The value at 3 MV corresponds to about 3  $\mu\text{m}$  in thickness even in high  $Z$  materials such as tungsten and gold. The same behavior of all lattice defects as that in bulk materials can be observed in situ with 3 MV class HVEMs, even when the atomic number of the constitutive atoms is extremely high.

### 1.4.3 In Situ Experiments with High Voltage Electron Microscopes

#### 1.4.3.1 Specimen Treatment Devices for In Situ Experiments

As previously stated, phenomena representative for bulk materials can be investigated dynamically with HVEMs. Therefore, to carry out in situ experiments, various specimen treatment devices have been improved [6, 21]. At present, specimen manipulations such as stretching, alternating stressing, constant-load deformation, electron irradiation, treatments in various environments, and their combinations can be carried out in a wide temperature range from about 5 to 2300 K. All of these devices are of top-entry type and



**Figure 5.** In situ deformation devices [6]. (a) Stretching devices for deformation at high (right) and low (left) temperatures, respectively. These devices are placed on a goniometer stage. (b) Cross-sectional view of an alternating stressing device using an electrostrictive vibrator.

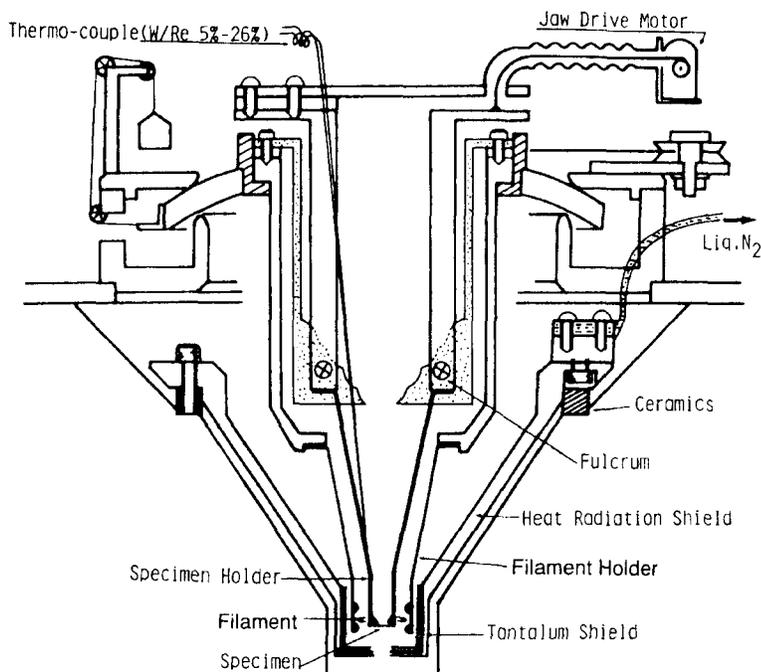
mounted on a universal goniometer stage whose maximum tilting angle is  $10^\circ$  or more [6]. Furthermore, most of these specimen holders can be easily exchanged using an air-lock system.

Figures 5–8 are some examples of those specimen treatment devices for in situ experiments [6]. Since the objective pole piece of the HVEMs is concave to decrease the chromatic and spherical aberrations, the path necessary for the specimen transfer is generally very large. However, the specimen chamber is so large that the various devices mentioned above can easily be mounted on the specimen stage, also a sufficient thermal and electromagnetic shielding as well as a long movement of the specimen holder jaws becomes possible.

Figures 5a and 5b show deformation devices for stretching and alternating stressing in fatigue deformation, respectively. Most of these devices can deform the

specimens up to about 1 kg in loading in a temperature range from about liquid nitrogen temperature to about 1200 K. Figure 6 shows a sectional diagram of a deformation stage for use at very high temperatures. The specimen is set on the tip of a tantalum cylindrical cone, and power for heating is supplied to the tantalum cone by electron beam heating with a tungsten ring-filament. Using this stage, the specimen can be tilted around any axis up to  $8^\circ$  during specimen treatment, and deformed to 50% elongation in a temperature range from room temperature to 2300 K. This stage has been effectively used for studying high temperature characteristics of refractory materials including various ceramics and their composites, as will be mentioned in a later section.

Figure 7 shows a sectional diagram and an outside view of a liquid helium temperature stage, on which the specimen can be tilted up to  $8^\circ$  around any axis during

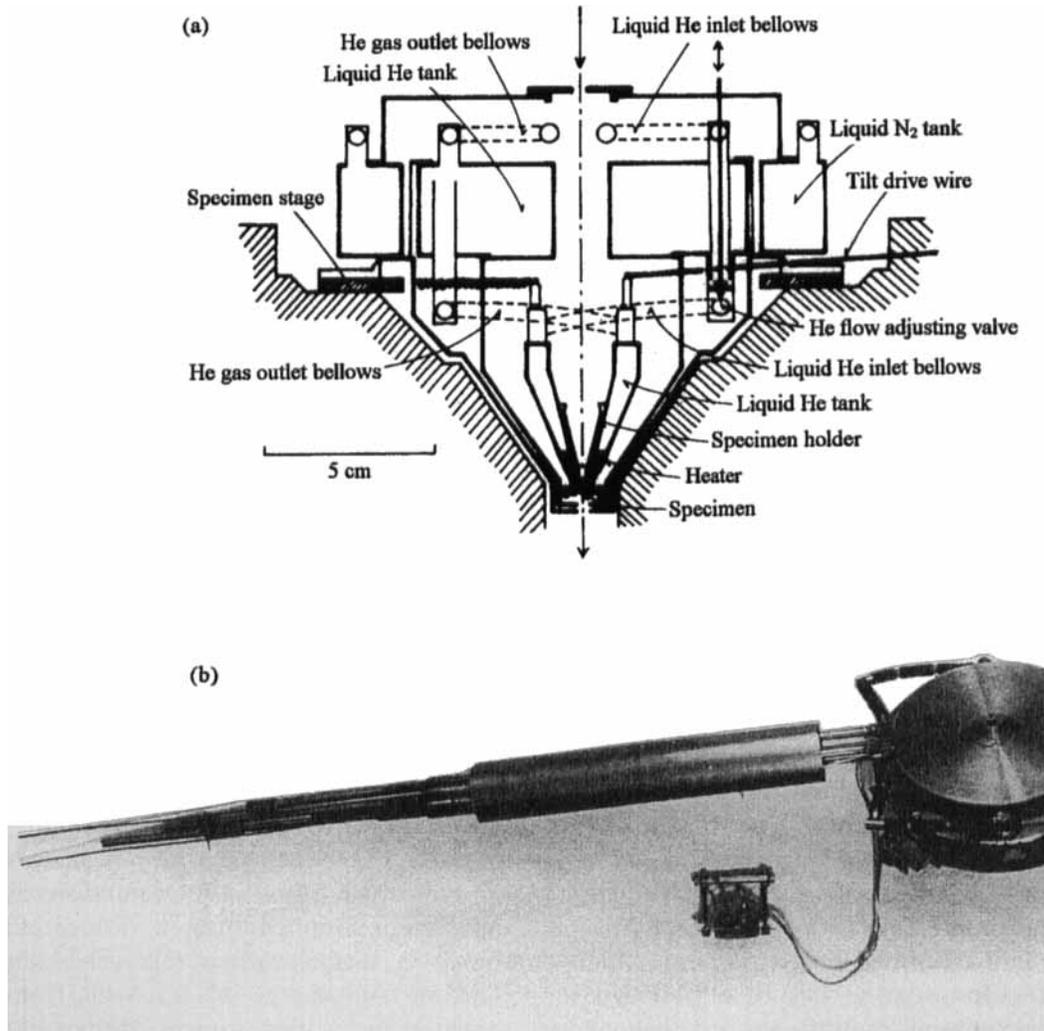


**Figure 6.** A very high temperature deformation device [6]. The specimen can be stretched up to 50% strain and tilted around any axis up to  $8^\circ$  in a temperature range from room temperature to about 2300 K.

the experiment. The stage was improved further to allow the specimens to stretch up to about 50% strain in a temperature range from room temperature to about 10 K.

Figure 8a shows a sectional diagram of a universal environmental cell that can be used for both biological and physical specimens [6]. An enlargement of the main part of the cell that consists of three films x, y, and z can be seen in Fig. 8b. The cell films are vapor-deposited aluminum, silicon oxide, and carbon films, depending on the purpose, and they are supported by metal meshes (300 and 400 mesh in size). Gases or liquids are circulated through the capsule by two pipes and the capsule is pressurized up to 300 kPa. The specimen holder can be exchanged by an air-lock system. In Fig. 8b, the specimen is set either directly on the center nickel mesh (y) or on a vapor-deposited film coated

onto the nickel mesh. The specimen is heated by directly passing an electric current through the nickel mesh, and cooled by connecting the mesh with a liquid nitrogen tank. The specimen also can be heated to high temperatures ( $>1300$  K) by passing a current directly through the specimen. Films x, y, and z are separated  $30\text{--}50\ \mu\text{m}$  from each other to allow gas exchange between these gaps. Furthermore, the specimen drift is less than  $0.1\ \text{nm s}^{-1}$  at room temperature. The environmental cell [6] has been effectively used for studying: (a) the sintering process of fine metal powders which are sensitive to oxidation and reduction; (b) chemical reactions, such as oxidation and inner oxidation, among gases, liquids, and solids; (c) the determination of the atomistic structure of liquids; (d) the observation of wet biological and microbiological specimens and their movement in salt water.

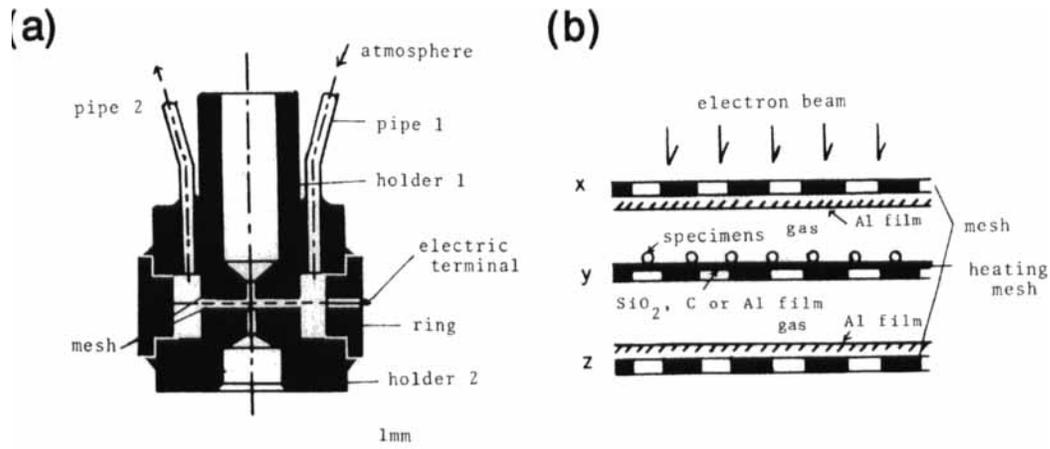


**Figure 7.** A sectional diagram of a liquid helium temperature stage [6]. The specimen can be tilted up to  $8^\circ$  around any axis during cooling.

### 1.4.3.2 Applications to Materials Science

As stated, to carry out in situ experiments the specimen thickness must be larger than the mean free path of related lattice defects, otherwise doubtful conclusions will result. Practically, at least 1.5 MV

class HVEMs are necessary for studying bulk representative dislocation behavior in materials whose atomic number is larger than about 25 (i.e., most of industrially important metals) [6, 31]. In the case of 3 MV HVEMs, the maximum observable thickness is always larger than the critical thickness for various phenomena even in very high  $Z$  materials.



**Figure 8.** (a) Sectional diagram of a universal environmental cell [6]. (b) Enlargement of a main part of the cell, which contains the specimen. This part consists of three films x, y and z. Maximum atmospheric pressure 300 kPa; maximum specimen temperature 1300 K.

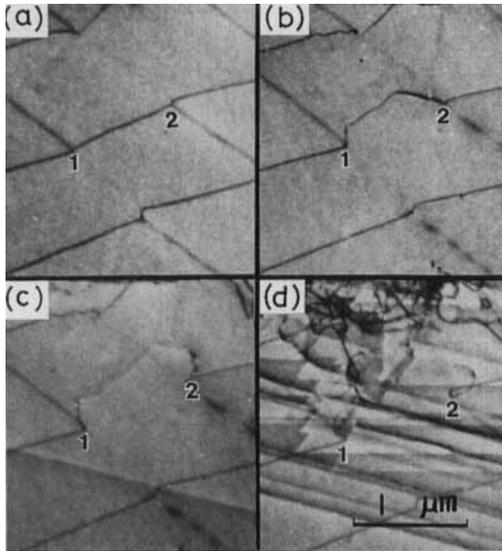
### Deformation [6, 21–31]

In situ deformation studies give valuable information on the following phenomena: (a) the behavior of individual dislocations detected by their slip traces; (b) the shape of dislocations under stress from which the magnitude of the local stress can be estimated; and (c) dislocation interactions and interaction between dislocations and microstructures. Based on these studies, new information has been obtained on the mechanisms of the various phenomena listed below.

#### (1) *Uniaxial Deformation*

At the onset of deformation, long dislocation segments of many slip systems are activated irrespective of their Schmid's factors immediately before the apparent yield stress. When the applied stress approaches the apparent yield stress the number of activated dislocation sources increases. In this process, the following facts become evident: (a) The source

dislocation cross-slips so frequently that the emitted dislocations have many super-jogs, as seen in Fig. 9 [6]. (b) Since the shear stress necessary to move the emitted dislocations rapidly increases as a result of (a), the pole dislocations are also moved outwards when the stress slightly increases, even when their Schmid factors are very small. Thus, only several dislocations are multiplied from a dislocation source in metals with a relatively high stacking fault energy [6]. (c) Aside from grown-in dislocation sources, dislocation sources are newly formed not only by double cross-slip of moving dislocations but also by the interaction between dislocations of different families. They are also formed by the interaction between a mobile dislocation and other microstructures such as small precipitates. These newly formed dislocation sources increase in number with increasing stress. Since the emitted dislocations from the sources have a large number of super-jogs in general, the passage through forest dislocations



**Figure 9.** Activation of a grown-in dislocation source due to cross slip in aluminum [6]. The micrograph was taken at 0.5 MV. When the dislocation source (1–2) begins to move, the dislocation becomes zigzag in shape by cross-slip in (b). Exact shape of the activated dislocation source can be determined from a zigzag slip trace in (c). (d) Slip traces of the dislocations emitted from single ended dislocation sources 1 and 2.

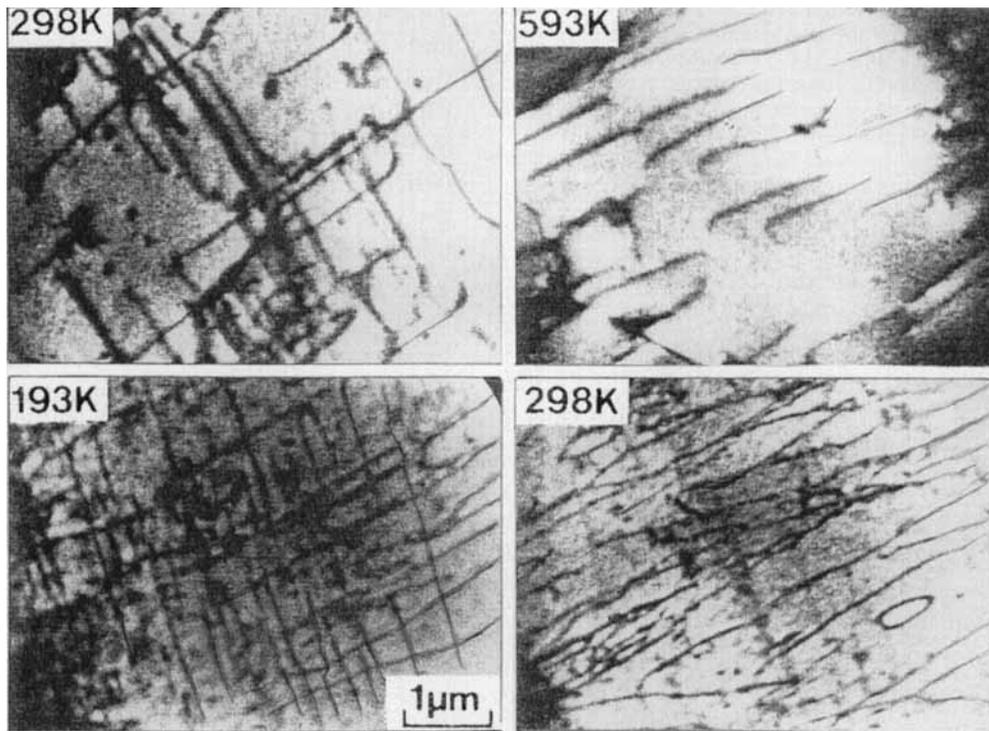
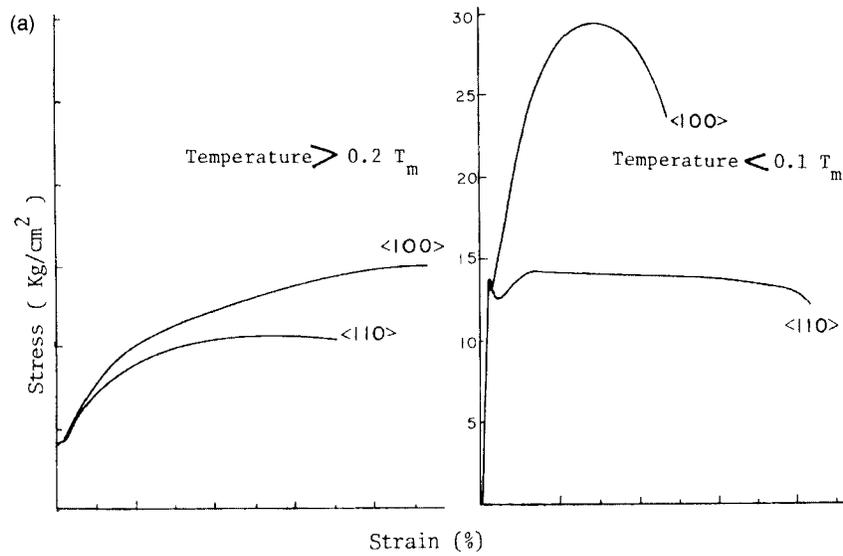
does not present a strong resistance against mobile dislocations until the spacing among forest dislocations becomes smaller than that of the super-jogs. (d) The formation mechanisms of various dislocation structures in stages I, II and III were also verified in detail, and the work hardening mechanism was investigated as a function of the type of crystals, the specimen orientation and the stacking fault energy. In these results, the role of conjugate slip in the work hardening process was verified using crystals with very low stacking fault energy [6]. Namely, the conjugate slip plays an important role in the work hardening in stage II, but it contributes to the dynamic recovery in stage III.

### (2) Deformation Under Alternating Stress

The behavior of individual dislocations is also revealed in this process, and it was found that the irreversible motion of screw dislocation plays an important role in fatigue deformation. In addition, the formation process and the dislocation structures of ladder structures were investigated in detail. These results provided useful information about the mechanism of fatigue fracture in materials [6].

### (3) Various Influences on Deformation

Deformation is very sensitive not only to the specimen orientation but also to both deformation temperature and impurity atoms. A typical example of temperature dependence is the deformation of  $\langle 100 \rangle$  and  $\langle 110 \rangle$  oriented b.c.c. crystals, as shown in Fig. 10 [6]. The upper figures in Fig. 10 show the stress-strain curves taken at temperatures of  $>0.2 T_m$  and  $<0.1 T_m$  (where  $T_m$  is the absolute temperature of the melting point), respectively. The micrographs in Fig. 10b were taken from VTR films showing dislocation movement in the  $\langle 100 \rangle$  oriented tungsten (upper pair) and iron (lower pair) crystals, respectively. From these micrographs it is concluded that the work hardening rate markedly decreases when only one slip system is operated in each individual local region of both crystals. That is, even in the  $\langle 100 \rangle$  crystal, in which four slip systems are expected to be activated geometrically, only one slip system is operative at temperatures higher than  $0.2 T_m$ , as seen at the right hand side of the micrographs in Fig. 10b. Furthermore, the temperature dependence of serrated yielding (the so-called Portevin–Le Chatelier effect) in various alloys has been investigated [6].



**Figure 10.** Temperature dependence of the work-hardening rate in the  $\langle 100 \rangle$  and  $\langle 110 \rangle$  b.c.c. crystals [6] (a). Micrographs were picked up from VTR films. Upper and lower series of micrographs show dislocation structures in the  $\langle 100 \rangle$  oriented tungsten and iron crystals taken at the same homologous temperatures, respectively (b).

#### (4) *Other Processes*

Processes of the following phenomena have also been investigated in detail: (a) dislocation behavior in the vicinity of boundaries, (b) precipitation hardening, (c) deformation processes in various ceramics, (d) electron irradiation softening of b.c.c. metals due to enhancement of both multiplication and movement of dislocations [6], (e) the effect of hydrogen gas on the dislocation behavior in b.c.c. metals, (f) dislocation channeling in particle irradiated specimens, and so on. Furthermore, it was shown that so-called 'super plasticity' easily and rapidly (with a strain rate of about  $10^{-2}$ ) occurs when the grain diameter is kept in the range of 0.1  $\mu\text{m}$ .

#### **Annealing Phenomena** [6, 19, 21–31]

These phenomena are generally determined by the diffusion of both point defects and solute atoms. When dislocations are involved in the phenomena as in the case of recovery and recrystallization, the specimen thickness must be large enough to keep the same distribution of dislocations as that in bulk specimens (i.e., a few  $\mu\text{m}$ ) [6]. The critical thickness, however, remarkably decreases at low temperatures when the phenomena are only correlated with diffusion of the point defects and solute atoms, as in the case of Guinier–Preston (G.P.) zone formation, because of a remarkable decrease of the mean free path of these defects.

##### (1) *Recovery and Recrystallization*

Various deformed specimens were annealed in situ, and the rearrangement

of dislocations, subgrain coalescence and the nucleation of new grains have been investigated in detail [6]. The results show that the misfit angle of boundary, which determines the boundary energy, increases with decreasing the dislocation density, which is directly related to the stored energy, within a recovered region during the nucleation process [32].

##### (2) *Sintering Process*

This phenomenon is very sensitive to the size of metal and ceramic powders. The coalescence among powders plays an important role when the powder size is  $\sim 0.1 \mu\text{m}$  in diameter or less [6]. Sintering and successive grain growth of ceramic composites have been also investigated in detail [19]. In ceramic composites whose melting point is higher than 2000 K, their phase diagrams can be determined by in situ annealing.

##### (3) *Precipitation Process*

The formation processes of G.P. zones [33] and precipitates have been investigated in situ in various metals and alloys, and useful information has been obtained for their formation mechanisms.

##### (4) *Melting Process*

The dislocation behavior during melting, and a peculiar phenomenon at solid–liquid interfaces [34] are important topics in the process.

#### **Martensitic Transformations** [19, 21–31, 35]

Phenomena related to this type of transformation are divided into two

groups, spontaneous and strain induced. Their formation mechanisms have been extensively investigated in both metals and ceramics by in situ experiments, especially by in situ deformation at low temperatures. From this new mechanisms of f.c.c.  $\rightarrow$  h.c.p. and f.c.c.  $\rightarrow$  b.c.c. transformations have been proposed. Furthermore, very toughened ceramics showing both plastic deformation and superplasticity at room temperature have been investigated from a view point of martensitic transformation [19].

#### **Radiation Effects [6, 21–31]**

The study of these phenomena is one of the important applications of in situ experiments, and the behavior of the primary and secondary defects induced by irradiation with high energy electrons and ions have been investigated extensively. This phenomenon is also correlated with both irradiation softening [6] of b.c.c. crystals and irradiation hardening of metals and alloys. Furthermore, the formation of nonequilibrium phases has also been carried out by the radiation effects, as mentioned later.

#### **Environment-Material Interactions [6, 21–31]**

The following subjects have been investigated by using both open- and closed-type environmental cells: (a) inner oxidation in metals and alloys, (b) hydrogen embrittlement of iron crystals, (c) chemical amorphization of metals [6], (d) crystallization

of materials with low melting points and crystal growth of various oxides, (e) observation of wet biological and living microbiological specimens [6], and others. Closed-type environmental cells have been effectively used in this research field, because the pressure of atmospheres can be widely changed up to 300 kPa so that liquid materials can also be used [6].

#### **Magnetic Domain Measurement [31]**

Various sorts of magnetic domain structures and interaction between domain walls and microstructures have been investigated by changing the strength of the magnetic field during in situ observation.

#### **Quantitative Studies [21–31, 36]**

By using three-dimensional measurements, the distribution of microstructures such as complex dislocation structures, radiation induced secondary defects, small precipitates, and inclusions have been quantitatively determined. Furthermore, the strength of applied and internal stresses in local regions can be estimated from both the change in the radius of curvature and the movement of dislocations when the applied stress is changed. The dislocation velocity can also be measured as a function of the Burgers vector and the applied stress [36]. Furthermore, atomistic structures of very fine precipitates can also be accurately estimated after sufficient growth by in situ annealing.

### 1.4.4 New Research Fields Using High Voltage Electron Microscopy

The technical difficulty of keeping the same stability of electron energy at high accelerating voltages must be resolved in the construction of HVEMs. As a result, the functional features of conventional electron microscopes have been very much improved. In particular, high resolution electron microscopes operating at high voltages from 300 kV to 1.5 MV, compact HVEMs, high vacuum electron microscopes, high resolution analytical microscopes have been developed. Using a 300 kV electron microscope, high quality many-beam lattice images of diamond and silicon crystals have been obtained [37]. Furthermore, very stable high energy electrons can also be used in other studies, such as the formation of various non-equilibrium phases, various kinds of microfabrication, and the generation of synchrotron orbital radiation induced by the rosette motion of electrons around the nucleus of each atom.

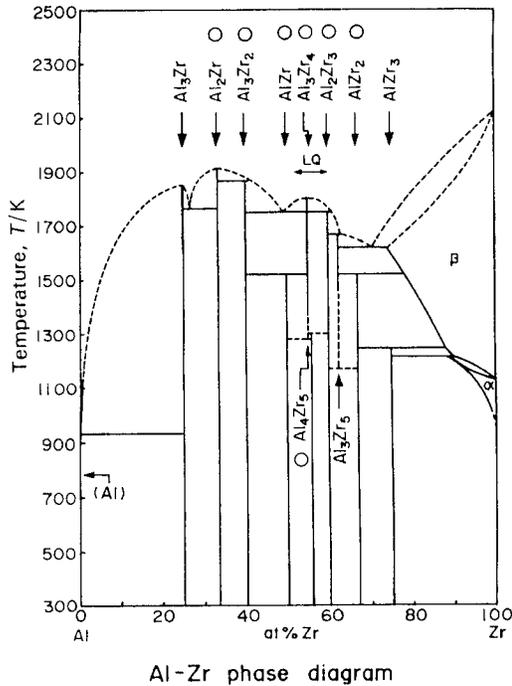
One of the most important applications is the formation of non-equilibrium phases to cause new properties. Such nonequilibrium phases are synthesized by changing (a) the atomic structures and (b) the chemical composition. Extreme examples of (a) and (b) are the formation of amorphous solids and implantation of foreign atoms into various solids. Using the high energy electron irradiation induced (EII) method, various conditions for the formation of nonequilibrium solid phases can be easily and precisely controlled. Furthermore, compared with ion implantation and ion and neutron mixing methods,

this method has the following advantages: (a) small irradiation damage, (b) small temperature rise (a few degrees) in the specimens, (c) deep implantation (more than 1 mm) of foreign atoms, and (d) very fine-scale implantation (i.e., about 1 nm in diameter). Furthermore, the foreign atom implantation can be carried out at any location within the material when the electron energy is sufficient for the implantation.

The formation mechanisms of these nonequilibrium phases can be dynamically investigated on an atomic scale by in situ experiments using sufficiently thick specimens. Today, modern electron microscopy, with the aid of HVEMs, forms a powerful tool for the characterization and identification of materials as well as finding an indispensable place in the 'micro-laboratory', in which various sorts of specimen treatments, including the formation of nonequilibrium phases, can also be carried out precisely on an atomic scale.

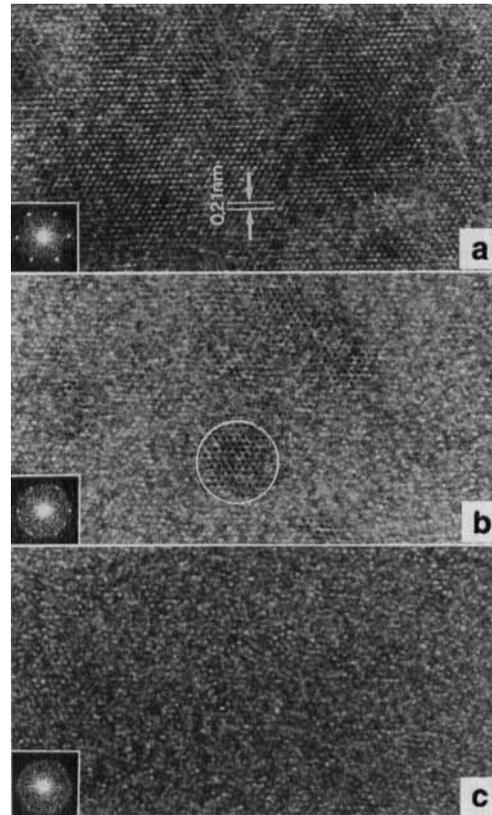
#### 1.4.4.1 Crystalline–Amorphous Transition

Figure 11 shows a crystalline–amorphous transition in the aluminum–zirconium system [32, 38]. Open circles represent the phases which have shown solid amorphization by the EII-method, and LQ indicates the composition range in which amorphization occurs by the liquid quenching method. From this, it can be easily recognized that the EII-method is superior to the liquid-quenching method in controlling the conditions of amorphization. The EII-method has been applied to examine the possibility of solid



**Figure 11.** Crystalline–amorphous transition in the Al–Zr system [38]. Open circles represent the phases that have shown solid amorphization by EII. LQ indicates the composition range in which amorphization by the liquid quenching method takes place.

amorphization for more than 60 intermetallic compounds and about 30 ceramics. Figure 12 shows an example of the amorphization induced by electron irradiation at 2 MV in an Ni-50 at % Ti shape-memory alloy [32], which is very difficult to amorphize by quenching from the liquid phase. The bright spots in (a) and dark ones in (b) represent atom chains whose direction is parallel to the beam direction. It should be noted in Fig. 12b that periodic arrays of spots can be observed in very limited regions. The existence of a critical size for such a region can be concluded, because the periodic arrays of atom chains are only observed in blocks larger than a critical size. The critical size depends on



**Figure 12.** High resolution electron micrographs corresponding to start of solid amorphization (a). (b) Partly amorphitized. (c) completely amorphitized [32].

the bonding mode and bond strength of the constitutive atoms (i.e., it decreases with increasing degree of covalent bonding and with decreasing coordination number of the atom) [32]. The critical size for the Ni-50 at % Ti alloy is about 3 nm in diameter, as shown by the white circle in Fig. 12b. Based on these experimental results, various factors determining amorphization have been clarified, and general rules for amorphization have been proposed [32, 39]. The general rules can be applied to all methods of amorphization such as chemical amorphization, liquid

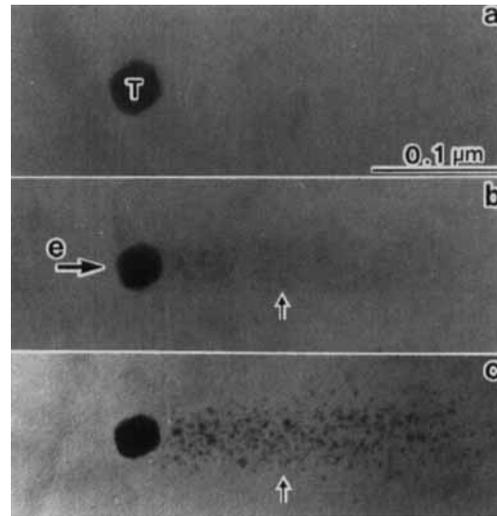
quenching, vapor deposition, mechanical alloying, sputtering, particle irradiation and so on.

#### 1.4.4.2 Electron-Irradiation-Induced Foreign Atom Implantation

The electron irradiation induced (EII) method has been used to implant foreign atoms into solids. A combination of solute and solvent atoms can be selected depending on the purpose, such as for the formation of supersaturated solid solutions, atom clusters and amorphous solids. For example, Si, Ge, Sn, and Pb atoms were implanted into an Al-matrix by the EII-method using 2 MV electrons at 160 K [32]. In these cases, amorphous phases are formed by mixing both Si and Ge atoms into the Al matrix. Small clusters of implanted atoms are also formed by slight annealing after the electron irradiation. Furthermore, it is concluded that the implantation speed markedly increases when amorphization occurs by mixing foreign atoms, as in the case of Al-Si and Al-Ge alloys.

Figures 13a and 13b were taken before and after implantation of Pb atoms into an Al crystal at 175 K, and Fig. 13c after post-annealing for 36 ks at 573 K [40]. It can be noted in Fig. 13b that the diffraction contrast does not appear so clearly in spite of the large difference of more than 20% in atom size between Pb and Al atoms. These implanted Pb atoms precipitate as ultra-fine particles by post-annealing, as shown in Fig. 13c.

When the foreign atom implantation by EII is carried out at slightly higher tem-



**Figure 13.** HVEM images taken (a) before and (b) after implantation of Pb atoms into an Al Crystal at 2 MV and 175 K. (c) Image taken after post-annealing at 573 K for 36 ks [40].

peratures (e.g., at 293 K in Al-Pb alloys), the implantation of solute atoms hardly occurs, even after a large dosage of electrons [40]. It is therefore concluded that the implantation results from repeated short displacement of foreign atoms, and that the mobility of solute atoms is closely related to the behavior of point defects induced in the substrate [32]. Implanted atoms interact with the point defects, in other words, oversized atoms with vacancies and undersized atoms with interstitials. Hence, they are moved out to the nearby point defect sinks, such as the surface or interfaces, when the substrate temperature increases sufficiently to move the related point defects, that is, vacancies for oversized atoms and interstitials for undersized atoms. For example, in the case of oversized atoms, they are hardly made to implant when the substrate temperature is high enough to move the vacancies. From Fig. 13c it is seen that a number of voids

appear together with ultrafine Pb particles in the implanted regions after postannealing [32]. This means that aluminum vacancies induced by the electron irradiation tightly interact with implanted Pb atoms at low temperatures at which the vacancies do not move. This is the reason why the lattice distortion of the Al matrix, resulting from a large atomic size difference, is markedly relieved by the buffer effect of the vacancies, as recognized from slight increase of the diffraction contrast in implanted regions in Fig. 13b.

Furthermore, it is emphasized in Fig. 13 that the implanted regions are formed exactly along the incident direction of electrons so that the implantation occurs inside a cylinder parallel to the incident beam direction and having a cross-section corresponding to the size of preexisting Pb particle. This is due to electron channeling occurring in the Al matrix during the implantation [32].

To reiterate: HVEMs has great advantages when compared with other methods using ions and neutrons; they are an indispensable tool for making various new nonequilibrium phases by amorphization and/or foreign atom implantation.

### 1.4.5 Conclusions

The present review has emphasized that electron channeling by simultaneous reflection becomes very important at high voltages, and that various specimen treatment devices have also been improved as well as stabilization of the high accelerating voltage of the electrons. As a result, not only in situ experiments of various phenomena occurring in materials but

also the formation of new functional materials as nonequilibrium phases have been carried out precisely by high voltage electron microscopy. Stable high energy electrons, obtained with HVEMs, can also be applied to various research fields such as ultrafine fabrication and the generation of synchrotron orbital radiation as well as to the construction of compact HVEMs for high-resolution structure imaging, precise analytical microscopes, etc.

One example in which HVEMs have shown a great advantage has been in the recent investigations on the anomalous behavior of atom clusters in the nanometer range as a new type of nonequilibrium phase [39].

Finally, because of the importance of information on valence electrons in materials science, there are many advantages to be accrued by combining HVEMs with other techniques such as Auger electron spectroscopy or STS. This will undoubtedly provide a fruitful path for the future.

### 1.4.6 References

- [1] W. Bollman, *Phys. Rev.* **1956**, *103*, 1588.
- [2] P. B. Hirsch, R. W. Horn, M. J. Whelan, *Phil. Mag.* **1956**, *1*, 677.
- [3] H. Fujita, Y. Kawasaki, E. Furubayashi, S. Kajiwara, S. Taoka, *Jpn. J. Appl. Phys.* **1967**, *6*, 214.
- [4] H. Fujita, *J. Phys. Soc. Jpn.* **1966**, *21*, 1605.
- [5] H. Fujita, H. Yamada, Proceedings of the International Conference on the *Strength of Metals and Alloys*, *Jpn Inst. of Metals, Tokyo, Japan* **1968**, p. 943.
- [6] H. Fujita, *J. Electr. Microsc. Techn.* **1982**, *3*, 243.
- [7] G. Dupouy, F. Perrier, *J. Microscopie* **1962**, *1*, 167.
- [8] H. Hashimoto, *J. Appl. Phys.* **1964**, *35*, 277.
- [9] K. Kobayashi, M. Ohara, Proceedings of the 6th International Conference on *Electron Microscopy*, Kyoto, Japan **1966**, *1*, 579.

- [10] F. Fujimoto, S. Takagi, K. Komaki, H. Koike, Y. Uchida, *Radiation effects* **1972**, *12*, 153.
- [11] A. Howie, *Phil. Mag.* **1966**, *14*, 223.
- [12] C. J. Humphreys, L. E. Thomas, L. E. Lally, R. M. Fisher, *Phil. Mag.* **1971**, *23*, 87.
- [13] F. Nagata, K. Fukuhara, *Jpn. J. Appl. Phys.* **1967**, *6*, 1233.
- [14] D. Watanabe, R. Uyeda, M. Kogiso, *Acta Cryst.* **1968**, *A24*, 249.
- [15] M. Kiritani, *Ultramicroscopy*, North-Holland **1991**, *39*, 135.
- [16] K. Urban, A. Seeger, *Phil. Mag.* **1974**, *30*, 1395.
- [17] P. B. Hirsch, *J. Phys. Soc. Jpn.* **1962**, *B-II*, 143.
- [18] H. Fujita, N. Sumida, T. Tabata, *J. Phys. Soc. Jpn.* **1973**, *35*, 224.
- [19] H. Fujita, *J. Electr. Microsc. Techn.* **1989**, *12*, 201.
- [20] H. Fujita, *JSIJ International* **1990**, *30*, 70.
- [21] E. P. Butler, *Rep. Prog. Phys.* **1979**, *42*, 833.
- [22] *High Voltage Electron Microscopy*, Academic Press, London and New York (Ed: P. R. Swann, C. J. Humphreys, M. J. Goringe) **1973**.
- [23] Proceedings of the 4th International Conference on *High Voltage Electron Microscopy*, Toulouse, France **1975**.
- [24] Proceedings of the 5th International Conference on *High Voltage Electron Microscopy*, Kyoto, Japan **1977**.
- [25] Proceedings of the 6th International Conference on *High Voltage Electron Microscopy*, Antwerp, Belgium **1980**.
- [26] Proceedings of the 7th International Conference on *High Voltage Electron Microscopy*, Berkeley, USA **1983**.
- [27] Annual Proceedings of the Electron Microscopy Society of America, San Francisco, CA **1983**.
- [28] R. Sandstrom, B. Lehtinen, R. Lagneborg, *Scand. J. Metallurgy*, **1975**, *4*, 1741.
- [29] E. P. Batler, K. F. Hale (Eds), *Dynamic Experiments in the Electron Microscope, Practical Methods in Electron Microscopy*, North-Holland, Amsterdam **1981**.
- [30] Proceedings of the 1st Osaka International Symposium on HVEM, *In Situ Experiments with HVEM*, Research Center for UHVEM, Osaka University, Osaka **1985**.
- [31] Proceedings of the 2nd Osaka International Symposium on HVEM, *New Directions and Future Aspects of HVEM; Ultramicroscopy*, North-Holland, Elsevier Science Publishers B.V. **1992**.
- [32] H. Fujita, *Materials Trans. JIM* **1990**, *31*, 523.
- [33] H. Fujita, C. Lu, *Materials Trans., JIM* **1992**, *33*, 892; 897.
- [34] H. Saka, K. Sasaki, T. Ohashi, I. Ohtsuka, T. Kamino, M. Tomita, Reference [31] **1992**, 110.
- [35] K. Shimizu, K. Otsuka, A. Nagasawa (Eds), Special Issue on Matensitic Transformations, *Materials., JIM* **1992**, *33*, 165.
- [36] T. Imura, H. Saka, N. Yukawa, *J. Phys. Soc. Jpn.* **1969**, *26*, 327.
- [37] H. Fujita, N. Sumida, *Hitachi Inst. News* **1986**, *19*, 3.
- [38] H. Mori, H. Fujita, Proceedings of the International Symposium on *Non-Equilibrium Solid Phases of Metals and Alloys. Suppl. to Trans. JIM* **1988**, *29*, 93.
- [39] H. Fujita, *Materials Trans., JIM* **1994**, *35*, 563.
- [40] H. Fujita, H. Mori, Proceedings of the International Symposium on *Non-Equilibrium Solid Phases of Metals and Alloys. Suppl. to Trans. JIM* **1988**, *29*, 37.

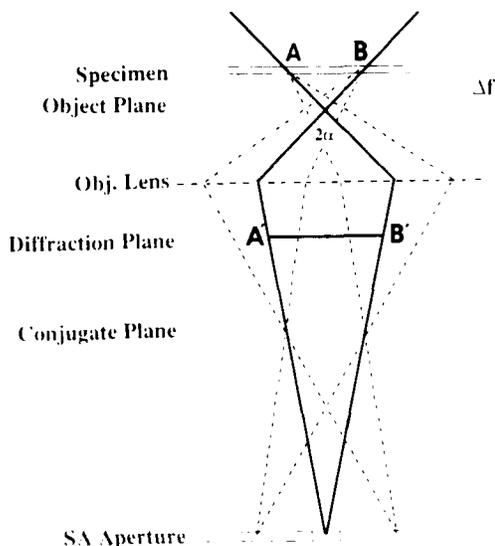
# 1.5 Convergent Beam Electron Diffraction

## 1.5.1 Introduction

Convergent beam electron diffraction (CBED) is perhaps best known as a method for determining crystal point and space groups, and, when combined with elemental analysis using X-rays or electron energy loss spectroscopy, it can be used to determine the composition and structure of small crystals. However, CBED has proved a versatile technique with a much wider range of applications. In this chapter we first explain what information is contained in CBED patterns. We then survey some recent developments including the quantitative determination of crystal bonding and atomic coordinates, the phasing of reflections using coherent diffraction and the analysis of crystal defects and interfaces. For more in-depth accounts of the techniques and applications, the reader is referred to three standard texts [1–3].

CBED is now widely accessible, owing to the commercial development of analytical electron microscopes which have small probe optics and vacuum engineering to minimize probe contamination. For a summary of the instrumental requirements for CBED and an explanation of experimental techniques, the reader is

referred to the article by Vincent [3]. In the following we refer to two techniques. In conventional CBED, the incident beam is brought to a focus on the specimen, the probe size being chosen such as to eliminate any variation in sample thickness, orientation or structure across the probe diameter. In large angle convergent beam electron diffraction (LACBED), see Fig. 1, the probe focus is appreciably above or below the specimen such that, in the object plane, the straight-through and diffracted beams are brought to foci which are separated laterally. By inserting a small selected area aperture (usually in the range 1–5  $\mu\text{m}$ ) a single diffracted beam can be selected such that the resulting diffraction pattern has only one disc. Although this has the apparent disadvantage of almost certainly introducing specimen variations across the incident probe, LACBED has some important advantages over CBED. First, whereas the total convergence angle in CBED is usually limited to avoid disc overlap, to a maximum of about  $2 \times 10^{-2}$  rad, the convergence angle in LACBED may be five to ten times greater, limited only by electron optics. Second, the selected area aperture acts, like an objective aperture in imaging, to filter out inelastic background. Jordan et al. [4] have shown that, by using a

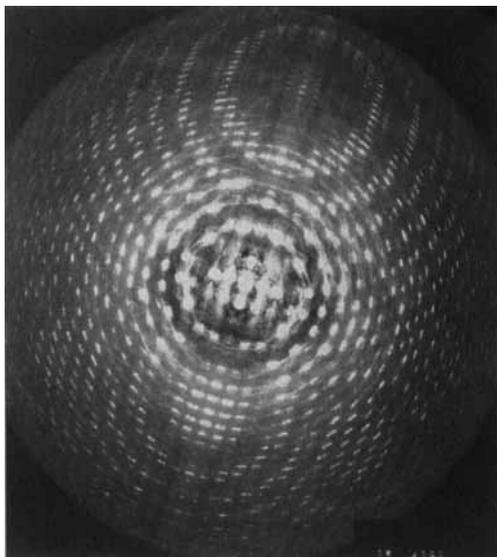


**Figure 1.** Schematic diagram of LACBED optics with the specimen  $\Delta f$  above the focus of the incident probe.

large defocus of the incident probe ( $\Delta f$  in Fig. 1) and a small selected area aperture ( $2\ \mu\text{m}$ ), the beam acceptance angle may be reduced to less than  $10^{-3}$  rad. Almost all the thermal diffuse background and some of the plasmon scattering can thereby be removed from the resulting LACBED pattern. Finally, in LACBED, since any illuminated region of the specimen sees only a subset of incident beam directions, a shadow image of the illuminated region appears on the final LACBED disk. The spatial resolution of this image is given by the minimum probe size in the limit of geometrical optics. The resulting mixture of real space and diffraction information has proved particularly important in understanding diffraction from defects such as dislocations where rapidly varying strain fields are involved (see Sec. 1.5.7 of this Chapter).

## 1.5.2 Geometry and Crystal Symmetry Determination

The Ewald sphere construction in the reciprocal lattice is by far the most convenient way of understanding the geometry of electron diffraction [5]. As a result of the very short wavelength of high energy electrons, the radius of the Ewald sphere is very much greater than the spacing between the reciprocal lattice points, typically 50 or 100 times greater. When the electron beam is incident along a zone axis direction of a crystal structure, a very flat Ewald sphere passes very close to a large patch of reciprocal lattice points surrounding the origin of a reciprocal lattice plane (zero layer plane) perpendicular to direction of incidence. With increasing scattering angle, the Ewald sphere eventually curves away from this zero layer plane and intersects successive layers of the reciprocal lattice, parallel to the zero layer plane, in a series of concentric circles of increasing radius  $G$ . Thus, the diffraction pattern, viewed at a sufficiently low magnification (camera length) consists of a central patch of reflections with additional large angle diffraction arranged in concentric circles of increasing radius (Fig. 2). The successive layers of the reciprocal lattice, where these reflections appear, are known as Laue zones; the zero layer is known as the zero order Laue zone (ZOLZ) and the others are called higher order Laue zones (HOLZ). The number of HOLZ rings visible depends on the angular view available, the microscope operating voltage and the spacing  $H$  of the successive layers of the reciprocal lattice along the zone axis direction. For electrons with wavenumber  $k$  we have, to a



**Figure 2.** Convergent beam electron diffraction pattern of NbSe<sub>3</sub> taken at 120 kV with the specimen cooled by liquid helium. The concentric circles of increasing radius at the center of the pattern are based on the chosen zone axis of incidence. Two other secondary zone axes are clearly visible along the mirror line through the center of the pattern and other zone axes off the mirror line may also be discerned.

good approximation

$$G \approx \sqrt{2kH} \quad (1)$$

In practice, large angle scattering is strongly affected by the Debye–Waller factor and so the visibility of the successive HOLZ rings may be limited by thermal or static disorder in the sample under investigation.

The existence of HOLZ diffraction provides firm evidence of the stacking sequence along the direction of electron propagation. The lack of HOLZ diffraction is an indication of disorder along this direction [6]. This might be planar-disorder as in the case of layer-structured materials with weak interlayer bonding, line-disorder as from a high dislocation

density or point-disorder. The relative weakness of HOLZ diffraction is the justification for ignoring it, as in the case of the so-called projection approximation. Neglect of HOLZ diffraction is equivalent to projection of the atomic structure along the zone axis direction. That this is only a first approximation and sometimes a gross oversimplification is evident from Fig. 2. One further point of considerable importance is that the Bravais lattice may be obtained by examination of CBED patterns showing HOLZ diffraction without overlap of diffraction orders [7]. To simplify the analysis it is preferable to perform the analysis at a high symmetry zone axis if possible. A mesh of lines is created in the ZOLZ with nodes at the reflections and a further mesh is constructed from the first order Laue zone (FOLZ) reflections. By the coincidence or not of the nodes of the two meshes the Bravais lattice may be determined.

The nature of HOLZ diffraction is generally rather different from that of ZOLZ diffraction. To a reasonable first approximation it can be treated kinematically, a point that will be discussed in more detail in Sec. 1.5.5 of this Chapter. In this approximation, the width of the diffraction peak, corresponding to a particular reflection  $g$ , will depend on the magnitude of its associated structure factor  $F_g$ . For large values of  $g$  as in HOLZ diffraction,  $F_g$  is small and so narrow diffraction maxima will occur. The locus of all points within a given cone of electron convergence that satisfy the Bragg condition for that particular reflection is essentially a straight line (in fact it is given by the intersection of two spheres of very large radius,  $k$ , one on the origin and one on the reciprocal lattice point  $g$  from the origin).

Thus HOLZ diffraction takes electrons away from the cone of incidence along a narrow line perpendicular to  $\mathbf{g}$  and their Bragg reflection produces a bright line of the same length, width and orientation in the appropriate HOLZ ring. The dark line in the central disk of the CBED pattern is known as a HOLZ deficiency line and the associated bright line in the HOLZ ring is called a HOLZ excess line [8]. Because of their narrow angular width, the HOLZ lines have a number of useful applications. In the case of very fine lines, resulting from a distant FOLZ ring, computations of HOLZ line patterns based on kinematic diffraction theory are likely to be valid with two important reservations [9]. First, for reasons to be discussed in this Chapter, Sec. 1.5.5, the microscope operating voltage must be allowed to vary within about  $\pm 3\%$  of the apparent value (we shall use the term pseudokinematic to describe HOLZ diffraction on account of this factor). Second, where two lines  $\mathbf{g}$  and  $\mathbf{g}'$  intersect at a small angle and  $F_{\mathbf{g}-\mathbf{g}'}$  is large, then strong dynamical coupling of the HOLZ lines corresponding to  $\mathbf{g}$  and  $\mathbf{g}'$  will occur near their intersection.

The first use of HOLZ lines is in lattice parameter measurements [10]. One can avoid zone axis directions to reduce the dynamical effects that introduce the pseudokinematic approximation rather than the true kinematic approximation, or else concentrate on changes of lattice parameter caused by temperature or composition changes so that dynamical corrections are unlikely to be important. The second use of HOLZ lines is in strain measurement [11]. Some words of caution are necessary in this context. One of the chief attractions of the technique is the ease with which HOLZ-line distortion

may be detected. However, accurate measurements of strains from such observations are rather hard to achieve for a variety of reasons. First, the measurements are being made on thin films where considerable stress relaxation will have occurred and has to be taken into account. Second, the process of making the specimen may have introduced additional stresses that are not immediately obvious (such as are caused by amorphous surface films or surface oxides.) Third, differential expansion effects caused by albeit small beam heating of an interface region between dissimilar materials may lead to spurious results. Finally, to obtain fine HOLZ lines it is generally necessary to have specimen thicknesses of  $< 100$  nm and the strain to be measured should remain constant along the column explored by the electron beam, to within the accuracy of the measurements. It is common to achieve measurements of strain or lattice parameters to about  $10^{-3}$  by the HOLZ line method. Greater accuracy down to about  $2 \times 10^{-4}$  can be achieved in favorable cases but energy loss to plasmon excitation then becomes a fundamental limitation.

CBED patterns evidently contain a wealth of symmetry information and this aspect of the subject is now well-understood and widely exploited [12]. The dynamical nature of electron diffraction far from being a complication, is the very strength on which the technique is based. As a result, the absence of inversion symmetry is detectable *straightforwardly*, the symmetry of the patterns may be related directly to the 32 point groups rather than the 11 Laue groups (created when Friedel's law is obeyed, as in normal X-ray diffraction). The theoretical analysis of symmetry

is based on very general and widely applicable approximations of high energy electron diffraction. Tables have been produced that allow one to go directly from experimental observations to point group and space group determinations. In favorable cases this is an extremely efficient process and the information gathered at one zone axis can be sufficient to perform an unambiguous space group determination. More commonly, information has to be gathered at two or more zone axes to complete the analysis. Fortunately, the inclination of the specimen in tilting to a zone axis has, generally, very little influence on pattern symmetry.

The nature of electron diffraction symmetry is more subtle than may be apparent at first. The symmetry of the direct beam may be higher than that of the whole diffraction pattern; individual orders of reflection can have internal symmetry that is important in determining which of the 31 different diffraction groups a particular pattern belongs to. The steps required in identifying the diffraction group and from that the crystal point group are discussed in several previous works [13].

Having determined the crystal point group, the presence of screw axes or glide planes may be inferred by the study of X-ray forbidden reflections in well-oriented zone axis patterns. These appear in electron diffraction as a result of multiple scattering but nevertheless lines of absent intensity occur within these reflections because of exact cancellation between the various contributing diffraction paths. The individual diffraction paths may be paired together such that each path contributes an equal amplitude to the reflection precisely out of phase with the contribution

from the other path. The resulting lines of absence are variously called dynamic absences, Gjønnes-Moodie lines, or dark bars. Accounts of their use in space group determination are given in several references [14].

### 1.5.3 Bloch Wave Theory

To understand the details of CBED patterns it is generally necessary to make reference to a form of dynamical diffraction theory. The simplest of these is the so called two-beam theory where only the incident beam and one diffracted beam are considered significant [15]. According to this theory, ignoring absorption, the intensity of the direct beam,  $I_0$ , is given by

$$\begin{aligned} I_0 &= 1 - 4C_0^{(1)2} C_0^{(2)2} \sin^2 \frac{\pi t}{\xi_g} \sqrt{1 + w_g^2} \\ &= 1 - I_g \end{aligned} \quad (2)$$

where  $I_g$  is the diffracted beam intensity,  $w_g = s_g \xi_g$  is called the deviation parameter from the  $g$  Bragg condition,  $\xi_g \propto 1/F_g$  is the so-called extinction length,  $s_g \sim g\Delta\theta$  where  $\Delta\theta$  is the angular deviation from the  $g$  Bragg condition and  $t$  is the thickness of the specimen.

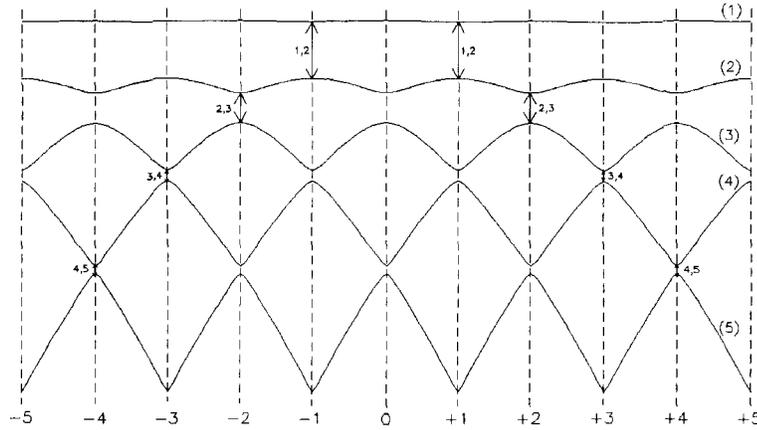
We can also write

$$I_g = 4C_0^{(1)2} C_0^{(2)2} \sin^2 \frac{\pi t}{\xi_g^{eff}} = \frac{\sin^2(\pi t / \xi_g^{eff})}{1 + w_g^2} \quad (3)$$

where  $\xi_g^{eff} = \xi_g / \sqrt{1 + w_g^2}$  and  $\xi_g^{eff}$  is called the effective extinction distance. Its value decreases symmetrically for  $\pm\Delta\theta$ .

The basis of this theory is the Bloch wave formulation of dynamical diffraction.

**Figure 3.** Dispersion surface for a  $(nnn)$  systematic row of silicon at 300 kV. The dispersion spheres appear as parabolas in this diagram and are asymptotic to the dispersion surface away from Brillouin zone boundaries. These parabolas are centered on each of the reflections  $n$ . At the  $n$  Bragg position, the effective extinction distance is the reciprocal of the gap labelled as  $n, n + 1$ . This diagram should be related directly to the LACBED result in Fig. 4.

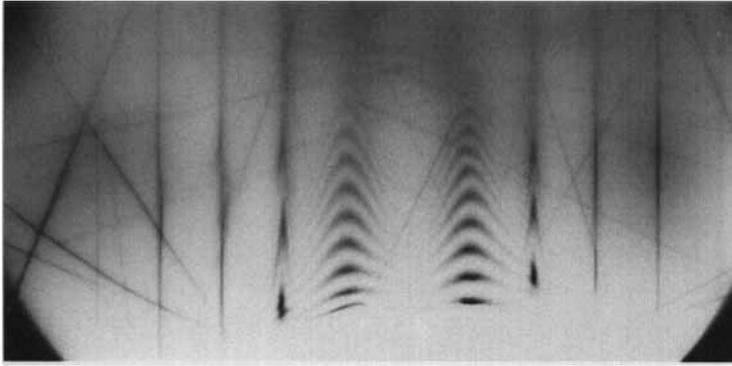


It is particularly convenient for understanding many of the diffraction effects that occur. The results of this form of diffraction theory are conveniently represented in terms of a so-called dispersion surface constructed from dispersion spheres (of radius proportional to the free electron energy) centered on each of the reciprocal lattice points. The spheres intersect at Brillouin zone boundaries, where electron diffraction occurs, causing standing (Bloch) wave formation and splitting of the constant energy surfaces near the circles of intersection of the spheres (Fig. 3). These effects are the direct equivalent of Fermi surface and forbidden band formation for conduction electrons.  $C_0^{(1)}$  and  $C_0^{(2)}$  in the expressions above are the Bloch wave amplitudes of the two branches approaching an intersection point of the constant energy spheres.

A particular attraction of the Bloch wave formulation is that only a small number of excitation amplitudes are significant at any given orientation and the same or an even smaller number show significant orientation variation within the angular range of a CBED disk. In the

simplest case, which often applies, only two branches need to be considered.

Let us consider the application of two beam theory to the LACBED pattern formed for silicon  $(111)$  systematic diffraction from a wedge shaped crystal (Fig. 4). The thickness of the specimen increases vertically up the page and the position midway between the two sets of curved fringes is the symmetric position with the direction of incidence perpendicular to the  $(111)$  systematic row. The series of dark vertical lines indicate the Bragg positions of higher order reflections  $(nnn)$ . They are marked by integers on the corresponding dispersion surface shown in Fig. 3. At the  $\pm 1$  Bragg positions the fringes have a maximum spacing that corresponds to the reciprocal of the gap opened up at the first Brillouin zone boundary between branches (1) and (2) of the dispersion surface. At the  $\pm 2$  Bragg positions, a series of widely-spaced arrow-like fringes are visible. These correspond to the results of another two branch theory, this time for the smaller gap opened up at the second Brillouin zone boundary between branches (2) and (3) of



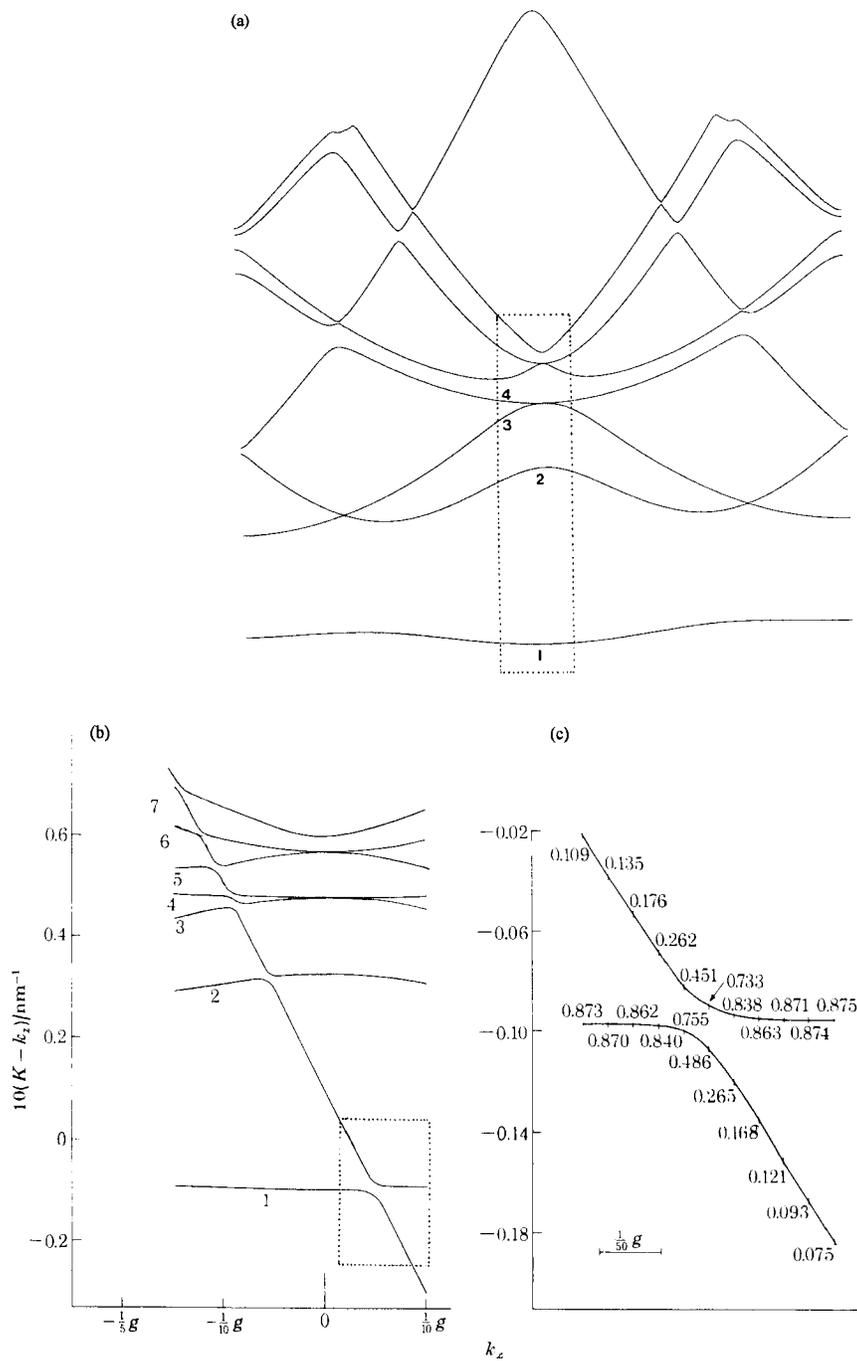
**Figure 4.** Experimental LACBED wedge/rocking curve for the  $(nnn)$  systematic row of silicon at 300 kV. Lines inclined to the vertical are nonsystematic reflections, not taken into account in the calculated results shown in Fig. 3. They clearly show the expected lack of vertical mirror symmetry. Thickness fringes may be observed clearly at the  $\pm 1$ , and  $\pm 2$  Bragg conditions: they can also be detected (just) at the  $\pm 3$  Bragg condition.

the dispersion surface. Because the gap is smaller, the extinction length is longer and the angular range over which the local two-branch theory applies is narrower than the first case. In the case of the  $\pm 3$ ,  $\pm 4$  and  $\pm 5$  Bragg conditions the gaps opened up between the relevant dispersion surface branches are so small that very long extinction lengths occur. Hence the first maximum is not reached before absorption takes over, resulting in diffuse scattering at the top edge of the pattern, and only dark lines corresponding to the first extinction minima are observed. The lines become narrower and fainter as  $n$  increases.

In addition to demonstrating the value of a series of local two-branch approximations over local orientation ranges, there are several further interesting points about this micrograph. According to kinematical diffraction theory, for spherical atoms with harmonic lattice vibrations, there would be no extinction at the  $\pm 2$  Bragg positions. The extinction observed is essentially the result of dynamical diffraction along the systematic row, with a small additional contribution from bonding effects and anharmonicity. According to

many beam dynamical theory, the extinction length at the  $\pm 1$  Bragg positions is 103.7 nm and at  $\pm 2$  it is 190.0 nm. In fact the ratio between the two extinctions obtained from measurements on the micrograph, is closer to 2.5. This difference is believed to be the result of nonsystematic interactions that perturb the purely systematic diffraction assumed in the calculations. The presence of non systematic reflections may be observed clearly in the micrograph and they break the mirror symmetry along the center line of the pattern as expected. The  $-6$  reflection is not observed and it is expected to be very weak because, like the  $(222)$  reflection, it can only appear through systematic dynamical interactions if bonding effects and anharmonicity are ignored.

This same theory is readily adapted to making accurate specimen thickness measurements from fringe patterns in CBED from systematic rows of reflections [16]. The one dimensional example given above is easily adapted to describe two dimensional (zone axis or cross-grating) diffraction [17]. Those Bloch states with high probability density of electrons on the atom strings have low potential energy



**Figure 5.**  $\{1\bar{1}\bar{1}\}$  section through a Si(111) zone axis dispersion surface at 100 kV. The diagram is inverted relative to Fig. 3 in order that the lowest potential energy state (1s) is at the bottom. (a) Large region of the dispersion surface centered on the origin for the case of two dimensional (zero layer) diffraction, (b) boxed region of (a) is shown in greater detail with a single HOLZ  $(\bar{5}\bar{5}11)$  reflection added to the calculation ( $g = 1/a[\bar{1}\bar{1}2]$ ). The intersections of the HOLZ sphere with the zero layer branches causes local splitting and hybridization as detailed in (c) for the boxed region in (b).

and high kinetic energy. They represent electron states bound in the atom-string potential wells which, in the simplest case, have cylindrical symmetry. These bound-states are analogous to atomic bound-states and have principal and angular momentum quantum numbers  $n/l$  where  $l \leq n - 1$  (e.g., 1s, 2p). The states are ordered in the atom-string potential first by their principal quantum number, then by their angular momentum quantum number, the higher the  $l$  value the lower the energy of the state. It follows that the two lowest-energy bound-states are 1s and 2p. However, except in the case of rather deep atom string potential wells, there is only one bound state (1s) and the nearly free states, just above the top of the well, have a different order, with 2s below 2p in energy (Fig. 5). This latter situation is very common and gives rise to the concentric ring pattern often observed at zone axes of simple materials. The rings are produced by two-branch interference between the 1s and 2s states exactly analogous to the state of affairs described by Eq. 2. The wells may be made deeper by a number of factors and it is possible [18] to define a string-strength parameter  $S$  that includes these factors to a first approximation

$$S = \frac{\gamma Z A}{d} \quad (4)$$

where  $\gamma$  is the relativistic mass factor,  $Z$  is the mean atomic number in the repeat distance,  $d$ , along the atom string and  $A$  is the area of the Wigner-Seitz cell associated with the atomic string. The  $\gamma$  factor arises because we can regard the relativistic effect, on increasing the microscope operating voltage, as multiplying the potential that the electron experiences by the

increase of its relativistic mass.

$$\left( \frac{-\hbar^2}{2m_0^2} \nabla^2 + \frac{m}{m_0} V(\mathbf{r}) \right) \psi(\mathbf{r}) = E\psi(\mathbf{r}) \quad (5)$$

It follows from this argument that an atom string potential well that is not quite deep enough to hold two bound states at 100 kV may be able to do so at some higher operating voltage, in consequence of the increase of relativistic mass [19]. For this to happen interchange of the order of the 2s and 2p states has to occur and it does so through a so-called accidental degeneracy (i.e., not symmetry determined) just at the point where the 2s level touches the top of the string potential well. This accidental degeneracy is known as the critical voltage effect and the voltage at which the degeneracy occurs may be determined experimentally with great accuracy [20]. At higher voltages the two  $n = 2$  states are bound and reversed in order, 2p below 2s. There is an exactly equivalent critical voltage effect in one dimensional electron diffraction but these critical voltages tend to occur at rather high microscope operating voltages [21].

Three-dimensional diffraction involves a relaxation of the projection (atomic string) approximation and requires the addition of free electron dispersion spheres centered on HOLZ reflections. It is frequently valid to consider the effect of a particular HOLZ reflection  $g$  in isolation from other HOLZ reflections by adding just one sphere centered on the particular HOLZ reciprocal lattice point to the dispersion surface construction [22]. Because of the large distance of  $g$  from the zone axis the  $g$  dispersion sphere intersects the zero layer dispersion surface at an appreciable angle and wherever it cuts a particular zero layer branch splitting occurs [9] and

a new two-branch situation is created (Fig. 5). It is precisely because the HOLZ line that results from the intersection between the HOLZ dispersion sphere and the branch of the zero layer dispersion surface is displaced from the kinematical position that the term pseudokinematical was adopted earlier and a degree of voltage variation had to be introduced in computer matching of HOLZ-line patterns. The splitting introduced between the zero layer branch and the HOLZ dispersion sphere is rather small [9] corresponding to an extinction length ( $\xi_H$ ) of several hundred nanometers. It follows that we can often make the kinematical approximation to Eq. (2), that is,

$$\sin \frac{\pi t}{\xi_H} \approx \frac{\pi t}{\xi_H}$$

In this sense, therefore, the HOLZ diffraction is kinematic and also in the sense that the splitting that occurs may be calculated by perturbation theory [22]. However, the displacement of the zero layer branch from the dispersion sphere on the origin is a dynamical effect that depends on the strength of the atom-string potential. In this sense HOLZ diffraction is not kinematical. Moreover, there may be several appreciably excited branches of the zero-layer dispersion surface, each giving a separate HOLZ line, displaced from, but approximately parallel to, the others. This HOLZ line splitting or fine structure is also a purely dynamical effect, permitting direct visualization of the strongly excited zero layer branches. Moreover, the extinction length  $\xi_H$  may be greatly reduced (by something like a factor of five) from the value deduced from kinematical theory ( $\xi_H^{kin} \propto 1/F_H$ ), where  $F_H$  is the HOLZ reflection structure factor) and hence

there is another sense in which HOLZ diffraction is not kinematical. Indeed, if it were not for the zone axis reduction of the HOLZ extinction lengths, HOLZ lines would probably not be visible in zone axis CBED patterns.

### 1.5.4 Crystal Structure Determination by Convergent Beam Electron Diffraction

For thin films of crystals which contain principally light elements, structure determination by plane wave electron diffraction has been based upon the techniques developed for the analysis of kinematic X-ray data [23]. Recent advances include the introduction of direct methods and maximum entropy criteria to phase high order reflections beyond the point resolution limit in lattice images [24]. However, this approach is not reliable for crystals which contain heavier atoms, where many of the low order reflections are associated with extinction distances comparable to the crystal thickness, even for CBED patterns acquired with the beam focused onto a crystal with a local thickness of a few nanometers. A partial solution, at least for inorganic crystals of moderate complexity, is offered by analysis of HOLZ reflections, which intercept the Bragg condition in zone axis CBED patterns, usually have extinction lengths much larger than the crystal thickness, are not located within strong systematic rows, and are associated with structure factors sensitive to small changes in the atomic parameters.

The basis for a quantitative analysis of HOLZ intensities is given by the Bloch

wave formalism for the solution of the Schrödinger equation in the projection approximation. Along major zone axes, the exit wave function is a phased sum over relatively few excited Bloch states, often described as molecular combinations of localized two-dimensional atomic states. The eigenvalues associated with a near-degenerate molecular cluster of 1s type Bloch states localized on a symmetry-related set of atom strings in the projected potential may be quite distinct from the eigenvalues of other localized or channeling Bloch states. It follows that the components of the electron wave vector along the zone axis also differ, and that the inner excess lines observed within the disks of HOLZ reflections correspond to diffraction from Bloch state clusters localized on subsets of atom strings, with a related shift of excess and deficiency lines away from their kinematic positions.

The Bloch wave formalism is equally useful for constructing a model of the factors which control the relative intensities of HOLZ excess lines. In the kinematic approximation, the amplitude of a HOLZ reflection is proportional to the relevant structure factor of a two-dimensional conditional potential,  $U^n(\mathbf{R})$ , constructed by a Fourier sum over structure factors in layer  $n$  of the reciprocal lattice, where  $\mathbf{R}$  is a position vector within the projected cell. For  $n = 0$ ,  $U^0(\mathbf{R})$  is the usual projected potential, which is real by definition, but the general conditional potentials are complex with phase factors proportional to atom positions along the projection axis, being simply an expression in real space of the information implicit in the structure factors. The final step is to consider the effect on the HOLZ amplitudes of diffraction not from the incident plane wave, but

from a cluster of Bloch states. It may be shown [25] that a modified version of the conditional potential is retained, defined by multiplication of  $U^n(\mathbf{R})$  with the relevant Bloch state at every position within the projected cell. In a quasikinematic limit, the amplitudes of associated HOLZ excess lines remain proportional to the Fourier coefficients of the modified potential which may include contributions from only a single subset of atoms.

Provided that an approximate structure is available for calculation of the zone axis Bloch states, the intensities of suitable HOLZ excess lines may be used to refine the atomic parameters. To improve the statistical accuracy, it is important to combine data acquired over a wide voltage range, equivalent to an annulus of HOLZ reflections. The atomic parameters are refined by standard X-ray procedures, which minimize the mean squared difference between observed and calculated intensities based upon a kinematic expression derived from the partial structure factors. Typical reliability factors for HOLZ refinements (defined as the mean fractional difference between observed and calculated intensities) are larger than equivalent X-ray values although the standard deviation of the atomic parameters is impressively small, often equivalent to an accuracy of 0.001 nm. The explanation for this apparent dichotomy is that only the high order reflections are refined. The intensities of individual reflections are very sensitive to small changes in the atomic parameters but are also subject to perturbation by coupling between adjacent HOLZ reflections that is not included in the quasikinematic approximation described above.

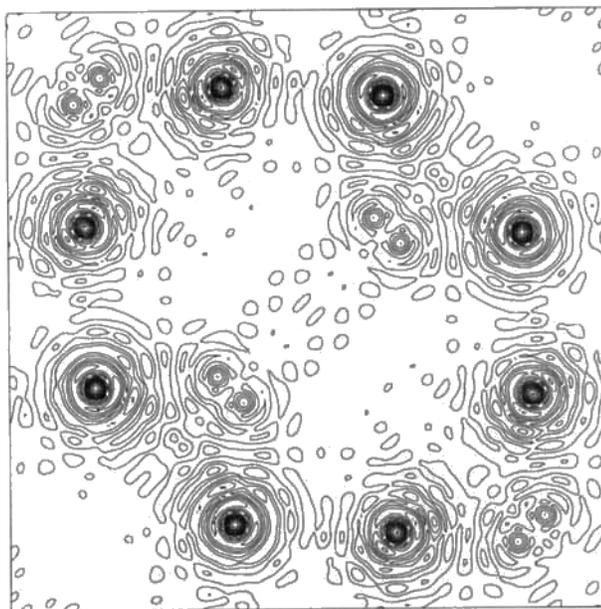
Examples of structure refinement [26–28] include the analysis of displacive

superlattices where the symmetry of the subcell is broken to produce a domain structure not easily accessible to X-ray or neutron techniques. A further application is to metastable crystal structures produced either by interface reactions or by rapid quenching of liquid or vapor phases. For crystals which have no resemblance to known phases, there remains the basic problem of deducing the structure. Some progress has been made by adapting the arguments outlined above for HOLZ diffraction. If the inner excess lines show strong variations of their relative intensities around a major axis, then it is reasonable to assume that the relevant Bloch states are localized, and that the diffracted intensities may be analyzed by X-ray procedures.

Patterson transforms of the HOLZ intensity data were used to deduce interatomic vectors and subsequently refine the atomic parameters in metastable alloys [29,

30]. The correlation peaks were very sharp, being the transform of a restricted set of reciprocal lattice vectors with short wavelengths. This approach fails when the peaks are not visible above the background ripple generated by the truncated data set, combined with the inevitable noise associated with dynamical coupling between HOLZ reflections. A further problem common to all analyses based on localized Bloch states is that weak atom strings, often corresponding to projections of atoms with low atomic number, are not detected. An alternative approach based upon an electron-optical equivalent to the X-ray precession geometry has been explored where the Ewald sphere is rocked through the Bragg condition for an annulus of HOLZ reflections [31]. An example for a rare earth pyrogermanate is shown in Fig. 6, where the Er and Ge positions agreed with the X-ray data, but the oxygen positions were obscured by nonsystematic perturbations.

**Figure 6** . Example of the projected Fourier transform of a precession pattern for  $\text{Er}_2\text{Ge}_2\text{O}_7$  (space group  $P4_{1,3}2_12$ ) calculated from the intensities of an annulus of FOLZ reflections excited around the  $c$  axis. The peaks corresponding to Er and Ge atoms in general (8-fold) positions were refined from the experimental data. The background ripple around and between peaks is associated with truncation of the Fourier series when summed over the FOLZ reflections. (Courtesy of M. E. Sleight.)



### 1.5.5 Quantitative Convergent Beam Electron Diffraction

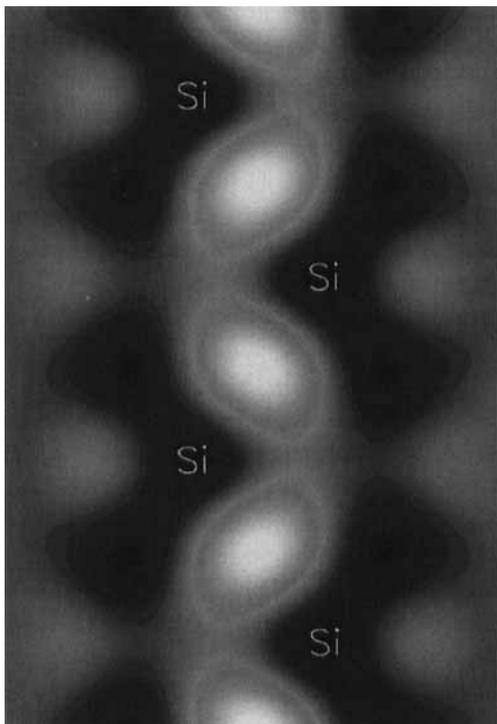
Unlike the analysis of kinematic X-ray and neutron diffraction data, there exists no general method for the inverse solution of the Schrödinger equation to retrieve a complete set of Fourier coefficients,  $V_g$  of the crystal potential. Nevertheless, there are several strong incentives, both theoretical and practical, for the quantitative comparison of calculated and experimental CBED intensities. These include the introduction of field emission sources with high brightness and also commercial electron spectrometers to remove the inelastic component from diffraction patterns. When combined with the absence of any requirement for perfect crystals larger than nanometer dimensions, and the continued decrease in the cost of dedicated computers, it has become possible to match many-beam dynamical calculations with experimental CBED patterns, treating the structure factors, thermal parameters and thickness as free variables which are adjusted to minimize the disparity between experiment and theory. In principle, we may infer the composition, atom positions, vibrational amplitudes and the bond charge distribution from a complete knowledge of the crystal potential or the equivalent charge distribution calculated via Poisson's equation.

However, the time required for each dynamical refinement cycle increases rapidly when more beams are included, and has imposed an effective division between structure determination by traditional methods applied to quasikinematic estimates of intensities for a large number of high order reflections (see Sec. 1.5.4 of

this Chapter), and the exact calculations applied to simple inorganic crystals with known structures. Even for these crystals, the majority of structure factors are reliably estimated by calculation from the form factors for neutral isolated atoms, and only the phase and amplitude of the lowest order components in the crystal potential are varied to seek a global minimum in phase space between observed and calculated intensities, subject to a statistical  $\chi^2$  criterion for acceptable agreement between experiment and theory.

As discussed in detail by Spence and Zuo [2], the purpose of this comparison is to measure with high accuracy the lowest order coefficients of the crystal potential, which are slightly perturbed by the redistribution of charge associated with bonding between atoms. A further advantage of using electrons for measurement of bond charge distributions is that the necessary conversion from voltage to charge reduces the statistical error, which is a vital consideration because bonding represents only a small fraction of the natural overlap between the valence wave functions when atoms are assembled into a crystal. The corresponding maximum change in the lowest order electron structure factors is less than 1%. Problems which are not yet fully addressed include the treatment of thermal diffuse scattering which remains unfiltered from experimental patterns, the accurate measurement of thermal Debye-Waller factors, which appear in the conversion from voltage to charge distributions, the improvement of models for the imaginary component of the crystal potential which simulates the loss of electrons by inelastic processes, and the accurate estimation of the local composition in compounds.

Measurement of the complex structure factor for a single reflection is based upon the two-beam diffraction geometry, where all other higher order reflections with significant excitation are included in calculations. Alternatively, dynamical interactions between low order reflections perturbed by bonding are enhanced within a single zone axis pattern, and may represent a more efficient strategy for refinement of the structure factors [32] (Fig. 7). As yet, published results have



**Figure 7.** Charge distribution in crystalline Si at 300 K, calculated by matching calculations to a (110) CBED pattern, and seen through a (110) planar section intercepting the nuclei and covalent bonds. To enhance visibility, the neutral atom charge distribution has been subtracted, leaving bonds as regions of excess negative (light tones) charge, balanced by dark electron-deficient regions, centered on the atoms and also extended into tetrahedral lobes in opposition to the bonds.

been confined to a few semiconductors, alloys and oxides, but the best quality data agree closely with the extensive X-ray measurements on Si and Ge, and are consistent with band structure calculations. A separate parameter not measured by pattern matching is  $V_0$ , the mean crystal potential which is a sensitive measure of charge redistribution in crystals. For self-consistency, measurements of structure factors and thermal parameters must agree with independent estimates derived from thickness-independent contrast features in CBED patterns, notably local (three-beam) coupling between nonsystematic reflections and critical voltage contrast associated with accidental degeneracy of Bloch state eigenvalues.

### 1.5.6 Coherent Convergent Beam Electron Diffraction

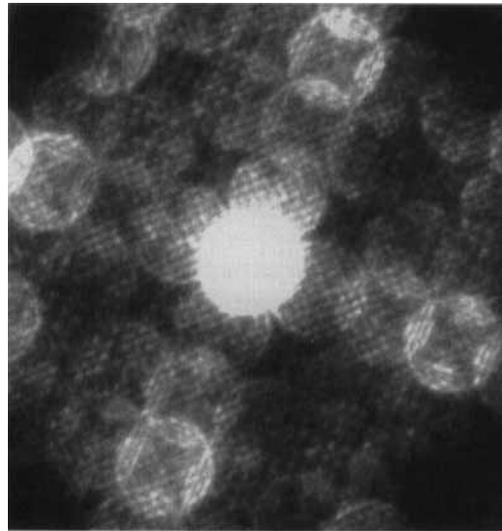
Provided that adjacent disks do not overlap, CBED patterns from a perfect crystal with plane parallel surfaces normal to the beam are indistinguishable, being independent not only of the probe size, position and defocus, but also of the probe coherence. The latter result follows from the absence of interference between adjacent beams within any disk of a CBED pattern. The situation is entirely different if the convergence angle is increased so that the discs overlap. The random phase relations between beams emitted by the conventional electron source used in most transmission microscopes ensures that only the intensities of adjacent discs are added in the overlap regions, whereas the overlap intensities in the CBED pattern formed by

an ideal coherent source are determined by the phased addition of the diffracted amplitudes. Further, the contrast within overlaps is dependent on the probe position within the projected cell. If the probe is scanned across a set of lattice planes with spacing  $d$  and reciprocal lattice vector  $g = 1/d$ , then the relative phase at any position in the overlap of a pair of reflections related by the difference vector  $\mathbf{g}$  cycles through  $2\pi$  in distance  $d$  with a corresponding periodic variation of the interference contrast.

For beam convergence angle  $2\alpha$ , adjacent discs overlap if  $2\alpha < \lambda/d$ , equivalent in real space to the condition that the size of an ideal, diffraction-limited focused probe must be comparable with or less than the planar spacing. This result follows by application of the Airy criterion, where  $\Delta = 0.61\lambda/\alpha$ , or  $\Delta < 1.22d$ . When the probe is defocused by  $\Delta f$  relative to the specimen plane, the beam remains coherent but a regular phase variation of the beam is introduced across the lattice planes, and sinusoidal contrast with angular period  $d/\Delta f$  is observed in the overlaps, equivalent to interference contrast between virtual diffracted images of the probe.

For a field emission source imaged onto the specimen by electron lenses, the probe size is increased by other factors, including the physical (incoherent) source size, instrumental instabilities and spherical aberration. Until recently, observations of coherent CBED contrast have been limited principally to scanning transmission microscopes equipped with field emission sources. As discussed by Cowley [1] and Spence and Zuo [2], there is a direct relation between the conditions for observation of coherent diffraction con-

trast discussed above, and the corresponding resolution of lattice fringes in the signal collected by a detector when the probe is scanned across the specimen. Applications to materials science include the detection of site-symmetry elements in nanometer-sized particles and the analysis of fault vectors on a similar length scale. More generally, there are prospects for 'super-resolution' based upon Fourier analysis of the diffraction contrast from a scanned coherent probe in the limit of a thin phase object [33]. In a related development, coherent zone axis patterns (Fig. 8) obtained from conventional transmission microscopes equipped with field emission sources [34, 35] offer the prospect of enhanced resolution in the diffraction plane by stepwise phasing of reflections



**Figure 8.** Coherent zone axis CBED pattern for  $\text{BaCuO}_2$  (space group  $\text{Im}\bar{3}\text{m}$ ,  $a \approx 1.3$  nm), aligned with the beam parallel to a cube axis. The convergence angle is adjusted so that CBED discs overlap with first and second nearest neighbors, and the probe is slightly defocused to show interference fringes in both sets of overlaps. (Courtesy of Dr. P. Spellward.)

beyond the limit on point resolution set by spherical aberration. The close relation between coherent diffraction and holography is confirmed by the observation of interference fringes in CBED disks overlapped by an electron biprism [36], where the limits on probe size and coherence are relaxed.

### 1.5.7 Studies of Imperfect Crystals

In the last few years it has become clear that CBED, and more particularly LACBED, are powerful methods of studying crystal imperfections. The essential point is that imperfections influence the crystal rocking curve. This can be seen most easily by considering the kinematical expression for the two-beam diffraction function,

$$\phi_g(s) \propto \int_0^t F_g(z) \exp[-2\pi i(sz + \mathbf{g} \cdot \mathbf{R})] dz \quad (6)$$

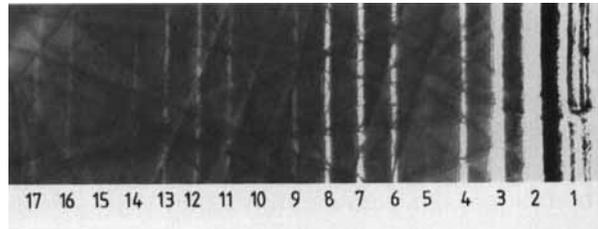
where  $F_g$  is the structure factor, assumed to be depth-dependent,  $\mathbf{R}$  is any displacement field due to, for example, a dislocation or an interface,  $s$  may include any uniform crystal strain and  $t$  is the sample thickness along the beam direction. CBED and LACBED allow us to profile the

rocking curve  $I_g = |\phi_g|^2$  for a selected reflection. Since diffraction angles are small and the Ewald sphere is nearly flat, the rocking curve, at least for low order reflections, is scanned approximately along the incident beam direction. The resulting rocking curve is mapped into the appropriate diffraction disc.

To analyze imperfect crystals it has proved most useful to examine two-beam rocking curves. Extensive studies have been carried out on semiconductor multilayer structures. A review of this topic is found in the literature [37]. Studies of plan view samples have proved most effective since, in this geometry, the essential diffraction information lies along the specimen normal around each fundamental reflection, the direction to which we are most sensitive when samples are flat. Figure 9 shows an example for a periodic multilayer where the rocking curve reveals superlattice satellite reflections,  $n\mathbf{q}$ , due to variations in  $F_g(z)$ . The large number of satellite reflections visible underlines both the efficient filtering of inelastic background achieved in LACBED, and the elimination of spatial averaging which would attenuate higher order satellites [37]. The weakness of the satellites at  $n \approx 5, 10, 15$  can be related to the ratio of adjacent layer thicknesses [38].

The rocking curves from periodic multilayers can be understood in some detail using kinematical theory since most

**Figure 9.** 200 dark field LACBED pattern from a 9.5 nm  $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}/35$  nm InP plan view superlattice sample showing superlattice sidebands (numbered). (After R. Vincent et al. [38].)



information is at large deviation parameters. Layer thicknesses and periods can, for example, be determined to near monolayer accuracy. However, for a quantitative fit, dynamical theory is required. Very detailed agreements between dynamical simulations and experiment have been achieved, revealing, in particular, some structure-sensitive systematic row interactions [39, 40].

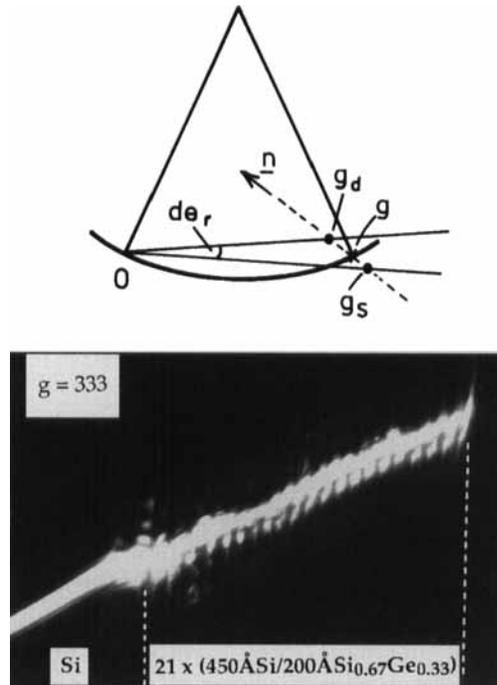
LACBED has also been used extensively to profile single buried layers in semiconductors using the plan view geometry. In this case, information on the buried layer appears as a modulation of the single crystal oscillations described by Eq. (3). Both layer composition and strain can be profiled, the latter generating asymmetry in the rocking curve between  $\pm s_g$ . Single atomic steps can be detected [41] and layers down to one or two monolayers analyzed [42]. By combining LACBED with complementary imaging methods, a detailed picture of layer topology at the atomic level can be derived.

LACBED has now been used to profile strains in a variety of structures. In strained multilayers the effect of strain is to cause a rotation of crystal planes inclined to the interfacial plane. The relative rotation between planes in adjacent lattice matched layers is given approximately by

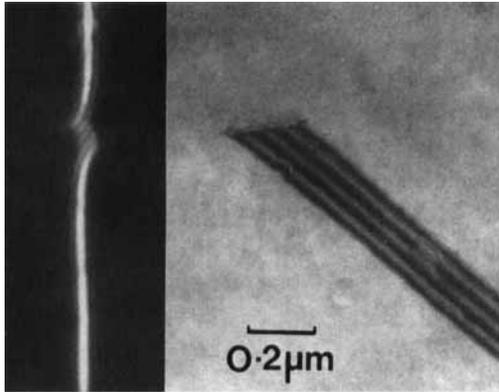
$$\delta\theta = \frac{\eta(1 + \nu)}{2(1 - \nu)} \sin 2\theta \quad (7)$$

where  $\eta$  is the natural mismatch,  $\nu$  is Poisson's ratio and  $\theta$  is the unstrained angle of inclination of the diffracting planes [43]. In reciprocal space this corresponds to a splitting of common reflections along the growth direction (see Fig. 10). This splitting can be observed in plan-view

specimens if the sample is tilted to select reflections from inclined planes. The method has also been applied to cross-sectional samples. Although LACBED is not sensitive to this splitting when the layers are edge-on, since the splitting is then perpendicular to the incident beam direction, a rotation of a few degrees from the edge-on orientation gives a component of the spot separation along the beam direction. This is observed as an angular displacement between diffraction contours. Figure 10 shows an example from recent work on Si/SiGe [44]. Quantitative studies are possible; interpretation can be complicated by surface relaxation in the cross-sectional geometry, although, in



**Figure 10.** Layer strains in a Si/SiGe superlattice observed by the displacement of a (333) contour. The schematic diagram shows how the diffraction vectors  $g_d$  and  $g_s$  from the Si and SiGe layers are displaced along the specimen normal  $n$ .



**Figure 11.** A dark field LACBED pattern taken in the reflection  $g = (0\bar{1}02)$  showing the image of a partial dislocation bounding a stacking fault in CdMnTe. The fault itself has no effect on the contour ( $g \cdot R = \text{integer}$ ). (After D. Cherns et al. [48].)

principle, such effects can be taken into account [45].

LACBED has proved particularly successful as a method of analyzing dislocations. It was first shown by Cherns and Preston [46] that relatively high order diffraction contours dissociate on crossing dislocations. An example is shown in Fig. 11. Under two-beam conditions, the number  $n$ , of subsidiary maxima in a dark field pattern, or subsidiary minima in the corresponding bright field pattern, is simply given by

$$g \cdot b = n + 1 \quad (8)$$

for integer  $g \cdot b$  values. Moreover, the direction of the displacement depends on the sign of  $b$ . Since the characteristics of the dislocation images depend principally on the long-range displacement of the diffracting planes, they are relatively insensitive to imaging parameters such as dislocation depth and foil thickness, and to the presence of impurities near dislocation cores. Breakdowns of the rule do occur where dislocations are close to the foil

surfaces and under strongly dynamical conditions [47]. The LACBED method of analyzing dislocations has proved particularly attractive because it gives both the magnitude and the sign of the Burgers vector which are difficult to obtain by imaging methods. Since high order contours can be used, the method is particularly suitable for partial dislocations with small Burgers vectors [48]. A key factor is the choice of reflections which put adjoining faults out of contrast. This is done in fig. 11 which shows part of the analysis of a partial dislocation terminating a stacking fault in CdMnTe. The reflection  $g \sim (0\bar{1}02)$  puts the fault out of contrast ( $g \cdot R = \text{integer}$ ); Eq. (8) can then be applied directly giving  $g \cdot b = 4$  (i.e.,  $n = 3$ ). The result is consistent with a Burgers vector of  $b = \frac{1}{3}(1\bar{1}1)$ .

Although dislocations introduce a strain field which varies with depth  $z$ , a stacking fault or interface can introduce a rigid body displacement  $R$ . It can easily be shown that a fractional  $g \cdot R$  in Eq. (6) introduces an asymmetry into the rocking curve, which is reversed when the sign of is reversed. This was used by Morniroli and Steeds [49] as a quick and reliable method of distinguishing intrinsic and extrinsic faults in Si, through the observation of HOLZ deficiency lines in a conventional CBED pattern. The method has been extended to the analysis of interface displacements, notably  $\text{CoSi}_2/(111)\text{Si}$ , where a more quantitative dynamical analysis was applied [50].

## Acknowledgements

One of us (JWS) wishes to thank Dr. Martin Saunders for assistance in preparing Figs. 3 and 5.

## 1.5.8 References

- [1] *Electron Diffraction Techniques*, Vol. 1 (Ed.: J. M. Cowley), Oxford University Press, Oxford **1992**.
- [2] J. C. H. Spence, J. M. Zuo, *Electron micro-diffraction*, Plenum Press, New York, **1992**.
- [3] *J. Electron Microsc. Tech.* **1989**, *13*, (1) and (2).
- [4] I. K. Jordan, C. J. Rossouw, R. Vincent, *Ultramicroscopy* **1991**, *35*, 237.
- [5] M. H. Loretto, *Electron Beam Analysis of Materials*, Chapman & Hall, London, **1984**.
- [6] J. W. Steeds, *Introduction to Analytical Electron Microscopy*, (Eds.: J. J. Hren, J. I. Goldstein, D. C. Joy, Plenum Press, New York, **1979**, 387.
- [7] J. P. Morniroli, J. W. Steeds, *Ultramicroscopy* **1992**, *45*, 219.
- [8] J. W. Steeds, J. P. Morniroli, *Reviews in Mineralogy*, **1992**, *27*, 37.
- [9] P. M. Jones, G. M. Rackham, J. W. Steeds, *Proc. R. Soc.* **1977**, *A354*, 197.
- [10] C. Sung, D. B. Williams, *J. Electron Microsc. Tech.* **1991**, *17*, 95.
- [11] See references in [10] and in Sec. 5.4. of [2].
- [12] B. F. Buxton, J. A. Eades, J. W. Steeds, G. M. Rackham, *Phil Trans. R. Soc.* **1976**, *A281*, 171.
- [13] a) M. Tanaka, *J. Electron. Microsc. Tech.* **1989**, *13*, 27; b) M. Tanaka, *Acta. Cryst.* **1994**, *A50*, 261.
- [14] J. W. Steeds, R. Vincent, *J. Appl. Cryst.* **1983**, *16*, 317.
- [15] P. B. Hirsch, A. Howie, R. B. Nicholson, D. W. Pashley, M. J. Whelan, *Electron Microscopy of Thin Crystals*, Robert Krieger, Florida, **1977**.
- [16] R. C. Ecob, *Scr. Metall.* **1986**, *20*, 1001.
- [17] D. M. Bird, *J. Electron. Microsc. Tech.* **1989**, *13*, 77.
- [18] J. W. Steeds, P. M. Jones, J. E. Loveluck, K. Cooke, *Phil. Mag.* **1977**, *36*, 309.
- [19] J. W. Steeds, *Ultramicroscopy* **1988**, *26*, 1.
- [20] H. Matsuhata, J. W. Steeds, *Phil. Mag.* **1987**, *B55*, 17.
- [21] J. R. Sellar, D. Imeson, C. J. Humphreys, *Acta. Cryst.* **1980**, *A36*, 686.
- [22] B. F. Buxton, *Proc. R. Soc.* **1976**, *A350*, 335.
- [23] B. K. Vainshtein, B. B. Zvyagin, A. S. Avilov in *Electron Diffraction Techniques*, Vol. 1 (Ed.: J. M. Cowley), Oxford University Press, Oxford **1992**, Chap. 6.
- [24] D. L. Dorset, *Adv. Electron. Electron Phys.* **1994**, *88*, 111.
- [25] D. M. Bird, *J. Electron Microsc. Tech.* **1989**, *13*, 77.
- [26] R. Vincent, D. M. Bird, J. W. Steeds, *Phil. Mag.* **1984**, *A50*, 745.
- [27] D. M. Bird, S. McKernan, J. W. Steeds, *J. Phys. C* **1985**, *18*, 499.
- [28] R. Vincent, R. L. Withers, *Phil. Mag. Lett.* **1987**, *56*, 57.
- [29] R. Vincent, D. R. Exelby, *Phil. Mag. Lett.* **1991**, *63*, 31.
- [30] R. Vincent, D. R. Exelby, *Acta Cryst.* **1995**, *A51*, 801.
- [31] R. Vincent, P. A. Midgley, *Ultramicroscopy* **1994**, *53*, 271.
- [32] D. M. Bird, M. Saunders, *Ultramicroscopy* **1992**, *45*, 241.
- [33] B. C. McCallum, J. M. Rodenburg, *Ultramicroscopy* **1993**, *52*, 85.
- [34] W. J. Vine, R. Vincent, P. Spellward, J. W. Steeds, *Ultramicroscopy* **1992**, *41*, 42.
- [35] R. Vincent, W. J. Vine, P. A. Midgley, P. Spellward, J. W. Steeds, *Ultramicroscopy* **1993**, *50*, 365.
- [36] R. A. Herring, G. Pozzi, T. Tanji, A. Tonomura, *Ultramicroscopy* **1995**, *60*, 153.
- [37] D. Cherns, *Inst. Phys. Conf. Ser.* **1991**, *117*, 549
- [38] R. Vincent, J. Wang, D. Cherns, S. J. Bailey, A. R. Preston, J. W. Steeds, *Inst. Phys. Conf. Ser.* **1987**, *90*, 233.
- [39] I. K. Jordan, A. R. Preston, L. C. Qin, J. W. Steeds, *Inst. Phys. Conf. Ser.* **1989**, *98*, 131.
- [40] D. Cherns, R. Touaitia, A. R. Preston, C. J. Rossouw, D. C. Houghton, *Phil. Mag.* **1991**, *A64*, 597.
- [41] I. K. Jordan, D. Cherns, N. Grigorieff, M. Hockly, P. C. Spurdens, M. R. Aylett, E. C. Scott, *Inst. Phys. Conf. Ser.* **1991**, *117*, 563.
- [42] N. Grigorieff, D. Cherns, M. J. Yates, M. Hockly, S. D. Perrin, M. R. Aylett, *Phil. Mag.* **1993**, *A68*, 121.
- [43] D. Cherns, C. J. Kiely, A. R. Preston, *Ultramicroscopy* **1988**, *24*, 355.
- [44] Y. Atici, D. Cherns, *Ultramicroscopy* **1995**, *51*, 435.
- [45] X-F. Duan, D. Cherns, J. W. Steeds, *Phil. Mag.* **1994**, *A70*, 1091.
- [46] D. Cherns and A. R. Preston, *Proc. of 11th Int. Congress on Electron Microsc. Kyoto*. (Eds.: T. Imura, S. Maruse and T. Suzuki) **1986**, *1*, 721.
- [47] C. T. Chou, A. R. Preston, J. W. Steeds, *Phil. Mag.* **1992**, *A65*, 863.
- [48] D. Cherns, J.-P. Morniroli, *Ultramicroscopy* **1994**, *53*, 167.
- [49] J.-P. Morniroli, J. W. Steeds, *Inst. Phys. Conf. Ser.* **1991**, *119*, 417.
- [50] M. A. Al-Khafaji, D. Cherns, C. J. Rossouw, R. Hull, *Inst. Phys. Conf. Ser.* **1991**, *119*, 51.

## 1.6 Low-Energy Electron Microscopy

### 1.6.1 Introduction

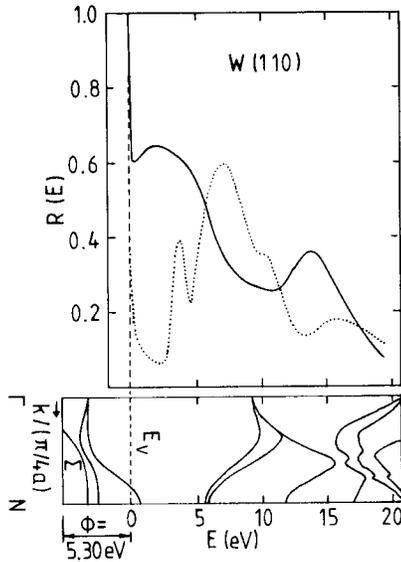
Low-energy electron microscopy (LEEM) is a surface-imaging method in which the surface is illuminated at normal incidence by slow electrons and the elastically reflected electrons are used to form an image of the surface. It is based on the high back-scattering cross-section which all materials have at low energies. In particular, in crystalline materials the back-scattered electrons are concentrated in one or more diffraction beams. One of them, usually the specular beam [the (00) beam], can be used to image the surface, wholly analogous to the manner in which the image is generated in the transmission electron microscope by the transmitted or by diffracted beams. LEEM is particularly useful for the study of processes on metal and semiconductor single-crystal surfaces. It is easily combined with other surface-imaging methods such as mirror electron microscopy (MEM) or photoemission electron microscopy (PEEM; see Chap. 6, Sec. 1 of this Volume). The lateral resolution of LEEM is limited by the aberrations of the cathode objective lens, whose cathode is the specimen, to the 1–10 nm range. The field of view can be as large as 100  $\mu\text{m}$ , depending upon the lens used, but the depth resolution is in favorable cases in

the subangstrom range. Many surface science and technology problems do not require a better resolution but a large field of view. Therefore, LEEM has a wide range of applications.

### 1.6.2 Theoretical Foundations

LEEM can be best understood by looking at the low-energy electron diffraction (LEED) pattern which is formed at the back focal plane of the cathode lens. Even in the absence of lens aberrations the LEED pattern is not the Fourier transform of the potential distribution in the surface layer of the specimen with which the incident electron wave interacts. This is due to the strong electron-specimen interaction which makes the first Born approximation invalid. As a consequence, the kinematical diffraction theory is insufficient for the calculation of the diffraction beams. Rather, the dynamical theory has to be used as described in various textbooks [1–4].

However, presently no quantitative information is drawn from intensity differences in LEEM images. Only the magnitude of the difference which determines the contrast is important. As an example,



**Figure 1.** Specular reflection coefficient of the W(110) and W(100) surfaces as a function of energy (top) and band structure of tungsten along the [110] direction ( $\Sigma$  direction) [6].

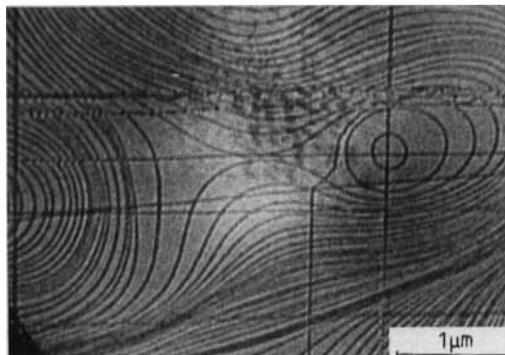
Fig. 1 shows the intensity  $I_{00}(E)$  of the specularly reflected beam from a W(110) and a W(100) surface as a function of energy  $E$  [5, 6]. The two  $I_{00}$  curves differ drastically in their energy dependence, so that it is always possible to find strong contrast between (110) and (100) oriented grains, for example at about 2 or 14 eV. On the other hand, no contrast may be seen if the energy is chosen wrongly, for example at about 6 or 11 eV. The figure also shows that reflectivities as high as 60% may be obtained. Such high values are connected with band gaps or regions with a low density of states (steep bands) in the band structure of the crystal along the propagation direction of the wave. For energies in band gaps no wave propagation is possible, and the wave is strongly damped and totally reflected: the crystal is a reactive medium. Ideally, the reflectivity should be 100%, but inelastic scattering,

which is predominantly in the forward direction, and phonon and defect scattering decrease it to the observed value. The missing electrons appear in the specimen current [6]. Similar considerations apply to energy ranges with a low density of states, but with increasing energy, part of the intensity is lost into other diffraction channels. For example, at 13.6 eV where  $I_{00}(E)$  has a maximum again on the W(110) surface, 10 diffraction beams in addition to the specular beam are excited as easily seen by an Ewald construction (taking the inner potential  $V_0 = 17.4$  eV into account). In addition, elastic back scattering decreases and inelastic scattering increases with energy and reduces  $I_{00}$  even more.

Diffraction contrast is not only strong between grains with different orientations but also between regions on a single-crystal surface which differ in surface structure, such as between reconstructed and unreconstructed surface regions or clean and adsorbate-covered regions, because these regions also differ in their  $I_{00}(E)$  dependence. Examples which will be discussed later are the contrast between  $(1 \times 1)$  and  $(7 \times 7)$  structures on Si(111) or between clean Mo(110) and copper monolayer islands on this surface. A frequent situation in surface science is the presence of an ordered overlayer in two or more azimuthal orientations. They are indistinguishable at normal incidence because the diffraction conditions are identical. Contrast can be obtained by changing the diffraction conditions, that is, by tilting the beam, or by imaging with one of the nonspecular overlayer beams. The resulting contrast has been termed tilted bright field and dark field contrast, respectively, while the contrast described previously has been called bright field contrast.

The contrast discussed up to now depends upon the periodic arrangement of the atoms in the surface region of the specimen. One of the main goals of surface microscopy is, however, to image deviations from this periodic arrangement, that is, surface defects. The most important defects are monoatomic steps, because many surface properties and processes depend strongly upon them. LEED is uniquely suited for the determination of the average step distribution which determines the profiles of the diffraction spots [7, 8]. The profiles are analyzed in terms of the kinematic theory because the back-scattered intensity may be separated into a dynamic and a kinematic term,  $|F|^2$  and  $|G|^2$ , respectively. The structure factor  $|F|^2$  contains the information on the size and shape of the two-dimensional unit cell and the atoms in it, the lattice factor  $|G|^2$  the size and shape of the periodic arrangement of the unit cells. A step terminates this arrangement by introducing a vertical phase shift which changes  $|G|^2$ . In bright field LEEM the step contrast can be easily understood in terms of this phase shift. A step of height  $d$  produces a path difference  $2d$  between the waves reflected from the adjoining terraces. When  $(2n + 1)\lambda/2 = 2d$ , the waves are out of phase and interfere destructively upon defocusing or as a consequence of the limited phase contrast transfer function of the objective lens. By determining the  $\lambda$  values—that is, the energies  $E = 1.5/\lambda^2$  (in electronvolts when  $\lambda$  is in nanometers)—at which the contrast is strongest, the step height can be determined with high accuracy. LEEM thus allows the determination of the local step distribution. An example is shown in Fig. 2.

Another contrast which is also a consequence of path differences between



**Figure 2.** LEEM image of the step distribution on an Mo(110) surface. The rounded steps were formed by sublimation during the preceding high-temperature treatment, while the straight steps are the result of glide during cooling. The electron energy is 14 eV [9].

different areas on the specimen is the quantum size effect contrast which appears in thin films with parallel top and bottom boundaries. It is the electron-optical counterpart to reflection-reducing or -enhancing layers in optics: the reflectivity has a maximum when  $2d = n\lambda_f$  and a minimum when  $2d = (2n + 1)\lambda_f/2$  ( $\lambda_f$  is the wavelength in the film). Thus, the thickness, and thickness differences, can be determined with high accuracy. Finally, it has to be kept in mind that the specimen is the cathode in a cathode lens. This means that the specimen is in a high field, and any deviation from planarity will cause a field distortion which will deflect the electrons. These deflections produce a topographic contrast which is frequently very helpful. However, if the field distortions of the surface features overlap, image interpretation becomes difficult if not impossible.

An important aspect of any surface microscopy is the information depth. In LEEM, as in any other surface probe which uses slow electrons, the information depth is, in general, determined mainly by

inelastic scattering but also by thermal diffuse scattering and elastic back scattering. Elastic back scattering is important mainly at very low energies at which, frequently, only the specular beam can be excited, as discussed above, and a band gap exists. The information depth is then of the order of the wavelength of the electron, which is 3–4 Å at zero energy for realistic mean inner potentials due to the exponential damping in total reflection. Thermal diffuse scattering, in particular multiphonon scattering which causes scattering predominantly far from the (0,0) beam, becomes important at high temperatures.

Inelastic scattering due to collective and one-electron excitations dominates the information depth at most temperatures. The inelastic mean free path has a minimum value of 3–5 Å around 50 eV which is predominantly caused by plasmon excitation. Toward lower energies—below the plasmon excitation threshold—it increases rapidly, and toward higher energies more slowly. The attenuation length in the LEEM energy range (0 eV to about 200 eV) is, thus, ranges from a few angstroms to 10 Å. In materials with fully occupied d bands, higher values are possible at very low energies because the possibilities of low-energy excitations are rather limited in them. If the d band is only partially occupied, low-energy excitations with high cross-section are possible, and the inelastic mean free path may decrease to less than 3 Å (see this Chap. 4, Sec. 1.1.6 of this Volume). On the other hand, in large band gap materials much larger values may occur. Depth information may also be obtained indirectly, for example on buried interfaces via the strain field of the interfacial dislocation, which may

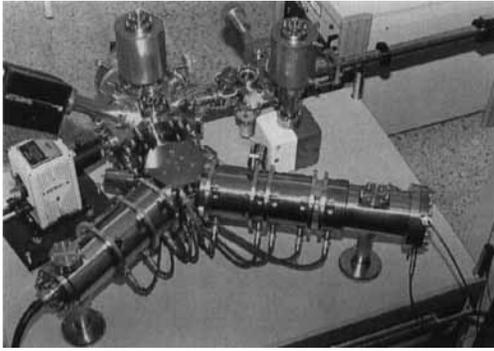
give sufficient diffraction contrast at layer thicknesses of the order of 100 Å.

Additional information on electron–specimen interaction and on contrast formation may be found in the literature [10–12].

### 1.6.3 Instrumentation

A LEEM instrument differs in several aspects from conventional electron microscopes. The very nature of its purpose, the imaging of well-defined surfaces, requires an ultrahigh vacuum (UHV) with a base pressure in the low  $10^{-10}$  to high  $10^{-11}$  mbar range. Similarly to the mirror microscope, illuminating and imaging must be separated if an image is to be produced. Otherwise only diffraction is possible. Finally, the specimen must be at a high negative potential or the complete optical system must be at high positive potential if the specimen should be at ground potential. Otherwise, an LEEM instrument makes use of the usual components of standard electron microscopes, albeit using UHV technology. The first (ill-fated) system was built with glass technology using electrostatic lenses and 90° deflection. Until recently, all subsequent systems were all-metal systems with magnetic lenses and 60° deflection [9, 13, 14]. The desire to have the specimen at ground potential has led to the development of electrostatic lens systems, the first of which made use of double 45° deflections [15, 16]. More recently, instruments with 90° deflection have been proposed again [17, 18].

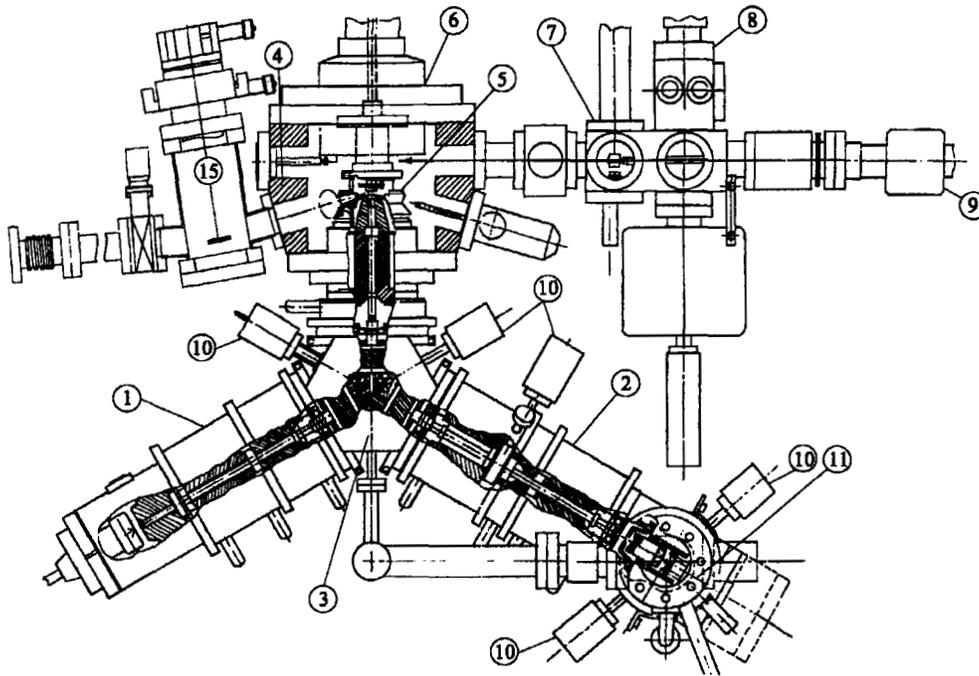
As an example, the most sophisticated instrument at present will be described



**Figure 3.** Overall view of a versatile LEEM instrument [12].

[13, 19]. Figure 3 shows its overall view, and Fig. 4 its cross-section, with an energy filter added. The design concept of the instrument is based on the desire to use it

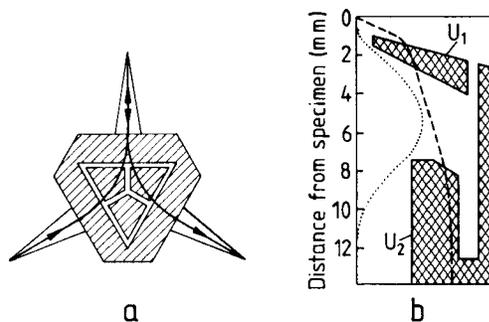
not only for LEEM but also for Auger electron emission microscopy (AEEM) and secondary electron emission microscopy (SEEM). In these two imaging modes the electron energies in the illumination and in the imaging sections are different: the illumination system must allow a wide range of illumination conditions and the imaging section must be able to accommodate an imaging energy filter. For efficient SEEM, primary beam energies up to about 500 eV are needed because the secondary electron yields of most materials peak at a few hundred electronvolts, while the secondary electron energy distribution peaks between a few electronvolts and 10 eV. In AEEM, primary electrons up to a few kiloelectronvolts are necessary in



**Figure 4.** Cross-section of the instrument shown in Fig. 3 but with an energy analyzer added: 1, illumination column; 2, imaging column; 3, beam separator; 4, specimen chamber; 5, cathode lens; 6, specimen manipulator; 7, specimen preparation chamber; 8, air lock; 9, transfer rod; 10, aperture manipulators; 11, energy filter; post-filter projective and detection system; 15, refocusing mirror chamber for synchrotron radiation.

order to be able to detect all elements, while the characteristic energies of the Auger electrons range from a few tens of electronvolts to about 1 keV.

For LEEM, and in particular for LEED, the specimen should be illuminated with a parallel beam; intensity is of minor concern because of the strong elastic back scattering. For AEEM and SEEM, parallel illumination is unimportant, but a high illumination intensity is required, at least for AEEM, because the Auger electron signals are weak and ride on a strong background. Without high intensity, excessive image acquisition times would be necessary for images with a sufficient signal-to-noise ratio. In order to achieve this flexibility an  $\text{LaB}_6$  cathode is used plus a double condenser and a transfer lens which images the demagnified cross-over either onto the back focal plane of the objective lens for LEEM and LEED or— together with the objective lens—onto the specimen for AEEM and SEEM. In the LEEM and LEED modes the beam aperture angle in the back focal plane and, therefore, the illuminated area on the specimen is controlled by an aperture in the beam separator. The beam separator, which is designed to be double focusing, consists of several sections which can be excited independently of each other so that primary and secondary electrons with different energies can be deflected by the same angle ( $60^\circ$ ) [20]. The 'close-packed prism array' used in the instrument in Figs. 3 and 4 is schematically shown in Fig. 5a with the central ray path indicated. The objective lens is a magnetic triode lens, that is, a magnetic lens in which the pole piece next to the specimen is electrically isolated (Fig. 5b) so that its potential can be varied from ground potential to near-



**Figure 5.** (a) Double-focusing separator for incident and emitted beams with different energies. (b) Magnetic triode cathode objective lens. The axial electrostatic potential and magnetic field are indicated by dashed and dotted lines, respectively.

specimen potential. This allows the study of surfaces at low fields. Other cathode lenses can and have been used as well [12]. The contrast aperture is not placed in the back focal plane of the objective lens—where the LEED pattern is located—as is usually done in standard electron microscopes—for two reasons: (1) for dark field imaging, that is, imaging with nonspecular beams, the aperture has to be shifted laterally and would intercept the incident beam for large shifts, and (2) the aperture would limit the intensity in the AEEM and SEEM illumination modes. Therefore, the contrast aperture is placed further downstream in an image of the LEED pattern. The objective lens is focused in such manner that the image of the specimen is in the center of the separator, where it is, in principle, achromatic. For selected-area LEED, a field-limiting aperture can be placed in this position. However, in most applications the illumination aperture is sufficient for this purpose, and the port for the manipulator of the field limiting aperture is used for a fluorescent screen with which the primary beam can be inspected (cathode

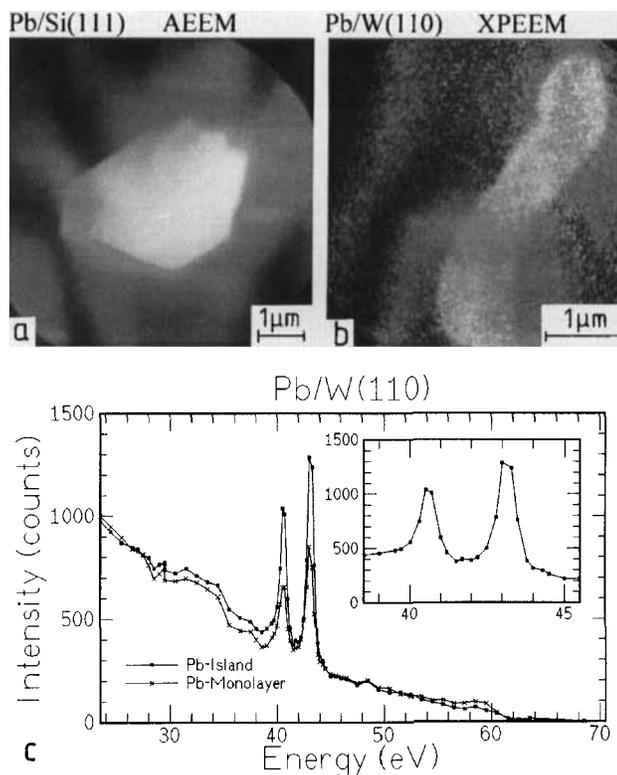
inspection, illumination system alignment). The contrast aperture in the image plane of the LEED pattern limits not only the angular aperture used in further image formation but acts also to a certain extent as a rough energy filter due to the (small) chromatic aberration of the deflection field.

The imaging system after the deflection field must allow the placement of either the image or the LEED pattern at the input side of the energy filter. For proper matching and magnification selection, several lenses are necessary [13, 19]. The energy filter is a  $90^\circ$  electrostatic sector in which the electrons are decelerated to about 1/10th of the energy which they have in the rest of the instrument except in the cathode field region. The final energy-filtered image is obtained by a double-gap projective lens on a double-channel plate-fluorescent screen detector. Of course, if the instrument is to be used only for LEEM, LEED, or mirror and emission electron microscopy with very slow electrons such as thermionic electron emission microscopy (TEEM) or near-ultraviolet PEEM, no energy filter, fewer lenses and a simpler beam separator are needed, and the instrument can be much simpler, as is the case for other instruments [9, 21]. In addition to the optical components mentioned up to now, the instrument contains a number of deflection systems for beam alignment and three stigmators, one in the objective lens, and one each in the illumination and imaging columns. The last two not only correct lens astigmatism but also the residual cylinder lens action of the beam separator.

The objective lens is located in a large specimen chamber with six ports pointing at the specimen position for mounting

accessories for in situ experiments (gas sources, evaporators, etc.) and for illumination with a high-pressure mercury short arc lamp for PEEM. The specimen is mounted on an eccentric  $x$ - $y$ - $z$  manipulator on which it can be heated to  $1600^\circ\text{C}$ . Specimen exchange is through an airlock via a specimen preparation chamber. The system is pumped by sputter ion and titanium sublimation pumps, and has a base pressure in the  $10^{-11}$  mbar range in the specimen chamber after a bake-out at  $180^\circ\text{C}$ . LEEM image acquisition is usually via a video camera, but a still camera can be used for high image quality. Spectroscopic images are recorded with a slow scan Peltier-cooled charge-coupled device camera. Figure 6a shows an electron beam-excited AEEM image, and Fig. 6c the local photoelectron spectrum which was obtained from a series of synchrotron radiation-excited photoelectron images taken in small energy increments (0.5 eV) and with good energy resolution (0.5 eV) such as the one shown in Fig. 6b.

Provided that vibrations, current and high-voltage fluctuations, and alternating current fields are low enough, the resolution of an LEEM is limited mainly by the lens aberrations of the cathode lens, the aberrations of the beam separator and—in a spectroscopic instrument—the aberrations of the energy filter. With a well-designed beam separator and energy filter and a well-aligned system the cathode lens determines the ultimate resolution, in particular its acceleration field region. In an idealized situation the lens can be divided into two sections: a homogeneous acceleration field, which forms a virtual image behind the specimen, and a subsequent conventional electron lens, which forms a real image of the virtual image. The



**Figure 6.** Specimen characterization in a spectroscopic LEEM. Lead layer on (a) an Si(111) surface and (b, c) a W(110) surface. (a) Electron-excited AEEM image of a lead crystal (primary energy 2.5 keV, imaging energy 94 eV, 5 s exposure); (b) synchrotron radiation-excited XPEEM image; (c) X-ray ( $h\nu = 60.7$  eV) photoelectron spectrum of a  $0.4 \mu\text{m}^2$  area on a lead crystal, and of a  $0.8 \mu\text{m}^2$  area of a lead monolayer obtained from a sequence of images (b) (20 s/image) [23].

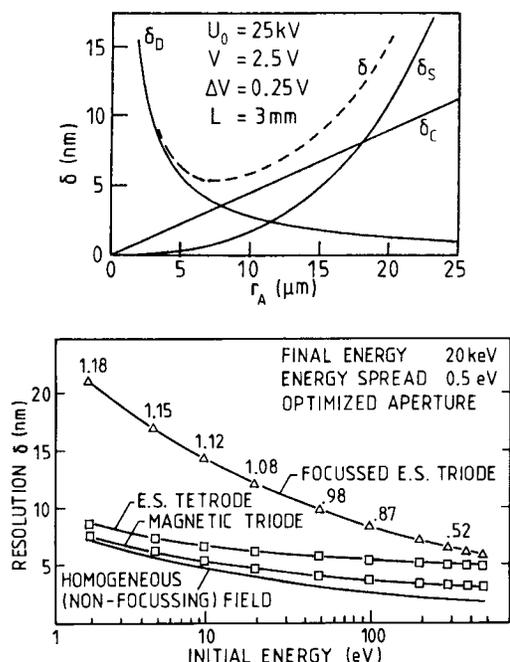
aberrations of the unavoidable acceleration field are dominating. They can be easily calculated analytically. For a fixed aperture radius  $r_A$  at the exit of the homogeneous field the spherical and chromatic aberrations are proportional to  $\Delta E/F$  and  $E/F$ , respectively, where  $F$  is the field strength at the cathode,  $E$  the average start energy of the electrons at the specimen, and  $\Delta E$  the energy width [22]. The actual values for a specific example are shown in Fig. 7a. For comparison, Fig. 7b shows the dependence of the resolution upon the start energy for three cathode lens types used in LEEM instruments for an energy width of 0.5 eV [24]. Reduction of the energy width to  $\Delta E = 0.1$  eV reduces the resolution  $\delta$  somewhat, but  $\delta = 4$  nm is about the limit at a start energy

of 2 eV. A further improvement of  $\delta$  is possible with elaborate correction elements [17] but has not been realized yet.

## 1.6.4 Areas of Application

### 1.6.4.1 Clean Surfaces

Because of its high depth resolution, LEEM is very well suited for the study of surface topography. The imaging of monoatomic steps has already been mentioned (Fig. 2), but etch pits, hillocks, or screw dislocations can be imaged as well. The real-time and high-temperature capabilities of LEEM also allow the life study



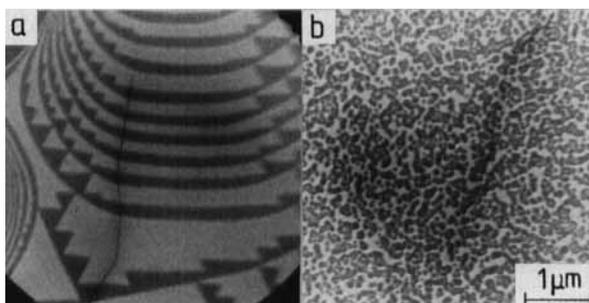
**Figure 7.** Resolution of the cathode lens. (a) Dependence of spherical aberration resolution  $\delta_S$ , chromatic aberration resolution  $\delta_C$  and diffraction disk  $\delta_D$  resolution upon aperture radius  $r_A$  of a homogeneous field of length  $L$  [22]. (b) Resolution of the electrostatic triode, electrostatic tetraode, and magnetic triode (Fig. 6) as a function of start energy. The values on the upper curve are the field strengths at the specimen [24].

of changes in topography due to surface diffusion and other processes such as sublimation. For the same reasons, the microstructural aspects of structural phase transitions are particularly attractive for LEEM. A well-known example is the  $\text{Si}(111)-(1 \times 1) \leftrightarrow (7 \times 7)$  phase transition, which can occur in a wide variety of morphologies with little difference visible in laterally averaging techniques such as LEED (Fig. 8) [25]. Domain growth on reconstructed surfaces [26a] or faceting of vicinal surfaces [26b] are other problems well-suited for LEEM.

### 1.6.4.2 Adsorption Layers

Many foreign atoms or molecules on surfaces have overall attractive lateral interactions even if the nearest-neighbor interactions are repulsive. Therefore, they form islands which produce structural contrast. This is also the case when the adsorbate and substrate have the same lateral periodicity as in submonolayers of copper on  $\text{Mo}(110)$ . The contrast between island and substrate is due to differences in back-scattering cross-section and phase. In general, however, the adsorbate layer has a lateral periodicity different from that of the substrate and, as a consequence, forms a number of equivalent domains. Dark field imaging with superstructure reflections characteristic for a certain domain then allows them to be imaged individually. An example is shown in Fig. 9. The two  $(5 \times 1)$  domains which appear dark in the bright field image (Fig. 9b) are easily distinguished by imaging with their corresponding  $(1/5, 0)$  order beams. The correlation between domain boundaries and steps is also visible.

Most adsorbates experience order-disorder transitions with increasing temperature before desorption occurs. Many of these transitions appear to be continuous in laterally averaging studies, for example with LEED. Gold adsorbates on  $\text{Si}(111)$  are again a good example [27]. Another is the misfit flip phase transition in the copper double layer on  $\text{Mo}(110)$  [28]. This transition layer is also a demonstration of the usefulness of being able to image at high temperatures, which allows one to follow the kinetics of the strain-induced surface restructuring of the original flat surface into the structure shown in Fig. 10 [29] and the subsequent

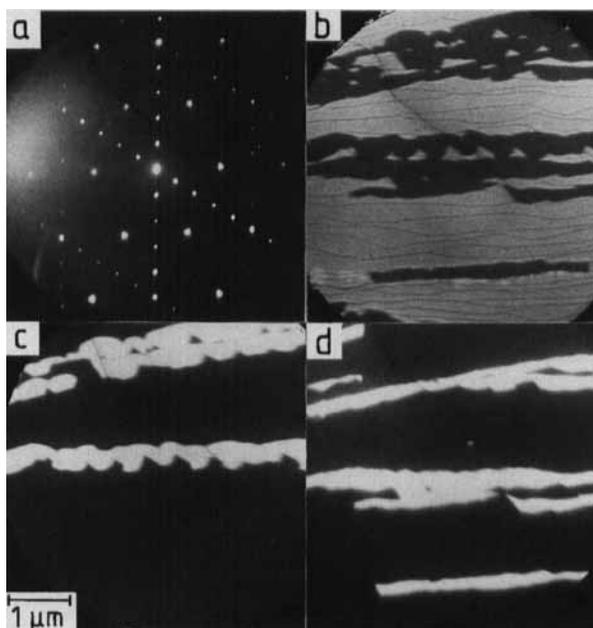


**Figure 8.** LEEM images taken during the Si(111)-(1 × 1) ↔ (7 × 7) phase transition. Both surfaces show a good (7 × 7) LEED pattern below the transition temperature (electron energy 11 eV) [25].

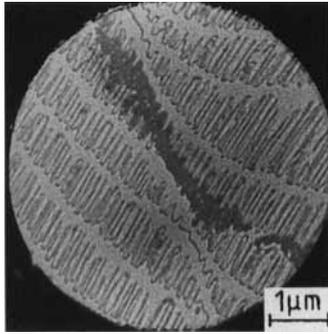
straightening of the steps during copper desorption. Desorption from the two-phase region (coexisting two-dimensional islands and gas) can also be easily distinguished by LEEM from desorption from the single-phase region (two-dimensional gas only). By slowly lowering the temperature, the condensation of two-dimensional islands from the two-dimensional gas can be studied. An example is shown in Fig. 11 [30]. During cooling, the gold atoms did not diffuse to the steps and condense there

but formed numerous strongly elongated islands with one of the two crystallographically equivalent orientations extending over large areas, indicating island–island interactions through the adsorbate-induced strain field in the substrate.

If an adsorbate is deposited locally, then surface diffusion can be studied by following the propagation of the boundary between characteristic adsorbate structures which differ in coverage. Similarly, reaction fronts between two different



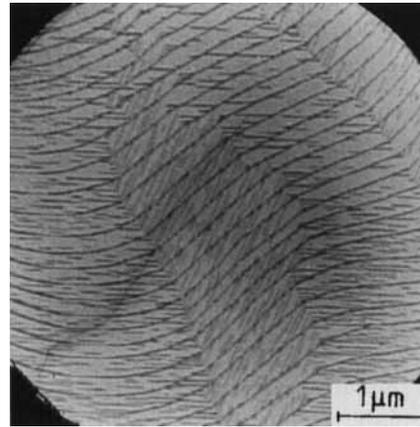
**Figure 9.** (a) LEED pattern (30 eV), (b) bright field image and (c, d) two dark field images taken with different 1/5 order reflections of a mixture of the  $(\sqrt{3} \times \sqrt{3})$  R30° and the (5 × 1) structure on Si(111) in two domain orientations (electron energy 6 eV) [27].



**Figure 10.** Step break-up of an Mo(110) surface covered with a copper double layer at about 900 K (electron energy 3 eV) [28].

adsorbates can be used to follow chemical reactions on surfaces [31]. This can be done also by MEM [32] and PEEM [33] if characteristic work function differences exist between the reacting adsorbates, a situation frequently encountered such as in the case of co-adsorbed oxygen and carbon monoxide.

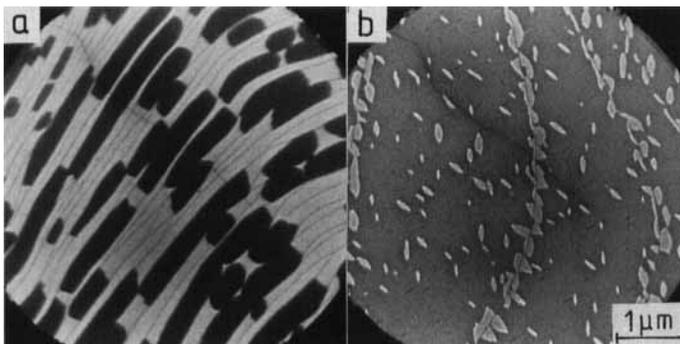
Adsorption of impurities from the interior of the crystal, that is, segregation, is also a good subject for LEEM. This is illustrated in Fig. 12 by the segregation of carbon to the (110) surface of molybdenum, where it forms  $\text{Mo}_2\text{C}$  crystals in a large morphological variety, depending upon the carbon concentration, cooling rate, and other parameters. Figure 12



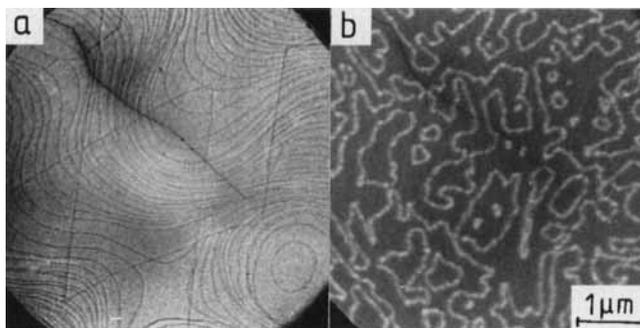
**Figure 11.** Two-dimensional condensation of gold from the two-dimensional gas phase on Mo(110) (electron energy 14 eV) [29].

also shows the influence of the electron energy on the contrast: at 3 eV the  $\text{Mo}_2\text{C}$ -covered regions are dark and the clean molybdenum surface bright, while at 9 eV the reverse is true.

Another frequent surface phenomenon is the formation of two-dimensional compounds by interactions with reactive gases. The early stages of oxidation have been studied with laterally averaging techniques for decades, but little is known about the processes on a microscopic level. Although not much work has been done with LEEM, a good deal of insight has already



**Figure 12.**  $\text{Mo}_2\text{C}$  formation on Mo(110) by carbon segregation from the bulk: electron energies of (a) 3 eV and (b) 9 eV [29].

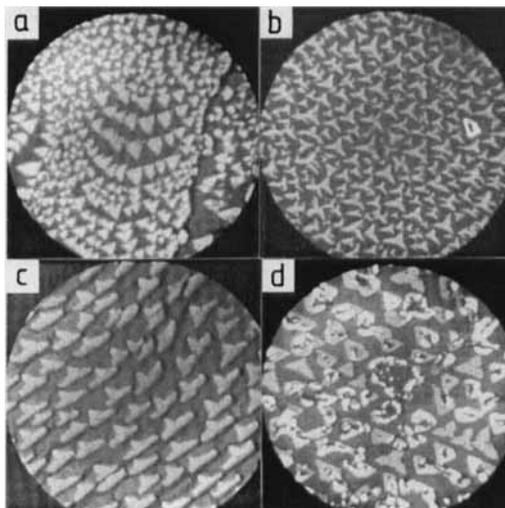


**Figure 13.** Clean W(100) surface ((a) electron energy 7 eV) and onset of two-dimensional oxide formation ((b) electron energy 16 eV) [34].

been gained. While in many surface processes, steps play an important role, this does not seem to be the case in the high-temperature oxidation of W(100) at low pressure. Once domains of chemisorbed oxygen have formed—which involves a considerable amount of surface rearrangement—the domain boundaries are the weak points that determine the further oxidation process. Figure 13 shows the initial step structure on a clean W(100)

after a high-temperature flash and the transition from the chemisorbed layer to the two-dimensional oxide which starts at the domain walls and spreads from them rapidly across the surface [34].

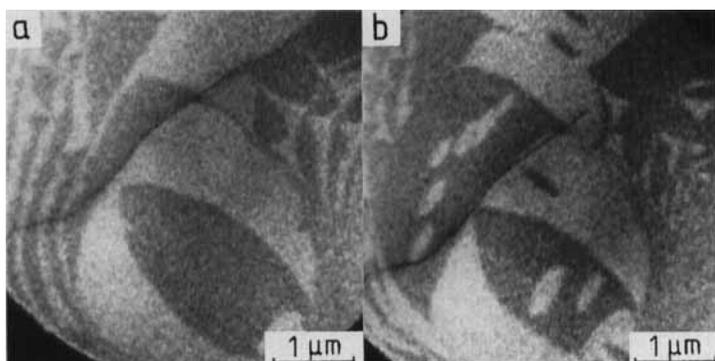
Two-dimensional silicon nitride formation in the early stages of the nitridation of Si(111) in ammonia at high temperatures proceeds quite differently. The nucleation clearly occurs at steps, and the  $(8 \times 8)$  nitride structure then spreads across the terraces in a manner similar to the growth of the  $(7 \times 7)$  structure on the clean surface. Again, depending upon nitridation conditions and the purity of the surface, a wide variety of morphologies and kinetics is observed. A typical frame from a video taken during this process is shown in Fig. 14. The combination with scanning tunneling microscopy has proven very useful in understanding structure and growth of the two-dimensional nitride [35].



**Figure 14.** Nucleation and growth of two-dimensional nitride with an  $(8 \times 8)$  structure on Si(111) by reaction with ammonia at 1150 K (electron energy 3 eV). Video frames are from growth experiments [35].

### 1.6.4.3 Thin Films

One of the major application areas of LEEM is the study of the early stages of the growth of thin films, in particular of how surface imperfections influence the growth and how reaction layers grow.

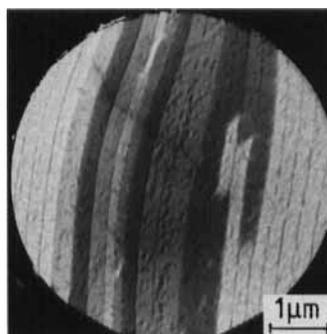


**Figure 15.** Homoepitaxy of silicon on Si(100) at (a) very low supersaturation and (b) low supersaturation (electron energy 5 eV). Video frames are from a growth experiment [36].

Most metal–metal and metal–semiconductor film–substrate combinations alloy or react, at least at high temperatures, but there are also some less complicated systems, the simplest being homoepitaxy. One of the most important processes is the homoepitaxy of silicon on Si(100), which has also been studied with LEEM. At low supersaturation, at which no nucleation can take place on the terraces, new terraces nucleate only at imperfections, and spread over the existing terraces in a step–flow growth mode. The growth rate is highly anisotropic, reflecting the large difference between the step energies of steps parallel and perpendicular to the dimer rows on this  $(2 \times 1)$ -reconstructed surface. This leads to the almond-shaped terraces seen in Fig. 15a. At somewhat higher supersaturation, nucleation on the terraces occurs (Fig. 15b), and the images rapidly become too complex for analysis [36].

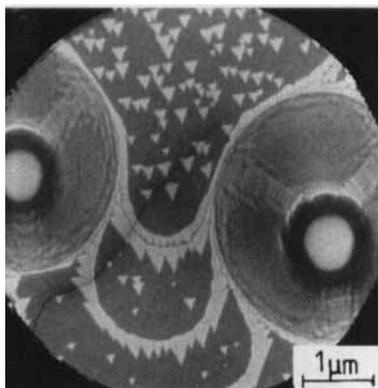
The next simplest growth process is heteroepitaxy without alloying or reaction. Typical examples are copper, silver, and gold on W(110). Below 400–500 K these layers grow in a quasi-monolayer-by-monolayer mode, but at the upper temperature limit they grow at significantly different rates on different terraces. Some terraces have preferred nucleation sites

from which the monolayers grow rapidly along the terraces, obviously incorporating atoms that diffuse across the steps from other terraces. This produces considerable thickness variations from terrace to terrace, which can be studied well by making use of the quantum size effect (Fig. 16) [37a]. At high temperatures, only one or two adsorbed layers are formed on which flat three-dimensional crystals grow. An example can be found in the chapter on SPLEEM (see Chap. 5, Sec. 3 of this Volume), in which also the agglomeration of low temperature-deposited layers upon annealing are discussed as well as alloying in miscible film–substrate systems.

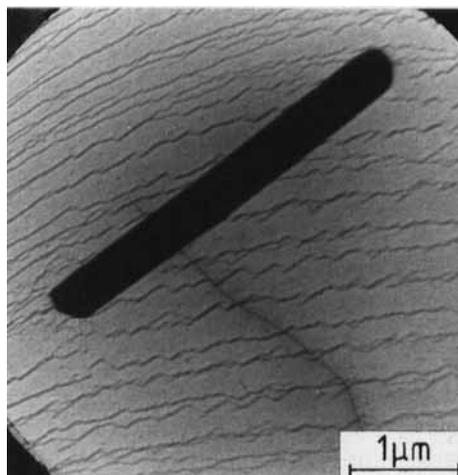


**Figure 16.** Terrace-limited monolayer-by-monolayer growth of copper on Mo(110) imaged via quantum size effect contrast (electron energy 4 eV) [37a].

The most complex film growth process involves reaction with the substrate. The growth and disintegration of silicide layers on silicon is an important example of this type.  $\text{CoSi}_2$  layers have attracted considerable attention because of their possible applications in microelectronics. An LEEM study showed that these layers tend to grow as three-dimensional crystals with a clean silicon surface in between. The  $\text{CoSi}_2$  crystals thus act as scavengers for cobalt atoms. Heating such a layer to a high temperature causes it to break up into many small three-dimensional crystals, from which a few large ones form via Oswald ripening at very high temperatures.  $\text{CoSi}_2$  sublimes more slowly than silicon and, therefore, large hillocks grow, with  $\text{CoSi}_2$  crystals on top (Fig. 17). This example shows that three-dimensional structures can also be imaged quite well, in spite of the field distortion caused by them, provided they are well separated. An additional condition for seeing structure in LEEM is that the faces of the three-dimensional crystal may not be tilted very much. This is illustrated in Fig. 18 with a

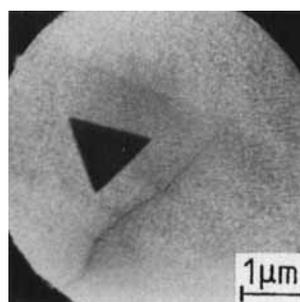


**Figure 17.** Hillocks on Si(111) with  $\text{CoSi}_2$  crystals on top, formed by faster sublimation of the substrate (electron energy 10 eV) [25].



**Figure 18.**  $\text{Cu}_3\text{Si}$  crystal surrounded by a two-dimensional copper silicide layer on Si(111) (electron energy 5 eV) [37b].

$\text{Cu}_3\text{Si}$  crystal obtained by copper deposition on Si(111). The structure of the crystal is identified by selected-area LEED, which simultaneously shows that the crystal has tilted top faces. The specular beams from these faces are off-axis and, therefore, are intercepted by the contrast aperture so that the crystal appears dark [37b]. In contrast to  $\text{CoSi}_2$  crystals, the  $\text{Cu}_3\text{Si}$  crystal is not surrounded by clean silicon but by a two-dimensional step faceting, which causes strong step faceting and



**Figure 19.** Au-Si eutectic 'particle' on an Si(111) surface at about 900 K [27].

agglomeration. Not all particles which are dark at all energies in LEEM are faceted. The dark triangle in Fig. 19 is not a crystal but a liquid Au–Si eutectic on an Si(111) surface which has well-defined equilibrium boundaries. It is dark because there is no well-defined specular beam due to the lack of crystallinity [27].

### 1.6.5 Discussion

When should LEEM be used for surface or thin film studies, and when not? The situations when it should not be used are obvious: when a routine resolution of better than 10 nm is needed, when the specimen is not UHV compatible because of high vapor or dissociation pressures, when it is very electron beam-sensitive due to electron-stimulated dissociation or desorption, or when it is very insulating or very rough. Another limitation is the gas pressure at the specimen. These situations are, however, not as LEEM-exclusive as they appear. Further improvements in objective and beam separators should allow a skilled operator to reach the middle 1 nm range for suitable objects. Specimen cooling will reduce the pressure problems, and the use of electron energies below the thresholds for dissociation or desorption, which are easily accessible in LEEM, will eliminate specimen damage. Charging of insulating samples can be frequently eliminated by heating, sometimes by choosing electron energies at which the secondary electron emission coefficient is unity. The only remedy for roughness is smoothing the surface by proper preparation procedures. As far as admissible

pressures are concerned, LEEM images have been taken at pressures up to the  $10^{-6}$  mbar range without noticeable image deterioration. This is possible because beam separator, illumination, and imaging columns are pumped separately from the specimen chamber, and the narrow constrictions between the chamber and the columns allow the maintenance of a large pressure gradient. It is mainly high-voltage stability which limits the maximum operating pressure. A major prerequisite is, of course, that the system is always clean, otherwise carbon contamination will make well-defined surface studies impossible.

The major competitors to LEEM in surface studies are high-energy reflection electron microscopy (REM; see Chap. 4, Sec. 1.2 of this Volume) and standard scanning electron microscopy (SEM; Chap. 4, Sec. 2 of this Volume). Both types of instruments are available with surface science-compatible pressures, at least in the specimen region, and bake-out possibilities. Advantages of REM are higher resolution perpendicular to the direction of the electron beam and access to the specimen at nearly normal incidence. A great disadvantage is the strong foreshortening in the beam direction caused by the grazing incidence of the electron beam. The major advantages of SEM are that it can be used for rough surfaces, that local spectroscopy is much faster in a scanning electron microscope equipped with an energy analyzer than in the corresponding LEEM instrument. However, image acquisition in an LEEM is much faster than in an SEM because of the parallel, versus sequential, detection. Also, the surface sensitivity is usually poorer in SEM than in LEEM and the

contrast is more limited. The last two disadvantages can be reduced by going to very low energies, for example by decelerating the electrons immediately before the specimen (scanning LEEM, SLEEM [38, 39]), but this introduces one of the main disadvantages of LEEM, the field distortions on rough samples. A great advantage of LEEM compared to the other two techniques is its easy combination with other imaging modes such as TEEM, PEEM, or SEEM and, in particular, MEM. MEM is very useful on non-crystalline surfaces, for example before the removal of amorphous surface layers such as SiO<sub>2</sub> on crystalline silicon or the (chemo)mechanical polishing layer on single-crystal surfaces. LEED, which is easily obtained by removing the contrast aperture and changing the excitation of the intermediate lens, is indispensable for chemical identification via the diffraction pattern.

Finally, a comparison with the various forms of scanning probe microscopies such as scanning tunneling microscopy (STM) or atomic force microscopy (AFM) is necessary. STM and AFM have an intrinsically higher resolution but are much slower due to their mechanical scanning image acquisition. When operated at high resolution, their field of view is small and vice versa. Operation over a wide temperature range is still a problem, but present STM and AFM instrument developments show that this can be overcome. Another, albeit minor, restriction is in deposition studies, because of the shadowing by the scanning tip. Experience has shown that STM and LEEM are more complementary than competing techniques [35, 40].

## 1.6.6 Concluding Remarks

LEEM is still a relatively young technique. Only a few instruments presently exist, and the number of studies and publications is still limited, as the latest review [12] shows. For more recent work, see Michely et al. [41], Tromp and Michely [42] and Altman and Bauer [43].

## 1.6.7 References

- [1] J. B. Pendry, *Low Energy Electron Diffraction*, Academic Press, New York **1974**.
- [2] M. A. Van Hove, S. Y. Tong, *Surface Crystallography by LEED*, Springer, Berlin **1979**.
- [3] L. J. Clarke, *Surface Crystallography*, Wiley, Chichester **1985**.
- [4] M. A. Van Hove, W. H. Weinberg, C.-M. Chen, *Low-Energy Electron Diffraction*, Springer, Berlin **1986**.
- [5] H.-J. Herlt, R. Feder, G. Meister, E. Bauer, *Solid. State Commun.* **1981**, 38, 973.
- [6] H.-J. Herlt, Ph.D. thesis, TU Clausthal, **1982**.
- [7] M. Henzler, *Appl. Surf. Sci.* **1982**, 11/12, 450; M. Henzler, *Appl. Phys.* **1984**, A34, 205.
- [8] H.-N. Yang, G.-C. Wang, T.-M. Lu, *Diffraction from Rough Surfaces and Dynamic Growth Fronts*, World Scientific, Singapore **1993**.
- [9] W. Telieps, E. Bauer, *Ultramicroscopy* **1985**, 17, 57.
- [10] E. Bauer, W. Telieps in *Surface and Interface Characterization by Electron Optical Methods* (Eds.: A. Howie, U. Valdre), Plenum, New York **1988**, p. 195.
- [11] E. Bauer in *Chemistry and Physics of Solid Surfaces VII* (Eds.: R. Vanselow, R. Howe), Springer, Berlin **1990**, Chap. 12.
- [12] E. Bauer, *Rep. Prog. Phys.* **1994**, 57, 895.
- [13] L. H. Veneklasen, *Ultramicroscopy* **1991**, 36, 76.
- [14] E. Bauer, *Surf. Sci.* **1994**, 299/300, 102.
- [15] K. Grzelakowski, T. Duden, E. Bauer, H. Poppa, S. Chiang, *IEEE Trans. Magn.* **1994**, 30, 4500.
- [16] K. Grzelakowski, E. Bauer, *Rev. Sci. Instrum.* **1996**, 67, 742.
- [17] H. Rose, D. Preikszas, *Optik* **1992**, 92, 31.
- [18] R. M. Tromp and M. Mankos, personal communication.

- [19] L. H. Veneklasen, *Rev. Sci. Instrum.* **1992**, *63*, 5513.
- [20] V. Kolarik, M. Mankos, L. Veneklasen, *Optik* **1991**, *87*, 1.
- [21] R. M. Tromp, M. C. Reuter, *Ultramicroscopy* **1991**, *36*, 91.
- [22] E. Bauer, *Ultramicroscopy* **1985**, *17*, 51.
- [23] C. Koziol, T. Schmidt, G. Lilienkamp, E. Bauer, to be published.
- [24] J. Chmelik, L. Veneklasen, G. Marx, *Optik* **1989**, *83*, 155.
- [25] E. Bauer, M. Mundschaue, W. Swiech, W. Teliieps, *J. Vac. Sci. Technol.* **1991**, *A9*, 1007.
- [26] (a) R. J. Phaneuf, N. C. Bartelt, E. D. Williams, W. Swiech, E. Bauer, *Surf. Sci.* **1992**, *268*, 227; (b) R. J. Phaneuf, N. C. Bartelt, E. D. Williams, W. Swiech, E. Bauer, *Phys. Rev. Lett.* **1993**, *71*, 2284.
- [27] W. Swiech, E. Bauer, M. Mundschaue, *Surf. Sci.* **1991**, *253*, 283.
- [28] E. Bauer, *Appl. Surf. Sci.* **1992**, *60/61*, 350.
- [29] M. Mundschaue, E. Bauer, W. Swiech, *Catalysis Lett.* **1988**, *1*, 405.
- [30] M. Mundschaue, E. Bauer, W. Teliieps, W. Swiech, *Surf. Sci.* **1989**, *213*, 38130.
- [31] B. Rausenberger, W. Swiech, W. Engel, A. M. Bradshaw, E. Zeitler, *Surf. Sci.* **1993**, *287/288*, 235.
- [32] W. Swiech, B. Rausenberger, W. Engel, A. M. Bradshaw, E. Zeitler, *Surf. Sci.* **1993**, *294*, 297.
- [33] S. Jakubith, H. H. Rotermund, W. Engel, A. von Oertzen, G. Ertl, *Phys. Rev. Lett.* **1990**, *65*, 3013.
- [34] M. S. Altman, E. Bauer, *Surf. Sci.* **1995**, *347*, 265.
- [35] E. Bauer, Y. Wei, T. Müller, A. Pavlovskaa, I. S. T. Tsong, *Phys. Rev. B* **1995**, *51*, 17891.
- [36] W. Swiech, E. Bauer, *Surf. Sci.* **1991**, *255*, 219.
- [37] (a) M. Mundschaue, E. Bauer, W. Swiech, *J. Appl. Phys.* **1989**, *65*, 581; (b) M. Mundschaue, E. Bauer, W. Swiech, *J. Appl. Phys.* **1989**, *65*, 4747.
- [38] I. Müllerova, M. Lenz, *Ultramicroscopy* **1992**, *41*, 399.
- [39] I. Müllerova, L. Frank, *Scanning* **1993**, *15*, 193.
- [40] L. Li, C. Koziol, K. Wurm, Y. Hong, E. Bauer, I. S. T. Tsong, *Phys. Rev.* **1994**, *50*, 10834.
- [41] T. Michely, M. C. Reuter, M. Copel, R. M. Tromp, *Phys. Rev. Lett.* **1994**, *73*, 2095.
- [42] R. M. Tromp, T. Michely, *Nature* **1995**, *373*, 499.
- [43] M. S. Altman, E. Bauer, *Surf. Sci.* **1995**, *344*, 51.

# 1.7 Lorentz Microscopy

## 1.7.1 Magnetic Domains

Specimens of ferromagnetic materials are usually not magnetized uniformly throughout their volume, but are subdivided into domains. Within each domain, the magnetization is uniform, but different domains are magnetized in different directions. The domains are separated by boundaries called *domain walls*. Within a domain wall separating two domains, the magnetization direction gradually changes away from that in one domain towards that in the other.

The subdivision into domains is a result of the requirement for the free energy of the solid to be a minimum. The shape, size and arrangement of the domains, as well as the width of the domain walls, are determined by the four magnetic contributions to the total energy: the exchange, magnetostatic, anisotropy and magnetoelastic energies. When an external magnetic field is applied, the balance of the energy terms changes, resulting in a rearrangement of the domains. This rearrangement can take place by the displacement of domain walls (wall motion) or by rotation of the magnetization of the domains. Since these processes are to some extent irreversible, the change of average magnetization of the specimen as a function of the applied

magnetic field is subject to hysteresis. The reader is referred to standard textbooks on magnetism for details. However, it is clear that the magnetic domain structure plays an important role in determining the magnetic properties, and it is therefore important to study the magnetic domain structure.

Many techniques have been developed for imaging magnetic domains. These techniques can be divided into the following categories:

- (1) Techniques using small magnetic particles.
- (2) Polarized light microscopy.
- (3) Electron optical techniques.
- (4) X-ray techniques.
- (5) Neutron beam techniques.
- (6) Scanning probe techniques.

None of the techniques is able to image the magnetic domains throughout the volume of a bulk specimen. Some techniques provide images of domains on the surfaces of specimens, and others can image domains in thin films. Among the electron optical techniques there are both types. In this article, we are concerned with the use of conventional transmission electron microscopes for domain imaging. The relevant techniques, which are collectively described as Lorentz microscopy, fall into the category of thin film techniques.

### 1.7.2 The Effect of a Magnetic Specimen on Electrons

Electrons carry an electric charge  $e$ , and when an electron passes through a region in which a magnetic induction  $\mathbf{B}$  is present, it experiences a force. Classically, this force can be expressed as

$$\mathbf{F} = -|e|\mathbf{v} \times \mathbf{B} \quad (1)$$

provided no electric field is present, where  $\mathbf{v}$  is the velocity of the electron. If an electron travels for a distance  $t$  through a region in which a magnetic induction  $\mathbf{B}$  perpendicular to  $\mathbf{v}$  is present, it will be deflected by an angle  $\beta$  given by

$$\beta = \frac{eBt}{mv} \quad (2)$$

where  $m$  is the mass of the electron. (Note that  $B$  and  $v$  are shorter forms of  $|\mathbf{B}|$  and  $|\mathbf{v}|$ , respectively.) This equation assumes that  $\beta$  is small, which is always true in the cases we are concerned with. Another form of Eq. (2) is

$$\beta = \frac{eB\lambda t}{h} \quad (3)$$

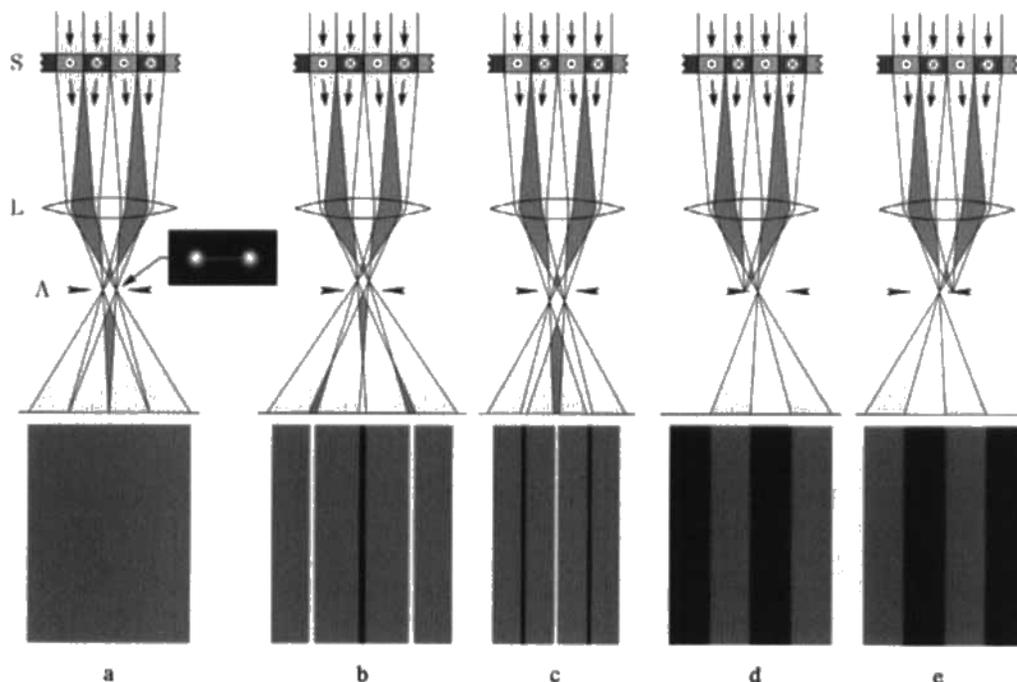
where  $\lambda$  is the electron wavelength and  $h$  is Planck's constant. Equation (3) is in a convenient form even for very fast electrons, since relativistic effects can be taken into account by using the correct expression for  $\lambda$ .

An upper limit to  $\beta$  is set by the fact that  $B$  is limited by the properties of existing materials, and  $t$  is limited by the penetration of electrons. In practice,  $\beta$  is never more than about  $10^{-4}$  rad. It is important to note that this value is about two orders of magnitude smaller than the Bragg diffraction angles from low-index atomic planes in typical metals. The presence of

magnetic domains therefore only causes a small perturbation to the electron diffraction pattern due to the crystalline nature of specimens.

### 1.7.3 Methods of Observing Magnetic Domains in the TEM

Since the effect of the magnetic induction is to deflect the electrons, the intensity of an electron beam passing through a magnetic specimen is unaffected by the presence of magnetic domains. The presence of the magnetic induction only changes the direction in which the electrons emerge from the specimen, but not the number or even the speed of the electrons. A perfect, in-focus image of the specimen therefore contains no information about the domain structure. There are, however, three ways in which the presence of domains can be observed. The basic principles are all illustrated in Fig. 1. Figure 1a shows how the featureless in-focus image is formed. The top part of the figure shows the electron ray diagram, together with a cross-section of the specimen, the objective lens and the objective aperture. The specimen shown has domains magnetized perpendicular to the incident beam, alternately out of, and into, the plane of the paper (illustrated by the arrow heads and tails, respectively). The electrons are incident normally on the specimen, illuminating it uniformly. After they have passed through the specimen, the electrons are deflected either to the right or to the left, depending on the domain magnetization direction, and after passing through the objective lens they form an image projected onto the screen.



**Figure 1.** Schematic diagrams illustrating the principle of Lorentz microscopy. The upper part of each diagram shows the electron ray paths, and the lower part shows the type of image that would be obtained in different imaging conditions. The specimen *S* is illuminated by a parallel beam of electrons incident from above, travelling in the direction of the arrows. *L* is the objective lens, *A* the objective aperture, and *I* the image plane. In (a) the image of the specimen at *I* is in focus. The inset shows the appearance of the central diffraction spot. The objective lens is overfocused in (b) and underfocused in (c), and the objective aperture is displaced in opposite directions in (d) and (e).

However, the screen is uniformly illuminated, as illustrated schematically at the bottom of the figure. (Note that in a real electron microscope, there would be two or more additional lenses producing a highly magnified final image. There would not be a screen in the position indicated, but the image plane of the objective lens would be the object plane of the next lens below.)

Although the in-focus image shows no evidence of the presence of domains in the specimen, the domains have an effect on the diffraction pattern. Figure 1a includes a simulation of the central part

of the diffraction pattern, in which the undiffracted spot (the spot formed by the part of the electron beam unaffected by the crystal structure of the specimen) has split into two subspots. These subspots correspond to the two magnetization directions present in the specimen.

In Fig. 1b the objective lens is overfocused, so that the image plane is slightly above the plane *I*. In the image shown in the bottom part of the figure, we now observe a dark line across the center, and two bright lines on either side. The ray diagram shows that these lines appear in

positions where the number of incident electrons has increased or decreased, respectively, as a result of the defocusing. The lines are in positions where the domain walls are imaged. The contrast reverses when the objective lens is underfocused, as shown in Fig. 1c. The image plane is now below the plane I, and in the plane I we observe a bright line in the center and two dark lines on either side. (Defocusing the objective lens alters the magnification of the image because of the change in the lens current, and not because of the magnetic effects of the specimen. These changes of magnification are somewhat exaggerated in Figs. 1b and 1c.)

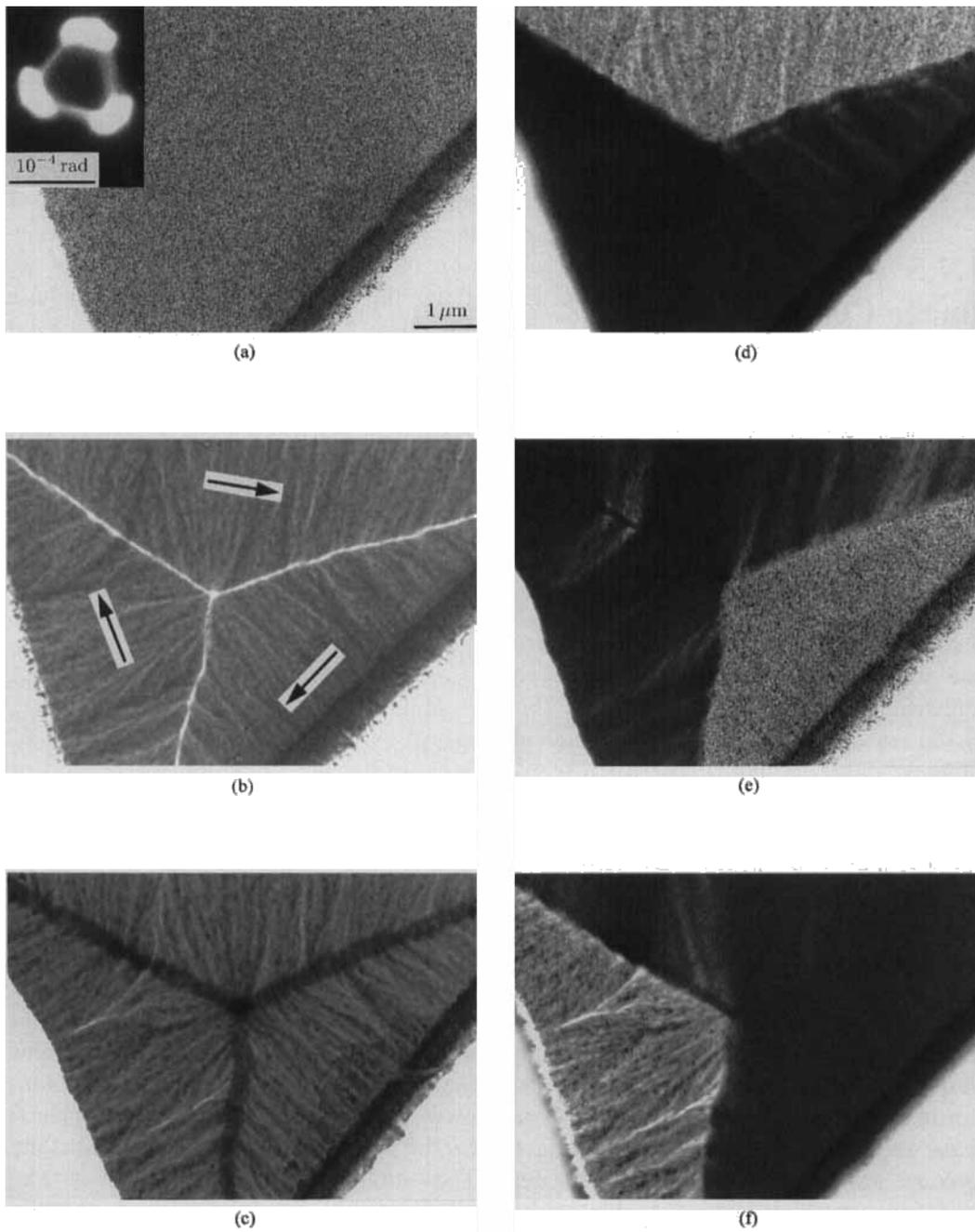
A second method of imaging the domain structure is to displace the objective aperture from its position centered on the optic axis of the microscope, as shown in Figs. 1d and 1e. In Fig. 1d the aperture has been moved to the right, and it now blocks out electrons that have passed through domains magnetized into the paper (marked with crosses). The parts of the image corresponding to these domains now appear dark. In Fig. 1e the aperture has been moved to the left, blocking out electrons that have passed through domains magnetized out of the paper (marked with dots).

The defocus method of imaging domains is often called the *Fresnel method* (for reasons explained in Sec. 1.7.5 in this Chapter), and the displaced aperture method is usually called the *Foucault method*, since it is analogous to a method of testing the quality of optical components developed by Foucault in the 19th century. Use of the Fresnel method of Lorentz microscopy was first reported by Hale, Fuller and coworkers [1,2], and the first

reports of the use of the Foucault technique were published soon after [3,4]. Reviews of Lorentz microscopy are given by Jakubovics [5] and Chapman [6].

### 1.7.4 Examples of Domain Images

Figure 2 shows transmission electron micrographs of a Co/Cu multilayer film, consisting of 40 cobalt layers of thickness 2 nm interleaved with copper layers of thickness 0.5 nm. Figure 2a is an in-focus image, which shows no evidence of a magnetic domain structure. The fine-scale contrast is due to the grain structure of the film. The inset shows the central part of the diffraction pattern, which is split into three main spots, indicating that there are three different magnetization directions in this region of the specimen. Figures 2b and 2c show images taken with the objective lens overfocused and underfocused, respectively. In Fig. 2b, three bright domain walls are visible, which become dark in Fig. 2c. The magnetization direction (shown by arrows in Fig. 2b) in the three domains can be determined on the assumption that the walls do not carry any magnetic charges (i.e., there is no change in the component of magnetization perpendicular to the wall across any of the walls). The magnetization directions are confirmed by the magnetization ripple (see Sec. 1.7.7). The presence of the ripple causes the three spots in the diffraction pattern to become slightly elongated. Figures 2d–2f show images obtained by displacing the objective aperture so that in each case only



**Figure 2.** Micrographs of a  $(\text{Co}_{2\text{nm}}/\text{Cu}_{0.5\text{nm}})_{40}$  multilayer film: (a) in focus; (b,c) with the objective lens overfocused and underfocused, respectively; (d–f) with the objective aperture displaced so that in each case, only one of the three spots in the central part of the diffraction pattern (shown as an inset in (a)) is inside the aperture. The micrographs were taken by J. D. Kim.

one of the three subspots seen in the inset in Fig. 2a is inside the aperture. In each of these Foucault images, one of the three domains is bright and the other two are dark.

### 1.7.5 Theory of Image Contrast

The first attempts to explain the contrast at domain walls in defocused images [2] were based on geometrical optics. It was also demonstrated that the wave nature of the electrons can have a strong effect on the image [7–9]: if the incident illumination is sufficiently coherent, a set of interference fringes appears in bright domain wall images. This effect is analogous to the interference fringes produced by a Fresnel biprism in light optics, and this is the reason why the defocus method of imaging domains in Lorentz microscopy is called the ‘Fresnel method’. Subsequently, it was shown that the wave theory was generally applicable to Lorentz microscopy, and that the intensity distribution in both Fresnel and Foucault images could be calculated from the Abbe theory of microscopy [10]. It was also shown [11] that in certain circumstances, the geometrical theory could be used to obtain reasonable approximations to the results of the wave theory. These approximations have become less useful, because fast Fourier transform methods, which enable image intensities to be calculated rapidly from the wave theory, have become easily available. We therefore give a brief outline of the wave theory.

The deflection of electrons, Eq. (2), caused by the classical Lorentz force, is equivalent to a phase change. In the following discussion, we assume that the surfaces of the specimen lie in the  $xy$  plane, the incident electrons are travelling along the  $-z$  direction, and the  $y$  component of the magnetic induction,  $B_y$ , is only a function of  $x$ . We denote the integral of  $B_y$  along the  $z$  direction by  $B_i$ . The phase difference between the electron wave at  $x = 0$  and  $x = X$  is

$$\phi(X) = -\frac{|e|\hbar}{\hbar} \int_0^X B_i(x) dx \quad (4)$$

where  $\hbar = h/2\pi$ . If the magnetic induction is confined to the volume of the specimen,  $t$  can be taken to be the specimen thickness. According to the Abbe theory, the amplitude in the back focal plane of the objective lens (diffraction plane) is the Fourier transform of the transmission function of the specimen, and it can be expressed as

$$A_d(\theta) = \int_{-\infty}^{\infty} \times \exp\left\{i\left[\phi(x) - \frac{2\pi(\theta + \theta_i)x}{\lambda}\right]\right\} dx \quad (5)$$

where  $\theta$  is the angular variable in the diffraction plane parallel to  $x$ , assuming that the specimen is illuminated by a plane wave of electrons arriving at an angle  $\theta_i$  to the normal to the specimen surface. The amplitude of the diffraction pattern is, in general, modulated by a complex function  $f(\theta)$ , which includes the effects of defocus, finite objective aperture size and lens aberrations. The amplitude  $A_i(\xi)$  in the image is the Fourier transform

of  $A_d(\theta)f(\theta)$ :

$$A_i(\xi) = \int_{-\infty}^{\infty} A_d(\theta)f(\theta) \times \exp\left(-\frac{2i\pi\theta\xi}{\lambda}\right) d\theta \quad (6)$$

where  $\xi$  is a coordinate in the image plane parallel to  $x$ . (Note that, since the limits of the integral are infinite, we should really use  $\tan \theta$  in the integrand. However, as any appreciable contribution to the amplitude comes from small  $\theta$ , we can neglect the difference between  $\theta$  and  $\tan \theta$ .) The image intensity is then given by

$$I_i(\xi) = |A_i(\xi)|^2 \quad (7)$$

The effect of a finite source size can be taken into account by performing a convolution with the intensity of illumination arriving in different directions.

The imaging modes in Lorentz microscopy can be described by using the general function

$$f(\theta) = a(\theta) \exp\left[-\frac{i\pi}{\lambda}\left(\frac{1}{2}C_s\theta^4 + \zeta\theta^2\right)\right] \quad (8)$$

where  $a(\theta)$  is the aperture function,  $C_s$  is the spherical aberration coefficient and  $\zeta$  is the defocus, with  $\zeta > 0$  corresponding to overfocus. In the one-dimensional case so far considered, we take  $a(\theta) = 1$  for  $\theta_1 \leq \theta \leq \theta_2$  and  $a(\theta) = 0$  elsewhere, where  $\theta_1$  and  $\theta_2$  correspond to the edges of the aperture. The two special cases of interest are: (i) the Fresnel case, in which we neglect the aperture, that is we take  $a(\theta) = 1$  for all  $\theta$ ; and (ii) the Foucault case, in which  $\zeta = 0$  and either  $\theta_1$  or  $\theta_2$  is finite. The effects of spherical aberration are usually negligible, because of the small values of the Lorentz deflection.

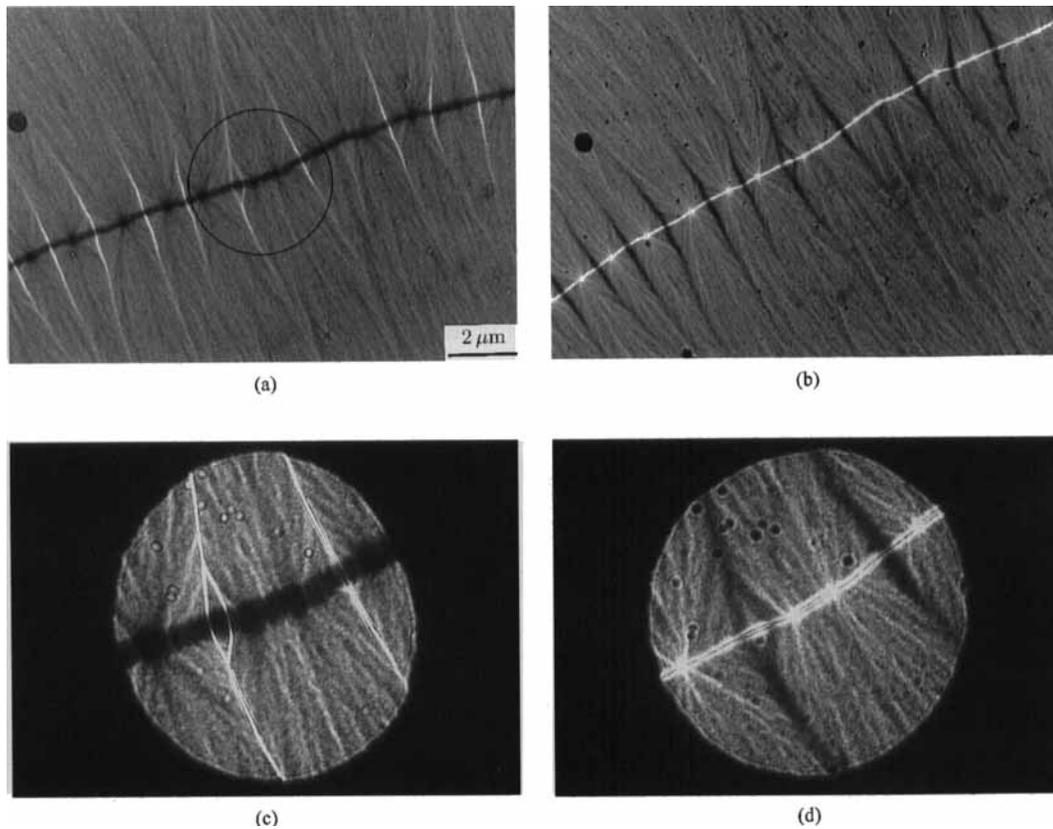
The above equations can easily be extended to two dimensions, that is to the case where the magnetic induction varies in both the  $x$  and  $y$  directions [6].

### 1.7.6 Interference Effects

Figure 3 shows an area containing a cross-tie wall [12, 13] in a polycrystalline cobalt film of thickness 67 nm. Figures 3a and 3b show two images of the same area taken with the objective lens overfocused and underfocused, respectively. Figures 3c and 3d show the part of the area circled in Fig. 3a, taken with a more parallel incident beam, that is with a more coherent illumination. It is seen that the bright domain walls in each image now break up into interference fringes.

### 1.7.7 Determination of Magnetization Distributions

In general, there is no simple relationship between the intensity distribution in Fresnel or Foucault images and the magnetization distribution or even the magnetic induction distribution in the specimen. In a specimen containing uniformly magnetized domains separated by domain walls that are narrow compared with the size of the domains, the Fresnel method can be used to find the position of the domain walls, and the Foucault method can be used to find the direction of magnetization of the domains. In cases where the domains are not uniformly magnetized but contain magnetization ripple [2], the



**Figure 3.** Micrographs of a polycrystalline cobalt film of thickness 67 nm. The objective lens is overfocused in (a) and underfocused in (b). Micrographs of the area circled in (a), taken with highly coherent illumination and with the objective lens defocused in opposite directions, are shown in (c) and (d). The micrographs were taken by J. D. Kim.

magnetization direction of the domains can be determined from either Fresnel or Foucault images, since the ripple is perpendicular to the magnetization (Fig. 2). However, details of the magnetization distribution, such as the structure of domain walls, cannot be deduced directly from Lorentz micrographs. At best, it is possible to determine domain wall widths by matching the observed intensity distribution to a series of computed profiles (e.g., [14]), or to detect the well-known asymmetry [15, 16] of some domain walls in thin films [17–19]. Magnetization distributions

can be mapped quantitatively by using a suitably modified scanning transmission electron microscope (TEM) [20], but this topic is outside the scope of the present article. It has, however, been shown recently [21] that an equivalent mapping method can be implemented in a conventional TEM by combining series of Foucault images. Such methods are becoming feasible with the use of digital image-acquisition facilities and computer control of electron microscopes. Other new methods for mapping magnetization distributions by transmission electron

microscopy have also been developed recently [22].

### 1.7.8 Practical Implementation of Lorentz Microscopy

Magnetic lenses used in TEMs produce magnetic fields approximately parallel to the axis of the microscope, which are strong enough to saturate most specimens. In order to study magnetic domains, it is therefore necessary to reduce the magnetic field acting on the specimen. In some cases, acceptable results can be achieved by switching off the objective lens and using the remaining lenses to produce an image, but the magnification and resolution available are usually insufficient. In TEMs with top-entry specimen stages, it is sometimes possible to achieve a sufficient reduction of the magnetic field by moving the specimen away from the objective lens and increasing the focal length of the lens. In TEMs with side-entry specimen stages, it is not possible to alter the specimen position. However, specially designed objective pole-pieces are available for some TEMs, which concentrate the focusing fields well below the specimen, and thus enable any standard side-entry specimen holder to be used [22, 23]. These special pole-pieces can achieve high magnification and resolution.

Reduction of the magnetic field acting on the specimen is the only modification to a standard TEM needed for Fresnel imaging. For Foucault imaging, there is the additional difficulty that if the focal length of the objective lens is increased, the objective aperture is no longer in the back-focal plane. When the objective aperture is

displaced, its edge casts a fairly narrow shadow over the screen, and the conditions to obtain Foucault contrast are only satisfied within this shadow. To overcome this problem, it is necessary to move the plane of the aperture away from the lens. Doole et al. [24] give an example of a successful solution to this problem.

It is sometimes useful to combine Lorentz imaging with in situ experiments. The use of side-entry specimen holders makes such experiments particularly convenient, since this arrangement enables all standard specimen holders, such as heating and cooling holders, to be used. The study of magnetic specimens often requires the application of controlled magnetic fields, in order to study the motion of domain walls. However, applied magnetic fields affect not only the specimen, but also the electron beam. A field applied in the plane of the specimen deflects the beam, since the field acts as an electron optical prism. Although it is possible to compensate for this deflection by applying magnetic fields in the opposite direction away from the specimen, the fields also introduce aberrations, which ultimately limit the magnitude of the field that can be applied. For a successful design of a set of in situ magnetizing coils for use with a top-entry magnetizing stage, see Taylor [25]. If it is required to apply magnetic fields perpendicular to the specimen, the field of the objective lens can sometimes be used, the specimen being moved along the axis of the microscope [26].

### Acknowledgements

The author is grateful to J. D. Kim for the micrographs shown in Figs. 2 and 3, and to A. K. Petford-Long for helpful comments on the manuscript.

## 1.7.9 References

- [1] M. E. Hale, H. W. Fuller, H. Rubinstein, *J. Appl. Phys.* **1959**, *30*, 789.
- [2] H. W. Fuller, M. E. Hale, *J. Appl. Phys.* **1960**, *31*, 238.
- [3] H. Boersch, H. Raith, *Naturwissenschaften* **1960**, *46*, 576.
- [4] H. W. Fuller, M. E. Hale, *J. Appl. Phys.* **1960**, *31*, 1699.
- [5] J. P. Jakubovics, in *Electron Microscopy in Materials Science, Third Course of the International School of Electron Microscopy* (Eds.: E. Ruedl, U. Valdrè), Commission of the European Communities **1975**, 1303.
- [6] J. N. Chapman, *J. Phys. D: Appl. Phys.* **1984**, *17*, 623.
- [7] H. Boersch, H. Hämisch, D. Wohlleben, K. Grohmann, *Z. Phys.* **1960**, *159*, 397.
- [8] H. Boersch, H. Hämisch, D. Wohlleben, K. Grohmann, *Z. Phys.* **1961**, *164*, 55.
- [9] H. Boersch, H. Hämisch, K. Grohmann, D. Wohlleben, *Z. Phys.* **1960**, *167*, 72.
- [10] D. Wohlleben, *J. Appl. Phys.* **1967**, *38*, 3341.
- [11] J. P. Guigay, R. H. Wade, *Phys. Stat. Sol.* **1968**, *29*, 799.
- [12] E. E. Huber, D. O. Smith, J. B. Goodenough, *J. Appl. Phys.* **1958**, *29*, 294.
- [13] S. Methfessel, S. Middelhoek, H. Thomas, *IBM J. Res. Dev.* **1960**, *4*, 96.
- [14] D. C. Hothersall, *Phil. Mag.* **1969**, *20*, 89.
- [15] A. E. LaBonte, *J. Appl. Phys.* **1969**, *40*, 2450.
- [16] A. Hubert, *Phys. Stat. Sol.* **1969**, *321*, 519.
- [17] D. C. Hothersall, *Phys. Stat. Sol. B* **1972**, *51*, 529.
- [18] C. G. Harrison, K. D. Leaver, *Phys. Stat. Sol. A* **1972**, *12*, 413.
- [19] C. G. Harrison, K. D. Leaver, *Phys. Stat. Sol. A* **1973**, *15*, 415.
- [20] J. N. Chapman, R. Ploessl, D. M. Donnet, *Ultramicroscopy* **1992**, *47*, 331.
- [21] A. C. Daykin, A. K. Petford-Long, *Ultramicroscopy* **1995**, *58*, 365.
- [22] J. N. Chapman, A. B. Johnston, L. J. Heyderman, S. McVitie, W. A. P. Nicholson, B. Bormans, *IEEE Trans. Magn.* **1994**, *30*, 4479.
- [23] K. Tsuno, M. Inoue, K. Ueno, *Mat. Sci. Eng. B* **1989**, *3*, 403.
- [24] R. C. Doole, A. K. Petford-Long, J. P. Jakubovics, *Rev. Sci. Instrum.* **1993**, *64*, 1038.
- [25] R. A. Taylor, in *Electron Microscopy 1980*, Vol. 4, Proc. 6th Int. Conf. on High Voltage Electron Microscopy, Antwerp (Eds.: P. Brederoo, J. Van Landuyt), Seventh European Congress on Electron Microscopy Foundation, Leiden **1980**, 38.
- [26] E. L. Houseman, J. P. Jakubovics, *J. Magn. Mater.* **1983**, *31*, 1007.

# 1.8 Electron Holography Methods

## 1.8.1 Principle and Problems of Conventional Transmission Electron Microscopy

### 1.8.1.1 Modulation of the Electron Wave

As early as 1932, Glaser [1] showed that, for electrons moving with an energy  $eU$  in a medium with an electrostatic potential field  $V_{el}(\mathbf{r})$  and a magnetic potential field  $\mathbf{A}(\mathbf{r})$ , the index of refraction is

$$n\left(\mathbf{r}, \frac{d\mathbf{r}}{ds}\right) = \sqrt{\frac{[U + V_{el}(\mathbf{r})]^*}{U^*} - \sqrt{\frac{e}{2m_0U}} \mathbf{A}(\mathbf{r}) \frac{d\mathbf{r}}{ds}} \quad (1)$$

Evidently, the electric contribution depends only on the position  $\mathbf{r}$ , whereas the magnetic potential contribution depends on both  $\mathbf{r}$  and the direction  $d\mathbf{r}/ds$  of the electron path. Comparing two paths  $s_1$  and  $s_2$  through a medium, an electron wave undergoes a phase shift

$$\phi = \frac{2\pi}{\lambda_0} \left\{ \int_{s_2} n ds - \int_{s_1} n ds \right\} \quad (2)$$

*Electric Potential,  $V_{el}(\mathbf{r})$*

In the case of the electric potential

distribution  $V_{el}(\mathbf{r})$  representing an object, the phase shift with respect to vacuum is given by

$$\phi = \frac{2\pi}{\lambda_0} \int_{\text{obj}} (n - 1) ds \quad (3)$$

where the integral has to be taken along the electron path  $s$ .

Since  $V \ll U$ , this can be written as

$$\phi = \frac{2\pi}{\lambda_0} \frac{1}{2U^*} \frac{m}{m_0} \int_{\text{obj}} V(\mathbf{r}) dz \quad (4)$$

or

$$\phi = \sigma V_p \quad (5)$$

with the interaction parameter

$$\sigma = \frac{2\pi\lambda_0 em}{h^2} \quad (6)$$

and the projected potential

$$V_p(x, y) = \int_{\text{obj}} V_{el}(\mathbf{r}) dz \quad (7)$$

The projected potential may represent a single atom, an amorphous thin film, or a crystal that is thin compared to the extinction distance. In thicker crystals, the effects of dynamic interaction have to be taken into account, which need, for example, multislice calculations instead of the simple potential projection.

*Magnetic Potential, A(r)*

Integration of

$$\phi = \frac{2\pi}{\lambda_0} \sqrt{\frac{e}{2m_0 U^*}} \times \left[ \int_{s_2} \mathbf{A}(\mathbf{r}) \frac{d\mathbf{r}}{ds} ds - \int_{s_1} \mathbf{A}(\mathbf{r}) \frac{d\mathbf{r}}{ds} ds \right] \quad (8)$$

with  $\mathbf{B} = \text{rot } \mathbf{A}$  and the magnetic flux  $\Phi_m = \oint \mathbf{A} d\mathbf{r}$  enclosed in the closed path  $s$  yields

$$\phi = \frac{e}{\hbar} \Phi_m \quad (9)$$

This phase shift occurs even if the electron path does not lead through an area with a magnetic field  $\mathbf{B}$ , that is even if no Lorentz force is acting on the electrons. This *Aharonov-Bohm effect* was discovered theoretically by Ehrenberg and Siday [2] and by Aharonov and Bohm [3].

An electron wave transmitted through an object undergoes a phase shift  $\phi(x, y)$  according to the electric or magnetic potential distribution. In addition, the amplitude  $a(x, y)$  may be modulated due to the following causes: scattering into angles larger than the collection angle of the subsequent optics, interference effects at the exit face of the object and to inelastic scattering (inelastically scattered electrons are no longer coherent with the elastically scattered ones).

In conclusion, the electron wave at the exit face of an object is given by the 'object function'

$$o(x, y) = a(x, y) \cdot \exp[i\phi(x, y)] \quad (10)$$

### 1.8.1.2 Propagation of the Electron Wave through the Electron Microscope

The electron wave leaving the object propagates in space according to the wave

equation, thereby changing its appearance in amplitude and phase in a rather complicated manner. Nevertheless, once the wave is known completely over a sufficiently wide surface it can be determined at every point in space by means of the Kirchhoff diffraction integral (KDI). Most familiar special cases frequently used in optics are the Fresnel approximation of the KDI delivering the Fresnel diffraction pattern, or, more importantly, the far field approximation of the KDI which allows one to determine the wave at the Fraunhofer diffraction condition by means of a Fourier transform.

The object wave propagates through the electron microscope according to the wave equation. Therefore, taking into account the properties of the imaging system, one can track the wave from the object to the image plane. In general, the image wave may be written as

$$b(x, y) = \int o(\tilde{x}, \tilde{y}) G(\tilde{x}, \tilde{y}, x, y) d\tilde{x} d\tilde{y} \quad (11)$$

Usually, the imaging process is considered isoplanar, that is  $G(\tilde{x}, \tilde{y}, x, y)$  depends only on the differences  $(\tilde{x} - x)$  and  $(\tilde{y} - y)$ , and hence the image wave is found as a convolution of the object wave with the 'point spread function' (PSF)  $(x, y) = G(0, 0, x, y)$

$$b(x, y) = o(x, y) \otimes \text{PSF}(x, y) \quad (12)$$

In Fourier space this reads

$$\text{FT}(b) = \text{FT}(o) \cdot \text{WTF}(\mathbf{R}) \quad (13)$$

$\text{WTF}(\mathbf{R}) = \text{FT}^{-1}(\text{PSF})$  is the wave transfer function, which depends on the spatial frequency vector  $\mathbf{R} = (R_x, R_y)$ . With an aberration-free imaging system,  $\text{WTF}(\mathbf{R}) \equiv 1$  holds, and hence the PSF would be a delta function (i.e., an

arbitrarily sharp point). Then, apart from a magnification factor, the image wave and object wave would match perfectly in both amplitude and phase.

However, electron lenses cannot be made free from aberrations; most severe are the spherical aberrations, which are unavoidable with the usual electron lenses [4], and chromatic aberration. In particular at increasingly high resolution there are a variety of additional aberrations to be considered, which can be classed as either coherent or incoherent aberrations.

*Coherent* aberrations are those aberrations which may be present under perfectly coherent illumination. These are spherical aberrations, defocus, astigmatism and axial coma. They enter the imaging process by means of a wave aberration  $\chi(\mathbf{R})$  which describes the resulting deviation of the wavefront in Fourier space from the ideal one. Coherent aberrations act like a phase plate in Fourier space. They do not destroy information. To understand their effect, one must distinguish between aberrations which are symmetrical and those which are antisymmetrical in  $\mathbf{R}$ . Those *symmetrical* in  $\mathbf{R}$  (i.e.,  $\chi_s(\mathbf{R}) = \chi_s(-\mathbf{R})$ ), such as spherical aberrations, defocus and two-fold astigmatism, act like a Zernike phase plate with a thickness which is variable in  $\mathbf{R}$ . Consequently, symmetric aberrations exchange amplitudes and phases in the image wave with respect to the object wave according to the  $\sin[\chi_s(\mathbf{R})]$  and  $\cos[\chi_s(\mathbf{R})]$  functions. The aberrations *antisymmetrical* in  $\mathbf{R}$  (i.e.,  $\chi_a(\mathbf{R}) = -\chi_a(-\mathbf{R})$ ), such as three-fold astigmatism and axial coma, produce a lateral shift  $\chi_a(\mathbf{R})$  of the elementary cosinoidal gratings associated in real space with each spatial frequency; this may severely degrade the image wave.

*Incoherent* aberrations always show up due to the restricted coherence of electron illumination. Due to the illumination aperture  $\theta_c$  (i.e., lateral coherence) and the energy spread  $\Delta E$  (i.e., longitudinal coherence) the envelope functions  $E_{\text{coh}}(\mathbf{R})$  and  $E_{\text{chrom}}(\mathbf{R})$  dampen the Fourier spectrum of the image wave. Incoherent aberrations act like an aperture in Fourier space cutting away the contributions of the higher spatial frequencies, and hence destroy information in that they prevent object information from showing up in the image plane.

In summary, the wave transfer function is given by

$$\text{WTF}(\mathbf{R}) = E_{\text{coh}}(\mathbf{R}) \cdot E_{\text{chrom}}(\mathbf{R}) \cdot \exp[-i\chi(\mathbf{R})] \quad (14)$$

with

$$\chi(\mathbf{R}) = \chi_s(\mathbf{R}) + \chi_a(\mathbf{R}) \quad (15)$$

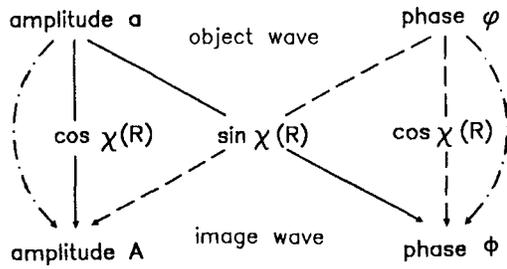
### 1.8.1.3 Problems in Conventional Transmission Electron Microscopy

A conventional electron image represents the intensity distribution given by

$$I(x, y) = b(x, y) \cdot b^*(x, y) \quad (16)$$

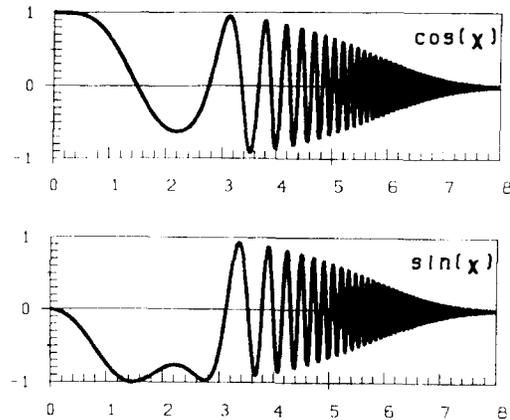
which is falsified with respect to the image wave  $b(x, y)$  in several respects.

$I(x, y)$  only represents the squared amplitude of the image wave, whereas the phase of the image wave is lost in a micrograph. Therefore the desirable object information should be directed into the amplitude of the image by means of a suitable selection of the wave aberration. Generally, however, the information found in the image amplitude is a mixture



**Figure 1.** Schematic transfer of the object wave into the image wave by the objective lens with a wave aberration  $\chi(R)$  symmetric in  $R$ . The Scherzer focus optimizes the  $\sin[\chi_s(R)]$  paths. Aberration-free imaging means  $\cos \chi \equiv 1$  at vanishing 'cross-talk'  $\sin \chi \equiv 0$ .

of both object amplitude and object phase transferred according to the transfer functions  $\cos[\chi_s(\mathbf{R})]$  and  $\sin[\chi_s(\mathbf{R})]$ , respectively (Fig. 1). Therefore the image intensity cannot be interpreted uniquely in terms of object amplitude or phase. Only if a weak phase object is assumed does  $I(x, y)$  represent solely the object phase—but distorted by the *phase contrast transfer function*  $\sin[\chi_s(\mathbf{R})]$  (Fig. 2). Under



**Figure 2.** Wave transfer function  $WTF(\mathbf{R}) = E_{\text{coh}}(\mathbf{R}) \cdot E_{\text{chrom}}(\mathbf{R}) \cdot \exp[-i\chi(\mathbf{R})]$  at Scherzer focus. The point resolution is defined by the first zero of the  $\sin \chi$  function. The information limit at about  $6 \text{ nm}^{-1}$  corresponds to a damping to  $\exp(-1)$ . (100 kV, energy width 0.5 eV,  $C_c = 1.2 \text{ mm}$ ,  $C_s = 1.2 \text{ mm}$ , illumination aperture 0.05 mrad.)

the assumption that only the spherical aberration (coefficient  $C_s$ ) and defocus  $Dz$  contribute to the wave aberration, the distortion can be minimized by selecting the Scherzer focus; then the function  $\sin[\chi_s(\mathbf{R})]$  has a shape such that it approaches unity over a broad range of spatial frequencies ranging from  $R_{\text{min}} = 0.385 / (C_s^{0.25} \lambda_0^{0.75})$  to  $R_{\text{Scherz}} = 1.5 / (C_s^{0.25} \lambda_0^{0.75})$  [5]. Therefore, large-area phase structures with  $R < R_{\text{min}}$  are not visible. Spatial frequency components with  $R > R_{\text{Scherz}}$  cannot be interpreted intuitively because of the contrast reversals and the zeros; therefore  $R_{\text{Scherz}}$  is called the *Scherzer point-resolution limit* of the electron microscope.

Nonlinearity also degrades the image information. Whereas, even under aberrated imaging conditions, the image wave is always linearly related to the object wave, this is not true in general for the image intensity. Due to the squaring of the amplitude, the intensity contains spatial frequencies which are not present in the object wave; the most common example of such frequencies are the half-spacings observed with strongly scattering crystals. They arise due to mutual interference of the reflections in addition to the regular interference of each reflection with the undiffracted beam. In addition, the inelastically scattered electrons contribute a background to the high-resolution image which is not included in the above description. The interpretation of such micrographs is severely impeded.

As the use of modern field emission guns greatly reduces the effect of the incoherent aberrations, an information limit of  $R_{\text{lim}} \approx 2R_{\text{Scherz}}$  is found. Of course, the point resolution  $R_{\text{Scherz}}$  should now be expanded to the information limit  $R_{\text{lim}}$  by

reducing  $C_s$ , which is difficult to achieve. Besides, the problem of large-area phase contrast is worsened because the Scherzer band is shifted to higher spatial frequencies. In a fully corrected imaging system, no phase contrast shows up at all. Therefore, aberration-free imaging is only desirable in combination with well-elaborated wave optical techniques for analysis of the image wave to extract all accessible information about the object structure. However, the wave optical facilities in the electron microscope are relatively poor in that the parameter accessible to the operator is defocus. It is this state of affairs that holography improves considerably.

## 1.8.2 Holography: Basic Principles

The above described problems in conventional electron microscopy arise mainly because the wave optical imaging process is irreversibly brought to an end by recording the image intensity and thus intrinsically losing the image phase. Since the image phase information is missing, an a posteriori correction (e.g., of focus or astigmatism) in a micrograph is absolutely impossible. In order to accomplish this, we have to consider the wave optics: if the complete image wave was available in terms of amplitude and phase then, again according to the laws of wave optics, it could be back-propagated through an optical system to recover the object wave again.

Gabor [6] extended this idea to the scheme of holography, showing how a wave could be recorded and then 're-created'

from the record: if one superimposes a coherent reference wave  $r$  on the image wave  $b$ , the resulting interference pattern represents a hologram given by the intensity distribution

$$I_{\text{hol}} = (r + b)(r + b)^* \quad (17)$$

Considering the hologram recorded on a photographic plate as a light optical diffraction grating, Gabor found one of the diffracted waves to be a copy wave of the recorded one:

Thus, if one illuminates the photographic record with an optical imitation of the electronic (reference) wave, only that part of the primary (image) wave will be strongly transmitted which imitates the modified wave both in phases and amplitudes. (Gabor, 1948 [6])

In fact, under illumination with the reference wave  $r$ , three waves leave the hologram:

$$r \cdot I_{\text{hol}} = r(|r|^2 + |b|^2) + r^2 b^* + |r|^2 b \quad (18)$$

where the last term is a copy of the recorded wave  $b$ .

Once the wave  $b$  is re-created in the plane where the hologram was taken, it can be propagated to any desirable plane (e.g., back to the level of the object). The point is that the recorded wave can be propagated by means of any medium obeying the wave equation; for example, the electron wave can be re-created and propagated as a light wave which, apart from a different wavelength, obeys the same wave equation. Thereby one can combine the advantages of electron microscopy (i.e., short electron wavelength allowing high resolution and strong interaction of the electrons at atomic

dimensions) with those offered by well-established light optics for highly sophisticated wave optical handling (e.g., with the Zernike phase contrast method) and an analysis of amplitude and phase.

During the subsequent years of development of light-optical holography it was shown that the wave optical analysis of the object wave can be performed on the reconstructed wave (e.g., phase contrast [7], holographic interferometry [8], phase enhancement [9], and a posteriori focusing and correction of aberrations by means of a corresponding lens [10]).

### 1.8.3 Electron Holography: Aspects of Realization

The general scheme for electron holography is outlined in Fig. 3. The electron wave found on the final screen of an electron microscope is recorded by a

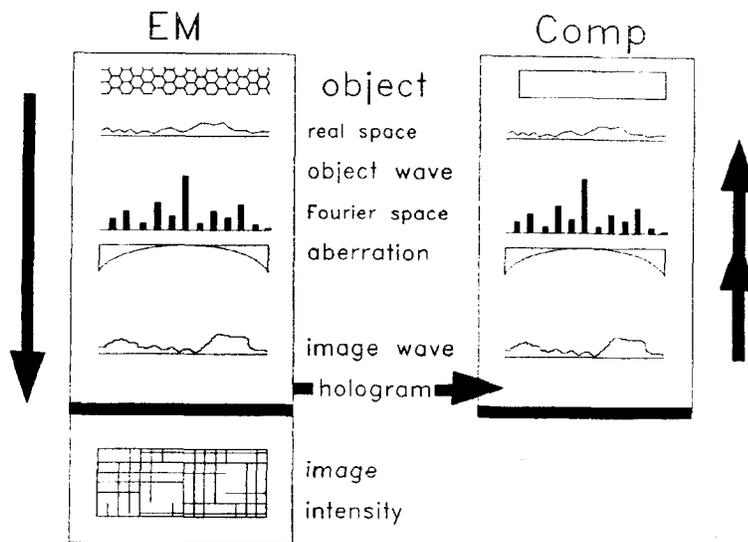
detector and transferred to the reconstruction device. There the wave is re-created and back-propagated through an optical system which is equivalent to the electron microscope. Finally, the wave can be analyzed at the stage of the object and of its Fourier spectrum.

The key is to take high quality holograms in the first instance. This allows one to re-create the captured wave to a high degree of fidelity for further processing. Therefore, the electron optical facilities for doing wave optics with electrons had to, and still have to, be developed and optimized for this purpose, and restrictions peculiar to electron optical technology have to be considered.

#### 1.8.3.1 Coherence

In light optics, holographic methods did not reach full success until the laser, as a perfectly coherent light source, became

**Figure 3.** Scheme of electron holography. In the electron microscope, instead of the intensity, the complex electron wave is recorded in a hologram and transferred to the reconstruction device, which may consist of a light optical bench or a computer. There the wave is re-created and back-propagated to the Fourier plane or the object plane for further wave optical analysis.



available, because a well-defined phase relation between the wavefield to be recorded and the reference wave is indispensable. Indeed, for electron holography, the key point is the coherence of electron radiation which, as electron sources are incoherent, can only be achieved by means of the usual coherence conditions.

Because of the chromatic aberration of electron lenses, temporal coherence is essential because it dampens the wave transfer function of the electron microscope. Fortunately, from the point of view of interferometry and holography, this poses virtually no additional restriction since the coherence length amounts to several 10 000 wavelengths [11]. Path-length differences in electron interference experiments, however, seldom exceed about 1000 wavelengths.

The main problem occurs with lateral coherence. Using incoherent sources, a plane at a large distance may be illuminated coherently so that, if two previously separate points  $P_1$  and  $P_2$  are superimposed, an interference contrast  $0 < V < 1$  is found; usually, the contrast decreases with increasing distance between  $P_1$  and  $P_2$ .  $V$  is the modulus of the degree of coherence which, according to the van Cittert–Zernike theorem, is distributed in the illuminated plane according to the Fourier transform of the intensity distribution of the source. In combination with the brightness  $B_0$  of the electron gun one finds that the total coherent current available for producing an interference pattern with contrast  $V$  is limited by [12]

$$I_{\text{coh}} = -B_0 \cdot \lambda_0^2 \ln(V) \quad (19)$$

As a property of the electron gun,  $I_{\text{coh}}$  is optically invariant along the microscope column.

Taking holograms at a sufficient contrast  $V$ , the current density is accordingly small. Therefore, to collect the electron dose needed for a low noise level, the exposure time is considerably higher than in conventional electron microscopy; this, in turn, is limited by residual instabilities, again reducing  $V$ . In any case, high brightness electron guns (e.g. field emission guns which have a brightness about 1000 times higher than thermionic guns) are indispensable.

### 1.8.3.2 Geometries for taking Electron Holograms

In the general formula describing a hologram

$$I_{\text{hol}} = (r + b)(r + b)^* \quad (20)$$

it is left open in which way the object-modulated wave  $b$  and the reference wave  $r$  are superimposed, and where the hologram is taken.

The first point to consider is whether the two waves are running essentially parallel (in-line holograms) or are superimposed at some angle (off-axis holograms). This determines how the reconstructed waves leaving the hologram

$$r \cdot I_{\text{hol}} = r(|r|^2 + |b|^2) + r^2 b^* + |r|^2 b \quad (21)$$

are mutually distributed in space. If  $r$  and  $b$  have been superimposed in parallel, then  $r^2 = |r|^2$  holds and the three reconstructed waves travel essentially in the same direction. Here one faces the famous ‘twin-image problem’:  $b$  and  $b^*$  overlap and cannot be separated, and hence disturb each other unless the field of view is very

small. Furthermore, certain restrictions apply in that the interaction with the object should be sufficiently small to ensure that the (undiffracted) reference wave, which is also partly transmitted through the object, is not modulated.

For these reasons, holography was brought a huge step forward by Leith and Upatnieks [13], who proposed off-axis holography, that is where  $r$  and  $b$  are superimposed at some angle. The reconstructed  $b$  and  $b^*$  waves are then separated from the remainder by the overlapping angle in opposite directions. One wave,  $b$  or  $b^*$ , can be cut out and back-propagated through the optical system.

The next point to consider is where the holograms should be recorded. In principle, the hologram can be recorded anywhere in the wavefield, for example in the image plane (image plane holograms), at some distance (Fresnel holograms), at a large distance (Fraunhofer holograms) or in the back focal plane of a lens (Fourier holograms). In any case, after corresponding back-propagation, the object wave should look the same, irrespective of where the hologram was taken. The main criterion for taking the hologram is given by the achievable resolution. The PSF containing all the information about one object point spreads into large areas, the area increasing with the distance to the object. If the PSF exceeds the coherent domain of the reference wave it is no longer recorded completely in the hologram, and hence the wave cannot be reconstructed at full resolution.

Reviewing the holographic arrangements realized in electron optics, in 1992 Cowley listed 20 forms of electron

holography [14]. Some of these are taken into account below.

### 1.8.3.3 In-Line Holography

Using the above described terminology, the first electron holography experiments (by Haine and Mulvey in 1952) were conducted as in-line Fresnel holography [15]. Their pioneering work verified the basic idea of electron holography; however, it suffered from the comparably poor electron optics technology, and the results revealed the problems of this technique (see above). Meanwhile, using a very coherent nanotip field emission source, Fink and coworkers [16, 17] have taken up this technique again, hoping to get rid of the twin-image by sophisticated numerical reconstruction; a detailed analysis of this method is given by Spence et al. [18].

In-line Fraunhofer holography was realized by Tonomura et al. [19], showing the value of this technique for holographic reconstruction in that the twin-image problem is avoided. More recent applications [20–22] show a resolution approaching 1 nm. However, this technique is restricted to weak and very small objects (e.g., to certain biological objects) [23].

Fourier holography has been realized experimentally [24], but has not yet been developed further. Coherent convergent beam electron diffraction can be considered a holographic technique in that the partly overlapping diffraction discs in the Fourier plane show interference patterns revealing the phases of the diffracted waves [25–28]. Methods for exploiting the available holographic information for

high-resolution structure analysis are under development.

### 1.8.3.4 Off-Axis Electron Holography

Off-axis techniques need a beam splitter to split an incoming wave into one going through the object and one which passes by, subsequently serving as a reference wave. Beam splitting can be achieved by amplitude division or wavefront division.

#### *Amplitude Division*

Amplitude division of electron waves is accomplished by diffraction at crystals, as applied in the electron interferometer set up by Marton et al. [29]. Matteucci et al. [30] used this principle for off-axis holography, which was recently experimentally developed further by Ru [31]. The main advantage over wavefront-division techniques is the generally lower requirement for lateral coherence.

#### *Wavefront Division*

The electron biprism developed by Möllenstedt and Düker [32] is the most successful device for electron interferometry; it uses the wavefront division principle.

A series of electron interference experiments has been conducted; for example, for:

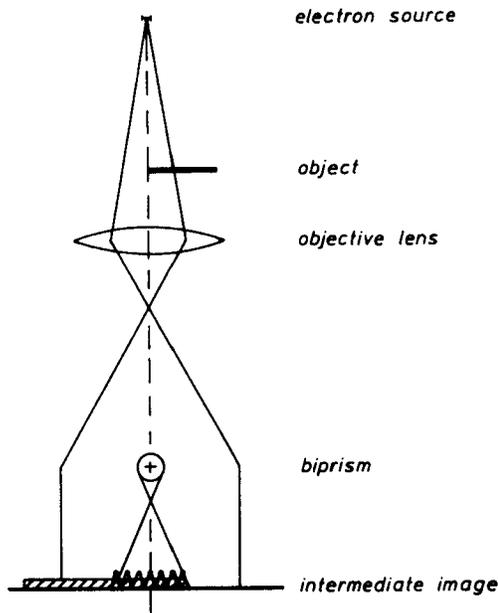
- (a) Measurement of the de-Broglie wavelength [32].
- (b) Electron interference microscopy [33, 34].
- (c) Phase shift by mean inner potential of thin films [35].

- (d) Experimental proof of the Aharonov–Bohm effect [36].
- (e) Flux quantization in superconducting hollow cylinders [37, 38].
- (f) Determination of contact potentials [39].
- (g) Study of lateral [40, 41] and longitudinal [11, 42] coherence.
- (h) Subatomic height differences and surface potentials [43].
- (i) Phase shift due to magnetization in thin films [44, 45].
- (j) Sagnac effect [46].

These experiments set the basis for electron holography in that they developed the deep understanding and thorough experience in experimental handling of electron wave optics. In fact, Möllenstedt and Wahl, highly experienced in electron interferometry, conducted the first experiment on off-axis electron holography using the electron biprism as a beam splitter (Fig. 4).

#### *Off-Axis Fresnel Holography*

Möllenstedt and Wahl [47], following the Gabor idea of lensless imaging, took an off-axis Fresnel hologram without any objective lens and reconstructed the image wave by means of a laser beam. Further progress was achieved by Tomita et al. [48] using a strongly defocused objective lens for taking the hologram; they also discussed the role of defocus for the achievable resolution. Saxon [49] performed the first experiments on a posteriori correction of aberrations. Recently, Mankos et al. [50] realized off-axis Fresnel holography in a scanning transmission electron microscope (STEM) for the investigation of magnetic microstructures. In general, however, off-axis Fresnel



**Figure 4.** Taking off-axis holograms with the electron biprism. The wave to be recorded and the empty reference wave passing the biprism filament on opposite sides are superimposed to form the hologram. Interference fringes arise which are shifted aside and modulated in contrast by the phase and amplitude of the recorded wave, respectively.

holography could not even roughly reach a lateral resolution comparable to conventional electron microscopy.

#### *Off-axis Image Plane Holography*

By means of a careful study of the role of coherence for the different geometries employable for holography, Weingärtner et al. [51] showed that the highest resolution can be obtained only from holograms taken at about zero defocus, that is in the image plane of the object. Consequently, according to Bryngdahl and Lohmann [52], the interference microscopy results obtained by Möllenstedt and Buhl [33] and Faget and Fert [34] were in fact image plane holograms, although they

had not been recognized as such. Interestingly, up to now image plane holography has been the most successful electron holographic technique.

## 1.8.4 Off-Axis Image Plane Holography

### 1.8.4.1 Principles

The procedure of image plane holography (i.e., taking the holograms and reconstructing the wave) was systematically developed and studied by Wahl [53]. The electron biprism inserted between the objective lens and the intermediate image plane superimposes a tilted empty reference wave of modulus 1 to the image wave

$$b(x, y) = A(x, y) \exp[i\Phi(x, y)] \quad (22)$$

In the image plane, a hologram

$$I_{\text{hol}}(x, y) = 1 + A^2(x, y) + 2VA(x, y) \times \cos[2\pi R_c x + \Phi(x, y)] \quad (23)$$

results. The image wave is riding on a carrier frequency  $R_c$  which has to meet the condition  $R_c \geq 3R_{\text{max}}$  where  $R_{\text{max}}$  is the maximum spatial frequency of the image wave contributing to the hologram, and  $V$  is the contrast of the hologram fringes. The hologram is recorded by means of an electron detector.

At the beginning of the procedure the reconstruction of the image wave was performed light optically. In the back-focal plane of a reconstruction lens, the Fourier transform of the hologram is

given by

$$\begin{aligned}
 \text{FT}[I_{\text{hol}}] &= \delta(R) + \text{FT}[A^2] && \text{Center band} \\
 &+ \text{FT}[A \exp(i\Phi)] \\
 &\otimes \delta(R - R_c) && +1 \text{ Sideband} \\
 &+ \text{FT}[A \exp(-i\Phi)] \\
 &\otimes \delta(R + R_c) && -1 \text{ Sideband}
 \end{aligned}
 \tag{24}$$

Cutting out one sideband, after centering around  $R = 0$  and inverse Fourier transformation, one obtains the image wave or the conjugate complex, respectively. At this stage, the wave optical tools can be applied (e.g., to extract amplitude and phase separately). Wahl demonstrated the possibilities of phase analysis by superimposing a plane wave reconstructed from another hologram, or from a second ‘empty’ hologram recorded on the same photographic plate (‘double exposure hologram’) [54]. Likewise he showed the facility for the generation of a focal series from one hologram.

In particular, at high resolution the condition  $R_c \geq 3R_{\text{max}}$  poses severe problems with regard to recording the fine fringe spacing needed. Therefore, as a special development of image plane holography, the technique of phase-shifting holography is adopted from light optics [55–57]. Despite the fact that the fringes are coarser than the image details, by using at least three holograms taken at different fringe positions the image wave can be retrieved.

### 1.8.4.2 Further Developments

#### *Producing the Hologram*

Since the work of Wahl, the steps of recording the holograms and reconstructing the

image wave have been improved considerably. The indispensable highly coherent field emission sources, which were previously used only in special experiments in electron interferometry and holography [20, 22, 41, 43, 49, 58, 59], are now generally available. Microscopes for generating holograms have been developed that are dedicated to special purposes; for example, for magnetic specimens at low ambient magnetic fields [60, 61]; equipped with a helium cold stage for superconducting specimens [62]; with a superconducting lens system (e.g., for biologic specimens) [63]; or for the needs of atomic resolution [64].

The geometry of the hologram is given by the position of the biprism along the optical axis, and by the magnification factor in the subsequent intermediate image. The fringe spacing can be selected by altering the filament voltage. The hologram width also depends on the filament voltage, and can also be varied by manipulating the distance to the subsequent intermediate image plane and hence can be set by varying the excitation of the intermediate lens. To achieve greater flexibility with regard to fringe spacing and hologram width, in addition to the usual position of the biprism in the Selected Area aperture, Hitachi offers a second port for the electron biprism between the two intermediate lenses. A rotatable biprism [65] can be very helpful to bring the hologram into line with an object detail. For high resolution holography, the axial position of the biprism should be optimized [66].

#### *Recording the Hologram*

The detector has to be examined according to the following criteria:

- (a) The modulation transfer function (MTF) decreases the contrast of the hologram fringes.
- (b) The output signal must be *linearly* related to the electron current density distribution; otherwise the amplitude (not the phase) is distorted, resulting in artifacts under a posteriori focusing or correction of aberrations.
- (c) The detection quantum efficiency (DQE) increases the noise in the hologram.
- (d) A high dynamic range enables a high dose to be collected without running into saturation effects.
- (e) Geometric distortions produce artificial bending of the recorded wave.
- (f) The time delay between hologram recording and reconstruction, for example by darkroom processing.

With regard to most of these criteria, the still often used photoplates are not the best choice; darkroom processing takes a lot of time, the results are only reproducible with exact processing procedures, and the correction of nonlinearity, which is indispensable for many applications, is an awkward task [67]. Modern CCD cameras [68] are much more favorable in all these respects [69, 70], except in the available pixel number limiting the recordable field of view. CCD cameras for electron microscopy are available with  $1000 \times 1000$  pixels, and recently  $2000 \times 2000$  pixel cameras have been marketed by Photometrics and GATAN.

#### *Reconstruction of the Electron Object Wave*

The light optical reconstruction proposed by Gabor has the advantage of high reconstruction speed. In addition, the usual techniques for phase contrast and

interferometry (e.g., by means of a Mach–Zehnder interferometer [71, 72]) can readily be applied. However, in particular for quantitative analysis and the correction of aberrations, numerical reconstruction is a far superior technique. Furthermore, with the rapid development in computing power, numerical reconstruction can now be performed in almost real-time. Hybrid techniques have been developed [73] which, by means of a TV camera and a liquid crystal display (LCD), transfer the hologram to an optical bench in order to proceed very quickly with the optical techniques applicable for time-dependent phenomena. In another set-up, phase-plates for focusing and aberration correction are generated by a computer and, using an LCD arranged in the Fourier plane, conveyed to the optical bench.

For a purely numerical reconstruction, the holograms are recorded by means of a CCD camera, digitized and fed to a computer [74]. Usually, for wave optical processing, image processing software is applied (e.g., IMAGIC or SEMPER) which is equipped with user-written subroutines for the particular purpose. Special developments in hardware and software allow automated or operator-controlled reconstruction within a couple of seconds [75–77]. For very careful correction of aberrations, a special program allows the operation of a fast computer like a microscope in that arbitrary values for focus, spherical aberration, etc., are selected and, after about 1 s, the corresponding amplitude and phase images of the reconstructed wave are displayed [78]. A new approach to numerical reconstruction using neural networks is described by Heindl et al. [57].

### 1.8.4.3 Holography at Low and Medium Resolution

For the analysis of structures consisting of spatial frequencies smaller than about  $R_{\max} = 0.1 \text{ nm}^{-1}$ , the aberrations need not be taken into account if the electron microscope is aligned with care. The contribution from a spherical aberration to the wave aberration  $\chi(\mathbf{R})$  can be neglected; likewise, defocus at roughly Gauss focus does not contribute much either. Consequently, without too much effort, at low resolution  $\cos \chi \approx 1$  and  $\sin \chi \approx 0$  (i.e., ideal aberration-free imaging conditions can be reached), and the reconstructed wave agrees perfectly with the object wave. At medium resolution, up to about half Scherzer resolution  $R_{\text{schertz}}$ , a Gabor focus  $Dz_{\text{gab}}$  of  $0.56Dz_{\text{schertz}}$  should be selected, which gives a sufficiently aberration-free object wave up to  $R_{\max} = 0.884R_{\text{schertz}}$  without the need to correct for aberrations [79]. Therefore, the main and most essential benefit of electron holography at low and medium resolution lies in the superb wave optical facilities for the analysis of amplitudes and phases, as shown by many applications.

#### *Applications: Magnetic Structures*

The contour lines of equal phases modulo  $2\pi$  representing the phase distribution can be interpreted such that between two successive lines the magnetic flux  $\Phi_0 = h/e$  is enclosed. In the case of homogeneous magnetic fields restricted to the object, the contour lines show the lines of the magnetic field  $\mathbf{B}$  [80]. If stray fields are present, interpretation of the phase distribution in terms of field lines needs care because stray fields in the direction of the

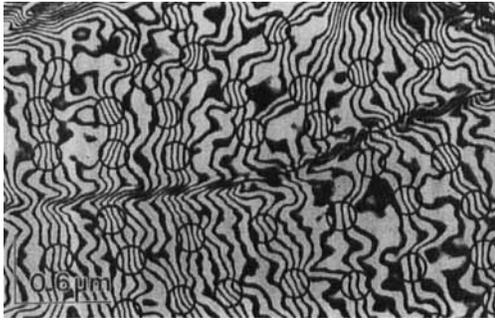
electron path sum together, and far-reaching stray fields may likewise modulate the reference wave. The power of holography for the analysis of magnetic microstructures has been shown by many examples [50, 81–83]. Certainly, the main progress in the holographic investigation of magnetic microstructures was achieved by Tonomura's group. The reader is referred to [84], where a detailed overview can be found. Typical applications are the analysis of:

- (a) Magnetization in thin films and the structure of magnetic domain walls.
- (b) Magnetization in and stray field around single and multiple domain particles.
- (c) Stray field distribution around tiny magnets produced by microlithography.
- (d) Confirmation of the Aharonov–Bohm effect at micrometer dimensions.
- (e) Stray field around ferromagnetic tips [60, 85].

The results obtained with superconductors, for example the study of single fluxons and, most recently, the study of vortex lattices in niobium at different ambient magnetic field strengths (Fig. 5) [86]. The experimental results are in good agreement with the theoretical description of the phase shift produced by a vortex [87, 88], and are very impressive.

#### *Applications: Electric Structures*

The phase distribution of the reconstructed wave represents the projected potential along the electron paths. Therefore, in the case of a three-dimensional potential distribution changing strongly along the electron path, the interpretation



**Figure 5.** Phase image of flux-line lattice of niobium specimen at 4.5 K and 100g magnetic field. The phase contour lines represent the projected magnetic lines of force indicating the vortices at the densified locations [86]. (Courtesy of A. Tonomura and Elsevier Publishers.)

of the phase lines may be rather complicated [89]. Again, far-reaching stray fields may affect the reference wave. In the case of electric fields limited to a well-defined object area, the phase distribution correlates with the potential distribution. Typical applications of holography in this field aim at the analysis of:

- (a) Stray fields around Latex spheres [90] and charged microtips [91].
- (b) Potential distributions over pn junctions [92–94].
- (c) Structure of silica gate oxide layers of field-effect transistors (FETs) [95].
- (d) Mean inner potentials of inorganic [96] and organic [97] materials.
- (e) Dynamic phase shifts in crystals [98].
- (f) Inelastic mean free path evaluation of the amplitude [99].
- (g) Ferroelectric domain walls [100].
- (h) Morphology of nanometer-sized particles [101, 102].
- (i) Cross-sections of different fullerene materials [103].
- (j) Potential distributions across grain boundaries [104].

- (k) Theoretical investigations of linear charged dislocations [105].
- (l) Composition profiles across heterogeneous interfaces [106].

In the field of biological objects, the use of holography can obviate the need for staining [107, 108]; at present, however, due to the radiation damage problem, noise still prohibits results equivalent to those obtained with conventional microscopy from being obtained. Hopefully, the use of a superconducting objective lens [63] will solve this problem.

The analysis of surfaces by reflection holography is made possible by superimposing waves which are Bragg-reflected at different areas of a crystal surface. The reconstructed wave shows a high sensitivity of the electron phase also for surfaces with subatomic topology [109, 110].

Another special technique is convergent beam interferometry, where two selected discs of the convergent beam diffraction pattern are superimposed by means of an electron biprism [111]. This method is still under development.

#### 1.8.4.4 Holography at High Resolution: Surmounting the Point Resolution

Beyond Gabor resolution, both  $\cos(\chi)$  and  $\sin(\chi)$  begin to oscillate heavily. Therefore, neither the amplitude nor the phase of the reconstructed wave can be interpreted in terms of the amplitude and phase of the object. Therefore, aberrations must be corrected during reconstruction prior to wave optical analysis [112, 113].

Correction of aberrations is accomplished by the application of a numerically generated wave transfer function  $WTF_{\text{num}}(\mathbf{R})$  to the Fourier spectrum of the reconstructed image wave. Then the reconstructed object wave is given by

$$o_{\text{rec}}(x, y) = \text{FT}^{-1}\{\text{FT}[o(x, y)] \cdot WTF_{\text{mic}}(\mathbf{R})/WTF_{\text{num}}(\mathbf{R})\} \quad (25)$$

with the wave transfer functions

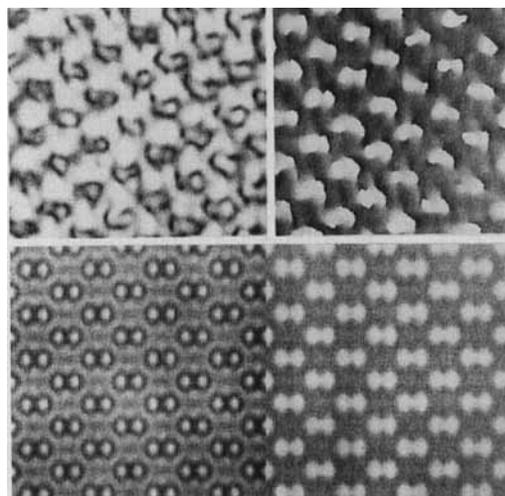
$$WTF_{\text{mic}}(\mathbf{R}) = E_{\text{coh}}(\mathbf{R}) \cdot E_{\text{chrom}}(\mathbf{R}) \cdot \exp[-i\chi_{\text{mic}}(\mathbf{R})] \quad (26)$$

for the electron microscope and  $WTF_{\text{num}}(\mathbf{R})$  numerically modeling the microscope accordingly [114].

Of course, the goal is that  $o_{\text{rec}}(x, y)$  approaches the object wave  $o(x, y)$  as closely as possible, for instance to reach a resolution of 0.1 nm. This means that the highest possible spatial frequencies must be caught in the hologram, no additional artifacts must be introduced by the method, and the reconstruction and correction procedures have to be performed as carefully as possible. With a 100 kV electron microscope, a resolution of 0.2 nm has been achieved [115]. A thorough analysis of the holographic steps shows that a desirable resolution of 0.1 nm will be possible at an electron energy of 300 kV, if the following points are adhered to [116]; experimentally, about 0.13 nm has been reached up-to-date (Fig. 6).

#### Making the Hologram

The information limit  $R_{\text{lim}}$  of the electron microscope must be larger than  $R_{\text{max}}$ , the highest spatial frequency of interest.



**Figure 6.** Object wave of silicon in the  $\langle 110 \rangle$  orientation reconstructed from a 300 kV hologram. The dumbbells with a spacing of 0.136 nm can be discerned in both the amplitude (left) and phase (right) images. (Top) Holographic reconstruction; (bottom) corresponding simulated results. (From [136].)

Therefore, for the chromatic envelope function  $E_{\text{chrom}}(\mathbf{R})$  a small coefficient  $C_c$  of chromatic aberration and a small energy spread is needed; with a field emission gun and usual  $C_c$  values, at 300 kV,  $E_{\text{chrom}}(\mathbf{R})$  reaches beyond  $R = 10 \text{ nm}^{-1}$ . However, the coherence envelope function

$$E_{\text{coh}}(\mathbf{R}) = \exp\left\{-\left[\frac{\theta_c}{\sqrt{\ln 2} \cdot \lambda} \text{grad}\chi_{\text{mic}}(\mathbf{R})\right]^2\right\} \quad (27)$$

where  $\text{grad}[\chi_{\text{mic}}(\mathbf{R})]$  depends on the focus and  $\theta_c$  is the illumination semi-aperture. With modern microscopes one usually finds, at Scherzer focus,  $R_{\text{lim}} \leq 7 \text{ nm}^{-1}$ , and hence holographically a resolution of about 0.1 nm cannot be reached at Scherzer focus.

Surprisingly, for high resolution the hologram must have a minimum width of

$w \geq 4\text{PSF}$ , where

$$\text{PSF} = \text{grad}[\chi_{\text{mic}}(\mathbf{R})]_{\text{max}}/\pi \quad (28)$$

is the diameter of the PSF in the image plane of the electron microscope;  $\text{grad}[\chi_{\text{mic}}(\mathbf{R})]_{\text{max}}$  is the maximum value of  $\text{grad}[\chi_{\text{mic}}(\mathbf{R})]$  for  $|\mathbf{R}|$  in the range  $[0, R_{\text{max}}]$ . Evidently, in contrast to Scherzer focus, which optimizes the  $\chi_{\text{mic}}(\mathbf{R})$  function for phase contrast in conventional microscopy, in holography the optimum focus

$$Dz_{\text{opt}} = 0.75C_s \left( \frac{R_{\text{max}}}{k} \right)^2 \quad (29)$$

has to be chosen to minimize  $\text{grad}[\chi_{\text{mic}}(\mathbf{R})]_{\text{max}}$  [79]. Then, compared to the Scherzer focus, with the PSF taking the value

$$\text{PSF} = 0.5C_s \left( \frac{R_{\text{max}}}{k} \right)^3 \quad (30)$$

the PSF is reduced by a factor 4; additionally, the information limit  $R_{\text{lim}}$  is increased by a factor of about 1.6 and hence exceeds the desirable threshold of  $10 \text{ nm}^{-1}$  [64].

Taking account of the condition  $R_c \geq 3R_{\text{max}}$ , at optimum focus one finds that a minimum number of fringes

$$N_{\text{fringe}} = R_c w = 6C_s \frac{R_{\text{max}}^4}{k^3} \quad (31)$$

is needed for a holographic resolution  $R_{\text{max}}$ . With a 300 kV microscope, typical values are  $w \geq 20 \text{ nm}$  and  $N_{\text{fringe}} = 600$  for a resolution of 0.1 nm.

### *Recording the Hologram*

Because of the drawbacks of the photoplate, in particular the nonlinearity producing artificial distortions of the recorded wave, for high resolution holography the

choice is a CCD camera. Here, the critical point is the available pixel number, which may limit the holographic resolution. The reason for this is that each fringe has to be sampled by at least 4 pixels [117]. Therefore the number of needed pixels is

$$N_{\text{pix}} \geq 4N_{\text{fringe}} = 24C_s \frac{R_{\text{max}}^4}{k^3} \quad (32)$$

and, by inversion, one finds

$$R_{\text{max}} = \left[ \frac{N_{\text{pix}} k^3}{24C_s} \right]^{1/4} \quad (33)$$

Related to the Scherzer resolution

$$R_{\text{schertz}} = 1.5 \left[ \frac{k^3}{C_s} \right]^{1/4} \quad (34)$$

of the microscope a holographic improvement factor of resolution

$$\frac{R_{\text{max}}}{R_{\text{schertz}}} = 0.3N_{\text{pix}}^{1/4} \quad (35)$$

results, given by the pixel number of the CCD camera. From a hologram taken with a microscope at point resolution  $R_{\text{schertz}}$  and a pixel number  $N_{\text{pix}}$ , the resolution of the reconstruction is, at best,  $R_{\text{max}}$  [12]. A different approach is discussed in [118].

### *Reconstruction of the Object Exit Wave by Correction of Aberrations*

The carrier frequency  $R_c$  and orientation of the hologram fringes determine the position of the sideband in Fourier space. In general, the center of the sideband does not coincide exactly with a pixel, and thus strong streaks arise which severely hamper the cutting out of the sideband for further processing and quantitative analysis in Fourier space. Furthermore, the reconstructed wave would be tilted with respect

to the optical axis. Therefore, the sideband must be centered around the optical axis with subpixel accuracy. Since interpolation would produce severe artifacts, special methods for subpixel centering, which involve a precise determination of the hologram parameters, have been developed [78, 119].

Next, for the correction of aberrations, the wave transfer function

$$\text{WTF}_{\text{num}}(\mathbf{R}) = E_{\text{num}}(\mathbf{R}) \cdot \exp[-i\chi_{\text{num}}(\mathbf{R})] \quad (36)$$

has to be created such that it models precisely the electron microscope:

$$\text{WTF}_{\text{num}}(\mathbf{R}) \equiv \text{WTF}_{\text{mic}}(\mathbf{R}) \quad (37)$$

delivers  $o_{\text{rec}}(x, y) \equiv o(x, y)$ . Modeling the envelope function  $E_{\text{num}}(\mathbf{R})$  is not critical; however, the wave aberration  $\chi_{\text{num}}(\mathbf{R})$  must be dealt with extremely carefully.

Ideally, the resulting wave aberration describing the overall transfer of the object exit wave to the reconstructed object wave should vanish. In reality, however, a residual aberration  $\delta\chi(\mathbf{R}) = \chi_{\text{num}}(\mathbf{R}) - \chi_{\text{mic}}(\mathbf{R})$  will distort the result by the transfer functions  $\cos(\delta\chi)$  and  $\sin(\delta\chi)$ . Freedom from aberrations can be assumed if  $\cos(\delta\chi) \geq 0.86$  and  $\sin(\delta\chi) \leq 0.5$ , that is if  $|\delta\chi| \leq \pi/6$  holds over the entire range  $[0, R_{\text{max}}]$  [116].

In the most simple case, only spherical aberration and defocus have to be considered

$$\delta\chi(\mathbf{R}) = 2\pi k \left[ 0.25 \delta C_s \left( \frac{R}{k} \right)^4 - 0.5 \delta Dz \left( \frac{R}{k} \right)^2 \right] \quad (38)$$

where  $\delta C_s$  and  $\delta Dz$  are the imprecision occurring in the determination of the

spherical aberration and defocus that are effective in the hologram; to reach  $|\delta\chi| \leq \pi/6$ , they simultaneously have to be smaller than

$$|\delta C_s| \leq \frac{k^3}{6R_{\text{max}}^4} \quad (39)$$

and

$$|\delta Dz| \leq \frac{k}{12R_{\text{max}}^2} \quad (40)$$

At 300 kV and  $R_{\text{max}} = 10 \text{ nm}^{-1}$ , the values are  $\delta C_s = 2 \mu\text{m}$  and  $\delta Dz = 0.4 \text{ nm}$ .

In the general case, with a comparable precision, one has also to determine:

- (a) The wavenumber  $k$ .
- (b) The gauging of the spatial frequencies.
- (c) The two- and three-fold astigmatism [120].
- (d) The axial coma.

Altogether this makes 10 parameters that are needed to model the electron microscope in the computer, all of which have to be determined for the specific hologram. However, the usual methods of diffractometry do not give sufficient accuracy, not even with sophisticated methods [121]; in addition, these data represent the microscope at the time of measurement, but not the aberrations caught in the specific hologram. For refinement, one can make use of the reconstructed wave by means of, for example:

- (a) Minimizing the amplitude contrast at pure phase objects [122, 123].
- (b) Minimizing the amplitude contrast in interatomic space [124].
- (c) Matching the amplitude and phase of ideal crystalline areas with simulations.
- (d) Evaluating the phases in the Fourier spectrum [122].

*Quality of the Reconstructed Object Wave*

Besides resolution, the degree of fidelity with the object exit wave and the signal-to-noise ratio properties need to be considered.

*Artifacts.* In holography there is a considerable risk of catching artifacts [125] from the following sources. Fresnel diffraction at the biprism produces parasitic amplitude and phase structures, which cannot easily be removed. Also, vignetting occurs due to the biprism, in that parts of the field of view are transferred as a single-sideband image with falsified amplitude and phase. Another serious artifact comes about because of the restricted field of view and the resulting broadening of reflections in reconstructed Fourier space. Large problems may arise due to geometric distortions of the hologram (e.g., from projector lenses) which falsify the phase; even worse, they restrict isoplanacy in the reconstructed wave, and hence may limit the applicability of the correction procedure and resolution. These artifacts can be avoided by taking wider holograms, and by correcting distortions by means of an empty reference hologram carrying the information needed, respectively [74].

*Signal-to-Noise Properties.* Correction of aberrations means a posteriori opening of the objective aperture. Therefore, the signal will be enhanced and there is some hope of detecting single atoms (e.g., gold or even oxygen with a phase shift of  $2\pi/12$  and  $2\pi/60$ , respectively). However, noise must be accordingly low [126–129].

The phase detection limit from a hologram is given by

$$\delta\phi = \frac{\text{SNR}}{V} \sqrt{\frac{2}{N'}} \quad (41)$$

The contrast  $V = |\mu| \cdot V_{\text{inst}} \cdot V_{\text{inel}} \cdot V_{\text{MTF}}$  is given by the degree of coherence, the instabilities, the inelastic interaction and the MTF of the detector.  $N' = N \cdot \text{DQE}$  means the number of electrons per resolved area reduced by the DQE of the detector.

Inserting

$$N = j_{\text{coh}} \epsilon t / R_{\text{max}}^2 \quad (42)$$

$$j_{\text{coh}} = \frac{-2B_0 \ln(|\mu|)}{\pi k^2 w^2} \quad (43)$$

and

$$w = 2C_s \left[ \frac{R_{\text{max}}}{k} \right]^3 \quad (44)$$

one obtains

$$\delta\phi = \frac{2\sqrt{\pi} C_s R_{\text{max}}^4 \text{SNR}}{|\mu| \cdot V_{\text{inel}} \cdot V_{\text{inst}} \cdot V_{\text{MTF}} \cdot k^2 \sqrt{-B_0 \ln(|\mu|)} t \epsilon \cdot \text{DQE}} \quad (45)$$

Note that  $C_s$  is an essential limiting factor, despite the fact that it is corrected at the end. For optimum signal-to-noise ratio (SNR) one has to take into account:

- That  $|\mu| \sqrt{-\ln(|\mu|)}$  is maximum at  $|\mu| = 0.61$ .
- The brightness  $B_0$  of the electron gun must be a maximum.
- There must be a trade-off between exposure time  $t$  and  $V_{\text{inst}}$ .
- The need for extremely low AC magnetic stray fields, mechanical stability of biprism and the object, and vibrations and acoustic noise.
- That a priori correction of spherical aberration [130] could help greatly.

With special care,  $\delta\phi \leq 2\pi/60$  at  $\text{SNR} = 3$  and  $R_{\text{max}} = 10 \text{ nm}^{-1}$  is within reach.

### Wave Optical Analysis

The reconstructed object wave can be analyzed systematically and quantitatively wave optically from the amplitude and phase image. In our experience, the phase image is often much easier to interpret in that, for example, there are no half-spacings and local crystal tilt can readily be seen by asymmetric phase structures at atomic columns. Phase jumps at the extinction thickness are readily discernible. Application of a numerical selected-area aperture with a diameter down to less than 1 nm and Fourier transform yields nano-diffraction patterns for a local analysis of crystal tilt and thickness effects. Both the amplitude and the phase of Fourier spectra can be studied; there is no phase problem with holography. Selecting single reflections or groups by using an aperture allows one to display the object wave under arbitrary 'selective imaging' conditions such as bright field, dark field, or diffraction contrast, and to investigate, for example, the scattering phases for each reflection [131].

Since all this is done using one hologram, the data obtained can be mutually related in a definite way. In addition to the availability of the phases, it is a great advantage that during the analysis neither the 'microscope' or the 'specimen' change (e.g., by shift of alignment, drift, contamination, or radiation damage). Applications to the study of real crystal structures are reported in [132–134].

### Outlook

With the CM30FEG-Special Tuebingen microscope dedicated to high resolution electron holography ( $C_s = 1.2$  mm,  $B_0 = 5 \times 10^8$  A cm<sup>-2</sup> sr<sup>-1</sup>,  $R_{\text{scherz}} = 5$  nm<sup>-1</sup>,

$R_{\text{lim}} = 10$  nm<sup>-1</sup>) and a 1000 × 1000 pixel CCD-camera, a holographic resolution of  $R_{\text{max}} = 7.5$  nm<sup>-1</sup> was achieved [135, 136]. After upgrading with the UltraTwin lens ( $C_s = 0.62$  mm,  $R_{\text{scherz}} = 5.7$  nm<sup>-1</sup>), and a 2000 × 2000 pixel CCD camera yielding a resolution improvement factor of 2, a holographic resolution of  $R_{\text{max}} = 10$  nm<sup>-1</sup> seems realistic.

## 1.8.5 Summary

After more than 40 years of development, electron holography has reached a state which shows the high potential for micro-structure analysis beyond the limits of conventional electron microscopy. This has already been proven by a wide range of applications in materials science.

At present, off-axis image plane holography is the most advanced holographic technique, at both low and high resolution. Presumably, the other techniques under development will also show special areas of applications where they will allow unique features to be observed, and new techniques will arise with the general progress in electron wave optics.

## 1.8.6 References

- [1] W. Glaser, *Z. Phys.* **1933**, 80, 451.
- [2] W. Ehrenberg, R. E. Siday, *Proc. Phys. Soc. B* **1949**, 62, 8.
- [3] Y. Aharonov, D. Bohm, *Phys. Rev.* **1959**, 115, 485.
- [4] O. Scherzer, *Z. Phys.* **1936**, 101, 593.
- [5] O. Scherzer, *J. Appl. Phys.* **1949**, 20, 20.
- [6] D. Gabor, *Nature* **1948**, 161, 777.
- [7] D. Gabor, G. W. Stroke, D. Brumm, A. Funkhouser, A. Labeyrie, *Nature* **1965**, 208, 1159.

- [8] L. O. Heflinger, R. F. Wuerker, R. E. Brooks, *J. Appl. Phys.* **1966**, *37*, 642.
- [9] O. Bryngdahl, *J. Opt. Soc. Am.* **1969**, *59*, 142.
- [10] D. Gabor, *Proc. R. Soc. London* **1951**, *64*, 449.
- [11] H. Schmid, PhD. Thesis, Tübingen **1985**.
- [12] H. Lichte, *Ultramicroscopy* **1993**, *51*, 15.
- [13] E. Leith, J. Upatnieks, *J. Opt. Soc. Am.* **1962**, *52*, 1123.
- [14] J. M. Cowley, *Ultramicroscopy* **1992**, *41*, 335.
- [15] M. E. Haine, T. Mulvey, *J. Opt. Soc. Am.* **1952**, *42*, 763.
- [16] H.-W. Fink, H. Schmid, H. J. Kreuzer, A. Wierzbicki, *Phys. Rev. Lett.* **1991**, *67*, 1543.
- [17] H. J. Kreuzer, H.-W. Fink, H. Schmid, S. Bonev, *J. Microsc.* **1995**, *178*, 191.
- [18] J. C. H. Spence, X. Zhang, W. Qian, in *Electron Holography* (Eds.: A. Tonomura, L. F. Allard, G. Pozzi, D. C. Joy, Y. A. Ono), Elsevier, Amsterdam **1995**, 267.
- [19] A. Tonomura, A. Fukuhara, H. Watanabe, T. Komoda, *Jpn. J. Appl. Phys.* **1968**, *7*, 295.
- [20] N. Bonnet, M. Troyon, P. Gallion, in *Proc. 9th Int. Cong. Electron Microscopy*, Microscopical Society of Canada, Toronto **1978**, vol. I, pp. 222–223.
- [21] T. Matsumoto, T. Tanji, A. Tonomura, in *Electron Holography* (Eds.: A. Tonomura, L. F. Allard, G. Pozzi, D. C. Joy, Y. A. Ono), Elsevier, Amsterdam **1995**, 249.
- [22] J. Munch, *Optik* **1975**, *43*, 79.
- [23] T. Matsumoto, T. Tanji, A. Tonomura, *Optik* **1995**, *100*, 71.
- [24] R. Lauer, *Optik* **1984**, *66*, 159.
- [25] J. Cowley, in *Electron Diffraction Techniques* (Ed.: J. M. Cowley), Oxford Science Publications, Oxford **1992**, vol. I, 439.
- [26] J. W. Steeds, P. A. Midgley, P. Spellward, R. Vincent, in *Electron Holography* (Eds.: A. Tonomura, L. F. Allard, G. Pozzi, D. C. Joy, Y. A. Ono), Elsevier, Amsterdam **1995**, 277.
- [27] J. A. Lin, J. M. Cowley, in *Proc. 43rd Meeting of the Electron Microscopical Society of America*, Electron Microscopical society of America, **1985**, 136.
- [28] J. M. Cowley, M. A. Gribelyuk, *MSA Bull.* **1994**, *24*, 438.
- [29] L. Marton, *Phys. Rev.* **1952**, *85*, 1057.
- [30] G. Matteucci, G. F. Missiroli, G. Pozzi, *Ultramicroscopy* **1981**, *8*, 109.
- [31] Q. Ru, in *Electron Holography* (Eds.: A. Tonomura, L. F. Allard, G. Pozzi, D. C. Joy, Y. A. Ono), Elsevier, Amsterdam **1995**, 343.
- [32] G. Möllenstedt, H. Düker, *Z. Phys.* **1956**, *145*, 377.
- [33] R. Buhl, *Z. Phys.* **1959**, *155*, 395–412.
- [34] J. Faget, Ch. Fert, *C. Phys.* **1987**, *11*, 285.
- [35] M. Keller, *Z. Phys.* **1961**, *164*, 274.
- [36] W. Bayh, *Z. Phys.* **1962**, *169*, 492.
- [37] H. Wahl, *Optik* **1970**, *30*, 577.
- [38] B. Lischke, *Z. Phys.* **1970**, *239*, 360.
- [39] E. Krimmel, G. Möllenstedt, W. Rothemund, *Appl. Phys. Lett.* **1964**, *5*, 209.
- [40] T. Hibi, S. Takahashi, *Z. Phys.* **1969**, *3*, 132.
- [41] R. Speidel, D. Kurz, *Optik* **1977**, *49*, 173.
- [42] F. Lenz, G. Wohland, *Optik* **1984**, *67*, 315.
- [43] H. Lichte, G. Möllenstedt, *J. Phys. E: Sci. Instrum.* **1979**, *12*, 941.
- [44] A. Tonomura, *Jpn. J. Appl. Phys.* **1972**, *11*, 493.
- [45] G. Pozzi, G. F. Missiroli, *J. Microsc.* **1973**, *18*, 103.
- [46] F. Hasselbach, M. Nicklaus, *Phys. Rev. A* **1993**, *48*, 143.
- [47] G. Möllenstedt, H. Wahl, *Naturwissenschaften* **1968**, *55*, 340.
- [48] H. Tomita, T. Matsuda, T. Komoda, *Jpn. J. Appl. Phys.* **1972**, *11*, 143.
- [49] G. Saxon, *Optik* **1972**, *35*, 359.
- [50] M. Mankos, P. de Haan, V. Kambarsky, G. Matteucci, M. R. McCartney, Z. Yang, M. R. Scheinfein, J. M. Cowley, in *Electron Holography* (Eds.: A. Tonomura, L. F. Allard, G. Pozzi, D. C. Joy, Y. A. Ono), Elsevier, Amsterdam **1995**, 329.
- [51] I. Weingärtner, W. Mirandé, E. Menzel, *Optik* **1969**, *30*, 318.
- [52] O. Bryngdahl, A. W. Lohmann, *J. Op. Soc. Am.* **1968**, *58*, 141.
- [53] H. Wahl, Thesis, Tübingen **1975**.
- [54] H. Wahl, *Optik* **1974**, *39*, 585.
- [55] W. D. Rau, H. Lichte, K.-H. Herrmann, *Optik* **1989**, *83* (Suppl 4), 79.
- [56] Q. Ru, G. Lai, K. Aoyama, J. Endo, A. Tonomura, *Ultramicroscopy* **1994**, *55*, 209.
- [57] E. Heindl, W. D. Rau, H. Lichte, *Ultramicroscopy*, in press.
- [58] W. Brünger, *Z. Phys.* **1972**, *250*, 263.
- [59] A. Tonomura, T. Matsuda, T. Komoda, *Jpn. J. Appl. Phys.* **1978**, *17*, 1137.
- [60] B. Frost, Thesis, Tübingen **1993**.
- [61] T. Hirayama, J. Chen, Q. Ru, K. Ishizuka, T. Tanji, A. Tonomura, in *Electron Holography* (Eds.: A. Tonomura, L. F. Allard, G. Pozzi, D. C. Joy, Y. A. Ono), Elsevier, Amsterdam **1995**, 145.
- [62] A. Tonomura, N. Osakabe, T. Matsuda, T. Kawasaki, J. Endo, S. Yano, H. Yamada, *Phys. Rev. Lett.* **1986**, *56*, 792.
- [63] U. Weierstall, H. Lichte, *Ultramicroscopy*, in press.
- [64] H. Lichte, P. Kessler, F. Lenz, W.-D. Rau, *Ultramicroscopy* **1993**, *52*, 575.
- [65] K. Yada, K. Shibata, T. Hibi, *J. Elec. Microsc.* **1973**, *22*, 223.

- [66] H. Lichte, *Ultramicroscopy*, in press.
- [67] E. Völkl, F. Lenz, Q. Fu, H. Lichte, *Ultramicroscopy* **1994**, 55, 75.
- [68] J. H. C. Spence, J. M. Zuo, *Rev. Sci. Instrum.* **1988**, 59, 2102.
- [69] I. Daberkow, K.-H. Herrmann, Libin Liu, W. D. Rau, *Ultramicroscopy* **1991**, 38, 215.
- [70] O. L. Krivanek, P. E. Mooney, *Ultramicroscopy* **1993**, 49, 95.
- [71] A. Tonomura, J. Endo, T. Matsuda, *Optik* **1979**, 53, 143.
- [72] K.-J. Hanszen, *Adv. Electron. Electron Phys.* **1982**, 59, 1.
- [73] J. Chen, T. Hirayama, G. Lai, T. Tanji, K. Ishizuka, A. Tonomura, in *Electron Holography* (Eds.: A. Tonomura, L. F. Allard, G. Pozzi, D. C. Joy, Y. A. Ono), Elsevier, Amsterdam **1995**, 81.
- [74] W. D. Rau, H. Lichte, E. Völkl, U. Weierstall, *J. Comput. Assist. Microsc.* **1991**, 3, 51.
- [75] W. D. Rau, *MSA Bull.* **1994**, 24, 459.
- [76] E. Voelkl, L. F. Allard, B. Frost, *J. Microsc.*, **1995**, 180, 39.
- [77] M. Lehmann, E. Völkl, F. Lenz, *Ultramicroscopy* **1994**, 54, 335.
- [78] M. Lehmann, H. Lichte, in *Electron Holography* (Eds.: A. Tonomura, L. F. Allard, G. Pozzi, D. C. Joy, Y. A. Ono), Elsevier, Amsterdam **1995**, 69.
- [79] H. Lichte, *Ultramicroscopy* **1991**, 38, 13.
- [80] H. Wahl, B. Lau, *Optik* **1979**, 54, 27.
- [81] B. Lau, G. Pozzi, *Optik* **1978**, 51, 287.
- [82] G. Matteucci, G. F. Missiroli, J. W. Chen, G. Pozzi, *Appl. Phys. Lett.* **1988**, 52, 176.
- [83] G. Matteucci, M. Muccini, *Ultramicroscopy* **1994**, 53, 19.
- [84] A. Tonomura, *Springer Ser. Opt. Sci.* **1993**, 70.
- [85] G. Matteucci, M. Muccini, U. Hartmann, *Appl. Phys. Lett.* **1993**, 62, 1839.
- [86] J. E. Bonevich, K. Harada, H. Kasai, T. Matsuda, T. Yoshida, G. Pozzi, A. Tonomura, in *Electron Holography* (Eds.: A. Tonomura, L. F. Allard, G. Pozzi, D. C. Joy, Y. A. Ono), Elsevier, Amsterdam **1995**, 135.
- [87] A. Migliori, G. Pozzi, *Ultramicroscopy* **1992**, 41, 169.
- [88] A. Migliori, G. Pozzi, A. Tonomura, *Ultramicroscopy* **1993**, 49, 87.
- [89] J. W. Chen, G. Matteucci, A. Migliori, E. Nichelatti, G. Pozzi, M. Vanzi, *Phys. Rev. A* **1989**, 40, 3136.
- [90] B. G. Frost, L. F. Allard, E. Völkl, D. C. Joy, in *Electron Holography* (Eds.: A. Tonomura, L. F. Allard, G. Pozzi, D. C. Joy, Y. A. Ono), Elsevier, Amsterdam **1995**, 169.
- [91] G. Matteucci, G. F. Missiroli, M. Muccini, G. Pozzi, *Ultramicroscopy* **1992**, 45, 77.
- [92] S. Frabboni, G. Matteucci, G. Pozzi, M. Vanzi, *Phys. Rev. Lett.* **1985**, 55, 2196.
- [93] C. Capiluppi, A. Migliori, G. Pozzi, *J. Microsc. Microanal. Microstruct.*, in press.
- [94] M. R. McCartney, B. Frost, R. Hull, M. R. Scheinfein, D. J. Smith, E. Voelkl, in *Electron Holography* (Eds.: A. Tonomura, L. F. Allard, G. Pozzi, D. C. Joy, Y. A. Ono), Elsevier, Amsterdam **1995**, 189.
- [95] W.-D. Rau, F. H. Baumann, P. K. Roy, J. A. Rentschler, A. Ourmazd, *Appl. Phys. Lett.*, in press.
- [96] M. Gajdardziska-Josifovska, M. R. McCartney, W. J. de Ruijter, D. J. Smith, J. K. Weiss, J. M. Zuo, *Ultramicroscopy* **1993**, 50, 285.
- [97] M. Libera, J. Ott, Y. C. Wang, in *Electron Holography* (Eds.: A. Tonomura, L. F. Allard, G. Pozzi, D. C. Joy, Y. A. Ono), Elsevier, Amsterdam **1995**, 231.
- [98] K.-J. Hanszen, *J. Phys. D: Appl. Phys.* **1986**, 19, 373.
- [99] M. R. McCartney, M. Gajdardziska-Josifovska, *Ultramicroscopy* **1994**, 53, 283.
- [100] Xiao Zhang, D. C. Joy, Y. Zhang, T. Hashimoto, L. Allard, T. A. Nolan, *Ultramicroscopy* **1993**, 51, 21.
- [101] E. Völkl, L. F. Allard, B. Frost, in *Electron Holography* (Eds.: A. Tonomura, L. F. Allard, G. Pozzi, D. C. Joy, Y. A. Ono), Elsevier, Amsterdam **1995**, 103.
- [102] A. K. Datye, D. S. Kalakkad, E. Völkl, L. F. Allard, in *Electron Holography* (Eds.: A. Tonomura, L. F. Allard, G. Pozzi, D. C. Joy, Y. A. Ono), Elsevier, Amsterdam **1995**, 199.
- [103] L. F. Allard, E. Völkl, S. Subramoney, R. S. Ruoff, in *Electron Holography* (Eds.: A. Tonomura, L. F. Allard, G. Pozzi, D. C. Joy, Y. A. Ono), Elsevier, Amsterdam **1995**, 219.
- [104] X. Lin, V. Ravikumar, R. Rodrigues, N. Wilcox, V. Dravid, in *Electron Holography* (Eds.: A. Tonomura, L. F. Allard, G. Pozzi, D. C. Joy, Y. A. Ono), Elsevier, Amsterdam **1995**, 209.
- [105] D. Cavalcoli, G. Matteucci, M. Muccini, *Ultramicroscopy* **1995**, 57, 385.
- [106] J. K. Weiss, W. J. der Ruijter, M. Gajdardziska-Josifovska, M. R. McCartney, D. J. Smith, *Ultramicroscopy* **1993**, 50, 301.
- [107] H. Lichte, U. Weierstall, *Inst. Phys. Conf. Ser.* **93** **1988**, 3, 325.
- [108] K. Aoyama, G. Lai, Q. Ru, in *Electron Holography* (Eds.: A. Tonomura, L. F. Allard, G. Pozzi, D. C. Joy, Y. A. Ono), Elsevier, Amsterdam **1995**, 239.
- [109] N. Osakabe, T. Matsuda, J. Endo, A. Tonomura, *Ultramicroscopy* **1993**, 48, 483.

- [110] H. Banzhof, K.-H. Herrmann, *Ultramicroscopy* **1993**, *48*, 475.
- [111] R. A. Herring, G. Pozzi, in *Electron Holography* (Eds.: A. Tonomura, L. F. Allard, G. Pozzi, D. C. Joy, Y. A. Ono), Elsevier, Amsterdam **1995**, 287.
- [112] H. Lichte, *Ultramicroscopy* **1986**, *20*, 293.
- [113] H. Lichte, *Adv. Opt. Electron Microsc.* **1991**, *12*, 25.
- [114] Q. Fu, H. Lichte, E. Völkl, *Phys. Rev. Lett.* **1991**, *67*, 2319.
- [115] A. Harscher, G. Lang, H. Lichte, *Ultramicroscopy* **1995**, *58*, 79.
- [116] H. Lichte, *Ultramicroscopy* **1992**, *47*, 223.
- [117] F. Lenz, E. Völkl, Stochastic limitations to phase and contrast determination in electron holography, *Proc. XIIth Int. Cong. Electron Microscopy*, San Francisco Press, Inc., Seattle **1990**, 228.
- [118] K. Ishizuka, *Ultramicroscopy* **1994**, *53*, 297.
- [119] E. Völkl, L. F. Allard, *MSA Bull.* **1994**, *24*, 466.
- [120] O. L. Krivanek, *Ultramicroscopy* **1994**, *55*, 419.
- [121] D. Typke, K. Dierksen, *Optik* **1995**, *99*, 155.
- [122] Q. Fu, H. Lichte, *J. Microsc.* **1995**, *179*, 112.
- [123] K. Ishizuka, T. Tanji, A. Tonomura, T. Ohno, Y. Murayama, *Ultramicroscopy* **1994**, *53*, 361.
- [124] M. A. Gribelyuk, J. M. Cowley, *Ultramicroscopy* **1993**, *50*, 29.
- [125] H. Lichte, D. Geiger, A. Harscher, E. Heindl, M. Lehmann, D. Malamidis, A. Orchowski, W. D. Rau, *Ultramicroscopy*, in press.
- [126] H. Lichte, K.-H. Herrmann, F. Lenz, *Optik* **1987**, *77*, 135.
- [127] F. Lenz, *Optik* **1988**, *79*, 13.
- [128] W. J. de Ruijter, J. K. Weiss, *Ultramicroscopy* **1993**, *50*, 269.
- [129] A. Harscher, H. Lichte, *Ultramicroscopy*, in press.
- [130] M. Haider, G. Braunshausen, E. Schwan, *Optik* **1995**, *99*, 167.
- [131] H. Lichte, E. Völkl, K. Scheerschmidt, *Ultramicroscopy* **1992**, *47*, 231.
- [132] E. Schweda, J. Strähle, H. Lichte, E. Völkl, *Ultramicroscopy* **1994**, *53*, 53.
- [133] T. Tanji, K. Ishizuka, *MSA Bull.* **1994**, *24*, 494.
- [134] A. Orchowski, H. Lichte, *Ultramicroscopy*, in press.
- [135] H. Lichte, W.-D. Rau, *Ultramicroscopy* **1994**, *54*, 310.
- [136] A. Orchowski, W. D. Rau, H. Lichte, *Phys. Rev. Lett.* **1995**, *74*, 399.

# 1.9 Spin-Polarized Low-Energy Electron Microscopy

## 1.9.1 Introduction

Spin-polarized low-energy electron microscopy (SPLEEM) is a method for imaging the magnetic microstructure of surfaces and thin films with slow specularly reflected electrons. It is based on the fact that electron scattering is spin-dependent via the spin–spin and spin–orbit interactions between the incident electron and the specimen. In specular reflection only the spin–spin interaction occurs. If the specimen has regions with preferred spin alignment at length scales at or above the resolution limit of low-energy electron microscopy (LEEM), then these regions can be imaged via the contribution of the spin–spin interaction to the total scattering potential. Thus, magnetic and structural information is obtained simultaneously. SPLEEM is easily combined with low-energy electron diffraction (LEED), mirror electron microscopy (MEM) and the various types of emission microscopy (photo electron, secondary electron or thermionic electron emission microscopy). Lateral and depth resolution, information depth, and field of view are comparable to that of LEEM (see Chap. 4, Sec. 1.6 of this Volume). Therefore,

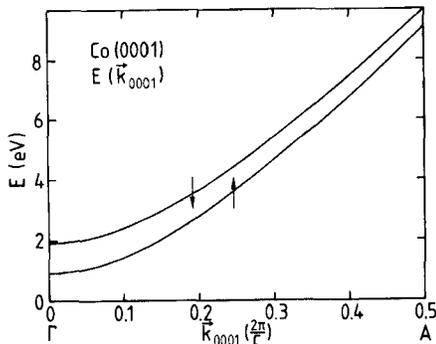
SPLEEM is an excellent method for the study of the correlation between magnetic structure, microstructure, and crystal structure.

## 1.9.2 Theoretical Foundations

Polarized electrons [1] have been used for some time in the study of the structure and magnetism of surfaces by spin-polarized low-energy electron diffraction (SPLEED) [2–5]. SPLEED is a laterally averaging method, but if the sample is used as a cathode in a cathode lens electron microscope the diffracted electrons may be used for imaging the surface in the same manner as in a standard LEEM instrument. The difference between SPLEEM and LEEM results from the fact that the incident beam is spin polarized in SPLEEM. The exchange interaction  $V_{\text{ex}} = \sum_i J(\mathbf{r} - \mathbf{r}_i) \mathbf{s} \cdot \mathbf{s}_i$ , where  $\mathbf{s}$ ,  $\mathbf{s}_i$  are the spins of the incident and target electron,  $\mathbf{r}$ ,  $\mathbf{r}_i$  their positions, and  $J(\mathbf{r} - \mathbf{r}_i)$  the exchange coupling strength, does not average to zero in regions with preferred spin alignment  $\mathbf{s}_i$  because of the polarization  $\mathbf{P}$  of the incident beam ( $\mathbf{s}$ ). If  $\mathbf{M}$  is the

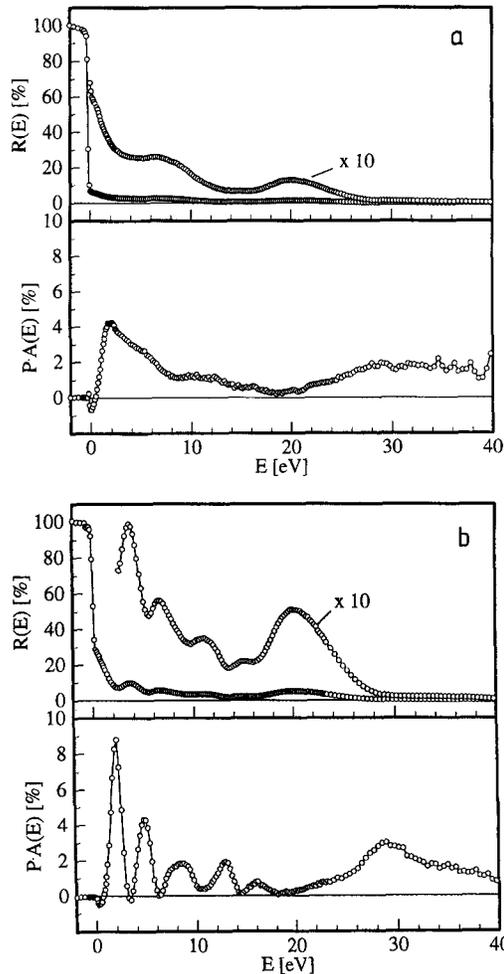
magnetization resulting from the preferred alignment of the target spins  $s_j$ , then  $V_{\text{ex}} \sim \mathbf{P} \cdot \mathbf{M}$ , which causes a  $\mathbf{P} \cdot \mathbf{M}$ -dependent contribution to the reflected intensity  $I = I_0 + I_{\text{ex}}$ . If  $\mathbf{P}$  is reversed,  $V_{\text{ex}}$  and  $I_{\text{ex}}$  change sign while the polarization-independent contribution  $I_0$  is unaffected. The difference of the intensities  $I^\pm = I_0 + I_{\text{ex}}^\pm$  of images taken with  $\pm \mathbf{P}$ , usually normalized with the sum of  $I^\pm$  and with the degree of polarization  $P = |\mathbf{P}| \leq 1$ ,  $(I_{\text{ex}}^+ - I_{\text{ex}}^-)/(I^+ + I^-)P = I_{\text{ex}}^+/I_0P = A_{\text{ex}}$ , is called the exchange asymmetry, and gives an image of the  $\mathbf{M}$  distribution in the sample. The direction of  $\mathbf{M}$  can be easily determined by maximizing  $A_{\text{ex}}$  by rotating  $\mathbf{P}$  parallel/antiparallel to  $\mathbf{M}$ . Extraction of the magnitude of  $\mathbf{M}$  requires an  $A_{\text{ex}}$  analysis in terms of a dynamical SPLEED theory or an empirical calibration.

SPLEEM is particularly useful for the study of crystalline ferromagnetic materials. These have a spin-dependent band structure, with majority-spin and minority-spin bands usually separated by a few tenths of an electron volt (eV) to about 1 eV. Figure 1 shows such an exchange-split band structure [6]. The



**Figure 1.** Band structure of cobalt above the vacuum level in the [0001] direction. The Fermi energy is 5.3 eV below the vacuum level [6].

[0001] direction is a frequently encountered orientation in cobalt layers. Below the two energy bands is a large energy gap. Electrons with energies in such a gap cannot propagate in the crystal and are totally reflected (see also Chap. 4, Sec. 1.6 of this Volume). This is true for majority-spin electrons up to 1 eV, and for minority-spin electrons up to 2 eV. Between these two energies there is an increasing excess of minority-spin electrons in the reflected intensity because majority-spin electrons can penetrate into the crystal. Thus,  $A_{\text{ex}}$  and, therefore, magnetic contrast are large. Above 2 eV,  $A_{\text{ex}}$  rapidly decreases because both types of electron can now penetrate into the crystal. This is clearly seen in Fig. 2a, in which the intensity  $I_{00}$  and measured exchange asymmetry  $PA_{\text{ex}}$  of the specularly reflected beam from a thick [0001]-oriented single crystalline cobalt layer are plotted as functions of energy [7]. One of the currently most important fields in magnetic materials is the study of ultrathin magnetic layers. These layers frequently show pronounced quantum size effect (QSE) oscillations in  $A_{\text{ex}}$ . They can be understood by inspecting Fig. 1: to every energy  $E$  above 2 eV there are two  $k \equiv k_{\perp}$ . Although  $k_{\perp}$  is not a good quantum number in very thin films, it is still defined well enough to allow thickness-dependent standing waves in the layer, which occur at different thickness for fixed  $E$  or at different  $E$  for fixed thickness for majority-spin and minority-spin electrons, causing oscillations and a significant enhancement of  $A_{\text{ex}}$  as seen in Fig. 2b. Such standing waves in the layer are formed not only in energy regions in which the substrate has a band gap but generally whenever there is poor matching of the wave functions at the interface. The



**Figure 2.** Specular intensity and exchange asymmetry of [0001]-oriented cobalt layers. (a) Thick layer and (b) six monolayers on a W(110) surface [7].

QSE effect in  $A_{cx}$  makes determination of the magnitude of  $M$  difficult, but is very helpful for contrast enhancement.

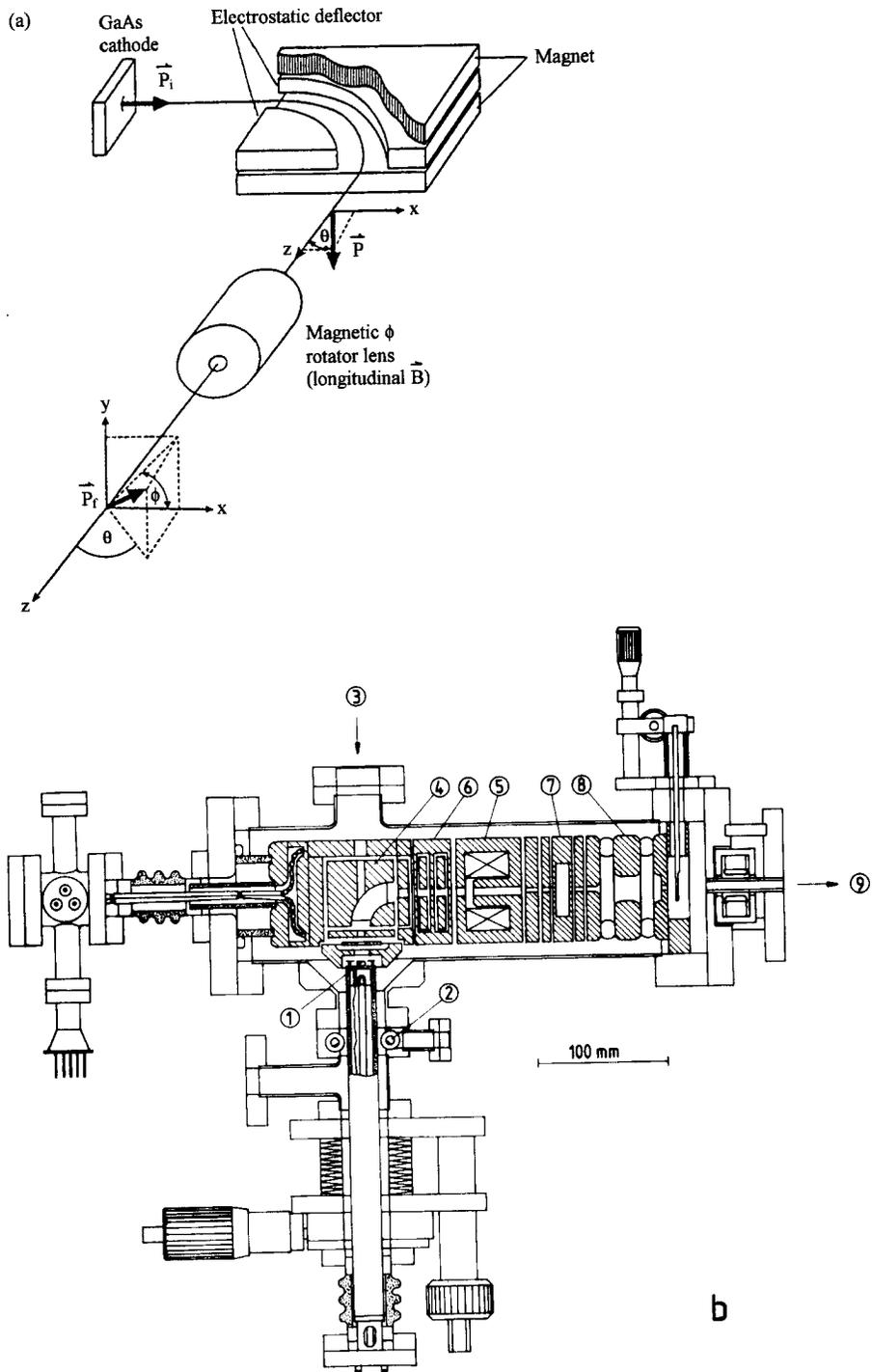
The information depth of SPLEEM is the same as in LEEM if spin-dependent inelastic scattering is neglected. Then the inverse attenuation length of electrons with energies of a few electronvolts above the vacuum level in d metals is  $1/\lambda \approx 1 + 0.8(5 - n) \text{ nm}^{-1}$ , where  $n$  is the number

of occupied d orbitals of one spin state [8]. Thus, there is no universal behavior but rather a pronounced material dependence. For cobalt ( $n = 4.1$ )  $\lambda \approx 6 \text{ \AA}$ , for iron ( $n = 3.5$ )  $\lambda = 4.5 \text{ \AA}$ , while for gold ( $n = 5$ )  $\lambda = 10 \text{ \AA}$ . In addition to the limitation by inelastic scattering, the information depth is strongly limited in bandgaps, for example on Co(0001) below 1 eV.

### 1.9.3 Instrumentation

A SPLEEM instrument is very similar to a LEEM instrument (see Chap. 4, Sec. 1.6 of this Volume) but differs in the illumination system and in the cathode objective lens. The magnetic field of most magnetic lenses at the specimen position is usually large enough (of the order of  $10^{-4} \text{ T}$ ) to influence magnetic specimens. Therefore, an electrostatic lens is called for, preferentially a tetrode lens which has a resolution comparable to that of a magnetic lens (see Chap. 4, Sec. 1.6 of this Volume). The magnetic sector is unavoidable, but disturbs the polarization of the incident beam only a little because of the low field and short path in it. The illumination system is usually electrostatic, and incorporates the polarization manipulator which is needed to adjust  $P$  collinear or normal to  $M$ . In the present instruments [9, 10] it consists of a crossed  $B-E$  field  $90^\circ$  deflector and a magnetic field lens as schematically shown in Fig. 3a [11]. The actual configuration of the illumination system is shown in Fig. 3b [10].

The spin-polarized electrons are produced by photoelectron emission from GaAs by left or right circular polarized light from a diode laser. The GaAs(100)



**Figure 3.** Polarization manipulator. (a) Functional scheme [9]. (b) Configuration with  $90^\circ$  sector and a magnetic field lens at high voltage [10]: 1, GaAs cathode; 2, cesium dispensers; 3, laser beam; 4,  $E-B$  field  $90^\circ$  deflector; 5, magnetic rotator lens; 6, double deflector; 7, double condenser; 8, immersion transfer lens; 9, to beam separator.

single crystal surface is covered with a Cs–O layer in order to approach a negative electron affinity surface. From this surface the electrons are excited into the spin–orbit–split conduction band of GaAs, from which they can escape without having to overcome a barrier. The maximum theoretical polarization obtainable from such an emitter is 50% but in practice 20–25% is usual. Contrary to the LaB<sub>6</sub> emitter used in LEEM which has thousands of operating hours, the activation of GaAs emitters has a useful life of the order of hours to days depending upon the vacuum, and activation is still more an art than a science. Strained alloy multilayers on GaAs basis are now available with much higher polarizations, but the current and reliability are still inferior to the standard GaAs emitter.

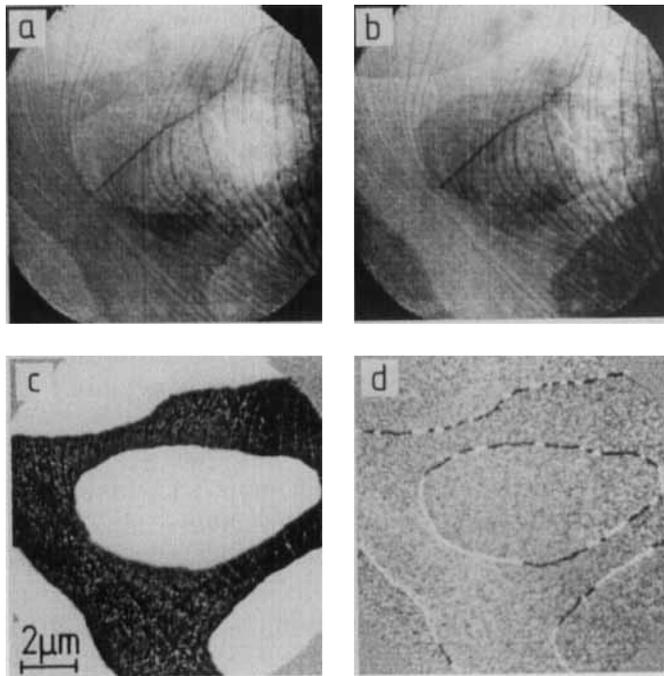
The resolution in SPLEEM is comparable to that in LEEM, but the signal-to-noise ratio is poorer due to the need for image subtraction if purely magnetic images are desired. Inasmuch as  $I_{\text{ex}} \ll I_0$ , longer exposures (in the 1 s range) are needed for a good signal-to-noise ratio. Further instrument development should allow image acquisition times of the order of 0.1 s. The field of view is up to about 50  $\mu\text{m}$ , similar to LEEM.

### 1.9.4 Areas of Application

The magnetic domain structure on the surface of bulk magnetic materials is generally too large grained for SPLEEM studies. Exceptions are, for example, closure domains on Co(0001) surfaces or domain walls [12]. The major field of SPLEEM is, therefore, the study of thin

films and superlattices whose magnetic properties depend strongly upon film thickness and structure, in particular upon interface structure.

One of the questions is to what extent the substrate influences the magnetic properties. Cobalt layers on W(110) illustrate this influence. On clean W(110), cobalt grows with the closest-packed plane parallel to the substrate, with h.c.p. stacking, at least at seven monolayers ((0001) orientation), on the W<sub>2</sub>C-covered surface in f.c.c. packing with the (100) orientation [13]. In both cases the magnetization is in-plane and has a pronounced uniaxial anisotropy with the easy axis in the h.c.p. layer parallel to W[110] [14]. In both cases good magnetic contrast is obtained at three monolayers, although the magnetic moment is still small at this thickness. The magnetic domains are initially small but rapidly coalesce with increasing film thickness [13]. Steps on the substrate surface have no influence on the size and shape of the domains and apparently neither on the location of the vertices in the Néel-type domain walls which can be imaged with  $\mathbf{P}$  perpendicular to  $\mathbf{M}$ . Sometimes the vertices are located at steps seen in the unsubtracted images of Fig. 4, but more often they occur on terraces [15]. In contrast to cobalt on W(110), which is in-plane magnetized from the very beginning, cobalt on Au(111) has initially out-of-plane magnetization which switches to in-plane at 4–5 monolayers. The exact thickness at which this transition occurs depends upon the state of the Au(111) surface. In particular, when cobalt is grown on an Au(111) double layer on W(110), the step distribution can have a strong influence on domain size and shape [16a].

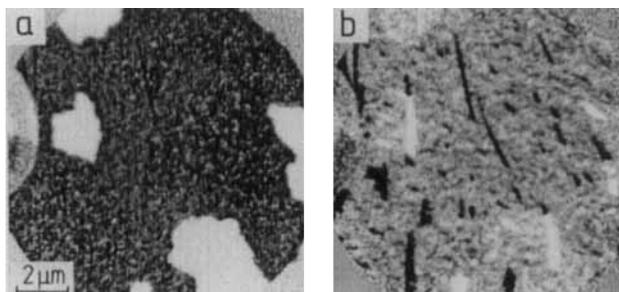


**Figure 4.** Magnetic domain images of a six monolayer thick cobalt layer on W(110). Images with  $P$  (a) parallel and (b) antiparallel to  $M$ , respectively. (c) Difference image between (a) and (b). (d) Difference image between two images with  $P$  perpendicular to  $M$ . Electron energy 2 eV [15].

The in situ study capabilities of SPLEEM instruments (vapor deposition, heating, cooling, etc.) also allow the study of changes in the magnetic structure as a function of film thickness or temperature. For example, during growth of cobalt on clean W(110) at an elevated temperature, for example 700 K, flat three-dimensional single-domain cobalt crystals form on a cobalt monolayer. When the crystals grow to join each other with increasing thick-

ness, the domains rearrange, which can be followed by SPLEEM quasi-life [14]. Another example is the change in magnetization when a continuous cobalt layer breaks up into small crystals during annealing [14].

Nonmagnetic overlayers on magnetic layers can have a strong influence on their magnetic structure. An example is gold on Co(0001) layers. Figure 5a is the SPLEEM image of a 10 monolayer thick

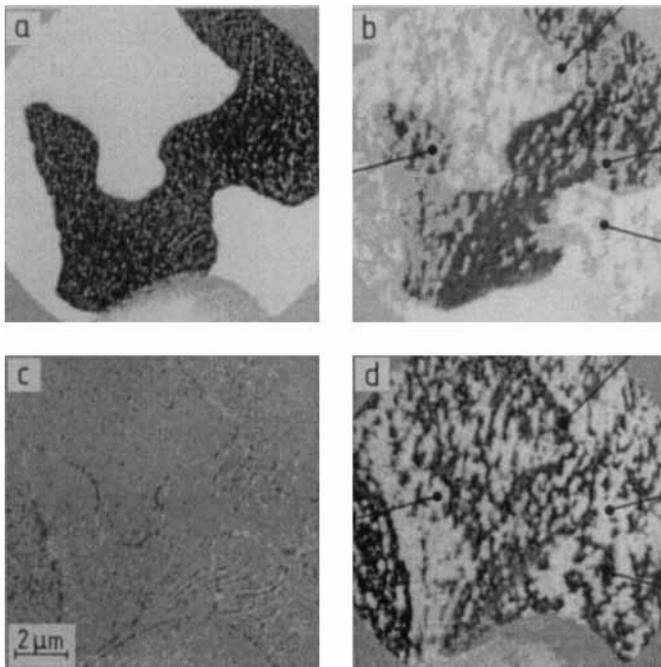


**Figure 5.** Influence of a gold overlayer on the magnetic structure of a 10 monolayer thick cobalt layer on W(110). (a) SPLEEM image of the cobalt layer and (b) SPLEEM image of the same area covered with two monolayers of gold. Electron energy 1.5 eV [16b].

cobalt layer deposited on W(110) at 400 K. One monolayer of gold deposited onto this cobalt layer at room temperature has no influence on the domain structure, but when two gold monolayers are deposited, most of the cobalt layer switches to out-of-plane magnetization—so that  $M \perp P$  and  $A_{\text{ex}} = 0$ —but on some terraces the original in-plane magnetization remains (Fig. 5b), even if additional gold is deposited [16b]. Figure 5 illustrates very well how the topography of the substrate can be propagated through a deposited layer and influence its magnetic properties. In contrast to gold, copper overlayers have no influence on the domain structure in cobalt layers up to the largest overlayer thickness studied (14 layers). Only pronounced QSE  $A_{\text{ex}}$  oscillations are seen as a function of copper layer thickness [17].

One of the most interesting subjects in magnetism has been for some time the

magnetic coupling between magnetic layers through nonmagnetic layers. Frequently it oscillates between ferromagnetic (FM) and antiferromagnetic (AFM) coupling, and the causes of the AFM coupling and its occasional absence has been the subject of much discussion. Here, SPLEEM can make an important contribution because of its significant information depth at very low energies. As an example, Fig. 6 shows SPLEEM images of a cobalt layer on W(110) and of a cobalt–copper–cobalt sandwich grown on it by sequential deposition of four monolayers of copper and six monolayers of cobalt. According to most magnetic studies, the two cobalt layers should be antiferromagnetically coupled through four monolayers of copper. Analysis of many SPLEEM images similar to Fig. 6, however, shows that there is only locally occasional AFM coupling, but also FM



**Figure 6.** Magnetic coupling between two cobalt layers through a copper layer. (a, c) Bottom cobalt layer only and (b, d) complete cobalt–copper–cobalt sandwich.  $P \parallel M$  in (a) and  $P \perp M$  in (b) in bottom layer. The solid circles denote identical domains in the top layer. Electron energy 1.5 eV [15].

coupling, and other  $M$  directions in the top layer, which also has much smaller domains [15]. These domains extend frequently across the boundaries of the domains in the lower cobalt layer which can be seen through the copper and the top cobalt layer.

### 1.9.5 Discussion

When should SPLEEM be used for the study of the magnetic microstructure of materials, and when not? Section 3.4 of this Chapter has already mentioned a number of interesting applications, but there are many more such as the study of magnetic phase transitions as a function of temperature, magnetic switching phenomena in pulsed fields, or magnetization processes in fields perpendicular to the surface. Fields parallel to the surface can only be applied in a pulsed manner because of the beam deflection which they cause, while fields normal to the surface require only refocusing. The areas for which SPLEEM is not well suited are essentially the same as in LEEM (see Chap. 4, Sec. 1.6 of this Volume): rough surfaces, high vapor pressure materials, etc.

There are many other magnetic imaging methods. Most of them do not image magnetization but the internal or external magnetic field distribution caused by the magnetization distribution such as Lorentz microscopy (see Chap. 4, Sec. 1.8 of this Volume), electron holography (see Chap. 4, Sec. 1.9 of this Volume), or magnetic force microscopy (see Chap. 7, Sec. 3 of this Volume). These techniques are to a large extent complementary to

SPLEEM, as is MEM [18], which can be easily combined with LEEM. The most important competitor to SPLEEM is scanning electron microscopy with polarization analysis (SEMPA; see Chap. 5, Sec. 2 of this Volume), which images the magnetization distribution in the specimen via polarization analysis of the secondary electrons. The advantages of SEMPA over SPLEEM are an easy combination with electron spectroscopy, unimportance of the crystallinity of the specimen for high brightness, and the absence of a high electric field at the specimen surface. Some of the advantages of SPLEEM over SEMPA are rapid image acquisition, and easy combination with LEED and various emission microscopies.

### 1.9.6 Concluding Remarks

At present, SPLEEM is still too young a technique to allow an extensive discussion of its advantages and disadvantages. Only two instruments are in operation at the time of writing. Commercial instruments will probably not be available before 1997.

### 1.9.7 References

- [1] J. Kessler, *Polarized Electrons*, 2nd ed., Springer, Berlin **1985**.
- [2] D. T. Pierce, R. J. Celotta, *Adv. Electron. Electron Phys.* **1981**, *65*, 219.
- [3] R. Feder, *J. Phys. C*, **1981**, *14*, 2049; R. Feder, *Phys. Ser.* **1983**, *74*, 47.
- [4] R. Feder (Ed.), *Polarized Electrons in Surface Physics*, World Scientific, Singapore **1985**.
- [5] J. Kirschner, *Polarized Electrons at Surfaces*, Springer, Berlin **1985**.
- [6] J. Noffke, personal communication.

- [7] K. Wurm, M.S. thesis, TU Clausthal **1994**.
- [8] H. C. Siegmann, *J. Phys.: Condens. Matter* **1992**, *4*, 8395.
- [9] K. Grzelakowski, T. Duden, E. Bauer, H. Poppa, S. Chiang, *IEEE Trans. Mag.* **1994**, *30*, 4500.
- [10] T. Duden, Ph.D. thesis, TU Clausthal 1996.
- [11] T. Duden, E. Bauer, *Rev. Sci. Instrum.* **1995**, *66*, 2861.
- [12] M. S. Altman, H. Pinkvos, J. Hurst, H. Poppa, G. Marx, E. Bauer, *MRS Symp. Proc.* **1991**, 232, 125.
- [13] H. Pinkvos, H. Poppa, E. Bauer, J. Hurst, *Ultramicroscopy* **1992**, *47*, 339.
- [14] H. Pinkvos, H. Poppa, E. Bauer, G.-M. Kim in *Magnetism and Structure in Systems of Reduced Dimensions* (Eds.: R. F. C. Farrow, B. Dieny, M. Donath, A. Fert, B. D. Hermsmeier), Plenum, New York **1993**, p. 25.
- [15] E. Bauer, T. Duden, H. Pinkvos, H. Poppa, K. Wurm, *J. Magn. Magn. Mat.*, in press.
- [16] (a) M. S. Altman, H. Pinkvos, E. Bauer, *J. Magn. Soc. Jpn.* **1995**, *19* (Suppl. S1), 129; (b) M. S. Altman, H. Pinkvos, E. Bauer, unpublished findings.
- [17] H. Poppa, H. Pinkvos, K. Wurm, E. Bauer, *MRS Symp. Proc.* **1993**, *313*, 219.
- [18] H. Bethge and J. Heydenreich, *Elektronenmikroskopie in der Festkörperphysik*, Springer, Berlin **1982**, p. 196.

## **2 Scanning Beam Methods**

## 2.1 Scanning Electron Microscopy

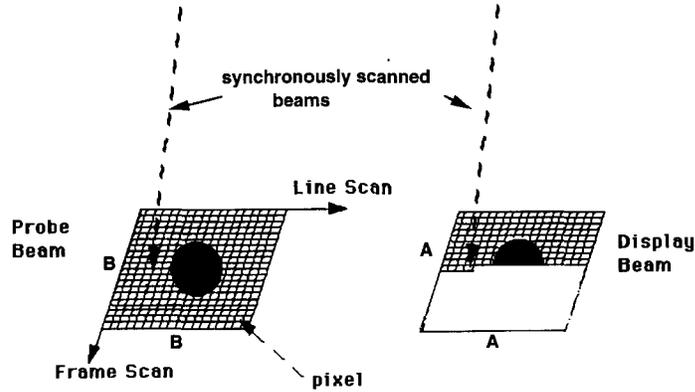
### 2.1.1 Introduction

The scanning electron microscope (SEM) is the most widely used of all electron beam instruments. It owes its popularity to the versatility of its various modes of imaging, the excellent spatial resolution of its images, the ease with which the micrographs that are generated can be interpreted, the modest demands that are made on specimen preparation, and its 'user-friendliness'. At one end of its operating range the SEM provides images which can readily be compared to those of conventional optical microscopes, while at the other end its capabilities are complementary to instruments such as scanning tunneling (STM) or atomic force (AFM) microscopes. While its resolution can now approach 0.5 nm, rivaling that of a transmission electron microscope, it can handle specimens as large as production size silicon wafers.

The SEM had its origins in the work of von Ardenne [1, 2] who added scanning coils to a transmission electron microscope. A photographic plate beneath the electron transparent sample was mechanically scanned in synchronism with the beam to produce the image. The first recognizably modern SEM was described by Zworykin et al. [3]. This instrument

incorporated most of the features of current instruments, such as a cathode-ray-tube display and a secondary electron detector, and achieved a resolution of 5 nm on solid specimens. In 1948 Oatley [4] and his students commenced their work on the development of the SEM leading in 1965 to the first commercial machine the Cambridge Scientific Instruments Mark 1 'Stereoscan'. There are now seven or eight manufacturers of these instruments in Europe, the USA, and Japan, and it is estimated that about 20 000 SEMs are in use worldwide.

The SEM is a mapping, rather than an imaging, device (Fig. 1) and so is a member of the same class of instruments as the facsimile machine, the scanning probe microscope, and the confocal optical microscope. The sample is probed by a beam of electrons scanned across the surface. Radiations from the specimen, stimulated by the incident beam, are detected, amplified, and used to modulate the brightness of a second beam of electrons scanned, synchronously with the first beam, across a cathode ray tube display. If the area scanned on the display tube is  $A \times A$  and the corresponding area scanned on the sample is  $B \times B$  then the linear magnification  $M = A/B$ . The magnification is therefore geometric in origin and may be changed by varying the area



**Figure 1.** Schematic illustration of the basic mapping principle of the scanning electron microscope.

scanned on the sample. The arrangement makes it possible for a wide range of magnifications to be obtained, and allows rapid changes of magnification since no alterations to the electron-optical system are required. There is no rotation between the object and image planes, and once the instrument has been focused on a given area the focus need not be changed when the magnification is varied. To a first approximation the size of the finest detail visible in the image will be set by the size of the probe scanning the specimen. Multiple detectors can be used to collect several signals simultaneously which can then be displayed individually, or combined, in perfect register with each other. It is this capability in particular which makes the SEM so useful a tool since multiple views of a sample, in different imaging modes, can be collected and compared in a single pass of the beam.

### 2.1.2 Instrumentation

Figure 2 shows the basic components of the SEM. These can be divided into two main categories, the electron-optical and

detector systems, and the scanning, processing, and display systems. The electron-optical components are often described as being the 'column' of the instrument while the other items are the 'console' of the machine. The source of electrons is the gun which produces them either thermionically, from tungsten or lanthanum hexaboride cathodes, or from a field emission source. These electrons are then accelerated to an energy which is typically in the range from 500 eV to 30 keV. The beam of electrons leaving the gun is then focused on to the specimen by one or more condenser lenses. Although either electrostatic or electromagnetic lenses could be employed all modern SEMs use electromagnetic lenses. Typically, the final objective lens has been of the pin-hole design with the sample sitting outside the magnetic field of the lens since this arrangement gives good physical access to the specimen. However, in this arrangement the specimen is 10 to 20 mm away from the lens which must therefore be of long focal length and correspondingly high aberration coefficients. In modern, high performance, instruments it is now common to use an immersion lens [5], in which the sample sits inside the lens at the center of

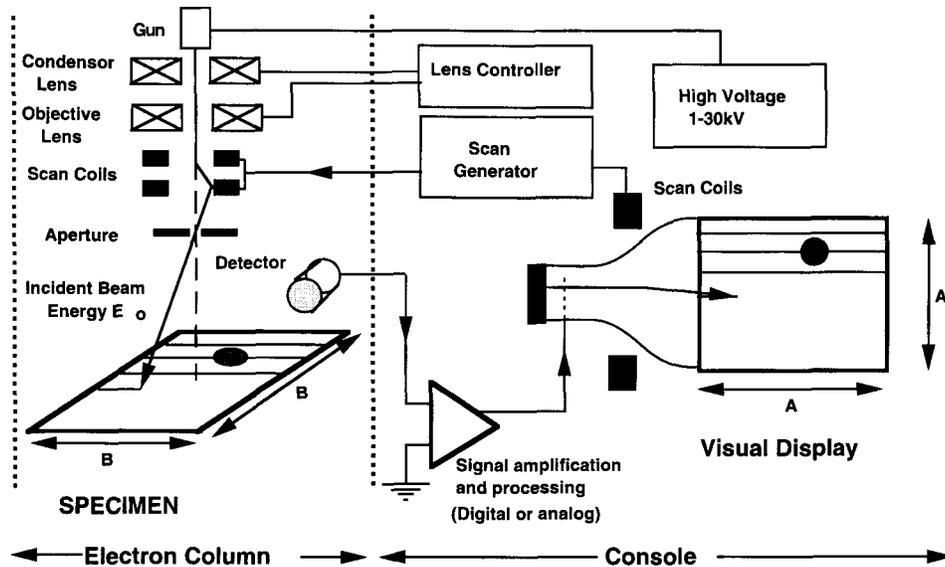


Figure 2. Basic components of the scanning electron microscope.

the lens field, or a 'snorkel' lens [6] in which the magnetic field extends outside of the lens to envelope the sample. Although the immersion lens gives very good performance and, by making the sample part of the lens structure, ensures mechanical stability, the amount of access to the specimen is limited. The snorkel lens on the other hand combines both good electron-optical characteristics with excellent access for detectors and stage mechanisms.

The coils that scan the beam are usually incorporated within the objective lens. A double scan arrangement is often employed in which one set of coils scans the beam through some angle  $\theta$  from the axis of the microscope while a second set scans the beam through an angle  $2\theta$  in the opposite direction. In this way all scanned beams pass through a single point on the optic axis allowing for the placement of a defining aperture without any constriction

of the scanned area. The scan pattern, or 'raster', produced on the specimen, is usually square in shape and is made up of 1000 horizontal lines each containing 1000 individual scanned points or 'pixels'. The final image frame thus contains  $10^6$  pixels, although for special activities such as focusing or alignment frames containing only  $256 \times 256$  pixels may be used. A variety of detectors are provided in the SEM, including an Everhart-Thornley [7] scintillator detector for secondary electrons, some type of detector for back-scattered electrons and, often, a detector for fluorescent X-rays to permit chemical microanalysis with the beam. Further details of these devices are given below.

Signals from any of the detectors are amplified and presented to the display screens in the console. The electronics provide control of the amplification, DC offset, and bandwidth of the signal. Increasingly the detector output is passed

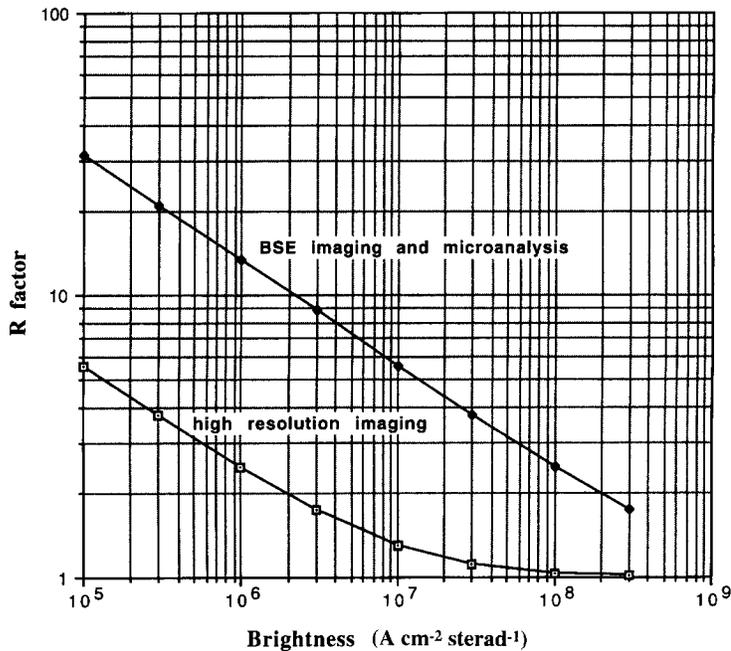
through an analog to digital converter (ADC) and then handled digitally rather than as an analog video signal. This permits images to be stored, enhanced, combined, and analyzed using either an internal or an external computer. Although the majority of images are still recorded on to photographic film, digital images can be stored directly to magnetic or magneto-optic discs and hardcopy output of the images can then be obtained using laser or dye-sublimation printers. The scan drive currents to the scan coils may also be digitized to provide precision control of the beam position although analog scans are still most commonly encountered. Typically scan repetition rates ranging from 15 or 20 frames per second (TV rate) to one frame in 30 to 60 s (photographic rate) are provided. In addition individual pixels, or arrays of pixels, within an image field may be accessed if required.

### 2.1.3 Performance

It is usual to define the performance of a microscope in terms of the spatial resolution that it can produce. In the case of the SEM the attainable resolution is determined by a number of factors including the diameter  $d$  of the electron-beam probe that can be generated, the current  $I_b$  contained in that probe, the magnification of the image, and the type of imaging mode that is being used. Over most of the operating energy range (5–30 keV) of the SEM the probe size and beam current are related by an expression of the form [8]

$$d = (C_s^{1/4} \lambda^{3/4}) \left[ 1 + \frac{I_b}{\beta \lambda^2} \right]^{3/8} \quad (1)$$

where  $\lambda$  is the wavelength of the electrons ( $\lambda \approx 1.226 E_0^{-1/2}$  nm where  $E_0$  is the incident electron energy in eV),  $\beta$  is the brightness of the electron gun in  $\text{A cm}^{-2} \text{steradian}^{-1}$ , and  $C_s$  is the spherical aberration coefficient of the objective lens. Equation (1) contains two groups of terms; the first of which ( $C_s^{1/4} \lambda^{3/4}$ ) can be taken as being the minimum spot size that the lens can produce; the second term then multiplies this limiting probe size by a factor which depends on the current  $I_b$  required in the beam. Note from Eq. (1) that the obtainable spot size  $d$  is always larger than the minimum value except in the limiting case when  $I_b$  is zero, so for any finite beam current the probe size must be enlarged and, hence, the spatial resolution of the image will be worsened. The degree to which this occurs will depend both on the current  $I_b$  required and on the gun brightness  $\beta$ . Figure 3 plots the value of this multiplying factor as a function of gun brightness for two beam currents, firstly  $I_b = 10 \times 10^{-12}$  A, a typical value for high resolution secondary electron imaging, and secondly  $I_b = 1 \times 10^{-9}$  A, a value suitable for backscattered imaging, electron channeling, or other special modes of imaging. For a brightness of  $10^5 \text{ A cm}^{-2} \text{ sr}^{-1}$ , such as encountered with a tungsten cathode thermionic gun [9], the multiplying factor is 5.6 for imaging currents and over 30 for the higher current case. In such a case the imaging performance of the SEM is dominated not by the lens but by the relatively poor performance of the electron gun. If the brightness is increased to  $10^6 \text{ A cm}^{-2} \text{ sr}^{-1}$ , for example, by using a lanthanum hexaboride cathode in the gun [9], then the multiplying factors become 2.5 and 1.3, respectively, showing that improving the performance of the gun



**Figure 3.** *R* factor modifying minimum probe size against brightness for high resolution SE imaging ( $I_B = 10^{-11}$  A) and backscattered imaging or microanalysis ( $I_B = 10^{-9}$  A).

dramatically enhances the resolution of the microscope. Finally, if the gun brightness is further increased to  $10^8 \text{ A cm}^{-2} \text{ sr}^{-1}$  by using a field emission source [9] then the factor is close to unity for both modes of operation considered. In this case the probe forming performance is no longer limited by the brightness of the source but is controlled by the properties of the lens. (Note that since electron optical brightness  $\beta$  increases linearly with beam energy  $E_0$ , while the electron wavelength  $\lambda$  varies as  $E_0^{-1/2}$ , the value of the quantity  $\beta\lambda^2$  is independent of the energy actually used.) For a modern SEM  $C_s$  is typically a few millimeters; thus minimum probe sizes of 1 or 2 nm are available. At low beam energies (below 5 keV) additional effects including the energy spread of electrons in the beam must also be considered, but the general conclusions discussed above remain correct.

A second limitation to the performance comes from the fact that the scan raster divides the image into a finite number of pixels, typically 1000 lines each of 1000 pixels. If the size of the display CRT is  $A \times A$ , and the instrumental magnification is  $M$  then the size of each pixel referred to the sample is  $A/(1000M)$ . For example, if  $A$  is 10 cm and  $M \approx 20$ , the effective pixel size on the sample is then 5  $\mu\text{m}$ . Each pixel represents a single piece of information in the image, so no detail smaller than the pixel size can be resolved even though the probe size may be significantly smaller than this value. At low magnifications the SEM is therefore pixel limited in its resolution; for example, it is not until a magnification of 10 000 $\times$  that the pixel size falls below 10 nm and so becomes comparable in size to the probe diameter discussed above.

Finally, the spatial resolution of an image may be limited by the lateral spread

of the electron interactions that produce the desired image information. The extent of this varies widely with the mode of operation and with the nature of the specimen, from a nanometer or less for secondary electron imaging in some circumstances, to a micrometer or more for backscattered or EBIC imaging. In summary, therefore, the resolution of the SEM cannot be defined by a single number but depends on many different factors. In the past the performance of the SEM has been mainly limited by gun brightness, lens quality, and other instrumental parameters, but with current instruments employing field emission guns and immersion lenses fundamental electron interaction effects now probably define the performance more than any specific attribute of the microscope itself.

## 2.1.4 Modes of Operation

### 2.1.4.1 Secondary Electron Imaging

Secondary electrons (SE) are those electrons emitted by the specimen, under irradiation by the beam, which have energies between 0 and 50 eV. Because of their low energy the SE only travel relatively short distances in the specimen (3–10 nm) and thus they emerge from a shallow ‘escape’ region beneath the surface. There are two cases in which an SE can be generated and subsequently escape from the specimen: first, when an incident electron passes downwards through the escape depth, and secondly, as a backscattered electron leaves the specimen and again passes through the escape region. Secondary

electrons produced in the first type of event are designated SE1 and, because they are generated at the point where the incident beam enters the specimen, it is these which carry high resolution information. The other secondaries are called SE2, and these come from a region whose size is of the order of the incident beam range in the sample. Since this can be 1  $\mu\text{m}$  or more at high energies and it can be seen that the SE2 carry low resolution information. The SE1 and SE2 signals cannot be separated by any device because they are identical in their properties but when the SEM is operated at high magnifications the area scanned by the beam is less than the area from which the SE2 signal is generated so the SE2 signal is effectively independent of the beam position and forms a constant background to the SE1 signal. The variations in the SE1 signal can then be isolated from the SE2 variations, although since the SE1 component is only about one-half to one-third as strong as the SE2 component this means that the high resolution information is diluted by the low resolution background. The yield  $\delta$  of secondary electrons ( $\delta$  = number of SE per incident electron) varies with the energy of the incident electron beam. At high energies (10 keV or more) the yield is typically only 0.1 to 0.2, but as the beam energy is reduced the yield rises rapidly and may exceed unity for energies of the order of 1 keV. SE imaging is, therefore, preferentially performed at a low rather than a high accelerating voltage.

Secondary electron imaging is the most common mode of operation of the SEM and it has been estimated that 95% of all published SEM images have been recorded with the SE signal. The importance and utility of this mode is the result

of several factors:

- (i) Secondary electrons are easy to collect.
- (ii) Secondary electrons carry information about the surface topography of the specimen. Information about surface chemistry, and magnetic and electric fields may also be obtainable on suitable specimens.
- (iii) SE images can, in most cases, be interpreted readily without specialist knowledge.
- (iv) The SE image can display information at a spatial resolution of 1 nm or better under properly optimized conditions.

The practical key to the success of SE imaging has been the detector originally described by Everhart and Thornley [7], shown in Fig. 4. The secondary electrons are allowed to strike a scintillator material after having been accelerated to an energy

of about 10 keV by a positive potential applied to the front face of the detector. To prevent this potential deflecting the incident electron beam the scintillator is often surrounded by a Faraday cage biased to about +200 V to attract the secondaries and made of open metal mesh. The light produced in the scintillator, which is usually either a plastic or a rare-earth doped YAG crystal, is transferred down a light-pipe made of quartz and leaves the vacuum chamber of the SEM through a window. On the other side of the window the light enters a photomultiplier tube where it is reconverted to an electronic signal. This arrangement produces a very high amplification of the original SE signal while only adding insignificant amounts of noise, a wide dynamic range (because of the logarithmic characteristic of the photomultiplier tube), and a rapid response to changes in the signal intensity. It is cheap

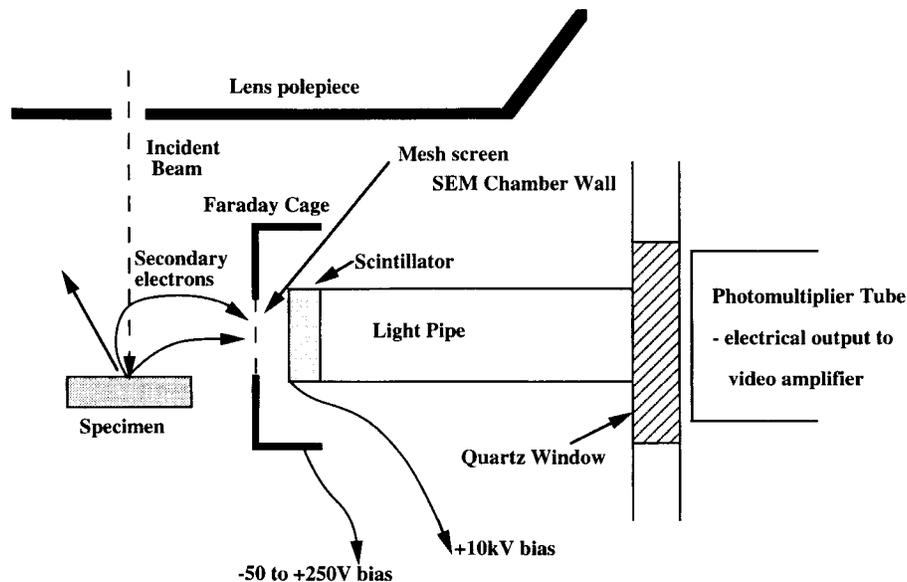
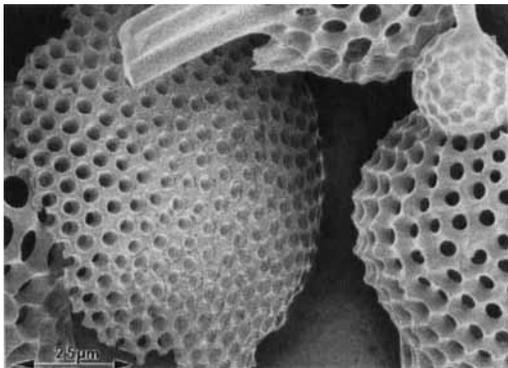


Figure 4. Everhart-Thornley SE detector and specimen chamber geometry.

to produce, and the detector may be placed anywhere within the specimen chamber of the SEM and still collect 50 to 60% of the SE leaving the specimen. On some advanced instruments the SE detector may be placed above the objective lens and out of line of sight of the sample [10]. The magnetic field of the lens then collects the SE which pass back through the lens and are then extracted by the detector. This arrangement provides good collection efficiency and is well suited for high resolution imaging.

The general characteristics of SE signals are illustrated in Fig. 5 which shows a collection of diatoms. The image has a marked three-dimensional effect caused by the lighting and shadow effects which decorate each object, and each edge is outlined by a fine bright line which adds definition to the detail. The ability to view topographic detail in the image occurs because the yield of secondary electrons varies with the angle between the incident electron beam and the local surface normal [9]. Areas which are at a high angle to the beam are bright (large signal) compared to



**Figure 5.** Secondary electron images of Radiolarium. Recorded in Hitachi S-4500 field emission SEM at 5 keV beam energy. Magnification: 800 $\times$ .

those faces which are normal to the beam (small signal), and faces looking towards the detector (which is in the top right hand corner of the micrograph) are in general brighter than those facing away from the detector although, because the SE are readily deflected by the electrostatic field from the detector, both faces are easily visible. This type of behavior is analogous to that of visible light as described by Lambert's cosine law [9]. If a source of light were placed on the SE detector then an observer looking down from the electron gun would see the brightest illumination on those surfaces tilted towards the light, and less light on those surfaces facing upwards towards the observer. Using this analogy it is easy to interpret secondary electron images in a reliable and consistent way. The bright edges in the image occur because secondary electrons can escape through two faces, rather than just one, in the vicinity of an edge and therefore the signal is anomalously high in such regions.

Even at much higher magnifications the appearance and interpretation of the SE image remain much the same. Figure 6 shows an image of the magnetic media on the surface of a computer hard disc. Although the detail is now only of a few nanometers in scale the image and its relationship to the surface topography can be understood in the same way as that described above. The fact that SE images can be consistently and easily interpreted over a wide range of magnifications has been a significant factor in the popularity of the SEM. This is in marked contrast to the situation on, for example, a scanning tunneling microscope (STM) which has similar spatial resolution but for which the multiplicity of possible effects contributing the image is such that

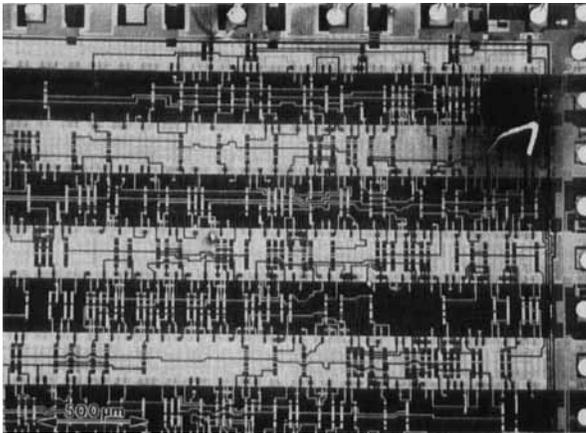


**Figure 6.** High resolution image of magnetic disc media surface recorded at 30 keV in Jeol JSM890 field-emission SEM.

interpretation is difficult and unreliable. At the highest resolutions (below 1 nm) the details of the electron–solid interactions that produce the SE must be considered and image interpretation becomes more complex and research is still in progress in this area [12].

Secondary electron images can also carry information about other properties of the specimen. One such mode is voltage contrast which is illustrated in Fig. 7. The micrograph shows a heart pacemaker chip, powered up and running but removed from its usual protective casing, which is

being examined in the SEM in the SE mode. In addition to the normal topographic contrast, large scale regions of uniform bright and dark contrast are visible. The bright areas are those which have a potential which is negative with respect to ground, while the dark areas are positive with respect to ground. The origin of such contrast is straightforward. When an area is negative then the collection field from the SE detector is increased and a higher fraction of the SE are collected. An area that is positive experiences a lower collection field from the detector, and also



**Figure 7.** Voltage contrast from integrated circuit. Recorded at 5 keV in Hitachi S-800 FE SEM.

has a tendency to recollect its own secondary electrons so the SE signal from such areas is lower. In this simple form the technique gives a qualitative view of static (DC) potential distributions but, by improvements in instrumentation, it is possible to study potentials which may be varying at frequencies up to 100 MHz or more [13], and to measure the potentials with a voltage resolution of  $\pm 10$  mV and a spatial resolution of  $0.1 \mu\text{m}$ . This unique ability of the SEM to measure voltages in real-time, from small areas, and without requiring any mechanical contact has been of considerable value in the development of semiconductor technology and represents a significant fraction of the usage of this instrumentation.

Contrast can also be generated by the presence of magnetic fields outside of a sample, for example, from the leakage fields that exist above the surface of a uniaxial magnetic material such as cobalt. In this case the contrast is produced by the Lorentz deflection of the SE after they leave the specimen. A field in one direction will deflect more SE towards the detector, while a field in the opposite sense will deflect the SE away from the detector [9]. In a more sophisticated application of this approach the incident electron beam is polarized, and the change in polarization of the emitted SE, measured by special detectors, is used to produce the signal contrast [14].

#### 2.1.4.2 Backscattered Electrons

Backscattered electrons (BSE) are defined as being those electrons emitted from the specimen which have energies between

50 eV and the incident beam energy  $E_0$ . Unlike the secondary electrons which are produced as the result of the incident electron irradiation, backscattered electrons are incident electrons which have been scattered through angles approaching  $180^\circ$  within the sample and consequently leave the sample again. The yield  $\eta$  of BSE ( $\eta$  = number of BSE per incident electron) varies monotonically with the atomic number  $Z$  of the specimen, increasing from about 0.05 for carbon to about 0.5 for gold. At high incident beam energies, therefore, the number of BSE produced is greater than the number of SE but despite this fact backscattered electron imaging has received less attention and use than SE imaging until recently. This is because of the practical problem of efficiently collecting the BSE. Since the energy of the BSE is of the order of  $E_0/2$  they are much more difficult to deflect towards a detector than the SE. Consequently the detector must be placed in a suitable position above the specimen to intercept the BSE, and must be physically large enough to collect a high fraction of the signal. Several successful detector designs are now in current use, including scintillator systems similar to the Everhart-Thornley SE detector discussed above, solid-state detectors, and electron-multiplier devices such as the microchannel plate.

Because the yield of BSE varies with the atomic number of the specimen the most widespread use of backscattered electrons is atomic number, or  $Z$ -contrast, imaging. Figure 8 shows a backscattered image of a sample of a 5000 year old Assyrian glass. Although the SE image showed the material as being homogeneous, the backscattered image reveals the presence of numerous precipitates and stringers

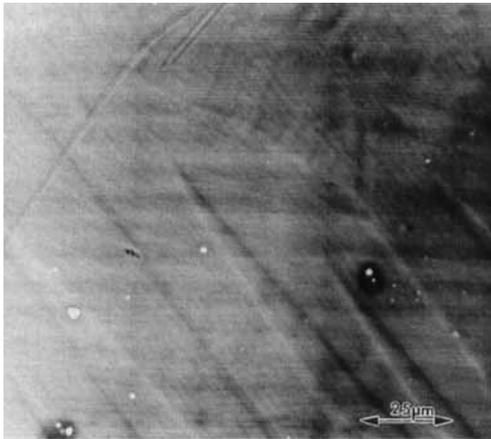


**Figure 8.** BSE Z-contrast image of Assyrian glass showing Cu<sub>2</sub>O dendrites. Recorded on Cambridge S250 SEM at 10 keV.

which were later determined to be Cu<sub>2</sub>O, probably dissolved from the vessel in which the glass was melted. The phases are readily distinguished here because the glass, essentially SiO<sub>2</sub>, for which the mean atomic number is 10, produces significantly fewer backscattered electrons than do the precipitates for which the mean atomic number is 22. In cases such as this BSE imaging provides a quick and convenient method of examining the distribution of chemistry within a material and qualitatively separating regions of high and low atomic number. The technique can also be used to examine diffusion gradients across a boundary, to scan for the presence of unexpected contaminants, and to look for evidence of multiphase structures. Atomic number contrast imaging has also been widely applied in biology [15] by attaching high atomic number elements such as gold to active molecular groups which preferentially bind at specified locations on a cell. A backscattered image of the material then shows the heavy metals as bright spots against the dark background of the predominantly carbon matrix. By comparing the BSE image with the SE image the

binding locations can be unambiguously located.

Despite several commercial attempts to justify the procedure, Z contrast imaging cannot be used as a substitute for proper microanalytical techniques. First, for all but pure elements, there are an infinite number of ways in which the same average atomic number can be achieved by combining different materials. Second, although it is widely assumed that the variation of backscattering yield with mean atomic number is monotonic for a compound there is no solid experimental evidence proving this and there are, in fact, examples where anomalous behavior has been reported [16]. Third, BSE detectors respond not only to the number of backscattered electrons but also to their energy and as a result two materials giving the same backscattered signal may have different backscattered yields compensated by the difference in the mean energy of the BSE. The technique should thus be regarded as a valuable diagnostic and observational mode but should not be used quantitatively without very careful preparation and calibration [17].



**Figure 9.** Magnetic contrast from domain walls in Fe-3%Si transformer core material. Recorded on Cambridge S250 at 30 keV.

The backscattering coefficient of a material can also be modified by other factors, including its magnetic configuration [18]. Figure 9 shows the backscattered image from Fe-3%Si transformer core material which has cubic magnetic anisotropy. The characteristic 'fir tree' magnetic domain structure of this type of material is clearly visible outlined by the black and white lines crossing the micrograph. The contrast in this case arises because of the Lorentz deflection of the incident beam within the specimen in the vicinity of the domain boundaries. For some directions of the magnetic flux across the domain the incident beam suffers a small additional downwards deflection which lowers the backscattering coefficient, while for flux in the opposite sense the incident electrons are deflected slightly upwards towards the surface and the backscattering yield is increased.

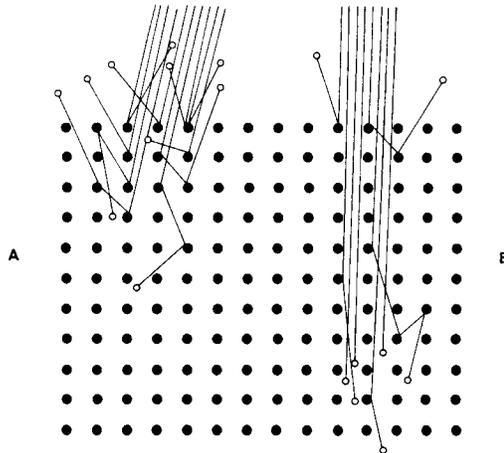
Unlike secondary electrons which come from the surface region of the specimen, backscattered electrons emerge from a volume which is of the order of one-third

of the incident beam range in depth and radius. To a first approximation, the range  $R$  (in nm) is given by

$$R = \frac{75E_0^{5/3}}{\rho} \quad (2)$$

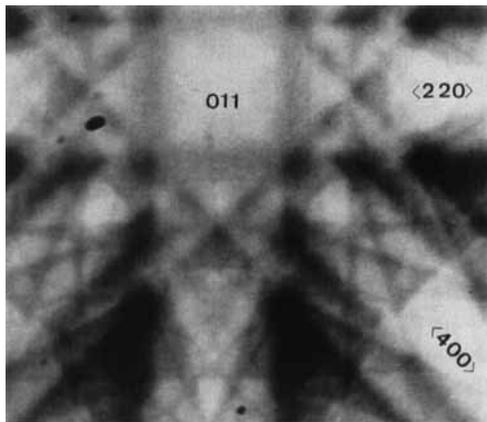
where  $E_0$  is the beam energy in keV and  $\rho$  is the density of the specimen in  $\text{g cm}^{-3}$ . Thus for beam energies of 10 keV, and typical densities, the backscattered signal comes from a region of the order of a fraction of a micrometer or more in breadth and depth. Consequently, the backscattered image contains information about the interior of the specimen rather than the surface but at the expense of the spatial resolution which is generally lower than that of the corresponding SE signal. For example, backscattered images can reveal the presence of buried voids or cracks in materials that are not visible at all in the SE image.

Unique to the backscattered image is information related to the crystalline nature of the sample. The origin of this contrast is illustrated schematically in Fig. 10. If the incident electron beam enters a crystal at a random angle of incidence then backscattering will occur in the normal way, and with the usual yield. But if the incident beam is aligned along a symmetry direction of the lattice then the incident electrons tend to channel between the lattice planes and as a result they penetrate more deeply into the specimen and the backscattering yield is reduced. Although this particle model is not physically realistic a detailed analysis using dynamical diffraction theory [19] confirms the existence of such a phenomenon. Thus if the angle of incidence between the electron beam and a crystal is varied then the backscattering yield is modulated at angles corresponding to the symmetry directions



**Figure 10.** Schematic illustration of the origin of electron channeling contrast (a) random incidence direction, (b) incidence along a symmetry direction.

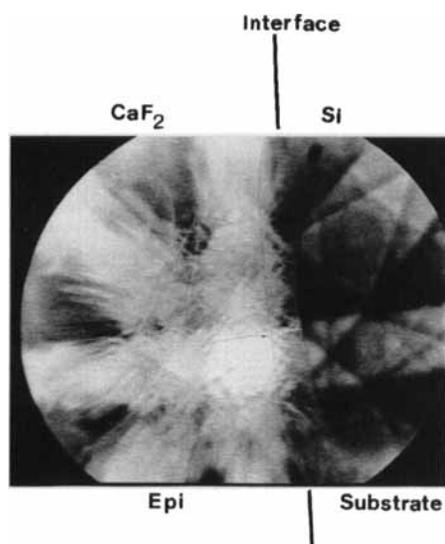
of the crystal. The image, or electron channeling pattern, produced in such a case is illustrated in Fig. 11 for a single crystal of InP. Here the variation in the angle of incidence has been achieved simply by scanning the large crystal at low magnification, so that between extremes of the field of view the angle between the beam and the surface normal varies by about  $\pm 8^\circ$ . The



**Figure 11.** Electron channeling pattern from InP recorded at 25 keV on Cambridge S250 SEM.

micrograph therefore contains both spatial and angular information. The small circular features scattered over the micrograph are conventional images of dust particles on the surface, while the larger scale linear structures are crystallographic in origin and are related to the angle of incidence. The broad bands crossing the micrograph horizontally, vertically, and diagonally, are the channeling contrast from lattice planes in the  $\{200\}$  and  $\{220\}$  type zones in the crystal. These bands cross with four-fold symmetry about a point whose angular position marks the direction of the  $(011)$  pole of the crystal. Moving the crystal laterally will not change its symmetry (unless some crystallographic boundary is crossed) and hence the channeling pattern will not change, but if the crystal is tilted or rotated then the pattern moves as if rigidly fixed to the lattice. Note that in the channeling condition the observed contrast comes only from regions within two or three extinction distances of the surface (i.e., typically 30 to 50 nm at 20 keV) not from the full depth from which the BSE can emerge [19]. The quality of the crystal in the surface region is therefore important, and samples must be chemically or electrochemically polished to give good patterns.

The geometry and appearance of the channeling pattern is similar to that of a Kikuchi pattern in transmission electron microscopy (i.e., a gnomonic projection) and contains similar information about the crystal. The angular width of the bands is twice the appropriate Bragg angle for the given lattice spacing and electron wavelength and so for a known accelerating voltage lattice spacings may be deduced from the measured width after the angular scale of the pattern is obtained by using



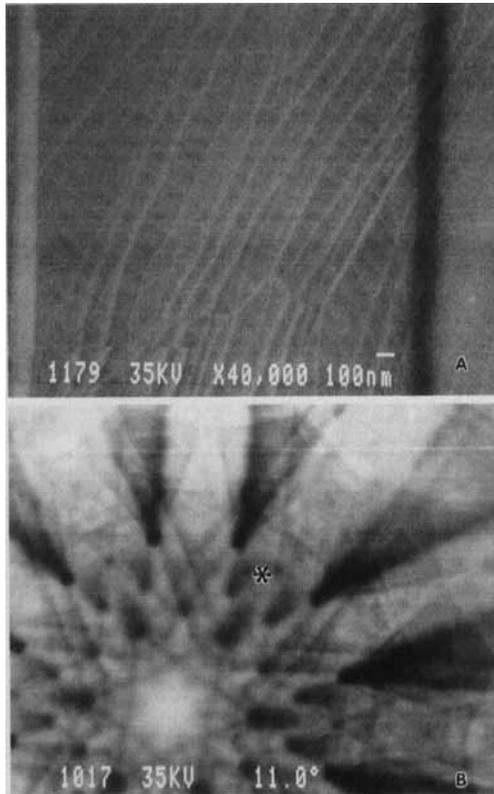
**Figure 12.** Selected area channeling pattern across the epitaxial interface between  $\text{CaF}_2$  and silicon. Recorded on Cambridge S250 at 25 keV.

a crystal of known spacings. From the lattice spacings and interplanar angles the pattern can then be indexed, and the symmetry, orientation and space group of the crystal can be determined.

By a modification of the scan arrangement it is possible to obtain selected area channeling patterns containing information from areas as small as  $1\ \mu\text{m}$  in diameter [19]. This is achieved by allowing the beam to rock about a fixed point on the surface but not scan laterally. Figure 12 shows an example of the utility of this arrangement. The sample is of calcium fluoride  $\text{CaF}_2$  grown epitaxially on silicon and the sample has been cross-sectioned to reveal the interface. The continuity of the selected area channeling pattern across the interface region confirms that epitaxy has been achieved but a comparison of the patterns on the two sides of the interface shows that while the pattern from the silicon contains sharp, well defined, detail

the corresponding pattern from the  $\text{CaF}_2$  is much less well defined. Although the major features of the pattern are still discernible the fine detail is absent. This effect is attributable to the presence of dislocations, and hence of strain, in the  $\text{CaF}_2$  layer. The variations in lattice orientation around each dislocation have the same effect as superimposing two patterns of slightly different orientation. Consequently the detail in the pattern becomes broader and less distinct and, at high enough dislocation densities, disappears completely. The channeling pattern can therefore be analyzed to determine the quality of the crystal as well as its other attributes. This technique has been successfully applied to studies of deformation, annealing, superplasticity, and to the study of stress concentrations associated with corrosion cracking [19].

In the limit the technique can be used to visualize individual defects within the crystal as shown in Fig. 13. If the crystal, here a bulk sample of  $\text{MoS}_2$ , is oriented with respect to the beam so as to avoid any strong reflections then the lattice distortions which occur around a dislocation may be sufficient to rotate the lattice into a strong channeling condition and so produce visible contrast as shown in the micrograph. Each of the bright lines represent a single dislocation threading through the crystal. In this particular example several separate sets of defects are evident. Such images obey the same rules as diffraction contrast micrographs in the TEM and so the Burger's vector of dislocations may be determined by imaging in several different orientations [20]. Because specimens need not be thinned to produce dislocation contrast in this mode, and because the specimen as a result has only



**Figure 13.** (a) Defect image in MoS<sub>2</sub> recorded at 25 keV on Jeol JSM890 FM SEM. The star on the channeling pattern (b) indicates the orientation to which the crystal was set for the micrograph.

one rather than two free surfaces, this technique is more rapid than the corresponding TEM methods and uses a specimen which is more physically realistic. It can also be usefully employed even at very low dislocation densities where the chance of finding a defect in the field of view of a TEM would be very small.

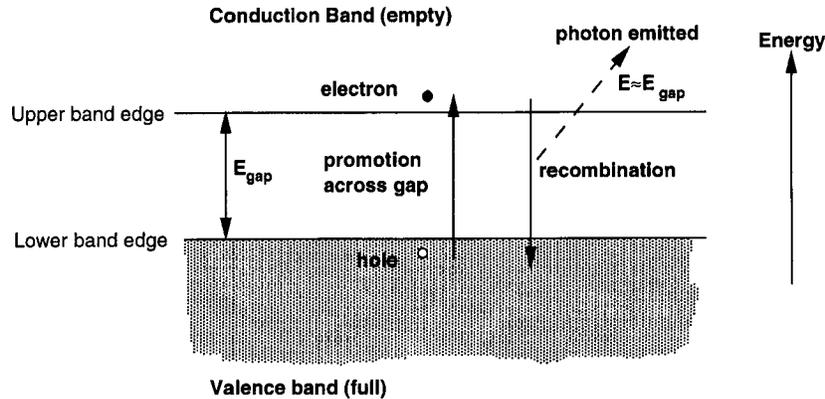
### 2.1.4.3 Special Techniques

Unlike the techniques already discussed the imaging modes that will now be

discussed are specific to a particular class of materials, those which are semiconductors. Because of the technological and economic importance of these materials the capabilities of the SEM in this field have proven to be of great value.

A semiconductor is a material, Fig. 14, in which the valence band and the conduction band energy levels are separated by a band gap. Depending on the position of the Fermi energy level the conduction band is therefore either empty of electrons, or contains only a very small number and the material has little or no electrical conductivity. If a fast electron is injected into the semiconductor then some of the energy that it deposits in the specimen can be used to promote electrons across the band from the valence band to the conduction band. Since the valence band was initially completely filled the removal of an electron also produces a hole in this band. Because the electron is negatively charged and the hole has a positive charge they drift through the material together and so it is convenient to describe them as being an electron-hole pair. The energy  $E_{ch}$  required to form one electron-hole pair is about three times the band gap, for example, in silicon  $E_{ch}$  is 3.6 eV. A single incident electron of energy  $E_0$  can then create about  $E_0/E_{ch}$  electron-hole pairs, that is approximately 3000 pairs for each 10 keV incident electron.

In the absence of any external stimulus the electrons and holes will drift through the lattice, staying physically close to each other so as to maintain overall electrical neutrality, and within a short time, typically  $10^{-15}$  to  $10^{-12}$  s after the initial excitation each electron will drop back into a hole releasing its excess energy in one or more ways including the production



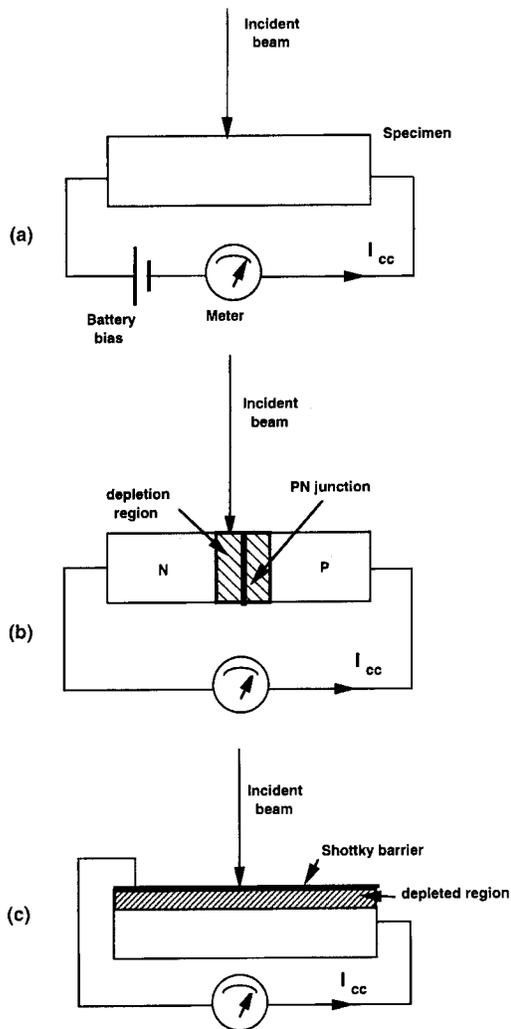
**Figure 14.** Band gap diagram for a semiconductor.

of cathodoluminescence (CL) which is discussed below. However, a semiconductor – unlike a conductor – can sustain a potential difference across itself. If a voltage is applied across the material then the resultant electric field will cause the holes to move towards the negative potential and the electrons towards the positive potential. This motion of charge carriers constitutes a current flow and hence the incident electron beam has produced conductivity in the semiconductor. If the incident beam is turned off the current flow will cease. This phenomenon is therefore referred to as electron beam induced conductivity (EBIC).

*(a) Charge Collection Microscopies (EBIC)*

A field can be produced across a semiconductor in three ways as shown schematically in Fig. 15. The first way, Fig. 15a, is to apply a potential from an external source. This mode is often referred to as  $\beta$ -conductivity and has not found widespread use because it offers no advantages over either of the other techniques

available [13]. The first practical technique for employing electron–hole pair carriers as a signal source is that shown in Fig. 15b which uses the depletion field which exists around a p–n junction in a semiconductor. In the arrangement shown the junction is short circuited through the external conductor. With no incident beam of electrons no current flows, but there is a potential difference between the p and n sides of the junction of typically 0.5 to 1 V. The field associated with this potential difference extends for a distance which depends on the resistivity  $\rho$  of the material but which is typically a few micrometers on either side of the physical position of the junction. This region is called the depletion zone because it can contain no mobile charge carriers. If the electron beam is allowed to fall on to the semiconductor far away from the junction then, although electron–hole pairs are being generated, they are in a field-free region and so will recombine without any net charge flow. But if the beam is placed within the depletion region then the field will separate the electrons and holes, a net motion of charge will occur, and a current  $I_{\text{cc}}$  will flow around



**Figure 15.** Schematic illustration of three ways of performing EBIC: (a) external battery bias; (b) p-n junction; or (c) Schottky surface barrier.

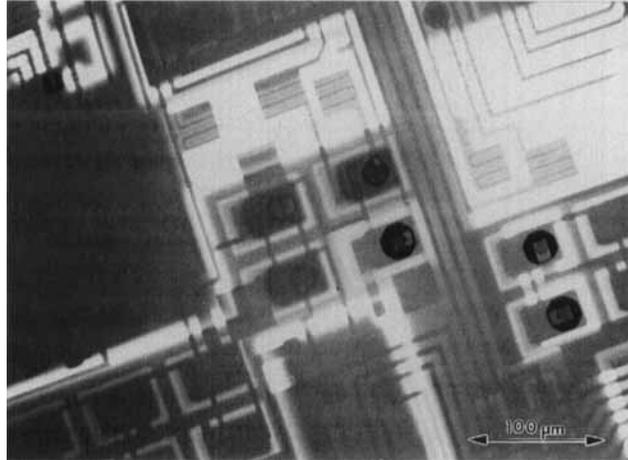
the external circuit. This current will be  $e(E_0/E_{ch})$  where  $e$  is the electronic charge and is referred to as the electron beam induced current (EBIC). (To avoid confusion it is better to describe all of these modes as being 'charge collected microscopies'.) If the beam is placed outside of the junction region but close to it then some of the electrons and holes will diffuse

into the depletion region, be separated, and generate a signal. If the distance between the incident beam and the junction is  $X$  then the fraction  $f(X)$  of the carriers which diffuse to the junction is

$$f(X) = \exp(-X/L) \quad (3)$$

where  $L$  is the minority carrier diffusion length (i.e., the diffusion length for the holes in the n-type material, or for the electrons in the p-type material). The charge collected signal  $I_{cc}$  therefore falls away exponentially on either side of the junction at a rate dependent on the value of  $L$ , which is typically 1 to 10  $\mu\text{m}$  for common semiconductors.

This behavior was first observed experimentally by Everhart and Wells [21] and provides a convenient and powerful way of examining the electrically active regions of semiconductor devices. An integrated circuit can be used to produce an EBIC image by connecting a suitable amplifier, that is, one that is capable of responding to currents in the 1 nA to 1  $\mu\text{A}$  range, across the + and - power lines into the device. Figure 16 shows an image formed in this way from a shift register device, together with the corresponding SEI image. Junctions from all of the transistors and diodes in the field of view and within the penetration depth of the incident beam contribute to the observed contrast in the EBIC image; this technique is therefore widely applied when reverse engineering a chip as it provides a detailed layout of the structure of the device. Two extremes of the signal are visible, bright white and dark black, corresponding to current flowing from either a p-n or an n-p junction since in these two cases  $I_{cc}$  has the same magnitude but opposite sign. None of the topographic contrast visible in the SE



**Figure 16.** EBIC image of p–n junctions in an integrated circuit. Recorded at 15 keV in Hitachi S-800 FE SEM.

image appears in the EBIC image because the collection of the electron–hole pair signal is dominated by the internal electric fields of the device rather than by anything happening at the surface. At the magnification with which this image was recorded the edges of the p and n regions look sharp, but if the magnification were to be increased then the boundaries would be found to be blurred because of the diffusion of the electrons and holes in the semiconductor. Although this sets a limit to the spatial resolution of the EBIC image it is also a valuable tool because if the variation of  $I_{cc}$  with beam position is measured then by using Eq. (3) the minority carrier diffusion length  $L$  can be deduced. Since  $L$  is greatly affected by the processing that the semiconductor has undergone the ability to measure the diffusion length on a micrometer scale provides an invaluable diagnostic tool in device fabrication, failure analysis, and quality assurance testing.

The limitation of using a p–n junction to collect the electron–hole pairs is that efficient collection of the signal only occurs within a few micrometers on either side of

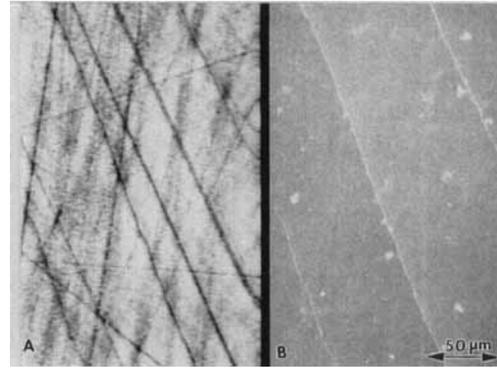
the junction and its depletion region. The technique is made more useful, as shown in Fig. 15(c), when a Schottky barrier is deposited on to the surface of the semiconducting material. The depleted region now extends beneath the whole area of the barrier, and downwards from the surface to a depth dependent on the resistivity of the material. By depositing a barrier several millimeters in size on to a semiconductor large areas of the material can therefore be examined. The Schottky barrier is a metal film, usually titanium or chromium, evaporated on to the atomically clean surface of the material [13] and acts in the same way as a p–n junction. The EBIC signal is collected by making electrical contacts to the barrier itself and to the semiconductor. The benefit of this approach is that it can be used to look at a material before it is processed into a device. If the material were perfect then the EBIC image would be of uniform brightness. But any factors which modify the electrical characteristics of the semiconductor will produce visible contrast effects. Figure 17 shows the EBIC image from a GaAs wafer. The circular area is



**Figure 17.** EBIC image of doping variations in GaAs wafer. Recorded in Cambridge S250 SEM at 15 keV.

the outline of the Schottky barrier region, and the shadow of the electrical contact on to the barrier is also visible. The variations in brightness across the material are due to impurities in the wafer. Although these are only present at a concentration of  $10^{16} \text{ cm}^{-3}$  (i.e., 1 part in  $10^7$ ) they vary the resistivity of the material and hence the depletion depth beneath the barrier. If the range of the incident electron beam [Eq. (2)] is greater than the maximum depletion depth then an increase in the depletion depth will increase the signal collected, and vice versa. The extreme sensitivity and speed of this technique makes it ideal for the characterization of as-grown semiconductor crystals [22].

Electrically active defects in a crystal also produce contrast in the EBIC image, as shown in Fig. 18. In this micrograph from a wafer of silicon deformed 0.4% at  $670^\circ \text{C}$  a network of dark lines can be seen. Each line is the trace of an individual dislocation in the material. Because of the presence of dangling bonds at the core of the defect, the dislocation is electrically active and acts as a recombination



**Figure 18.** (a) EBIC image of individual crystallographic defects in thin film of Si regrown by hot-wire over an  $\text{SiO}_2$  layer. (b) Corresponding SE image showing the linear grain boundaries. Images recorded in Cambridge S250 at 15 keV.

site. If the electrons and holes recombine at the defect they do not contribute to the external measured current  $I_{cc}$  and so the signal intensity falls. The width of the defect line image depends on the local diffusion length in the material as well as on the depth beneath the surface at which the defect is lying, but it is typically a fraction of a micrometer. The ability of the SEM to image defects in a bulk wafer in this way is of great importance, especially in modern semiconductor materials where the defect density may be only 1 to  $100 \text{ cm}^{-2}$ . Conventional techniques such as transmission electron microscopy can only examine areas of the specimen a few micrometers in diameter, and the chance of finding a defect within such a small area is very low. Because the EBIC defect images are relatively broad they can easily be seen at low imaging magnifications and large areas of the specimen can be rapidly examined. In an important extension of this technique the defects can be further characterized by the technique of deep level transient spectroscopy (DLTS) in

which the electron beam is used as a source of charge carriers to populate all of the trapping levels of the semiconductor. A measurement of the transient current which flows when the beam is switched off, due to thermal desorption of carriers from the traps, as a function of the specimen temperature, provides a spectrum which both characterizes different types of traps (electron or hole) and their energy within the band gap [13].

### (b) *Cathodoluminescence*

The charge collection mode discussed above is a powerful tool for the examination of semiconductor materials and devices. However, it can only be applied if the material contains p–n junctions or if a Schottky barrier can be fabricated on the sample surface, and for many materials of interest, such as a II–VI compound semiconductor, this is not readily possible. In these cases, and also for the study of geological and organic materials, the technique of cathodoluminescence (CL), in which the light emitted from a sample is studied, is of value. The luminescence signal is the result of electronic transitions between quantum mechanical energy states separated by a few electron volts, and two basic types can be distinguished: intrinsic or edge emissions, and extrinsic or characteristic luminescence.

Intrinsic luminescence is associated with the recombination of electrons and holes across the gap and occurs as a band of excitation with its intensity peak at a photon energy

$$h\nu \approx E_{\text{gap}} \quad (4)$$

where  $E_{\text{gap}}$  is the band-gap energy of the material. In direct gap semiconductors

(such as InP, GaAs, CdS) this transition usually results in radiation being produced, but in indirect gap semiconductors (such as Si, Ge, GaP) the simultaneous emission of a photon and a phonon is required to conserve momentum in the transition and the probability of this process is small so the emission is relatively weak.

Extrinsic luminescence depends on the presence of impurities and can occur in both direct and indirect semiconductors. The emission bands in these cases are activated by the presence of impurity atoms or other defects and the luminescence is characteristic of the atom with which it is associated. Extrinsic radiation is much more intense than intrinsic radiation and varies about linearly with the concentration of impurity atoms present. A phosphor is an extreme example of extrinsic luminescence production.

Electron beam excitation leads to emission by both of these modes in all types of semiconductors. The brightness dependence  $I_{\text{CL}}$  of the cathodoluminescence signal depends on both the beam energy  $E_0$  and the beam current  $I_b$  through an expression of the form

$$I_{\text{CL}} = f(I_b)(E_0 - E_d)^n \quad (5)$$

where the functional dependence on the beam current is about linear, and  $n$  is typically between 1 and 2.  $E_d$  is the ‘dead voltage’ of the material, that is, the beam energy below which no luminescence is produced, typically 2 to 5 keV. This ‘dead voltage’ is the result of the competition between surface recombination and diffusion effects on the electron–hole pairs produced near the sample surface and so is temperature and material sensitive. In general the intensity of the CL signal also

improves as the sample temperature is lowered, especially the weak band-edge emissions. For this reason it is desirable to be able to cool the sample to liquid nitrogen or even liquid helium temperatures. Except for a few exceptional materials the intensity is never very high, so the prerequisites for successful CL operation in the SEM are a highly efficient system for the collection of light, the ability to generate high incident beam currents at moderately high beam energies (10–25 keV), and some provision for sample cooling.

The collection of the CL signal requires care. Although simple systems in which a light guide, or a light sensitive detector, are allowed to look directly at the specimen have often been used these are not reliable because most materials show some fluorescence under electron impact and consequently it is the backscattered electrons, rather than the photons, hitting the collector that generate most of the observed signal. Figure 19 shows schematically a suitable arrangement for performing CL studies in the SEM [23]. The emitted light is collected by an ellipsoidal mirror, with

the specimen placed at one focus of the ellipse and the light guide at the other. With this arrangement the light pipe is shielded from the possibility of any impact by incident or scattered electrons, and the solid angle of light collection from the specimen approaches  $2\pi$  steradians. At high beam energies (1000 keV or more) it may be necessary to use a double mirror system because X-rays generated by scattered electrons on the mirror can also result in the production of spurious CL [24]. If the specimen is to be cooled to cryogenic temperatures then the system is made still more complex by the need to provide radiation shields around the sample, and the need to cool the mirror. Light emerging from the collector can then either be passed to a photomultiplier, for panchromatic imaging, or into a spectrometer for spectral analysis. The luminescence signal is finally detected using photon counting electrons fed either into a scaler or a multichannel analyzer. Although the emission for some common semiconductors is in the wavelength range 300 to 800 nm, for most materials the radiations

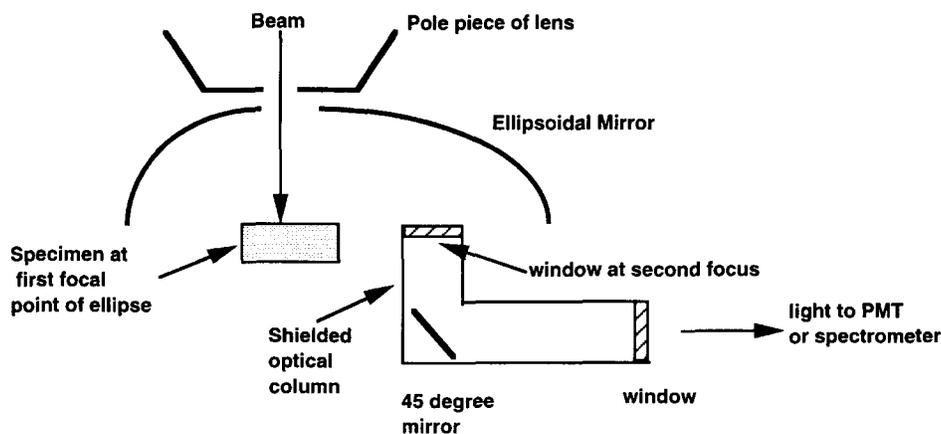
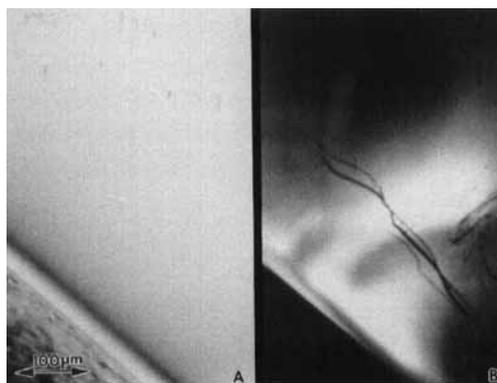


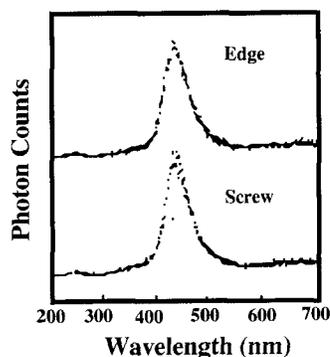
Figure 19. Schematic layout for doing CL in SEM.

of interest occur in the infrared region with wavelengths from 1 to 10  $\mu\text{m}$  or more. It is, therefore, necessary to choose light guides, window materials, and detectors which are suitable for the wavelengths of interest [23].

There are two basic modes of CL operation. In the first, all of the emitted radiation is collected and amplified for display. The CL image produced as the beam scans is therefore a measure of the variation in the total luminescence produced. This mode is simple to set up and is suitable both for the observation of contrast due to extrinsic (dopant produced) CL as well as for intrinsic effects. Figure 20 shows the image of dislocations in a GaAs wafer produced in this way. As for the case of EBIC the dislocation is visible because electron-hole pairs recombine on it instead of producing photons by radiative recombination, the defect therefore appears darker than the surrounding perfect material. The spatial resolution of this kind of image is limited by the diffusion of the carriers and by the rather large incident probe size needed to achieve a high enough



**Figure 20.** CL image of dislocations in GaAs recorded using panchromatic radiation at liquid nitrogen temperatures in Cambridge S250 SEM.



**Figure 21.** CL spectra from individual dislocations in diamond. Adapted from Yamamoto et al. [24].

beam current (here about  $0.1 \mu\text{A}$ ) at the specimen, not by optical diffraction (i.e., the Abbé limit) as would be the case in a conventional optical microscope. A resolution of from 0.1 to  $0.5 \mu\text{m}$  is usually possible.

The other main mode is to collect the spectrum of CL emission from the sample through a grating or prism spectrometer. Figure 21 shows spectra recorded from individual dislocations [24] under the electron beam. Both the wavelength of the peak intensity and the shape of the luminescence peak are seen to vary. These parameters depend on the exact electronic nature of the defect and can therefore be used as a diagnostic. In addition the spectrum allows identification of the characteristic band-edge radiation, and hence a determination of the band-gap of the material, as well as emission peaks due to dopants and impurities. The chemical sensitivity of such an analysis is extremely high, since effects due to dopants below 1 ppb are readily detectable [13]. If sufficient signal is available then imaging can be combined with spectroscopy by forming images from a specified range of wavelengths [24] so allowing the direct

identification of specific features on the specimen.

In summary, cathodoluminescence is a technique which offers many important capabilities for the examination of semiconductors and insulators. The major drawback of the mode has been the complexity of the equipment needed to perform it satisfactorily but the advent of commercial systems may solve this problem and make the technique more accessible.

## 2.1.5 Conclusions

The scanning electron microscope is a uniquely versatile and powerful tool for the characterization and visualization of materials. It combines high spatial resolution with the ability to look at samples of a practical size, and offers a wide range of imaging modes with which to attack a variety of questions. When the ease of specimen preparation for the SEM and its general user-friendliness are also taken into account the future for this instrument looks promising.

## Acknowledgements

Oak Ridge National Laboratory is managed by Martin Marietta Energy Systems Inc. under contract DE-AC05-84OR21400 with the US Department of Energy.

## 2.1.6 References

- [1] M. Von Ardenne, *Z. Phys.* **1938**, 109, 553.
- [2] M. Von Ardenne, *Z. Techn. Phys.* **1938**, 19, 407.
- [3] V. K. Zworykin, J. Hillier, R. L. Snyder, *ASTM Bulletin* **1942**, 117, 15.
- [4] C. W. Oatley, *J. Appl. Phys.* **1982**, 53, R1.
- [5] T. Nagatani, S. Saito, M. Sato, M. Yamada, *Scanning Microscopy* **1987**, 1, 901.
- [6] T. E. Mulvey, C. D. Newman, *Institute of Physics Conference Series* **1973**, 18, p. 16.
- [7] T. E. Everhart, R. F. M. Thornley, *J. Sci. Instrum.* **1960**, 37, 246.
- [8] K. C. A. Smith, *Proceedings of the 5th Annual SEM Symposium* (Ed: O. Johari), IITRI, Chicago **1972**, p. 1.
- [9] J. I. Goldstein, D. E. Newbury, P. Echlin, D. C. Joy, A. D. Romig, C. E. Lyman, C. E. Fiori, E. Lifshin, *Scanning Electron Microscopy and X-ray Microanalysis*, Plenum Press, New York **1992**, Chapter 2.
- [10] H. Kimura, H. Tamura, *Proceedings of the 9th Annual Symposium on Electron, Ion, and Laser Beams* **1967**, p. 198.
- [11] D. C. Joy, *Ultramicroscopy* **1991**, 37, 216.
- [12] D. C. Joy, *Proceedings of the 52nd Annual Meeting MSA* (Ed: G. W. Bailey, A. J. Garratt-Reed), San Francisco Press, San Francisco, **1994**, p. 1032.
- [13] D. B. Holt, D. C. Joy, *SEM Microcharacterization of Semiconductors*, Academic Press, London **1989**, Chapters 7 and 9.
- [14] K. Koike, K. Hayakawa, *Japan. J. Appl. Phys.* **1984**, 23, L178.
- [15] E. de Harven, D. Soligo, *Am. J. Anat.* **1986**, 175, 277.
- [16] M. D. Ball, M. Wilson, S. Whitmarsh, in *Electron Microscopy and Microanalysis 1987* (Ed: L. M. Brown), Institute of Physics, London **1987**, p. 185.
- [17] P. C. Sercel, J. A. Lebens, K. J. Vahala, *Rev. Sci. Instrum.* **1989**, 60, 3775.
- [18] J. Tixier, R. Philibert, *Micron* **1969**, 1, 174.
- [19] D. C. Joy, D. E. Newbury, D. E. Davidson, *J. Appl. Phys.* **1982**, 53, R81.
- [20] D. C. Joy, in *High Resolution Electron Microscopy of Electron Defects* (Ed: R. Sinclair, D. J. Smith, U. Dahmen), Materials Research Society, Warrenton, PA, **1990**, p. 199; see also J. T. Czerneska, N. J. Long, E. D. Boyes, P. B. Hirsch, *Phil. Mag. Lett.* **1990**, 62, 227.
- [21] T. E. Everhart, O. C. Wells, C. W. Oatley, *J. Electron. Control* **1959**, 7, 97.
- [22] H. J. Leamy, *J. Appl. Phys.* **1982**, 53, R51.
- [23] B. G. Yacobi, D. B. Holt, *Cathodoluminescence Microscopy of Inorganic Solids*, Plenum Press, New York **1990**, Chapter 1.
- [24] N. Yamamoto, J. C. H. Spence, D. Fathy, *Phil. Mag.* **1984**, 49, 609.

## 2.2 Scanning Transmission Electron Microscopy

### 2.2.1 Introduction

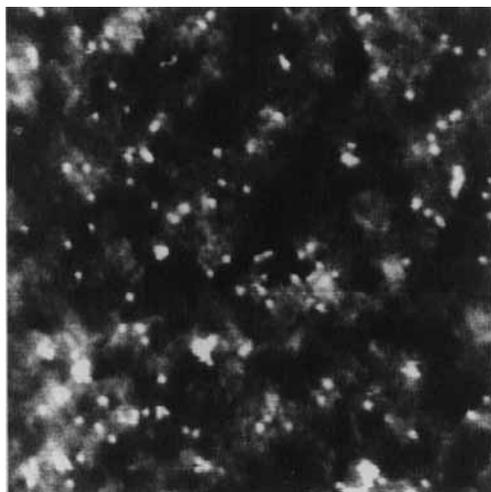
Although the concept of scanning transmission electron microscopy (STEM) was not new, the possibility of employing the scanning principle for high-resolution imaging of thin specimens in transmission was first realized by Crewe, who introduced the use of field-emission guns for this purpose [1, 2]. Dark-field images, obtained with an annular detector in a STEM instrument, showed the first clear electron microscopy images of individual heavy atoms [3] (Fig. 1). From that time, STEM has developed as an important alternative to conventional, fixed beam transmission electron microscopy (CTEM), with special advantages for many purposes.

In a STEM instrument, a fine electron probe, formed by using a strong objective electron lens to demagnify a small source, is scanned over the specimen in a two-dimensional raster (Fig. 2a). The electron probe is necessarily convergent: the convergence angle is, ideally, inversely proportional to the minimum probe size which determines the microscope resolution. On any plane after the specimen, a convergent beam electron diffraction pattern is formed. Some part of this diffraction pattern is collected in a detector, creating a signal which is displayed on a cathode

ray tube screen to form the image using a raster scan matched to that which deflects the incident electron beam. If the detector samples the directly transmitted beam (i.e., if it comes within the central spot of the diffraction pattern) a bright-field image is formed. Detection of any part, or all, of the electrons scattered outside the central beam gives a dark-field image.

The use of a field emission gun (FEG) for high resolution STEM is necessary in order to provide sufficient signal strength for viewing or recording images in a convenient time period. The effective electron source of a FEG has a diameter of 4–5 nm as compared to 1–5  $\mu\text{m}$  for guns with  $\text{LaB}_6$  or tungsten hair-pin filaments. The amount of demagnification needed to form the electron probe is not large. Because the FEG source has a brightness which is a factor of  $10^4$  or  $10^3$  greater than that of a W hair-pin filament or a  $\text{LaB}_6$  pointed filament, the total current in the electron beam is greater when beam diameters of less than about 10 nm are produced. The current in a beam of 1 nm diameter is typically about 0.5 nA. With a FEG, bright- or dark-field STEM images can be recorded in a few seconds or even at TV rates (i.e., 30 frames per second).

As suggested by Fig. 2b, the essential components of a STEM imaging system are the same as for a CTEM instrument,

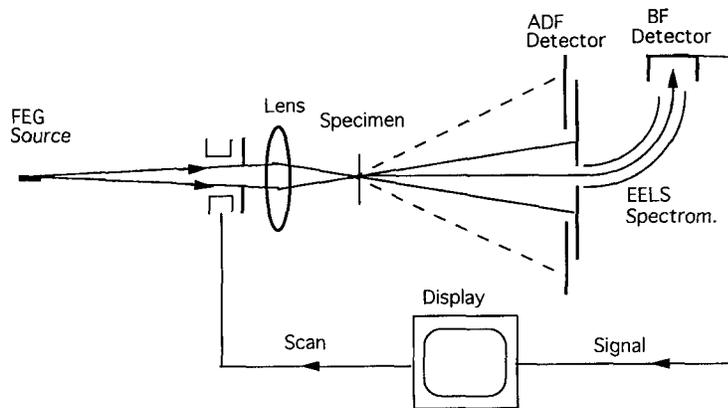


**Figure 1.** Annular dark-field image of uranium atoms on an amorphous carbon film. Full width of image: 31.5 nm. (Courtesy of J. Wall, Brookhaven National Lab.).

but with the electrons traveling in the opposite direction. In this diagram condenser and projector lenses have been omitted and only the essential objective

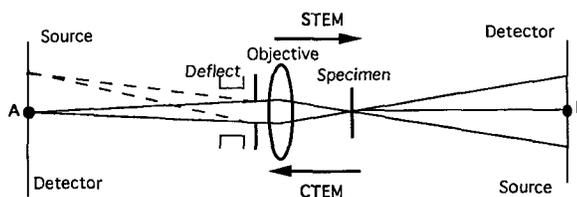
lens, which determines the imaging characteristics, is included. The STEM detector replaces the CTEM electron source. The STEM gun is placed in the detector plane of the CTEM, and the scanning system effectively translates the STEM source to cover the CTEM recording plate.

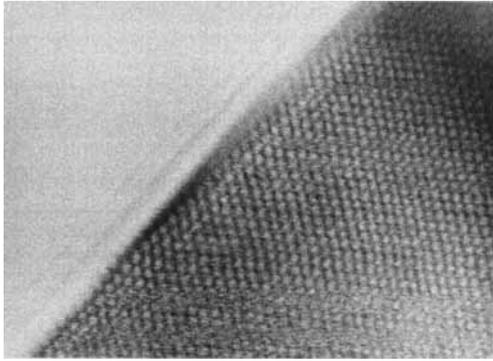
It was pointed out by Cowley [4] (see also Zeitler and Thompson [5]) that application of the reciprocity principle implies that, for the same lenses, apertures and system dimensions, the image contrast in STEM must be the same as for CTEM. The reciprocity principle applies strictly to point emitters and point detectors. It states that, for any essentially scalar system, the wave amplitude at a point B due to a point source at A is identical to the wave amplitude at A due to a point source at B. For an incoherent source of finite diameter and a finite incoherent detector, the reciprocity principle may be considered to apply to each point of the source and each point of



**Figure 2.** (a) Diagram of the essential components of a STEM instrument.

(b) Diagram suggesting the reciprocity relationship between STEM (electrons going from left to right) and CTEM (electrons going from right to left).





**Figure 3.** Bright-field STEM image of a small MgO crystal in [1 1 0] orientation [95].

the detector separately. Thus the effect on the image intensity of increasing the STEM detector size is the same as that of increasing the CTEM source size, and so on. For a particular STEM configuration the image contrast can often be inferred by finding the equivalent CTEM geometry, for which the imaging theory has most probably been well established. Thus for bright-field STEM imaging with a small detector, the image is the same as for BF CTEM with the same small angle of convergence of the incident beam (Fig. 3). A large source size for STEM would degrade the image resolution in the same way as a CTEM detector having poor resolving power.

Practical experimental considerations, however, lead to clear advantages or disadvantages of the STEM instruments, relative to CTEM for some imaging modes. For example, dark-field images can be obtained with high collection efficiency in STEM by collecting all the electrons scattered outside of the incident beam spot of the diffraction pattern. The equivalent CTEM configuration would require an incident beam coming from all directions outside the cone of the objective

aperture, which is difficult to realize and very inefficient in its use of the incident electrons.

Important differences in the fields of application of STEM and CTEM arise from the different form taken by the image signal. For CTEM a two-dimensional detector such as a photographic plate is used to record intensities at all image points in parallel. In STEM the image information is produced in serial form as a time-dependent voltage or current variation. For many years this gave STEM the unique possibility of online image processing to manipulate the image contrast for special purposes. Now the use of CCD detectors allows a serial read-out and online image processing for CTEM also; however, for STEM further possibilities exist. Thus in STEM several detectors may be used simultaneously to produce signals which may be added, subtracted, multiplied or otherwise manipulated.

A variety of STEM signals may be obtained in addition to the bright-field or dark-field signals derived from the elastic scattering of electrons in the specimen. STEM instruments are normally fitted with an energy-loss spectrometer which not only allows microanalysis of very small specimen regions by electron energy-loss spectrometry (EELS), but also allows images to be formed with electrons that have lost particular amounts of energy, characteristic of particular elements or of particular electronic excitations. Also images may be formed by detecting secondary radiations, such as low-energy secondary electrons, Auger electrons, or characteristic X-rays, as will be discussed below. The serial nature of all these image signals provides possibilities

for quantitative comparisons and correlations of information on specimen compositions, crystallography and morphology which can be very valuable.

The serial mode of imaging in STEM has some practical disadvantages. The recording times for images are usually longer than for CTEM: up to 20 s. The image quality can be degraded by fluctuations in the emission from the field-emission tip, resulting from fluctuations of the work function of the tip surface as molecules of residual gas are absorbed or desorbed. Such fluctuations may give the images a streaky appearance. Also there may be a steady decay of emission current so that the correlation of signals over time is difficult. These effects may be minimized if the gun is operated in very high vacuum, but it is difficult to eliminate them entirely. On the other hand, specimen drift can have an adverse effect in CTEM, smearing out the image and degrading the resolution, whereas in STEM, if the image is recorded with a single slow scan of the image field, the effect of a specimen drift is to produce only a slight distortion of the image.

### **2.2.2 Scanning Transmission Electron Microscopy Imaging Modes**

Bright-field STEM images from thin specimens, obtained with a small axial detector, show the same contrast features as for CTEM, as expected from the reciprocity relationship. Phase-contrast effects including the reversal of contrast with defocus and Fresnel fringes, and amplitude contrast due to diffraction and inelastic

scattering, are similar (Fig. 3). If the detector aperture is very small, to simulate the case of parallel-beam CTEM, the image signal is generally too weak and noisy for convenience. Hence, there is a tendency to use larger detector angles with the result that the image contrast is somewhat reduced (although it can be restored by online image processing) and the resolution is slightly improved, although not so readily calculated (see next Section).

For STEM, an EELS detector is normally present so that the BF image may be formed with all electrons transmitted through the specimen or else with only those electrons which have lost less than about 1 eV in energy. For thin specimens, the difference in these two cases is small. For specimens of thickness comparable with the mean free path for inelastic scattering (of the order of 100 nm for 100 keV electrons), the difference becomes significant. The resolution and contrast of the images are degraded for thick specimens by two factors.

First, there is a loss of resolution from the geometric effect of multiple scattering: the point of origin of the scattered electron becomes indefinite because after a second scattering process, for example, it cannot be determined whether the electron comes from the first scattering point or from the second. The different scattering processes appear to occur at different lateral positions. This effect is the same for STEM and CTEM. Second, the inelastic scattering produces a spread of energies of the transmitted electrons. For a 100 nm thickness for 100 keV electrons, for example, the average energy loss is about 30 eV and there is an average spread of energies of about the same magnitude. For CTEM, the chromatic aberration of the objective

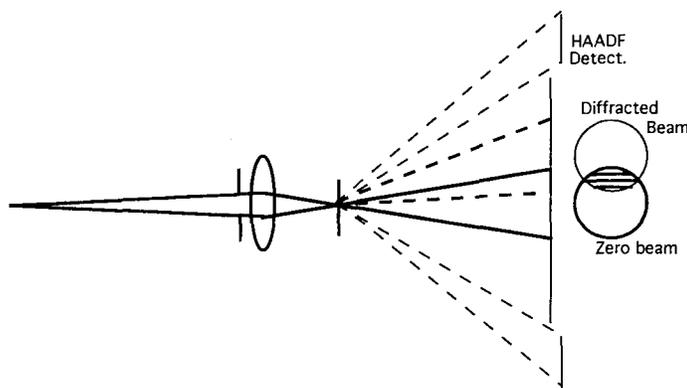
lens then degrades the resolution, but in STEM, because the objective lens comes before the specimen, the effect of such an energy spread is negligible. Hence, in general, the resolution and/or contrast is better for STEM than for CTEM for a given specimen thickness: or, for a given resolution level, the thickness of specimens that can be used is greater for STEM than for CTEM [6, 7]. Because the chromatic aberration effect decreases with accelerating voltage, being proportional to the ratio of the energy spread and the incident beam energy, the advantage of STEM decreases, but even so, a STEM operating at 0.5 MeV has the same 'penetration' as a CTEM at 1 MeV [8].

In the original work of Crewe et al. [3], the detector used was an annular one designed to collect all electrons scattered out of the incident beam. Simplifying approximations which are good for isolated heavy atoms, but may break down for groups of atoms [9], suggest that the image intensity is then proportional to  $Z^{3/2}$ , where  $Z$  is the atomic number of the elements present and the image resolution should be better than for bright-field by a factor of 1.4 or 1.5. The efficiency of

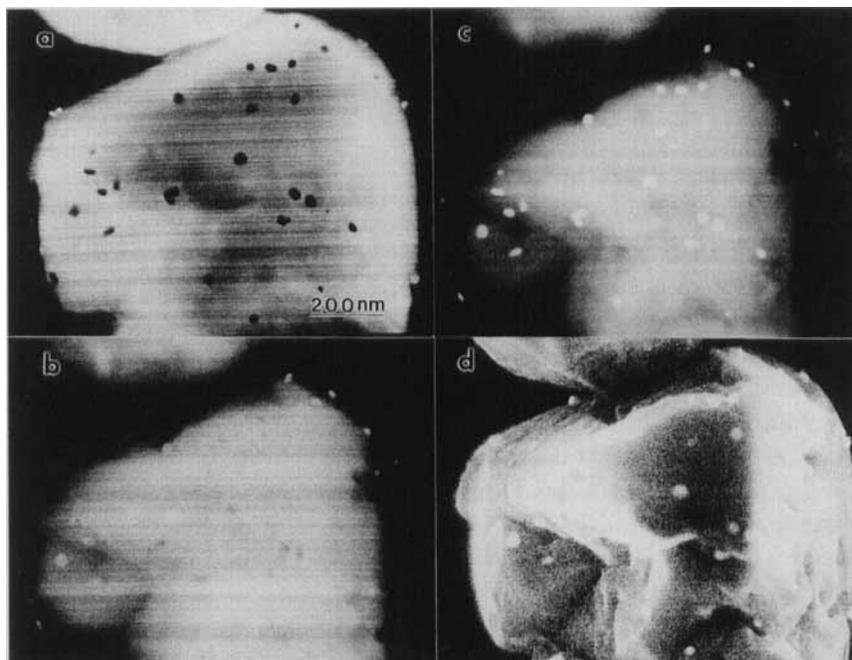
collection of the dark-field signal is much better than for dark-field CTEM, for which only a small part of the diffraction pattern can be collected by the objective aperture and used to form the image.

The  $Z$ -dependence of the annular dark-field (ADF) image suggested its use for detecting heavy-atom particles in a matrix of light-atom material such as occur, for example, in supported metal catalysts. However, if the light-atom material is microcrystalline, as is often the case, the variations of diffraction intensities produce large fluctuations in the ADF image intensity, obscuring the heavy atoms. Howie [10] suggested that this effect could be avoided if the inside hole of the annular detector is made so large that the lower-angle region, where the diffraction spots occur, is not included (Fig. 4). When a high-angle annular dark-field (HAADF) detector is used, good  $Z$ -dependent contrast is obtained, independent of crystallinity [11] (Fig. 5).

The signal collected comes, in part, from the high-angle elastic scattering which has an average value depending on the square of the atomic scattering factor,  $f(u)$ , where  $u$  is the angular variable, equal



**Figure 4.** Diagram of a STEM system suggesting the formation of overlapping diffracted beam disks, with interference fringes in the area of overlap, and the collection of high-angle scattered radiation to form HAADF images.



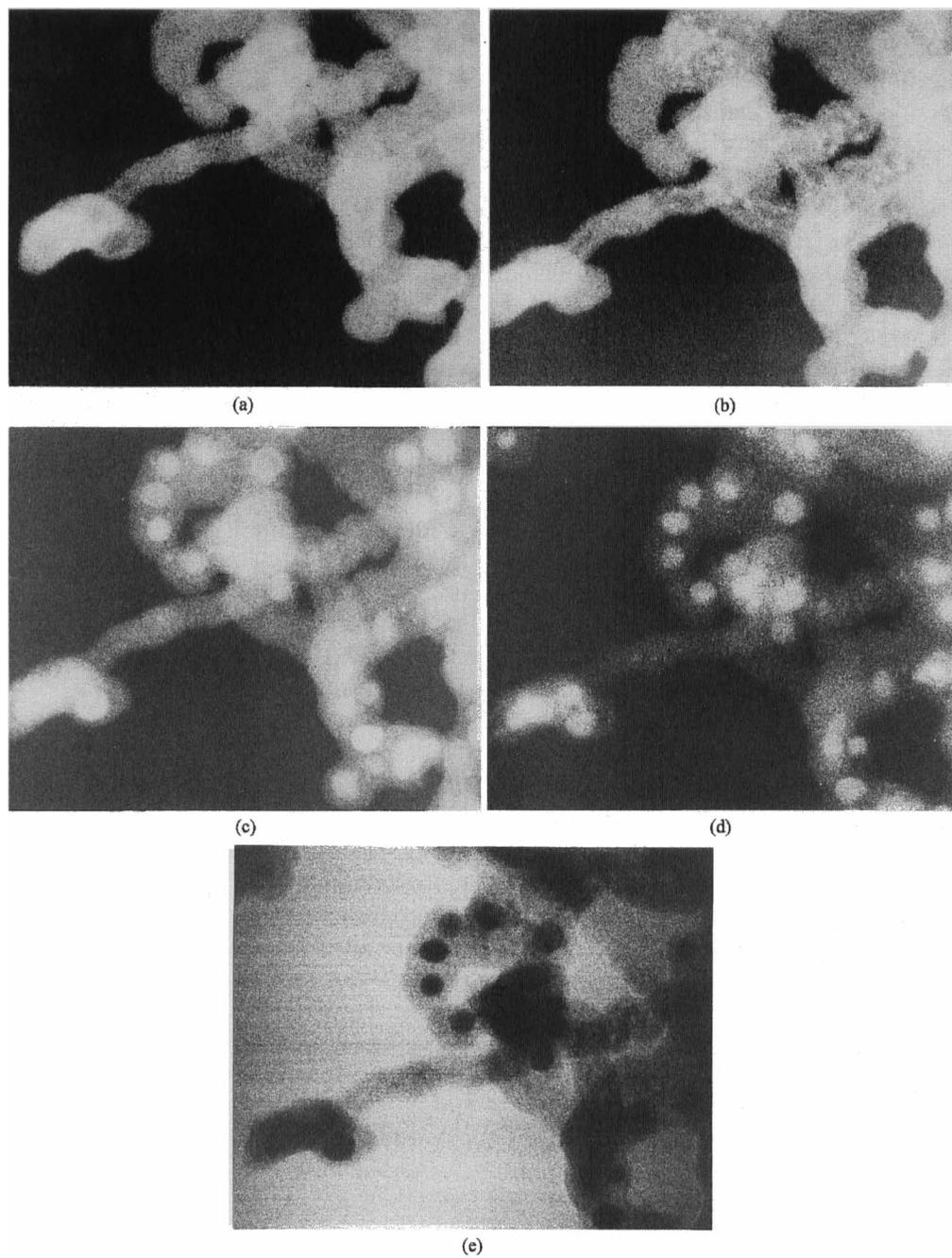
**Figure 5.** (a)–(c) ADF STEM images of Pt particles on a thick NaY-zeolite crystal with inner collector angle 20 mrad for (a) and increasing for (b) and (c), plus a SEM image (d), of the same specimen showing the Pt particles on one surface. Note the reversals of contrast of the Pt particles as the thickness and detection angles change. (Courtesy of J. Lui [96]).

to  $(2/\lambda) \sin(\phi/2)$  for a scattering angle  $\phi$ . A predominant part of the signal, however, comes from the thermal diffuse scattering which is intrinsically a high-angle scattering because, even for first-order scattering, the intensity depends on the square of  $u \cdot f(u)$  [12, 13]. More recently it has been shown that the HAADF mode can give good high-resolution imaging of crystals in principal orientations in which differences in  $Z$  of the columns of atoms parallel to the beam can be clearly indicated [14, 15].

It is a virtue of STEM that the form of the detector may be varied readily to give special types of information in the images. In Fig. 5 it is seen that, for a relatively thick specimen, the influence of multiple scattering to higher angles can lead to

striking changes of contrast as the inner diameter of the ADF detector is changed. For thin specimens, if a thin annular detector is used, with only about 10% difference between inside and outside diameters, images may be obtained from different ranges of diffraction angles so that components of the specimen which diffract differently may be distinguished (Fig. 6).

Circular detectors in the bright-field area, split into two semicircular halves, were proposed by Dekkers and de Lang [16]. The signals from the two halves may be added or subtracted. A simple, geometrical optics description serves to give a useful approximation to the image intensities. A linear increase or decrease in the projected potential in the specimen acts



**Figure 6.** STEM images of a specimen of carbon particles formed in a high-pressure arc with Mn present. (a)–(c), Images with a thin annular detector for average  $d$ -spacings of 0.6, 0.3 and 0.1 nm, showing, respectively, the amorphous carbon, the graphitic carbon and the Mn carbide particles. (d) HAADF image, and (e) bright-field image.

like a prism for electrons, deflecting the incident beam one way or the other. If the deflection has some component in the direction perpendicular to the slit between the two semicircular detectors, the signal in one detector is increased and the signal in the other is decreased. The difference signal then corresponds to the gradient of the projected potential of the specimen and is either positive or negative. The sum of the signals, proportional to the projected potential for a weak phase object, can be used as a reference signal.

This differential imaging mode is useful in some cases for detecting particles in a matrix since it gives sharply defined positive or negative signals at the edges of the particle. It can also be used to detect magnetic fields in a specimen which deflect the incident beam. Chapman et al. [17] have developed the method, called differential phase contrast imaging, into a highly effective means for studying the magnetic fields and their changes in the domain structures of thin ferromagnetic films. By using a circular detector split into four quadrants, and adding or subtracting the signals from the various quadrants, they could define the directions and strengths of the magnetic fields. One difficulty with this approach is that the signal from the magnetic field variation may be confused by signals from the variation of projected potential arising from variations of specimen thickness or structure. A further subdivision of the detector into eight regions, with two concentric sets of four quadrants, allows the separation of the magnetic signal, which occurs mostly at lower angles of scattering, from the structural signal [18]. Also, even more complicated multiple detectors have been proposed [19].

### 2.2.3 Scanning Transmission Electron Microscopy Theory

Following Cowley [20], we may write relatively simple expressions to derive the form of the image contrast for various STEM detectors if we confine our treatment to thin objects for which the effect of the specimen is well represented by a transmission function,  $q(x, y)$ , which multiplies the incident electron wave-function. The electrostatic potential within the specimen modifies the energy and wavelength of the electrons, acting like a refractive index slightly greater than unity, and so modifies the phase of the electron wave by an amount proportional to the projection of the potential distribution in the incident beam direction

$$\int \phi(x, y, z) dz \equiv \phi(x, y)$$

Then the transmission function is

$$q(x, y) = \exp[-i\sigma\phi(x, y)]$$

where  $\sigma$  is the interaction constant, equal to  $2me\lambda/h^2$ . This is the so-called phase-object approximation, valid only for very thin specimens but without the limitation of the weak-phase object approximation which assumes that  $\sigma\phi(x, y) \ll 1$  and is valid for only light atom materials. In the following, for convenience, we deal with only one-dimensional functions. The extension to two dimensions is obvious.

For the very small effective source size of a field emission gun, the convergent beam incident on the specimen may usually be assumed to be completely coherent, as if coming from a point source, represented by a delta function. The coherence function at the objective aperture

position is given, according to the Zernike–van Cittert theorem, by the Fourier transform of the intensity distribution of the effective source. For the usual dimensions of an FEG illumination system, the width of the coherence function is of the order of several millimeters and so is much greater than the diameter of the objective apertures used for optimum STEM imaging. For the very much larger source sizes common for thermal electron sources, the width of the coherence function at the objective aperture is normally only about 1  $\mu\text{m}$ , so that it is a good assumption that electron waves arriving at the specimen from all parts of a 50 or 100  $\mu\text{m}$  aperture are completely incoherent.

If the electron source is assumed to be a delta function, the electron wave coming through the objective lens is given by the transfer function of the lens,  $T(u)$

$$T(u) = A(u) \exp[i\chi(u)] \quad (1)$$

where  $u = (2/\lambda) \sin(\phi/2)$  for scattering angle  $\phi$  and the aperture function  $A(u) = 1$  if  $u < u_0$  and 0 if  $u > u_0$ , and the phase factor,

$$\chi(u) = \pi \Delta \lambda u^2 + \frac{\pi}{2} C_s \lambda^3 u^4$$

where  $\Delta$  is the defocus (negative for underfocus, i.e., for a weakening of the objective lens) and  $C_s$  is the spherical aberration constant. In the approximation normally used for electron microscopy at the current levels of resolution, the other higher-order or off-axis aberrations, which would add further terms in the phase function, are ignored and it has been assumed that the astigmatism has been corrected. Then the wave amplitude incident on the specimen is given by Fourier transform

of  $T(u)$  as the spread function  $t(x)$  which represents the smearing-out of the ideal image of the source due to the defocus and aberrations of the lens. Translating the incident beam by an amount  $X$ , the wave transmitted through the specimen is  $q(x) t(x - X)$ . Fourier transforming and squaring this function gives the intensity distribution on the detector plane as

$$I_X(u) = |Q(u) * T(u) \exp(2\pi i u X)|^2 \quad (2)$$

Here the  $*$  denotes the convolution operation defined by

$$f(x) * g(x) \equiv \int f(X) g(x - X) dX$$

If the detector has a transmission function,  $D(u)$ , the signal detected as a function of the incident beam position is

$$S(X) = \int D(u) I_X(u) du \quad (3)$$

It is useful to distinguish the transmitted beam from the scattered waves, so we put  $q(x) = 1 - p(x)$ , of which the Fourier transform is  $Q(u) = \delta(u) - P(u)$ . Substituting in Eq. (3) then gives

$$\begin{aligned} S(X) = & \int D(u) A(u) du \\ & + \int D(u) |P(u) * T(u) \\ & \times \exp(2\pi i u X)|^2 du \\ & - \int D(u) T^*(u) \exp(-2\pi i u X) \\ & \times [P(u) * T(u) \exp(2\pi i u X)] du \\ & - \text{c.c.} \end{aligned} \quad (4)$$

where c.c. indicates the complex conjugate of the previous term.

If one makes use of the relation

$$\begin{aligned} Q(u) * T(u) \exp(2\pi i u X) \\ = \exp(2\pi i u X) [q(X) * t(X) \\ \times \exp(-2\pi i u X)] \end{aligned}$$

Eq. 4 can be written as

$$\begin{aligned} S(X) = \int D(u) A(u) du \\ + \int D(u) |p(X) * t(X) \\ \times \exp(-2\pi i u X)|^2 du \\ - p(X) * \left[ t(X) \int D(u) T^*(u) \right. \\ \left. \times \exp(-2\pi i u X) du \right] - \text{c.c.} \quad (5) \end{aligned}$$

In the expressions on the right hand side of Eqs. (4) and (5), the first term comes from the incident beam. The second term comes from the scattered waves, both inside and outside the incident beam cone. It is a term of second order in  $p(X)$  and so represents an integration over intensities. The third term and its complex conjugate are of first order in  $p(X)$ , derived from the coherent interference effect of the superimposed incident and scattered waves and so contribute to the intensity only within the central beam spot defined by  $A(u)$ .

From these expressions, it is straightforward to derive the image intensities obtained for various forms of the detector function,  $D(u)$ . For example, for a very small detector located on the axis, a good approximation is  $D(u) = \delta(u)$ . Then Eq. (5) reduces to

$$S(X) = |q(X) * t(X)|^2 \quad (6)$$

which is identical with the expression for bright-field CTEM with a parallel beam

illumination. The specimen transmission function is smeared-out by the spread function and its modulus is squared. In the weak phase object approximation, the intensity of the image is given simply as  $I(X) = 1 + 2\sigma\phi(X) * s(X)$ , where  $s(X)$  is the Fourier transform of  $\sin \chi(u)$ , the imaginary part of the transfer function, so a direct smeared-out representation is given of the projected potential.

For an annular detector which collects all the electrons scattered outside the central beam spot, the product  $D(u) T(u)$  is zero so that only the second term remains in Eqs. (4) and (5). In general, this term is not easily evaluated, but if the approximation is made that the contribution of scattered intensity from within the central beam spot is proportional to the signal from outside the central beam spot (as is approximately the case for scattering from a single isolated atom) then one can assume that  $D(u)$  is a constant and

$$\begin{aligned} S(X) &= \int |P(u) * T(u) \exp(2\pi i u X)|^2 \\ &= \int |p(X) t(x - X)|^2 dX \\ &= |p(X)|^2 * |t(X)|^2 \quad (7) \end{aligned}$$

where the second equality follows from Parseval's theorem and the result implies that the square of the scattering function  $p(X)$  is imaged with a spread function equal to the intensity distribution of the incident beam. The annular dark field (ADF) imaging has hence been described as 'incoherent' imaging. For the weak phase-object approximation,  $\sigma\phi(X) \ll 1$ , the image intensity depends on  $\phi^2(X)$ , so that a small bright dot should appear in the image for a positive peak in potential

(e.g., an isolated atom) or for a negative peak (e.g., a hole in thin perfect crystal). On the assumption that  $t(X)$  is a Gaussian function, the resolution for the ADF image is seen to be better than for the bright-field image by a factor of  $\sqrt{2}$ .

The assumption that  $D(u)$  may be replaced by a constant is a reasonable one when the scattering from the object gives an intensity distribution which falls off smoothly from the center of the diffraction pattern as is the case for single isolated atoms or for amorphous materials. It may fail significantly, however, for particular cases such as two closely-spaced atoms, for which the diffraction pattern has sinusoidal oscillations of periodicity comparable with the dimensions of the aperture [9].

If  $D(u)$  represents an axial, circular hole, smaller than the central beam disc, the intensity distribution in the image can be evaluated by computer calculations [21]. Some indication of the effect can be seen from the form of the last term in Eq. (5). If the integral gives a function which falls away from a central peak, multiplying  $t(X)$  by this function has the effect of making the spread function narrower and hence of improving the resolution. However, it turns out that as the detector hole diameter is increased, the constant decreases for weak phase objects and becomes zero for a detector diameter equal to that of  $A(u)$ . On the other hand, the second term of Eq. (4) or Eq. (5), for  $D(u) = A(u)$ , becomes equal to the negative of Eq. (7), so that a bright-field image is given with the same resolution as the dark-field image although with relatively poor contrast because of the constant background intensity from the first term of Eq. (4) or Eq. (5) [22].

To improve the bright-field resolution even further, and to make use of the first-order terms in the intensity expressions, it is necessary to find a detector function,  $D(u)$ , such that the integral in the third term of Eq. (5) represents a peaked function much sharper than  $t(X)$ . It was shown [20], for example, that if  $D(u)$  represents a thin annulus, with inside and outside diameters differing by only about 10% and an average radius equal to the optimum bright-field aperture size, the bright-field resolution should be improved by a factor of 1.7. This thin-annular bright-field STEM mode is equivalent, by reciprocity, to the CTEM bright-field mode using a hollow-cone incident beam, which has been explored extensively [23, 24].

For thicker crystals (thickness greater than 2–5 nm for 100 keV electrons) a simple transmission function, such as the phase-object approximation, cannot be used. As in CTEM, the diffraction pattern and image intensities must be calculated using one of the approximations to the three-dimensional, many-beam dynamical diffraction theory such as the Bloch-wave or multislice methods [25, 26]. For small detector aperture sizes it is convenient to make use of the reciprocity relationship and make the calculations for the equivalent CTEM configuration with a nearly parallel incident beam.

For imaging with large or more complicated detector configurations or for the convergent-beam diffraction patterns, one approach is to make dynamical calculations for each incident beam orientation and then add the diffraction amplitudes or intensities (depending on the assumptions of coherence of the incident beam) for each diffracted beam direction. In an alternative

approach, a multislice calculation is made for an incident beam amplitude described by the spread function,  $t(x, y)$ . Since  $t(x, y)$  is nonperiodic, it is placed within an artificially large unit cell corresponding to a large number of the crystal lattice unit cells, making use of the assumption of periodic continuation [27]. The effect of imposing this large periodicity on the structure in real space is to sample the continuous scattering function in reciprocal space at a finely spaced lattice of points. Hence, the number of effective 'diffracted beams' becomes very large, although not too large to be handled conveniently by modern computers.

For the relatively simple case of lattice fringe imaging, we may consider the case of a periodic structure, periodicity  $a$ , giving a one-dimensional row of diffraction spots with an objective aperture size such that the diffraction spot discs overlap. The diffraction pattern amplitude is then described by

$$\Psi(u) = T(u) \exp(2\pi i u X) * \sum_h F_h \delta(u - h/a) \quad (8)$$

From this may be derived the intensity distribution which, as will be discussed later, includes all the coherent interactions of the diffracted beams with the incident beam and with each other. In the region of overlap of diffraction spot disks, patterns of interference fringes appear (Fig. 4). For the moment, however, we confine our attention to the midpoints of the regions of overlap. The midpoint of the region of overlap of the disks due to the  $h$  and  $h + 1$  reflections comes at  $u = +\frac{1}{2}a$  relative to the  $h$  reflection and at  $u = -\frac{1}{2}a$  relative to the  $h + 1$  reflection, so that the intensity at

that point is

$$\begin{aligned} |\Psi(u + \frac{1}{2})|^2 &= |F_h \exp[i\chi(+\frac{1}{2}a) + \pi i X/a] \\ &\quad + F_{h+1} \exp[i\chi(-\frac{1}{2}a) - \pi i X/a]|^2 \\ &= |F_h|^2 + |F_{h+1}|^2 + 2|F_h| \\ &\quad \times |F_{h+1}| \cos(2\pi X/a + \alpha) \quad (9) \end{aligned}$$

where  $\alpha$  is the phase difference between the two reflections, since  $\chi(u)$  is symmetrical. Hence, if a small detector is placed at the midpoint of the overlap region, the STEM image shows sinusoidal fringes of period  $a$  and a shift of the fringes relative to an origin point which indicates the relative phase of the diffraction spots [28]. It has been proposed that, on this basis, it may be possible to derive the relative phases of all diffraction spots in a diffraction pattern. In the case of the weak phase-object approximation, this information would allow the summing of the Fourier series with coefficients  $F_h$  to give the projection,  $\phi(x, y)$ , of the crystal potential. In this way one could make an unambiguous structure analysis of a thin crystal, avoiding the 'phase problem' which hinders the structure analysis of crystals based on kinematical X-ray or electron diffraction data [29, 30]. The practical difficulties which have prevented the realization of this scheme include the difficulty of obtaining sufficiently thin single crystals of nontrivial structure which are sufficiently resistant to radiation damage by the incident beam.

## 2.2.4 Inelastic Scattering and Secondary Radiations

In addition to the elastic scattering, the beam in a STEM instrument undergoes a

number of inelastic scattering processes. Thermal diffuse scattering with energy losses of about 0.02 eV, produces a broad diffuse background to the electron diffraction pattern extending to high angles of typically  $10^{-1}$  radians, with peaks of intensity around the Bragg diffraction spots. In the bulk of conducting samples, collective electron oscillations, or plasmons, are generated resulting in energy losses in the range of 5 to 30 eV and an angular distribution of the energy loss electrons of  $10^{-3}$  radians or less. At the specimen surfaces, surface plasmons are generated with somewhat smaller energy losses and comparable scattering angles. For non-conducting or conducting specimens, electrons from the outer shells of atoms, or from the conduction bands of solids, may be excited with the same range of energy losses of 5 to 30 eV and a somewhat broader distribution of scattering angles which is still much narrower than for elastic scattering.

The excitation of inner shell electrons of the specimen atoms gives energy-loss peaks in the range of about 50 to several thousand eV with energies depending on the atomic number and the electron shell (K, L, M, ...) involved. The dependence of the inner-shell energy losses and of the energies of the emitted X-rays on the nature of the atoms present provide the basis for microchemical analysis techniques.

Spectroscopy of the transmitted electrons using electron energy-loss spectroscopy (EELS) is performed by passing the electrons transmitted through the specimen through a magnetic quadrant spectrometer. The X-rays emitted from the specimen are detected by placing the small energy-sensitive detector used for

energy dispersive spectrometry (EDS) as close to the specimen as possible. These microanalytical methods, as applied to the thin samples in CTEM or STEM instruments, are described in Secs. 1. and 2.4 of this Chapter. Also the signal from electrons having a particular energy loss, or from the X-rays of a characteristic wavelength, may be used to form an image which displays the distribution of a particular element in the sample.

The particular virtue of the STEM instruments for these purposes arises because, with a cold field-emission gun giving a high intensity of electrons concentrated into a very small probe, the microanalysis may be performed for specimen regions of diameter as small as 1 nm or less and the images showing distributions of the elements may be produced in a convenient time without excessive noise. Although the total electron current from an FEG source is much less than from a thermal source, the much greater brightness of the FEG source ensures that the electron current that can be concentrated in a beam of small diameter is greater for beam diameters of less than a few tens of nanometers. Thus STEM instruments are much more effective if the microanalysis is to be carried out, or atom-specific imaging is to be performed, with high spatial resolution. Discussion and example of analysis and imaging with high spatial resolution have been given for EELS [31, 32] and for X-rays [33].

The signal strengths for energy losses due to plasmons and outer-shell excitations are much higher than for inner shells, but the energy loss distributions are in general rather diffuse and not often clearly characteristic of particular materials. A few materials such as elemental Al and Si

give sharp plasmon peaks which have been used for discrimination in special circumstances [34]. Also the relationship of the energy loss spectra to the electronic states of solids and their surfaces has been explored [35]. The use of very fine beams in a STEM instrument has suggested the possibility of exploring the energy losses, and hence the nature of the excitations of specimen surfaces, by passing the beam close to the edge of a particle or parallel to the face of a small crystal [36, 37]. Such observations have been compared with theoretical models of surface potential distributions and their interactions with high-energy beams [38].

Apart from the characteristic X-rays, the secondary emissions from materials struck by high-energy electrons include low-energy secondary electrons, which provide the basis for SEM imaging (see Sec. 2.1 of this Chapter) and Auger electrons which may be detected to provide scanning Auger electron microscopy (SAM). Both the secondary electrons and the Auger electrons which are detected, being of relatively low energy, arise from atoms close to the specimen surfaces. The secondary electron images give information on surface morphology and are dependent, to a limited extent, on the composition and structure of the near-surface layers of atoms (see Fig. 5). The Auger electrons, which are derived from inner-shell excitations, have energies which are characteristic of the atomic species and so provide information on the compositions of the surface layers. When applied in microscopes having an ultrahigh vacuum specimen environment and the means for providing atomically clean surfaces, SEM and SAM become powerful tools for surface science [39].

Modern SEM instruments, equipped with field-emission guns and operating usually at about 15 to 30 keV, have shown surface morphology with a spatial resolution of 1 nm or better [40]. In a 100 keV instrument, resolutions of 0.6 nm have been demonstrated [41]. There is thus very little basis for the argument that, because the excitation processes leading to secondary electron emission are delocalized, the SEM resolution must be limited to about 10 nm or more. An explanation for the high resolution attainable in SEM with a FEG source has been provided by the electron coincidence measurements of Scheinfein et al. [42]. With a suitable configuration of detectors in a STEM instrument it is possible to obtain SEM images from either the top or the bottom surface of a specimen (or both) in parallel with normal bright-field or dark-field transmission images [43] (Fig. 13).

Since energy spectroscopy of Auger electrons (AES) and SAM imaging allow determinations of the chemical compositions and atomic distributions of the few top layers of atoms on a specimen surface, they can be valuable for ensuring the cleanliness of surfaces with a sensitivity of a fraction of a monolayer of impurity and for following the details of surface chemical reactions. The conventional methods of exciting and detecting Auger electrons using cylindrical or hemispherical electron energy analyzers, are limited by the noisiness of the signals to spatial resolutions of 10 nm or greater. In a UHV STEM instruments, the high-brightness source and the use of special high-efficiency electron detectors have allowed spatial resolutions of 1 nm or better to be achieved [44].

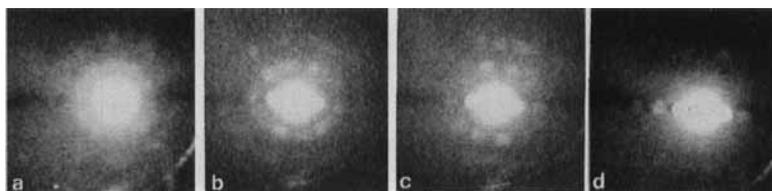
## 2.2.5 Convergent-Beam and Nanodiffraction

The diffraction patterns that form on the detection plane of a STEM instrument are necessarily convergent beam electron diffraction (CBED) patterns with diffraction spot disks of diameter determined by the objective aperture size. With a two-dimensional detector such as a phosphor or CCD system, these patterns can be observed and recorded and used, in the same way as CBED patterns obtained in a CTEM instrument, for accurate determination of crystal lattice constants or crystal thickness or for absolute determinations of space-group symmetries (see Sec. 1.5 of this Chapter). The use of a field-emission gun, rather than a thermal emission gun, introduces two new aspects. First, the region from which the CBED pattern is obtained may be much smaller; 1 nm or less in diameter rather than 10 nm or more. Second, the FEG source gives an essentially coherent incident convergent beam, rather than an essentially incoherent one, so that the effects of coherent interference may be seen in the patterns.

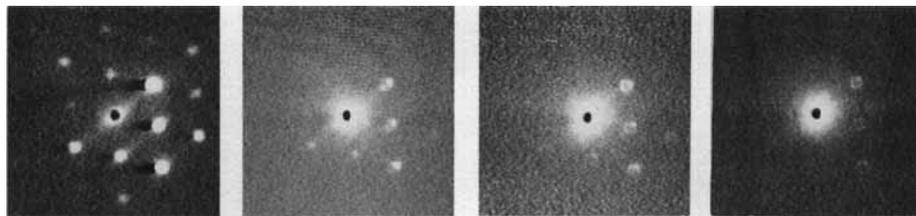
For an ideally perfect thin crystal, there is no difference in the diffraction patterns for coherent and incoherent radiation if the diffraction spot disks do not overlap.

If they do overlap, interference fringes appear in the region of overlap if the incident beam is coherent [45,46]. The coherent interference effects in diffraction patterns will be considered in detail in the next part of this Section. For now we concentrate on the possibilities for obtaining diffraction patterns from regions 1 nm or less in diameter (i.e., on nanodiffraction).

Such patterns can be obtained from any selected region of a STEM image merely by stopping the scan at the designated spot, or else series of patterns can be recorded at up to TV rates (i.e.,  $30\text{ s}^{-1}$ ), as the electron beam is scanned slowly across the specimen, by use of a low-light-level TV camera and a video cassette recorder (VCR) [47]. Diffraction patterns such as those of Fig. 7 were obtained with a nominally 1 nm beam; that is, with a beam convergence angle of about 10 mrad so that the beam at the specimen has a half-width at half-height of 0.7 nm and a diameter, measured to the first zero intensity, of 1.6 nm. For a larger objective aperture size the beam diameter at the specimen can be made as small as about 0.2 nm, but then the diffraction spots are so large that they overlap for even small unit cell sizes and it is no longer possible to recognize and measure the spot geometries easily. For large unit cell sizes, a smaller aperture size is preferred.



**Figure 7.** Nanodiffraction patterns obtained as a beam of diameter 1 nm is scanned across a carbon nanotube of diameter approx. 20 nm showing the variation of the structure from the middle of the tube at (a) to one edge of the tube at (d) [92].



**Figure 8.** Nanodiffraction patterns obtained as a beam of diameter 1 nm is scanned across the edge of an MgO crystal showing the open circle form of the diffraction spots when the beam covers the crystal edge.

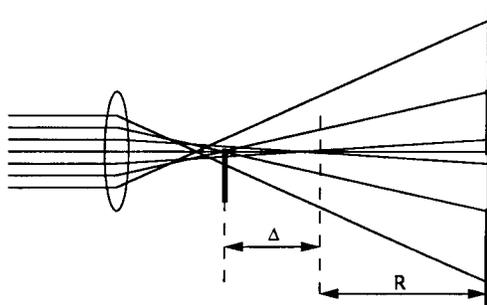
Imperfections in crystals, including crystal edges, faults or other defects, give scattering in between the Bragg diffraction spots of the pattern. Then coherent interference effects, characteristic of the type of defect, are seen. When the region of the specimen illuminated by the incident beam includes the straight edge of a crystal, it is to be expected from kinematical diffraction theory for a parallel incident beam that all diffraction spots are accompanied by streaks perpendicular to the crystal edge. For a coherent convergent incident beam, however, the diffraction disks become hollow circular rings, as in Fig. 8. A crystal fault plane, or an out-of-phase domain boundary in a superlattice structure, gives a similar form for spots corresponding to the particular set of crystal lattice planes for which the defect represents a discontinuity [48, 49]. The spots take on rather more complicated shapes if the beam illuminates part or all of a small crystal of regular shape [50]. The spot shape depends on the wedge angle if the boundary of the crystal is given by two planar surfaces meeting at an edge. Also the lateral extent of the edge is significant.

It follows that nanodiffraction can provide a very effective means for studying very small regions of crystals, such as those containing individual defects, or crystals

which have dimensions in the range of 1 to 10 nm, such as, for example, the small metal crystals in supported metal catalyst samples. The nanodiffraction is often coupled with STEM imaging, especially in HAADF mode, which provides a very effective means for locating small heavy-metal particles, and with EELS or EDS microanalysis for determining the particle composition.

### 2.2.6 Coherent Nanodiffraction, Electron Holography, Ptychology

As the objective aperture size in a STEM instrument is increased, the diffraction spots in the diffraction pattern from a crystal increase in diameter, overlap, interfere and become less distinct, until the periodicity of the crystal is no longer evident. Then the incident beam diameter is much less than the crystal unit cell dimensions, as viewed from the incident beam direction. The pattern may be considered as given only by that grouping of atoms illuminated by the beam. The symmetry and intensity distribution in the pattern change as the beam is moved about within the unit cell [51].

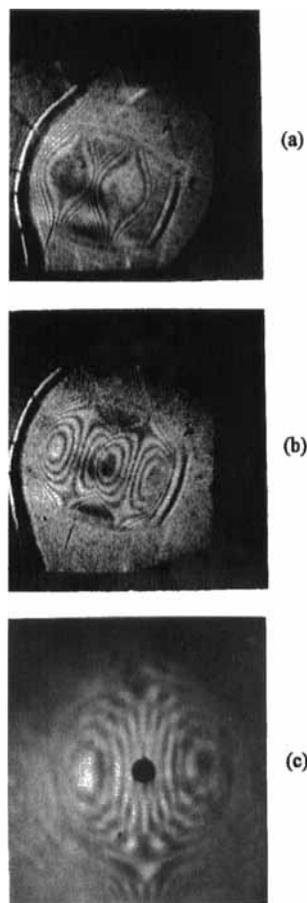


**Figure 9.** Geometrical optics ray diagram of the shadow-image formation with strong spherical aberration. The infinite-magnification circle is formed when adjacent rays cross over at the specimen plane. Within this circle, the image of the straight-edge is inverted.

For a very large, or no, objective aperture, the diffraction pattern turns into what appears to be a 'shadow image' of the specimen with a magnification, on axis, proportional to the inverse of the defocus value. Off-axis, the magnification is affected by the spherical aberration of the objective lens as suggested by Fig. 9. In over-focus (i.e., with the objective lens current increased) the magnification decreases uniformly from an axial maximum. For under-focus, the magnification increases from the axial value and goes to infinity for some particular angle (depending on the defocus and  $C_s$  values), before reversing sign and decreasing with increasing angle. Figure 9 refers only to radial magnification which becomes infinite when two beams which are adjacent in the radial direction have a crossover at the specimen plane. An infinity in magnification in the circumferential direction, given when adjacent beams having the same distance from the axis cross over at the specimen position, occurs at a radius in the image  $\sqrt{3}$  times as great. This infinite magnification gives rise to a very obvious infinite-magnification circle, which is of

great practical use for determining the defocus and for correcting the lens astigmatism [52].

For thin crystalline specimens, the defocused images show patterns of lattice fringes which becomes distorted by the spherical aberration, astigmatism or other aberrations to give the characteristic bowed shapes of Ronchi fringes (Fig. 10), similar to those familiar from light optics



**Figure 10.** Electron Ronchi fringes formed in the shadow image (or in-line hologram) of a crystal. (a), (b) Ronchi fringes from the (200) planes of MgO at different defocus values; (c) Ronchi fringes from the 0.8 nm planes of beryl, showing the zero-contrast ellipses [54].

and used for the testing of large telescope mirrors [53]. In STEM instruments, such patterns are highly useful for correcting astigmatism and for determining  $C_s$  and defocus values with high accuracy [54, 55].

The shadow images formed with a stationary beam in STEM instruments contain all the coherent interference effects between the transmitted incident beam and the beams scattered by the specimen. They are exactly what Gabor [56, 57] called holograms. Gabor proposed the method of holography as a means for overcoming the limitation of electron microscope resolution due to the irreducible aberrations of the objective lens. He realized that the aberrations produce a perturbation of the phase of the waves interfering to form the image and could possibly be removed if the phases of the waves could be determined by interference with a reference wave of known form. He considered that if a thin object was illuminated by a very small source, the directly transmitted ('unscattered') wave could serve as a reference wave and the effects of its interference with the relatively weak waves scattered by the specimen would be recorded in the hologram.

Gabor proposed a method for reconstructing the original transmission function of the object from the hologram by use of a light-optical system having a built-in spherical aberration to correct for the aberration of the electron-optical lens, thereby enhancing the resolution of electron microscopes. Gabor's suggested technique became feasible only after the development of field-emission guns provided sufficiently small, bright sources of electrons. It was then realized using digitized holograms and computer processing in place of the optical reconstruction methods [58].

The main difficulty associated with the Gabor scheme of holography is that, because it is the intensity distribution of the hologram which is recorded, the reconstructed image is accompanied by a defocused, aberrated 'conjugate' image which confuses the background. Later, Leith and Upatnieks [59] introduced the idea that, instead of the incident, transmitted wave, a wave passing through vacuum, outside of the object, could act as the reference wave and form a pattern of interference fringes with the wave scattered by the specimen to establish the relative phases of the scattered beams. This opened the way for the major development of holography with light waves as well as other radiations, including high energy electrons, as described below. However, the possible use of the coherent diffraction patterns and shadow-images, or in-line holograms, for image resolution enhancement still has considerable appeal.

In normal CTEM or STEM imaging, one image signal is obtained for each point of the image. However, in a STEM system, a complete two-dimensional array of data, including interference information on relative phases of scattered beams, is given for each incident probe position. In principle, it should be possible to make use of all the information contained in these diffraction patterns to greatly enhance the amount of information derived concerning the specimen. The term ptychology has been introduced to describe various methods by which this might be done [29].

It was pointed out above, that the interference of adjacent overlapping spots in the diffraction pattern of a thin single crystal could give relative phases and so allow structure analysis of the crystal. Kohnert et al. [60] proposed a more

sophisticated method by which the diffraction patterns from a series of closely-spaced positions of the incident beam could be Fourier transformed to give a series of correlated real-space autocorrelation functions from which the structure of a crystal and defects in the structure might be derived. This process was carried out to give a reconstruction of the [1 1 0] projection of a Si crystal lattice with a resolution of 0.1 nm or better.

A related approach to the problem has been proposed by Rodenburg and Bates [61]. From Eq. (2), if the diffraction pattern is recorded as a function of the two-dimensional vector,  $\mathbf{u}$ , at all the incident beam positions defined by the two-dimensional vector,  $\mathbf{X}$ , a four-dimensional function is created which, in the projection-function approximation used for Eqs. (4) and (5), may be written

$$\begin{aligned} I(\mathbf{u}, \mathbf{X}) &= |T(\mathbf{u}) \exp(2\pi i \mathbf{u} \cdot \mathbf{X}) - P(\mathbf{u}) \\ &\quad * T(\mathbf{u}) \exp(2\pi i \mathbf{u} \cdot \mathbf{X})|^2 \\ &= |T(\mathbf{u})|^2 - \int T^*(\mathbf{u}) P(\mathbf{U}) T(\mathbf{u} - \mathbf{U}) \\ &\quad \times \exp(-2\pi i \mathbf{U} \cdot \mathbf{X}) d\mathbf{U} - \text{c.c.} \quad (10) \end{aligned}$$

Fourier transforming this function with respect to  $\mathbf{X}$  gives the four-dimensional function,

$$\begin{aligned} G(\mathbf{u}, \rho) &= |T(\mathbf{u})|^2 \delta(\rho) - T^*(\mathbf{u}) \\ &\quad \times T(\mathbf{u} - \rho) P(\rho) - \text{c.c.} \quad (11) \end{aligned}$$

Then, because  $|T(\mathbf{u})|^2 = 1$  if the aperture limitation is neglected, the section of this function represented by  $G(\rho/2, \rho)$  is just  $\delta(\mathbf{u}) - P(\mathbf{u}) - P^*(\mathbf{u})$ , or, in the weak phase-object approximation, the Fourier transform of  $1 + 2\sigma\phi(\mathbf{X})$ .

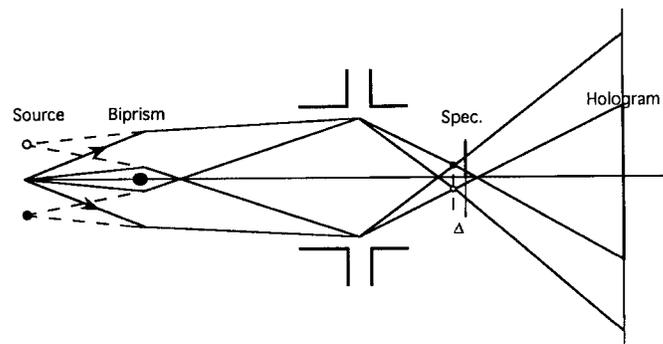
The validity of this approach has been verified by Rodenburg et al. [62] using

light-optical analogue experiments and STEM experiments with moderate resolution. Complications arise, for example, from the fact that the aperture limitation cannot be ignored, but evidence that the method may provide considerable enhancement of resolution has been given.

## 2.2.7 Holography

The concept of holography can be applied with a STEM instrument in a number of different ways including both in-line and off-axis forms [63]. As pointed out above, the shadow images formed with a stationary incident beam are just the holograms of Gabor's original proposal of the holography concept [56, 57]. If the shadow images are recorded with large defocus and so with moderate magnification, the coherent diffraction effects are visible. In particular, Fresnel fringes occur at any discontinuities and the effects are seen of any deflection of the incident beam by potential field gradients. This is the basis for the Fresnel imaging of the magnetic domain boundaries in thin films of ferromagnetic materials [64]. Reconstruction from such Fresnel images to give the phase variations in the object with high spatial resolution is possible but, as in the original Gabor form of holography, the desired image is confused by an unwanted, defocused conjugate image.

If the beam in a STEM instrument is scanned over the specimen and the diffraction pattern, or near-focus shadow image, is recorded for a closely-spaced set of beam positions, reconstruction, with correction of aberrations, may be made for each recorded pattern and correlation between



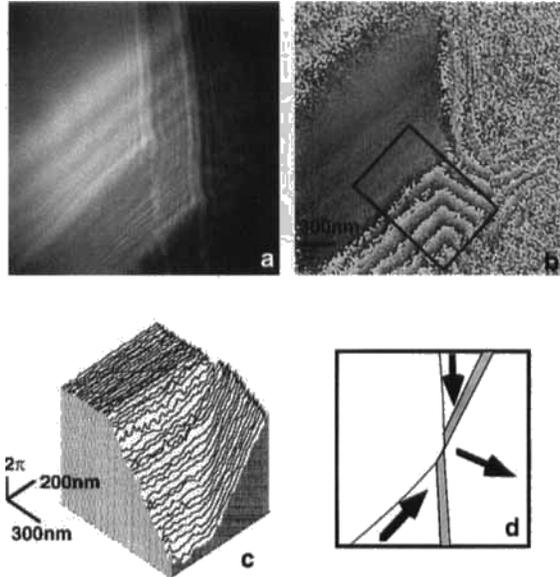
**Figure 11.** Diagram of the arrangement for off-axis STEM holography. The insertion of an electrostatic biprism near the source produces two coherent effective sources and two probes at the specimen plane, giving interference fringes in the plane of observation.

patterns can be used to reduce, and effectively eliminate, the effects of the conjugate images [55]. An alternative approach is that of Veneklasen [65], who proposed that by making a detector with a detector function  $D(u) = T(u)$ , an on-line reconstruction of the wave-function may be made, giving a STEM image directly proportional to the projected potential if the weak phase object approximation is valid. Analysis of this scheme suggests that the desired image will be accompanied by a weak, defocused and aberrated conjugate image, which may not be a serious impediment, and that the difficulty of making a detector function having the complex form of  $T(u)$  may be avoided [20]. However, this scheme has not yet been realized in practice.

An off-axis form of STEM holography, equivalent in some respects to the off-axis form of CTEM holography (Sec. 1.8 of this Chapter), may be realized, in either stationary-beam or scanning-beam form, if an electrostatic biprism is inserted in the illumination system before the specimen. The biprism has the effect of producing two coherent virtual sources so that two mutually coherent probes are formed at the specimen level, with a variable separation so that one may pass through the

specimen and the other may pass through vacuum and act as a reference wave (Fig. 11). The essential difference from the CTEM off-axis holography scheme is that the pattern of interference fringes is formed in the diffraction plane rather than the image plane; but equivalent reconstruction schemes, giving equivalent enhancement of image resolution by correction for aberration effects, are possible [66].

For greatly defocused shadow images, the off-line holography scheme gives a set of interference fringes superimposed on what can be regarded as an image plane and reconstruction of the phase and amplitude distribution of the wave transmitted through the specimen can be made as for the CTEM case. This has been the basis for a very effective means for the imaging and quantitative measurement of the magnetic fields in thin films and around small particles of ferromagnetic materials, as illustrated in Fig. 12 [67]. The spatial resolution of the phase distributions derived from the holograms is limited by the Fresnel diffraction effects if no correction is made for the effects of defocus, but this limited resolution may be sufficient for many observations of domain boundaries. However, if the reconstruction process



**Figure 12.** Off-axis hologram in a STEM instrument with a stationary beam obtained from a thin ferromagnetic film (a), and the reconstruction (b), showing the variation of the magnetic field around a domain boundary. The contrast goes from black to white for phase changes of  $2\pi$  (i.e., it is a 'wrapped' image). (c) An unwrapped image; (d) the magnetic field directions [67].

includes a correction for the defocus effect, the resolution achieved can be 1 nm or better [68].

For the scanning mode in STEM, with a biprism used to produce an interference fringe pattern, a reconstruction of the phase and amplitude components of the specimen wave can be achieved, without the complication of a conjugate image, by use of a special detector configuration such as an elaboration of that proposed by Veneklasen [65].

A much simpler form of detector, consists of a set of fine parallel lines, has been used effectively by Leuthner et al. [69]. The central part of the set of fine interference fringes forming the hologram, or an optically magnified and projected image of it, falls on a grating of black and transparent lines of matched spacing and the transmitted intensity is detected with a photomultiplier. Any variation of the phase of the specimen wavefunction relative to that of the reference wave gives a shift of the

fringes and hence a variation of the intensity transmitted through the grating. The averaged intensity of the fringes gives the amplitude of the wave function. This approach has given phase distributions and contoured maps of potential distributions with a spatial resolution equal to that for bright-field imaging of the STEM instrument used.

For all forms of holography, whether in CTEM or STEM, the process of reconstruction to enhance the resolution of the image relies on an accurate knowledge of the imaging parameters. For the current level of attainable image resolution, the most important parameters to be considered are the astigmatism, the defocus and the spherical aberration coefficient. For any improvement of resolution by holographic means, or for the quantitative interpretation of images obtained by any method, the requirements for accurate knowledge of these parameters become very exacting [70]. For STEM, the

observation of shadow images, particularly for the periodic structures of crystals, provides a very convenient means for the correction of astigmatism and the measurement of defocus and spherical aberration.

In underfocused shadow images, as suggested by Fig. 4, there is a well-defined circle of infinite magnification. In particular, for the straight edge of a specimen, this circle is clearly defined and is bisected by the image of the edge with an inversion of the image from the inside to the outside of the circle. The light-optical analogue of this is the well-known 'knife-edge' test used to detect astigmatism of optical lenses. Astigmatism distorts the circle into an S-shaped curve. The astigmatism in the STEM objective lens is corrected by removing such a distortion in the shadow image. An even more stringent test for astigmatism is given by the observation of the Ronchi fringes in the shadow images of thin crystals in that any astigmatism leads to a distortion of the characteristic form of the fringes near focus, illustrated in Fig. 10 [71].

The form of the Ronchi fringes is strongly dependent on both the defocus and  $C_s$ . It was pointed out by Lin and Cowley [54] that these parameters can be determined separately by making observations under special conditions. Thus, for the Ronchi fringes from crystal lattice planes of moderately large spacing ( $\sim 0.7$ – $1.0$  nm) it is seen that the fringe contrast goes to zero and reverses on a set of concentric ellipses (Fig. 10c). The ratio of the squares of the major axes of any two consecutive ellipses depends only on the spherical aberration constant,  $C_s$ , which can therefore be determined independently of the defocus. Once the value of

$C_s$  is determined in this way, the defocus can be derived from the shape of the Ronchi fringes and the number of fringes within the infinite magnification circle [55].

## 2.2.8 STEM Instrumentation

Apart from the early instruments made in individual laboratories, following the original designs of Crewe et al. [1–3], all of the current dedicated STEM instruments are those made by the one commercial manufacturer, VG Microscopes (UK). STEM attachments are available for a number of commercial CTEM instruments but, even when a field-emission gun is provided, the STEM performance tends to be limited by the essential features of the design which are optimized for the CTEM performance. The main imaging field of the objective lens comes after, rather than before, the specimen. Probe sizes at the specimen level of 1 nm or less can be achieved but the intensity within such a probe is usually much less than for the dedicated STEM instrument and insufficient for convenient nanodiffraction or microanalysis. Hence our discussion here will be limited to the one commercial class of dedicated STEM machines.

The cold field emission gun is operated at up to 100 keV (or in a few cases, up to 300 keV) and is placed at the bottom of the column for the sake of mechanical stability. Since the effective source size for a FEG is about 4 nm, it is clear that mechanical vibration of the field-emission tip with an amplitude of 1 nm or less may have an adverse effect on the achievable resolution. Since a large demagnification of the effective electron source is not required, only

two weak condenser lenses are used before the strong objective lens which forms the small probe on the specimen. The condenser lenses allow the choice of a range of probe sizes and intensities to suit the particular application.

The insertion of an objective aperture between the pole pieces of the objective lens, as is done for a CTEM instrument, is often considered undesirable because it limits the space available for specimen manipulators and also may produce unwanted X-ray signals, complicating the interpretation of EDS spectra and X-ray images. Hence, a virtual objective aperture is placed in the space before the condenser lenses.

In many STEM instruments, the electrons passing through the specimen are allowed to diverge until they reach the detector plane, with no lens action except for that of the weak postspecimen field of the asymmetrical objective lens. However, for the convenient observation of diffraction patterns or shadow images with variable magnification, or for holography, it is convenient to have at least two postspecimen lenses. Some recent instruments, in fact, have a symmetrical objective lens so that the postspecimen field of the lens is strong, plus two intermediate lenses and a strong projector lens. The postspecimen lens system is then equivalent to that of a normal CTEM instrument and there is great flexibility for the various detector configurations.

The easy access to the top of the STEM column provides the necessary flexibility in the construction of the detector system. The minimum detection system includes a phosphor or scintillator screen for observation of the diffraction pattern, recorded by means of a low light-level TV camera

with a VCR, and preferably a similar, alternate screen with an aperture which allows a selected part of the diffraction pattern to pass through to the EELS spectrometer for bright-field or dark-field imaging with filtered, elastically scattered electrons, for imaging with selected energy-loss electrons, or for EELS microanalysis of selected regions of the specimen. In addition there should be one, or preferably several, interchangeable, annular detectors to provide flexibility in ADF, HAADF imaging or special bright-field imaging modes.

For quantitative recording of diffraction patterns, shadow-images or holograms, the scintillator-TV combination is replaced, interchangeably by a CCD camera system. With current CCD systems the scan rate is slow so that one image frame is recorded in one second or often longer and no provision is made for switching to a fast scan. Since fast scans at TV rates are essential while searching for the desired specimen area, or for focusing and stigmating, it is important to be able to switch quickly from a TV to the CDD detector. Provision is also desirable for introducing specially shaped detectors or masks to implement the several suggested imaging modes involving detectors of non-standard configuration, such as those of Leuthner et al. [69] or Veneklasen [65]. A detector system designed to have all these desirable features with reasonably efficient signal collection has recently been described [20] but is undergoing redesign for further improvement.

The usual asymmetrical design for the STEM objective lens has some advantages in allowing easier access to the specimen region for the EDS X-ray detector, which, for efficiency of signal collection, needs to

be as close to the specimen as possible. Also it is relatively easy with this lens configuration to hold specimens in the space outside the strong magnetic field of the objective lens so that their magnetic configurations can be observed under controlled conditions of externally applied magnetic field [64, 67].

The signal-handling capabilities needed for STEM have been well-developed in many cases. Digital scans may be substituted for the conventional analog scans. The images produced in the scanning mode may be displayed directly on cathode ray tubes and recorded with parallel photographic monitors. Signals from several detectors may be displayed and recorded simultaneously, or may be combined by addition, subtraction or multiplication and recorded for special purposes. Alternatively, the signals may be digitized and recorded for further processing, on-line or off-line, by attached computer systems. To overcome the problem of a fluctuating emission from the electron gun to some extent, a reference signal may be obtained from, for example, the electrons striking the objective aperture blade but not transmitted: then the ratio of the image signal to this reference signal is recorded.

For the stable operation of a cold field-emission gun, the vacuum in the gun chamber must be better than  $10^{-10}$  torr and the column vacuum must be  $10^{-8}$  or better to prevent excessive backstreaming of gas molecules into the gun. The STEM instruments have a column vacuum of better than  $10^{-8}$  for this reason and also in order to minimize the effects of contamination of the specimen. For poorer vacuum, or with specimens which have not been adequately cleaned, the contamination rate can be very

high when the incident beam illuminates only a very small part of the specimen. Then the migration of organic molecules along the specimen surface is not prevented by the flooding of a large specimen area around the area of interest by a broad electron beam, as is common in CTEM. With care, however, in a STEM instrument the contamination rate can be reduced to a level for which no accumulation of contaminant is observable, even with a focused probe, over a period of many hours.

With special care in the construction and operation, STEM instruments can have a column vacuum of better than  $10^{-9}$  torr, and in at least one case, an instrument has been made with a completely UHV system, with better than  $10^{-10}$  in the whole system, including an elaborate auxiliary chamber for specimen preparation and manipulation [72]. This instrument, the so-called MIDAS system (a microscope for imaging, diffraction and analysis of surfaces), was specifically designed for research on surface structures with high spatial resolution. The various STEM imaging modes, nanodiffraction and EELS microanalysis can be performed on surfaces and thin films which are clean to the exacting standards demanded in surface science.

In addition, SEM and SAM signals can be collected and analyzed with respect to their energy distributions, with high spatial and energy resolutions, through the application of magnetic 'parallelizers'. The low-energy electrons emitted from the sample in the high magnetic field of the objective lens spiral around the lens axis as they drift out of the field. If the field is properly shaped, the spiral angle of the electrons decreases with the magnetic field (i.e., the electrons are parallelized) and

when they emerge into field-free space they are confined to a cone a few degrees wide, ideally suited for feeding into a low-energy electron spectrometer for energy analysis and detection with high efficiency [73]. The spatial resolution for SAM images obtained in this way has reached the level of 1 nm, an order of magnitude better than for other instruments [44].

As in the case of CTEM, images of the surfaces of crystals can be obtained in STEM instruments if a bulk specimen is mounted so that the incident beam strikes an extended, nearly flat surface at a grazing angle of incidence. The diffraction pattern formed on the detector plane is then a convergent beam reflection high energy electron diffraction (RHEED) pattern, similar to a normal RHEED pattern except that the individual diffraction spots are replaced by diffraction disks. If a region of high intensity within one of the strong RHEED spots is selected to give the STEM signal, the scanning reflection electron microscopy (SREM) image produced resembles the corresponding REM image formed in a CTEM instrument (see Sec. 1.2 of this Chapter).

It is interesting to note that the highest intensity regions within an extended convergent-beam reflection high-energy electron diffraction (CBRHEED) spot often correspond to incident beam orientations for which strong surface resonance effects occur. As in REM, the image intensity is highly sensitive to small defects on the crystal surface, such as steps one atom high or more, the strain fields of emerging dislocations or other crystal faults, and variations of the structure or composition of the topmost one or two layers of atoms on the surface [74]. As compared with REM, SREM has the advantage that

SEM and SAM signals are produced at the same time, or subsequently with higher angles of incidence, to give complementary information on the surface morphology and composition. The only modification required for a STEM instrument to be used for SREM is that the specimen should be mounted with an extended flat face nearly parallel to the incident beam.

## 2.2.9 Applications of Scanning Transmission Electron Microscopy

Although various applications of STEM will be included in Volume 2 of this Handbook, a few will be mentioned here to illustrate the capabilities of the technique and also illustrate the nature of the information that it can provide.

In the initial work by Crewe's group [75] it was clearly established that it was possible with ADF imaging to obtain clear images of individual heavy atoms on light atom supports (Fig. 1). The heavy atoms could be as light as Ag [76]. Several possible applications immediately suggested themselves. It was considered, for example, that if heavy atoms were attached to particular known sites on biological macromolecules it would be possible to deduce something about the structures of the molecules from the images of the heavy atoms even if the molecules themselves were destroyed by radiation damage in the course of the imaging. Some limited success was achieved in this endeavor, particularly by Beers et al. [77]. However, questions arose as to whether the heavy atoms could remain in their original sites

while the molecules around them disintegrated.

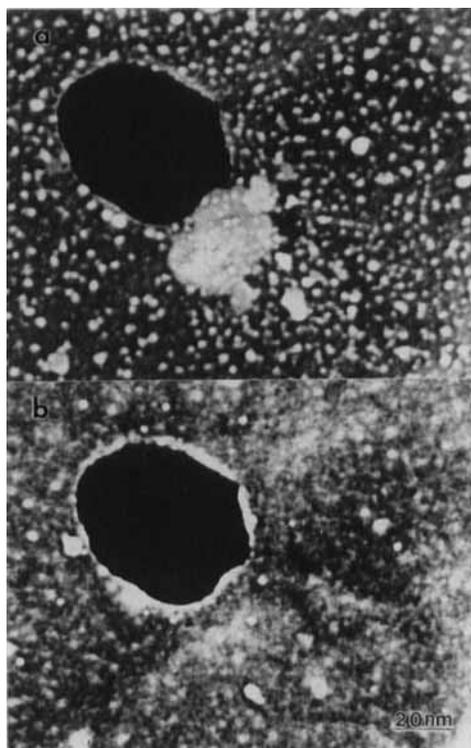
In observations of heavy atoms sitting on thin films of amorphous carbon, it appeared that the heavy atom positions tended to change from one image scan to the next [78]. Experiments were made to determine whether the atom movements resulted from energy pulses given to the atoms by the incident electron beam by inelastic scattering events, or whether they derived from random thermal vibration excitations. The conclusion was that the latter was more likely.

The quantitative nature of the ADF image signal in a STEM instrument gave rise to the suggestion that measurements of image intensities could be used to determine the masses of small particles and, in particular, of individual macromolecules [79]. To a good approximation, the intensity collected by the annular detector is proportional to the total scattering from a nearly amorphous object, since the intensity in the diffraction pattern falls off smoothly with scattering angle. The mass per unit area of the specimen could be derived from measurements of the scattered intensity by comparison with the scattering cross section per unit mass, derived from the theoretical atomic scattering cross sections. Measurements made on particles of known structure gave good agreement with the masses derived from other methods. On this basis, a large number of valuable mass determinations for a wide variety of biological particulates have been made, particularly by Wall (who initially proposed the method) and associates [80]. See also [81].

A widely explored application of STEM has been for the study of small particles and particularly of the small particles of

heavy metals in light-atom supports, such as the active components in supported metal catalysts. These particles have sizes typically in the range 1–10 nm. When such particles are embedded in the near-amorphous support material, or even when they are placed on the usual amorphous carbon supporting films, it is difficult to measure, or even detect them in bright-field CTEM or STEM images because the phase-contrast noise from the amorphous material gives random intensity fluctuations on much the same scale. The ADF, and especially the HAADF, imaging mode can reveal such particles clearly even when they are completely obscured in BF images by relatively large thickness of the support material.

Once a particle is detected in an HAADF STEM image, it may be selected for analysis by nanodiffraction, EELS or EDS, when the incident beam is stopped at that point. In addition, in specially equipped instruments, the HAADF image may be correlated with SEM or SAM images of the same specimen area to determine whether the particle is on the top or bottom surface of the sample or in its interior. This information is often of importance in that it can indicate the extent to which the catalyst particle may be exposed to the reacting gaseous atmosphere. Figure 13 shows SEM images from the top side and bottom side of a sample of gold particles on a thin carbon support [43]. Transmission images give no indication as to the relative positions of the gold particles and the carbon film in the beam direction. The SEM images, obtained simultaneously with detectors above and below the specimen, indicate clearly that the gold particles are all on the bottom side of the supporting film, as are some



**Figure 13.** SEM images obtained from the entrance (a) and exit (b) surfaces of a carbon film with small metal particles and some light-atom material on the entrance side (Courtesy of J. Liu [96]).

light-atom particles, visible in one image but not the other. It may be noted that weak, diffuse images of the gold particles appear in the SEM images even on the 'wrong' side of the thin carbon film, presumably because some of the low-energy secondary electrons can penetrate through the film.

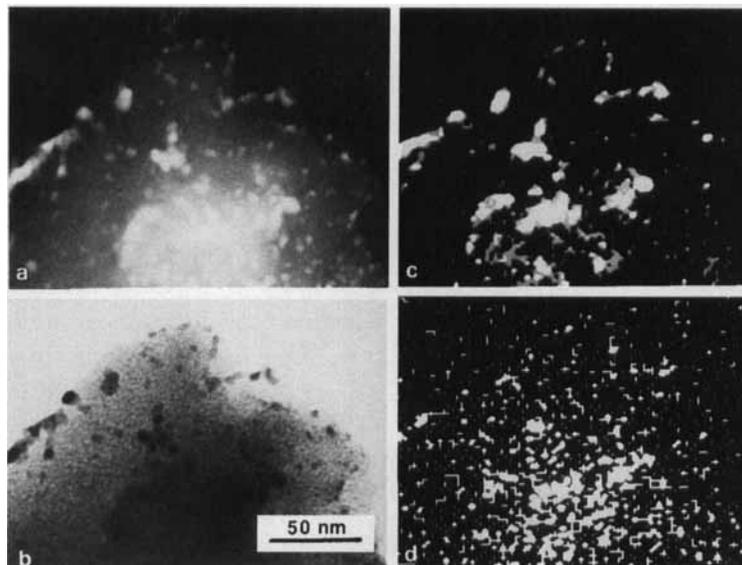
The problem of locating heavy-metal atoms or small clusters within the channels in the structures of zeolites becomes more significant as the use of zeolites in industrial catalysts becomes more extensive. A major difficulty for the application of electron microscopy techniques to such materials is that the zeolite structural

framework is rapidly destroyed by electron irradiation. It is necessary to make use of minimum exposure techniques so that the image information is obtained from the first electrons to strike the sample area of interest. This can readily be achieved in STEM for which only the area of the specimen actually being scanned is exposed to the electron beam. Initial location of interesting specimen areas can be done at low magnification with low irradiation doses. Final focusing and stigmation can be done on adjacent areas before the beam is switched to the region of interest where a single scan is made and used to record the image. A comparison of CTEM and HAADF STEM imaging by Rice et al. [82] showed the latter to be more effective in revealing few-atom clusters or small metal particles in zeolites. An additional advantage is that nanodiffraction patterns from the particles could be made immediately after the image, to give some information on the relative orientations of the small metal crystals and the walls of the zeolite channels [83].

In Fig. 14, an EDS image obtained with the characteristic X-rays from small Pd particles on a carbon support [84], it is evident that particles as small as 2–5 nm in diameter may be detected, especially when a digital imaging technique is applied. With the intrinsically better detection efficiencies of electron energy-loss imaging, even smaller particles may be imaged [85]. In addition, for very small regions, it is also possible to detect and make use of the fine structure of the energy-loss peaks which is characteristic of the particular valence states or bonding configurations of the atoms [86] (Sec. 1.3 of this Chapter).

Small particles of metals are often single crystals. However, there are many CTEM

**Figure 14.** (a) Annular dark-field dedicated STEM image of Pd particles on a carbon support; (b) bright-field STEM image; (c) Pd L- $\alpha$  X-ray image with background X-ray intensity subtracted; (d) background X-ray image of the same area with signal intensity expanded ten times more than (c). Digital images (c) and (d) were smoothed, which caused line artifacts in the low-intensity image (d). Original magnification = 400 000 $\times$ . (Courtesy of C. E. Lyman [84]).



observations to suggest that particles of noble metals, and particularly of gold, in the size range 10–100 nm are often twinned or multiply twinned (see, for example, Allpress and Sanders [87]). They may form decahedral particles, with five tetrahedrally shaped regions of perfect crystal related by twinning on (111) planes, or icosahedral particles with twenty tetrahedral, mutually twinned regions. Considerations of the energetics of the small particles suggest that the multiply twinned forms may be the equilibrium forms for very small particles, but the evidence from CTEM is inconclusive for particle sizes below about 10 nm. For one sample, formed by co-sputtering of gold and polyester, the STEM images showed a range of particle sizes down to about 2 nm or less. Nanodiffraction from the individual particles revealed that for 3–5 nm particles about half were multiply twinned, but the proportion was even less for smaller particles in the range 1.5–3 nm [88]. This result was not necessarily in

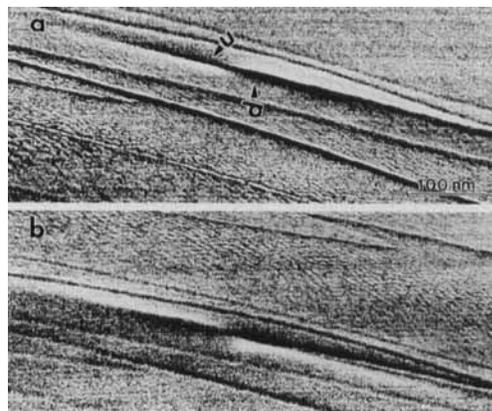
disagreement with the theoretical studies which referred only to isolated particles with no perturbation of surface energies by support material.

The possibility of recording nanodiffraction patterns at TV rates by using a low light-level TV camera and VCR has allowed detailed studies of several small-scale structures. The structure of the individual planar defects in diamond, thought to be associated with the aggregation of nitrogen atoms, was investigated in this way by Cowley et al. [89]. Patterns were recorded at intervals of 0.02 nm as a beam of 0.3 nm diameter was scanned across a defect with an instrumental magnification of  $2 \times 10^7$ . The observed intensities were compared with dynamical diffraction calculations based on the various models which had been proposed for the defect. The best fit was obtained for a model due to Humble [90] which contained no nitrogen atoms.

Similar series of nanodiffraction patterns were obtained with a beam of

diameter about 1 nm scanned across various carbon nanotubes, formed in a carbon arc discharge in high-pressure He and shown by Iijima [91] to be made up of concentric cylinders of graphitic sheets with internal diameters from about 2 to 10 nm and external diameters from 3 to 20 nm (Fig. 7). It was known that the individual sheets had a helical structure and several different helical angles could be included in one nanotube. It had been assumed that the cylinders had circular cross section. Because a regular graphite crystal structure cannot be consistent with a superposition of graphite sheet cylinders with regularly increasing circumference, it was concluded that the lateral arrangement of the sheets must be disordered. From the nanodiffraction patterns it was evident that, for many nanotubes, the cylinders were not circular but polygonal, and probably pentagonal, in cross section [92]. It was seen that, in the regions of the flat faces of the cylinders, the graphitic sheets were ordered as in the crystalline graphite structure. The nanotubes having polygonal cross section included non-helical ones and ones having one or more helical angles.

The scanning reflection electron microscopy (SREM) mode has been applied in various studies of surface structure (Liu and Cowley [74]). Single-atom-high surface steps and the strain fields of emerging dislocations have been observed with the same characteristic contrast variations with diffraction conditions as in REM (see Sec. 1.2 of this Chapter). Long-period surface superstructures on oxygen-annealed cleavage faces of MgO have been observed and studied (Fig. 15). Interesting differences in image contrast have been seen for SREM and the allied



**Figure 15.** Scanning reflection electron microscopy (SREM) image of the (100) face of MgO crystal showing the fringes due to a surface superlattice in two directions and the reversal of contrast of the streaks due to the strain field around an emerging dislocation as the defocus is reversed from (a) to (b). (Courtesy of J. Liu [74]).

techniques of SEM imaging and the SREM imaging mode using a high-angle annular detector, in analogy with the transmission HAADF mode [74]. For the latter two modes there is little dependence of the image contrast on diffraction conditions and the image shows mostly the surface morphology with different characteristic intensity variations in the two cases.

As in the case of STEM, the SREM technique has the virtue that the beam scan may be stopped at any point of the image so that nanodiffraction or microanalysis with EELS or EDS may be performed on any chosen feature. In studies of the deposition of Pd metal on MgO surfaces, for example, nanodiffraction patterns from individual particles of about 2 nm diameter sitting on the surface showed them to be single crystals of Pd but, under continued electron beam irradiation, these crystals were seen to be gradually covered by a layer of different

material, identified by nanodiffraction as single crystal PdO [93]. The SREM technique has also been used by Milne [94] for studies of the oxidation of copper surfaces. The images revealed the form of the copper oxide crystallites growing on the copper surfaces and their nucleation sites relative to the steps and facets of the copper crystal cut at a small angle to the (1 0 0) surface and annealed to give a surface reconstruction.

It is evident that, for both transmission and reflection modes, the STEM instruments have particular capabilities which make them invaluable for some special purposes, particularly when the combination of imaging, nanodiffraction and microanalysis provides a much more complete assessment of a specimen than can be obtained by any other approach. The full exploitation of the coherent interference effects observable in the diffraction patterns remains as a challenge for the future.

## 2.2.10 References

- [1] A. V. Crewe, J. Wall, *J. Mol. Biol.* **1970**, *48*, 373.
- [2] A. V. Crewe, *Rep. Progr. Phys.* **1980**, *43*, 621.
- [3] A. V. Crewe, J. Wall, J. Langmore, *Science* **1970**, *168*, 1333.
- [4] J. M. Cowley, *Appl. Phys. Letts.* **1969**, *15*, 58.
- [5] E. Zeitler, M. G. R. Thompson, *Optik* **1970**, *31*, 258, 359.
- [6] J. R. Sellar, J. M. Cowley in *Scanning Electron Microscopy 1973* (Ed.: O. Johari), IIT Research Institute, Chicago **1973**, p. 243.
- [7] T. Groves, *Ultramicroscopy* **1975**, *1*, 15.
- [8] A Strojnik in *Microscopie Electronique à Haute Tension* (Eds: B. Jouffrey, P. Favard), Société Française de Micros. Electronique, Paris **1976**.
- [9] J. M. Cowley, *Ultramicroscopy* **1976**, *2*, 3.
- [10] A. Howie, *J. Microsc.* **1979**, *17*, 11.
- [11] M. M. J. Treacy, A. Howie, *J. Catal.* **1980**, *63*, 265.
- [12] Z. L. Wang, J. M. Cowley, *Ultramicroscopy* **1990**, *32*, 275.
- [13] P. Zu, R. F. Loane, J. Silcox, *Ultramicroscopy* **1991**, *38*, 127.
- [14] S. J. Pennycook, D. E. Jesson, *Ultramicroscopy* **1991**, *37*, 14.
- [15] S. Hillyard, J. Silcox, *Ultramicroscopy* **1993**, *52*, 325.
- [16] N. H. Dekkers, H. de Lang, *Optik* **1974**, *41*, 452.
- [17] J. M. Chapman, P. E. Batson, E. M. Waddell, R. P. Ferrier, *Ultramicroscopy* **1978**, *3*, 203.
- [18] I. R. McFayden, J. M. Chapman, *Electron Microsc. Soc. Amer. Bull.* **1992**, *22*, 64.
- [19] M. Hammel, H. Kohl, H. Rose in *Proc. XIIIth Internat. Congress Electron Microsc.* (Eds.: L. D. Peachey, D. B. Williams) San Francisco Press, San Francisco **1990**, Vol. 1, p. 120.
- [20] J. M. Cowley, *Ultramicroscopy* **1993**, *49*, 4.
- [21] J. M. Cowley, A. Y. Au in *Scanning Electron Microscopy, Vol. 1* (Ed.: O. Johari), SEM Inc., Chicago **1978**, p. 53.
- [22] J. Liu, J. M. Cowley, *Ultramicroscopy* **1993**, *52*, 335.
- [23] H. Rose, *Ultramicroscopy* **1977**, *2*, 251.
- [24] W. Kunath, H. Gross, *Ultramicroscopy* **1985**, *16*, 349.
- [25] C. J. Humphreys, E. G. Bithell in *Electron Diffraction Techniques, Vol. 1* (Ed.: J. M. Cowley), Oxford University Press **1992**, p. 75.
- [26] J. Barry in *Electron Diffraction Techniques, Vol. 1* (Ed.: J. M. Cowley), Oxford University Press **1992**, p. 170.
- [27] J. M. Cowley, J. C. H. Spence, *Ultramicroscopy* **1981**, *6*, 359.
- [28] J. C. H. Spence, J. M. Cowley *Optik* **1978**, *50*, 129.
- [29] W. Hoppe, *Ultramicroscopy* **1982**, *10*, 187.
- [30] R. Nathan in *Digital Processing in Biomedical Imaging* (Eds.: K. Preston, M. Onoe), Plenum Press, New York **1976**, p. 75.
- [31] P. E. Batson, *Ultramicroscopy* **1992**, *47*, 133.
- [32] F. P. Ottensmeyer, J. W. Andrew, *J. Ultrastruct. Res.* **1980**, *72*, 336.
- [33] D. B. Williams, J. R. Michael, J. I. Goldstein, A. D. Romig, Jr. *Ultramicroscopy* **1992**, *47*, 121.
- [34] M. Scheinfein, A. Muray, M. Isaacson, *Ultramicroscopy* **1985**, *16*, 237.
- [35] H. Raether, *Excitation of Plasmons and Interband Transitions by Electrons*, Springer Tracts in Modern Physics, Vol. 88, Springer, New York **1980**.
- [36] Z. L. Wang, J. M. Cowley, *Ultramicroscopy* **1987**, *21*, 77.
- [37] J. M. Cowley, *Phys. Rev. B* **1982**, *25*, 1401.
- [38] N. Zabala, P. M. Echenique, *Ultramicroscopy* **1990**, *32*, 327.

- [39] R. H. Milne, G. G. Hembree, J. S. Drucker, C. J. Harland, J. A. Venables, *J. Microsc.* **1993**, *170*, 193.
- [40] K. Kuroda, S. Hosoki, T. Komoda, *Scanning Microscopy* **1987**, *1*, 911.
- [41] J. Liu, J. M. Cowley, *Scanning Microscopy* **1988**, *2*, 1957.
- [42] M. R. Scheinfein, J. S. Drucker, J. K. Weiss, *Phys. Rev. B* **1993**, *47*, 4068.
- [43] G. G. Hembree, P. A. Crozier, J. S. Drucker, M. Krishnamurthy, J. A. Venables, J. M. Cowley, *Ultramicroscopy* **1989**, *31*, 111.
- [44] J. Liu, G. G. Hembree, G. E. Spinnler, J. A. Venables, *Ultramicroscopy* **1993**, *52*, 369.
- [45] J. M. Cowley, *Ultramicroscopy* **1979**, *4*, 435.
- [46] W. J. Vine, R. Vincent, P. Spellward, J. W. Steeds, *Ultramicroscopy* **1992**, *41*, 423.
- [47] J. M. Cowley in *Scanning Electron Microscopy 1980, Vol. 1* (Ed.: Om Johari), SEM Inc., Chicago **1980**, p. 61.
- [48] J. Zhu, J. M. Cowley, *Acta Crystallogr. A* **1982**, *38*, 718.
- [49] J. Zhu, J. M. Cowley, *J. Appl. Crystallogr.* **1983**, *16*, 171.
- [50] M. Pan, J. M. Cowley, J. C. Barry, *Ultramicroscopy* **1989**, *30*, 385.
- [51] J. M. Cowley, *Sov. Phys. Crystallogr.* **1981**, *26*, 549.
- [52] J. M. Cowley, *Ultramicroscopy* **1979**, *4*, 413.
- [53] V. Ronchi, *Appl. Optics* **1964**, *3*, 437.
- [54] J. A. Lin, J. M. Cowley, *Ultramicroscopy* **1986**, *19*, 31.
- [55] S.-Y. Wang, J. M. Cowley, *Micros. Res. Techs.* **1995**, *30*, 181.
- [56] D. Gabor, *Nature* **1948**, *161*, 777.
- [57] D. Gabor, *Proc. R. Soc. Lond.* **1949**, *A197*, 454.
- [58] J. A. Lin, J. M. Cowley, *Ultramicroscopy* **1986**, *19*, 179.
- [59] E. N. Leith, J. Upatnieks, *J. Opt. Soc. Amer.* **1962**, *52*, 1123.
- [60] J. Konnert, P. D'Antonio, J. M. Cowley, A. Higgs, H. J. Ou, *Ultramicroscopy* **1989**, *30*, 371.
- [61] J. M. Rodenburg, R. H. T. Bates, *Philos. Trans. R. Soc. Lond.* **1992**, *A339*, 521.
- [62] J. M. Rodenburg, B. C. McCallum, P. D. Nellist, *Ultramicroscopy* **1993**, *48*, 304.
- [63] J. M. Cowley, *Ultramicroscopy* **1992**, *41*, 335.
- [64] I. R. McFayden, J. N. Chapman, *Electron Micros. Soc. Amer. Bull.* **1992**, *22*, 64.
- [65] L. H. Veneklasen, *Optik* **1975**, *44*, 447.
- [66] M. A. Gribelyuk, J. M. Cowley, *Ultramicroscopy* **1992**, *45*, 115.
- [67] M. Mankos, M. R. Scheinfein, J. M. Cowley, *J. Appl. Phys.* **1994**, *75*, 7418.
- [68] M. Mankos, P. de Haan, V. Kambersky, G. Matteucci, M. R. McCartney, Z. Yank, M. Scheinfein, J. M. Cowley in *Electron Holography* (Eds.: A. Tonomura et al.), Elsevier Science BV **1995**, p. 329.
- [69] Th. Leuthner, H. Lichte, K.-H. Herrmann, *Phys. Stat. Solidi (a)*, **1989**, *116*, 113.
- [70] H. Lichte, *Ultramicroscopy* **1993**, *51*, 15.
- [71] J. M. Cowley, M. Disko, *Ultramicroscopy* **1980**, *5*, 469.
- [72] J. A. Venables, J. M. Cowley, H. S. von Harrah in *Inst. Phys. Conf. Series, No. 90, Chapter 4*, IOP Publishing, Bristol **1987**, p. 81.
- [73] P. Kruit, J. A. Venables, *Ultramicroscopy* **1988**, *25*, 183.
- [74] J. Liu, J. M. Cowley, *Ultramicroscopy* **1993**, *48*, 381.
- [75] J. P. Langmore, J. Wall, M. S. Isaacson, *Optik* **1973**, *38*, 335.
- [76] M. Retsky, *Optik* **1974**, *41*, 127.
- [77] M. D. Cole, J. W. Wiggins, M. Beer, *J. Mol. Biol.* **1978**, *117*, 378.
- [78] M. Isaacson, D. Kopf, M. Utlaut, N. W. Parker, A. V. Crewe, *Proc. Nat. Acad. Sci. USA* **1977**, *74*, 1802.
- [79] M. K. Lamvik, J. P. Langmore, *Scanning Electron Microscopy 1977, Vol. 1* (Ed.: Om Johari), IIT Research Institute, Chicago **1977**, p. 401.
- [80] A. V. Crewe, J. Wall in *Proc. 29th Annual Meeting, Electron Micros. Soc. Amer.* (Ed.: C. J. Arceneaux), Claitors Publ. Divis., Baton Rouge **1971**, p. 24.
- [81] G. J. Brakenhoff, N. Nanninga, J. Pieters, *J. Ultrastructure Research* **1972**, *41*, 238.
- [82] S. B. Rice, J. Y. Koo, M. M. Disko, M. M. J. Treacy, *Ultramicroscopy* **1990**, *34*, 108.
- [83] M. Pan, J. M. Cowley, I. Y. Chan, *Ultramicroscopy* **1990**, *34*, 93.
- [84] C. E. Lyman, *Ultramicroscopy* **1986**, *20*, 119.
- [85] E. Freund, J. Lynch, R. Szymansky, *Ultramicroscopy* **1986**, *20*, 107.
- [86] P. Rez in *Physical Methods of Chemistry, Vol. 4, Microscopy* (Eds.: B. W. Rossiter, J. F. Hamilton), John Wiley & Sons, New York **1991**, p. 203.
- [87] J. A. Allpress, J. V. Sanders, *Surf. Sci* **1965**, *7*, 1.
- [88] W. B. Monosmith, J. M. Cowley, *Ultramicroscopy* **1984**, *12*, 177.
- [89] J. M. Cowley, M. A. Osman, P. Humble, *Ultramicroscopy* **1984**, *15*, 311.
- [90] P. Humble, *Proc. R. Soc. Lond.* **1982**, *A381*, 65.
- [91] S. Iijima, *Nature* **1991**, *354*, 56.
- [92] M. Liu, J. M. Cowley, *Ultramicroscopy* **1994**, *53*, 333.
- [93] H. J. Ou, J. M. Cowley, *Phys. Status Solidi (a)* **1988**, *107*, 719.

- [94] R. H. Milne in *Reflection High Energy Electron Diffraction and Reflection Electron Imaging of Surfaces* (Eds.: P. K. Larsen, P. J. Dobson), Plenum Press, New York and London **1988**, p. 317.
- [95] J. M. Cowley, *Ultramicroscopy* **1989**, 27, 319.
- [96] J. Liu, J. M. Cowley, *Ultramicroscopy* **1991**, 37, 50.

## 2.3 Scanning Transmission Electron Microscopy: Z Contrast

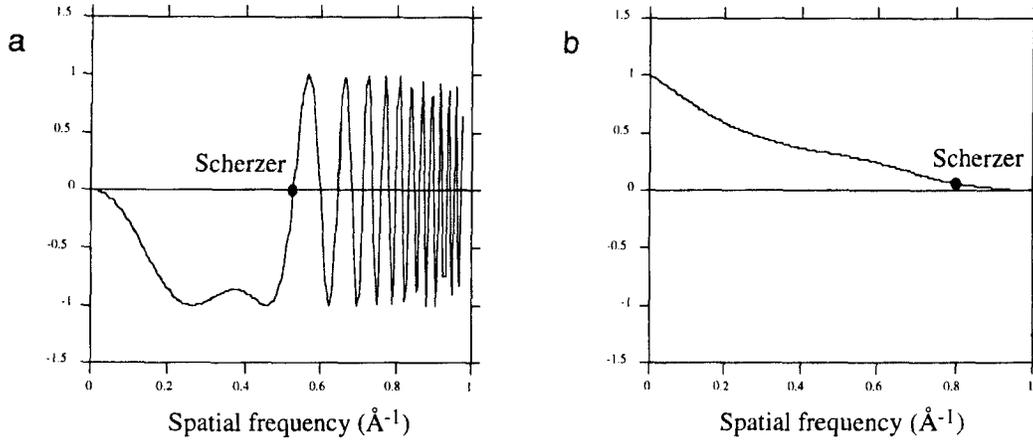
### 2.3.1 Introduction

Z-contrast scanning transmission electron microscopy (STEM) realizes the alluring goal of direct structure retrieval through its incoherent, compositionally sensitive image, at a present resolution of 1.26 Å. Furthermore, it facilitates atomic resolution microanalysis, as the STEM probe can be positioned over known atomic columns for the collection of X-ray or electron energy loss spectroscopy (EELS) data. This powerful combination of column-by-column imaging and spectroscopy opens up many important but previously inaccessible materials problems for fundamental study. Such investigations require incoherent imaging conditions to be established for both the imaging and the analytical signals. In this section, we examine the theoretical basis for incoherent imaging with the three distinct classes of electrons, coherent scattering, thermal diffuse scattering, and inelastically scattered electrons. The power of this approach to structure determination is illustrated through a variety of applications; determining interface and grain boundary structures in semiconductors and ceramics, imaging the morphology of

nanocrystalline materials, and elucidating the atomic scale structure of highly dispersed supported catalysts.

An incoherent image is defined as the convolution of two positive quantities, one representing the specimen, which we refer to as the object function  $O(\mathbf{R})$ , and one representing the resolution of the imaging system, which in our case is the probe intensity profile  $P^2(\mathbf{R})$ . The transfer function is then the Fourier transform of the probe intensity profile, which is also generally positive. In Fig. 1, the optimum contrast transfer functions for coherent and incoherent imaging conditions are compared, assuming a 300 kV microscope with an objective lens  $C_s$  of 1 mm. The difference between the transfer function of the incoherent mode and the oscillating transfer function of the coherent mode is evident. Because  $P^2(\mathbf{R})$  is always sharper than  $P(\mathbf{R})$ , the second important property of incoherent imaging is its substantially enhanced resolution; the Scherzer resolution condition for incoherent imaging is  $0.43C_s^{1/4}\lambda^{3/4}$ , significantly higher than that for coherent imaging,  $0.66C_s^{1/4}\lambda^{3/4}$  [1].

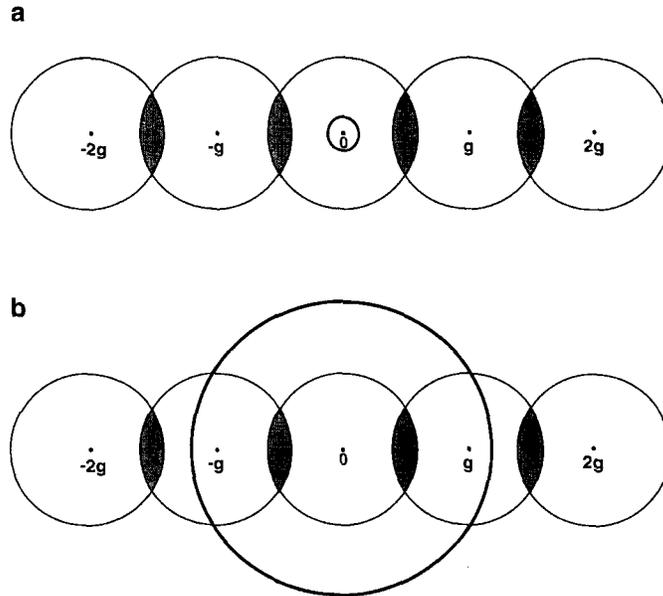
These resolution criteria are deduced assuming the objective aperture can be set to the optimum size appropriate to the imaging conditions used. For a



**Figure 1.** Contrast transfer functions for a 300 kV microscope with a 1 mm  $C_s$  objective lens for (a) coherent and (b) incoherent imaging conditions. Curves assume Scherzer optimum conditions: (a) defocus  $-505 \text{ \AA}$ , (b) defocus  $-438 \text{ \AA}$ , aperture cut-off  $0.935 \text{ \AA}^{-1}$ .

specific aperture size, incoherent dark field imaging gives *double* the resolution of bright field coherent imaging, as shown in Fig. 2. Image contrast results from the regions of overlapping disks that reach the respective detector. In a bright field image,

spacings must be below the objective aperture radius for the interference region to fall on the small axial detector, exactly as in conventional transmission electron microscopy (TEM). With an annular detector, interference is detected from



**Figure 2.** Regions of overlapping convergent beam disks for a diffraction vector greater than the objective aperture radius. (a) An axial bright field detector shows no contrast, while (b) regions of overlapping disks fall on the annular detector. The incoherent dark field image has double the resolution of the bright field image.

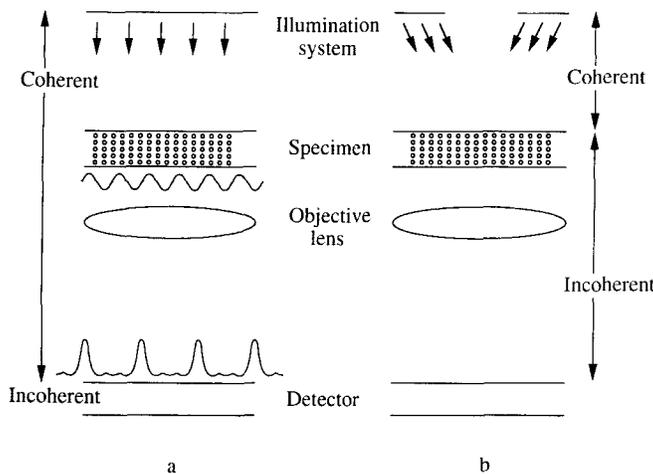
spacings up to the aperture diameter. The dark field image therefore shows double the resolution of the bright field image.

A less obvious but equally important advantage of incoherent imaging is that its intuitive relation to the object is maintained under dynamic diffraction conditions. Modern computer-based approaches to phase contrast imaging are successfully extending resolution beyond the Scherzer condition, but direct inversion will be limited to thin objects in which dynamical effects are small. In general, coherent imaging methods must rely on the use of preconceived structure models and 'goodness of fit' measures. Incoherent images are highly intuitive; an atomic column does not reverse contrast even if its image overlaps that of a neighboring column. Column intensities do not increase linearly with column length, due to absorption effects or some partial coherence, but column *locations* can still be determined to high accuracy by direct inspection of the image.

The difference between the two imaging modes is perhaps more apparent in Fig. 3,

comparing coherent and incoherent imaging optics in a conventional TEM geometry. Phase contrast imaging (Fig. 3a) involves the coherent transfer of a plane wave fast electron through the specimen and objective lens system to the detector, at which point the intensity is taken. To determine the specimen potential requires first the reconstruction of the exit face wave function, and, second, a solution of the inverse dynamic diffraction problem. Usually, therefore, interpretation must be made through simulations of model structures, with the unavoidable risk that the correct model may not be considered.

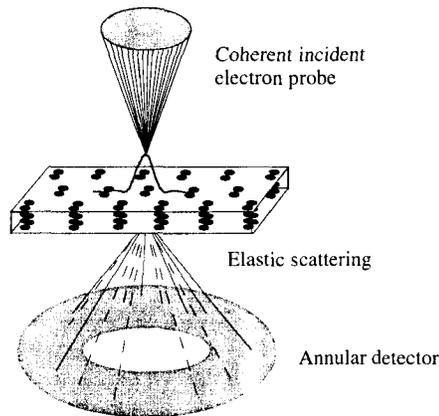
These problems are overcome if the coherence of the imaging process is broken, as in Fig. 3b. This requires illumination over a large range of scattering angles, which, through the increased contribution of nuclear Rutherford scattering, is also the reason for the strong atomic number (*Z*) sensitivity of the images. Then the specimen is made effectively self-luminous, and the same imaging optics now provide an incoherent image with



**Figure 3.** Schematic comparing (a) coherent and (b) incoherent imaging optics for a transmission electron microscope.

the transfer characteristics of Fig. 1b. Contrast reversals can theoretically occur in the presence of aberrations [2], but are rare, and occur only at low amplitude. Therefore, strong image features that do not reverse contrast on changing objective lens defocus can be directly interpreted as the positions of atomic columns. There are no proximity effects at interfaces, no contrast oscillations with increasing thickness, and beam broadening is reduced. There is no phase problem associated with incoherent imaging, and no problem with dynamical diffraction. The reliance on preconceived structure models is removed and direct structure inversion becomes a reality; quantitative methods such as maximum entropy [3] allow atomic column coordinates to be retrieved with an accuracy of a few tenths of an angstrom.

The need for a wide range of scattering angles means that incoherent imaging is more efficiently achieved with the reciprocally equivalent STEM geometry shown in Fig. 4, where the annular detector col-



**Figure 4.** Schematic of the STEM optics for incoherent imaging.

lects elastically and thermally scattered electrons. Most inelastically scattered electrons pass through the central hole, and can therefore be detected *simultaneously*, which was one of the important motivations for the original development of the scanning transmission electron microscope. Indeed, the first Z-contrast images were obtained by taking the ratio of elastic to inelastic signals [4]. Such methods proved useful in biological studies, but were unsuitable in materials science due to the need for zone axis crystals to achieve atomic resolution, and the unavoidable presence of strong dynamical scattering in most samples of practical interest. How this is overcome represents a fascinating study in the effects of quantum mechanical superposition, as discussed in Sec. 2.3.5 of this Chapter. It is the preferential selection of tightly bound s-type Bloch states that leads to the strong columnar channeling and makes incoherent imaging possible in thick crystals. Thin crystal, weak scattering results can be simply extended into the thick crystal regime. First, therefore, we examine the theory of incoherent imaging in thin crystals using electrons scattered by coherent, thermal, and inelastic mechanisms.

### 2.3.2 Incoherent Imaging with Elastically Scattered Electrons

For clarity, we will retain the general form  $P(\mathbf{R} - \mathbf{R}_0)$  of the STEM probe amplitude profile, where  $\mathbf{R}$  is a transverse positional coordinate on the specimen entrance surface, and  $\mathbf{R}_0$  is the probe position. The expansion of the probe as a phase

aberrated spherical wave is well known,

$$\begin{aligned}
 P(\mathbf{R} - \mathbf{R}_0) &= \int_{\text{objective aperture}} \exp\{i[\mathbf{K} \cdot (\mathbf{R} - \mathbf{R}_0) \\
 &\quad + \gamma(\mathbf{K})]\} d\mathbf{K} \quad (1)
 \end{aligned}$$

where  $\mathbf{K}$  is the transverse component of the incident wavevector  $\chi$ , and

$$\gamma = \frac{\pi}{\lambda} (\Delta f \theta^2 + \frac{1}{2} C_s \theta^4) \quad (2)$$

is the transfer function phase factor for a defocus  $\Delta f$  and spherical aberration coefficient  $C_s$ . The scattering angle  $\theta = K/\chi$  and  $\lambda = 2\pi/\chi$  is the electron wavelength. For a very thin specimen, effects of probe dispersion and absorption may be ignored, and the scattered amplitude  $\psi_s$  obtained immediately from the first Born approximation,

$$\begin{aligned}
 \psi_s(\mathbf{K}_f) &= \frac{m}{2\pi\hbar^2} \int \exp[i\mathbf{K}_f \cdot \mathbf{R}] \\
 &\quad \times V(\mathbf{R}) P(\mathbf{R} - \mathbf{R}_0) d\mathbf{R} \quad (3)
 \end{aligned}$$

where  $V(\mathbf{R})$  is the projected potential. Integrating the scattered intensity  $|\psi_s|^2$  over all final states  $\mathbf{K}_f$ , using the identity

$$\begin{aligned}
 &\int \exp[-i\mathbf{K}_f \cdot (\mathbf{R} - \mathbf{R}')] d\mathbf{K}_f \\
 &= (2\pi)^2 \delta(\mathbf{R} - \mathbf{R}') \quad (4)
 \end{aligned}$$

gives the image intensity

$$I(\mathbf{R}_0) = \int O(\mathbf{R}) P^2(\mathbf{R} - \mathbf{R}_0) d\mathbf{R} \quad (5)$$

which represents a convolution of the probe intensity profile  $P^2(\mathbf{R})$  with an object function  $O(\mathbf{R})$  given by

$$O(\mathbf{R}) = \sigma^2 V^2(\mathbf{R}) \quad (6)$$

where  $\sigma = \chi/2E$  is the interaction constant. Therefore, provided *all* scattered electrons could be collected, we see immediately that incoherent imaging would be obtained with a resolution controlled by the incident probe intensity profile. Note that the Fourier components of the object function are given by

$$\tilde{O}(\mathbf{K}) = \int O(\mathbf{R}) \exp(-i\mathbf{K} \cdot \mathbf{R}) d\mathbf{R} \quad (7)$$

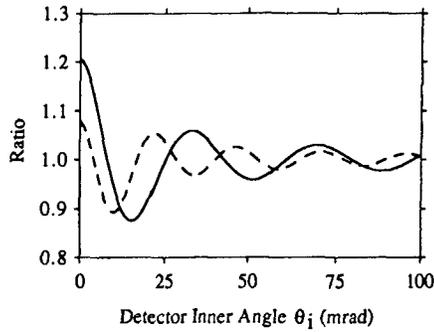
which can be written as a self-convolution of elastic scattering form factors  $f(\mathbf{K})$ ,

$$\tilde{O}(\mathbf{K}) = \frac{1}{\chi^2} \int f(\mathbf{K} + \mathbf{K}') f(\mathbf{K}') d\mathbf{K}' \quad (8)$$

A similar reciprocal space integral occurs in the object functions for diffuse scattering and inelastic scattering presented later.

Unfortunately, it is not practically possible to detect all scattered electrons, which has led in the past to a widely held but fallacious view that incoherent imaging could not be achieved at high resolution. Over the angular range of the objective aperture the scattered beam cannot be physically distinguished from the unscattered beam. For high-resolution imaging, the objective aperture is necessarily large, so that a significant fraction of the total scattering will inevitably occur within the same angular range. If a hole is cut in the detector to exclude the unscattered beam, some of the scattered beam will also be lost, and the requirements for incoherent imaging cannot be satisfied [5, 6]; this is referred to as the hole-in-the-detector problem.

However, at the cost of reduced signal intensity, this problem may be circumvented to any desired degree of accuracy. Suppose it was possible to collect all the scattered intensity up to some cut-off



**Figure 5.** Ratio of detected intensity to the incoherent result for two silicon atoms spaced 1.0 Å (solid line) and 1.5 Å (dashed line) apart for an annular detector with inner angle  $\theta_i$ .

wavevector  $K_c$ . Due to the angular dependence of atomic form factors, with increasing  $K_c$  the signal would increase and eventually saturate. If  $K_c$  is chosen to collect a large fraction  $f$  of the total scattering, sufficient to provide incoherent imaging to our desired accuracy, then collecting all the scattering above  $K_c$  will also give an incoherent image, but with the reduced signal level of  $1 - f$ . In this way, a high angle annular detector can collect a small but *representative* fraction of the total scattering and circumvent the hole-in-the-detector problem. Figure 5 shows how the intensity detected from two silicon atoms approaches the incoherent limit as the inner angle of the annular detector is increased [7]. Significant deviations occur for small inner angles, but at the second minimum in the figure the deviation from incoherent imaging is only  $\sim 5\%$ . This gives a useful criterion for the minimum inner detector angle necessary for the incoherent imaging of two objects separated by  $\Delta R$ ,

$$\theta_i = 1.22\lambda/\Delta R \quad (9)$$

The object function is no longer the full potential  $V^2(\mathbf{R})$ , but involves only the

sharp part of the potential that generates the high-angle scattering,

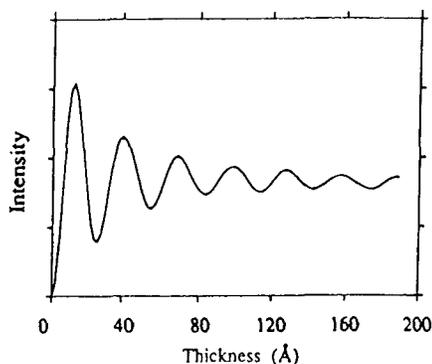
$$O(\mathbf{R}) = \sigma^2 |V(\mathbf{R}) * d(\mathbf{R})|^2 \quad (10)$$

Here,  $d(\mathbf{R})$  is the Fourier transform of the annular detector function  $D(\mathbf{K})$ , which is unity over the detector and zero elsewhere. For incoherent imaging to be a good approximation,  $d(\mathbf{R})$  must be significantly narrower than  $P(\mathbf{R})$ . It then localizes contributions to the scattered intensity in a manner similar to the action of the  $\delta$  function in Eq. (4). The difference, however, is that the detector acts as a spatial frequency filter for the potential  $V(\mathbf{R})$ , selecting only the sharp nuclear contribution. This is the origin of the strong  $Z$  sensitivity to the image. It is of course very convenient that the atomic potential does include a wide range of spatial frequencies, since this approach would fail for a specimen comprising a single sine wave potential.

Provided the criterion (9) is met, images of single atoms or small clusters will show no reversals in contrast with lens defocus, and will show enhanced resolution compared to a bright field image, characteristics that were indeed seen experimentally [8]. However, these incoherent characteristics refer only to the transverse plane. Although coherence in the transverse plane is limited to scales smaller than the width of  $d(\mathbf{R})$ , which can be conveniently arranged to be below the intercolumnar spacings we wish to image, for a column of  $n$  atoms aligned with the beam direction, the object function becomes

$$O(\mathbf{R}) = \sigma^2 |nV(\mathbf{R}) * d(\mathbf{R})|^2 \quad (11)$$

giving an intensity proportional to the square of the number of atoms in the column. This shows that each atom within



**Figure 6.** The intensity of zero-layer coherent scattering reaching a 50–150 mrad annular detector calculated kinematically for a 100 kV probe of 10.3 mrad semiangle located over a dumbbell in Si (110).

a particular column contributes coherently to the total intensity scattered by that column. Re-evaluating the experimental data of Isaacson et al. [9] indicates the expected  $n^2$  behavior of intensity on column length [7].

For crystals with thickness  $t$  comparable to  $2\lambda/\theta_i^2$ , the amplitude scattered from the top and bottom of a column will no longer be in phase, and will destructively interfere. This results in the oscillatory thickness dependence, as shown in Fig. 6. Note that in the absence of thermal vibrations the intensity never rises above that of a thin crystal. The reason is clear from a reciprocal space argument; at this thickness it is no longer possible to ignore the curvature of the Ewald sphere, and with increasing column length the scattered intensity becomes more concentrated into the reciprocal lattice spot, thus missing the Ewald sphere. In practice, however, this behavior is masked by the increasing importance of thermal vibrations, which also mask any contribution from HOLZ lines.

### 2.3.3 Incoherent Imaging with Thermally Scattered Electrons

The need for high scattering angles naturally suggests that thermal diffuse scattering might represent an important, perhaps dominant contribution to the detected intensity. Furthermore, we might suppose that the thermal vibrations themselves could be effective in breaking the coherence of the imaging process, which was the original motivation for the use of a high angle annular detector [10]. We find that thermal vibrations make little difference in the transverse plane; coherence here is very efficiently broken by the detector geometry alone. Thermal vibrations are, however, very effective in breaking the coherence along the columns, dramatically suppressing HOLZ lines and avoiding the oscillatory behavior of coherent zero layer diffraction seen in Fig. 5.

Thermal vibrations are normally treated in the framework of an Einstein model of independently vibrating atoms; in this case the diffuse scattering is by definition generated incoherently. In reality, atomic vibrations are not independent, and positions of near-neighbor atoms will be correlated. In order to examine the effect of these short-range positional correlations on the image intensity we need a phonon model of lattice vibrations. First, however, we consider the Einstein model since it provides explicit atomic cross-sections for the diffuse scattering, although, for the correct high angle behavior, multiphonon scattering processes must be included. We therefore use the approach of Hall and Hirsch [11, 12], which naturally includes all multiphonon processes

by calculating the total time-averaged intensity distribution, and subtracting the Bragg reflections. Integrating this intensity over the Ewald sphere gives directly the Fourier components of the object function,

$$\begin{aligned} \tilde{O}^{\text{TDS}}(\mathbf{K}) &= \frac{1}{\chi^2} \int f(|s'|)f(|s-s'|) \\ &\quad \times [\exp(-Ms^2) - \exp(-Ms'^2)] \\ &\quad \times \exp[-M(s-s')^2] d\mathbf{K}' \quad (12) \end{aligned}$$

where  $s = K/4\pi$  as used in the conventional tabulations [13], and  $M$  is the usual Debye–Waller factor.

Now consider the important case of the high-angle annular detector. If the potential is much sharper than the probe intensity profile, we can remove the probe from the integral in Eq. (5), giving

$$\begin{aligned} I(\mathbf{R}_0) &\approx P^2(\mathbf{R} - \mathbf{R}_0) \int O^{\text{TDS}}(\mathbf{R}) d\mathbf{R} \\ &= P^2(\mathbf{R} - \mathbf{R}_0) \tilde{O}^{\text{TDS}}(0) \quad (13) \end{aligned}$$

where  $\tilde{O}^{\text{TDS}}(0)$  is simply the atomic diffuse scattering distribution integrated over the high-angle detector,

$$\begin{aligned} \tilde{O}^{\text{TDS}}(0) &= \frac{1}{\chi^2} \int_{\text{detector}} f^2(|s'|) \\ &\quad \times [1 - \exp(-2Ms'^2)] d\mathbf{K}' \\ &= \sigma_{\text{at}} \quad (14) \end{aligned}$$

the atomic cross-section for diffuse scattering into the annular detector. For sufficiently high angles the term in square brackets tends to unity, and the cross-section becomes the total elastic scattering intensity integrated over the detector. For a column of  $n$  atoms aligned with the beam direction, these cross-sections are

summed, giving an intensity which scales as  $n$  in the limit of a thin crystal where probe dispersion and absorption are negligible,

$$O(\mathbf{R}) = n\sigma_{\text{at}}\delta(\mathbf{R}) \quad (15)$$

We now consider the effects of correlated vibrations on the diffuse scattering. In the context of a high-angle annular detector, transverse incoherence is ensured by the detector geometry, so again we need only consider correlations for a column of atoms aligned with the beam direction. The important physical concept to emerge is the existence of a finite coherence length  $l_c$  along a column. Crystals of thickness  $t \ll l_c$  will show the  $n^2$  dependence characteristic of coherent scattering, as seen in Eq. (11). In crystals with  $t \gg l_c$ , the coherent component will be small, and columnar intensities will show the linear behavior predicted by the Einstein model of diffuse scattering. Note, however, that  $l_c$  cannot be obtained from the Einstein model of independently vibrating atoms; here the coherence length is by definition zero for the diffuse component, and  $t$  for the coherent component. In reality, the degree of correlation is highest for near neighbors, and reduces smoothly with increasing separation.

To examine the role of residual correlations it is necessary to consider phonons explicitly. Consider a column of atoms with instantaneous deviations  $u_i$  from their lattice positions  $\mathbf{r}_i = (\mathbf{R}_i, z_i)$  scattering to a high-angle detector, so that all scattering occurs in close proximity to the atomic sites, and transverse incoherence is ensured. An object function can then be written including all zero layer, HOLZ, and diffuse scattering

as [14]

$$I^H(\mathbf{R}_0) = \frac{1}{\chi^2} \sum_{m,n} \int_{\text{detector}} f^2(s) \times \exp\{i\mathbf{k} \cdot (\mathbf{r}_m - \mathbf{r}_n)\} \times \langle \exp\{i\mathbf{k} \cdot (\mathbf{u}_m - \mathbf{u}_n)\} \rangle d\mathbf{K} \quad (16)$$

where  $\mathbf{k} \approx (\mathbf{K}, K^2/2\chi)$  is the scattering vector and the angled brackets denote the time average. A treatment due to Warren [15] incorporates phonons by assuming a Debye dispersion relation, equipartition, and replacing the Brillouin zone boundary at  $q_B$  by a sphere of equal volume, whereupon the time average in Eq. (16) reduces to an analytical form,

$$W_{mn} = \langle \exp\{i\mathbf{k} \cdot (\mathbf{u}_m - \mathbf{u}_n)\} \rangle = \exp\left[2M \left(\frac{\text{Si}(2\pi q_B r_{mn})}{2\pi q_B r_{mn}} - 1\right)\right] \quad (17)$$

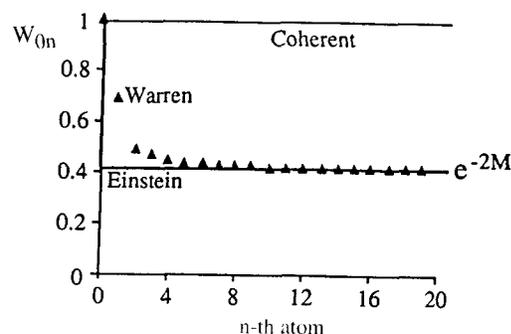
This factor expresses the degree of correlation between the  $m$ th and  $n$ th atoms in a column, where  $r_{mn} = |\mathbf{r}_m - \mathbf{r}_n|$  and  $\text{Si}(x)$  is the sine integral function

$$\text{Si}(x) = \int_0^x \frac{\sin u}{u} du \quad (18)$$

The partial coherence between near neighbors in the column is shown in Fig. 7, where  $W_{0n}$  is seen to fall smoothly from 1 at  $m = 0$  to  $e^{-2M}$  for large  $m$ . If we use an Einstein dispersion relation, with constant frequency for all modes and polarizations, we recover the Einstein model, with a correlation factor given by

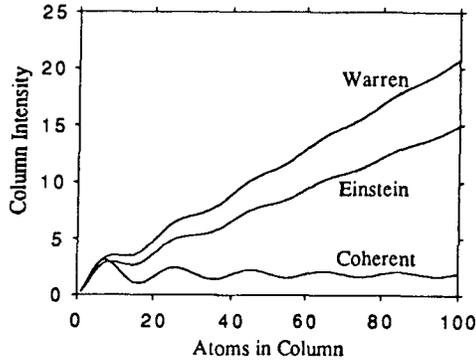
$$W_{0n} = \exp[2M(\delta_{0n} - 1)] \quad (19)$$

This is again 1 at  $m = 0$ , but now takes the value  $e^{-2M}$  for all other atoms, that is, near-neighbor correlations are specifically excluded.



**Figure 7.** Correlation factor between atoms separated by  $n$  spacings along a column calculated in the Warren approximation (triangles), the Einstein limit and the coherent limit for  $M = 0.45$  and  $\theta_i = 75$  mrad. The arrowed vertical line defines a packet within which partial coherence exists.

Physically, we can picture the columnar coherence volume as defining a packet surrounding each atom, within which partial coherence exists, but outside of which the atoms appear uncorrelated. Coherent interference effects are observed in the form of thickness oscillations for columns shorter than the length of the packet, while for longer columns the thickness dependence becomes linear, but with a slope that can be above or below the Einstein value depending on whether the interference effects within each packet are constructive or destructive in nature. Figure 8 shows the thickness dependence for a column of silicon atoms, showing the initial  $t^2$  dependence, some destructive interference, and the eventual linear behavior. Shown also are the coherent and diffuse components calculated in the Einstein model. In this case it can be seen that the coherence effects within a packet are constructive. Generally, however, the changes in columnar cross-section due to correlation effects are relatively small for high-angle detectors, but can be significant with low-angle detectors. This effect could be utilized to



**Figure 8.** Calculated thickness dependence of elastic scattering from a single column of silicon atoms calculated with a phonon model for diffuse scattering due to Warren [15], compared to the coherent and diffuse components calculated in the Einstein model (inner detector angle 50 mrad, atom separation 1.5 Å, Debye–Waller parameter  $M = 0.45$ ).

boost signal levels from beam-sensitive specimens or small clusters by tuning the packet length to match the specimen thickness.

### 2.3.4 Incoherent Imaging using Inelastically Scattered Electrons

The optical equation for incoherent imaging with inelastically scattered electrons was first derived by Rose [16], and expressed in a form explicit to STEM by Ritchie and Howie [17]. The derivation is somewhat more elaborate than the simple case of elastic scattering given above, but follows similar lines. We give here Ritchie and Howie's equation (11) for the image intensity when all scattered electrons are collected, which can again be written as a convolution of the probe intensity profile with an inelastic object

function  $O'(\mathbf{R})$ ,

$$I'(\mathbf{R}_0) = \int P^2(\mathbf{R} - \mathbf{R}_0) O'(\mathbf{R}) d\mathbf{R} \quad (20)$$

where

$$O'(\mathbf{R}) = \left( \frac{e^2}{\pi \hbar v} \right)^2 \times \left| \int \frac{\rho_{n0}(\mathbf{q}) \exp(-i\mathbf{K} \cdot \mathbf{R}) d\mathbf{K}}{q^2} \right|^2 \quad (21)$$

Here we write the momentum transfer  $\mathbf{q} = (\mathbf{K}, q_z)$  to recognize the minimum momentum transfer at zero scattering angle, and introduce the matrix elements  $\rho_{n0}$  for the transition from an initial state  $|0\rangle$  to a final state  $|n\rangle$ . Ritchie and Howie showed further that the matrix elements in Eq. (21) may be calculated in the semiclassical approximation, that is, assuming no deflection of the fast electron trajectory, which was convenient for their study of the excitation of surface plasmons (see also the discussion by Batson [18]).

Again, the object function can be expressed as a form factor in reciprocal space [19],

$$\tilde{O}'(\mathbf{K}) = \left( \frac{2e^2}{\hbar v} \right)^2 \int \frac{\rho_{n0}(\mathbf{q}') \rho_{n0}^*(\mathbf{q}' + \mathbf{K})}{q'^2 |\mathbf{q}' + \mathbf{K}|^2} d\mathbf{K}' \quad (22)$$

where

$$O'(\mathbf{R}) = \frac{1}{(2\pi)^2} \int \tilde{O}'(\mathbf{K}) \exp(i\mathbf{K} \cdot \mathbf{R}) d\mathbf{K} \quad (23)$$

For atomic resolution analysis we are primarily concerned with inner shell excitations, and here a hydrogenic model developed by Maslen and Rossouw is particularly useful [20–24]. Now, an inelastic excitation with scattering vector  $\mathbf{q}$  for the fast electron is modeled by excitation of the K-shell electron to the

continuum of the hydrogenic atom, representing ejection of a secondary electron with wavevector  $\kappa$ . As the secondary is not observed experimentally, the process is integrated over all secondary electron emission directions, and the matrix product  $\rho_{n0}(\mathbf{q}')\rho_{n0}^*(\mathbf{q}' + \mathbf{K})$  is replaced with

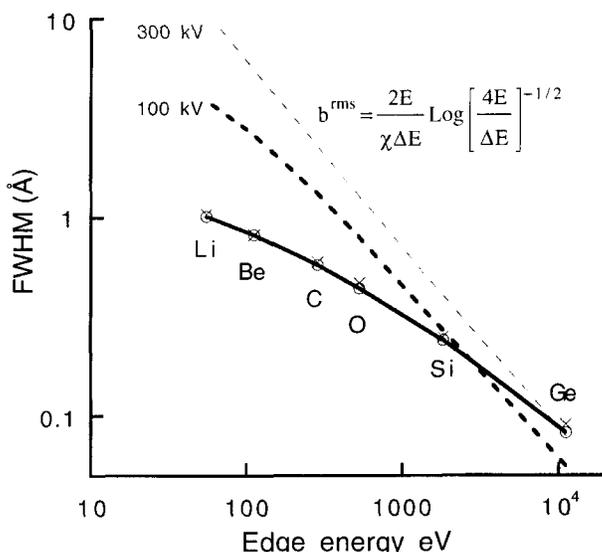
$$\int \rho_{n0}(\mathbf{q}', \kappa)\rho_{n0}^*(\mathbf{q}' + \mathbf{K}, \kappa) d\kappa \quad (24)$$

Analytical expressions have been given for these integrals [25,26], and their general properties have been considered by Kohl and Rose [27].

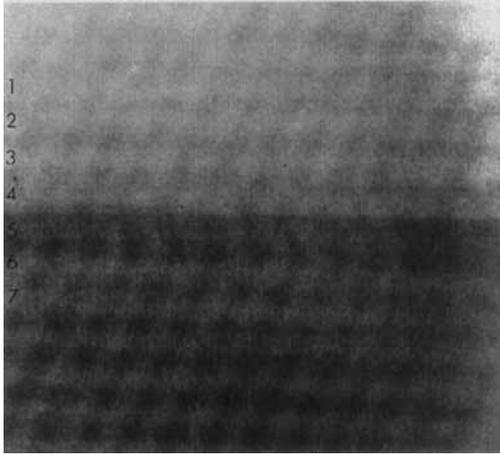
Figure 9 plots the full-width half-maxima of object functions calculated for a number of elements, assuming that all the scattered electrons are collected. Somewhat surprisingly it would appear that atomic resolution is possible even for the lightest elements. This is a significant difference from impact parameter estimates based on the angular widths of scattering distributions [28–30], and reflects the strong  $Z$  dependence of the hydrogenic

model. Furthermore, the quantum mechanical expression is insensitive to the beam energy, predicting object function widths that are practically identical for 100 and 300 kV electrons.

Although analytical expressions are available for L- and M-shell cross-sections, none exist for the Fourier components  $\tilde{O}'(\mathbf{K})$ . However, if we are only concerned with excitations near threshold it is perfectly reasonable to model the process as bound-to-bound transitions within the hydrogenic atom, as recently suggested by Holbrook and Bird [31]. They find widths for L-shell excitations that are very comparable to those for K-shell excitation at the same energy, findings that are consistent with recent experimental results. Batson [32] has found changes in the silicon L fine structure on moving the probe one or two atomic spacings, while Browning et al. [33] obtained atomic resolution at the cobalt L edge, using a  $\text{CoSi}_2$ -Si (111) epitaxial interface as an edge resolution



**Figure 9.** Full-width half-maxima of object functions for K-shell excitation of various elements calculated with a hydrogenic model for a 50 eV ejected secondary assuming all scattered electrons are collected. Circles are for 100 kV incident electrons, crosses for 300 kV. Dashed lines show impact parameters estimated from the angular widths of scattering distributions [30].

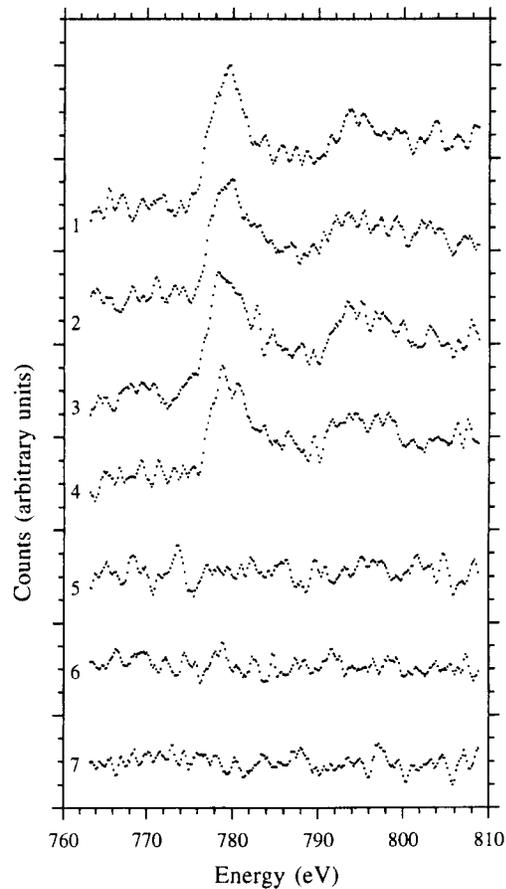


**Figure 10.** The Z-contrast image of a  $\text{CoSi}_2$ -Si (111) epitaxial interface used as an edge resolution test specimens for EELS. Cobalt atoms image bright; planes numbered correspond to the spectra in Fig. 11.

test specimen. Figure 10 is an image of this interface taken with a 100 kV VG Microscopes HB501UX, which reveals a structure not previously considered for this interface [34]. The last plane of the silicon is seen to be in a rotated configuration, which maintains the eightfold coordination of the cobalt. EELS spectra recorded from each plane in turn clearly demonstrated the atomic resolution capability. In Fig. 11 the cobalt L edge is seen to drop from 86 to 7% of its bulk value in moving a single plane across the interface, a distance of 2.7 Å.

### 2.3.5 Probe Channeling

In practice, specimens for which the Born approximation is valid are rarely, if ever, encountered. Even a single heavy atom is sufficient to invalidate the Born approximation, and for a typical microscope



**Figure 11.** EELS spectra obtained phase by plane across the interface of Fig. 10 showing atomic resolution at the cobalt L edge.

specimen one might expect dynamical scattering to complicate the situation and necessitate a full image simulation, a return to a reliance on model structures. However, experimentally it is found that the characteristics of incoherent imaging are preserved to large specimen thicknesses [35–38]. In fact, with a 300 kV scanning transmission electron microscope, structure imaging remains to quite remarkable thicknesses.

The reason for this lies in the channeling behavior of a coherent STEM probe, in

particular, the effect of the superposition required to describe the convergent probe. A plane wave entering a crystal forms the set of Bloch states,

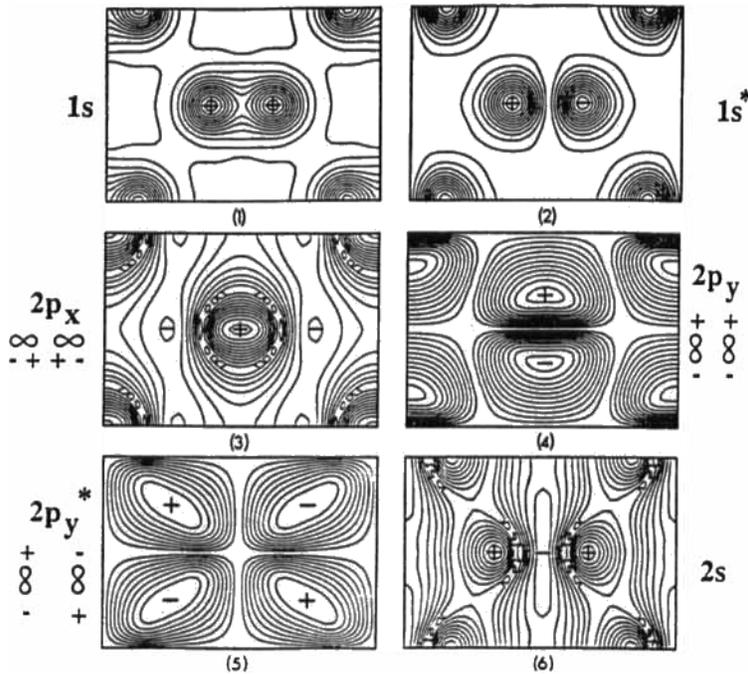
$$\begin{aligned} \psi(\mathbf{R}, z) = & \sum_j \epsilon^j(\mathbf{K}) \tau^j(\mathbf{K}, \mathbf{R}) \\ & \times \exp(-is^j(\mathbf{K})z/2\chi) \\ & \times \exp(-\mu^j(\mathbf{K})z) \end{aligned} \quad (25)$$

where  $\tau^j(\mathbf{K}, \mathbf{R}) = b^j(\mathbf{K}, \mathbf{R}) \exp(i\mathbf{K} \cdot \mathbf{R})$  are two-dimensional Bloch states of transverse energy  $s^j(\mathbf{K})$ , absorption  $\mu^j(\mathbf{K})$  and excitation coefficients  $\epsilon^j(\mathbf{K})$ . The states of greatest transverse energy take on the form of molecular orbitals about the atomic strings, as seen in Fig. 12. They propagate with slightly different wavevectors through the crystal thickness, and it is their interference that leads to the familiar

dynamical diffraction effects. Now to form a probe we make a coherent superposition of phase-aberrated plane waves over the objective aperture (see Eq. (1)), so that the wave function inside the crystal becomes

$$\begin{aligned} \psi(\mathbf{R} - \mathbf{R}_0, z) &= \sum_j \int_{\text{objective aperture}} \epsilon^j(\mathbf{K}) b^j(\mathbf{K}, \mathbf{R}) \\ &\times \exp(-is^j(\mathbf{K})z/2\chi) \exp(-\mu^j(\mathbf{K})z) \\ &\times \exp\{i[\mathbf{K} \cdot (\mathbf{R} - \mathbf{R}_0) + \gamma(\mathbf{K})]\} d\mathbf{K} \end{aligned} \quad (26)$$

The effect of this angular integration depends very strongly on the nature of the particular Bloch state. Tightly bound 1s states overlap little between neighboring



**Figure 12.** Bloch states for a 100 kV electron in silicon  $\langle 110 \rangle$ , with their molecular orbital assignments.

strings and are therefore nondispersive. They add in phase during the angular integration. In contrast, states such as 2s and 2p are significantly more extended and do overlap with states on neighboring columns. They then become dispersive and interfere destructively when forming the superposition. Calculations confirm that the intensity at the atomic columns is dominated by the tightly bound s-type states [35–37]. This explains how an incident probe propagating through the crystal becomes compacted around the columns themselves, taking on the character of the s states [39, 40]. It is in this way that we obtain a rather simple columnar channeling behavior.

For signals that originate from close to the atomic sites, such as the high angle elastic and diffuse scattering, a good approximation to the full wavefunction of Eq. (26) is to include only s states in the Bloch wave sum. Since they are nondispersive, they may be taken outside the angular integration, giving

$$\begin{aligned} \psi(\mathbf{R} - \mathbf{R}_0, z) &\approx b^{1s}(\mathbf{R}, 0) \exp(-is^{1s}(0)z/2\chi) \\ &\times \exp(-\mu^{1s}(0)z) \\ &\times \int_{\text{objective aperture}} \frac{\epsilon^{1s}(\mathbf{K})}{\epsilon^{1s}(0)} \\ &\times \exp\{i[\mathbf{K} \cdot (\mathbf{R} - \mathbf{R}_0) + \gamma(\mathbf{K})]\} d\mathbf{K} \end{aligned} \quad (27)$$

We once more recover the optical equation for incoherent imaging,

$$I(\mathbf{R}, t) = O(\mathbf{R}, t) * P_{\text{eff}}^2(\mathbf{R}) \quad (28)$$

where  $O(\mathbf{R}, t)$  is now the thickness-

dependent object function,

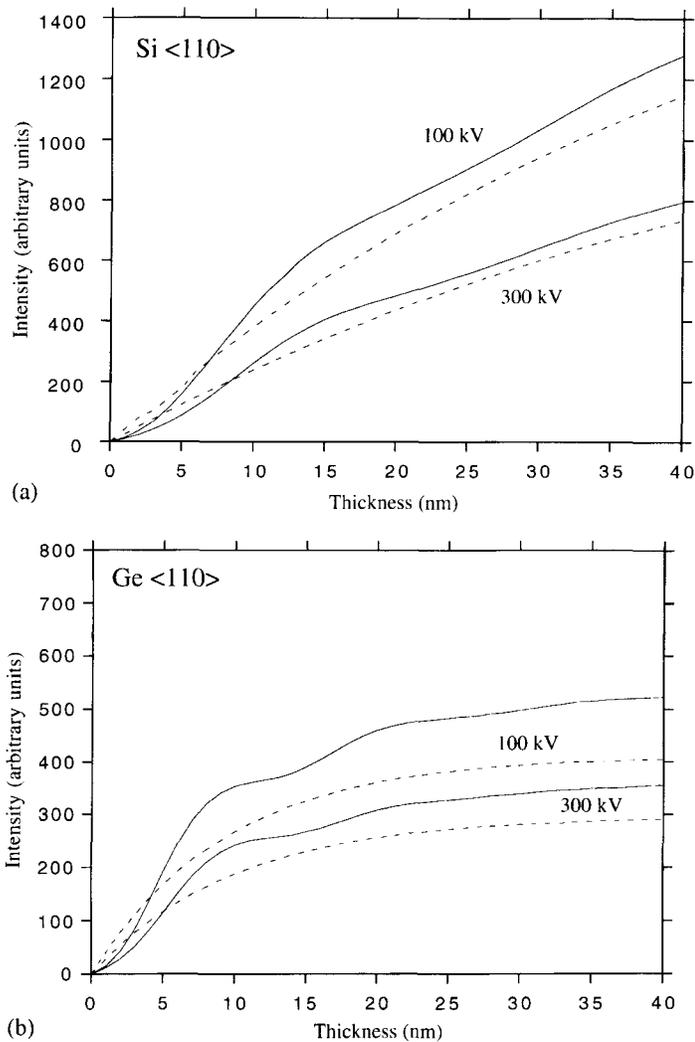
$$\begin{aligned} O(\mathbf{R}, t) &= \epsilon^{1s^2}(0) b^{1s^2}(\mathbf{R}, 0) \\ &\times \left( \frac{1 - \exp(-2\mu^{1s}(0)t)}{2\mu^{1s}(0)} \right) O^B(\mathbf{R}) \end{aligned} \quad (29)$$

in which the appropriate Born approximation object function  $O^B(\mathbf{R}) = O(\mathbf{R})$  or  $O^{\text{TDS}}(\mathbf{R})$  is scaled by the columnar channeling effect, and

$$\begin{aligned} P_{\text{eff}}^2(\mathbf{R}) &= \left| \frac{1}{\epsilon^{1s}(0)} \int_{\text{objective aperture}} \epsilon^{1s}(\mathbf{K}) \right. \\ &\times \exp\{i[\mathbf{K} \cdot (\mathbf{R} - \mathbf{R}_0) \\ &\left. + \gamma(\mathbf{K})]\} d\mathbf{K} \right|^2 \end{aligned} \quad (30)$$

is an effective probe intensity profile which includes the angular fall-off in s-state excitation. This fall-off is quite small over the range of a typical objective aperture, and only broadens the probe by around 10%. The same situation will hold for highly localized inelastic scattering, and may also hold approximately for less local object functions, although the degree of s-state dominance away from the atom sites has yet to be investigated.

This formulation models the imaging process as weak scattering out of the s-state wavefield. Note that since the s states do not broaden with depth into the crystal, the assumption of a nondispersive probe used in the Born approximation derivations is not violated. The thickness dependence of the object function is shown in Fig. 13 for silicon and germanium in the  $\langle 110 \rangle$  projection. As the s states are peaked at the atomic sites, they are absorbed quite rapidly, especially with high-Z columns as



**Figure 13.** Thickness dependence of the object function in (a) silicon  $\langle 110 \rangle$  and (b) germanium  $\langle 110 \rangle$  calculated using s states alone (dashed lines) compared to the full dynamical calculation (solid lines). Calculations assume an Einstein absorption model.

in the case of germanium. This leads to reduced contrast between heavy and light columns with increasing thickness. Also shown in Fig. 13 is the object function calculated with all Bloch states included, in which residual dynamical effects are seen at small thicknesses. A more accurate model for the object function [36] is to include the interference of the s states with the incident probe, which gives a better fit to the full dynamical curve.

This channeling behavior explains how incoherent imaging characteristics are maintained in thick crystals, and how weak scattering models can be simply modified to predict elastic or inelastic image behavior, even in the presence of the dynamical diffraction. An important additional benefit of imaging with only the highly localized Bloch states is that the object function remains highly local. This means that the columnar scattering

power is independent of the type and arrangement of surrounding atomic columns. Image interpretation remains highly intuitive, in contrast to coherent imaging in which interference with nonlocal interface states may complicate interpretation.

## 2.3.6 Applications to Materials Research

### 2.3.6.1 Semiconductors

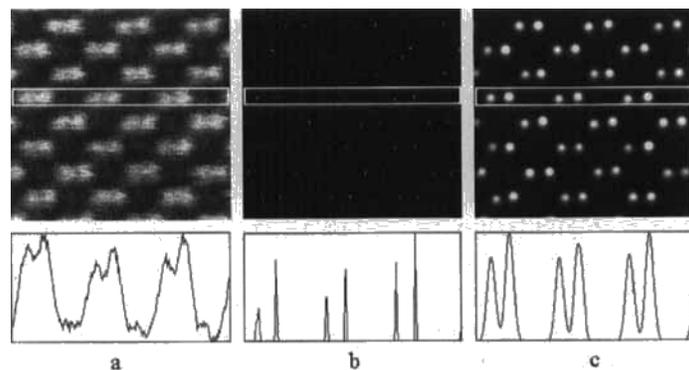
One of the alluring properties of a 300 kV scanning transmission electron microscope is a demonstrated probe size of 1.26 Å, less than the separation of the ‘dumbbells’ seen in the  $\langle 110 \rangle$  projection of diamond cubic semiconductors. In compound semiconductors, the two columns of the dumbbell are different species; the incoherence of the image, coupled with its  $Z$  contrast, therefore enables the sublattice polarity to be determined directly from the  $Z$ -contrast image [41].

Figure 14 shows GaAs imaged in the  $\langle 110 \rangle$  projection with the VG Microscopes HB603U scanning transmission electron

microscope. Although gallium and arsenic are only two atomic numbers apart in the periodic table, their scattering cross-sections differ by approximately 10%, depending upon the inner detector angle. Here, with an inner angle of 30 mrad, they are distinguishable in the raw image; the difference in scattering power is seen from the line trace below. This direct structure image is maintained up to thicknesses of the order of 1000 Å.

A structure image of this nature is a convolution of the probe intensity profile with the object; it should therefore be possible to invert the image directly by deconvolution. Unfortunately, it is well known that deconvolution procedures tend to enhance high-frequency noise, so we have employed a maximum entropy method which has an outstanding noise rejection capability [42]. The maximum entropy image is a blank page, that is, a uniform array of intensity with no information content. This is of course a poor fit to any actual image, but is used as a starting point in the maximum entropy iteration, which proceeds to find an image of constrained maximum entropy, which does provide a good fit to the image data. The probe intensity profile is

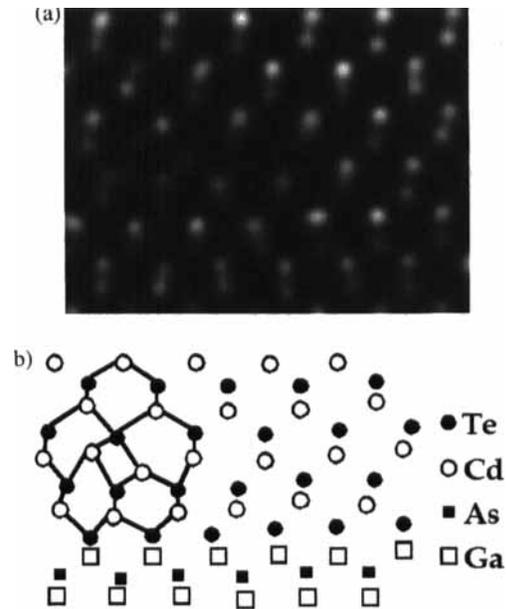
**Figure 14.** Images of GaAs  $\langle 110 \rangle$  in the HB603U 300 kV scanning transmission electron microscope. (a) Raw image, with arsenic columns showing the expected  $\sim 10\%$  higher scattering power, (b) a maximum entropy reconstruction of the object, and (c) a reconstructed image. Line traces below each image show the vertically averaged intensity within the rectangles outlined.



assumed as an input parameter, though the inversion is not critically dependent on the form assumed. From the line trace in Fig. 14a, it is a simple matter to estimate the probe width, and the maximum entropy object for a Lorentzian probe of this size is shown in Fig. 14b. The reconstructed positions of all gallium and arsenic columns are each within  $0.1 \text{ \AA}$  of their actual positions, and their relative intensity is as expected. This information is appreciated more easily by reconvoluting the object function with a small Gaussian, to give the reconstructed image of Fig. 14c. Notice how the raw image contains a secondary maximum in the channels between the dumbbells, arising from the tails of the probe intersecting the six surrounding dumbbells. These features are not part of the object, and are successfully removed from the reconstructed image.

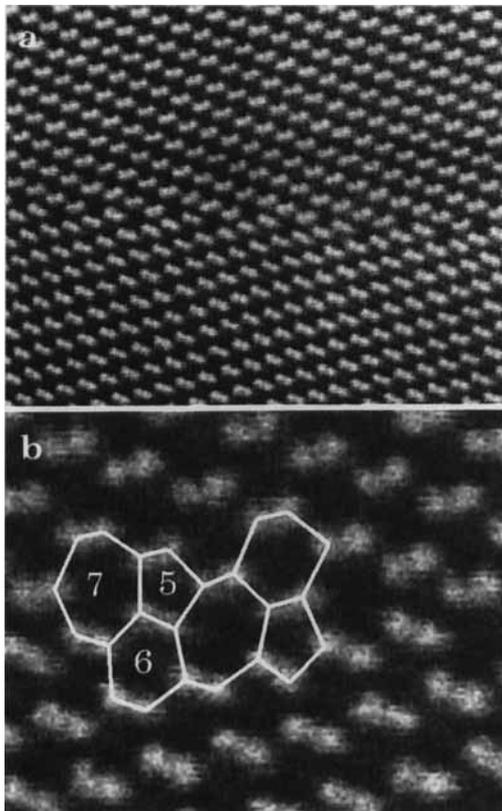
The greatest value of this direct inversion capability is that it can reveal unexpected features of the object that would not have been included in any model structure. Figure 15 shows an image of a perfect edge dislocation core at a CdTe(001)–GaAs(001) interface reconstructed in this way. It is clearly inconsistent with the five- and seven-membered ring structure of the Hornstra core [43], which is often assumed for these materials, since it shows a fourfold ring surrounded by distorted sixfold rings [41]. Such data can be used as a starting point for theoretical studies of dislocation properties.

Not all interfaces contain surprises, as demonstrated by Fig. 16, a section of a  $39^\circ$   $\langle 110 \rangle$  symmetric tilt boundary in silicon ( $\Sigma = 9, \{221\}\langle 110 \rangle$ ), viewed along the  $[110]$  direction [44]. The boundary is seen to consist of a periodic array of perfect edge dislocations with their line direction



**Figure 15.** A new core structure observed for an edge dislocation at a CdTe (001)–GaAs (001) interface. The core comprises four- and sixfold rings rather than the five- and sevenfold rings of the Hornstra structure.

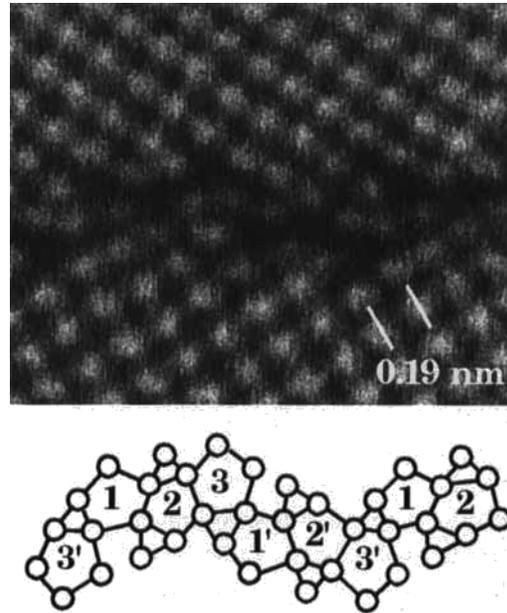
parallel to the  $\langle 110 \rangle$  tilt axis. This is seen clearly to comprise the five- and seven-membered ring structure first shown by Krivanek et al. [45]. In contrast, the symmetric  $\Sigma = 13, \{150\}\langle 001 \rangle$ , tilt boundary in silicon (Fig. 17) does show an unexpected arrangement of dislocation cores [46]. Here the boundary is a contiguous array of six dislocations, consisting of a perfect edge dislocation (labeled 1) and two perfect mixed dislocations arranged as a dipole (labeled 2 and 3), followed by the same three cores mirrored across the boundary plane (labeled 1', 2', and 3'). These cores are connected, but remain individually distinct, and tetrahedral bonding is retained through a combination of five-, six-, and seven-membered rings. This structure differs from the aperiodic



**Figure 16.**  $\Sigma = 9, \{221\}\langle 110\rangle$ , symmetric tilt boundary in silicon (viewed along the  $[110]$  direction), showing its five- and seven-membered ring structure.

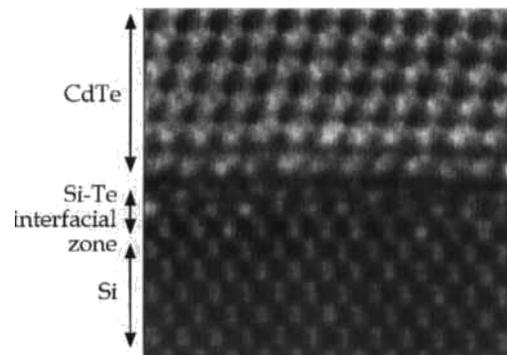
boundary containing multiple structures reported previously [47], and is not the structure predicted from the principle of continuity of boundary structure [48].

Structure images of this nature are greatly complemented by the ability to perform EELS at atomic resolution. Given our present understanding of the intrinsic width of inelastic object functions, atomic resolution is to be expected for all inner shells likely to be used in microanalysis, whether in a 100 kV or a 300 kV STEM instrument. A spectacular demonstration of the power of combined incoherent structure imaging and EELS is



**Figure 17.**  $\Sigma = 13, \{150\}\langle 001\rangle$ , symmetric tilt boundary in silicon showing an unexpected arrangement of dislocation cores comprising five-, six-, and seven-membered rings.

shown by the image of a CdTe(111)–Si(100) interface in Fig. 18. Grown by MBE, the image has been high-pass filtered to remove the high background from

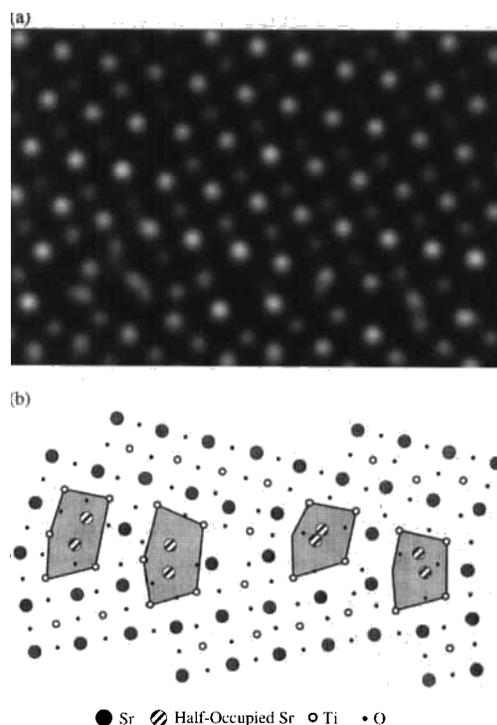


**Figure 18.** Complex atomic structure at a CdTe (111)–Si (100) interface grown by MBE. The film ends on the tellurium sublattice, but EELS revealed the presence of tellurium diffused in the first few monolayers of the silicon substrate. Tellurium-rich columns are seen bright.

the film, and the lattice polarity and interfacial structure are strikingly revealed. The film terminates at tellurium, as was expected, but the two monolayers below do not image in the positions expected for silicon. EELS reveals that these planes contain a substantial fraction of tellurium; in fact, certain columns are seen to be far brighter than can be accounted for by image noise, and must therefore be identified as tellurium-rich. That these effects are a real part of the material is confirmed by the fact that the cadmium EELS signal did indeed show atomic abruptness at the interface. Insights of this nature show directly how the growth procedure determines the interface structure. In turn, the interface structure controls the nucleation of extended defects, such as twins and dislocations, which can then propagate through the entire film and will dominate its electrical properties. With Z-contrast structure imaging and atomic resolution EELS, interface engineering becomes a real possibility.

### 2.3.6.2 Ceramics

An early application of combined Z-contrast imaging and atomic resolution EELS was to [001] tilt grain boundaries in  $\text{SrTiO}_3$  [49]. The strontium and titanium sublattices could be distinguished clearly even with the 100 kV STEM instrument, and EELS spectra could be taken from the boundary plane to compare with spectra from the bulk. In fact, the fine structure changed only a little, showing that the titanium at the boundary remained octahedrally coordinated to oxygen, though in a somewhat distorted configuration. The



**Figure 19.** (a) Reconstructed image and (b) atomic structure for a  $\Sigma = 85, \{920\}\langle 001 \rangle$  symmetric tilt boundary in  $\text{SrTiO}_3$  deduced using the 100 kV STEM. Strontium columns (seen bright) are spaced 3.9 Å apart in the bulk crystal.

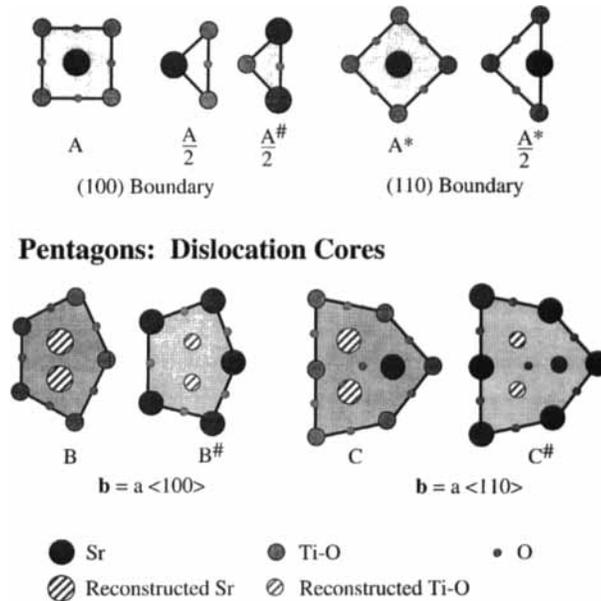
structure deduced for a  $\{920\}$  grain boundary is shown in Fig. 19, and revealed a number of interesting features. Most significantly, the columns shown hatched are located on the strontium sublattice, but clearly cannot both be present at such a small separation. The simplest model consistent with the experimental data is that these columns are half-filled, that is, sites exist in both halves of the bicrystal, but cannot be simultaneously occupied due to like-ion repulsion. Half-occupancy overcomes the problem, and gives two columns in projection consistent with the image. Alternatively, they can be considered as a single reconstructed strontium column.

Such sites give insight into how impurities may segregate into the boundary plane; they have significant excess volume and could be attractive sites for high atomic radius impurity species.

These two half-occupied strontium columns form the core of a grain boundary dislocation with Burgers vector  $\langle 100 \rangle$ . This Burgers vector corresponds to a single lattice plane in the bulk, that is, a  $\langle 100 \rangle$  plane can terminate from one side of the boundary or the other, but not both simultaneously. This causes the structure to be microscopically asymmetric, even though the boundary is macroscopically symmetric, a form of microfacetting. The reasons for it can be understood from simple elasticity arguments [50].

Examining  $36^\circ$  and  $67^\circ$  symmetric tilt boundaries revealed further structural units, again all characterized by the presence of half-filled columns. For these

boundaries, the structural units terminated lattice planes from each side of the boundary, preserving the microscopic symmetry of the atomic structure. The set of structural units assigned originally [50] allowed structure models to be constructed for any symmetric  $[001]$  tilt grain boundary using the principle of continuity of boundary structure. However, if we regard unit cells and fragments thereof as simple spacer units (Fig. 20), it is possible to account for all the observed structures using the three pentagonal cores B, B<sup>#</sup> and C. The original  $\{920\}$  boundary is given by the sequence  $AB^{(1)}AB^2A\dots$  where the labels 1 and 2 represent  $\{100\}$  planes terminating from different halves of the bicrystal. With A\* the  $36^\circ$  boundary contains the same B units, as well as additional units of the same geometry but on the other sublattice, that is, a core containing titanium half-columns (B<sup>#</sup> in



**Figure 20.** Structural units for all symmetric and asymmetric  $(001)$  tilt grain boundaries in  $\text{SrTiO}_3$ .

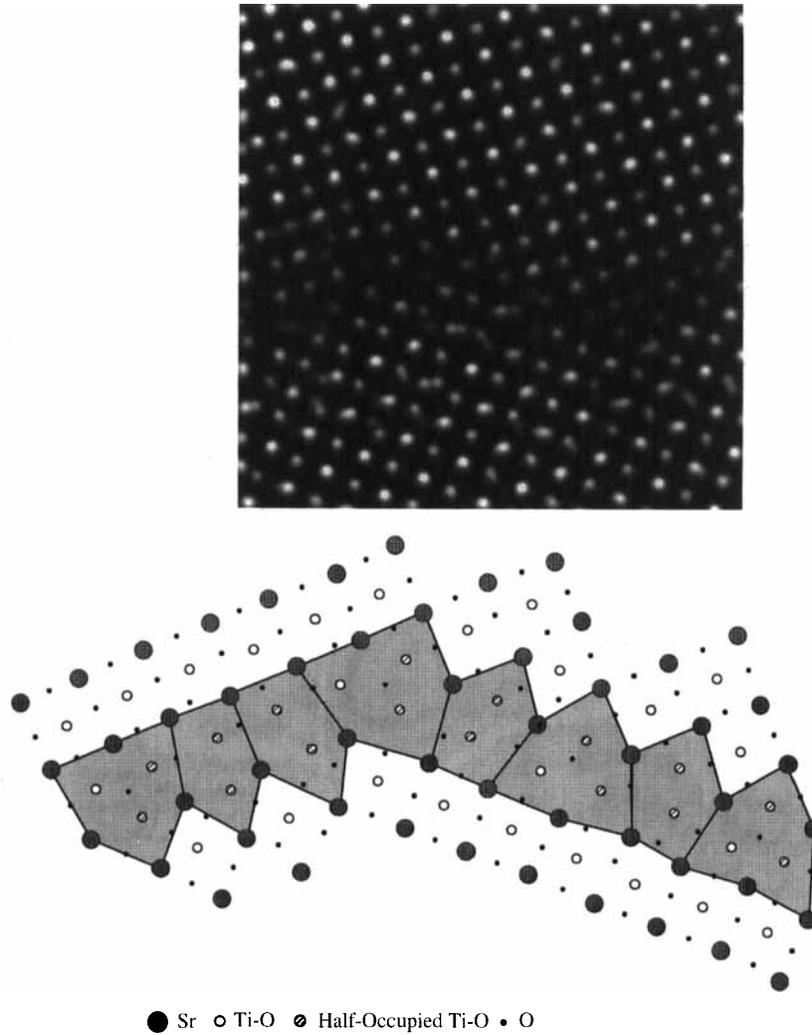
Fig. 20). Thus, only a single type of lattice dislocation, in these two variants, is needed to account for all symmetric boundaries. In this description, continuity of boundary structure is no longer maintained but the number of units required is reduced.

One advantage of this structural unit description is that the same core can be used in its asymmetric configuration to build asymmetric boundaries. Two different units are required, as it is known from crystal lattice geometry that two types of dislocations are needed to describe an asymmetric boundary. With the  $67^\circ$  grain boundary, cores with a  $\langle 110 \rangle$  Burgers vector are used (C in Fig. 20) in the sequence  $A^\#C\dots$ . Images of a  $45^\circ$  symmetric bicrystal [51] which had decomposed into a set of small asymmetric facets are shown in Fig. 21. Units with both  $\langle 100 \rangle$  and  $\langle 110 \rangle$  Burgers vectors are used, but the  $\langle 110 \rangle$  unit now has a core containing Ti half columns ( $C^\#$  in Fig. 20). In these asymmetric boundaries, the same two units coexist over the entire misorientation range, with different frequencies relative to the spacer unit cells.

Such insights into the structure of grain boundaries and likely impurity sites should enable the important link to be established between grain boundary structure and properties. A recent spectacular advance in this area has come from images of  $YBa_2Cu_3O_{7-x}$  grain boundaries, grown on  $SrTiO_3$  bicrystals. It might be expected that due to their similar crystal structures the same units exist in  $YBa_2Cu_3O_{7-x}$  grain boundaries as in  $SrTiO_3$ , and this is confirmed by Fig. 22. It is interesting to note that almost the entire grain boundary is asymmetric, waving back and forth about the orientation defined by the substrate.

To gain insight into the superconducting properties of such boundaries we have used a bond valence sum analysis to examine the region around the core where the copper valence differs from its bulk value necessary for superconductivity. The fact that the asymmetric cores contain partially occupied copper columns would be expected to have a significant effect on the copper valence, and this is borne out by the calculations. A substantial nonsuperconducting zone exists around each structural unit. For a low-angle boundary this leads to a rapid reduction in critical current for small orientation angles [52]. At about  $10^\circ$ , the dislocation strain fields overlap sufficiently to cause the boundary to appear as a 'wall' of nonsuperconducting material, through which the supercurrent must tunnel. Now we can understand the puzzling results in the literature reporting the exponential reduction of critical current with misorientation  $\theta$  [53, 54]. Unit C (see Fig. 20) has the larger Burgers vector, the most strain and the largest nonsuperconducting zone associated with it. Its frequency in the boundary plane increases as  $(1 - \cos\theta)$ , and the width of the nonsuperconducting zone inferred from bond valence sum analysis increases roughly linearly with misorientation. The exponential decrease in critical current is therefore naturally explained [55].

The implications of such studies for improving transport characteristics are clear. Attempts to improve the high-angle behavior should focus on engineering unit C (see Fig. 20) to 'repair' the copper valence; for applications to wires, efforts should be focused on unit  $B^\#$  since this is the dominant dislocation at small misorientation angles. Z-contrast structure

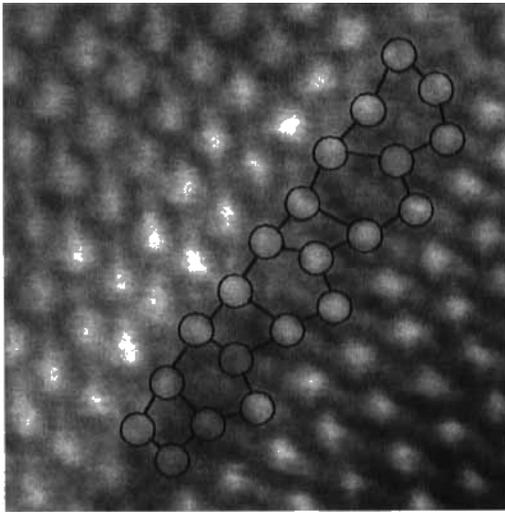


**Figure 21.** Reconstructed 100 kV image of a macroscopically symmetric  $45^\circ$  tilt boundary in  $\text{SrTiO}_3$  showing decomposition into facets, with deduced atomic structure.

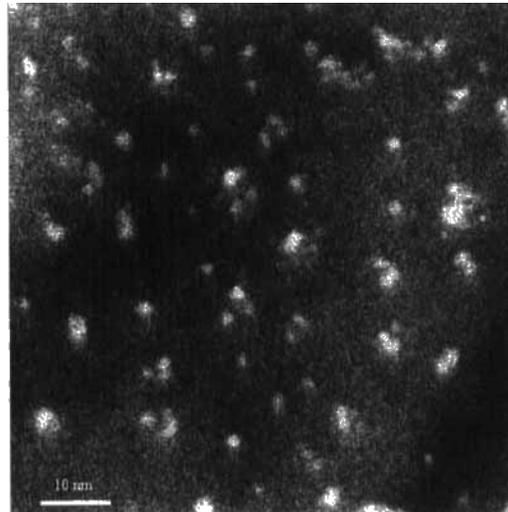
imaging with spectroscopy can form an invaluable tool for understanding grain boundary structure–property relationships, as a route to grain boundary engineering. In the world of high- $T_c$  materials, where grain boundaries are necessary to distinguish competing mechanisms [56], these techniques would seem to have particular potential.

### 2.3.6.3 Nanocrystalline Materials

Of increasing technological importance, nanocrystalline materials represent another area where  $Z$ -contrast methods offer major advantages. It is the combination of freedom from coherent interference effects, such as Fresnel fringes or the speckle pattern from an amorphous



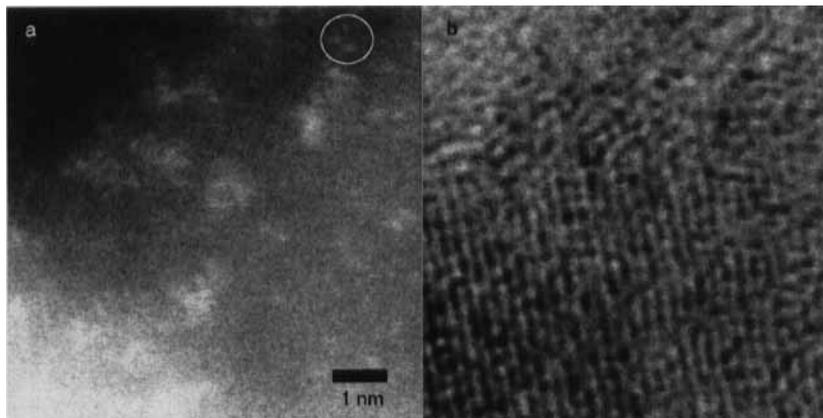
**Figure 22.** Z-contrast image (300 kV) of a predominantly asymmetric tilt boundary in  $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ , grown on a symmetric  $30^\circ$   $\text{SrTiO}_3$  bicrystal. Structural units are equivalent to units  $B^\#$  and  $C^\#$  for  $\text{SrTiO}_3$  (see Fig. 20).



**Figure 23.** Z-contrast images (300 kV) of silicon nanocrystals formed by implantation into  $\text{SiO}_2$  revealing nanometer-sized clusters.

support, with the Z-contrast that provides the insight. As an example, Fig. 23 shows images of nanocrystalline silicon, grown in amorphous  $\text{SiO}_2$  by an ion implantation and annealing procedure [57]. Although the Z-contrast between  $\text{SiO}_2$  and silicon

is not great, because it is the only source of image contrast, the small particles are clearly visible. Surprisingly, they were found to have an internal structure, appearing as clusters of nanometer-sized particles. The smallest of these clusters can contain at most a few tens of atoms.

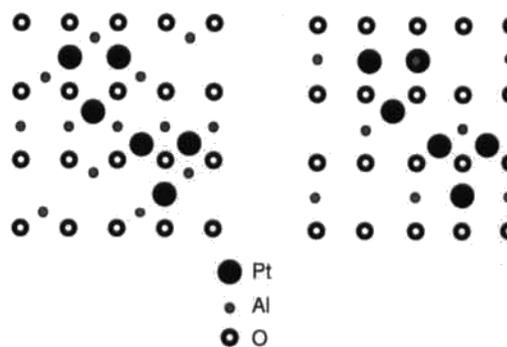


**Figure 24.** Images (300 kV) of a platinum catalyst on a  $\gamma\text{-Al}_2\text{O}_3$  support (a) phase contrast and (b) Z-contrast, obtained simultaneously. The catalyst is seen to comprise of platinum dimers and trimers with some unresolved three-dimensional clusters.

Such images can only be obtained with the 300 kV scanning transmission electron microscope. The higher accelerating voltage reduces beam broadening by a factor of three compared to a 100 kV STEM instrument, and its probe size is almost a factor of two smaller. The probe is successfully able to image nanometer-sized clusters buried hundreds of angstroms deep inside the  $\text{SiO}_2$ . Accurate information on cluster size is essential to solve the issue of whether quantum confinement effects are playing a major role in these materials.

### 2.3.6.4 Catalysts

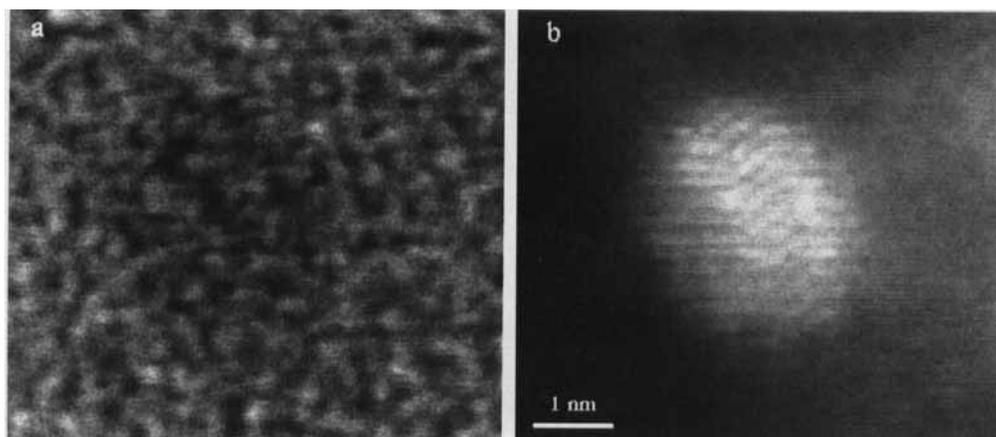
The ultimate example of Z-contrast sensitivity to date is the ability to image single platinum atoms on a  $\gamma\text{-Al}_2\text{O}_3$  support, as seen in Fig. 24. The small bright dots are visible well above the noise level of the substrate, and are arranged as dimers, trimers, and other configurations. The bright field image in this case shows strong  $\{222\}$  fringes, and the orientation of the surface can be deduced from its Fourier



**Figure 25.** Possible surface configurations of platinum on the two  $\{110\}$  surfaces of  $\gamma\text{-Al}_2\text{O}_3$ , deduced from Fig. 24.

transform. It is now possible to find surface sites that explain the observed spacings and orientations, as shown in Fig. 25 [58].

Such information is unobtainable by scanning tunneling microscopy since the support is insulating. Indeed, very little is known about the surface of  $\gamma\text{-Al}_2\text{O}_3$  at all, but such information is essential for understanding the mechanisms of cluster formation. Again, images such as these immediately suggest avenues for theoretical studies. Figure 26 shows bright field



**Figure 26.** (a) Phase contrast and (b) Z-contrast images of a rhodium catalyst on  $\gamma\text{-Al}_2\text{O}_3$  obtained with a 300 kV scanning transmission electron microscope.

and dark field images of a small, three-dimensional rhodium particle on  $\gamma$ -Al<sub>2</sub>O<sub>3</sub>. While the particle is barely detectable in the bright field image because of the amorphous speckle pattern from the carbon support, its internal surface is clear from the Z-contrast image. It is also possible to begin to deduce its external facet and step structure, again suggesting avenues for theoretical studies.

One of the most exciting possibilities would be to examine a set of samples following different chemical treatments to try to distinguish the atomic scale morphology of the catalytically active species itself. Theoretical modeling may then be able to reveal the actual reaction pathway.

## Acknowledgments

The authors are grateful to A. G. Norman, V. P. Dravid, C. W. White, J. G. Zhu, D. R. Liu, R. Gronsky, and Z. Weng-Sieh for provision of samples, and to T. C. Estes and J. T. Luck for technical assistance. This research was sponsored by the Division of Materials Sciences, US Department of Energy, under contract DE-AC05-84OR21400 with Lockheed Martin Energy Systems, and in part by the Laboratory Directed Research and Development Fund of Oak Ridge National Laboratory (ORNL) and appointments to the ORNL Postdoctoral Research Program administered by the Oak Ridge Institute for Science and Education.

## 2.3.7 References

- [1] O. Scherzer, *J. Appl. Phys.* **1949**, *20*, 20.
- [2] J. W. Goodman, *Introduction to Fourier Optics*, McGraw Hill, New York **1968**.
- [3] S. F. Gull, J. Skilling, *IEE Proc.* **1984**, *131F*, 646.
- [4] A. V. Crewe, J. Wall, J. Langmore, *Science* **1970**, *168*, 1338.
- [5] J. M. Cowley, *Ultramicroscopy* **1976**, *2*, 3.
- [6] C. Colliex, C. Mory in *Quantitative Electron Microscopy* (Eds.: J. N. Chapman, A. J. Cra-ven), Scottish Universities Summer School in Physics, Edinburgh, **1984**, p. 149.
- [7] D. E. Jesson, S. J. Pennycook, *Proc. R. Soc. London A* **1993**, *441*, 261.
- [8] A. V. Crewe, J. P. Langmore, M. S. Isaacson in *Physical Aspects of Electron Microscopy and Microbeam Analysis* (Eds.: B. M. Siegel, D. R. Beaman), Wiley, New York, **1975**, p. 47.
- [9] M. S. Isaacson, M. Ohtsuki, M. Utlaut in *Introduction to Analytical Electron Microscopy* (Eds.: J. J. Hren, J. I. Goldstein, D. C. Joy), Plenum Press, New York, **1979**, p. 343.
- [10] A. Howie, *J. Microsc.* **1979**, *117*, 11.
- [11] C. R. Hall, P. B. Hirsch, *Proc. R. Soc. London A* **1965**, *286*, 158.
- [12] D. M. Bird, Q. A. King, *Acta Crystallogr. Ser. A* **1990**, *46*, 202.
- [13] P. A. Doyle, P. S. Turner, *Acta Crystallogr. Ser. A* **1968**, *43*, 390.
- [14] D. E. Jesson, S. J. Pennycook, *Proc. R. Soc. London A* **1995**, *449*, 273.
- [15] B. E. Warren, *X-Ray Diffraction*, Dover, New York, **1990**.
- [16] H. Rose, *Optik* **1976**, *45*, 139, 187.
- [17] R. H. Ritchie, A. Howie, *Phil. Mag. A* **1988**, *58*, 753.
- [18] P. E. Batson, *Ultramicroscopy* **1992**, *47*, 133.
- [19] M. J. Whelan, *J. Appl. Phys.* **1965**, *36*, 2099.
- [20] V. W. Maslen, C. J. Rossouw, *Phil. Mag. A* **1983**, *47*, 119.
- [21] V. W. Maslen, C. J. Rossouw, *Phil. Mag. A* **1984**, *49*, 735.
- [22] C. J. Rossouw, V. W. Maslen, *Phil. Mag. A* **1984**, *49*, 749.
- [23] L. J. Allen, C. J. Rossouw, *Phys. Rev. B* **1990**, *42*, 11644.
- [24] L. J. Allen, *Ultramicroscopy* **1993**, *48*, 97.
- [25] V. W. Maslen, *J. Phys. B* **1983**, *16*, 2065.
- [26] C. J. Rossouw, V. W. Maslen, *Ultramicroscopy* **1987**, *21*, 173.
- [27] H. Kohl, H. Rose, *Adv. Electron. Electron Phys.* **1985**, *65*, 2065.
- [28] P. G. Self, P. Buseck, *Phil. Mag.* **1983**, *48*, L21.
- [29] A. J. Bourdillon, *Phil. Mag.* **1984**, *50*, 839.
- [30] S. J. Pennycook, *Ultramicroscopy* **1988**, *26*, 239.
- [31] O. F. Holbrook, D. M. Bird in *Proc. Microscopy and Analysis 1995*, Jones and Begall, New York **1995**, p. 278.
- [32] P. E. Batson, *Nature* **1993**, *366*, 727.
- [33] N. D. Browning, M. F. Chisholm, S. J. Pennycook, *Nature* **1993**, *366*, 143.
- [34] M. F. Chisholm, S. J. Pennycook, R. Jevasinski, S. Mantl, *Appl. Phys. Letts.* **1994**, *64*, 2409.
- [35] S. J. Pennycook, D. E. Jesson, *Phys. Rev. Lett.* **1990**, *64*, 938.

- [36] S. J. Pennycook, D. E. Jesson, *Ultramicroscopy* **1991**, 37, 14.
- [37] S. J. Pennycook, D. E. Jesson, *Acta Metall. Mater.* **1992**, 40, S149.
- [38] R. F. Loane, P. Xu, J. Silcox, *Ultramicroscopy* **1992**, 40, 121.
- [39] J. Fertig, H. Rose, *Optik* **1981**, 59, 407.
- [40] R. F. Loane, E. J. Kirkland, J. Silcox, *Acta Crystallogr. Ser. A* **1988**, 44, 912.
- [41] A. J. McGibbon, S. J. Pennycook, J. E. Angelo, *Science* **1995**, 269, 519.
- [42] A. J. McGibbon, S. J. Pennycook, *J. Microsc.* (in press).
- [43] J. Hornstra, *J. Phys. Chem. Solids* **1958**, 5, 129.
- [44] M. F. Chisholm, S. J. Pennycook in *Proceedings Microscopy and Microanalysis 1996*, San Francisco Press **1996**, p. 332.
- [45] O. L. Krivanek et al., *Phil. Mag.* **1977**, 36, 931.
- [46] M. F. Chisholm, M. Mostoller, T. Kaplan, S. J. Pennycook, *Phil. Mag.* (in press).
- [47] A. Bourret, J. L. Rouvire, *Springer Proc. Phys.* **1989**, 35, 8.
- [48] A. P. Sutton, V. Vitek, *Phil. Trans. Roy. Soc. A* **1983**, 309, 1.
- [49] M. M. McGibbon et al., *Science* **1995**, 266, 102.
- [50] N. D. Browning et al., *Interface Science* **1995**, 2, 397.
- [51] M. M. McGibbon et al., *Phil. Mag.* (in press).
- [52] M. F. Chisholm, S. J. Pennycook, *Nature* **1991**, 351, 47.
- [53] D. Dimos, P. Chaudhari, J. Mannhart, *Phys. Rev. B* **1990**, 41, 4038.
- [54] Z. G. Ivanov et al., *Appl. Phys. Lett.* **1990**, 59, 3030.
- [55] N. D. Browning, P. D. Nellist, D. P. Norton, S. J. Pennycook, *Nature* (submitted).
- [56] C. C. Tsuei et al., *Science* **1996**, 271, 329.
- [57] J. G. Zhu, C. W. White, J. B. Budai, S. P. Withrow, Y. Chen, *J. Appl. Phys.* **1995**, 78, 4386.
- [58] P. D. Nellist, S. J. Pennycook, *Science* (submitted).

“The submitted manuscript has been authored by a contractor of the US Government under contract No. DE-AC05-84OR21400. Accordingly, the US Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for US Government purposes.”

## 2.4 Scanning Auger Microscopy (SAM) and Imaging X-Ray Photoelectron Spectroscopy (XPS)

### 2.4.1 Introduction

The properties of solids are related to their electronic structure on the one hand, and to the presence of defects such as impurities, vacancies, dislocations, grain boundaries, etc. on the other hand. The surface is not thought of as containing lattice defects similar to those just mentioned; nevertheless it is also very influential on the properties of solids. Indeed, the space lattice, which is a three-dimensional infinite periodic array of points, is the geometrical representation of a perfect crystal and departures from this mathematical model are by definition lattice defects. The discontinuity which constitutes the surface is therefore a lattice defect. As a result of the presence of unsaturated bonds, the surface induces properties in the solid which are the direct result of its existence. In this manner the surface is very similar to the other lattice defects already mentioned above and abundantly discussed in this Handbook.

The surface is also the agent through which the solid interacts with its surroundings and is therefore of great importance

in, for example, microelectronics, catalysis, corrosion, etc. It has therefore been intensively studied since the early 1960s when UHV (ultrahigh vacuum) became readily available. Indeed, to obtain clean surfaces, vacua in the  $10^{-10}$  mbar range are necessary and these can now routinely be obtained.

When studying a surface one is interested in its structure and chemical composition, that is, the chemical bonding at the surface and in the surface layer. Yet, when studying a surface one does not restrict oneself to the uppermost layer only. It is important to examine further layers, even to a depth of a few nanometers thick, because properties related to the surface are also influenced by the subsurface layers.

A large number of techniques for studying these properties have been developed and, in this chapter, two of the oldest, but most widely used, techniques are described: namely, Auger electron spectroscopy (AES) and X-ray photoelectron spectroscopy (XPS) or ESCA. Both techniques are used to study the chemical composition of the surface and their information depth is limited to a few atomic

layers, making them ideally suited for this purpose. AES was developed about 1967 first using the LEED retarding grids as electron spectrometer, although soon afterwards the cylindrical mirror analyzer (CMA) was introduced and became commercially available. AES has now evolved into the most widely used surface analysis technique. It is found not only at universities and research institutes, but also in industry. XPS became commercially available around 1970, following the pioneering work of Siegbahn's group at Uppsala [5]. Its potential was immediately recognized and, consequently, XPS also evolved into a widely used surface analysis technique, especially after it was realized that the information depth of both techniques is practically equal. For a long time, however, the limited lateral resolution of XPS constituted a drawback for its use in, for instance, microelectronics. The recent developments described in this chapter have, however, drastically altered this perspective.

Lateral resolution in AES was less of a problem than in XPS because, in 1972, it was suggested that SEM and AES should be combined into a single instrument. This resulted in AES spectra with high lateral resolution, as will be further described. The scanning Auger microscope (SAM) consequently became available by the late 1970s and has subsequently found extensive use, particularly in industrial laboratories. Imaging XPS was introduced in the mid-1980s, and benefited from developments in the field of X-ray focalization and monochromatization and improved electron detection capabilities such as multichannel plates.

In broad terms, Sec. 2.4.2 of this Chapter summarizes the general principles and

techniques of AES and XPS. Section 2.4.3 describes, in some detail, the experimental aspects of both scanning Auger microscopy and imaging XPS. Finally, Sec. 2.4.4 describes the characteristics of the images obtained by these techniques.

## **2.4.2 Basic Principles of Auger Electron Spectroscopy (AES) and X-Ray Photoelectron Spectroscopy (XPS)**

### **2.4.2.1 Auger Electron Spectroscopy (AES)**

Many review papers in the literature describe in detail the different aspects of the Auger process. Consider, for example, the review paper by Fiermans and Vennik [1]. Only the main aspects necessary to understand what follows are summarized here.

#### *The Auger Process*

Auger electrons emerge from a solid when the latter is excited with energetic electrons (in the keV range), X-rays, ions, protons, etc. They are a consequence of radiationless rearrangements of the electrons in atoms in which a core hole has been created by the exciting radiation. The energy released during this radiationless rearrangement is transferred to an electron which emerges from the solid with a certain kinetic energy, the atom being left behind in a doubly ionized state. This electron is called the Auger electron, after the French

physicist, Pierre Auger, who first described this process in the 1920s.

#### Auger Notations

Taking into account the spin-orbit splitting, the electrons are arranged in an atom in the shells K(1s), L<sub>1</sub>(2s), L<sub>2</sub>(2p<sub>1/2</sub>), L<sub>3</sub>(2p<sub>3/2</sub>), M<sub>1</sub>(3s), M<sub>2</sub>(3p<sub>1/2</sub>), M<sub>3</sub>(3p<sub>3/2</sub>), . . . , a notation used in X-ray spectroscopy. Considering, for instance, the removal of a 1s electron from the K-shell, one has radiative KL<sub>2</sub> (and KL<sub>3</sub>) transitions giving rise to X-ray photons, and a number of radiationless transitions where the energy is transferred to a second electron. This can also be an electron in the L<sub>2</sub> (or L<sub>3</sub>) shell and the Auger electron is then called the KL<sub>2</sub>L<sub>2</sub>-electron. This, however, is not the only possibility and a whole series of KLL-processes has to be considered, each giving rise in L-S coupling to a number of spectroscopic terms (<sup>1</sup>S<sub>0</sub>; <sup>1</sup>P<sub>1</sub>, <sup>3</sup>P<sub>2,1,0</sub>; <sup>1</sup>S<sub>0</sub>, <sup>3</sup>P<sub>2,1,0</sub>, <sup>1</sup>D<sub>2</sub> for the 2s<sup>0</sup>2p<sup>6</sup>, 2s<sup>1</sup>2p<sup>5</sup> and 2s<sup>2</sup>2p<sup>4</sup> final states, respectively). Some of these terms are parity forbidden (2s<sup>2</sup>2p<sup>4</sup>: <sup>3</sup>P for instance) and in *j-j* coupling only six terms remain, namely: KL<sub>1</sub>L<sub>1</sub>, KL<sub>1</sub>L<sub>2</sub>, KL<sub>1</sub>L<sub>3</sub>, KL<sub>2</sub>L<sub>2</sub>, KL<sub>2</sub>L<sub>3</sub> and KL<sub>3</sub>L<sub>3</sub>.

In practice one often finds a set of terms determined by an intermediate coupling scheme. The notation including the symbol of the spectroscopic term considered, is generally adopted, e.g. KL<sub>2,3</sub>L<sub>2,3</sub>(<sup>1</sup>D). The same notation is used for Auger processes with initial holes in L, M, N, . . . shells (e.g., L<sub>3</sub>M<sub>4,5</sub>M<sub>4,5</sub>(<sup>1</sup>D), etc.).

When one or more levels participating in the process are situated in a composite valence band, the symbol V is used (e.g., KVV, L<sub>3</sub>VV, L<sub>3</sub>M<sub>2,3</sub>V, etc.).

Coster-Kronig transitions are Auger processes involving sublevels of the same

shell (e.g., L<sub>1</sub>L<sub>2,3</sub>M<sub>2,3</sub>). Their rate is strongly dependent on the energy separation between the sublevels considered. Super Coster-Kronig transitions are processes occurring in the same shell (e.g., M<sub>1</sub>M<sub>2</sub>M<sub>4,5</sub>, etc.).

#### Energies of the Auger Transitions

When electron holes (e.g., in a photoemission or in an Auger experiment) are created, a charge rearrangement takes place in the atom. It is, therefore, necessary to include in the Auger kinetic energy formula for an Auger transition involving, for instance, the *j*, *k* and *l* levels, terms related to multiplet splitting and electronic relaxation (see [2] for a more detailed account):

$$E_{\text{kin}}(jkl; X) = E_{\text{b}}(j) - E_{\text{b}}(k) - E_{\text{b}}(l) - F(kl; X) + R_{\text{c}} \quad (1)$$

wherein *X* is the multiplet state resulting from the coupling of the two holes *k* and *l* in the final state. The *E<sub>b</sub>*-terms are ground-state one-electron binding energies, for which experimental values can be used. In this way initial state chemical effects and one-hole electronic relaxation phenomena are automatically taken into account.

The two-electron *F*-term accounts for the fact that within the two-step model considered for the Auger process, upon emission of the Auger electron *l*, already one electron, the electron *k*, is missing. As a result, the binding energy of the electron *l* increases with respect to the ground-state value, *E<sub>b</sub>*(*l*). The *R<sub>c</sub>*-term is a supplementary term, called cross relaxation energy [2], which contains both an atomic and an extra-atomic contribution. It primarily accounts for the more repulsive medium

that is created for the electron  $l$  due to the relaxation that has taken place towards the electron hole  $k$ .

In this formula the  $F$  and  $R_c$  terms have to be calculated, which sometimes can be done for atomic and quasiatomic Auger transitions. Experimentally, however, a semiempirical formula is preferred, using experimentally determined and published electron binding energies.

Consider, for instance, an Auger transition  $KL_2L_3$ . To a first approximation, the following expression for kinetic energy is readily obtained:

$$E_{\text{kin}} = E_K - E_{L_2} - E_{L_3(L_2)} \quad (2)$$

Self-consistency is introduced in such semiempirical expressions through the Chung and Jenkins formula [3]:

$$E_{\text{kin}} = E_K(Z) - \frac{1}{2}[E_{L_2}(Z) + E_{L_2}(Z+1)] - \frac{1}{2}[E_{L_3}(Z) + E_{L_3}(Z+1)] \quad (3)$$

Using Eq. (3), kinetic energy values have been tabulated by Coghlan and Clausing [4] and these are now widely used.

The Auger energies measured are kinetic energies referred to the vacuum level of a spectrometer (Fig. 1). Since binding energies generally are referred to the Fermi level, it is necessary to add the spectrometer work function  $\Phi_{\text{anal}}$  to the measured Auger energy to obtain values comparable with those calculated.

The spectrometric work function is an unknown factor and, furthermore, can change from one experiment to another. Normally, the energy scale is calibrated using suitably chosen primary peaks and once the system is calibrated it is tacitly assumed that it is not modified in the course of experiments.

Since the primary energy of the exciting electrons in AES generally is in the range

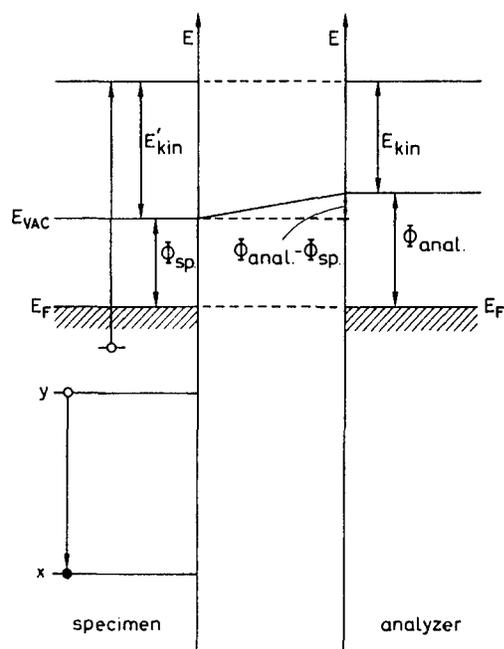


Figure 1. Energy-level diagram showing the relationship between actual and measured kinetic energies.

3–10 keV, only K-shell core holes of the light elements (C, O, ...) can be created, giving rise to the K-series Auger lines. For medium and heavy elements one excites the L-, M-, ... series Auger lines. Summarizing, the observed Auger electron kinetic energies are generally situated in the energy range 0–2000 eV, limiting the information depth. Indeed, the escape depth of these electrons, without energy loss, is limited to a few atomic layers. This point will be discussed in detail in the paragraph on quantitative analysis.

#### Chemical Shift and Line Shape of Auger Lines

Depending on the chemical environment of the ion in which the transition occurs, shifts in the positions of Auger peaks are commonly observed. The so-called

chemical shifts are generally larger in AES than in XPS (see below). In AES the observed shifts are a sum of binding energy shifts (the true chemical shift) and the change in total relaxation energy. Unfortunately, the latter is difficult to determine for most compounds. These chemical shifts are important when imaging the chemical composition. Therefore, this detail is discussed in the following account. The shape of an Auger line is also strongly dependent on the chemical environment and can also be used for chemical composition imaging as will be shown.

#### *Experimental Aspects of AES*

(a) *Energy analyzers:* These are abundantly discussed in the second paragraph of this chapter.

(b) *Sample preparation:* A specimen prepared in normal atmospheric conditions is covered by a layer of contamination. This layer gives rise to relatively strong carbon and oxygen spectral lines and reduces the intensity of the Auger lines of the elements present in the solid. To increase the signal-to-background ratio, the surface is normally cleaned by Ar<sup>+</sup>-bombardment. This, however, carries in it the danger of preferential sputtering of the solid under study, with the inherent danger of altering relative intensities in an elemental quantitative analysis. A better way to obtain atomically clean surfaces is via UHV in situ preparation involving scraping, fracture, cleavage, etc. However, this is not always possible.

*Beam effects:* Certain compounds are subject to rather severe beam effects in AES. The primary electron beam can, for

example, cause surface decomposition, electron assisted adsorption, electron assisted desorption, or induce formation of compounds (e.g., carbide formation on metals), which have all to be taken into account when performing the surface chemical composition analysis. The Ar<sup>+</sup> ion beam bombardment mentioned above is commonly used to record depth profiles. In such an analysis the specimen is subjected to both an energetic ion and electron beam bombardment, generally resulting in a modification of the original surface composition. This has clearly to be borne in mind when interpreting these profiles.

Another beam effect concerns nonconducting specimens. To obtain reliable Auger spectra, the specimens should be conducting and at ground potential. If this is not the case charging will immediately occur, rendering impossible any further measurement.

### **2.4.2.2 X-Ray Photoelectron Spectroscopy (XPS)**

#### *The XPS Process*

The XPS process is nothing more than the photoelectric process, whereby an electron is emitted from a solid when the latter is irradiated with X-rays. In the emitting atom an electron hole is created and charge rearrangement will occur because a nonrelaxed one-electron hole system is not an eigenstate of the corresponding Hamiltonian. A charge rearrangement will take place to screen-off the positive hole, thereby lowering the total energy of the system. As the charge rearrangement generally occurs in a time interval that is

comparable with the time of the photoemission process itself, the reduction of the total energy, with respect to the frozen orbital situation, leads to an increase of the kinetic energy of the outgoing photoelectron. Since the binding energy of an electron is defined as the difference between the kinetic energy of the incoming photon and the kinetic energy of the outgoing photoelectron, that is,

$$\begin{aligned}
 E_{\text{kin}} (\text{outgoing } e^-) &= h\nu (\text{incoming photon}) \\
 &\quad - E_b(i) (\text{binding energy of the } i\text{th electron})
 \end{aligned}$$

the increase of the kinetic energy of the detected electron results in a reduction, of the order of a few eV, of the effectively experimentally observed binding energy. These electronic screening effects occur in free atoms as well as in the condensed phase. As a result of the so-called extra-atomic relaxation, screening will normally be larger for the condensed phase than for free atoms. Consequently, in the discussion of chemical shifts (below), it should be borne in mind that these are not only determined by initial state charge transfer phenomena, but also by final state electronic screening effects.

When measuring the kinetic energy of the photoelectron with an analyzer, the work function,  $\Phi$ , of the latter must be taken into account, binding energies being referred towards the Fermi-level:  $E_{\text{kin}} = h\nu - E_b - \Phi$ .

For photoelectrons, a notation based on an X-ray notation, using the spectroscopic level (e.g., C1s, Si2p<sub>1/2</sub>, Si2p<sub>3/2</sub>, etc.) is adopted. To a first approximation, the initial states before photoemission are

characterized by completely filled shells, while after photoemission one electron hole is present. One consequently only expects the corresponding spin-orbit splitting spectroscopic terms, explaining the notation used above. If, however, the electron hole thus created can couple with other holes already present, multiplet splitting occurs and the photoline becomes much more complex than the single lines just described.

Since the FWHM value of the photoline is determined by, among other factors, the width of the exciting X-ray photon, narrow monoenergetic soft X-ray lines such as AlK $\alpha$  (1486.6 eV) or MgK $\alpha$  (1253.6 keV) are used as primary excitation. The kinetic energies of the photolines are consequently limited to these values. The escape depth of the photoelectrons is therefore comparable to the one discussed for Auger electrons. This point will be dealt with in more detail in Sec. 2.4.2.4 of this Chapter.

In an XPS spectrum, one notices the presence of core level photolines next to the valence band structure. In such a spectrum one normally can also clearly distinguish the presence of Auger lines. Indeed, upon creation of the core holes, deexcitation readily occurs, giving rise to the emission of Auger electrons, which are measured in the spectrometer (Fig. 2).

Determining binding energies and consequently identifying the chemical composition of a surface layer is the main subject of XPS. It is also known under the popular acronym ESCA (electron spectroscopy for chemical analysis) [5].

Analysis is based on the fact that each element has a unique spectrum and can

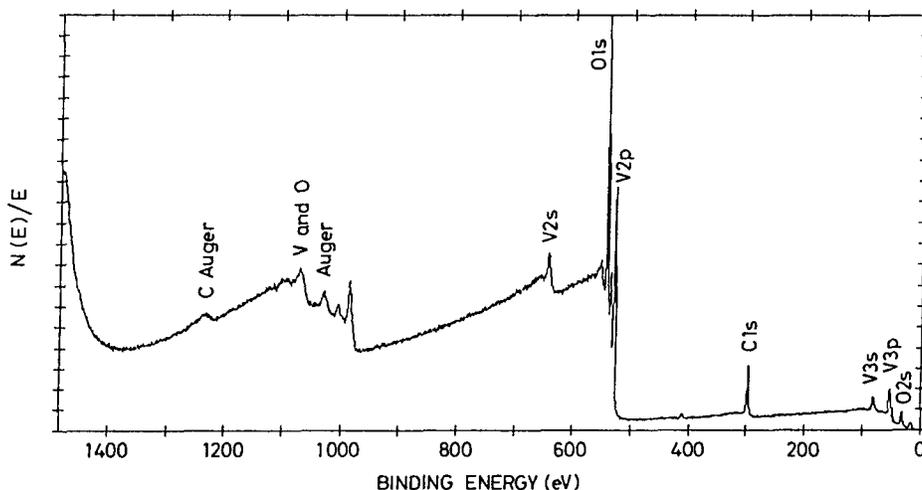


Figure 2. XPS spectrum of  $V_2O_5$  obtained with a CHA/lens combination (CHA in retarding mode).

therefore be used as a kind of fingerprint. It is also straightforward that a mixture of elements gives rise to a mixture of these unique spectra and, to a first approximation, quantitative data can be obtained from the peak areas. In general, the identification of the peaks poses no severe problems, providing that one is aware of the fact that satellites such as plasmon losses, shake-up, and shake-off satellites may be present.

#### Beam Effects

Beam effects are much less severe in XPS than in AES. Indeed, the soft X-ray beam induces only minor photoreduction effects on certain compounds. Charging is also less of a problem when measuring non-conducting samples. In XPS the sample charges positively as a result of the photoemission, inducing a shift towards higher binding energies and a general broadening of the photolines. This effect can be compensated by flooding the sample with very

low energy electrons produced by an auxiliary electron flood gun.

#### 2.4.2.3 Quantitative Analysis in AES and XPS

There are two possible approaches to the problem of quantitative analysis: (i) a purely formal *ab initio* approach, or (ii) a semi-empirical approach. The latter is routinely installed in most AES and XPS instruments. Inevitable approximations in these routines can lead to considerable errors. It should also be pointed out that AES and XPS are not trace analysis techniques, their detection limits being in the range 0.1–1% for most elements. The semi-empirical approach was discussed in detail a few years ago by Nebesny et al. [6] and we will rely heavily on this excellent paper in our discussion.

An electron beam excited Auger signal is given by a formula, introduced by Seah

[7] (and using the notation of Nebesny et al. [6]):

$$I_A = I_0 \times \sigma_{A(E_p)} \times (1 + r_{M(E_A, \alpha)}) \times T_{(E_A)} D_{(E_A)} \times \int_0^\infty N_A(z) \times \exp[-z/\lambda_{M,E} \cos \theta] dz \quad (4)$$

where  $I_0$  is the primary electron beam current ( $\text{A cm}^{-2}$ ),  $\sigma_{A(E_p)}$  is the electron impact ionization cross section at primary energy  $E_p$ , the factor  $(1 + r_{M(E_A, \alpha)})$  refers to Auger electrons at energy  $E_A$  produced by backscattered electrons in which the term  $r_{M(E_A, \alpha)}$  is the backscattering coefficient at angle  $\alpha$  to surface normal,  $T_{(E_A)}$  is the transmission efficiency of analyzer at  $E_A$  (Auger kinetic energy), and  $D_{(E_A)}$  is the

detector efficiency. The integral contains the analyte concentration ( $N_A$ ), modified by an exponential decay term, with decay constant  $\lambda$ , the electron escape depth at energy  $E_A$ , and  $\theta$  is the angle of analysis to surface normal.

The escape depth is defined as the depth from which an Auger electron (or for that matter a photoelectron) can escape from the solid without undergoing inelastic collisions whereby it may lose part of its energy and disappear in the background. Experimentally determined values of  $\lambda$  give rise to the so-called 'universal escape depth' curve from which it follows that the escape depth for electrons with kinetic energy limited to 2000 eV, is only a few atomic layers (0–2 nm) (Fig. 3). Assuming

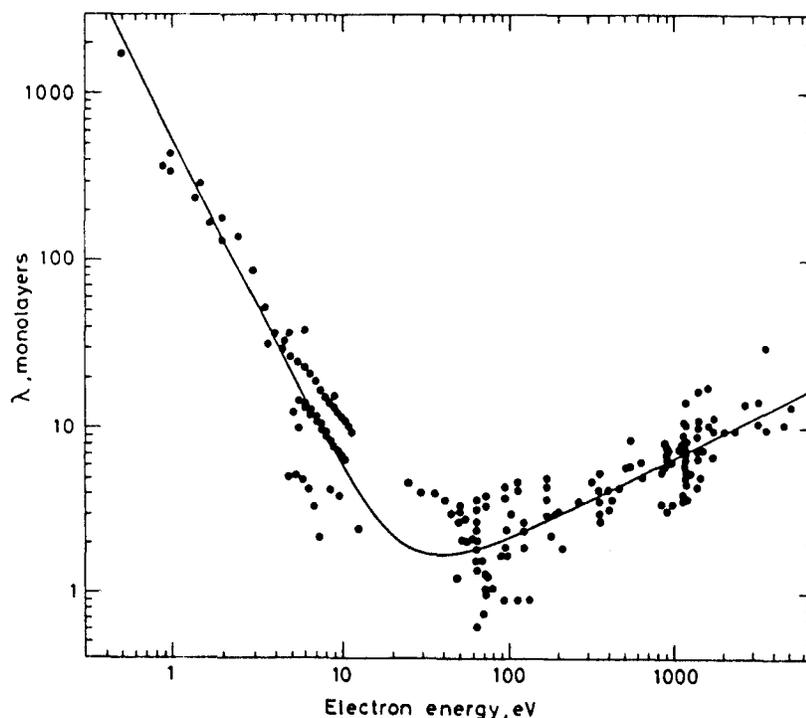


Figure 3. Escape depth in function of electron energy. Adapted from [7]. Reproduced by permission of M. P. Seah and John Wiley & Sons Ltd.

Auger emission occurs within  $z = 5\lambda$ , an information depth limited to maximum  $100 \text{ \AA}$  is obtained. This equation is an ab initio formula wherein the principal uncertainties arise from  $\sigma_{A(E_p)}$ ,  $r_M$ ,  $\lambda_{M,E_A}$  and the instrumentally dictated parameters  $T_{(E_A)}$  and  $D_{(E_A)}$ .

The easiest application of this formula is for submonolayer adsorbates, where the complete formula is employed [8]. When the solid is homogeneous within the analysis depth for all the elements of interest (A and B for instance), the uncertainties can be partially canceled by considering relative atomic values. In this case one obtains:

$$\frac{I_A}{I_B} = \frac{\sigma_A(E_A)}{\sigma_B(E_B)} \times \frac{(1 + r_{M(E_A)})}{(1 + r_{M(E_B)})} \times \frac{\lambda_{M,E_B}}{\lambda_{M,E_A}} \times \frac{T_{E_A}}{T_{E_B}} \times \frac{D_{E_A}}{D_{E_B}} \times \frac{N_A}{N_B} \quad (5)$$

If  $T$  and  $D$  are known to within a certain tolerance, these terms can then be canceled. If pure element standards are available for A and B, an atomic percentage of A and B can be computed from  $I_A/I_B$  and  $I_A^\infty/I_B^\infty$  (the Auger current ratio from the pure element standards) and a correction for matrix effects, a 'matrix factor'  $F_{A,B}^A$ . An alternative is to use standards that are near the suspected composition and have the same matrix as the unknown.

The relative atomic ratio  $(N_A/N_B)_{\text{unk}}$  of the unknown is computed from the formula:

$$\left(\frac{N_A}{N_B}\right)_{\text{unk}} = \left(\frac{I_A}{I_B}\right)_{\text{unk}} \times \left(\frac{I_B}{I_A}\right)_{\text{std}} \times \left(\frac{N_B}{N_A}\right)_{\text{std}} \quad (6)$$

One obtains (see Nebesny et al. [6] for more details) for the atomic percentage  $X_A$ :

$$X_{A(\text{unk})} = \frac{I_{A(\text{unk})}}{\sum_i I_{i(\text{unk})}} \times \frac{\sum_i I_{i(\text{std})}}{I_{A(\text{std})}} \times \frac{1}{X_{A(\text{std})}} \quad (7)$$

where  $(I_A/I_B)_{\text{unk}}$  and  $(I_A/I_B)_{\text{std}}$  are Auger signal ratios for the unknown and the standard, and  $(N_A/N_B)_{\text{std}}$  is the stoichiometry of the standard. Another option consists of using sensitivity factors for each element obtained from different sources. This option, however, has to be used with care since the factors can differ from one instrument to another.

As an example, the procedure employed by Physical Electronics Inc., and routinely installed in the instrument [9], can be mentioned. With sensitivity factors using, for instance, Ag as a standard, the procedure leads to a much simpler formula:

$$C_X = \frac{I_X}{I_{\text{Ag}} S_X D_X} \quad (8)$$

where  $C_X$  is the atomic concentration of element X,  $I_X$  is normally the peak-to-peak amplitude of the element X,  $I_{\text{Ag}}$  is the peak-to-peak amplitude for the Ag standard,  $D_X$  is a relative scale factor which can be calculated and  $S_X$  is the relative sensitivity of element X towards silver.

By using peak-to-peak amplitudes (discussed below) variations in the Auger line shape are neglected. For instance, when recording a depth profile, elemental concentration data using this formula can be transformed by statistical methods (factor analysis and MLCFA correlations) into more reliable chemical information [10]. In MLCFA (maximum likelihood

common factor analysis), the Auger peak shape is taken into account and the number of components building this shape is determined. Each component corresponds to a particular chemical environment of the element considered. We refer the reader to the related literature [10] for a more detailed account of this powerful approach to quantitative chemical composition determination.

For XPS similar formalisms have been developed, taking into account peak shapes and the presence of satellites. In XPS the area under the photoline is normally used for quantitative analysis; however, the question is how to obtain accurate peak areas both in AES and XPS. A particular problem concerns accurate background subtraction.

In AES the background arises from different processes [6]:

- energy loss by electrons from the source originally at higher kinetic energies;
- electrons from Auger emission at higher kinetic energy;
- intrinsic losses, mainly in free electron metals, due to relaxation effects;
- extrinsic energy losses such as plasmons.

The first two processes lead to an exponentially growing background towards lower kinetic energies, the so-called 'secondary cascade' on which the Auger electrons are superimposed.

In XPS the principal processes contributing to the background are:

- photoemission events of higher kinetic energy, leading to the 'secondary cascade';
- similar extrinsic losses due to plasmons;
- shake-up and shake-off satellites.

In both AES and XPS instruments, the background is normally subtracted following different steps: (i) dividing the spectral intensities by the kinetic energy; (ii) linearization of the log–log form of the data and subtraction; and (iii) the integral background method [11] is available in virtually all commercial instruments. In this technique the integral is calculated over the whole peak area and a calibrated curve thus obtained is subtracted from the peak area. Further details of (i)–(iii) are provided in [6]. Ideally, one could then deconvolute the experimental peak shape with an instrumental curve obtained with a primary peak at the peak value under study. However, this is not commercially available in the present instruments.

Concentrations and concentration profiles are then calculated from the peak areas obtained after carrying out the above background subtraction routine. One might, in principle, decompose the photoline into its different components when different valence states (chemical shifts) are present, and then calculate relative concentrations of the different chemical states of the element present. This approach is not routinely available in most commercial instruments as it involves rather long data acquisition times.

## 2.4.3 Scanning Auger Microscopy (SAM) and Imaging XPS

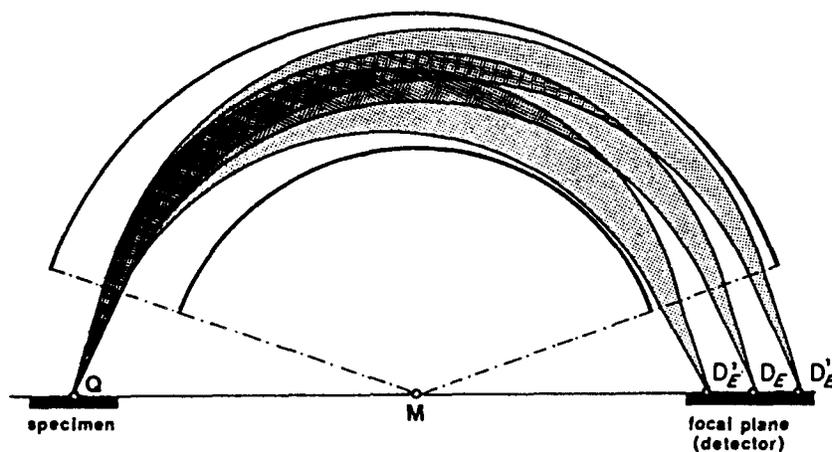
### 2.4.3.1 Basic Principles of Imaging

In microscopic imaging techniques, just as in classical optical systems, the object or

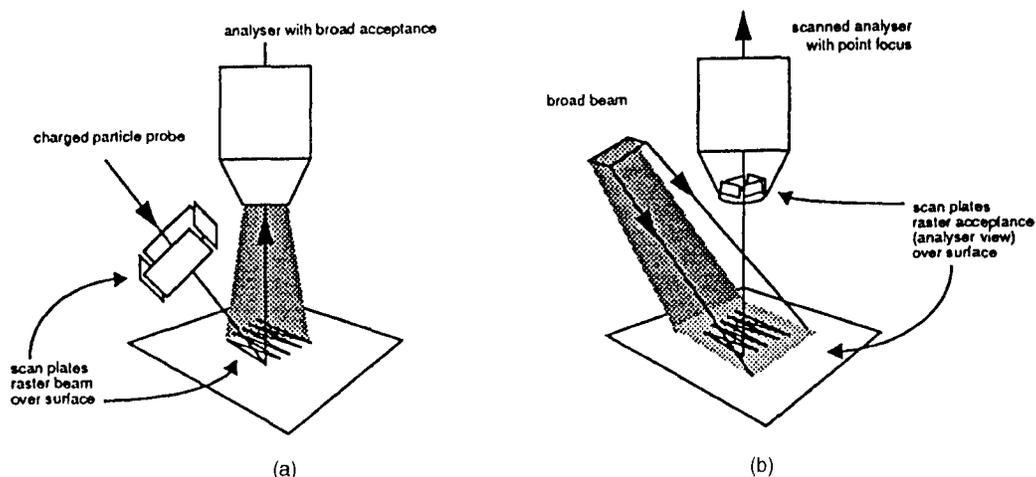
sample to be viewed is illuminated by a source, broad enough to illuminate the whole sample at once. The reflected light or emitted beam from the sample is then passed through a magnifying imaging device (cf. an optical lens) and hence a magnified image of an object or sample is obtained in which all details are seen simultaneously. In pure AES or XPS spectroscopy, the primary aim is to energy analyze either Auger and/or photoelectrons originating from a sample under the influence of the illumination or excitation of an electron or X-ray beam. It is, of course, the energy analysis that is essential to both AES and XPS, which gives the required elemental and binding energy oriented information. However, as in most energy analyzers, the energy dispersion results in a spatial displacement of the image as a function of energy (Fig. 4). The combination of energy analysis with imaging is not straightforward and in the simple scheme described one ends up with a 'picture' at the detector plane which is the convolution of energy information and

image information. Such a picture is neither interpretable nor usable for simple microscopic purposes.

The simplest solution to this problem is to step down from the concept of global or simultaneous imaging of the sample and to divide it into as many pixels as needed to obtain a good quality image: in other words, to obtain a good lateral resolution. Each pixel then acts as a point source for the energy analyzer. By performing a sequential energy analysis of each pixel, a picture of the sample can be reconstructed at a fixed electron energy. The question now is how to divide the sample into pixels. If the illuminating beam can be focused and steered or scanned over the sample as with charged particle beams, a sample is easily 'pixeled'. The area of the pixel or point source, seen by the energy analyzer, is then defined by the focusing properties of the illuminating or primary beam. The location of the pixel on the sample depends on the spatial position of the primary probe and is defined by deflecting voltages on scan plates between



**Figure 4.** Energy selection gives rise to a spatial displacement of the image. Adapted from N. Gurker, M. F. Ebel and H. Ebel, *Surf. Interface Anal.* **1983**, *5*, 13 (Fig. 5). Reproduced by permission of N. Gurker and John Wiley & Sons Ltd.



**Figure 5.** (a) Source defined imaging: the primary excitation source is scanned over the sample to be viewed. The analyzer, with an acceptance broad enough to see the scanned area, accepts the information of each pixel. (b) Detection defined imaging: the sample area under investigation is illuminated with the primary excitation source. The analyzer with restricted acceptance sees each pixel. Adapted from Kratos System Overview Axis 165. Reproduced by permission of Kratos Analytical Ltd.

source and sample. For this approach, an analyzer with an acceptance broad enough to see the scanned area on the sample is needed. This solution to energy dispersive imaging is called (Fig. 5a) 'source defined' imaging. If the primary beam is not steerable and hardly focusable, as is traditionally the case with, for instance, X-rays, another solution is required if charged particles are to be detected. By reducing the acceptance of the energy analyzer, the lateral dimensions of the area under investigation are reduced until the energy dispersive system sees a pixel or point source. By rastering this accepted area over the sample, using deflection plates in front of the analyzer (Fig. 5b), electrons are again sequentially collected from the different pixels. This approach to imaging is called 'detection defined' imaging and requires a broad and homogeneous primary excitation beam. As detailed later, it is between these two apparent extreme solutions that

other approaches to imaging in AES and XPS have developed into powerful analytical tools.

#### 2.4.3.2 General Aspects of Analyzers

Basically, there are three ways in which the energy of charged particles may be measured.

##### *Retarding Field Analyzers*

The idea behind the use of retarding fields is that only those particles with sufficient energy can reach a collector. Clearly, such retarding fields act as high pass energy filters and the collected particle current is given by

$$I(E_0) = K \cdot \int_{E_0}^{\infty} N(E) dE \quad (9)$$

where  $K$  is the instrumental constant,  $E_0$  is the cut off energy corresponding with a retarding potential  $V_0$ , and  $N(E)$  is the energy distribution of incident particles.

By modulating the retarding potential  $V_0$  with an alternating signal  $\Delta V_0$  and by synchronously detecting the AC part of the collected current, one effectively measures the first derivative of  $I$  with respect to  $E_0$  [12], that is,

$$\frac{dI(E_0)}{dE_0} \sim N(E_0) \tag{10}$$

In fact, this was how, in 1967, AES was introduced by using LEED optics as a retarding field analyzer. Considering the secondary electron emission distribution of Fig. 6, it should be noted that Auger peaks are rather small features. This is the reason why normally  $dN(E_0)/dE$  versus

$E_0$  spectra were recorded and amplified, using well established lock-in amplifier techniques. Synchronous detection of the second harmonics content of the current  $I(E_0)$  gives a signal proportional with the required first derivative  $dN(E_0)/dE_0$ . In this second harmonics mode of the lock-in modulation technique, the background is greatly reduced and electronic amplification can be considerably increased. Another approach to obtaining a bandpass instead of a high pass characteristic is to reaccelerate and selectively collect those particles which pass through the decelerating field at minimum kinetic energy. In other words, after passing the decelerating high pass filter, particles must subsequently pass through a low pass energy filter.

Spectrometers of this type [13–16] are available commercially for both AES and

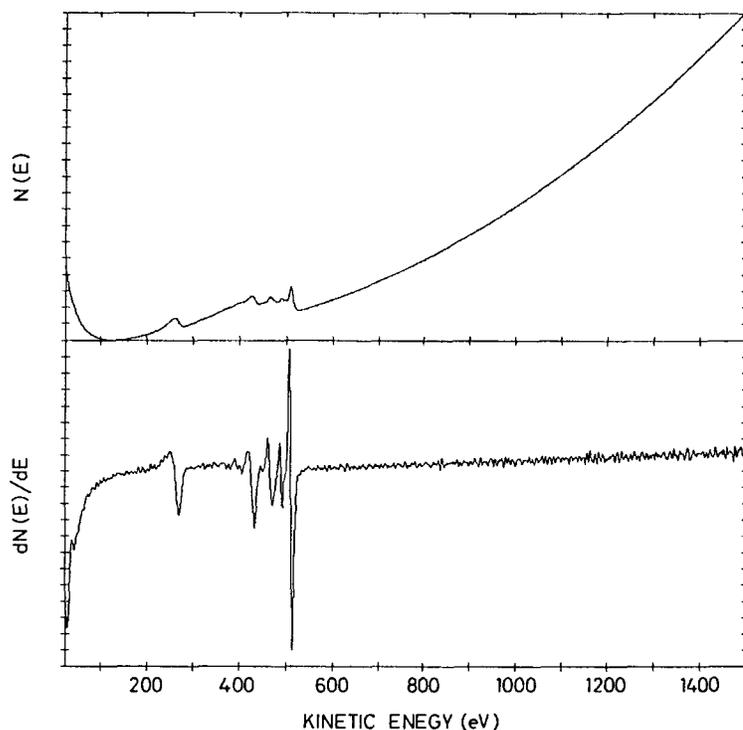


Figure 6. Auger spectrum of  $V_2O_5$ .

XPS and are characterized by a large solid acceptance angle  $\Omega$  and by a large étendue  $\lambda$ . The étendue  $\lambda$ , as proposed by Heddle [17], is defined as the product of accepted area  $S$  and of accepted solid angle  $\Omega$

$$\lambda = S\Omega \quad (11)$$

Due to the lack of good lateral resolution, these spectrometers can only be used for imaging purposes in the source defined mode, that is, with a well-defined primary excitation beam scanning the sample surface and yielding the required lateral resolution.

#### *Deflection Type Analyzers*

A second approach is by deflection in electric and/or magnetic fields. Historically, the oldest application of energy analyzers can be traced back to  $\beta$ -spectroscopy, in which different magnetic deflection schemes were developed. However, during recent decades, the pure electrostatic energy analyzer has become the more popular instrument for several reasons. Practical reasons, for instance, are to be found in the less bulky construction as compared with most magnetic instruments and also in the fact that pure electrostatic devices are better adapted for UHV environments as used in AES and XPS experiments.

There are also more fundamental reasons for the popularity of electrostatic analyzers. For example, it is much easier to define the boundary conditions in electrostatic systems as compared with magnetic layouts. Shielding magnetic fields is much more of a problem than shielding electrostatic fields. In pure electrostatic deflection devices, the trajectories of a particle are independent of mass and charge [18]. Mass and charge

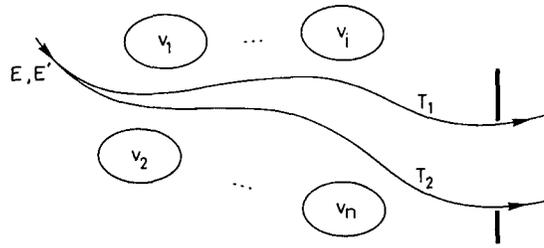
only intervene in the transit time of the particle along the followed ray path. Clearly, this property is of value in the energy analysis of ions having different masses and (multiple) charges; only the ratio of electrode potentials to the accelerating potential of the particles remains constant [18].

A strong impetus to the development of electrostatic energy analyzers has originated from ion beam research. Many different types of deflecting electrostatic energy analyzer have been developed and applied successfully. Without elaboration, the most important types are as follows:

- Parallel plate analyzers [19] and fountain analyzers [20]. Both are characterized by a uniform electric field between parallel electrodes.
- Cylindrical sector fields. These are best known through a special focusing configuration, namely, the  $127^\circ$  sector analyzer [21, 22], and cylindrical mirror analyzers (CMA) [23–29]. Both types are characterized by an inverse first power of  $r$  force field. The cylindrical mirror analyzer was first used in AES by Palmberg et al. [30] in 1969 and then became a standard tool in AES. It also has good transmission (i.e., ratio of emergent particle flux to incoming particle flux), of the order of 10% for moderate resolution (0.3%).
- Hemispherical analyzers [17, 31, 32]. These are characterized by an inverse second power of  $r$  force field.

#### *Time of Flight (TOF) Analyzers*

In contrast with ion spectroscopic techniques, time of flight (TOF) methods are not really used for energy analysis of



**Figure 7.** Trajectories  $T_1$  and  $T_2$  are the limiting trajectories through the output aperture.

electrons in AES or XPS. Due to the much smaller mass of an electron as compared with ions and taking into account the required relative energy resolution ( $\sim 0.1\%$ ), timing electronics are not readily available for a time of flight setup with realistic dimensions of the order of 1 m maximum. Furthermore, TOF techniques require fast pulsed excitation sources and, in any case, the required electronics are not readily available.

### 2.4.3.3 Energy Resolution of Deflecting Electrostatic Analyzers

Of course, an important figure of merit of an energy analyzer is the energy resolution although for deflecting electrostatic analyzers one is always confronted with the relative resolution,

$$R = \frac{\Delta E}{E} \quad (12)$$

The reason for this can be briefly explained as follows. Consider a system with  $n$  completely arbitrary shaped electrodes carrying potentials  $V_1, \dots, V_n$ . This potential  $V_i$  of the  $i$ -th electrode is measured with respect to a reference potential  $V_0$  and the measured potential differences are given by

$$v_i = V_i - V_0 \quad (13)$$

The reference potential is chosen such that at  $V_0$  the particles injected into the system under consideration are at rest, that is, at zero kinetic energy. This means that the kinetic energy is given by

$$E = q(V - V_0) = qv \quad (14)$$

if the particle is accelerated by the potential  $V$ . As already mentioned [18], the trajectories only depend on the ratio of electrode potentials to the accelerating potential. This means that in Fig. 7 trajectory  $T_1$  is characterized by

$$\frac{v}{v_i} = K_i \quad 1 \leq i \leq n \quad (15)$$

with  $v$  the reduced potential given by Eq. (14) and corresponding with a kinetic energy  $E$  of a particle entering the electrode system. Similarly, a particle with a kinetic energy  $E'$

$$E' = q(V' - V_0) \quad (16)$$

will follow a trajectory  $T_2$  through the electrode system and this trajectory is characterized by

$$\frac{v'}{v_i} = K'_i \quad 1 \leq i \leq n \quad (17)$$

From Eqs. (14) and (16) it follows that

$$\frac{v'}{v} = \frac{E'}{E} = K_0 \quad (18)$$

If  $\Delta E = E' - E$  it is obvious that  $R = \Delta E/E$  is also constant.

In other words, for given electrode potentials the trajectories  $T_1$  and  $T_2$  depend on the respective energies  $E$  and  $E'$ . If those trajectories are the limiting trajectories through an exit aperture as is the case for an energy spectrometer, the relative energy resolution  $R = \Delta E/E$  is a constant and depends on the system geometry.

Clearly, the higher the transmitted energy  $E$ , the higher the absolute energy resolution or energy spread  $\Delta E$ . An important consequence of the constant relative resolution,  $\Delta E/E$ , is that when measuring an energy distribution  $N(E)$  at the output of the analyzer it is not simply a measure of current, that is,

$$I(E) \sim N(E) \quad (19)$$

but, instead

$$I(E) \sim \Delta E \times N(E) \quad (20)$$

or

$$I(E) \sim E \times N(E) \quad (21)$$

To eliminate this problem, in other words to measure an energy distribution with constant  $\Delta E$ , it is necessary to shift the energy  $E$  of the incoming electrons, either by retarding or accelerating them in an appropriate electric field, by an amount  $E'$  such that the final energy is always at the fixed pass energy  $E_0$  of the analyzer, that is,

$$E - E' = E_0 \quad (22)$$

As the pass energy of the analyzer is fixed so is the absolute energy spread  $\Delta E$  during the measurement. As far as the electrostatic energy dispersive system is concerned, the measured current at the output is simply proportional to  $N(E)$ ; however, the

retardation also influences the output current of the complete spectrometer. Again, the background can be eliminated by measuring the first derivative of  $N(E)$  with respect to  $E$ . By modulating the deflecting voltages of the analyzer, this can be achieved by synchronously demodulating the output signal and using the first harmonics content. As in the nonretarding mode, the output current is, according to Eq. (21), proportional to the product  $E \times N(E)$ , the first derivative now becomes

$$\frac{dI}{dE} \propto E \frac{dN}{dE} + N(E) \quad (23)$$

To a good approximation, this reduces to

$$\frac{dI}{dE} \propto E \frac{dN}{dE} \quad (24)$$

which again is obtained by measuring the first harmonic with lock-in detection techniques. The sensitivity of these detecting techniques was greatly enhanced by using electron multipliers. In these analogue detection layouts, use is normally made of high current type electron multipliers. In contemporary instrumentation, analogue detection is frequently supplanted by fast pulse counting techniques in which each individual electron hitting the electron multiplier is detected and counted. This results in much better signal-to-noise ratios as compared with the described analogue techniques. As will be clear, pulse counting yields directly either  $N(E)$  or  $E \times N(E)$  as a function of  $E$ , depending on whether the retarding or nonretarding mode was used. Background subtraction or other data treatments are performed digitally by using the appropriate software.

### 2.4.3.4 Cylindrical Mirror Analyzer (CMA) versus the Concentric Hemispherical Analyzer (CHA)

From a historical point of view in AES spectroscopy, the cylindrical mirror analyzer (CMA) (Fig. 8) has nearly always been used as the energy dispersive element

whereas for XPS experiments a concentric hemispherical analyzer (CHA) (Fig. 9) has been the choice. A reason for this different approach is to be found in the fact that AES was always considered as an imaging or microscopic technique in which the sample area under investigation was excited by a well focused primary beam scanned over the sample surface. It is for this reason that this approach is called

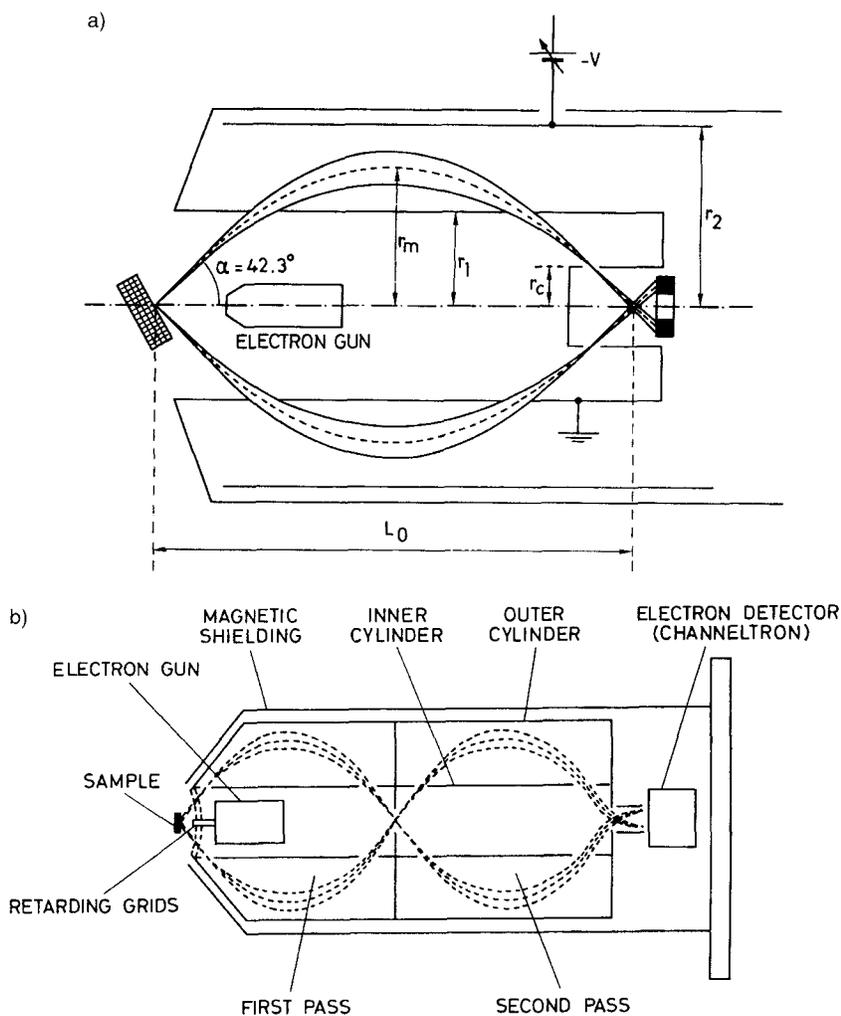
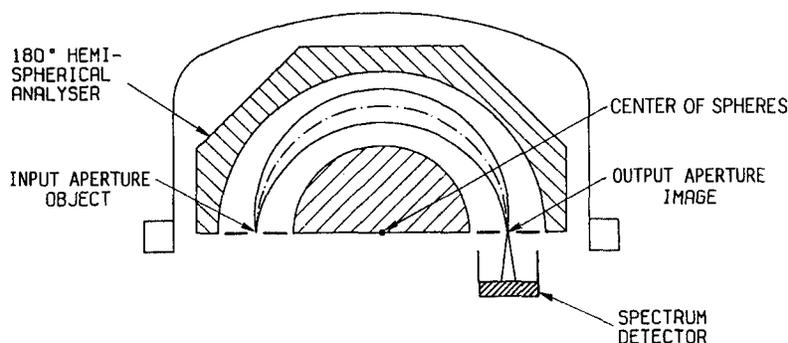


Figure 8. (a) Single-pass CMA; (b) double-pass CMA.



**Figure 9.** Concentric hemispherical analyzer. Object, image and center of the concentric spheres are on a straight line. This is a general property and not only holds for the 180° device. Adapted from E. Adem, R. Champaneria and P. Coxon, VG Scientific Limited, *Vacuum* **1990**, *41*, 1695 (Fig. 1).

scanning Auger microscopy (SAM). The smaller the excited area, the better the lateral resolution [33–36]. To collect as many emitted Auger electrons as possible from this small excited area, the important parameter is the solid angle  $\Omega$  under which the electrons are emitted. Under these conditions and for comparable physical dimensions, the CMA was the better choice compared with CHA. However, in XPS experiments in which the primary excitation is due to X-rays, which are much more difficult to focus, the excited sample area is much larger as compared to AES. The electrons to be energy analyzed are no longer collected from a point source, but from a much extended source. This means that no longer is  $\Omega$  the important parameter but instead the étendue  $\lambda = \Omega S$ , that is, the product of solid angle of acceptance and area seen by the energy dispersive system. Under these operating conditions a CHA has better performances than the CMA.

#### *The Focusing Properties of the CMA and CHA*

The CMA, as shown in Fig. 8(a) and consisting of two concentric cylinders,

has become very popular since Zashkvara and coworkers [37] showed that it has unexpected second order focusing properties for  $\alpha = 42.3^\circ$  and that refocusing of an object situated at A occurs at B for

$$L_0 = 6.1 r_1 \quad (25)$$

if

$$\frac{E_0}{qV} = \frac{1.3}{\ln(r_2/r_1)} \quad (26)$$

where  $E_0$  is the energy of the transmitted electrons, and  $V$  the potential difference between the two concentric cylinders with radii  $r_1$  and  $r_2$ .

Second order focusing means that in the Taylor expansion of the function

$$\Delta L = f(\Delta E, \Delta\alpha) \quad (27)$$

with  $\Delta L = L - L_0$ ,  $\Delta E = E - E_0$ ,  $\Delta\alpha = \alpha - \alpha_0$ , neither first nor second order terms in  $\Delta\alpha$  appear. For a CMA with  $\alpha_0 = 42.3^\circ$

$$\frac{\Delta L}{r_1} = 5.6 \left( \frac{\Delta E}{E_0} \right) - 15.4 (\Delta\alpha)^3 + \dots \quad (28)$$

which is a figure of merit of the focusing properties of the analyzer as a function of energy resolution  $\Delta E/E_0$  and half opening angle  $\Delta\alpha$ . The influence of  $\Delta\alpha$  on the

focusing properties is also called spherical aberration.

In their comparison of the CHA and CMA devices, Hafner et al. [38] showed that a similar Taylor expansion for the  $180^\circ$  CHA has a second order term in  $(\Delta\alpha)^2$  but no term in  $(\Delta\alpha)^3$ . Despite the absence of  $(\Delta\alpha)^3$  CHA only has first order focusing. Due to the different order of focusing, a comparison of the performance between the CMA and CHA as a function of  $\Delta\alpha$  using the 'reduced dispersion', which is the ratio between the numerical values of  $\Delta E/E_0$  to the first significant aberration term in  $\Delta\alpha$ , is impossible. For this reason, Hafner used the 'reduced aberration', being the ratio between the trace width (a measure of the spatial extension of the beam) and the relative resolution or dispersion, as a figure of merit.

It turns out that despite the second order focusing (not to be confused with double focusing which is used for instance in mass spectrometers of the Mattauch-Herzog type where ions are focused to a single point independently of the angle under which they are injected (geometrical focusing) and irrespective of their energy (energy focusing)) the CMA and CHA have comparable performances. However, as the CMA has a third order aberration term it is known from optics [39] that there is a region different from the image plane where the trace width is at a minimum. For a CMA, this region lies ahead of the image, which is found on the symmetry axis of the instrument, and takes for reasons of symmetry the form of a cylinder, with radius  $r_c$  (Fig. 8a), concentric with the deflecting cylinders. If the output aperture coincides with this fictitious cylinder, clearly the energy resolution is greatly enhanced as a smaller aperture can be used, while retain-

ing the same transmission. Under these operating conditions a CMA performs clearly better than a CHA when using the reduced aberration as a criterion. Because of its excellent performance, the CMA has been very popular as an energy dispersive device for both AES and XPS. Today, it is very successful in AES, wherein a very compact instrument can be built if the electron gun is fitted into the inner cylinder of the CMA.

#### *The CMA as used in XPS*

For XPS purposes, where traditionally the absolute energy spread  $\Delta E$  is of prime importance, retardation of incoming electrons to a fixed low pass energy  $E_0$  of the analyzer has become standard practice. The retardation is mostly obtained between hemispherical grids centered around the source area. However, grids suffer from several drawbacks such as scattering of the electrons, generation of secondary electrons strange to the experiment and responsible for extra background noise, and degradation of the angular properties of the beam due to lens action of each mesh of the grid. As an alternative to using grids, other solutions have been proposed such as special input lenses in front of the CMA [40]. However, because the accepted beam in a CMA is confined between two conical surfaces, the design of a retarding lens is not straightforward and, to our knowledge, this idea is not used in commercial equipment.

A major drawback of the CMA is its sensitivity to sample position, both in the axial direction and in the direction perpendicular to the axis. A misalignment of the sample would give a reduction in

transmission, a change in pass energy and a worsening of the resolution. These effects can be greatly reduced by using either a double pass CMA (Fig. 8b) or a single CMA combined downstream with a small  $90^\circ$  hemispherical analyzer. The output aperture of the first CMA acts as the input aperture of the second energy dispersive device and the pass energy of this energy selector no longer depends on the sample position. The double pass [41] scheme also enables the signal-to-noise ratio to be enhanced at the detector while retaining maximum luminosity  $L$  (luminosity is the product of the étendue  $\lambda$  with the transmission  $T$ , the latter being the ratio of emergent particle flux to incoming particle flux). This approach also facilitates the detection of the energy-filtered electrons because a simple electron multiplier or channeltron can be used instead of the more complex annular detector often recommended in a single pass CMA.

#### *Parallel Detection*

When using energy dispersive devices of the focusing type, such as the CMA or CHA, a real image is formed of the input aperture at the exit of the analyzer, which is spatially shifted in the energy dispersive direction depending on the energy under consideration. An energy analysis of the incoming particles can be obtained in the following two ways. Either one uses an exit aperture, which defines the resolution of the analyzer, and onto which the entrance aperture is imaged for one particular energy. By changing the field between the deflecting plates, the pass energy or the energy of the particles passing through the exit aperture, and which hit a detector, can be varied. By sweeping this pass

energy, the energy distribution of the incoming particle flux is probed sequentially. Or, one can eliminate the exit aperture and replace it by several detectors positioned in the dispersive direction of the analyzer. In the same manner as an exit slit, the physical dimension of each detector defines an energy spread  $\Delta E$  and, by using several detectors simultaneously, a region of the energy spectrum can be measured in parallel. Of course, the total energy region probed simultaneously by the different detectors is rather restricted and depends essentially on the physical dimensions of the analyzer and the mean energy transmitted (mean energy has the same meaning as pass energy in an analyzer with an exit aperture). Nevertheless, it is clear that under critical circumstances (i.e., bad signal-to-noise ratios) the principle of parallel detection alleviates the noise problem and reduces the measuring time inversely according to the number of detectors.

In modern equipment for AES and XPS the concept of parallel detection is fully implemented for many applications. As in a CHA, the energy dispersive direction coincides with the produced direction of the input aperture and the center of the deflecting spheres. There is also good accessibility for mounting several detectors. In a CMA, however, where the energy dispersive direction coincides with the symmetry axis of the device, parallel detection is not as easily implemented as in a CHA.

#### *Combination CHA-Input or -Transfer Lens*

As already outlined, CMA based systems have traditionally been used for small spot work such as with AES, whereas CHA

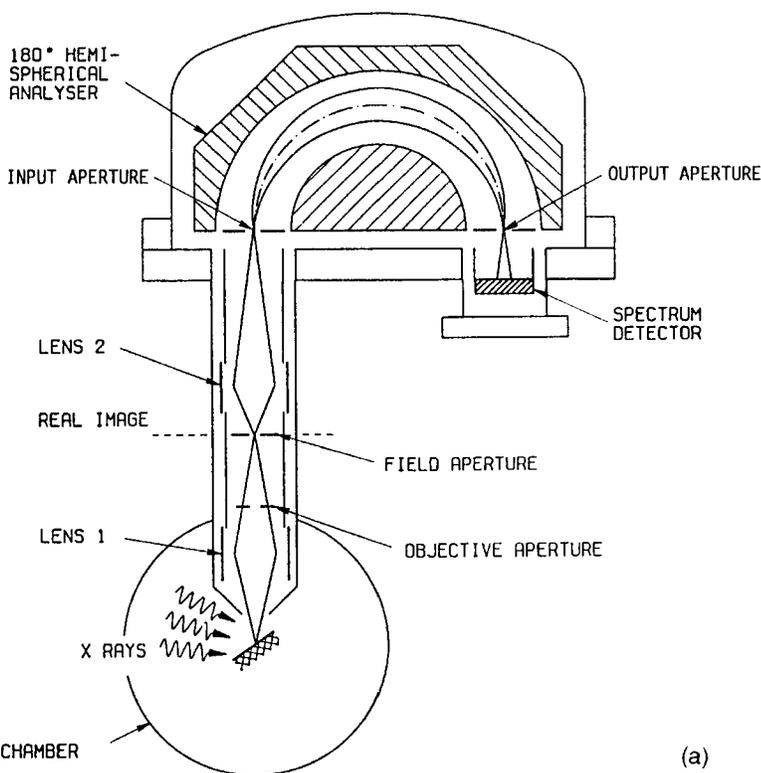
systems are more appropriate when dealing with extended sources as was normally the custom in XPS. However, CHA systems combined with input or transfer lenses are seriously challenging the best performing CMA systems, even for small spot work. The flexibility of the combination of CHA with transfer lens even allows easy switching between small spot and extended area analysis. To our knowledge, the first report in this respect was published by Nöller and coworkers [42] in which they compared a CMA with a CHA combined with an input retarding lens; both arrangements were used in high resolution AES. As shown by Heddle [17], the étendue of all deflection-type analyzers is increased when used in a retarding mode. According to various studies [43,44] this is due to the fact that by retarding the electrons before entering the analyzer, larger input and output apertures can be used for the same overall resolution of the whole spectrometer. However, as a consequence of the law of Helmholtz–Lagrange, this increase in étendue is partly counteracted by the action of the retarding field itself [17]. The overall result of retardation is a net improvement in étendue, and consequently an improvement in sensitivity which is better for a CHA than for a CMA. Due to the retarding field, it can be shown that the current at the detector in such a system is proportional to  $N(E)/E$ .

The fact that the CHA is easily combined with an input lens system, is related to the narrow pencil-like beam shape accepted by the CHA, a beam shape which is also easily transmitted through relatively simple cylindrical symmetric lenses without introducing too much aberration. By contrast, in a CMA system the

accepted beam has a conical profile with large opening angles, prohibiting the use of simple transfer lenses. Finally, when lateral resolution becomes important, as is the case in ‘detector defined’ imaging, the combination of CHA with transfer lens is certainly the best solution. A consequence of the flexibility of the CHA–lens combination with respect to SAM and XPS is that a new generation of combined SAM–XPS apparatus is now commercially available.

#### *The Input or Transfer Lens*

The input or transfer lens in front of the CHA has always been of the electrostatic type despite magnetic lenses being far superior if spherical aberrations are considered. As already outlined several reasons are responsible for this. On the one hand, there are dimensional constraints: as magnetic lenses are usually much bulkier the choice is in favor of simple and compact electrostatic lenses. On the other hand, to fully exploit the low aberration characteristic together with their large acceptance or opening angle, magnetic lenses should be placed as close as possible to the sample. This, of course, puts a severe constraint on the transport of other charged particles in the vicinity of the sample; for instance, with the sputtering ion beams or flood gun electrons used to avoid charge buildup on insulating samples. It is not possible to use magnetic fields in the transfer lens of the CHA with an AES setup because it is impossible to generate a primary beam parallel with the magnetic induction. Nevertheless, as we will see later, in XPS, when the primary excitation is due to X-ray photons, it is possible and advantageous to use some



**Figure 10.** (a) Combination of transfer lens and CHA defining the field of view and CHAMBER the acceptance angle.

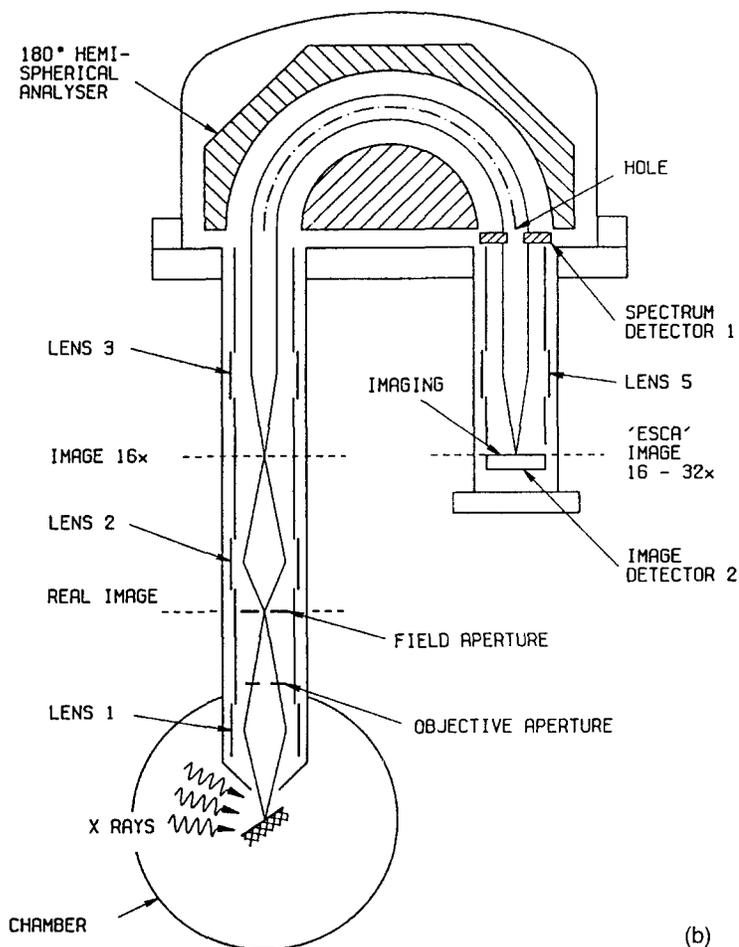
kind of magnetic lens close to the sample with the  $\mathbf{B}$  vector pointing towards the transfer lens.

The situation becomes even more complex during XPS experiments when flooding electrons are required. Simple straight-on flood guns cannot be used as they illuminate the sample under a direction different from the axis of the transfer lens or  $\mathbf{B}$  vector and special layouts, coaxial with the transfer lens or  $\mathbf{B}$  vector, are needed.

Generally, the purpose of the electrostatic transfer lens is threefold [45]. Namely, (i) it defines the field of view, (ii) it defines the acceptance angle, and (iii) it (possibly) retards the incoming electrons in the CHA. Figure 10a shows an input or

transfer lens in front of the CHA which, as can be seen, comprises two separate lenses. The sample surface is situated in the object plane of lens 1 and is imaged on the field aperture situated in the image plane of the same lens: that is, the sample and field aperture are conjugate. The area seen on the sample or the field of view is defined by the aperture area of the field aperture and by the magnification of the lens, depending on the potentials applied to the lens elements. The field of view area is simply the ratio of the area of the field aperture to the lens magnification.

The objective aperture, preceding of the field aperture, defines the acceptance angle of the incoming beam. The real image sharpness (or spatial resolution) at



**Figure 10** (b) Similar setup as in (a), but with additional Fourier lenses 3 and 5. Adapted from E. Adem, R. Champaneria and P. Coxon, VG Scientific Limited, *Vacuum* **1990**, *41*, 1695. Reproduced by permission of P. Coxon et al. and Pergamon Press.

the field aperture, which depends on lens aberrations, is enhanced by reducing the area of the objective aperture, just as in classical optics. However, the better the sharpness, the smaller the intensity or the number of electrons emitted from the field of view passing through the lens.

The second lens forms an image of the field aperture on the input aperture of the CEO. This lens provides extra magnification and also performs the decelerating function (without the use of grids) if the CEO is used in constant pass energy

mode. The degree of lens filling (i.e., the radial trajectory distribution) depends on the amount of deceleration which means that the spherical aberration of this lens also depends on the degree of retardation. The complete behavior of such a lens is best understood by trajectory simulations [46]. To reiterate: the deceleration just before entering the CEO vastly enhances the performance of the CEO combined with a transfer lens and thus makes this combination particularly attractive for XPS.

### 2.4.3.5 Imaging Techniques

As already discussed, there are two distinct methods used in AES/XPS imaging: the source defined (SD) approach and the detector defined (DD) approach. In the SD mode, lateral resolution is obtained by good focusing of the primary excitation source. In AES, focusing presents no problem since it originates from the field of scanning electron microscopy.

#### *Imaging AES or Scanning Auger Microscopy*

A scanning Auger microscope (SAM) is, in principle, a combination of a scanning electron microscope (SEM) and an Auger electron energy analyzer (CMA or CHA) [47]. Indeed, in a SEM a focused beam of relatively high energy (25–65 keV) electrons not only produces the backscattered electrons (BSE) and the true secondary electrons (SE) with which SEM images are formed, but also Auger electrons at the point of impact. Taking into account the limited escape depth of Auger electrons, only the uppermost layers contribute to the signal, although the penetration depth of the primary beam may be of the order of 1  $\mu\text{m}$ .

When an energy analyzer is introduced in such an instrument and calibrated on a fixed energy corresponding to a particular Auger transition, a composition image of the surface layer can be formed with a lateral resolution equal to or even better than the resolution of the SEM image. Indeed, the resolution of the SEM image is somewhat lower than the beam diameter as a result of backscattering electrons emerging from the specimen relatively far from the point of impact. It has been

shown, however, that Auger electrons emerge practically exclusively from the point of impact; therefore, the resolution is in principle given by the beam diameter [48]. In the commercially available dedicated SAM instruments the primary energy of the electron beam is limited to 10 keV to increase Auger electron production. Indeed, since the ionization cross section for Auger electron emission decreases when using higher primary energies, the lateral resolution is lowered by a considerable amount, in most instruments by around 100 nm.

The deflection plates of a conventional SAM are added to the electron gun optics to be able to raster the surface of the sample. In general, lateral resolutions sufficient to study problems in metallurgy or microelectronics are readily obtained and spatial distributions of elements in a surface can be studied. In principle, it is also possible to map a component of a composite Auger line when the element under examination is present in different chemical environments, as discussed above. The SAM has found widespread use in solving industrial problems on samples which are microscopically nonhomogeneous and sometimes exhibiting rather rough surfaces. Quantification of the surface concentrations on such samples has to be undertaken with great care, as was described by Prutton et al. from the York group [49]. We will return to these problems in more detail in what follows.

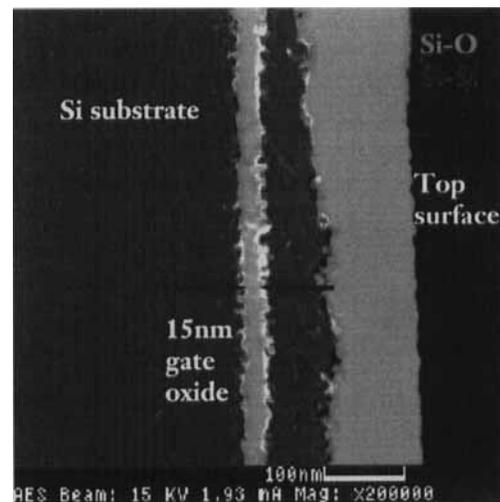
In most commercial SAM instruments a secondary electron detector (SED) is included to produce SEM images of the surface under study. Furthermore, absorption current images can also be obtained. Both are very useful in selecting the surface area which has to be studied with the SAM.

The higher Auger production efficiency, when working at lower primary energies, allows a reduction of the primary beam current ( $\sim 10$  nA) and consequently limits the beam effects. Additional to two-dimensional imaging of the surface chemical composition, called element mapping, line scans can also be used in combination with SEM images, producing SEM images combined with elemental composition lateral profiles.

Mapping in scanning Auger microscopy is a time-consuming procedure. A reduction in measuring time can be obtained by using electron emitting sources with a brightness as high as possible. In this respect the introduction of LaB<sub>6</sub> emitters has proven to be a leap forward in brightness as compared with the classic thermionic emitters. However, LaB<sub>6</sub> sources suffer from two drawbacks. First, their brightness, in other words the amount of current extracted from the filament tip per unit area and solid angle, decreases drastically when, for instance, hydrocarbons which can react with the LaB<sub>6</sub> filament are in the vicinity. In this respect, several oxygen treatments have been proposed which enhance the performance of the filament although there remains a degradation in brightness and in emitting stability. Second, the emitting tip of the LaB<sub>6</sub> filament is rather sensitive to ion backspattering. As soon as the pressure in the neighborhood of the filament rises above  $10^{-6}$  mbar there is sufficient restgas ionization to produce positive ions which are accelerated towards the filament and sputter the filament. This sputtering degrades the geometry of the electron emitting tip, decreasing the geometrical quality of the electron source.

The stability problems of LaB<sub>6</sub> emitters are largely shelved by using a new generation of electron emitters, namely, Schottky field emitters. These consist of a single crystal tungsten tip coated with a zirconia layer to lower the work function of the tip. Typical dimensions of the electron emitting surface are of the order of 250 nm. Compared with LaB<sub>6</sub> emitters, Schottky field emitters have higher brightness, lower energy spreads and longer lifetimes. However, due to the needle shape of the emitter, backspattering must again be eliminated completely requiring a dedicated pumping system in the vicinity of the Schottky emitter.

Figure 11 shows an example of an SAM image recorded using a Schottky electron emitter. The image was recorded using the Fisons Microlab 310-F field emission scanning Auger microprobe. It shows a 15 nm wide SiO<sub>2</sub> gate oxide layer, 200 nm below



**Figure 11.** A 15-nm wide SiO<sub>2</sub> gate oxide layer, 200 nm below the top surface of an Si wafer, as revealed using the Fisons Microlab SAM on a fractured edge of the wafer. Reproduced by permission of Fisons Instruments.

the top surface of an Si wafer. The image was obtained by recording the Si–Auger signal characteristic, respectively, for the bulk material and for the oxide.

### *Imaging XPS*

In scanning Auger microscopy, at electron energies of  $\sim 2$  keV, resolutions of the order of 50 nm are attainable with beam currents of the order of  $10^{-9}$  A. This corresponds to current densities of  $40 \text{ A cm}^{-2}$  and pinpoints one of the weaknesses of AES or SAM, namely sample damage. In XPS, the damage due to X-rays is much lower; however, it is much more difficult to focus X-rays.

*Generation of X-rays:* X-rays are classically generated by energetic electrons impinging on an anode or anticathode. The type of anticathode material defines the X-ray spectrum. Besides the characteristic lines, background bremsstrahlung is also generated although this can be reduced considerably by using appropriate filtering window materials. Nevertheless, bremsstrahlung levels of as high as 20% of the principal characteristic emission lines remain.

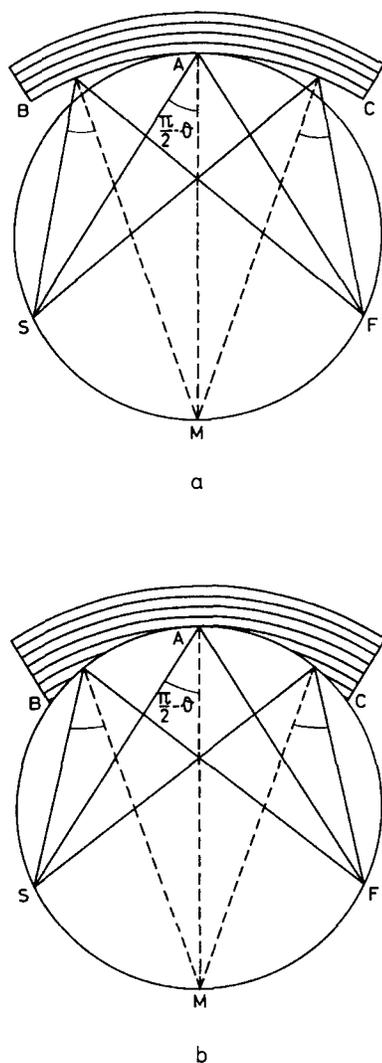
With the simple cathode–anticathode arrangement the emerging X-ray beam is usually very broad and so spatial resolution can only be obtained by collimating the beam. This is done at the expense of intensity. Besides intensity, there remain several other drawbacks and limitations for XPS work using a simple X-ray source. For high energy resolution XPS, there is often a lack of X-ray line sharpness due to satellite and ghost lines [50] which can complicate the measured XPS spectrum. Also, there is the remaining continuum

radiation which, on the one hand, can damage sensitive samples while, on the other hand, generating extra noise in the XPS detection equipment. A complete survey of these effects is given by Chaney [51].

*Focusing and monochromatizing the X-rays:* The benefits of using a crystal monochromator to eliminate the aforementioned problems has been discussed by Siegbahn [52] and coworkers. If a crystal lattice is bent as shown in Fig. 12a, with a radius equal to the diameter of the Rowland circle, X-rays originating from a point source on the Rowland circle can be diffracted and monochromatized from an extended area of the bent crystal, thus collecting X-rays in an opening angle proportional to the diffracting area. The Bragg relation,

$$n\lambda = 2d \sin \theta \quad (29)$$

with usual notation, is satisfied along the whole diffracting area. As the X-rays are also reflected under specular conditions, refocusing of the monochromatized X-rays also occurs at a point on the Rowland circle. Rowland first recognized the imaging properties of this setup for a spherical grating and it was Johann who first implemented these ideas in a practical arrangement for X-rays by bending a crystal lattice. As is clear from Fig. 12a, some focusing errors remain. These were corrected by Johansson in the layout shown in Fig. 12b. Johansson not only bent the crystal, but ground it to the correct shape. The practical implementation of this approach is named after Johansson or Guinier. Focusing X-ray monochromators such as these became a standard tool in XPS and are used in UHV



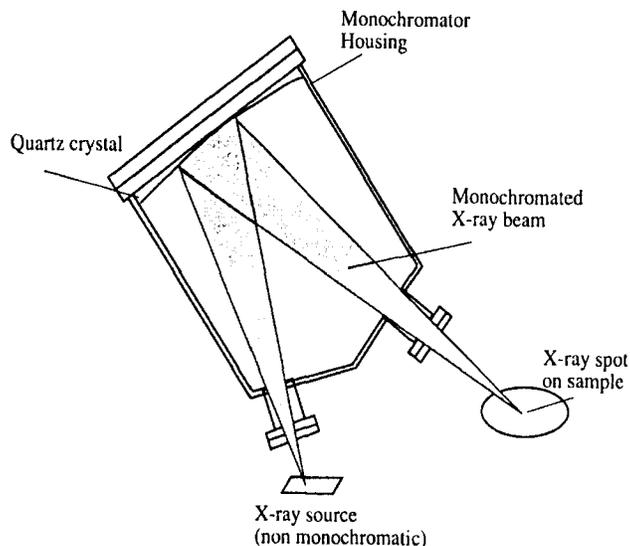
**Figure 12.** (a) The Johann solution: the bent crystal has a radius equal to the diameter of the Rowland circle; (b) the Johansson or Guinier solution.

apparatus as shown in Fig. 13. In practice, several bent quartz crystals are used to monochromatize and refocus as much X-ray flux as possible. To reduce astigmatism, the crystals are bent toroidally, not spherically. With this equipment, it is possible to obtain illuminated spot sizes of the order of 50–100  $\mu\text{m}$ .

*'Source defined' imaging in XPS:* With these dimensions as lateral resolution, some kind of source defined (SD) imaging work already becomes possible and the trivial answer to providing a sequential pixeling is performed by moving the sample stage with a precision X–Y manipulator, storing the XPS energy information pixel by pixel. This crude approach is, of course, extremely time consuming and certainly not practical for extensive use.

A promising approach to SD XPS imaging was given early on by Hovland [53]. This author coated the back of his thin samples with an aluminum layer which fulfilled the function of anticathode. A scanning electron beam impinging on the back of his sample, on the aluminum layer, generated a scanning X-ray beam, exciting the sample and producing the photoelectrons. However, due to the need for elaborate sample preparation and, in any case, unsatisfactory results, this idea was abandoned.

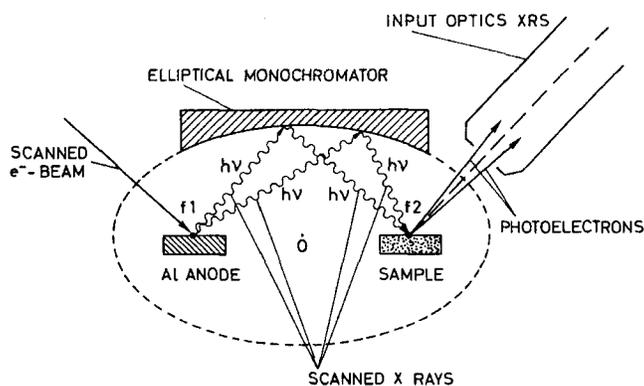
A breakthrough in the pixeling process has been made by Physical Electronics Inc. using elliptically bent crystals. The principle underlying this approach is shown in Fig. 14. At the foci of the ellipse, which defines the bent crystals lattice surface, are the anticathode and sample. By scanning the X-ray generating electron beam over the anticathode, this anticathode is sequentially pixelated. Each pixel emits X-rays which are monochromatized and reflected on the sample surface by the elliptically bent lattice. In other words, each electron beam defined pixel at the anticathode is X-ray imaged at the sample and by scanning the electron beam over the anticathode, a monochromatized and focused X-ray beam scanning the sample is obtained. This provides an elegant



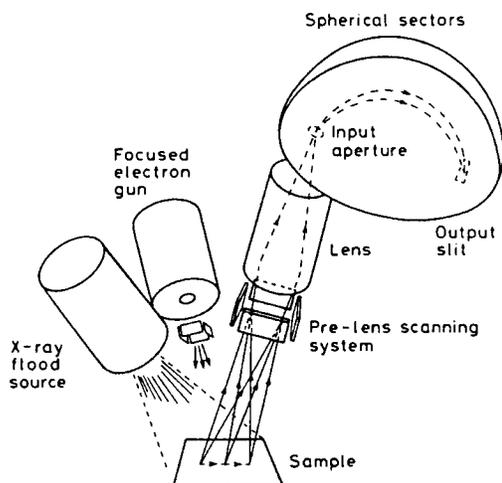
**Figure 13.** Practical focusing X-ray monochromator of the Johansson or Guinier configuration, ultrahigh vacuum adapted for XPS work. Adapted from Kratos Analytical, Axis product brief. Reproduced by permission of Kratos Analytical Ltd.

solution to SD XPS imaging. In principle the lateral dimensions of the X-ray spot are of the same order of magnitude as the dimensions of the electron beam on the anticathode. For XPS imaging, this leads to lateral resolutions of the order of  $10\ \mu\text{m}$ . Besides imaging, this system enables point mode measurements to be made (i.e., enables more standard XPS measurements to be made by keeping the small spot X-ray beam fixed on a particular position on the sample).

*'Detector defined' imaging in XPS:* This aforementioned and elegant solution to source defined XPS imaging strongly contrasts with the early attempts of detector defined (DD) XPS imaging in which an extended area of the sample is illuminated with X-rays and, by preference, also monochromatized. In the first attempts, Keast and Dowing [54] simply inserted a collimator between the sample and their energy dispersive system to achieve some degree of lateral resolution and again, by



**Figure 14.** A scanning electron beam on the Al-anticathode is transformed, with the aid of an elliptically bent crystal, into a monochromatized scanning X-ray beam focused on the sample. Adapted from Physical Electronics Quantum 2000 documentation. Reproduced by permission of Physical Electronics Inc.

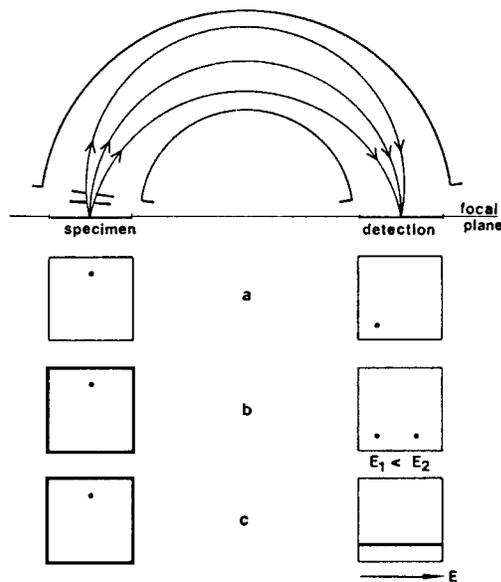


**Figure 15.** In the detection defined XPS imaging, pixeling can be obtained by the use of a pre-lens scanning system. Adapted from M. P. Seah and G. C. Smith, *Surf. Interface Anal.* **1988**, *11*, 69 (Fig. 3). Reproduced by permission of M. P. Seah and John Wiley & Sons Ltd.

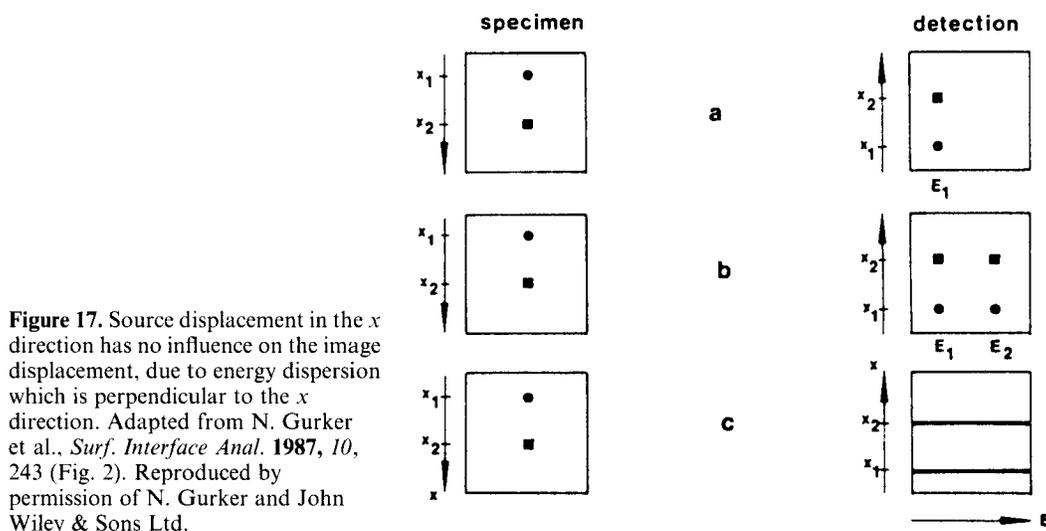
moving the sample stage, a time-consuming and trivial solution to detection defined XPS imaging can be attained.

A second logical step in DD XPS imaging has already been mentioned. Namely, this approach entails pixeling by using scanning deflection plates (Fig. 15) in front of the transfer lens of the CHA [55]. The transfer lens defines the field of view or the lateral resolution, whereas the scan plates provide for the pixeling by sequentially collecting the emitted electrons from the different virtual pixels. In this approach, only those electrons emitted from the pixel addressed by the scan plates are detected, while electrons emitted from other areas of the X-ray illuminated sample are not accepted by the combination of transfer lens and CHA. This is clearly a waste of information and translates itself into longer acquisition times for building up an image of sufficient quality.

To improve the detection efficiency in the DD imaging mode, Gurker et al. [56, 57] exploited the imaging properties of the CHA in combination with a two-dimensional position sensitive detector. As previously discussed, a CHA is a first order imaging device in two orthogonal directions. As a consequence, a point source at the entrance plane is imaged in the output plane as a spot with the same dimensions as the source. The object plane (input) and image plane (output) are conjugate. This is illustrated in Fig. 16 in which electrons are emitted from the same spot on the sample under three different energy conditions. When all electrons have the same energy, and if the pass energy of



**Figure 16.** Imaging properties of the CHA as a function of energy. (a) All electrons have the same energy; (b) electrons with two different energies  $E_1$  and  $E_2$  are emitted from the same spot; (c) electrons with a continuous energy distribution are emitted from the same source spot. Adapted from N. Gurker et al., *Surf. Interface Anal.* **1987**, *10*, 243 (Fig. 1). Reproduced by permission of N. Gurker and John Wiley & Sons Ltd.



**Figure 17.** Source displacement in the  $x$  direction has no influence on the image displacement, due to energy dispersion which is perpendicular to the  $x$  direction. Adapted from N. Gurker et al., *Surf. Interface Anal.* **1987**, *10*, 243 (Fig. 2). Reproduced by permission of N. Gurker and John Wiley & Sons Ltd.

the CHA is tuned close enough to that energy, these electrons are imaged as a single spot in the detection plane as is shown in situation (a). If, however, electrons with two different energies, situation (b), are emitted from the same spot, they will refocus in the detection plane at two different spot positions, depending on their respective energies. More generally, if electrons with a continuous energy distribution are emitted from the sample spot, they will refocus following a straight line in the detection plane. As shown in situation (c), this line is parallel with the energy dispersive direction of the CHA. Figure 17 shows the behavior of the system if electrons are emitted from two different spot positions. The spot positions on the sample are chosen such that they follow the  $x$  direction, that is, perpendicular to the energy dispersive directions of the CHA. In this way, the imaging function of the CHA is not convoluted with the energy dispersing function of the CHA and both kinds of information remain simultaneously accessible on a

two-dimensional position sensitive detector. In other words, by using a slit as aperture at the entrance, perpendicular to the energy dispersive direction of the CHA, a strip-like 'field of view' at the sample is projected on the detector again as a strip-like image. The one-dimensional spatial information is displayed following a direction perpendicular to the energy dispersive direction while the energy information is accessible in the parallel direction. Without yielding true two-dimensional imaging, this setup fully exploits the possibilities of the CHA combined with a two-dimensional position sensitive detector, because of the parallel detection scheme. Complete two-dimensional imaging can be achieved by moving the sample stage in the energy dispersive direction and taking successive linescans to produce a true XPS image at a fixed energy. A similar approach has been adopted by Allison [58] and Scienta Instrument A.B.

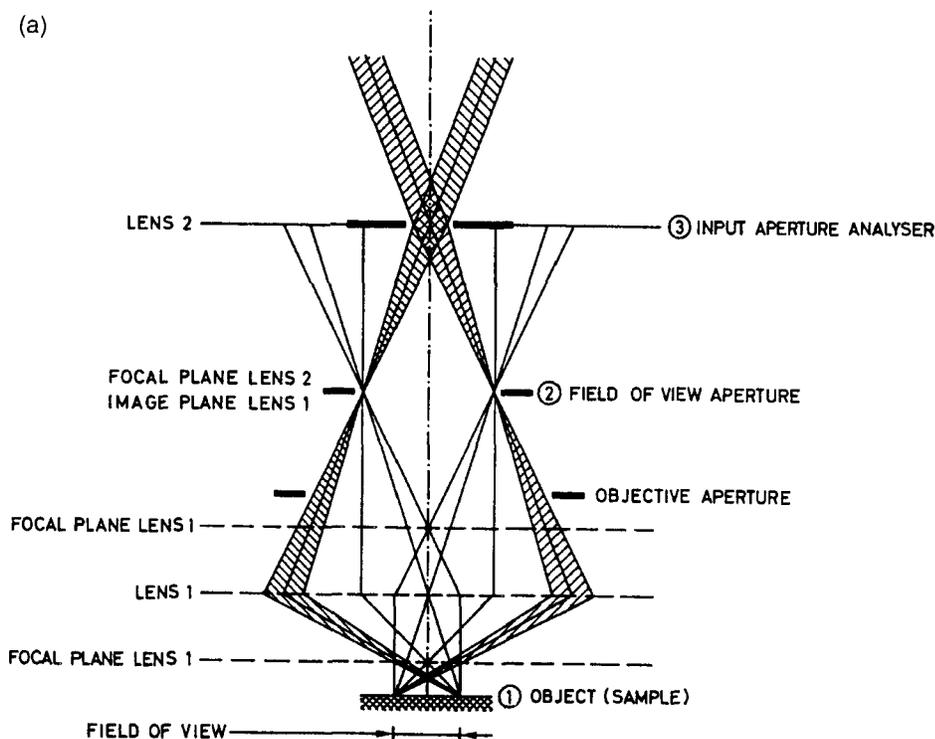
Recently, VG Scientific Ltd have pioneered a breakthrough on the side of DD

XPS imaging by fully exploiting the possibilities and potentialities of their transfer lens system combined with an extra lens at the exit of the CHA [59, 60]. Figure 10b gives a schematic picture of this lens combination. Compared with Fig. 10a, the only difference with the previously described transfer lens-CHA system is to be found in the extra 'Fourier' lenses 3 and 5. The Fourier lens gets its name from the fact that the real image of the sample produced by lenses 1 and 2 is situated in the focal plane of the Fourier lens 3. Similarly, the image detector, behind of the complete electron optical system, is in the focal plane of lens 5. As is well known from optics, a point source situated in the focal plane of a lens produces a parallel beam of light downstream of the lens and vice versa; exactly the same happens with the Fourier lenses of Fig. 10b. Electrons emitted from a single point in the image plane of lens 2, which coincides with the focal plane of Fourier lens 3, give rise to a parallel beam of electrons through the CHA. Similarly, a parallel beam of electrons entering Fourier lens 5 are refocused to a single point at the focal plane of lens 5 on to the image detector.

To understand how imaging is combined with energy selection we refer to Fig. 18a. For simplicity, the transfer lens consists of two idealized thin lenses: that is, a first lens combining the function of lens 1 and lens 2 in Fig. 10b, magnifying, defining the field of view, and the degree of lens filling or acceptance; and, a second lens acting as a Fourier lens. In other words, the field of view aperture of the first lens is the real object of Fourier lens 2 and is situated at the focal plane of this Fourier lens. As previously described, the

field of view aperture in the image plane of lens 1 defines the field of view on the sample. Narrowing the field of view aperture reduces the field of view on the sample. From each edge of this field of view, all electrons are of course refocused at the edge of the field of view aperture but only those electrons within the hatched beam of Fig. 18a can be transmitted through aperture 3 of the CHA as a parallel beam of electrons. Narrowing this input aperture 3 of the analyzer does not change the field of view but only narrows the opening angle of the hatched beam at the sample, again at the expense of intensity. The same reasoning holds for each point of the sample within the field of view: in other words, each point emits electrons in an opening angle defined by aperture 3 and produces a parallel beam of electrons into the CHA under a well-defined angle between the limiting angles of Fig. 18a corresponding with the edges of the field of view. Two important things are to be remembered: first, a specific point within the field of view is translated into a specific beam direction into the CHA; second, by reducing the size of aperture 3 no image information is lost, only intensity!

Figure 18b gives a simple overview of this transfer lens behavior. The double ended arrow on the sample defines the field of view and is transformed or magnified into a real image at the aperture field of view. Each point of this image is transformed into a particular beam direction in the CHA. As a  $180^\circ$  CHA is a 1 : 1 imaging device, all beam directions at the input aperture are transformed into similar beam directions at the output aperture. The input aperture of the CHA is correctly imaged on to the output aperture only for those electron energies at the pass energy



**Figure 18.** (a) Simplified picture of the functioning of the single input transfer lens 1 combined with the Fourier lens 2

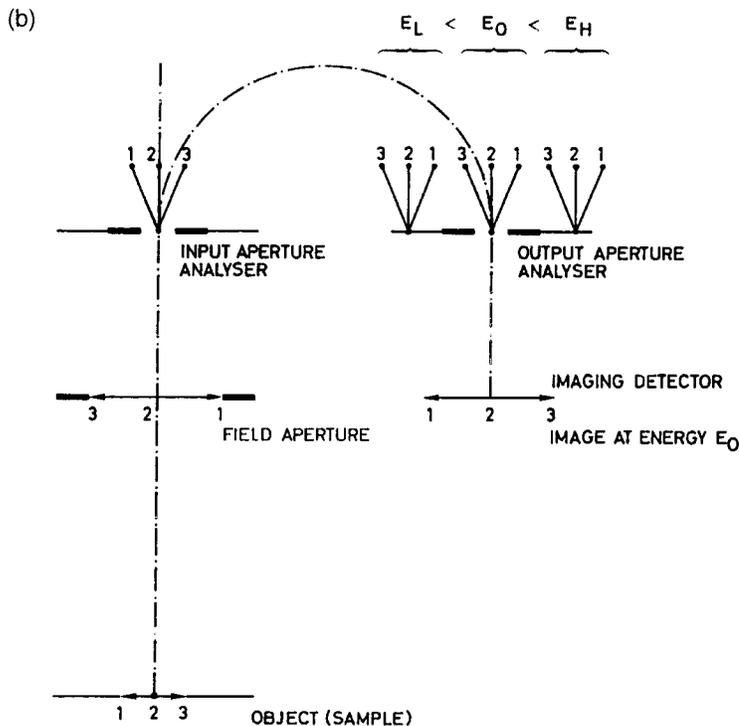
$E_0$ , giving the required energy dispersion. The energy resolution can be changed by changing the size of output and input aperture but, as discussed above, this has no influence on the image information, only on the intensity which translates itself into acquisition time of the image.

Finally, the Fourier lens at the output of the CHA transforms all directions back to a real image on the imaging detector, but at a well-defined energy.

Figure 19 shows an example of an XPS image, recorded using the VG Scientific Ltd ESCALAB 220i fitted with the XL lens. It shows the  $W4f_{7/2}$  image of  $5\ \mu\text{m}$  W bars with  $5\ \mu\text{m}$  spacing, on an Si substrate. The resolution is better than  $2\ \mu\text{m}$ .

### 2.4.3.6 Magnetic Fields in Imaging XPS

The idea of using magnetic fields for XPS imaging is not new. Already in the early 1980s, Beamson and coworkers [61,62] proposed a true XPS microscope in which a superconducting seven tesla magnet was used. Similar equipment, not typically for XPS purposes, has been developed with electromagnetics [63] and even with permanent magnets [64]. In these special arrangements nonhomogeneous magnetic fields [65] are essential. In the high field region each emitted electron spirals around a magnetic field line under



**Figure 18** (b) simplified overall picture of the imaging process as obtained with the aid of Fourier lenses at the input and output of the CEO.

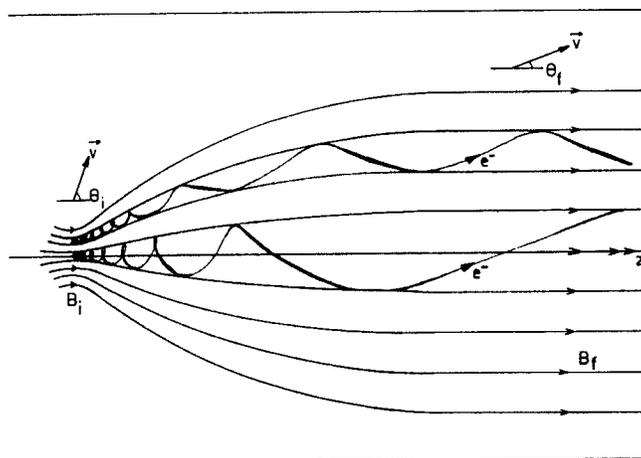
the influence of a Lorentz force. If the change in magnetic field strength during one such spiraling motion is compared with the average magnetic field, the total energy and angular momentum are



**Figure 19.** 5  $\mu\text{m}$  wide W bars, with 5  $\mu\text{m}$  spacing, on an Si substrate. Reproduced by permission of VG Scientific Ltd.

conserved quantities (adiabatic motion) [63]. The total conserved energy of an electron in such a magnetic field is the sum of an axial component, along the field lines, and a cyclotron component. Due to conservation of angular momentum, this cyclotron component of the electron energy decreases as the field strength decreases in the nonhomogeneous diverging magnetic field. The net effect is an increase in the axial component at the expense of the cyclotron energy component and results, as is shown in Fig. 20, in unspiraling and parallelization of the electron trajectories initially collected in a  $2\pi$  steradian solid angle in the high field zone. With this approach, Beamson obtained a true photoelectron microscope with very fast image acquisition and acceptable spatial resolution, but without any

**Figure 20.** Electrons emitted in the high field region of an inhomogeneous divergent magnetic field are parallelized and unspiraled in the low field region. Adapted from P. Kruit and F. H. Read, *J. Phys. E* **1983**, *16*, 313 (Fig. 1). Reproduced by permission of N. Kruit and The Institute of Physics (Great Britain).



inherent energy dispersive capability. By inserting a retarding field, high pass energy images could be obtained, but no true energy dispersive imaging has been reported. Although very promising, no widespread use has been made of such inhomogeneous magnetic fields for XPS imaging. This is probably due to several reasons. As already mentioned, with magnetic fields close to the sample, it is very difficult to provide for instance flooding electrons used to avoid charging of the sample. The intense magnetic fields as used in the Beamson approach will also compromise the correct functioning of any type of electrostatic energy analyzer used to obtain energy-selected images.

A much simpler, yet efficient, approach in XPS imaging relies on the use of a magnetic objective lens, situated as close to the sample as possible (Fig. 21). These lenses do not provide an acceptance angle as large as with the previously discussed nonhomogeneous magnetic fields. However, compared with electrostatic objective lenses, they constitute a much improved acceptance angle. Due to the low aberrations of magnetic lenses, this increase in

acceptance is obtained without loss of spatial or energy resolution. Of course this improvement in an opening angle is translated into a net reduction in acquisition time, that is, the time required to build up an energy selected image of good enough quality. The combination of such a magnetic objective lens with a Fourier imaging system and a CHA yields spatial resolutions [66] of the order of  $1\ \mu\text{m}$ .

## 2.4.4 Characteristics of Scanning Auger Microscopy Images

### 2.4.4.1 General Aspects

The problems encountered in correctly interpreting SAM images have been discussed in considerable detail by Prutton et al. [49]. This discussion summarizes the principal conclusions obtained by these authors. These authors recognize five basic complicating factors which might give rise to image artefacts:

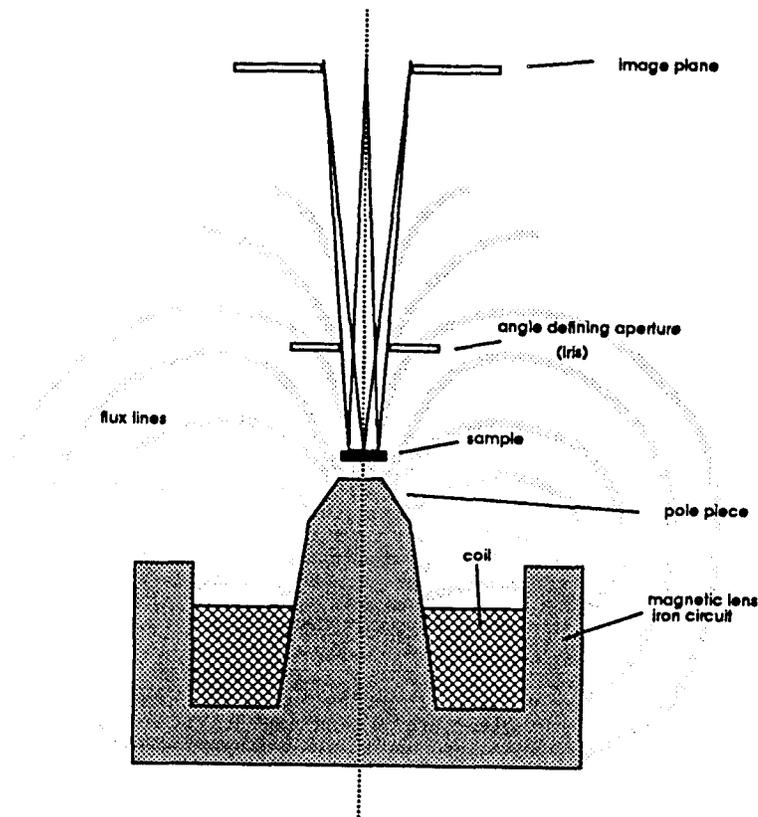


Figure 21. Magnetic objective lens as used in XPS. Adapted from Kratos Axis 165 System Overview. Reproduced by permission of Kratos Analytical Ltd.

- (i) Since in SAM imaging the contrast in the Auger image is often obtained from a comparison of peak heights in the Auger peak of an element and in the background, background slope effects could sometimes lead to an apparent Auger signal although the particular element is not even present.
- (ii) A uniform layer does not necessarily lead to a uniform Auger image, if the substrate under the layer is nonhomogeneous and the Auger backscattering factor varies from place to place.
- (iii) The changes in Auger signal due to surface topography or roughness may be so large that they are able to obscure the variations in chemical composition at the surface.
- (iv) In view of the large scan times that are necessary to obtain favorable signal-to-noise and signal-to-beam Auger signal, beam current fluctuations can occur during the measurements, thus spoiling the SAM image contrast.
- (v) Extreme surface topographics, such as sharp edges, can lead to erroneous chemical compositions as a result of shadowing and/or enhancement effects on the Auger signal.

Prutton et al. described how these effects, which we will now consider in detail, can be remedied in the MULSAM instrument (multispectral Auger microscope). The results obtained on a sample

consisting of a silicon substrate, a gold overlay, an SiGe alloy on top of part of the gold overlay, and the SiGe substrate are discussed.

This sample exhibits four surface regions:

- SiGe on top of the Si substrate
- SiGe on top of the Au overlay
- the Si substrate
- the Au overlay.

This sample should reveal three different regions when imaging the Si concentration in the surface layer: the brightest corresponding to the Si substrate, followed by the next brightest corresponding to Si, and finally a region without Si when the beam strikes the Au overlay.

#### 2.4.4.2 Background Slope Effects

When the image contrast is formed from the quantity  $N_1 - N_2$ , where  $N_1$  represents the counts in the peak and  $N_2$  in the background, or if a linearly extrapolated background is used, four regions are revealed instead of three and Si seems to be more concentrated in the Au overlayer, which, of course, is an erroneous result.

This artefact is explained by the background slope in the spectrum when the beam strikes the Au overlayer. Indeed,  $N_1$  and  $N_2$  were measured at 82 and 102 eV, respectively, which is a justifiable choice when measuring the SiL<sub>2,3</sub>VV Auger peak. The curvature of the background, however, when measuring on Au leads to a larger difference between the signals at 82 and 102 eV than on Si, although no peak is present. From this it is concluded that a more adequate background subtraction procedure is

necessary, for instance, by taking more calibration points on the curve. Unfortunately, this increases the measuring time considerably.

#### 2.4.4.3 Substrate Backscattering Effects

A difference in intensity of the Si signal in the region SiGe, respectively, above Si and Au, is observed, although the same Si concentration in the surface layer is present. This is due to a difference in the Auger backscattering factor  $R$  for both regions. Prutton et al. [49] described how, using backscattering electron detectors in the MULSAM instrument, allowed an estimation of the Auger backscattering factors in an image obtained simultaneously with the energy-analyzed images. When this correction is carried through a uniform intensity for the SiGe region is obtained and a quantitative analysis of the relative Si concentration gives a result in agreement with the expected value for the alloy SiGe.

#### 2.4.4.4 Topographic Effects

The angle of incidence  $\phi$ , measured relative to the surface normal of the primary electron beam, affects the Auger yield. The number of core ionizations within the Auger escape depth increases as  $1/\cos\phi$  on the one hand, and  $R$  varies with  $\phi$  since the energy and angular distribution of backscattered electrons is a function of  $\phi$ , on the other hand. Furthermore, the fraction of the total Auger yield collected by the electron spectrometer is

dependent upon the take-off angle  $\theta$ , also measured relative to the surface normal, because the Auger emission is not isotropic, but exhibits approximately a  $\cos\theta$  emission distribution. Therefore, not taking into account obstruction and shadowing or Auger electron diffraction effects, it is expected that the Auger yield will vary as

$$I_A = I_0 \frac{1}{\cos\phi} R(\phi) \cos\theta \quad (30)$$

Prutton and his colleagues continued by explaining how, in the MULSAM instrument, four BSE detectors placed coaxially around the primary beam, produced difference signals, which enable a topography correction to be made. This was tested on an anisotropically etched (001) slice and it was shown that when this sample, showing eight (331) faces, four (111) faces, and one (100) face, is covered by a carbon layer of ca 100 nm thickness, almost homogeneous carbon intensity can be obtained. The correction procedure is fully described in their paper [49].

#### 2.4.4.5 Beam Current Fluctuation Effects

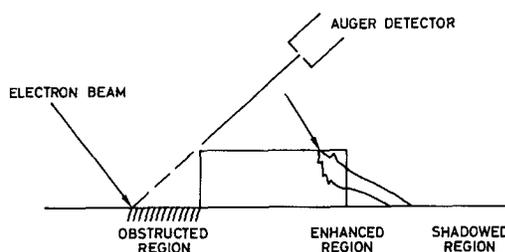
Again, Prutton and his colleagues describe how the acquisition of both a secondary electron image and a sample absorption current image is necessary to correct for beam current fluctuations. Both images are anticorrelated, a feature exploited to form a beam current image, which is then divided into any other images obtained at the same time, to correct them to a constant beam current.

#### 2.4.4.6 Edge Effects

Prutton and colleagues also considered the situation of Fig. 22 where the step side-walls of a metallic overlayer on an Si substrate are shown. On the substrate there will be an obstructed region and a shadowed region (for obvious reasons) and an enhanced region as a result of elastic and inelastic scattering effects. The former can give rise to electrons with a more grazing incidence and the latter to lower primary energy electrons. Both are favorable for increasing the Auger electron yield of the substrate in the region directly adjacent to the metallic overlayer and next to the shadowed region.

These are the well documented edge effects which have to be taken into account when interpreting the Auger images. Prutton et al. [49] explain how this could be achieved and we refer the reader to this paper for more details.

Correction of the different image artefacts, described above, is possible in principle. However, it necessitates a number of supplementary measurements as well as their accompanying software routines. These are not installed on all commercially available instruments and the operator



**Figure 22.** Different regions as discussed in text. Adapted from M. Prutton et al., *Surf. Interface Anal.* **1992**, *18*, 295 (Fig. 3). Reproduced by permission of M. Prutton and John Wiley & Sons Ltd.

should, therefore, be aware of the existence of these artefacts.

## 2.4.5 Conclusion

Scanning Auger microscopy is already in a well-developed state although imaging XPS is still in its infancy. It is, however, a rapidly evolving field as far as the introduction of new instruments is concerned. Characteristics of XPS images are therefore not considered in detail but case studies are discussed in Volume 2 of this Handbook. In this Chapter only scanning Auger microscopes have been discussed. Combinations of STEM instruments with Auger spectroscopy leads to instruments with impressive lateral resolutions. We refer the reader to the literature for a discussion of these combinations [67].

## 2.4.6 References

- [1] L. Fiermans, J. Vennik in *Advances in Electronics and Electron Physics* (Ed. L. Marton), Academic Press **1977**, Vol. 43, pp. 139–203.
- [2] R. Hoogewijs, L. Fiermans, J. Vennik, *Belvac News* **1987**, 3, 7; *Chem. Phys. Lett.* **1976**, 37, 87; *Chem. Phys. Lett.* **1976**, 38, 471; *J. Phys. C.* **1976**, 9, L103; *J. Electr. Spectr. & Rel. Phen.* **1977**, 11, 171; *Surf. Sci.* **1977**, 69, 273.
- [3] M. F. Chung, L. H. Jenkins, *Surf. Sci.* **1970**, 22, 479.
- [4] W. A. Coghlan, R. E. Clausing, *A Catalog of Calculated Auger Transitions for the Elements*, ORNL-TM 3567, Oak Ridge National Laboratory, Oak Ridge, TN **1971**.
- [5] K. Siegbahn, C. Nordling, A. Fahlman, R. Nordberg, K. Hamrin, J. Hedman, G. Johansson, T. B. Bergmark, S. Karlson, I. Lindgren, B. Lindberg, ESCA, *Atomic, Molecular and Solid Structure Studied by means of Electron Spectroscopy, ESCA*, Almqvist & Wiksells AB, Uppsala **1967**.
- [6] K. W. Nebesny, B. L. Maschhoff, N. R. Armstrong, *Anal. Chem.* **1989**, 61, 469.
- [7] M. P. Seah, in *Practical Surface Analysis, Part 1* (Ed: M. Briggs, M. P. Seah), Wiley, New York **1990**, p. 201.
- [8] J. A. Schoeffel, A. T. Hubbard, *Anal. Chem.* **1977**, 49, 2330.
- [9] L. E. Davis, N. C. MacDonald, P. W. Palmberg, G. E. Riach, R. E. Weber, *Handbook of Auger Electron Spectroscopy*, Physical Electronics Inc., **1976**.
- [10] P. De Volder, R. Hoogewijs, R. De Gryse, L. Fiermans, J. Vennik, *Surf. & Interface Anal.* **1991**, 17, 363; *Appl. Surf. Sci.* **1993**, 64, 41.
- [11] D. A. Shirley, *Phys. Rev. B.* **1972**, 5, 4709.
- [12] L. B. Leder, J. A. Simpson, *Rev. Sci. Instrum.* **1958**, 29, 571.
- [13] D. A. Huchital, J. D. Rigden, *Appl. Phys. Lett.* **1970**, 16, 348.
- [14] Ph. Staib, *J. Phys. E* **1972**, 5, 484.
- [15] D. A. Huchital, J. D. Rigden, *J. Appl. Phys.* **1972**, 43, 2291.
- [16] Ph. Staib, *J. Phys. E* **1977**, 10, 914.
- [17] D. W. O. Heddle, *J. Phys. E* **1971**, 4, 589.
- [18] See, for instance, P. Grivet, *Electron Optics, 2nd ed (Eng.)*, Pergamon Press, Oxford, **1972**.
- [19] G. A. Proca, T. S. Green, *Rev. Sci. Instrum.* **1970**, 41, 1778.
- [20] W. Schmitz, W. Mehlhorn, *J. Phys. E* **1972**, 5, 64.
- [21] D. Roy, J. D. Carette, *Rev. Sci. Instrum.* **1971**, 42, 1122.
- [22] H. Hughes, V. Rojanski, *Phys. Rev.* **1929**, 34, 284.
- [23] S. Aksela, *Rev. Sci. Instrum.* **1972**, 43, 1350.
- [24] S. Aksela, M. Karras, M. Pessa, E. Suoninen, *Rev. Sci. Instrum.* **1970**, 41, 351.
- [25] H. Z. Sar-El, *Rev. Sci. Instrum.* **1967**, 38, 1210.
- [26] H. Z. Sar-El, *Rev. Sci. Instrum.* **1968**, 39, 533.
- [27] H. Z. Sar-El, *Rev. Sci. Instrum.* **1971**, 42, 1601.
- [28] H. Z. Sar-El, *Rev. Sci. Instrum.* **1972**, 43, 259.
- [29] J. S. Risley, *Rev. Sci. Instrum.* **1972**, 43, 95.
- [30] P. W. Palmberg, G. K. Bohn, J. C. Tracy, *Appl. Phys. Lett.* **1969**, 15, 254.
- [31] E. M. Purcell, *Phys. Rev.* **1938**, 54, 818.
- [32] C. E. Kuyatt, J. A. Simpson, *Rev. Sci. Instrum.* **1967**, 38, 103.
- [33] J. Cazaux, *J. Microsc.* **1987**, 145, 257.
- [34] J. Cazaux, *Surf. Interface Anal.* **1989**, 14, 354.
- [35] M. Tholomier, D. Dogmane, E. Vicario, *J. Microsc. Spectrosc. Anal.* **1988**, 13, 119.
- [36] M. M. El Gomati, M. Prutton, B. Lamb, C. G. Tuppen, *Surf. Interface Anal.* **1988**, 11, 251.
- [37] V. V. Zashkvara, M. I. Korsunskii, O. S. Kosmachev, *Soviet Phys.-Tech. Phys. (English Transl.)* **1966**, 11, 96.

- [38] H. Hafner, J. A. Simpson, C. E. Kuyatt, *Rev. Sci. Instrum.* **1968**, 39, 33.
- [39] See, for instance, A. E. Conrady, *Applied Optics and Optical Design*, Dover Publications, New York **1957**, p. 120.
- [40] J. L. Gardner, J. A. R. Samson, *J. Electron. Spectrosc. & Relat. Phenom.* **1975**, 6, 53.
- [41] J. M. Watson, US Patent 3 783 280, 1 Jan. **1974**.
- [42] H. G. Nöller, H. D. Polaschegg, H. Schillalies, Proceedings of the 6th International Vacuum Congress **1974**; *Jpn. J. Appl. Phys.*, Suppl. 2, Pt. 1, **1974**.
- [43] J. C. Helmer, N. H. Weichert, *Appl. Phys. Lett.* **1968**, 13, 266.
- [44] P. J. Bassett, T. E. Gallon, M. Prutton, *J. Phys. E* **1972**, 5, 1008.
- [45] K. Yates, R. H. West, *Surf. Interface Anal.* **1983**, 5, 217.
- [46] B. Wannberg, A. Skollermo, *J. Electron. Spectrosc. & Relat. Phenom.* **1977**, 10, 45.
- [47] N. C. MacDonald, J. R. Waldrop, *Appl. Phys. Lett.* **1971**, 19, 315.
- [48] J. Cazaux, *Surf. Sci.* **1983**, 125, 335; M. M. El Gomati, A. P. Janssen, M. Prutton, J. A. Venables, *Surf. Sci.* **1979**, 85, 309.
- [49] M. Prutton, I. R. Barkshire, M. M. El Gomati, J. C. Greenwood, P. G. Kenny, R. H. Robert, *Surf. Interface Anal.* **1992**, 18, 295.
- [50] See, for instance, D. Briggs, M. P. Seah (Eds), *Practical Surface Analysis by Auger and X-ray Photoelectron Spectroscopy*, Wiley, New York **1983**, p. 128.
- [51] B. L. Chaney, *Surf. Interface Anal.* **1987**, 10, 36.
- [52] K. Sieghbahn, C. Nordling, A. Fahlman, R. Nordberg, K. Hamrin, J. Hedman, G. Johansson, T. B. Bergmark, S. Karlson, I. Lindgren, B. Lindberg, *ESCA, Atomic, Molecular and Solid Structure Studied by Means of Electron Spectroscopy*, Almqvist & Wiksells AB, Uppsala **1967**.
- [53] C. J. Hovland, *Appl. Phys. Lett.* **1977**, 30, 274.
- [54] D. J. Keast, K. S. Dowing, *Surf. Interface Anal.* **1981**, 3, 99.
- [55] M. P. Seah, G. C. Smith, *Surf. Interface Anal.* **1988**, 11, 69.
- [56] N. Gurker, M. F. Ebel, H. Ebel, *Surf. Interface Anal.* **1983**, 5, 13.
- [57] N. Gurker, M. F. Ebel, H. Ebel, M. Mantler, H. Hedrich, P. Schön, *Surf. Interface Anal.* **1987**, 10, 242.
- [58] D. A. Allison, T. F. Anater, *J. Electron Spectrosc.* **1987**, 43, 243.
- [59] P. Coxon, J. Krizek, M. Humpherson, I. R. M. Wardell, *J. Electron Spectrosc. & Relat. Phenom.* **1990**, 52, 821.
- [60] E. Adem, R. Champaneria, P. Coxon, *Vacuum* **1990**, 7, 1695.
- [61] G. Beamson, H. Q. Porter, D. W. Turner, *J. Phys. E* **1980**, 13, 64.
- [62] G. Beamson, H. Q. Porter, D. W. Turner, *Nature* **1981**, 290, 556.
- [63] P. Kruit, F. H. Read, *J. Phys. E* **1983**, 16, 313.
- [64] T. Tsubai, E. Y. Xu, Y. K. Bae, K. T. Gillen, *Rev. Sci. Instrum.* **1988**, 59(8), 1357.
- [65] T. Hsu, J. L. Hirshfield, *Rev. Sci. Instrum.* **1976**, 47, 236.
- [66] N. M. Forsyth, P. Coxon, *Surf. Interface Anal.* **1994**, 21, 430.
- [67] G. C. Hembree, J. A. Venables, *Ultramicroscopy* **1992**, 47, 109–120.

## 2.5 Scanning Microanalysis

### 2.5.1 Physical Basis of Electron Probe Microanalysis

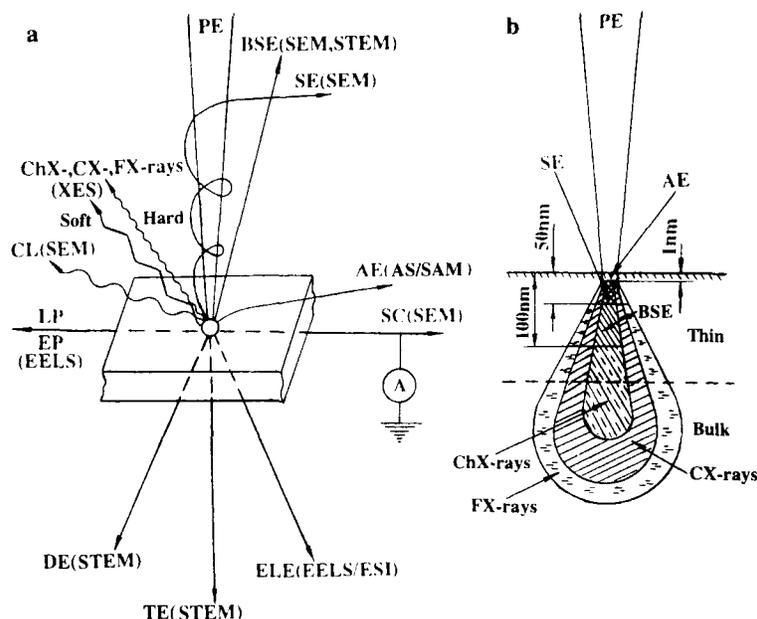
Classical electron probe microanalysis (EPMA) is carried out by determining the intensity of characteristic X-ray emission from an element of interest in the specimen and comparing it with that from a standard of known composition. Maintaining the same analysis conditions enables an X-ray intensity ratio to be obtained which, according to Castaing's first approximation, relates quantitatively to the mass concentration of the element [1].

#### 2.5.1.1 Electron Interactions in Solids

In general, electron microanalysis is based on the effects of elastic and inelastic scattering of an accelerated electron beam upon interaction with atoms and electrons of the material to be examined [2]. The final signals used for image formation and for analytical measurements are not normally the result of single scattering processes but of some electron diffusion caused by the gradual losses of electron energy and by some lateral spreading due

mainly to multiple elastic large-angle scattering. Figure 1 illustrates schematically the most important interaction processes and signals detected in different operating modes of the analytical electron microscope (a) and their information volumes (b). The complete energy spectrum comprises primary electrons with energy  $E_0$ , and emitted electrons, ions, heat, quanta, and internally generated signals such as: transmitted electrons ( $E \approx E_0$ ); wide-angle and narrow-angle elastically scattered and/or diffracted electrons ( $E \approx E_0$ ); backscattered electrons ( $50 \text{ eV} < E \leq E_0$ ); secondary electrons ( $0 < E \leq 50 \text{ eV}$ ); inelastically scattered electrons with energy loss  $\Delta E$  ( $E = E_0 - \Delta E$ ); sample current or absorbed electrons ( $E = E_F$ , where  $E_F$  is the Fermi level); Auger electrons ( $E < 10 \text{ eV}$ ); hard and soft characteristic and continuous X-ray photons ( $0 < h\nu < E_0$ ); cathodoluminescence ( $0 < h\nu < 1-3 \text{ eV}$ ); and electron plasmons and lattice phonons.

Usually only a comparatively small fraction of the characteristic X-rays isotropically emitted from the specimen is detected because of small solid angles of collection ( $10^{-3}$ – $10^{-1}$  sr). Moreover, inner-shell ionizations result in the emission not only of X-rays but also of Auger electrons, and the X-ray fluorescence yield

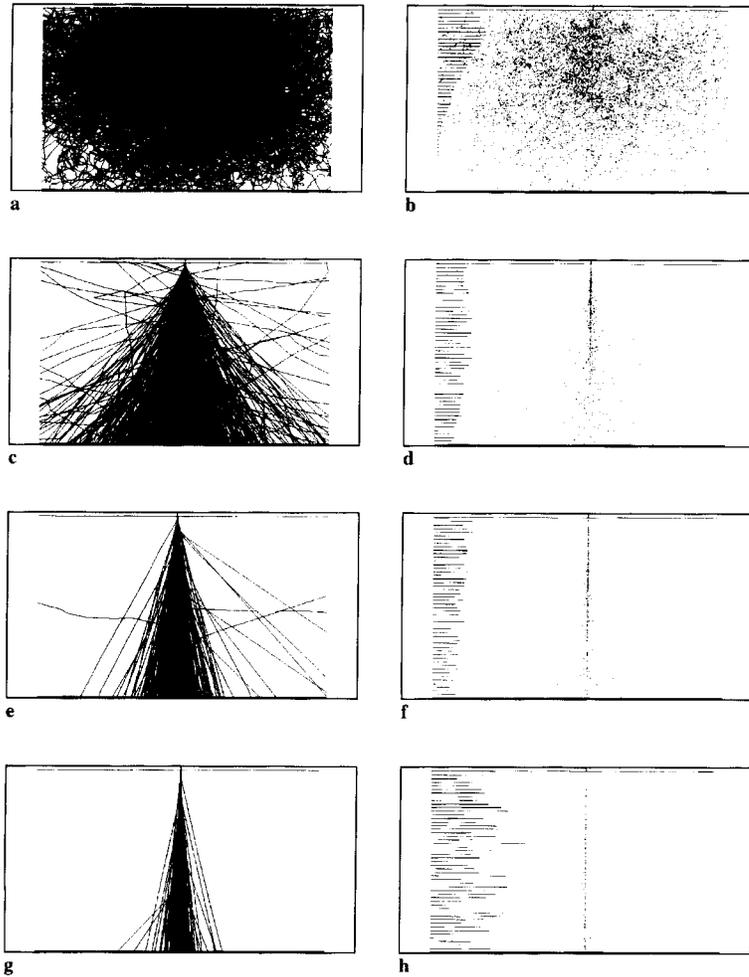


**Figure 1.** Main processes of electron beam interactions and signals detected in different operating and analytical modes of an analytical electron microscope (a) and schematic arrangement of their information volumes (b); PE, primary electrons; TE, transmitted electrons; DE, diffracted electrons; ELE, electrons with losses of energy; BSE, backscattered electrons; SE, secondary electrons, the information depth of SE is 5–50 nm; AE, Auger electrons, the information depth of AE is about 1 nm; SC, sample current; ChX-rays, characteristic X-rays; CX-rays, continuous X-rays; FX-rays, secondary X-ray fluorescence; CL, cathodoluminescence in ultraviolet, visible, and infrared regions; EP, electron plasmons; LP, lattice phonons. SEM, scanning electron microscopy; STEM, scanning transmission electron microscopy; XES, X-ray emission spectroscopy; AS, Auger spectroscopy; SAM, scanning Auger microscopy; EELS, electron energy-loss spectroscopy; ESI, electron spectroscopic imaging.

decreases with decreasing sample atomic number  $Z$ , being below 1% for light elements. A large fraction of about 10–70% of the inner-shell ionization processes leads to inelastically scattered electrons concentrated in small scattering angles which also pass through the objective diaphragm. Atomic electrons can be excited from an inner K, L, or M shell to unoccupied energy states above the Fermi level, resulting in a characteristic edge in the energy-loss spectrum. Complementary to X-ray spectroscopy, in specimens with thickness of about or smaller than the mean free electron path, the well-defined ionization edges, in particular those due to

the K shell excitation for elements with atomic number  $Z \leq 12$ , can be easily analyzed by electron energy-loss spectroscopy (EELS) [3].

The plasmons and inelastic intra- and interband excitations of the outer shell electrons near the Fermi level that normally can be observed with energy losses  $\Delta E$  smaller than 50 eV are influenced by chemical bonds and the electronic band structure, by analogy with optical excitations. In semiconductors the electron impact results in generation of electron-hole pairs and causes an electron-beam-induced current (EBIC). Electron-hole pairs can recombine with emission of

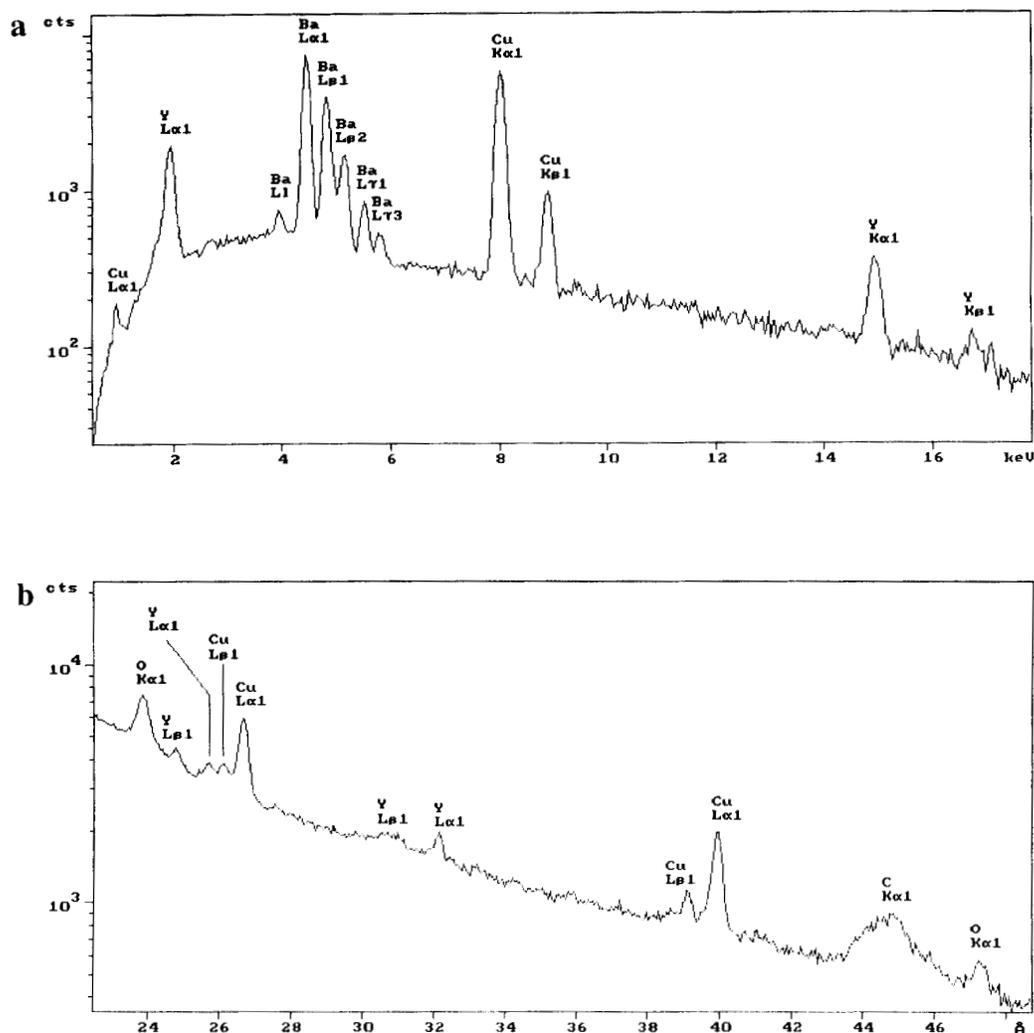


**Figure 2.** Single scattering Monte Carlo simulations of a lateral electron distribution (a, c, e, g) and Ag  $L_{\alpha}$  X-ray generation (b, d, f, h) in an AgBr tabular crystal of 100-nm thickness with a point electron source at 5 keV (a, b), 30 keV (c, d), 100 keV (e, f), and 300 keV (g, h) for 1000 trajectories. A histogram of the normalized yield of Ag  $L_{\alpha}$  emission  $\phi(\rho z)$  is plotted on the left-hand side in (b), (d), (f), and (h), starting at the crystal surface; the length of the bars shows the relative value of  $\phi(\rho z)$  at that depth.

luminescent photons in either the ultra-violet, visible, or infrared regions, or by nonradiative lattice phonons. Some fraction of  $E_0$  that is lost in the course of a complicated cascade of inelastic scattering processes is converted into phonons and/or heat and causes radiolysis, thermal damage, bond rupture, and loss of mass and crystallinity by sputtering of specimen matter.

Hence the actual mechanisms of electron–solid interactions may be rather complicated. Therefore, estimates of the

information volume of different signals based on Monte Carlo simulations (Fig. 2) are of importance, in particular, for understanding relationships between scattered electrons and X-rays [4]. The consequence of the gradual diminution of the electron energy is that the electrons have a finite depth range of the order of several nanometers up to tens of micrometers, depending on the value of  $E_0$  and the thickness and density of the specimen. The information depth and the lateral extent of the information volume



**Figure 3.** The X-ray emission spectra of a high temperature superconductor  $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$  ceramic recorded at an accelerating voltage of 25 kV. (a) EDS, 1 nA probe current; and (b-d) WDS, 200 nA probe current: (b) lead stearate crystal; (c) PET crystal; (d) LiF crystal.

governing the resolution of the corresponding operating modes contribute to each of the possible signals, decreasing considerably with a decrease in specimen thickness and an increase in its density. In this way the various electron-specimen interactions can generate a great deal of structural and analytical information in

the form of emitted electrons and/or photons and internally produced signals.

### 2.5.1.2 X-Ray Emission Spectra

Ionization of an inner electron shell by the inelastic impact results in a vacancy which

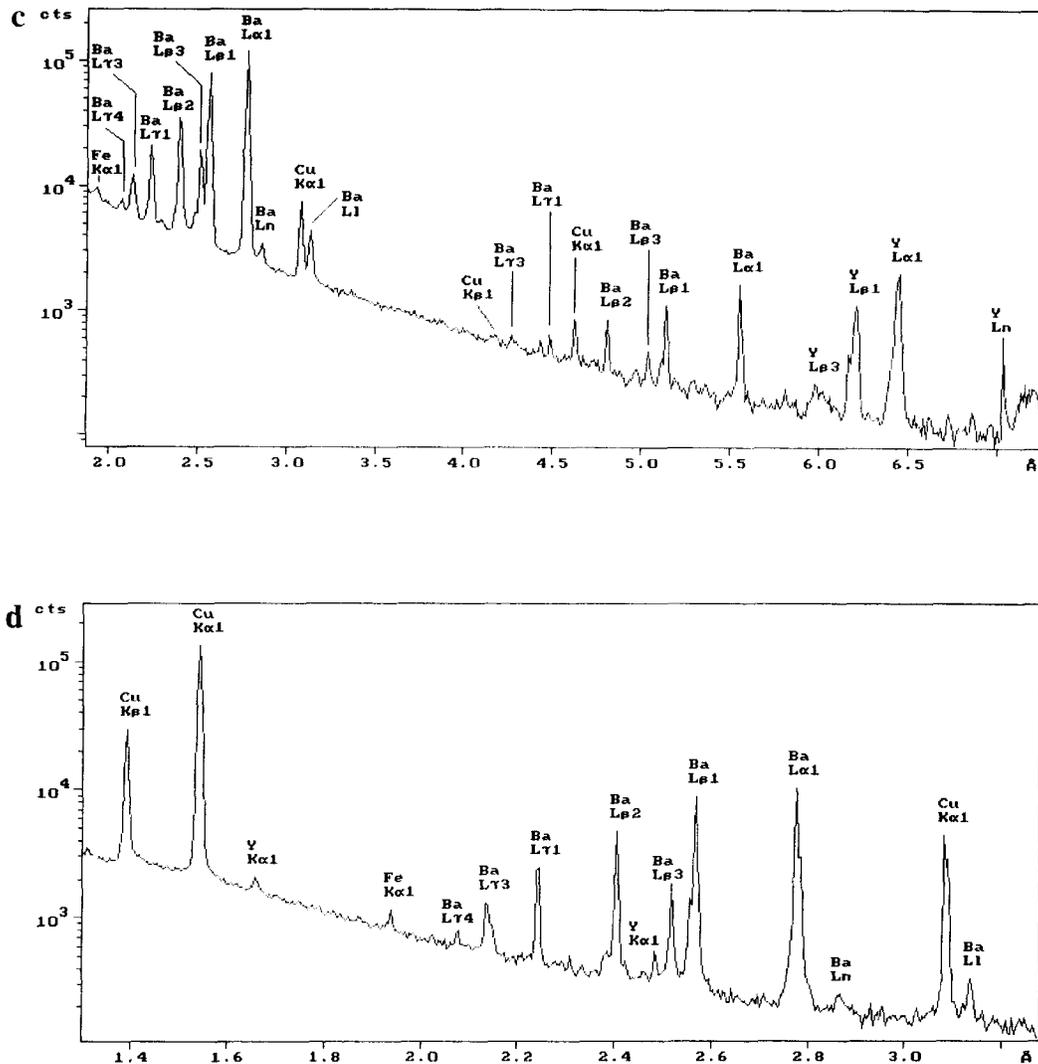


Figure 3 Continued

can be filled by an electron from a higher state. The energy difference can then be emitted either as a characteristic X-ray quantum or as an Auger electron. Hence the spectrum represents the balance between two energy levels; that is, generated intensities are the product of the state density and transition ratio governed mainly by selection rules. The X-ray

emission spectra of an advanced high temperature superconductor  $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$  ceramic are shown in Fig. 3. They consist of a background (continuum X-rays, bremsstrahlung) which extends up to the energy of the incident beam, together with superimposed discrete characteristic lines of Cu (CuK and CuL series), Y (YK and YL series), Ba (BaL series) and O (OK

series). The wavelength-dispersive spectroscopy (WDS) spectra (Fig. 3b–d) clearly demonstrate much better resolution than the energy-dispersive spectroscopy (EDS) spectrum (Fig. 3a). The peak of  $CK_{\alpha}$  at 4.47 nm (0.277 keV) in Fig. 3b belongs to a carbon conductive coating deposited on the specimen surface. The X-ray spectra in Figs 3c and 3d also show the presence of traces of Fe ( $FeK_{\alpha}$  at 0.194 nm; 6.398 keV) and Pt ( $PtL_{\alpha}$  at 0.131 nm; 9.441 keV), probably from crucible material.

### 2.5.1.3 Characteristic X-Ray Spectra

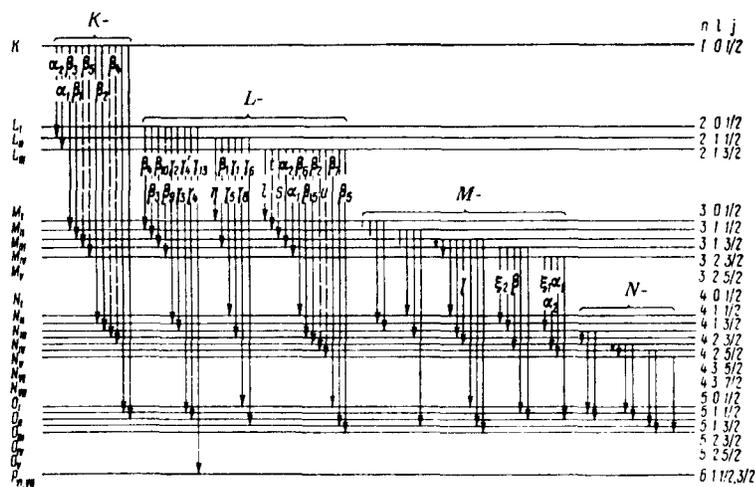
The relationship between the energy of characteristic X-ray emission lines and the atomic number of the element of interest is described by Moseley's law:

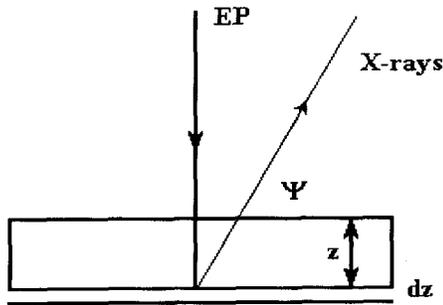
$$E = A(Z - 1)^2 \quad (1)$$

where  $A$  is constant within K, L, and M series, and  $Z$  is the target atomic number. When an inner shell electron is ejected

from the atom, the latter becomes ionized and goes to a higher energy state. The vacancy formed in this way must be filled by an electron from one of the outer levels. Electron transitions are regulated by the selection rule in which  $\Delta n \neq 1$ ,  $|\Delta l| = 1$ ,  $|\Delta j| = 0$ , where  $n$ ,  $l$ , and  $j$  denote the principal quantum number, azimuthal quantum number, and inner quantum number, respectively. An X-ray quantum can be emitted with a discrete energy corresponding to the difference in energy between the levels involved. Major X-ray emission lines together with their designation are shown in Fig. 4. However, transitions which do not satisfy the selection rules, so called 'forbidden' transitions, can in fact occur and do produce some emission lines, but their intensities are usually low. Bonding also somewhat affects inner shell electrons due to the change in the surrounding charge distribution. However, the range of characteristic lines which allow one to observe chemical effects is fairly limited. This is because the spectrum does not reflect precisely changes in electron states: there are no changes in

Figure 4. Schematic diagram showing common X-ray emission lines and their designations.





**Figure 5.** The intensity of the total X-ray emission originating from depth  $z$ . EP, electron probe;  $\psi$ , angle between emitted X-rays and the specimen surface.

X-ray spectra if the chemical effects on two different energy levels are the same.

The intensity of the total X-ray emission originating from depth  $z$  (Fig. 5) below the surface of a specimen with density  $\rho$  including any fluorescence contribution may be expressed [1] as

$$I = \phi(\Delta\rho z) \int_0^\infty \phi(\rho z) \times \exp(-\chi\rho z) d\rho z f(\chi)(1 + \gamma + \delta) \quad (2)$$

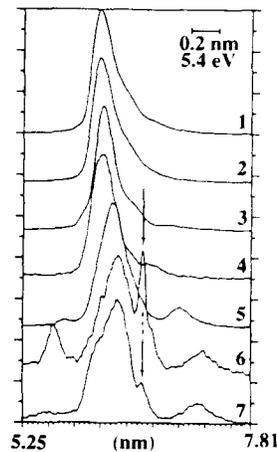
where  $\phi(\Delta\rho z)$  is the emission from an isolated thin film of mass thickness  $\Delta\rho z$ ,  $\chi = (\mu/\rho) \text{ cosec } \psi$ ,  $\rho z$  is the specimen mass thickness, and  $\mu$  is the linear absorption coefficient. The absorption factor is defined as

$$f(\chi) = \frac{\int_0^\infty \phi(\rho z) \exp(-\chi\rho z) d\rho z}{\int_0^\infty \phi(\rho z) d\rho z} \quad (3)$$

The fluorescence correction factor  $(1 + \gamma + \delta)$  includes the ratio of the fluorescence intensity to the primary characteristic X-ray intensity  $\gamma$ , and the corresponding ratio for the continuum fluorescence contribution  $\delta$ .

Although most of the characteristic X-ray emission can be explained on the basis

of transitions allowed by the selection rules, weak lines may appear which occur as satellites close to one of the principal lines. Their production has been explained by assuming that an atom may be doubly ionized by the incident radiation. The two ionizations have to occur virtually simultaneously because the lifetime of an excited state is very short ( $10^{-14}$  s). For example, an electron transition from the doubly ionized state in both K and L orbitals (Fig. 6), results in the emission of a single photon (the KL–LL transition). Its energy may be somewhat higher than that associated with an ordinary L–K transition owing to the fact that the original extra vacancy would have reduced the degree of screening of the nucleus by the electrons and thus increased their binding energy. Satellite lines then appear on the high-energy side of the  $K_{\alpha_1\alpha_2}$  peak at energies corresponding to the doubly ionized L suborbital. Two  $B_2O_3$  species in Fig. 6



**Figure 6.** Satellite lines (arrows), line shifts and change of shape due to chemical bonding effect in the  $BK_{\alpha}$  peak of boron containing compounds recorded with a stearate crystal at an accelerating voltage of 10 kV: (1) B in SiC; (2) pure B; (3)  $LaB_6$ ; (4) BP; (5) BN; (6) B–Si glass (14.9%  $B_2O_3$ ); (7) datolite (20.24%  $B_2O_3$ ). Courtesy of JEOL.

(borosilicate glass (6) and natural mineral datolite (7)) reveal the emergence of satellite peaks. However, the satellite intensity in spectrum (7) is much smaller than that in (6), probably due to chemical effects, although the content of  $B_2O_3$  is higher.

Satellites can also occur on the low-energy side of the line and may contribute to the asymmetry of the energy distribution of the band generated. Usually, satellites are relatively more intense for lighter elements because the lifetime of an excited state is longer and the probability of double ionization is higher. Spectral deconvolution accounting for the presence of low- and high-energy satellites and instrumental distortions is required in order to process experimental asymmetrical peaks of soft X-ray emission bands. It has been found that peak shape changes as a function of the excitation conditions, and the matrix composition is related to self-absorption phenomena [5].

### 2.5.1.4 Soft X-Ray Spectra

The soft X-ray range can be defined as extending from about 100 eV up to 1.5 keV. For light elements ( $4 \leq Z \leq 9$ ), in particular, there is no alternative other than using soft X-ray emission. However, soft X-rays may be produced not only due to electron transitions involving inner orbitals of the atoms but also as a result of transitions associated with outer orbitals containing valence electrons. The overlap of the valence energy states leads to a decrease in X-ray intensity from ionized atoms. Moreover, because the inner levels are relatively discrete compared with the valence band transitions

associated with outer orbitals, the low-energy X-ray lines sensitively reflect the energy states of valence electrons. Chemical bonding effects in the soft X-ray emission spectra are usually more pronounced in insulators than in conductors because the binding energies of valence electrons increase from metallic, through covalent, to ionic bonding. At the same time the energy of the soft X-ray emission decreases correspondingly. Line shifts and change of shape due to the chemical bonding (see Fig. 6) may be observed in the  $K_\alpha$  series from the light elements ( $Z < 10$ ), where the L shells involved in K-L transitions are incomplete, as well as in the L series from transition metals and their compounds ( $21 < Z < 28$ ). The  $BK_\alpha$  lines of boron in SiC (1), pure boron (2),  $LaB_6$  (3), BP (4), BN (5), and two  $B_2O_3$  containing species (6, 7) show changes in peak position and shape, as well as the above considered satellites. The valence of boron in each compound is +3 and, therefore, it cannot affect the spectra. Furthermore, differences between the  $BK_\alpha$  profiles correlate with the electronegativity of the neighboring element P (2.1), N (3.0), and O (3.5) which form the compounds with boron (2.0).

Undoubtedly, analysis of the fine structure of soft X-ray spectra recorded with the appropriate resolution can give not only the elemental composition of an object under study but also important information on electronic structure and bonding. So, analysis of soft  $FeL_{III,II}$  X-ray emission spectra of the mineral wüstite allowed an estimate to be made of the relative energies of the valence conduction band orbitals as well as the splitting of these orbitals in the crystal field and the size of their spin splitting [6]. Application

of low-energy X-rays to layered specimens enables one to obtain a reduced depth of ionization. It also permits one to eliminate most of the secondary X-ray fluorescence emission that occurs when lines of higher energy are employed, particularly for the EPMA of multiphase specimens [7]. On the other hand, in this case, quantitative analysis may be complicated by a number of problems such as contamination, coating, background subtraction, and line interference. Therefore, a cautious approach and a deeper insight into the physical processes involved are necessary.

### 2.5.1.5 X-Ray Continuum

Bremsstrahlung is produced as a consequence of the slowing down of electrons in the Coulomb field of atomic nuclei. The continuous X-rays form a background over a wide energy range  $0 \leq E \leq eU$  extending up to an energy corresponding to the conversion of the entire energy of an incident electron into a radiation photon in one single interaction. The energy distribution of bremsstrahlung may be expressed by the Kramers' equation [1]:

$$N(E) dE = bZ \frac{E_0 - E}{E} dE \quad (4)$$

where  $N(E) dE$  is the number of photons within the energy interval  $E$  to  $E + dE$ ,  $b = 2 \times 10^{-9}$  photons  $s^{-1}$   $eV^{-1}$  electron $^{-1}$  is Kramers' constant, and  $E_0 = eU$  is the incident electron energy in electron-volts. The angular distribution of the X-ray continuum is anisotropic.

The bremsstrahlung intensity can be used to calibrate the film thickness in the microanalysis of biological sections. It also

contributes to the background below the characteristic X-ray peaks, thereby decreasing the peak-to-background ratio; the latter can be improved by increasing the energy of the incident electron beam.

### 2.5.1.6 Overview of Methods of Scanning Electron Beam Analysis

EPMA instruments and modern analytical scanning electron microscopes, scanning transmission electron microscopes, and specially designed Auger-electron micro-analyzers equipped with corresponding analytical facilities in order to detect X-rays, inelastically scattered electrons, and/or Auger electrons are the most important electron-optical instruments for the analytical and structural investigation of the variety of bulk and thin samples. Table 1 contains some basic data characterizing state-of-art modern analytical scanning electron beam techniques. This information may also be useful in comparing different methods.

### 2.5.1.7 Electron Probe X-Ray Microanalyzers

The main task of the electron probe X-ray microanalyzer shown in Fig. 7a is to analyze the elemental compositions of flat, polished surfaces at normal electron incidence with a high analytical sensitivity. The ray diagram of such an instrument (Fig. 7b) is similar to that for a scanning electron microscope, but it contains an additional optical microscope to select the specimen points, profiles, and/or

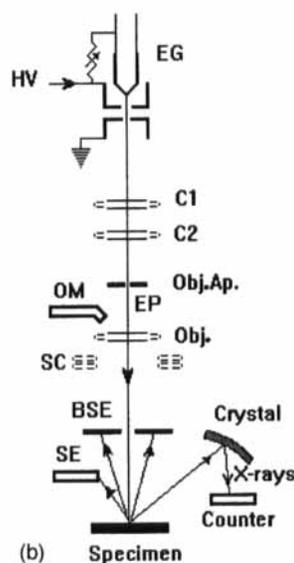
Table 1. Basic characteristics of scanning electron probe methods<sup>a</sup>

Method	Signals	Type of contrast	Resolution (nm) (magnification)	Information obtained	Information depth (nm)	Element range	Detection limit (at. %) (Accuracy (%))	Refs
SEM	SE	Topography, atomic number, orientation	0.6-7 (edge-to-edge) ( $\times 10^{-10}^6$ )	Structural (surface topography images)	5-50	-	-	2, 8
	BSE	Atomic number	1.5-15 ( $\times 10^{-10}^4$ )	Elemental (composition images), structural (surface topography images)	100	$\Delta Z < 0.1$	-	
	SC	Atomic number-orientation	3-100	Elemental (composition images), structural	1000	-	-	
STEM	TE, DE	Diffraction	0.2-2 (point-to-point) ( $\times 300-10^7$ )	Structural (bright-, dark-field images)				2, 9, 17-19
	TE+DE	Phase		Structural (lattice images)				
	DE, TE+DE	Z contrast		Structural (dark-field images)				
	BSE	Atomic number	3-10	Elemental	$t$ (film)	$\Delta Z < 0.1$	2-3	2
SAED/STEM	DE	-	1000	Structural (point, ring diffraction patterns)	-	-	(10)	
CBED/STEM	DE	-	5	Structural (microdiffraction patterns)	-	-	-	2
WDS	X-ray photons	-	1000 (bulk)	Elemental (spot spectra, profiles, mapping)	1000 (bulk)	4-92	0.01 (2-6)	1, 2, 9-12
EDS	X-ray photons	-	1000 (bulk) 0.6-1 (film)	Elemental (spot spectra, profiles, mapping)	1000 (bulk) $t$ (film)	(5)11-92	0.1 (2-6)	1, 9-15

EELS/ESI with $0 < E < 2 \text{ keV}$	ELE	-	1-10	Elemental (spot spectra, 2D mapping); chemical (ELNES, EXELFS) structural (ESI images, diffraction patterns), surface	$t$ (film)	(1)3-92	1-5 (10-20)	9, 11, 16, 17, 19, 20
AES/SAM	SE, AE	-	8-300	Elemental (spot spectra, mapping), chemical (spectra), structural (spectroscopic images)	1	3-92	0.1 (10)	2, 21
CL, UV-VIS, SEM IR photons $0 < h\nu < 1-3 \text{ eV}$	-	-	30-500					2, 24

<sup>a</sup>For abbreviations, see legend to Fig. 1.

(a)



**Figure 7.** General view (a) and ray diagram (b) of an electron microprobe. On the right-hand side in (b) the electron optical column with two vertical and one inclined wavelength-dispersive spectrometers. On the left-hand side is the data acquisition and processing system including a 32-bit SUN workstation with a 20 in. color screen, a 17 in. frame-store color monitor with a 1024×768 pixel observation screen, and a dedicated keyboard. EG, electron gun; HV, high voltage; C1, first condenser lens; C2, second condenser lens; Obj. Ap., objective lens aperture; OM, optical microscope; EP, electron probe; Obj., objective lens; SC, scanning coils; BSE, backscattered electron detector; SE, secondary electron detector. Courtesy of CAMECA.

areas to be analyzed and up to five wavelength-dispersive X-ray crystal spectrometers which can record different characteristic X-ray wavelengths, and often also an energy-dispersive X-ray spectrometer which can detect X-rays in a wide energy range. An electron probe (about 6 nm to 1  $\mu$ m in diameter), governed by the acceptable probe current ( $10^{-12}$ – $10^{-5}$  A), in the scanning mode is produced by a one-, two- or three-stage demagnification of the smallest cross-section of the electron beam after acceleration. Images are displayed on a cathode-ray tube (CRT) rastered in synchronism. The CRT beam intensity may be modulated by any of the different signals (i.e. secondary electrons, backscattered electrons, sample current or X-rays) that result from the electron–specimen interactions.

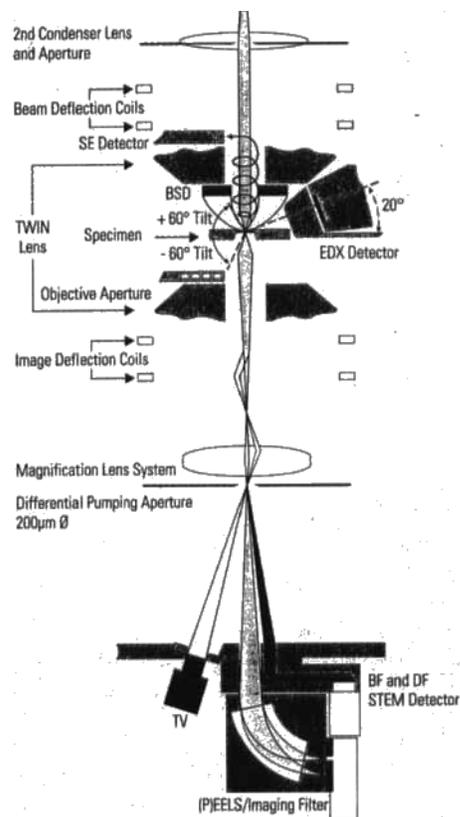
The combined wavelength- and energy-dispersive microanalyzer is a new-generation analyzer that controls its energy- and wavelength-dispersive spectrometers with the aid of a powerful computer multitasking workstation and presents X-ray data acquired by both spectrometers and images as a unified analysis result. The combination of WDS and EDS can increase the number of simultaneously detectable elements to 13 (5 with WDS and 8 with EDS). A TV display allows the optical microscope image and one of the scanning electron microscope modes (secondary electron or backscattered electron images) to be observed simultaneously.

Traditionally, the energy- and wavelength-dispersive systems in an electron probe X-ray microanalyzer or in an analytical scanning electron microscope are operated at a high voltage (15–40 kV). These conditions are sufficient to allow

excitation of X-ray emission of all relevant elements. However, with the increasing interest in detecting light elements and the availability of ultrathin-window or windowless detectors, the importance of the application of low voltages (several kilovolts and less than 1 kV) is also emphasized. With field-emission electron guns that produce sufficient beam currents even in the low-voltage range (5–2 kV), the lateral resolution for X-ray analysis may be improved significantly [22]. As a consequence, thin layers and small particles can be examined in the scanning electron microscope without interference of the bulk support.

### 2.5.1.8 Analytical Electron Microscopes

A remarkable capability of the scanning transmission electron microscope is the formation of very small electron probes less than 1 nm in diameter by means of a three-stage condenser-lens system. This enables the instrument to operate in the scanning transmission mode with a resolution determined by the electron probe diameter and sample thickness (thin samples). The main advantage of equipping a transmission electron microscope with a scanning transmission electron microscope attachment is the possibility to produce a very small electron probe, with which X-ray analysis can be performed on extremely small areas. X-ray generation in thin foils is confined to the small volume excited by the electron probe only slightly broadened by multiple scattering. Better spatial resolution is therefore obtainable for precipitates, or for segregation effects

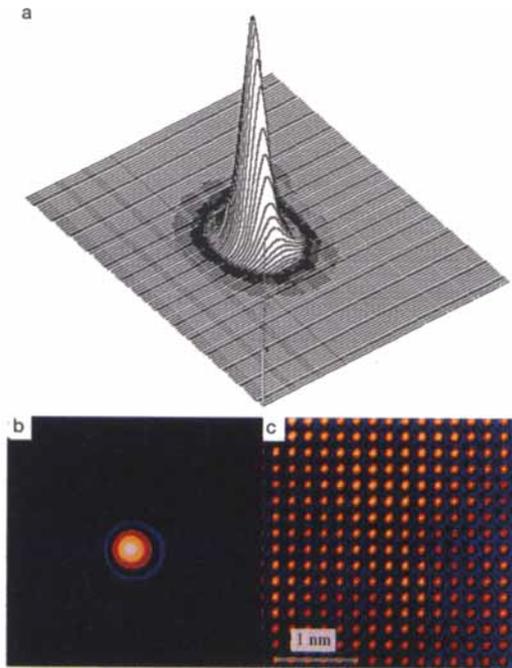


**Figure 8.** General view (a) and schematic diagram of the electron optics (b) of a dedicated analytical scanning transmission electron microscope. On the left-hand side in (b) is the high-resolution TV monitor and the electron optical column with side-entry airlock and windowless electron detector. On the right-hand side is the scanning observation attachment, control panel and X-ray electron detection system. Courtesy of Philips Electron Optics.

at crystal interfaces, than in an X-ray microanalyzer with bulk specimens, where the spatial resolution is limited to 100–1000 nm by the diameter of the electron interaction volume.

Figure 8 shows a dedicated analytical scanning transmission electron microscope. Normally such an instrument involves a field-emission gun, a probe-forming condenser-lens system, an objective lens, and an electron-detection system, often together with a parallel electron energy-loss spectrometer for separating

the currents of unscattered electrons, of elastically scattered electrons and of inelastically scattered electrons. Electron probes of 0.2–0.5 nm diameter (Fig. 9) can be formed, the spherical aberration of the lens being the limiting factor in this case. An advantage of scanning transmission electron instruments is that the contrast can be enhanced by collecting several signals and displaying differences and/or ratios of these by analog or digital processing. Single atoms of heavy elements on an appropriate thin substrate can be



**Figure 9.** Three-dimensional view of a 0.4 nm (FWHM) electron probe (a) and spot-size measurements at a current of 30 pA (b) using a slow-scan CCD camera. (c) The lattice image of gold in the [100] orientation, recorded and displayed at exactly the same magnification as the image in (b), shows the 0.2-nm spacing for calibration. Courtesy of Dr M. Otten, Philips Electron Optics, The Netherlands.

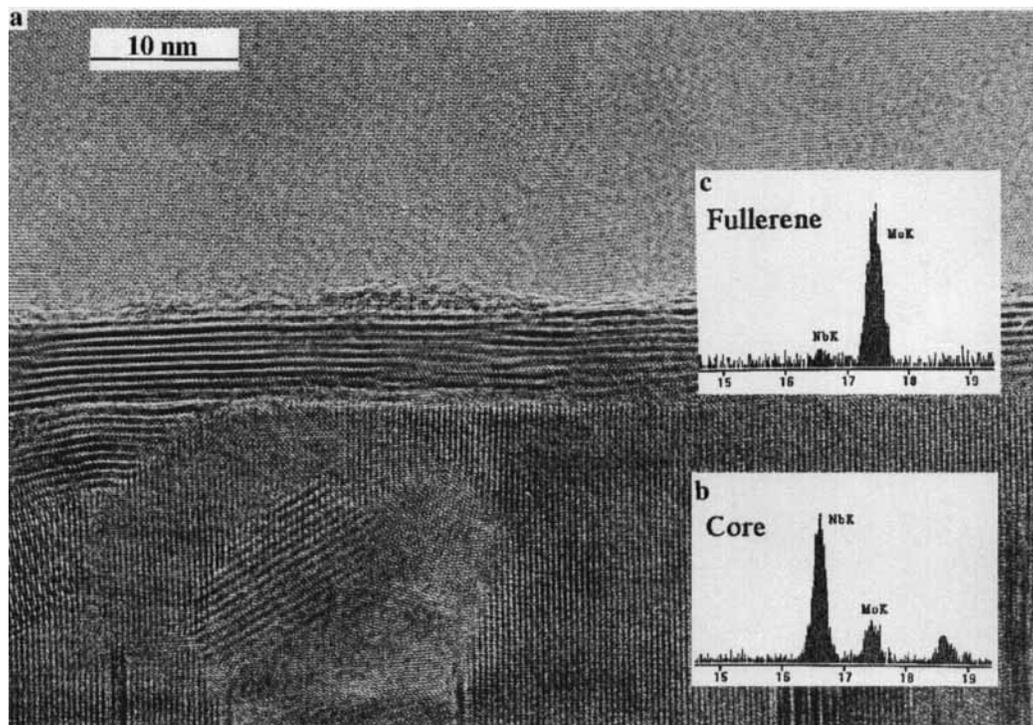
imaged using Z-contrast with a wide-angle annular dark-field semiconductor detector [19]. A higher contrast than in conventional TEM bright- or dark-field modes is achieved in this case.

X-ray microanalysis in scanning transmission electron microscopy (STEM) has a significant advantage over classical EPMA in that the lateral resolution can be improved by reducing the illumination area to less than 1–10 nm (Fig. 10). The fraction of continuous X-ray emission is lower than for bulk samples because of preferential emission in the forward direction. Recently, improvements in X-ray

detection have been achieved by implementing in a single 100 kV field-emission analytical scanning transmission electron microscope several design concepts, including: increasing the collection angle for a solid state detector up to 0.3 sr and the X-ray count rate up to 29 000 cps by the use of electrostatic blanking; and simultaneous X-ray collection from two detectors with equivalent view points and increasing the peak–background value for the intrinsic germanium detector up to 6300 [15]. Determinations of the minimum mass fraction (MMF) for Cr measured on a standard thin film showed that an MMF below 0.1–0.05 wt.% is possible. With dedicated STEM instruments, ultra-sensitive analysis of few atoms [13] and at subnanometer lateral resolution has been already realized [14, 19]. Due to the complementary nature of the information obtained, the simultaneous Z-contrast high-resolution imaging and X-ray and energy-loss spectroscopy in the dedicated analytical electron microscope provide a powerful tool for gaining deeper insight into the fundamental correlations between the atomic and electronic structure of materials at the atomic level [20].

### 2.5.1.9 Multipurpose Electron Probe Analytical Systems

Conventional EPMA utilizes electron excitation for carrying out chemical analysis. However, it is possible to use the electron gun in order to generate an X-ray source for exciting X-rays of sufficient intensity from the specimen; that is, X-ray fluorescence analysis which allows detection limits below 1 ppm to be obtained in an



**Figure 10.** High-resolution image (a) and X-ray energy-dispersive spectra of molybdenum-sulphide fullerene (b) surrounding a niobium-sulphide core (c). Courtesy of Dr M. Otten, Philips Electron Optics, The Netherlands.

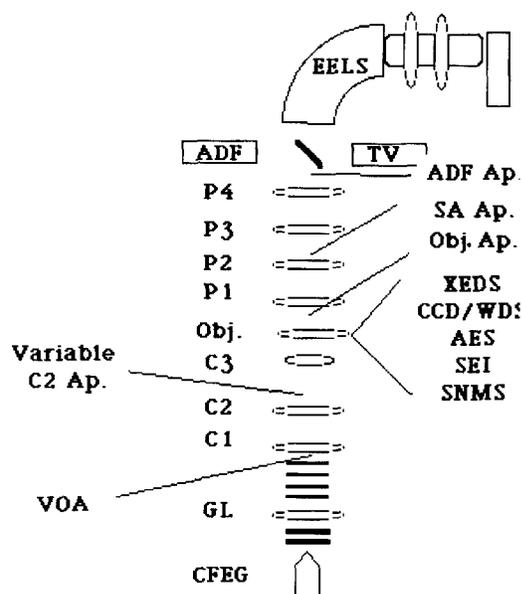
instrument supplied with an energy-dispersive spectrometer [23]. A microanalyzer may also be equipped with a cylindrical mirror Auger-electron spectrometer [21]; this, however, needs an ultrahigh vacuum. Auger-electron microanalyzers, in which the 1–10 keV electron gun is incorporated in the inner cylinder of the spectrometer, can work in the scanning mode so that an image of the surface is formed with secondary electrons or an element-distribution map, especially, of light elements is generated using Auger electrons. Digital image processing of multiple detector signals (secondary electrons, backscattered electrons, sample current, Auger electrons, X-rays, cathodoluminescence and/or light

image) and computer control with the aid of a powerful multitasking workstation, conversion and storage of data, on-line processing for immediate interpretation of images and spectra, and feedback to the instrument increase significantly the capability of modern analyzers.

Utilizing the scanning beam facilities of the electron probe X-ray microanalyzer or scanning electron microscope, panchromatic cathodoluminescence imaging and cathodoluminescence emission spectroscopy may be easily combined with X-ray energy- and wavelength-dispersive spectroscopies (EDS and WDS), and X-ray mapping and electron imaging to perform comprehensive spatially resolved

microanalysis of point defects in minerals and ceramics and dopant impurities in semiconductors [24]. The environmental scanning electron microscope, the natural extension of a conventional scanning electron instrument that can operate from high vacuum up to a pressure level which can maintain fully wet specimens, is promising for microanalytical studies of uncoated insulating specimens by EDS and cathodoluminescence in conjunction with morphological characterization by secondary electron and backscattered electron imaging [25].

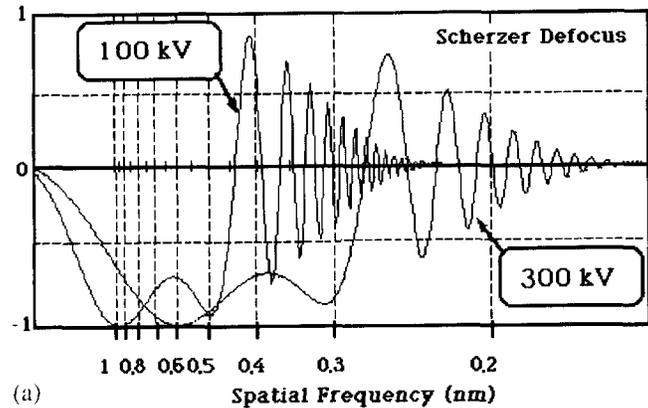
Recently an advanced 300 kV field-emission analytical electron microscope (Argonne National Laboratory) has been designed to attain the best possible analytical sensitivity, resolution, and versatility for EDS, EELS, Auger electron spectroscopy (AES), selected area electron diffraction (SAED), convergent beam electron diffraction (CBED), scanning transmission electron diffraction (ED), and reflected high energy ED consistent for state of the art materials research and still provide moderate imaging capabilities in conventional transmission electron microscopy (high-resolution electron microscopy (HREM)) (CTEM), STEM, and SEM modes. Basically, the system shown in Fig. 11 comprises a conventional field-emission gun with gun lens, a triple condenser, objective, and quadrupole projector. The expected gun brightness is approximately  $4 \times 10^9 \text{ \AA cm}^{-2} \text{ sr}^{-1}$  at 300 kV and the nominal image resolutions in both TEM and STEM bright-field modes should be better than 0.3 nm (point-to-point), and better than 0.2 nm in the high resolution annular dark-field STEM mode. Figure 12a shows plots of a calculated contrast transfer function for



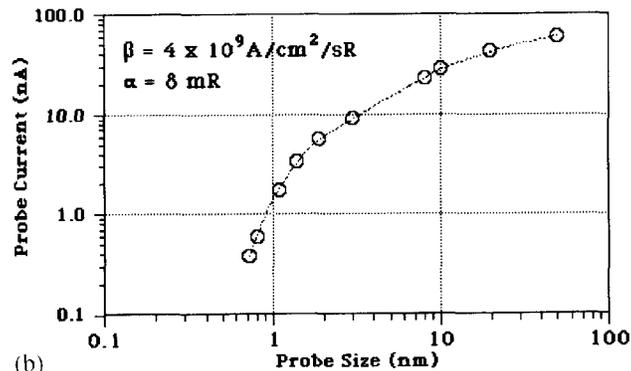
**Figure 11.** Schematic diagram of the electron optics of an advanced universal analytical electron microscope. CFEG, conventional field-emission gun; GL, gun lens; VOA, virtual objective aperture; C1, C2, C3, condenser lenses; Variable C2 Ap., variable C2 condenser apertures; Obj., objective lens; Obj. Ap., objective aperture; P1, P2, P3, P4, projective lenses; SA Ap., selected-area aperture; ADF, annual dark-field detector; EELS, serial/parallel EELS system; TV, TV video camera; XEDS, windowless X-ray energy-dispersive detector; CCD/WDS, slow scan CCD camera/WDS; AES, conventional hemispherical Auger spectrometer; SEI, secondary electron image detector; SNMS, mini secondary neutral mass spectrometer. Courtesy of Dr N. Zaluzec, Argonne National Laboratory, USA.

the objective lens at 300 and 100 kV, while Fig. 12b shows plots of the calculated probe current/size relationships at 300 kV. The four projectors allow imaging of the probe, which is less than 0.2 nm in diameter, under all conditions.

Specimen rotation about the primary tilt axis is limited to  $\pm 85^\circ$ . Image detection is accomplished by using a video TV camera in the CTEM and bright-field/annular dark-field STEM modes, using one of four operator-selected YAG screens, the signals



(a)



(b)

**Figure 12.** (a) Calculated contrast transfer function at 300 and 100 kV (Scherzer defocus,  $C_s = 4.4$  mm), and (b) probe size/current relationship at 300 kV for the advanced universal Auger electron microscope. Courtesy of Dr N. Zaluzec, Argonne National Laboratory, USA.

from which are flash digitized (8, 16, 32 bit) and routed to two independent frame stores ( $2K \times 2K \times 8$  bit). A comprehensive specimen-preparation chamber is interfaced directly to the column; this allows complete extensive cleaning, characterization, and preparation of the specimen surface by the use of sputter cleaning, a mini secondary ion mass spectrometer system, thin film evaporator, etc. Two  $400$  and two  $601s^{-1}$  ion getter pumps, four titanium sublimators, and one turbomolecular pump comprises the evacuation system for an ultrahigh vacuum environment ( $2 \times 10^{-11}$  to  $2 \times 10^{-9}$  torr).

Multilayer low/high  $Z$  material combinations have been employed in both beam- and nonbeam-defining apertures and at all

critical surfaces to minimize potential sources of uncollimated hard X-rays which give rise to the hole count phenomenon. The windowless energy-dispersive system has been optimized to maximize the subtending solid angle and allows retraction along a direct line-of-sight path to the specimen. This allows the instrument to achieve a continuously variable solid angle up to a maximum of  $0.3$  sr. A hemispherical Auger spectrometer with extraction optics is interfaced to the center of the objective lens and both serial and parallel EELS detection capabilities will be present. In addition, secondary and Auger electron spectrometers utilizing parallelizer optics are installed within the objective prefield. This universal AEM

**Table 2.** Comparison of WDS and EDS [1, 2, 11, 12, 15]

Item	WDS	EDS
Basic method	Wavelength dispersion by diffracting crystals	Energy dispersion by solid state detector
Available elemental range	${}^4\text{Be}$ to ${}_{92}\text{U}$	$({}^5\text{B})_{12}\text{Na}$ to ${}_{92}\text{U}$
Resolution (eV), ( $\text{MnK}_{\alpha}$ , FWHM (nm))	$\sim 20$ ( $\sim 0.7 \times 10^{-3}$ )	100–150 ( $\sim 6 \times 10^{-3}$ )
Probe current range (Å)	$10^{-9}$ – $10^{-5}$	$10^{-12}$ – $10^{-9}$
Detection limit (ppm)	50–100	1500–2000
Number of simultaneously analyzed elements	1	up to 25
X-ray acquisition rate (cps)	$10^2$ – $10^5$	$10^3$ – $3 \times 10^4$

system may be controlled either directly by the operator using conventional multi-function dials and switches, or the PC and a mouse directed interactive graphical user interface thus providing telepresence microscopy remote control and operation over networks.

### 2.5.1.10 X-Ray Emission Spectrometry

Analysis of X-rays emitted from the specimen involves measurements of their energies (or wavelengths) and intensities. There are two main commercially available spectroscopic systems which can be coupled to the scanning electron or scanning transmission electron microscope

to record X-ray quanta (i.e. WDS and EDS). Their principal features are compared in Table 2.

### 2.5.1.11 Wavelength-Dispersive Spectrometry

WDS utilizes the Bragg reflection of X-ray emission dispersed by an analyzing curved crystal from its lattice planes ( $2d \sin \theta = n\lambda$ ). Crystals with different lattice spacing (0.4–10 nm) (Table 3) are used to analyze the whole wavelength range from below 0.1 nm ( $\text{UL}_{\alpha} \simeq 0.091$  nm) to above 11 nm ( $\text{BeK}_{\alpha} \simeq 11.3$  nm). The spot irradiated by the electron beam on the specimen acts as an entrance slit, while the analyzing crystal and the exit slit are

**Table 3.** Analysis range of analyzing crystals

Crystal (abbr.)	$2d$ (nm)	Wavelength range (nm)	10 Na	20 Ca	30 Zn	40 Zr	50 Sn	60 Nd	70 Yb	80 Hg	90 Th	
STE	10.04	2.22–9.3	${}^5\text{B}$ to ${}^8\text{O}$ , ${}^{16}\text{S}$ to ${}^{23}\text{V}$									
TAP	2.576	0.569–2.38	${}^8\text{O}$ to ${}^{15}\text{P}$ , ${}^{23}\text{Cr}$ to ${}^{41}\text{Nb}$ ${}^{46}\text{Pd}$ – ${}^{79}\text{Au}$									
PET	0.8742	0.193–0.81	${}^{13}\text{Al}$ to ${}^{25}\text{Mn}$ , ${}^{36}\text{Kr}$ – ${}^{65}\text{Tb}$ , ${}^{70}\text{Yb}$ – ${}^{92}\text{U}$									
LIF	0.4027	0.0889–0.373	${}^{19}\text{K}$ – ${}^{37}\text{Rb}$ ${}^{48}\text{Cd}$ – ${}^{92}\text{U}$									
MYR	8.0	1.77–7.41	${}^5\text{B}$ to ${}^9\text{F}$ , ${}^{17}\text{Cl}$ to ${}^{25}\text{Mn}$									
LDEI <sup>a</sup>			${}^6\text{C}$ to ${}^9\text{F}$									
LDE2, <sup>a</sup> LDEN			${}^5\text{B}$ to ${}^8\text{O}$									
LDEB <sup>a</sup>			${}^4\text{Be}$ to ${}^5\text{B}$									

<sup>a</sup>Superlattice crystals.

mounted on a Rowland circle. The lattice planes of the crystal are bent so that their radius is  $2R$  and the crystal surface is ground to a radius  $R$ . Focusing allows one to obtain better separation of narrow characteristic lines and a solid angle of collection of nearly  $10^{-3}$ – $10^{-2}$  sr.

Behind the slit, X-ray photons are recorded by a proportional counter and their energy is converted to a voltage pulse. The detection efficiency of the Bragg reflection and of the proportional counter is about 10–30% [1, 2]. The number of electron-ion pairs generated in the counter is proportional to the quantum energy  $E = h\nu$ . The pulses are further amplified, discriminated in a single-channel analyzer, and counted by a scaler. The pulse intensity per second is indicated by a ratemeter. Advanced counters with ultrathin mylar windows are available to detect the weak  $K_{\alpha}$  lines of light elements ( $4 \leq Z \leq 11$ ) more efficiently than in EDS. WDS systems offer much better energy (wavelength) resolution and higher count rates ( $>50\,000$  cps) than EDS ones (see Fig. 3 and Table 2). For scanning across a chosen spectral region, the counter and the analyzing crystal should be moved by a pivot mechanism. Usually, X-ray micro-analyzers are equipped with several wavelength-dispersive spectrometers which enable different wavelengths to be recorded simultaneously.

### 2.5.1.12 Energy-Dispersive Spectrometry

With the energy-dispersive spectrometer a solid-state detector is positioned before the dispersing system to collect the

distribution of X-ray emission over a wide energy range (0.1–40 keV for a conventional Si(Li) detector, and even up to 80 keV for an intrinsic Ge detector) [10]. Dispersion of the signal detected then takes place by following processing using pulse-height amplification, pile-up rejection of possible coincidence of pulses, and sorting by a multichannel analyzer which relates the measured pulse height to the energy of the incoming photon. The entire assembly of a conventional detector including a field-effect-transistor preamplifier is cryo-cooled under vacuum at near liquid nitrogen temperature to minimize any thermally induced signals. The detection efficiency of an Si(Li) detector is nearly 100% over the range 3–15 keV [1, 2]. The decrease at low energies is caused by the absorption of X-rays in the thin Be window separating the high vacuum microscope column from the detector. Windowless detectors can record  $K_{\alpha}$  quanta from light elements up to  ${}_5\text{B}$ . At energies higher than 15 keV the decrease in efficiency is caused by the increasing probability of penetrating the sensitive layer of the detector without photoionization.

Recent developments in EDS detector fabrication have resulted in significant characteristic improvements in low-energy performance and sensitivities. Modern EDS detectors can also be exploited without permanent cooling and stay at room temperature without any degradation in performance. Newly designed ones are electrically cooled with an incorporated Peltier device or need no cooling ever, although still at the cost of some decrease in resolution.

Unlike WDS, where the irradiated point has to be adjusted on the Rowland circle, in EDS there is no need for any

mechanical adjustment, and therefore the spectrometer can be used much more effectively for profile or area analyses by scanning large and/or rough specimens. As an energy-dispersive spectrometer occupies a smaller space than a wavelength-dispersive system, it is commonly used in analytical SEM and STEM. A further advantage of EDS is that most of the characteristic X-ray lines can be recorded simultaneously in a short time. For this reason, X-ray microanalyzers that work with WDS are often also equipped with an electron-dispersive spectrometer under the control of the same computer, thus forming an integrated WDS-EDS system in order to provide more rapid and accurate assessments of the elemental composition of the specimen. This includes: simultaneous data acquisition of up to 13 elements with the combination of five wavelength-dispersive and one energy-dispersive spectrometer, more than twice the number measurable with WDS alone; microanalysis of light elements and micro-volume elements with a large beam current of  $10^{-9}$ – $10^{-5}$  A (WDS) and of heavy elements of a few percent under the same conditions as for SEM observations at a current of  $10^{-12}$ – $10^{-9}$  A (EDS) for increasing the total analysis efficiency; and preliminary evaluation of beam-sensitive samples with EDS.

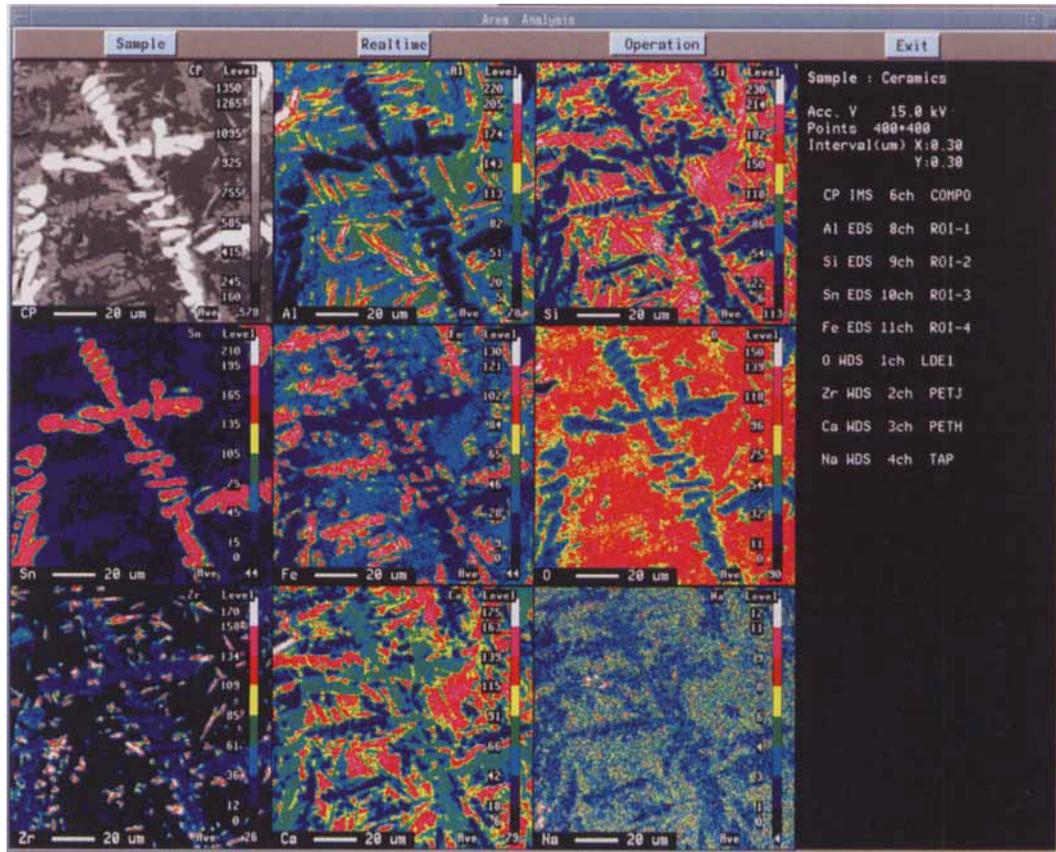
### 2.5.1.13 X-Ray Mapping

X-ray maps (Fig. 13) provide valuable information on the two-dimensional elemental distributions over bulk or thin samples. X-ray mapping can be undertaken with both EDS and WDS, where

incoming counts for elements of interest are fed back into the SEM or STEM system. The X-ray analyzer takes control over the electron beam via a special interface, driving the beam around in a rectangular frame and collecting the X-ray emission for each pixel separately. An external beam deflection interface allows X-ray mapping by beam rocking in a transmission electron microscope using standard hardware and software [26].

Color compositional maps provide multielement spatially resolved X-ray analysis within the chosen region of the sample. However, it is often difficult to visualize the compositional ranges and resulting inter-element correlations from the map, particularly for minor or trace constituents. As an alternative method, a composition–composition histogram displaying the numerical relationship between concentrations of the components in various points of the sample can be proposed [12].

The SEM images in secondary and/or backscattered electrons are normally used to choose regions for subsequent X-ray mapping. If features of interest are not differentiated in the monochrome, black and white electron image, they can easily be overlooked for further analysis. Using the X-ray spectrum detected by EDS, one can construct a color response directly related to the underlying elemental composition where the spectrum from each compound has a characteristic color [27]. Thus, the topography and elemental composition of a specimen are compressed into a single view. Furthermore, using this new technique it is not necessary to monitor the X-ray spectrum, set windows, or collect X-ray maps. Compositional data, even for samples with a rough surface, are automatically acquired into the SEM images



**Figure 13.** Wavelength- and energy-dispersive combined  $140 \times 140 \mu\text{m}^2$  area analysis of an Na-Al-Si-Ca-Fe-Zr-Sn-O ceramic ( $400 \times 400$  points). X-ray maps of eight constituent elements and a backscattered electron image ('compo' mode, CP) recorded at an accelerating voltage of 15 keV. Mapping was performed simultaneously by EDS (Al, Si, Fe, Sn) and WDS (O, Na, Ca, Zr). Courtesy of JEOL.

because the colors displayed always relate to the sample composition.

## 2.5.2 Introduction to Quantitative X-Ray Scanning Microanalysis

In order to quantify X-ray spectra, the measured intensity of a particular characteristic X-ray line from the specimen

should be compared with that from a reference standard of known composition. Hence by keeping instrumental settings (probe current, high voltage, detector efficiency, etc.) constant while the X-ray intensity readings are being taken, one can consider only the ratio of the X-ray intensities of line  $i$  of an element  $a$  ( $k_a^i$ ) measured in the specimen  $I_{\text{sp}}^i$  and the standard  $I_{\text{st}}^i$ :

$$\frac{I_{\text{sp}}^i}{I_{\text{st}}^i} = k_a^i \quad (5)$$

where the intensity ratio can be related to the mass concentration  $c_a$  of the analyzed element as  $k_a^i = f(c_a)$  and  $\sum c_a = 1$ . It is of course necessary to apply dead-time corrections related to the spectrometer.

An accurate quantitative evaluation can be performed if the  $k_a^i$  ratio is corrected for various effects: the atomic number correction for the differences between the electron scattering and penetration in the sample and the standard; the absorption correction for the difference in the absorption of the emitted X-rays as they pass through the sample or standard; and the fluorescence correction for the X-ray fluorescence generated by the X-ray emission in the specimen and in the standard.

### 2.5.2.1 ZAF Method

Following from Eq. (2), the ratio of X-ray intensities emitted from element  $a$  in the specimen and in the pure standard is given by

$$\frac{I_{sp}^a}{I_{st}^a} = \frac{\phi(\rho z)_{sp}^a \left[ \int_0^\infty \phi(\rho z) d\rho z \right]_{sp}^a}{\phi(\rho z)_{st}^a \left[ \int_0^\infty \phi(\rho z) d\rho z \right]_{st}^a} \times \frac{f(\chi)_{sp}^a (1 + \gamma + \beta)_{sp}^a}{f(\chi)_{st}^a (1 + \gamma + \beta)_{st}^a} \quad (6)$$

The first term in Eq. (6), which corresponds to the ratio of the emission of element  $a$  from an isolated thin layer of mass thickness  $d\rho z$  in the sample and in the standard, respectively, is proportional to the number of ionizations produced by the electron beam:

$$\frac{\phi(\rho z)_{sp}^a}{\phi(\rho z)_{st}^a} = \frac{Q\omega(Nc_a/A) d\rho z}{Q\omega(N/A) d\rho z} = c_a \quad (7)$$

where  $Q$  is the ionization cross-section of atoms  $a$ ,  $\omega$  is the X-ray emission yield,  $N$  is Avogadro's number,  $A$  is the atomic weight of  $a$ , and  $c_a$  is the mass concentration of element  $a$  in the specimen. Then one can easily deduce from Eq. (6) that the next three terms may be used to represent the atomic number, absorption, and fluorescence (ZAF) correction factors  $k_z$ ,  $k_A$ , and  $k_F$ , respectively. The corresponding corrections are considered in the ZAF method as independent multiplicative terms to the  $k$  ratios:

$$k_a^i = k_z k_A k_F c_a \quad (8)$$

### 2.5.2.2 Atomic Number Correction

The atomic number correction should be applied to the  $k$  ratio to compensate for the difference between the electron retardation and electron backscattering in the sample and standard.  $k_z$  for element  $a$  can be written as:

$$k_z^a = \frac{R_{sp}^a \int_{E_c}^{E_0} (Q/S)_{sp}^a dE}{R_{st}^a \int_{E_c}^{E_0} (Q/S)_{st}^a dE} \quad (9)$$

where the terms  $R$  and  $S$  refer to the electron backscattering and the electron stopping power, respectively, and  $Q$  is the ionization cross-section. The integral limits are from the incident electron beam energy  $E_0$  to the critical excitation band  $E_c$  for the chosen X-ray line of  $a$ . The electron stopping power, defined as the energy lost per unit electron path length in material of density  $\rho$ ,  $S = (-1/\rho) dE/dx$ , may be approximated as

follows

$$S = 78\,500 \frac{Z^a}{A^a} \frac{1}{E} \ln \frac{1.166E}{J^a} \quad (10)$$

where  $J = [10.04 + 8.25 \exp(-Z/11.22)]Z$  [12].

The expressions for the ionization cross-section  $Q$  have the general form

$$Q = C \frac{\ln U}{UE_c^2} \quad (11)$$

where  $C$  is a constant and  $U = E_0/E_c$  is the overvoltage.

The electron backscattering factor  $R = (I_t - I_b)/I_t$  defined as the fraction of total generated X-ray intensity excluding loss to backscatter electrons may be calculated using empirical expressions [28]. In a multielement system, the factor  $R$  for element  $i$  may be derived from the equation

$$R_i = \sum_j C_j R_{ij} \quad (12)$$

where  $i$  represents the element being measured and  $j$  denotes the elements in the specimen including  $i$ ;  $R_{ij}$  is the backscatter correction for element  $i$  in the presence of element  $j$ .

Several tabulations [29, 30] and fitted polynomial expressions [31, 32] are available to estimate values of  $R$  for various elements as a function of  $Z$  and  $U$ .

### 2.5.2.3 X-Ray Absorption Correction

The absorption correction factor  $k_A = [f(\chi)_{sp}^a]/[f(\chi)_{st}^a]$  reflects the attenuation of the X-ray intensity measured by the detector. In general, the correction should

be considered because the specimen and the standard are not identical. The basic formulation for the absorption term  $f(\chi)$  is given by the Philibert–Duncumb–Heinrich equation [33, 34]:

$$\frac{1}{f(\chi)} = \left(1 + \frac{\chi}{\sigma}\right) \left(1 + \frac{h}{1+h}\right) \frac{\chi}{\sigma} \quad (13)$$

where  $\sigma = (4.5 \times 10^5)/(E_0^{1.65} - E_c^{1.65})$  is Lenard's constant and  $h = 1.2(A/Z^2)$ .

The variable  $h$  is dependent on the composition and must be averaged for the various elements present in multielement specimens as follows:

$$h_i = \sum_j c_j h_j$$

In addition, the mass absorption coefficient  $\mu/\rho$  for the characteristic line of element  $i$  must be the weighted sum over all elements:

$$\left(\frac{\mu}{\rho}\right)_i^{sp} = \sum_j c_j \left(\frac{\mu}{\rho}\right)_i^j$$

Calculation of  $f(\chi)$  from Eq. (13) is most accurate for values greater than 0.7.

### 2.5.2.4 Fluorescence Corrections

The characteristic fluorescence correction is necessary, when the energy  $E_j$  of the characteristic X-ray line from an element  $j$  is greater than the critical excitation energy  $E_{c,i}$  for an element  $i$  and this difference  $E_j - E_{c,i} < 0.5$  keV. This may result in excitation of the characteristic emission of element  $i$  by the corresponding characteristic line of element  $j$ . The basic formulation of the characteristic fluorescence

correction is given by

$$F_i = \frac{\left(1 + \sum_j I_{i,j}^f/I_i\right)_{\text{sp}}}{\left(1 + \sum_j I_{i,j}^f/I_i\right)_{\text{st}}} \quad (14)$$

where  $I_{i,j}^f/I_i$  is the ratio of the characteristic intensity of element  $i$  excited by fluorescence to the electron-excited intensity. The total correction should be summed over all the elements in the specimen. In practice, when the standard is either a pure element or there is no significant fluorescence of element  $i$  by other elements in the standard, Eq. (14) can be simplified

$$F_i = \left(1 + \frac{\sum_j I_{i,j}^f}{I_i}\right)_{\text{sp}} \quad (15)$$

In addition, the calculation of the continuum fluorescence is relatively complicated, involving integration over the range of  $E_0$  to  $E_c$  for each element in the sample. When  $f(\chi) > 0.95$ ,  $c_i < 0.5$ , and  $\bar{Z}_{\text{sp}} \neq \bar{Z}_{\text{st}}$ , the correction can be as large as 2–4%, it should be included for highest accuracy [35].

### 2.5.2.5 $\phi(\rho z)$ Methods

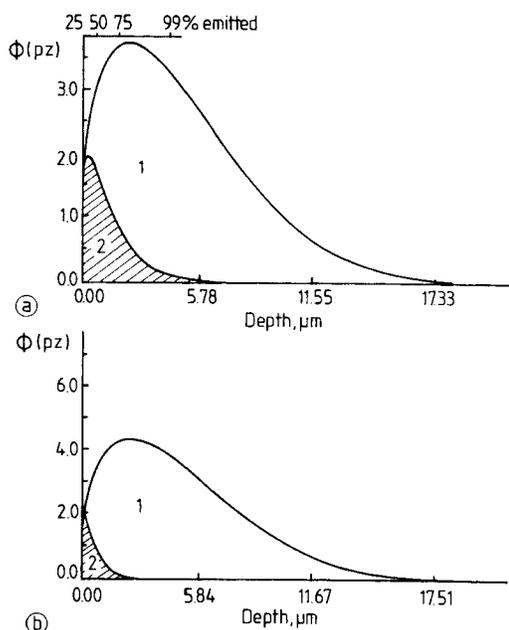
Quantitative microanalysis can be carried out more suitably by treating the atomic number and absorption factors together rather than as separate entities. The correction procedure based on the integration of X-ray distributions,  $\phi(\rho z)$  curves, is attractive for the analysis of low-energy X-ray lines, where  $f(\chi)$  is much less than 0.7 and the accuracy of the ZAF method is low. However, results of the analysis based on the  $\phi(\rho z)$  methods are dependent on the

accuracy of description of the experimental  $\phi(\rho z)$  curves (especially in the case of systems with unknown X-ray depth distributions) and/or fitting to the corresponding results of Monte Carlo simulation as well. Several researchers have proposed a number of different parametrizations of  $\phi(\rho z)$  curves [1, 36–39].

Among the proposed methods the Packwood–Brown model [40] is now widely accepted. This model, which is based on numerous experimental data, utilizes a modified surface-centered Gaussian function to describe the shape of  $\phi(\rho z)$  curves. The Gaussian nature of the curve implies some random distribution of electron-scattering events. The surface modification was applied to take into account that the incoming accelerated electrons require some finite distance range before becoming randomized. The general expression for  $\phi(\rho z)$ , which is drawn in Fig. 14 for the particular case of the AgL and BrL characteristic irradiation generated in a silver halide tabular crystal, is as follows:

$$\begin{aligned} \phi(\rho z) = & \gamma[\exp(-\alpha^2(\rho z)^2)] \\ & \times \left[1 - \left(\frac{\gamma - \phi(0)}{\gamma}\right)\exp(-\beta\rho z)\right] \end{aligned} \quad (16)$$

The Gaussian function is modified by a transient function to model an increase of the X-ray production away from the near-surface region. The coefficient  $\alpha$  relates to the width of the Gaussian function and  $\gamma$  relates to its amplitude. The term  $\beta$  in the transient is related to the slope of the curve in the subsurface region; this means the rate at which the focused electron beam is randomized through scattering in the sample. The intercept  $\phi(0)$  is related to the surface ionization potential. The  $\alpha$ ,  $\beta$ ,



**Figure 14.**  $\phi(\rho z)$  functions for AgL (a) and BrL (b) irradiation produced (1) and emitted (2) under an 80-keV electron beam in an  $\text{AgBr}_{0.95}\text{I}_{0.05}$  tabular crystal; calculated using the PROZA program [41].

and  $\gamma$  parameters are expressed in terms of several different experimental terms (elemental composition of the sample, incident electron energy, atomic number, etc.). Bastin et al. [41, 42] and Armstrong [43] have modified Eq. (16) to optimize the fit for different sets of experimental data.

### 2.5.2.6 Standardless Analysis

In this approach the measured standard intensities are substituted by calculated ones based on atomic data and empirical adjustments to experimental data. The better quantitative procedures were developed in the last decade in order to validate better standardless calculations. In particular, with the K lines a relative accuracy of 1–5% is reached [44]. Some

uncertainties exist with the L lines, and especially with the M lines where atomic data sets are still incomplete or not accurate enough. Further study of the effect of the incompletely filled inner shells is necessary to obtain better atomic data.

The X-ray intensity generated in a bulk sample (standard) for the simpler case of the K line [12, 45] may be expressed as

$$I = g\epsilon p\omega f(\chi) \frac{N}{A} R \int_{E_0}^{E_c} \frac{Q}{dE/d\rho s} dE \quad (17)$$

where  $g$  is a normalization factor which depends on the experimental parameters (beam current, acquisition time, solid angle of the detector, etc.),  $\epsilon$  is the detector efficiency,  $p$  is the transition probability,  $\omega$  is the fluorescence yield,  $f(\chi)$  is the absorption correction defined by Eq. (3),  $R$  is the backscatter factor, and  $Q$  is the ionization cross-section given in general form by Eq. (12). The additional intensity induced by the secondary X-ray fluorescence is not considered here. When calculating the number of ionizations for the L and M shells, both direct ionizations induced by the bombarding electrons and indirect ones caused by the nonradiative Coster–Kronig transitions between subshells prior to the emission of the X-ray line must also be taken into account. The Si(Li) detector efficiency may be given by

$$\begin{aligned} \epsilon = & \left\{ 1 - \exp \left[ - \left( \frac{\mu}{\rho} \right)_{\text{Si}} \rho d_{\text{Si,act}} \right] \right\} \\ & \times \exp \left[ - \left( \frac{\mu}{\rho} \right)_{\text{Be}} \rho d_{\text{Be}} \right] \\ & \times \exp \left[ - \left( \frac{\mu}{\rho} \right)_{\text{Au}} \rho d_{\text{Au}} \right] \\ & \times \exp \left[ - \left( \frac{\mu}{\rho} \right)_{\text{Si}} \rho d_{\text{Si,dead}} \right] \end{aligned} \quad (18)$$

where  $Be$ ,  $Au$ ,  $Si_{act}$  and  $Si_{dead}$  refer to the beryllium window, the gold contact and the active and dead silicon layers of the detector, respectively.

The  $k$  ratios obtained by comparing experimental spectra with calculated standard intensities may be used as concentrations of first approximation following the iteration correction [46].

### 2.5.2.7 Analysis of Thin Films and Particles

Thin films and particles are conveniently analyzed in AEM by using high-energy electron beams with accelerating voltages not less than 80 keV. Under these conditions the effect of electron backscattering is minimal and electron energy losses in the specimen are rather small. For specimens in which the thin-film criterion [47] is fulfilled (i.e. the X-ray absorption <3% and/or fluorescence <5%), the atomic-number correction can be neglected and the characteristic X-ray intensity of element  $i$  produced may be expressed [48] as:

$$I_i = K \epsilon_i c_i \omega_i Q_i a_i \frac{t}{A_i} \quad (19)$$

where  $K$  is a constant,  $\epsilon_i$  is the efficiency of the EDS detector to detect X-ray emission from element  $i$ ,  $c_i$  is the mass concentration of element  $i$ ,  $\omega_i$  is the fluorescence yield for the analyzed X-ray peak of element  $i$ ,  $a_i$  is the measured fraction of total series intensity,  $t$  is the film thickness, and  $A_i$  is the atomic weight of element  $i$ . To avoid the influence of the substrate, the particles should be deposited onto thin-film-supports of low atomic number such as carbon. Besides, the particles should be

scanned in a raster because, in general, the uniformity of the beam density cannot be provided when the probe diameter is comparable to or larger than the particle size.

In practice, it is suitable to measure the relative concentration of element  $i$  to another element  $j$  which may be given by [49]

$$\frac{c_i}{c_j} = k_{ij} \frac{I_i}{I_j} \quad (20)$$

where  $c_i$  and  $c_j$  are the mass concentrations of elements  $i$  and  $j$ , respectively,  $I_i$  and  $I_j$  are the corresponding X-ray line intensities, and  $k_{ij}$  is the Cliff-Lorimer sensitivity factor. The sensitivity factors can be related to an inner standard element as

$$k_{ij} = \frac{k_{i,is}}{k_{j,is}} = \frac{k_i}{k_j} \quad (21)$$

where  $k_{i(j),is} = k_{i(j)}$  are the relative sensitivity factors. The index is referred usually to silicon and iron commonly used as internal standard elements. The values of the sensitivity factors for K, L, and M line emission from various elements have been reported by several researchers [50–52].

Equation (20) may be expanded to take into account the effects of absorption and fluorescence:

$$\frac{c_i}{c_j} = k_{ij} \frac{I_i}{I_j} k_a \frac{1}{1 + I_f^i/I_0^i} \quad (22)$$

where  $I_f^i$  and  $I_0^i$  are the fluoresced and primary X-ray intensities of  $i$ , respectively, and  $k_a$  is the absorption factor given by [12]

$$k_a = \frac{(\mu/\rho)_{sp}^i}{(\mu/\rho)_{sp}^j} \cdot \frac{1 - \exp[-\chi^i(\rho t)]}{1 - \exp[-\chi^j(\rho t)]} \quad (23)$$

The parameterless extrapolation method proposed by Van Cappellen [53]

is based on measurements of the relative intensities  $I_i/I_j$  at various thicknesses following extrapolation to zero thickness ( $I_i = 0$ ) by plotting  $I_i/I_j$  versus  $I_i$ . Then, if the secondary emission of the elements is much smaller than the primary one, the mass concentration ratio can be determined as follows:

$$\frac{c_i}{c_j} = k_{ij} \lim_{I_i \rightarrow 0} \frac{I_i}{I_j} \quad (24)$$

### 2.5.3 Conclusions

Since the electron probe X-ray microanalyzer and its basic concepts were first developed in the 1950s by Raymond Castaing, scanning microanalysis using modern X-ray microanalyzers and analytical scanning electron and scanning transmission electron microscopes equipped with energy- and/or wavelength-dispersive spectrometers has been successfully applied to the elemental and compositional characterization of advanced materials in a variety of fields of materials science (metallurgy, ceramics, electronics, geology, etc.). The last decade has shown considerable progress in the theoretical description of fundamental electron–solid interactions and methodological development in scanning microanalysis, especially on fully quantitative and standardless procedures and automatic instrumentation, low-voltage microanalysis of light elements, and super-high sensitive analysis at nano- and subnanometer lateral resolution.

Nowadays, there is a tendency to develop integrated multipurpose AEM systems based on scanning electron

and/or scanning transmission electron microscopes equipped with multiple detector systems and to detect effects of inelastic interactions of electron beams within the object (WDS, EDS, EELS and electron spectroscopic imaging, ESI, AS/SAM). However, standard modes, in particular, electron and X-ray diffraction methods (SAED, CBED, electron channelling, Kikuchi patterns), dark field, and stereo-observations are of course still widely used. As a result, conventional SEM, STEM, and EPMA are evolving from instruments for observation or for elemental analysis only, with moderate imaging capabilities, into universal analytical electron microscopy systems for obtaining and treating various types of information on nanometer- and micrometer-sized parts of the object to be examined: its morphology, topography, crystal and defect structure, elemental composition, and electronic state of the matter.

The versatility of advanced electron probe analysis techniques combining high-resolution imaging in various modes, and the powerful analytical facilities and variety of signals and contrast effects available, establish them as outstanding tools for universal applications.

### Acknowledgements

The authors are grateful to Dr E. Van Cappellen, Philips Electron Optics, The Netherlands, and to Dr R. Herstens, JEOL (EUROPE) B.V., Belgium, for providing valuable scientific and technical information used in this paper, to Prof. D. Joy, University of Tennessee, USA, for providing the Monte Carlo simulation programs, and to R. Nullens, University of Antwerp (U.I.A.), Belgium, for assistance in recording of X-ray spectra.

## 2.5.4 References

- [1] V. D. Scott, G. Love, S. J. B. Reed, *Quantitative Electron-Probe Microanalysis*, 2nd edn., Ellis Horwood, New York **1995**, 19, 28, 37, 45, 61.
- [2] L. Reimer, *Transmission Electron Microscopy. Physics of Image Formation and Microanalysis*, 2nd edn., Springer Verlag, Berlin **1989**, 3, 5, 7, 221, 376, 404, 425, 428.
- [3] F. Hofer in *Energy-Filtering Transmission Electron Microscopy* (Ed.: L. Reimer), Springer Verlag, Berlin **1995**, Chap. 4.
- [4] D. E. Newbury, *Microbeam Anal.* **1992**, 1, 9.
- [5] G. Remond, C. Gilles, M. Fialin, O. Rouer, R. Marinenko, R. Myklebust, D. Newbury, *EMAS '95 4th European Workshop on Modern Developments in Applied Microbeam Analysis*, EMAS, St Malo, France **1995**, 117.
- [6] D. Timotijevic, M. K. Pavicevic, *Mikrochim. Acta* **1992**, 12(Suppl.), 255.
- [7] J. L. Pouchou, *EMAS '95 4th European Workshop on Modern Developments in Applied Microbeam Analysis*, EMAS, St Malo, France **1995**, 95.
- [8] D. C. Joy, *Microbeam Anal.* **1992**, 1, 19.
- [9] N. D. Browning, S. J. Pennycook, *Microbeam Anal.* **1993**, 1, 19.
- [10] D. B. Williams, J. I. Goldstein, C. E. Fiori in *Principles of Analytical Electron Microscopy* (Eds.: D. C. Joy, A. D. Romig, J. I. Goldstein), Plenum Press, New York **1986**, Chap. 4.
- [11] J. I. Goldstein, D. B. Williams, *Microbeam Anal.* **1992**, 1, 29.
- [12] J. A. Small in *Handbook of X-ray Spectrometry: Methods and Techniques* (Eds.: R. E. Van Grieken, A. A. Markowicz), Marcel Dekker, New York **1993**, Chap. 12.
- [13] G. Cliff, R. F. Devenish, P. J. Goodhew, R. J. Keyse, G. W. Lorimer, *Proc. 13th International Congress on Electron Microscopy*, Paris **1994**, Vol. 1, 719.
- [14] Y. Bando, *Proc. 13th International Congress on Electron Microscopy*, Paris **1994**, Vol. 1, 591.
- [15] C. E. Lyman, J. I. Goldstein, D. B. Williams, D. W. Ackland, S. Von Harrach, A. W. Nicholls, P. J. Statham, *J. Microsc.* **1994**, 176, 85.
- [16] L. Reimer in *Energy-Filtering Transmission Electron Microscopy* (Ed.: L. Reimer), Springer Verlag, Berlin **1995**, Chap. 1.
- [17] P. E. Batson, *Proc. 13th International Congress on Electron Microscopy*, Paris, **1994**, Vol. 1, 709.
- [18] M. Isaacson, M. Ohtsuki, M. Utlait in *Introduction to Analytical Electron Microscopy* (Eds.: J. J. Hren, J. I. Goldstein, D. C. Joy), Plenum Press, New York **1979**, Chap. 13.
- [19] N. D. Browning, M. F. Chisholm, S. J. Pennycook, *Nature* **1993**, 366, 143.
- [20] S. J. Pennycook, D. E. Jesson, N. D. Browning, M. F. Chisholm, *EMAS '95 4th European Workshop on Modern Developments in Applied Microbeam Analysis*, EMAS, St Malo, France, **1995**, 301.
- [21] M. A. Baker, J. E. Castle in *Materials Science and Technology: a Comprehensive Treatment* (Eds.: R. W. Cahn, P. Haasen, E. I. Kramer), VCH, Weinheim **1994**, Vol. 2B, Chap. 13.
- [22] C. E. Nockolds, *Microbeam Anal.* **1994**, 3, 185.
- [23] I. Pozsgai, *X-ray Spectrosc.* **1994**, 23, 32.
- [24] J. F. Bresse, G. Remond, B. Akamatsu, *EMAS '95 4th European Workshop on Modern Developments in Applied Microbeam Analysis*, EMAS, St Malo, France **1995**, 213.
- [25] G. D. Danilatos, *Mikrochim. Acta* **1994**, 114/115, 143.
- [26] J. M. Brock, *Philips Electron. Optics Bull.* **1994**, 133, 17.
- [27] P. Statham, *EMAS '95 4th European Workshop on Modern Developments in Applied Microbeam Analysis*, EMAS, St Malo, France **1995**, 408.
- [28] W. Reuter, *Proc. 6th International Conference on X-ray Optics and Microanalysis*, University of Tokyo Press, Japan **1972**, 121.
- [29] G. Springer, *Mikrochim. Acta* **1966**, 3, 587.
- [30] P. Duncumb, S. J. B. Reed in *Quantitative Electron-Probe Microanalysis*, NBS Special Publication No. 298, Department of Commerce, Washington, DC **1968**, 133.
- [31] H. Yakowitz, R. L. Myklebust, K. F. J. Heinrich, *NBS Tech. Note 796*, Department of Commerce, Washington, DC **1973**.
- [32] R. L. Myklebust, *J. Phys.* **1984**, 45(Suppl 2), C2-41.
- [33] J. Philibert in *3rd International Congress on X-ray Optics and Microanalysis*, Academic Press, New York **1963**, 379.
- [34] J. I. Goldstein, D. E. Newbury, P. Echlin, D. C. Joy, C. E. Fiori, E. Lifshin, *Scanning Electron Microscopy and X-ray Microanalysis*, Plenum Press, New York **1981**, 312.
- [35] R. L. Myklebust, C. E. Fiori, K. F. J. Heinrich, *NBS Special Publication No. 298*, Department of Commerce, Washington, DC **1968**, 197.
- [36] H. Yakowitz, D. E. Newbury, *SEM I, Proc. 9th Annual SEM Symposium*, IIT Research Institute, Chicago, IL **1976**, 151.
- [37] J. L. Pouchou, F. Pichoir, *J. Phys.* **1984**, 45, C2-17, C2-47.
- [38] J. L. Pouchou, F. Pichoir in *Microbeam Analysis — 1988*, San Francisco Press, San Francisco, CA **1988**, 315.

- [39] J. L. Pouchou, F. Pichoir in *Electron Probe Quantification* (Eds.: K. F. J. Heinrich, D. E. Newbury), Plenum Press, New York **1991**, 31.
- [40] R. H. Packwood, J. D. Brown, *X-ray Spectrom.* **1981**, *10*, 138.
- [41] G. F. Bastin, F. J. J. Van Loo, H. J. M. Heijligers, *X-ray Spectrom.* **1984**, *13*, 91.
- [42] G. F. Bastin, H. J. M. Heijligers, F. J. J. Van Loo, *Scanning* **1986**, *8*, 45.
- [43] J. T. Armstrong, *Microbeam Analysis — 1982* (Ed.: K. F. J. Heinrich), San Francisco Press, San Francisco, CA **1982**, 315.
- [44] J. L. Pouchou, *3rd European Workshop on Modern Developments in Applied Microbeam Analysis*, EMAS, Rimini, Italy **1993**, 67.
- [45] J. L. Labar, *Microbeam Anal.* **1995**, *4*, 65.
- [46] J. L. Pouchou, *Mikrochim. Acta* **1994**, *114/115*, 33.
- [47] J. I. Goldstein, J. L. Costley, G. W. Lorimer, S. J. B. Reed in *Scanning Electron Microscopy* (Ed.: I. O. Johari), SEM, Chicago, IL **1977**, 315.
- [48] J. P. Goldstein, D. B. Williams, G. Cliff in *Principle of Analytical Electron Microscopy* (Eds.: D. C. Joy, A. D. Romig, J. I. Goldstein), Plenum Press, New York, **1986**, Chap. 5.
- [49] G. Cliff, G. W. Lorimer, *J. Microsc.* **1975**, *103*, 203.
- [50] J. E. Wood, D. B. Williams, J. I. Goldstein, *J. Microsc.* **1984**, *133*, 255.
- [51] T. P. Schreiber, A. M. Wims in *Microbeam Analysis — 1981* (Ed.: R. H. Geiss), San Francisco Press, San Francisco, CA **1981**, 153.
- [52] P. J. Sheridan, *J. Electron. Microprobe Technol.* **1989**, *11*, 41.
- [53] E. Van Cappellen, *Microsc. Microanal. Microstruct.* **1990**, *1*, 1.

## 2.6 Imaging Secondary Ion Mass Spectrometry

### 2.6.1 Introduction

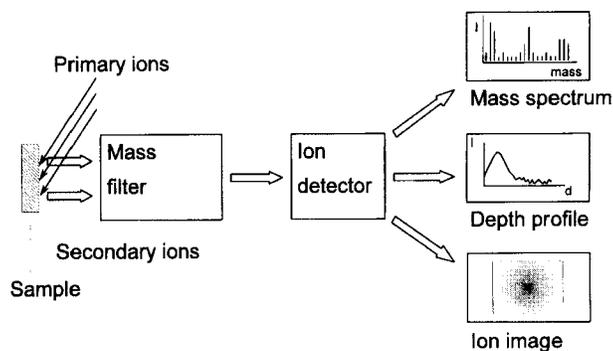
Secondary ion mass spectrometry (SIMS) is a surface and microanalytical technique for the investigation of the chemical composition of solids. The specimen to be analyzed is brought under high vacuum and bombarded with energetic ions (e.g., 10 KeV  $\text{Ar}^+$ ). The sputtering process results in the emission of secondary ions from the sample surface. These secondary ions are separated according to their mass-to-charge ratio by a mass spectrometer, and the mass selected secondary ions are registered by a suitable detector. SIMS is a surface analysis technique because the secondary ions generated by the sputtering process can only escape from the outermost layers ( $\sim 2$  nm) of the sample.

#### 2.6.1.1 Types of Secondary Ion Mass Spectrometry Measurements

SIMS measurements can be done in various ways. The chemical composition of the sample surface is studied by acquiring a mass spectrum, that is, the intensity of the generated secondary ions is registered as a function of their mass. The sputtering process causes an erosion of the surface.

Under prolonged bombardment the surface layer is removed, and deeper layers become exposed to the primary ion beam. By acquiring the intensity of one or more ion species as a function of time and thus as a function of depth, so-called depth profiles are registered. This ability to perform (large-area) surface analysis and depth profiling was originally the main reason for using SIMS as an analytical technique. This functionality proved to be essential in the development of semiconductor materials in the microelectronics industry [1, 2, 3].

Although the ability to map the lateral distribution of secondary ions was already present in one of the earliest commercial available SIMS instruments originally developed by Slodzian [4, 5], the demand to characterize complex structures with micrometer and submicrometer scale dimensions, pushed improvements in instrument design and added imaging capability as a third major functionality to SIMS. By acquiring the intensity of the mass-filtered secondary ions as a function of their location of origin, spatially resolved ion maps can be produced. These ion images reveal the two-dimensional chemical composition of the surface [6]. By combining image acquisition with depth profiling, that is, acquiring images as a function of depth, information on the



**Figure 1.** Schematic diagram of a secondary ion mass spectrometer set-up and the basic types of data acquisition: mass scan, depth profile and ion mapping

three-dimensional chemical composition of the sample can be obtained [7, 8]. The essential components of an SIMS instrument and the basic data acquisition modes are shown schematically in Fig. 1.

### 2.6.1.2 Dynamic and Static Secondary Ion Mass Spectrometry

When dealing with SIMS, one has to distinguish between dynamic and static modes. In a static SIMS experiment the sample receives a total primary ion dose of less than  $10^{15}$  ions  $\text{cm}^{-2}$ . Under these conditions the erosion of the surface is very limited and only the outermost layer of the substrate is sampled. Static SIMS has a very high surface sensitivity: 0.1% of a monolayer can be detected. It is possible to investigate organic layers: due to the low dose, large molecules can be sputtered from the surface with little or no fragmentation [9].

Dynamic SIMS uses a much higher primary beam current density, up to  $0.2 \text{ A cm}^{-2}$ , resulting in an erosion rate of several monolayers per second. The high material consumption results in dynamic

SIMS having a very high 'bulk' sensitivity. Elements present in the sample at very low concentration levels—parts per million and even parts per billion—can be detected. Depth profiles of typically  $1 \mu\text{m}$  into the sample can be acquired in less than 1 h. However, all information on organic constituents is lost due to the nearly complete fragmentation of the organic molecules under the intense ion bombardment.

### 2.6.1.3 Ion Microscope and Ion Microprobe

Ion images can be acquired in two different ways: in the ion microscope mode or in the ion microprobe mode [10]. In an ion microscope the sample is bombarded with a broad primary ion beam. The stigmatic ion optical design of the mass spectrometer preserves the lateral distribution of the secondary ions from the point of origin at the surface to the point of detection by a suitable position-sensitive detector. The detector registers a mass-filtered image of the secondary ions as they leave the sample. This is also called direct-imaging SIMS. The lateral resolution of an ion microscope is of the order of  $1 \mu\text{m}$ ,

**Table 1.** Comparison of SIMS with other imaging microanalytical and surface analytical techniques

Technique	Element range	Detection limits	Information depth	Lateral resolution	Type of information
Secondary ion mass spectrometry ion mapping (SIMS)	≥ H	ppm	0.3–2 nm	~100 nm	Elemental, isotopic, organic
Scanning Auger microscopy (SAM)	≥ Li	0.1%	0.3–2.5 nm	50 nm	Elemental, some chemical states
X-ray mapping in electron probe X-ray microanalysis (EPXMA)	≥ B	0.1%	0.5–2 μm	1 μm	Elemental
X-ray photoelectron microscopy (XPS)	≥ He	1%	1–3 nm	~30 μm	Elemental, chemical state

and is limited by the aberrations of the stigmatic secondary ion optics.

In an ion microprobe the sample is irradiated by a finely focused primary ion beam, and the mass-filtered secondary ions originating from the location of impact are measured. By scanning the beam over the surface and registering the signal as a function of beam position, an ion image is built up in a sequential manner, much in the same way as in a scanning electron microprobe or a scanning Auger microprobe. The lateral resolution is determined by the diameter of the primary beam, and can be as low as 20 nm.

Both dynamic and static SIMS can be done in the ion microscope and ion microprobe modes. Some instruments are dedicated ion microprobes, whereas others can be operated in both the microscope and in microprobe modes.

#### 2.6.1.4 Characteristics of Secondary Ion Mass Spectrometry

Being a mass spectrometric method, SIMS is able to detect all elements (actually all

isotopes) of the periodic table. The sensitivity, that is, the amount of signal generated for a given concentration, varies over orders of magnitude for different elements, and depends strongly on the sample composition and on the experimental conditions. Under favorable conditions detection limits in the part per million and even in the part per billion range are attainable. The high sensitivity of SIMS and the ability to detect all elements, isotopes, and organic molecular fragments are major advances compared to other surface analytical techniques such as Auger spectroscopy and X-ray photoelectron spectroscopy (XPS) [11]. Table 1 compares some characteristics of SIMS with other imaging surface and microanalytical techniques. The ability to map the distribution of chemical species on a microscopic scale with high sensitivity and isotopic discrimination is unique to SIMS [12]. Depth profiling in combination with imaging is a powerful technique to solve problems of multielement characterization of materials at the microscopic level [8], and forms the basis for three-dimensional reconstruction of elemental distributions within sputtered microvolumes of materials [13, 14].

A more detailed treatment of SIMS and its various aspects can be found in monographs [15, 16].

## 2.6.2 Secondary Ion Formation

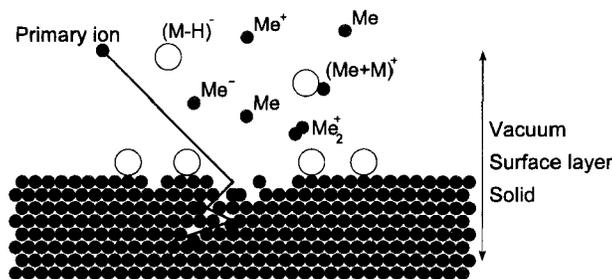
When a primary ion with an energy in the kiloelectronvolt range impacts on the surface of a solid it will collide with the sample atoms and gradually lose its energy until it comes to rest a few tenths of nanometers below the sample surface. The energy released in the collision cascade results in a disruption of the material; atoms are displaced and bonds are broken. Particles at and near the surface can receive sufficient kinetic energy to become ejected. Intact molecules, for example adsorbed on the surface, or fragments of these molecules as well as atoms constituting the sample surface or clusters of these atoms, can be ejected. This process is illustrated in Fig. 2. The majority of particles are released as neutral species, but some carry a positive or negative charge. These secondary ions are measured in SIMS.

The sputtering process itself is quite well understood [17]. For the formation of the secondary ion, that is, the ionization

process, several models have been put forward, but no general applicable concept is available [18–21]. The observed secondary ion current  $I_M$  for an ion M is given by

$$I_M^\pm = S\gamma^\pm C_M\alpha^\pm I_P$$

where  $S$  is the sputter yield,  $\gamma^\pm$  is the ionization probability,  $C_M$  is the concentration of species M in the sample,  $\alpha^\pm$  is the transmission efficiency of the instrument, and  $I_P$  is the primary ion current. The  $\bullet$  sign indicates whether positive or negative secondary ions are considered. The sputter yield  $S$  is defined as the total number of sputtered particles per incident primary ion, and depends on the primary beam parameters (mass, energy, and incident angle) as well as on the sample composition. Its value ranges typically between 1 and 10 sputtered particles per incident ion. The ionization probability is the fraction of particles M sputtered as positively (or negatively) charged ions. Its value can vary from  $\sim 0.1$  to as low as  $10^{-5}$ , and depends on the electronic structure of the species and on the chemical state in which they are before ejection. Electropositive elements (e.g., sodium and the other alkaline elements) have a higher positive ion yield than the noble elements (e.g., gold); the ion yield of silicon is much higher in  $\text{SiO}_2$  than in silicon. This complex dependence of the ionization probability on the



**Figure 2.** The sputtering process, as a result of the collision cascade of the impacting primary ions, causes, neutral, positive and negative particles to be released from the surface.

experimental conditions and on the chemical state of the sample (the so-called matrix effect) makes truly quantitative SIMS analysis extremely difficult [22, 23]. The ion yield of a species and thus the sensitivity can vary over orders of magnitude depending on the matrix.

The ion bombardment induces considerable changes in the near-surface region of the sample. The impinging ion becomes implanted at a certain depth below the surface. While colliding with the sample atoms, they get displaced and driven deeper into the sample (knock on and mixing), causing a degradation of the surface composition. The process is especially important in dynamic SIMS, and is responsible for the loss of depth resolution during depth profiling.

Whereas the minimal depth resolution of an SIMS measurement is determined by the escape depth of the secondary ions, the ultimate lateral resolution limit is determined by the diameter of the interaction area in the material. Monte Carlo calculations of the intensity of sputtered particles as a function of the radial distance from the impact of  $\text{Ga}^+$  ions onto aluminum yield an interaction range of less than 20 nm [24, 25]. A lateral resolution of 20 nm, approaching this theoretical limit, has been obtained with the UC-HRL scanning ion microprobe [26].

### 2.6.3 Instrumentation

SIMS instruments are rather complex machines costing in the order of US\$1 million and having a footprint of some  $5\text{ m}^2$  or more. The basic subdivision is between static and dynamic SIMS

instruments, although they can also be categorized according to the type of mass spectrometer and image formation.

The four main parts of an SIMS instrument are the primary ion guns, the sample chamber, the mass spectrometer, and the secondary ion detector and the image registration system. Most SIMS instruments use PCs or workstation computers to control all or nearly all instrument functions.

#### 2.6.3.1 Primary Ion Sources

The most important parameters of the primary ion gun are the type of ions, the accelerating voltage, the beam current, and the brightness of the source. Electron impact ion sources, able to produce  $\text{Ar}^+$  or  $\text{O}^+$  ions, and duoplasmatron sources, able to produce  $\text{Ar}^+$ ,  $\text{O}_2^+$ , or  $\text{O}^-$  ions, are low-brightness sources. They can produce beam currents of up to a few hundred milliamps, but the beam can only be focused down to a few micrometres.  $\text{Cs}^+$  ions can be produced with a caesium surface ionization source [27]. At ion currents below 100 pA, a submicrometer diameter beam can be obtained. The ultimate beam size and thus lateral resolution in ion microprobes is achieved with  $\text{Ga}^+$  liquid metal ion sources [28]. Compared to the chemically neutral  $\text{Ar}^+$  ions,  $\text{O}_2^+$  primary ions enhance the ion yield of electropositive elements by several orders of magnitude. The implanted oxygen increases the workfunction of the surface, which in turn increases the positive ion formation efficiency. In a similar way, by lowering the workfunction,  $\text{Cs}^+$  primary ions enhance the formation of

negative ions from electronegative elements [29].

### 2.6.3.2 Sample Chamber

To minimize surface contamination during the static SIMS measurements, an ultra-high vacuum ( $<10^{-6}$  Pa) is maintained in the sample chamber. Otherwise, residual gas molecules would deposit on the sample at a faster rate than particles would be removed from the surface by ion bombardment. Due to the continuous erosion of the surface, a dynamic SIMS instrument only requires a moderately high vacuum ( $<10^{-4}$  Pa). Sample holders can accommodate specimens of typically up to 1 cm in diameter, although complete wafers can be fitted in some instruments. The sample can be moved in the  $x$ ,  $y$ , and  $z$  directions, and rotated. Cooling of the sample at near-liquid nitrogen temperatures is often possible.

### 2.6.3.3 Mass Spectrometer

Three types of mass spectrometer are in use in SIMS: quadrupole mass filters, double-focusing magnetic sector instruments, and time-of-flight (TOF) mass analyzers [30].

Quadrupole mass spectrometers have a mass range of up to 1000 amu, but only unit mass resolution and a low transmission efficiency ( $<1\%$ ). For imaging they can only be used in ion microprobes.

For the magnetic sector mass spectrometer the mass range is limited to  $\sim 500$  amu. A mass resolution of up to 10 000 (10% valley) is attainable; however,

the transmission efficiency is inversely proportional to the mass resolution, and ranges for  $\sim 30\%$  at  $M/\Delta M = 800$  to less than 1% at  $M/\Delta M = 10\,000$ . This mass spectrometer is mainly found in dynamic SIMS instruments and can be used for ion microscopes as well as for ion microprobes.

In a TOF mass spectrometer the mass is determined by measuring the time for which the ions fly in a drift tube after being accelerated by a constant potential. To allow this timing, TOF mass spectrometers are used with pulsed ion sources. This mass analyzer combines a very large mass range of up to 10 000 amu, with a very high transmission efficiency ( $\sim 50\%$ ). The mass resolution is determined by the time uncertainty of the arrival of the ions, that is, the ion formation time (= primary ion pulse width) plus the spread in the initial kinetic energy distribution of the secondary ions. Using subnanosecond primary beam pulses and a drift tube with an electrostatic deflector, a mass resolution of 10 000 (full width at half-maximum, FWHM) at mass 30 is possible [31]. Static SIMS instruments are most often equipped with a TOF mass spectrometer because of the high transmission efficiency. They can be used in ion microprobes and ion microscopes. Table 2 summarizes some of the characteristics of the three mass analyzers.

### 2.6.3.4 Ion Detection and Image Registration

In an ion microscope the mass resolved and magnified secondary ion image is detected with a position-sensitive device. The first stage of the detection system

**Table 2.** Characteristics of different mass spectrometers used in SIMS instruments

Mass spectrometer	Mass resolution	Transmission efficiency (%)	Mass range (amu)	Imaging mode
Double focusing and magnetic sector	10 000 <sup>a</sup> 800 <sup>a</sup>	1 30	0–500	Ion microscope and Ion microprobe
Quadrupole	Unit mass	< 1	0–1000	Ion microprobe
TOF	10 000 <sup>b</sup>	~50	0–10 000	Ion microscope and Ion microprobe

<sup>a</sup> 10% valley.

<sup>b</sup> FWHM at  $M = 30$ .

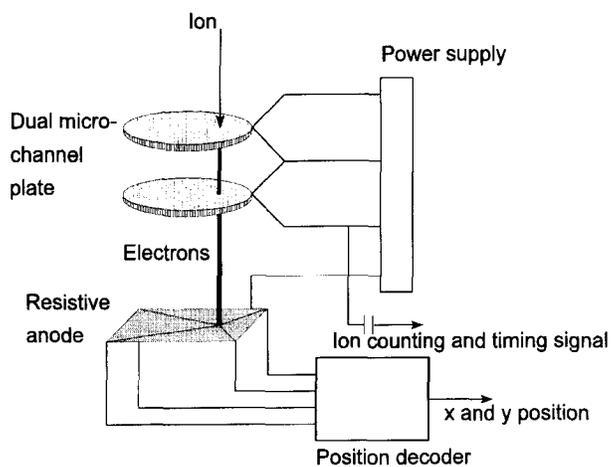
consists of a single or dual microchannel plate (MCP), which acts as an ion-to-electron converter and amplifier. The gain of the MCP depends on the applied voltage, and is limited to about 1000 for a single MCP and can be up to  $10^5$  (one ion producing  $10^5$  electrons) for a dual MCP. The localized electron cloud can be detected with a resistive anode encoder (RAE) or a charge-coupled device (CCD) image sensor or converted to photons using a phosphor screen. In the last case a video camera is used to capture the image and a frame grabber to convert it to a digital image.

The earliest recording of ion images was based on 35 mm film [32]. However, the

analog nature of photographic recording does not provide a direct measure of the ion intensities in these images.

In the case of the resistive anode encoder, the electrons from the MCP hit a resistive film (the resistive anode). The position of impact is calculated from the ratio of the charge collected at the four corners of the anode. The impact itself produces a count pulse corresponding to the detection of an individual secondary ion [33]. A schematic diagram of the RAE is shown in Fig. 3.

Video camera-based systems are frequently used for image registration [34–36] as well as digital cameras such as the charge injection device (CID) camera [13]

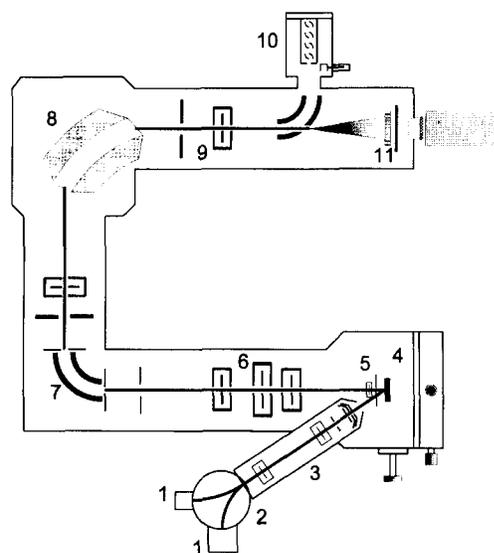

**Figure 3.** Schematic diagram of an RAE detector used as a position-sensitive ion counter in SIMS imaging.

and the CCD camera [37–39]. The lateral resolution of both the RAE and camera-based system is sufficient not to cause any additional degradation of the ion image.

In an ion microprobe the image registration is in principle much simpler. An ion multiplier is used to count the number of ions of a certain mass while the focused primary beam strikes the sample surface for a certain amount of time. By scanning the beam the image is built up sequentially. Count rates of up to  $10^6$  counts per second can be sustained. High secondary ion currents need to be measured with a Faraday cage. The dwell time per pixel can be varied from a few microseconds to several milliseconds, resulting in acquisition times between 1 s and tenths of minutes for one image. This is the case when using a quadrupole or magnetic sector mass analyzer. With a TOF mass spectrometer, all secondary ions generated by the impact of one pulse of primary ions are separated in time according to their mass and can be counted one after the other. In this way the images of all masses are built up simultaneously. However, the number of secondary ions per primary ion pulse is very small, zero to a few hundred at the maximum, so that a large number of pulses need to be fired at each location.

### 2.6.3.5 Typical Configurations

In the following paragraphs we will discuss in some detail two typical SIMS instrument configurations. The Cameca IMS instrument is a dynamic SIMS instrument, originally designed as an ion microscope but also capable of working as an ion microprobe [40, 41]. The Cameca ION-



**Figure 4.** Layout of the Cameca IMS secondary ion mass spectrometer, a dynamic SIMS instrument original designed as an ion microscope but also capable of working as an ion microprobe. 1, ion source; 2, primary beam mass filter; 3, primary column lenses; 4, sample chamber; 5, immersion lens; 6, transfer optics; 7, electrostatic analyzer; 8, magnetic sector; 9, projection lens; 10, electron multiplier and Faraday cage; 11, MCP and fluorescent screen.

TOF SIMS is an example of a static ion microprobe [9, 42, 43].

Figure 4 shows a schematic drawing of the Cameca IMS SIMS instrument. This instrument is a dynamic SIMS instrument for near-surface analysis, depth profiling, and ion imaging. It uses a magnetic sector mass spectrometer of a rather sophisticated ion–optical design that allows the instrument to operate in both the ion microscope and microprobe modes.

In the configuration shown in Fig. 4 the instrument is equipped with two primary ion sources. The duoplasmatron source can be used to produce  $O_2^+$ ,  $O^-$ , or  $Ar^+$  ions. This is a low-brightness source that, for ion imaging, is mainly used in the ion microscope mode. The second source is a

cesium surface ionization source, producing a  $\text{Cs}^+$  ion beam with sufficient brightness to allow ion microprobe operation with a lateral resolution of a few hundreds of nanometers. The primary ions are accelerated to an energy that is variable from 5 to 17.5 keV, and pass a magnetic filter which, depending on the magnetic field strength and direction, directs the appropriate ions into the primary column. Three electrostatic lenses focus the primary beam onto the sample. Beam-defining apertures and stigmators in the column determine the shape and current of the beam. At the end of the column, four deflector plates allow scanning of the beam over a sample area of up to  $500 \times 500 \mu\text{m}^2$ .

The sample is mounted vertically on the sample stage, which can be moved in the  $x$  and  $y$  directions with the aid of stepping motors. The vacuum in the sample chamber is  $\sim 10^{-4}$ – $10^{-6}$  Pa. For the measurement of positive secondary ions the sample is kept at a potential of +4.5 kV, and for negative ions at –4.5 kV.

The secondary ions produced at the sample surface are extracted by an immersion lens. The immersion lens front plate is held at ground potential. The transfer optics consist of three lenses, of which only one is energized. This lens focuses the secondary ion beam onto the entrance slit of the mass spectrometer. The mass spectrometer consists of a  $90^\circ$  electrostatic sector and a  $90^\circ$  magnetic sector in a modified Nier–Johnson geometry. By varying the width and position of the energy slit, situated behind the electrostatic sector, secondary ions within a certain narrow energy band ( $<127$  eV) are selected. The spectrometer lens transfers these energy-filtered secondary ions to

the magnetic sector. The magnetic field separates the energy-filtered ions into various spatially resolved beams according to the mass-to-charge ratio of the ions. For a given magnetic field strength only secondary ions with a certain mass will pass the exit slit. The width of this slit determines the mass resolution. By varying the magnetic field, ions with different mass can be selected. The mass range is 1–280 amu, or optionally to 500 amu. The mass resolution varies from 200 to a maximum of 10 000 (10% valley).

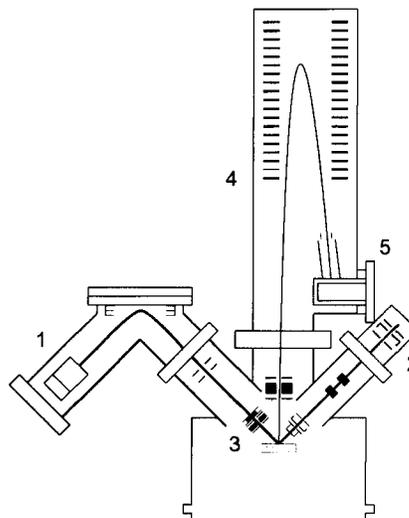
The projection lenses, situated behind the exit slit of the mass spectrometer, are used to postaccelerate the energy- and mass-filtered secondary ions and to direct them into the detection system. The beam can be focused and deflected into an electron multiplier or a Faraday cage to measure the ion current of the secondary ions. Alternatively, the beam can be projected onto an MCP–fluorescence screen assembly.

In the ion microscope mode, a primary ion beam with a diameter of up to  $200 \mu\text{m}$  and a current density of up to  $50 \text{ mA cm}^{-2}$  is rastered over the sample surface. The primary reason for scanning the beam is to get an even ‘illumination’ (and erosion) of the sample surface. The immersion lens forms a global (nonmass filtered) image of part of the bombarded surface area. The size of the imaged field (the area from which ions can enter the mass spectrometer) is determined by the transfer lens and the field aperture. Three different transfer lenses and their associated apertures allow image fields with diameters of 25– $400 \mu\text{m}$ , corresponding to a magnification between  $2000\times$  and  $50\times$ , respectively. The cross-over of the global image, as projected by the transfer optics, is situated

at the entrance slit of the mass spectrometer. Due to lens aberrations, not all secondary ions are focused at the cross-over at one point but rather within a disc of finite diameter. The contrast aperture, placed in front of the cross-over point, is used to reduce the aberration. A smaller aperture results in a higher lateral resolution at the expense of intensity. The global image itself is formed at a fixed distance beyond the slit. The electrostatic sector forms a second virtual global image. The spectrometer lens transfers this to a third virtual image inside the magnetic sector. The energy- and mass-filtered image inside the magnetic sector is transferred by the projection lenses onto the MCP and fluorescent screen.

The process of image formation in the ion microprobe mode is considerably simpler. In this mode the cesium primary ion source is often used to produce a primary ion beam with a diameter of a few hundred nanometers at a current of up to 100 pA. This focused primary ion beam is scanned over the sample area, and the secondary ions released at the point of impact of the primary beam are mass analyzed by the mass spectrometer and measured (counted) by the electron multiplier. For optimal performance in the ion microprobe mode the mass spectrometer must accept a very large fraction of the secondary ions. Since the system must only measure the arrival of the secondary ions and not their position, all apertures and slits are opened.

The Cameca TOF-SIMS instrument is typical of a static SIMS instrument equipped with a TOF mass spectrometer. The schematic diagram of this instrument is given in Fig. 5. One distinguishes again two ion sources, the sample chamber and



**Figure 5.** Typical layout of the Cameca TOF-SIMS, a static ion microprobe instrument. 1, electron impact source; 2, gallium liquid metal ion source; 3, sample chamber; 4, flight tube with reflectron; 5, ion detector.

the TOF mass spectrometer with its read-out electronics.

Pulsed ion sources, delivering very short bunches of ions at a high repetition rate, are used. Each bunch of primary ions creates a few secondary ions which are extracted into the TOF mass spectrometer. Their arrival time at the detector is a measure of their mass.

The electron impact ion source can deliver  $\text{Ar}^+$  and  $\text{O}^+$  ions with an energy of 11 keV. The  $90^\circ$  deflector is used as an electrodynamic mass filter and beam chopper. It also provides axial compression of the ion bunches. An Einzel lens is used for beam focusing. Beam alignment and scanning is done with the X and Y deflectors. The objective lens transfers the beam to the sample surface. Space charge effects impose limits to the beam current (the number of primary ions in each pulse), the pulse length, and the beam diameter.

If, for a given number of ions, the diameter is decreased, the length of the pulse increases. For a beam diameter of 25  $\mu\text{m}$  the minimum pulse length is 0.8 ns, corresponding to  $\sim 1800$  ions per pulse or a current of 3 pA. At this pulse length the TOF mass spectrometer has a mass resolution of  $M/M\Delta$  of about 7000 at mass 29. If the beam diameter is decreased to 3  $\mu\text{m}$ , the pulse length increases to 50 ns, resulting in a mass resolution of  $\sim 300$ , while each bunch contains about 300 ions. The maximum repetition frequency of this source is 10 kHz.

The gallium liquid metal ion source is used in combination with a two-lens column containing a stigmator and a raster unit. The accelerating voltage is 25 kV and the pulse frequency 40 MHz. This ion gun uses an axial electrostatic buncher to produce very short pulses. When the lenses are operated at intermediate cross-over, high lateral resolution can be achieved at high mass resolution. With a beam diameter of  $\sim 150$  nm and a pulse length of 1.5 ns (four ions or 500 pA) the mass resolution is  $\sim 5000$ . A higher beam current can be achieved with the same spot size by increasing the pulse length (100 ns = 266 ions) but at the expense of mass resolution. The highest lateral resolution is achieved with the beam collimated down to a diameter of 80 nm. With a pulse width of 100 ns (= 30 ions), unit mass resolution is achieved of up to 250 amu.

These pulsed ion sources produce a current density too low to erode the sample (static SIMS conditions). For depth profiling a third wide-beam ion source must be installed. Depth profiling is done in the dual-beam mode, alternating between analyzing the surface with the pulsed source and sputtering the sample

with, for example, a 50  $\mu\text{m}$  diameter  $\text{Ar}^+$  ion gun.

The ultra-high vacuum chamber has a base pressure of  $\sim 6 \times 10^{-8}$  Pa maintained by a turbomolecular pump to ensure an oil-free vacuum. A specimen manipulator can translate the sample in the  $x$ ,  $y$ , and  $z$  directions. The sample is at ground potential.

An Einzel lens is used to extract the secondary ions in combination with a quadrupole for secondary beam alignment. The extraction potential is  $\pm 2$ –3 kV.

The TOF mass spectrometer is of the 'reflectron' type with first-order energy focusing. The secondary ions enter the electrostatic reflector, where they are first decelerated and then accelerated again, making a nearly  $90^\circ$  turn. The energy of the secondary ions is the sum of their initial kinetic energy and the acceleration they receive when they were formed at or near the surface. Ions with an excess initial kinetic energy will penetrate deeper into the reflector and therefore travel in total a longer distance in the flight tube and arrive at nearly the same time at the detector as ions with the same mass but lower initial kinetic energy. A mass resolution in excess of 10 000 ( $M/\Delta M$ , FWHM) can be achieved in this way. The mass range extends up to 10 000 amu. Depending on the initial kinetic energy distribution of the second ions the transmission varies between 20 and 80%.

The detection system comprises a single MCP followed by a scintillator and a photomultiplier. The output of the photomultiplier goes to a multistop time-to-digital converter with a time resolution that is variable from 50 to 800 ps, which can detect up to 512 events (= secondary ions) per primary ion pulse. Up to 4000

elemental images can be recorded in parallel.

Two other important configurations are the University of Chicago scanning ion microprobe and the TOF SIMS instrument developed by Charles Evans and Associates.

The University of Chicago UC-HRL scanning ion microprobe is a dynamic SIMS instrument currently having the highest lateral resolution. It uses a Ga<sup>+</sup> liquid metal ion source mounted perpendicularly above the sample surface. Using an acceleration voltage of 40–60 kV, beam diameters of 20–85 nm carrying a current of 1.6–60 pA are obtained. Originally the secondary ions were analyzed with a quadrupole mass spectrometer [44]. Because of the low transmission of the quadrupole mass spectrometer (<0.1%), the instrument has been fitted with a magnetic sector mass spectrometer giving a transmission of ~20% at a mass resolution  $M/\Delta M = 1000$  [45].

The TOF secondary ion mass spectrometer developed by Charles Evans and Associates is a static SIMS instrument using a TRIFT mass spectrometer of unique design that allows both ion microscopy and ion microprobe operation in combination with a TOF analyzer [46]. It can be equipped with a large-diameter Cs<sup>+</sup> ion source for ion microscope analysis and

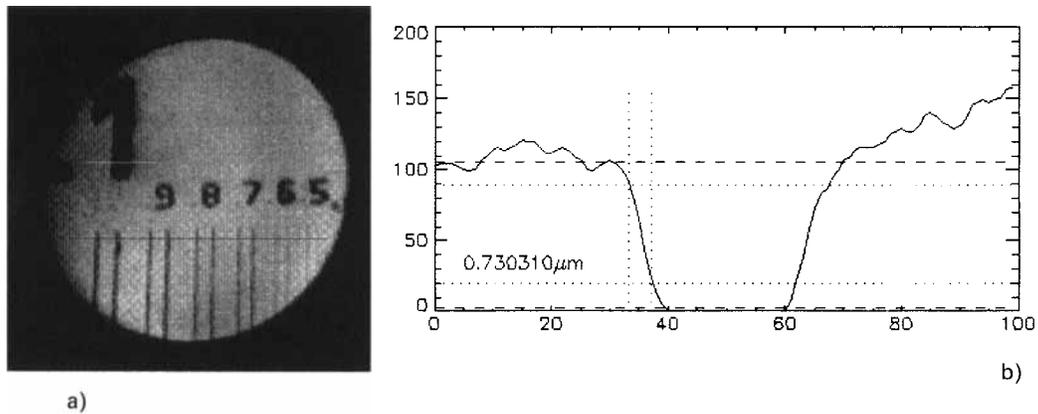
with a Ga<sup>+</sup> liquid metal ion source for imaging in the microprobe mode. The TRIFT mass spectrometer has stigmatic optics consisting of three electrostatic energy analyzers [47]. In the ion microscope mode, a mass resolution >5000 can be obtained, while the lateral resolution is limited to ~3 μm. In the ion microprobe mode a submicrometer lateral resolution is attainable at the cost of mass resolution. A time-to-digital converter and a resistive anode encoder are used as an ion counting/position-sensitive detector for the image registration. The instrument operates over a mass range from 1 to 10 000 amu.

#### 2.6.4 Comparison of Ion Microprobe and Ion Microscope Mode

Table 3 compares some of the features of ion microprobes and ion microscopes. As already stated, the lateral resolution of the ion microscope is limited to about 1 μm by the aberrations in the stigmatic ion optics. In the ion microprobe the lateral resolution is determined by the diameter of the finely focused primary ion beam that scans over the sample surface. With

**Table 3.** Comparison of the main features of ion microprobes and ion microscopes

Feature	Ion microprobe	Ion microscope
Primary ion current	0.1–100 pA	1 nA–1 μA
Beam diameter	<1 μm	100 μm
Analyzed area	10 μm	150 μm
Lateral resolution	~100 nm	~1 μm
Detection limits	~0.1%	~1 ppm

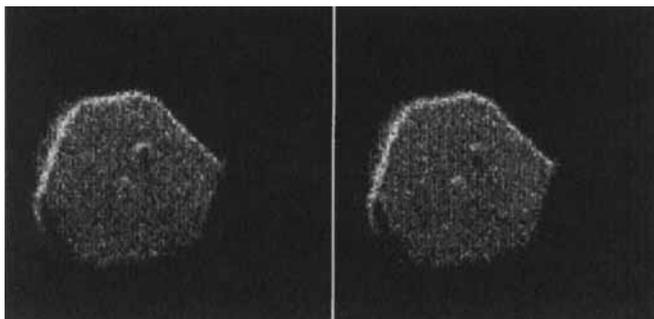


**Figure 6.** (a)  $\text{Si}^+$  ion image obtained from a silicon substrate with a gold test pattern (Cameca 4f in ion microscopy mode) and (b) line scan over the digit '1'. The image has a lateral resolution of  $0.7\ \mu\text{m}$  close to the maximum attainable resolution in the ion microscope mode.

high-brightness  $\text{Ga}^+$  liquid metal ion sources, a diameter and thus a lateral resolution of  $\sim 100\ \text{nm}$  is readily obtainable. Figure 6a shows the  $\text{Si}^+$  image obtained from a silicon substrate covered with a gold test pattern using a Cameca 4f instrument. The bars have a width from  $1\ \mu\text{m}$  down to  $0.5\ \mu\text{m}$ . To measure the lateral resolution [48], a line scan over the digit '1' was made (Fig. 6b). The intensity drops from 84 to 16% of the maximum over a distance of  $0.7\ \mu\text{m}$ , which is close to the maximum attainable resolution. Figure 7 shows the ion images of  $\text{Br}^-$

from a silver halide microcrystal acquired with a TOF ion microprobe instrument using a  $\text{Ga}^+$  liquid metal ion source with a pulse width of 100 ns at a mass resolution of 300. The image field is  $6 \times 6\ \mu\text{m}^2$ , and the lateral resolution is about 75 nm, clearly demonstrating the superior lateral resolution of the ion microprobe.

Other factors to be considered when comparing the ion microscope and the ion microprobe are mass resolution, analyzed area (or volume), and image acquisition time.



**Figure 7.**  $\text{Br}^-$  secondary ion images from a silver halide microcrystal obtained with a TOF SIMS ion microprobe using a  $\text{Ga}^+$  liquid metal ion source. The imaged field is  $6 \times 6\ \mu\text{m}$  and the lateral resolution is  $\sim 75\ \text{nm}$ . Because of the high lateral resolution of the ion microprobe, structures at the surface of this micrometer-sized object can be observed.

In the ion microscope the mass resolution is in principle independent of the lateral resolution, although high mass resolution is only obtained at the expense of instrument transmission and thus sensitivity. In ion microprobes equipped with a TOF mass spectrometer, the highest lateral resolution is only obtained at low mass resolution.

The main difference between the ion microprobe and the ion microscope modes of imaging is in the applied primary ion current. The total number of atoms consumed in an SIMS experiment in a time  $t$  is given by  $N_c \sim A\Phi_p St$ , where  $A$  is the analyzed area,  $\Phi_p$  is the primary current density, and  $S$  is the sputter rate [49]. The current densities in the ion microprobe and in the ion microscope are of the same order of magnitude, but in the microscope we can use a much broader beam and thus a much higher current. We can analyze an area of  $100 \times 100 \mu\text{m}^2$  in the microscope mode using a beam that covers the entire area, and obtain secondary ion images quasi-instantaneously. To analyze the same area in the ion microprobe mode we use a beam of  $1 \mu\text{m}^2$  and raster this beam over the surface. Since the current density is about the same in both experiments, we would need to measure 10 000 times longer to sputter the same amount of material (i.e., to obtain the same detection limit). Or, when using the same acquisition time, the ion microscope is orders of magnitude more sensitive than the ion microprobe. The ion microscope is therefore most suitable to analyze large areas with high sensitivity, but if submicrometer lateral resolution is required, the microprobe is the preferred instrument.

## 2.6.5 Ion Image Acquisition and Processing

### 2.6.5.1 Dynamic Range of Secondary Ion Mass Spectrometry Ion Images

Image acquisition is complicated by the dynamic range of SIMS, especially in ion microscopy. Secondary ion intensities can easily vary over 3–6 orders of magnitude. To observe high concentrations next to very low ones, an equally large dynamic range is required within one image. Due to both the large variation in concentration and sensitivity for different species, one also has to accommodate the large signal intensity variation from one image (or mass) to another.

The RAE has a dynamic range of about 50 000, which is sufficient for most practical applications. Moreover, this device operates in the ion-counting mode, so that the read-out is a direct measure of the true ion intensity [50].

The situation is quite different when using video camera-based systems. The intrinsic dynamic range is typically of the order of 200–1000 ( $\sim 8$  to  $\sim 10$  bits). Images of largely different intensity can be easily acquired by varying the gain of the MCP and/or the internal amplification stages of the camera. This can be done under computer control to obtain images of optimum brightness and contrast for each ion species [51]. However, in this way the direct relation between the ion intensity and the grey level values in the image is lost unless a sophisticated calibration scheme is set up. To overcome this problem, one can measure the total ion

current of the image with the electron multiplier or Faraday detector just after the acquisition of the image.

Within one image the dynamic range can be extended by real-time integrating the image and storing the 8-bit subimages which are not yet in overflow. This integration is possible because of the very low background of the camera ( $0.004 \text{ counts s}^{-1} \text{ pixel}^{-1}$ ). In this way, a dynamic range of about 4000 is possible [51].

### 2.6.5.2 Influence of Mass Resolution

Ion microscope and ion microprobe images are often acquired at low mass resolution ( $M/\Delta M \sim 300$ ). This resolution is sufficient to separate ions of different nominal mass; however, interference from molecular ions or organic fragments are very likely at this resolution. Figure 8a shows the secondary ion image at mass 194 acquired at low resolution with a Cameca 4f instrument taken from a Teflon-coated Pt–Ir wire with a diameter of  $30 \mu\text{m}$ . The wire was cut at an angle so that the Pt–Ir core is visible. In Fig. 8b and c the same area is imaged at high mass resolution. Figure 8b corresponds to

$^{194}\text{Pt}^+$  ions, and reveals the core of the wire, whereas Fig. 8c shows the distribution of an organic fragment due to the Teflon.

### 2.6.5.3 Image Sequences

Images of different ions are acquired sequentially with a magnetic sector instrument. The mass spectrometer is tuned to a certain mass, and the image is acquired (averaged or integrated). The mass spectrometer is then adjusted for the next pre-defined mass, and again an image is taken. The entire sequence takes up to several minutes, depending on the number of masses (typical 10–20) and the integration time per image. If image depth profiles are acquired, this cycle is repeated several times. An image depth profile can result in a very large amount of data; for example, if 50 cycles of 10 images each are taken, in total 500 images (of  $256 \times 256$  pixels) need to be stored. Such an image depth profile is acquired in about 1 h of instrument time and, depending on the applied primary beam current and imaged field, corresponds to a sampled volume with an area of  $100 \times 100 \mu\text{m}^2$  and a depth of  $1 \mu\text{m}$ .



**Figure 8.** (a) Image at mass 194 of a cross-section of a Teflon-coated Pt–Ir wire acquired with the Cameca 4f ion microscope at low mass resolution. Images of the same area at high mass resolution showing the contribution from (b) the

### 2.6.5.4 Interpretation and Processing of Ion Images

Ideally the contrast observed in an ion image should be only due to local concentration variations of the measured ion. However, the relation between elemental concentration at the surface and the observed grey level in the acquired image is often disturbed by spurious contrast mechanisms. The main sources of spurious contrast are: matrix contrast, orientation or crystallographic contrast, topographic contrast, and chromatic or energy contrast [11]. The last two are less important in ion microprobe images. Finally, the uneven response of the position-sensitive ion detector used in the ion microscope (e.g., gain variations across the MCP) may cause unwanted intensity variations in the image. To interpret ion images correctly, the concentration contrast must be separated from the other contrast sources [6].

Crystallographic contrast occurs in polycrystalline materials because the sputter rate and thus the secondary ion intensity depend on the orientation of the crystallographic planes with respect to the incident primary ion beam. An example of this contrast is given in Fig. 9.

If the surface exhibits a certain roughness, topographic contrast can be observed. Due to the oblique incidence of the primary ion beam, some parts of the surface may not receive any primary ions. Also, the fact that the sputter yield depends on the incident angle contributes to this type of contrast.

Chromatic contrast is caused by the local variation in energy distribution of the secondary ions and the limited energy



**Figure 9.**  $\text{Cu}^+$  secondary ion image of a pure copper sample, showing crystallographic contrast.

bandpass of the mass spectrometer. This contrast is most pronounced in the ion microscope, especially when charge build-up occurs on nonconducting samples.

Matrix contrast causes the secondary ion yield of the element of interest to be different in different locations of the sample even if the concentration is the same, due to local variations in the composition of the matrix [52–54]. This false contrast is the most difficult to recognize, and may seriously affect the interpretation of the ion images.

Depending on the type of analysis, various image-processing steps might be required. Normalizing the image of a trace impurity to the image of a matrix element can be done to eliminate contrast due to uneven illumination or camera response and effects of crystal orientation. The normalized images then better reflect the contrast due to the concentration variation [55]. Truly quantitative images are more difficult to obtain because of the high variation in sensitivity and the influence

of the matrix on the sensitivity. The sensitivity factor and the so-called matrix/ion species ratio (MISR) have been applied successfully on a pixel basis to obtain images where each grey value corresponds to the local concentration of the element [51].

Classical image-processing procedures such as contrast enhancement or grey level equalizing and the use of pseudo-color look-up tables are often employed to aid the visual interpretation of the ion images [56]. Image restoration techniques using Fourier filtering have been used to enhance the lateral resolution of ion microprobe images [57, 58]. However, the high noise level in the ion microprobe images (pixels on average may contain only a few counts) limits the applicability of this technique.

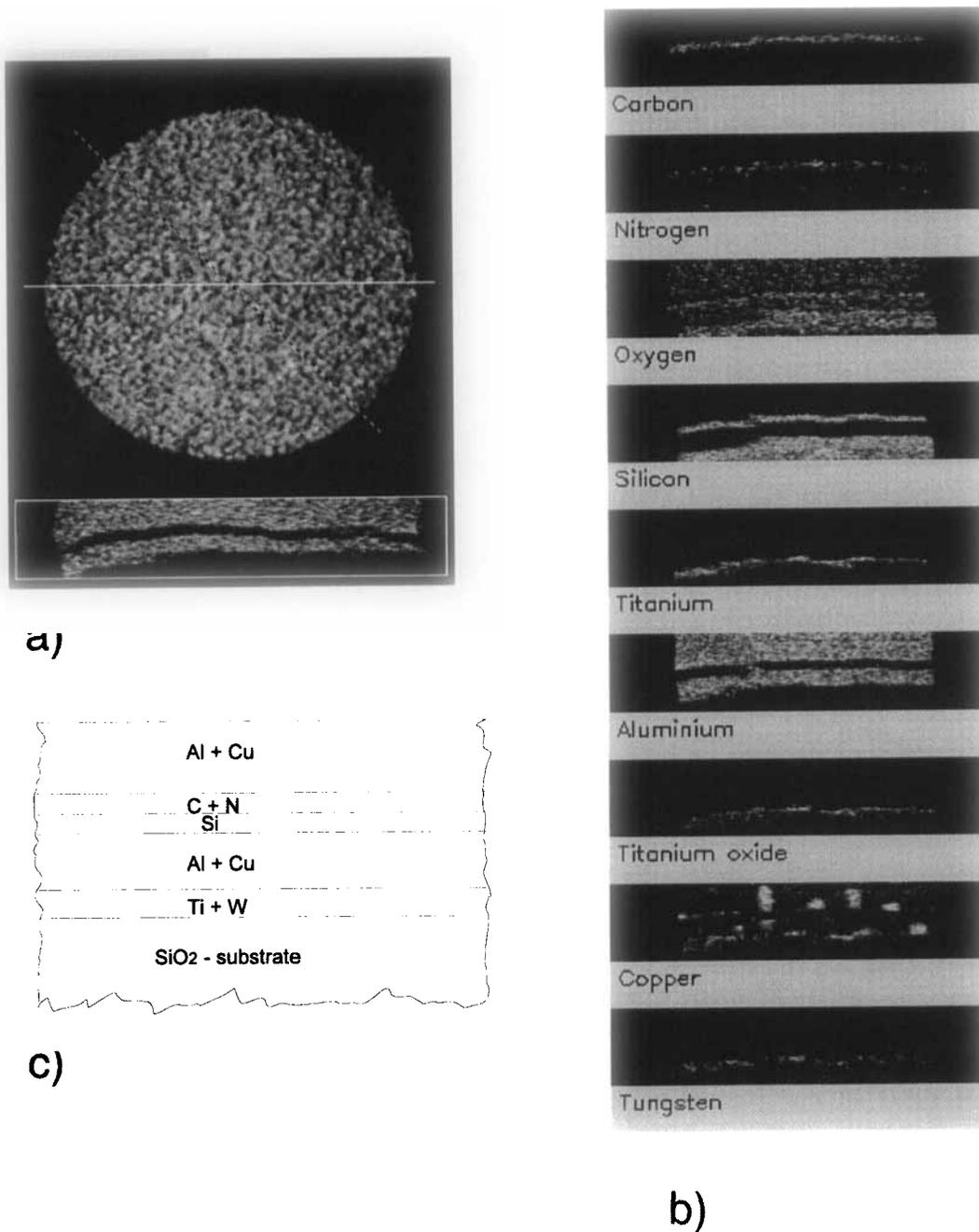
### 2.6.5.5 Analysis of Image Depth Sequences

The acquisition of sequences of images as functions of depth allows a number of interesting data-processing options [59–66]. The images for one specific ion can be arranged in a stack, where the  $x$  and  $y$  directions correspond to the lateral distribution, and the  $z$  direction to depth. This stack can be probed in different ways. In local (or retro) depth profiling, an area is selected on top of the stack, and the intensity of the pixels in this area is calculated for the different images in the stack. The result is a depth profile (intensity versus depth). The interesting aspect is that this process can be repeated on different regions of the sample area without reanalyzing the sample. This is especially useful

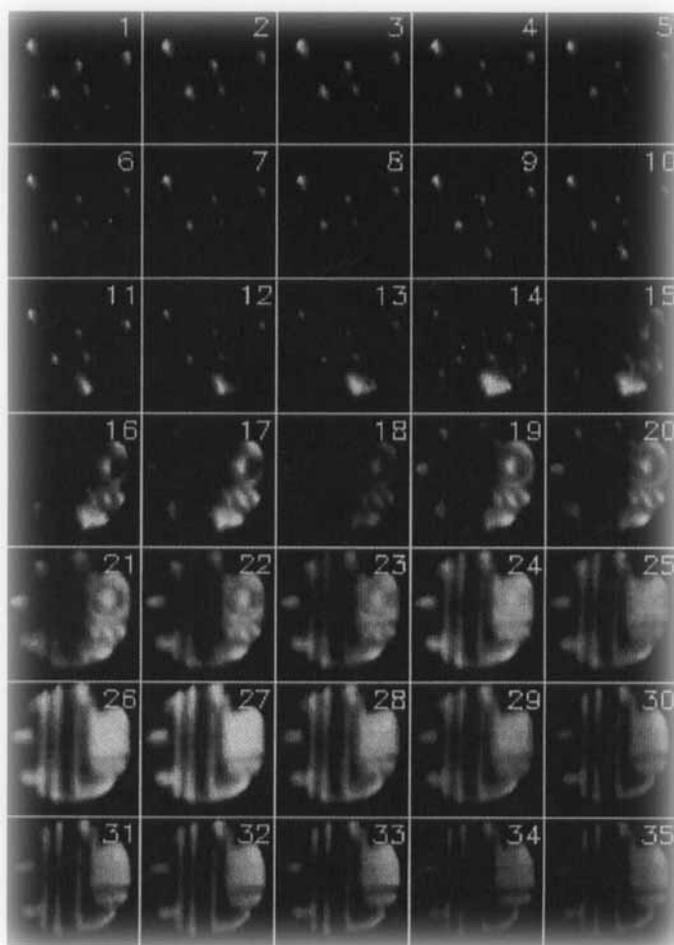
for the investigation of complex samples containing certain features buried below the surface. Cross-sections of the image stack allow the depth distribution of the various species to be viewed under different angles. Finally, a pseudo three-dimensional reconstruction of the interior structure can be made. Examples of these procedures are given below.

In the first example, the sample studied consisted of a  $1.5\ \mu\text{m}$  thick multilayer containing aluminum, copper, titanium, tungsten, and some other impurities on an  $\text{SiO}_2$  substrate. An area of  $150\ \mu\text{m}$  in diameter was analyzed in the ion microscope mode with the Cameca 4f SIMS instrument. An image depth profile of 9 masses ( $^{12}\text{C}^+$ ,  $^{14}\text{N}^+$ ,  $^{16}\text{C}^+$ ,  $^{30}\text{Si}^+$ ,  $^{49}\text{Ti}^+$ ,  $^{27}\text{Al}_2^+$ ,  $^{46}\text{Ti}^{16}\text{O}^+$ ,  $^{63}\text{Cu}^+$ , and  $^{184}\text{W}^+$ ) and 44 cycles was acquired in about 90 min, resulting in 396 images of  $256 \times 256$  pixels. Figure 10a shows the uniform distribution of aluminum at the surface. The full line indicates the location where a cross-section was made perpendicular to the surface. Figure 10b shows cross-sectional images through each of the nine image stacks. From these images one can clearly deduce the structure of the multilayer, which is schematically represented in Fig. 10c.

In the second example we consider the pseudo three-dimensional reconstruction of a sample consisting of aluminum structures (lanes and planes) deposited on an  $\text{SiO}_2$  substrate and covered with an  $\text{SiO}_2$  passivation layer. The passivation layer also contains aluminum inclusions. Thirty-five cycles of six masses ( $^{11}\text{B}^+$ ,  $^{23}\text{Na}^+$ ,  $^{24}\text{Mg}^+$ ,  $^{27}\text{Al}^+$ ,  $^{31}\text{P}^+$ , and  $^{39}\text{K}^+$ ) were acquired under conditions similar to the first example. Figure 11 shows the lateral distribution of aluminum as a



**Figure 10.** (a)  $\text{Al}^+$  image at the surface of a multilayer structure. The full line indicates the position where a vertical cross-section through the image stack is made. (b) Cross-sectional images obtained from the image stacks of nine different ions. (c) Schematic view of the reconstructed multilayer.

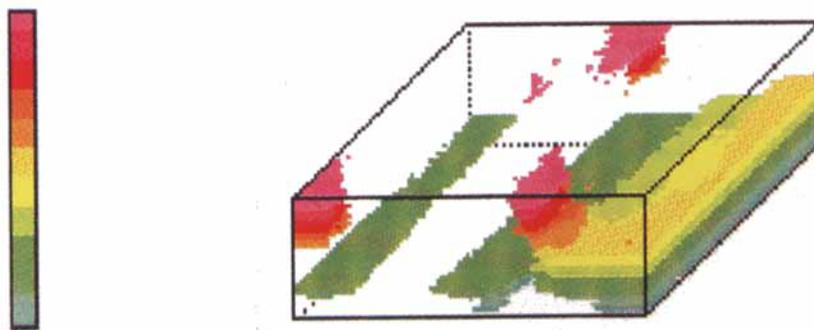


**Figure 11.** Lateral distribution of aluminum as a function of depth in a sample containing aluminum structures buried under an  $\text{SiO}_2$  passivation layer. Cycle 1 corresponds to the surface, and cycle 35 to  $\sim 1 \mu\text{m}$  below the surface.

function of depth. Cycles 1 (at the surface) to 11 show the aluminum inclusions in the passivation layer. From cycle 16 onwards, the buried aluminum structure becomes visible. From the stack of aluminum images, a pseudo three-dimensional reconstruction was made, as shown in Fig. 12. In this transparent view the grey scale corresponds to depth in the sample. Some of the aluminum inclusions in the passivation layer are in contact with the buried aluminum structure, causing device failure.

### 2.6.5.6 Analysis of Multivariate Ion Images

Images of different masses taken at the surface or at a certain depth can be arranged as a multivariate image stack. In this multivariate image stack the  $x$  and  $y$  directions correspond to the lateral distribution, while the  $z$  direction is now a spectroscopic dimension, that is, each image holds the same lateral (spatial) distribution of a different ion species.



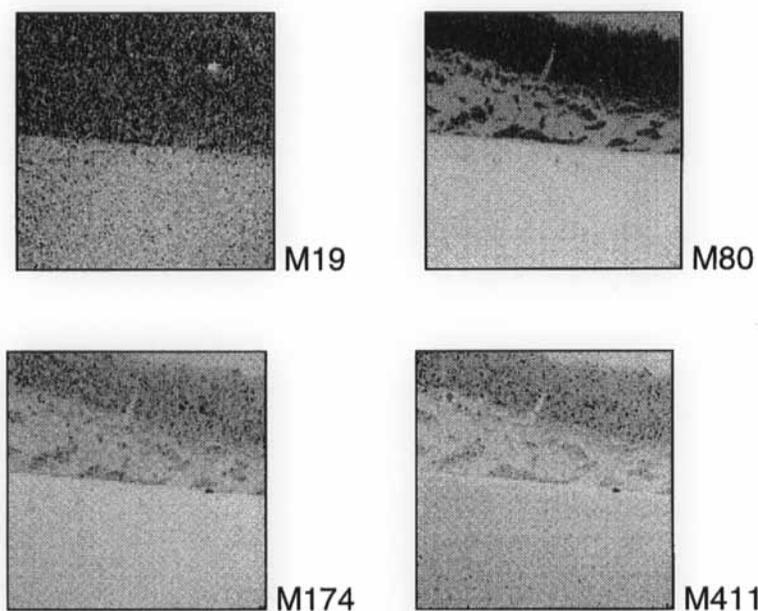
**Figure 12.** Reconstructed pseudo three-dimensional view of the aluminum distribution in the interior of a sample. In the transparent view the grey scale corresponds to depth in the sample.

Very often these images are correlated because species are present at the same location on the sample or are chemically related. Multivariate data analysis methods such as principal component analysis can be applied to this image stack [67–69]. The basic idea is to reduce the large number of correlated images to a smaller number of images (two or three) in such a way that the major portion of the information, or variation, present in all the images is condensed in these (principal component) images. This data reduction allows the user to observe the essential information. By making two-dimensional scatter plots of these first few principal component images, regions in the sample with similar chemical properties can be identified [70, 71]. This method is especially interesting for TOF-SIMS. The nature of the acquisition means that often a large number of images (20–100) are acquired simultaneously. Multivariate image analysis aids considerably in interpreting the large amount of data gathered.

As an example of the application of multivariate image analysis, we consider the study of a copper surface treated with butylbenzotriamide (BBT) [9]. BBT is a

corrosion inhibitor forming a self-assembling monolayer on metals. The formation of the passivation layer is affected by the oxidation state of the metal surface and by the presence of impurities. Figure 13 shows four negative secondary ion images out of a series of 21 images in total. The images were acquired with a TOF SIMS instrument using a  $1\ \mu\text{m}$  diameter pulsed  $\text{Ga}^+$  primary ion beam with a dose of  $2.3 \times 10^{12}$  ions  $\text{cm}^{-2}$ . In Fig. 13, mass 19 corresponds to  $\text{F}^-$ , mass 80 to  $\text{SO}_3^-$ , mass 174 to deprotonated BBT ( $\text{BBT-H}^-$ ), and mass 411 to a Cu–BBT adduct ( $\text{Cu}(\text{BBT-H})_2^-$ ). The other images at masses 12, 25, 32, 35, 38, 41, 42, 50, 81, 96, 97, 131, 143, 254, 255, 263, and 382 more or less resemble one of the four images shown.

Figure 14 shows the first three principal component images calculated from the 21 original images. The first two show a clear structure in the image, whereas the third (and following) principal component images contain mainly noise. When the two-dimensional histogram of the first two principal component images is made, three different regions, corresponding to pixels that have similar grey values, are observed (Fig. 15). Next, a mask is



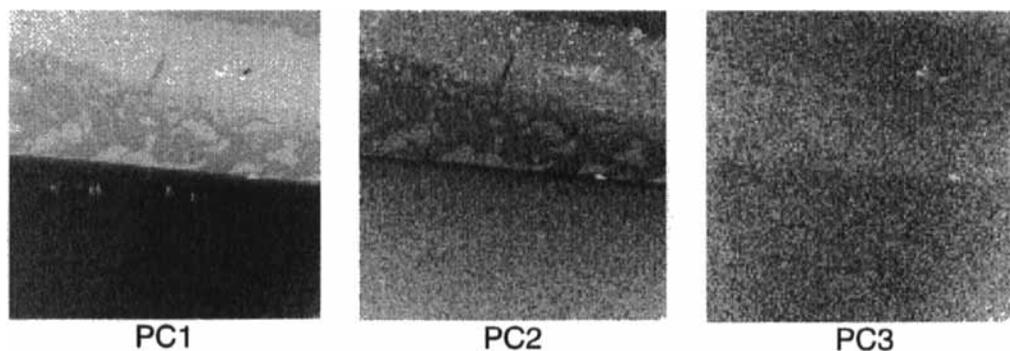
**Figure 13.** TOF-SIMS negative ion images from a copper surface treated with BBT; mass 19,  $F^-$ ; mass 80,  $SO_3^{2-}$ ; mass 174,  $(BBT-H)^-$ ; mass 411,  $(Cu(BBT-H)_2)^-$ .

constructed with the pixels that contributed to each of these three regions. In Fig. 16 these three masks are shown. Mask A corresponds to class A in Fig. 15, and shows the area on the sample where the BBT molecule is associated with the copper surface. Mask B (class B in Fig. 15) corresponds to unreacted BBT, and mask C (class C in Fig. 15) shows the area where only inorganic acids and other contaminants are present at the surface.

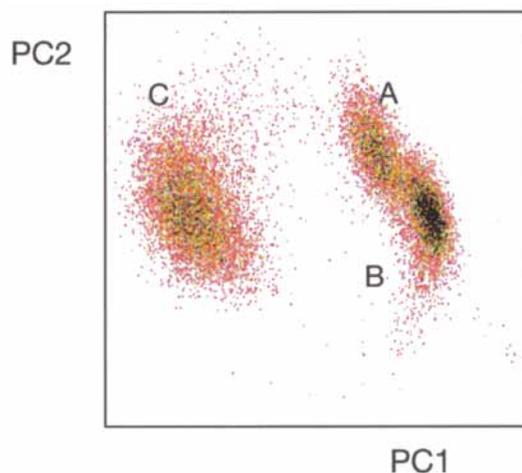
In this way it is possible to identify locations on the surface where specific types of chemical reactions have occurred.

## 2.6.6 Sample Requirements

Because one bombards the surface with charged particles and also extracts charged



**Figure 14.** First three principal component images calculated from the 21 ion images from a BBT-treated copper surface.



**Figure 15.** Two-dimensional histogram (scatter plot) of the first two principal component images shown in Fig. 14.

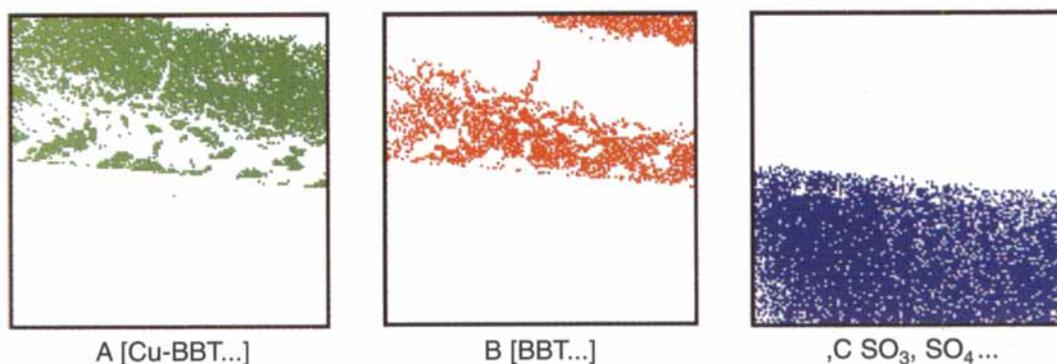
particles the samples should be conducting. Charge build-up will strongly affect the secondary ion extraction and cause considerable image degradation. Coating the sample with a conducting layer (gold), as is done in scanning electron microscopy (SEM), is less advantageous for SIMS. One has to sputter through the layer before reaching the sample surface. For true surface analysis this method is ruled out as the coating will destroy the surface

composition. However, insulating samples can be analyzed when a positive primary ion beam is used and the charge build-up is compensated by flooding the surface at the same time with low energetic electrons.

Surface roughness is an important factor in SIMS and SIMS imaging. The sputter yield depends on the incident angle of the primary ions. Moreover, a shadowing effect occurs because the primary beam hits the sample typically at an angle of  $40^\circ$ , and some parts of the rough surface will not receive any primary ions. This causes a topographic induced contrast in the image. The roughness of the surface also degrades the depth resolution in depth profiling because it is assumed that all ions escape from the same depth, which is evidently not the case for a rough surface. The sample surface does not become smoother under continued bombardment; in fact, the roughness increases.

### 2.6.7 Application Domain

SIMS imaging is applied in many diverse fields of science and technology where the



**Figure 16.** Masks obtained from the pixels that make up the three classes in Fig. 15. Masks A, B, and C correspond to classes A, B, and C, respectively, in Fig. 15.

more traditional SIMS methods are also commonly used: microelectronics, materials science [72, 73], geochemistry [74], biological and biomedical sciences [75, 76], and environmental studies. The imaging aspect of SIMS is of advantage whenever the spatial distribution of minor and trace elements needs to be investigated. In this respect, ion imaging complements SEM X-ray mapping with the added advantage of a much higher detection limit and the ability to perform three-dimensional analysis. Odom [49] provides a detailed overview of the use of imaging SIMS, with special emphasis on biological applications. The proceedings of the biannual SIMS conferences provide a good overview of the developments and applications of this imaging technique. It is interesting to note that at the 1982 SIMS conference about 8% of the contributions dealt with imaging aspects of SIMS whereas in 1991 this number had increased to more than 22%, clearly showing the growing importance of imaging.

Applications in microelectronics dealing with various aspects of manufacturing processes are reported in the literature: localization and identification of contaminants [77], study of device failure, three-dimensional characterization of implants, and the study of multilayer structures [78]. The ability of static SIMS to view the distribution of organic species at the surface is used to study contamination from cleaning solvents and other process chemicals employed in photolithography.

A considerable number of applications deal with the study of biological systems. Although this is probably the most difficult subject, the high sensitivity of imaging SIMS is of great value here, because bio-

chemical processes are often controlled by a low concentration of chemical species localized in subcellular structures. For the analysis of soft tissues adequate sample preparation is necessary to maintain the cell structure and the chemical composition after removing or immobilizing the cell fluid. One possibility is to employ cryomicrotomy after shock freezing of the sample. If the SIMS instrument is equipped with a cryotransfer system and a cryo-stage, the section can be analyzed directly [79]. Alternatively, the section can be freeze dried under carefully controlled conditions. Other special sample preparation techniques have been developed. The work of Lodding [80, 81] gives a detailed account of the analysis of hard tissues such as teeth and bone.

SIMS imaging is also used to study particles and other microscopic objects. Atmospheric particulate material of a few micrometers in diameter can be studied on an individual particle basis to investigate the chemical composition and to disclose surface enrichments [82, 83]. Silver halide microcrystals used in photographic material are typically 1  $\mu\text{m}$  in diameter and have a complex core-shell structure with different chemical compositions of the type  $\text{AgX}_x\text{Y}_y$ , where X and Y are  $\text{Cl}^-$ ,  $\text{Br}^-$ , and/or  $\text{I}^-$ . SIMS imaging using the ion microprobe mode is used to reveal the internal structure of these microscopic crystals and to give feedback on the production process and on the photographic properties of the material [84].

A fast growing area of SIMS imaging is the mapping of organic coatings and layers using static TOF SIMS. The ability to localize organic species allows the verification of the uniformity of organic coatings on a wide variety of materials and

the study of the chemistry involved in adhesion [9].

## 2.6.8 References

- [1] C. W. Magee, *Ultramicroscopy* **1984**, *14*, 55.
- [2] W. Katz, G. A. Smith, *Scanning Electron Micros.* **1984**, 1557.
- [3] K. T. F. Janssen, P. R. Boudewijn in *Analysis of Microelectronic Materials and Devices* (Eds.: M. Grasserbauer, H. W. Werner), Wiley, Chichester **1991**, p. 407.
- [4] A. Balise, G. Slodzian, *Rev. Phys. Appl.* **1973**, *8*, 105.
- [5] G. H. Morrison, G. Slodzian, *Anal. Chem.* **1975**, *47*, 933A.
- [6] W. Steiger, F. G. Rüdener, *Anal. Chem.* **1979**, *51*, 2107.
- [7] A. J. Patkin, B. K. Furman, G. H. Morrison in *Microbeam Analysis* (Ed.: D. B. Wittry), San Francisco Press, San Francisco **1980**, p. 181.
- [8] F. G. Rüdener, *Surf. Interface Anal.* **1984**, *6*, 132.
- [9] A. Benninghoven, *Angew. Chem., Int. Ed. Engl.* **1994**, *33*, 1023.
- [10] G. Slodzian, *Scanning Microsc.* **1987**, 1.
- [11] H. W. Werner in *Proc. Conf. on Electron and Ion Spectroscopy of Solids* (Eds.: L. Fiermans, J. Vennik, W. Dekeyser), Plenum Press, New York **1978**, p. 324.
- [12] D. M. Drummer, G. H. Morrison, *Anal. Chem.* **1980**, *42*, 2147.
- [13] S. R. Bryan, W. S. Woodward, R. W. Linton, D. P. Griffis, *J. Vac. Sci. Technol.* **1985**, *A3*, 2102.
- [14] H. Hutter, M. Grasserbauer, *Mikrochim. Acta* **1992**, *107*, 137.
- [15] A. Benninghoven, F. G. Rüdener, H. W. Werner (Eds.), *Secondary Ion Mass Spectrometry: Basic Concepts, Instrumental Aspects, Applications and Trends*, Wiley, New York **1987**.
- [16] R. G. Wilson, F. A. Stevie, C. W. Magee, *Secondary Ion Mass Spectrometry*, Wiley, New York **1989**.
- [17] P. Sigmund, *Phys. Rev.* **1969**, *184*, 383.
- [18] V. R. Deline, C. A. Evans, *Appl. Phys. Lett.* **1978**, *33*, 578.
- [19] C. A. Anderson, J. R. Hinthorn, *Anal. Chem.* **1973**, *45*, 1421.
- [20] P. Williams, *Surf. Sci.* **1979**, *90*, 588.
- [21] P. Williams, *Appl. Surf. Sci.* **1982**, *13*, 241.
- [22] K. Wittmack, *Appl. Surf. Sci.* **1981**, *9*, 315.
- [23] A. E. Morgan, H. A. M. de Grefte, H. J. Tolle, *J. Vac. Sci. Technol.* **1981**, *18*, 164.
- [24] M. T. Bernius, Y.-C. Ling, G. H. Morrison, *Anal. Chem.* **1986**, *58*, 94.
- [25] J. M. Chabala, R. Levi-Setti, Y. L. Wang, *Appl. Surface Sci.* **1988**, *32*, 10.
- [26] R. Levi-Setti, G. Crow, Y. L. Wang in *Secondary Ion Mass Spectrometry V* (Eds.: A. Benninghoven, R. J. Colton, D. S. Simons, H. W. Werner), Springer-Verlag, Berlin **1986**, p. 132.
- [27] H. A. Storms, K. F. Brown, J. D. Stein, *Anal. Chem.* **1977**, *49*, 2023.
- [28] J. M. Chabala, R. Levi-Setti, Y. L. Wang, *J. Vac. Sci. Technol.* **1988**, *B6*, 910.
- [29] M. Schuhmacher, H. N. Migeon, B. Rasser in *Secondary Ion Mass Spectrometry VIII* (Eds.: A. Benninghoven, K. T. F. Janssen, J. Tümpner, H. W. Werner), Wiley, Chichester **1992**, p. 49.
- [30] J. C. Vickerman, A. Brown, N. M. Reed (Eds.), *Secondary Ion Mass Spectrometry: Principles and Applications*, Clarendon Press, Oxford **1989**.
- [31] J. Schweiters, H. G. Cramer, U. Jürgens, E. Niehuis, T. Heller in *Secondary Ion Mass Spectrometry VIII* (Eds.: A. Benninghoven, K. T. F. Janssen, J. Tümpner, H. W. Werner), Wiley, Chichester **1992**, p. 497.
- [32] J. D. Fasset, J. R. Roth, G. H. Morrison, *Anal. Chem.* **1977**, *49*, 2322.
- [33] R. W. Odom, B. K. Furman, C. A. Evans, Jr., C. E. Bryson, W. A. Petersen, M. A. Kelly, D. H. Wayne, *Anal. Chem.* **1983**, *55*, 574.
- [34] B. K. Furman, G. H. Morrison, *Anal. Chem.* **1980**, *52*, 2305.
- [35] D. Newbury, D. Bright in *Microbeam Analysis* (Ed.: D. Newbury), San Francisco Press, San Francisco **1988**, p. 105.
- [36] F. P. Michiels, W. K. Vanhoolst, P. E. Van Espen, F. C. Adams in *Microbeam Analysis* (Ed.: P. E. Russell), San Francisco Press, San Francisco **1989**, p. 594.
- [37] J. D. Brown, *Scanning Microsc.* **1988**, *2*, 653.
- [38] J. L. Hunter, S. F. Corcoran, R. W. Linton, D. P. Griffis in *Microbeam Analysis* (Ed.: P. E. Russell), San Francisco Press, San Francisco **1989**, p. 597.
- [39] D. S. Mantus, G. H. Morrison, *Anal. Chem.* **1990**, *62*, 1148.
- [40] H. N. Migeon, C. Le Pipec, J. J. Le Goux in *Secondary Ion Mass Spectrometry V* (Eds.: A. Benninghoven, R. J. Colton, D. S. Simons, H. W. Werner), Springer-Verlag, Berlin **1986**, p. 155.
- [41] H. N. Migeon, B. Rasser, M. Schuhmacher, J. J. Le Goux in *Secondary Ion Mass Spectrometry VIII* (Eds.: A. Benninghoven, K. T. F. Janssen, J. Tümpner, H. W. Werner), Wiley, Chichester **1992**, p. 195.

- [42] E. Niehuis in *Secondary Ion Mass Spectrometry VIII* (Eds.: A. Benninghoven, K. T. F. Janssen, J. Tümpner, H. W. Werner), Wiley, Chichester **1992**, p. 269.
- [43] J. Schwieters, H. G. Cramer, T. Heller, U. Jürgens, E. Niehuis, J. Zehnpfenning, A. Benninghoven, *J. Vac. Sci. Technol.* **1991**, *A9*, 2864.
- [44] R. Levi-Setti, G. Crow, Y. L. Wang, *Scanning Electron Microsc.* **1985**, 535.
- [45] R. Levi-Setti, J. M. Chabala, K. Gavrilov, J. Li, K. K. Soni, R. Mogelevsky in *Secondary Ion Mass Spectrometry IX* (Eds.: A. Benninghoven, Y. Nikei, R. Shimuzi, H. W. Werner), Wiley, Chichester **1994**, p. 233.
- [46] B. Schueler, P. Sander, D. A. Reed, *Vacuum* **1990**, *41*, 1661.
- [47] B. Schueler in *Secondary Ion Mass Spectrometry VII* (Eds.: A. Benninghoven, C. A. Evans, Jr., K. D. McKeegan, H. A. Storms, H. W. Werner), Wiley, New York **1990**, p. 311.
- [48] M. Schumacher, H. N. Migeon, B. Rasser in *Secondary Ion Mass Spectrometry VII* (Eds.: A. Benninghoven, C. A. Evans, Jr., K. D. McKeegan, H. A. Storms, H. W. Werner), Wiley, New York **1990**, p. 939.
- [49] R. W. Odom in *Microscopic and Spectroscopic Imaging of the Chemical State* (Ed.: M. D. Morris), Marcel Dekker, New York **1993**, p. 345.
- [50] R. W. Odom, D. H. Wayne, C. A. Evans in *Secondary Ion Mass Spectrometry IV* (Eds.: A. Benninghoven, J. Okano, R. Shimizu, H. W. Werner), Springer-Verlag, Berlin **1984**, p. 186.
- [51] F. Michiels, W. Vanhoolst, P. Van Espen, F. Adams, *J. Am. Soc. Mass Spectrom.* **1990**, *1*, 37.
- [52] J. T. Brenna, G. H. Morrison, *Anal. Chem.* **1986**, *58*, 1675.
- [53] F. Michiels, F. Adams, *J. Anal. Atom. Spectrom.* **1987**, *2*, 773.
- [54] R. Canteri, L. Moro, M. Anderle in *Secondary Ion Mass Spectrometry VIII* (Eds.: A. Benninghoven, K. T. F. Janssen, J. Tümpner, H. W. Werner), Wiley, Chichester **1992**, p. 135.
- [55] A. N. Thorne, A. Dubus, F. Degréve, *Scanning Electron Microsc.* **1986**, 1255.
- [56] D. E. Newbury, D. S. Bright in *Secondary Ion Mass Spectrometry VII* (Eds.: A. Benninghoven, C. A. Evans, Jr., K. D. McKeegan, H. A. Storms, H. W. Werner), Wiley, New York **1990**, 929.
- [57] B. G. M. Vandeginste, B. C. Kowalski, *Anal. Chem.* **1983**, *66*, 557.
- [58] W. Hao, P. Van Espen, *Vacuum* **1994**, *45*, 447.
- [59] F. Michiels, W. Vanhoolst, W. Jacob, P. Van Espen, F. Adams, *Micron Microsc. Acta* **1987**, *18*, 249.
- [60] D. R. Kingham, A. R. Bayly, D. J. Fathers, P. Vohralic, J. M. Walls, A. R. Waugh, *Scanning Microsc.* **1987**, 463.
- [61] J. J. Lee, K. H. Gray, W. J. Lin, J. L. Hunter, R. W. Linton in *Secondary Ion Mass Spectrometry VIII* (Eds.: A. Benninghoven, K. T. F. Janssen, J. Tümpner, H. W. Werner), Wiley, Chichester **1992**, p. 505.
- [62] N. S. McIntyre, K. F. Taylor, G. R. Mount, C. G. Weisener in *Secondary Ion Mass Spectrometry VIII* (Eds.: A. Benninghoven, K. T. F. Janssen, J. Tümpner, H. W. Werner), Wiley, Chichester **1992**, p. 513.
- [63] A. J. Patkin, G. H. Morrison, *Anal. Chem.* **1982**, *54*, 2.
- [64] W. Steiger, F. G. Rüdener, G. Ernst, *Anal. Chem.* **1986**, *58*, 2037.
- [65] G. Gillen, R. L. Myklebust in *Secondary Ion Mass Spectrometry VIII* (Eds.: A. Benninghoven, K. T. F. Janssen, J. Tümpner, H. W. Werner), Wiley, Chichester **1992**, 509.
- [66] W. K. Vanhoolst, P. J. Van Espen, *Mikrochim. Acta* **1991**, 415.
- [67] K. Esbensen, P. Geladi, *Chemometrics Intelligent Lab. Systems* **1989**, *7*, 67.
- [68] P. Geladi, S. Wold, K. Esbensen, *Anal. Chim. Acta* **1986**, *191*, 473.
- [69] P. Van Espen, G. Janssens, W. Vanhoolst, P. Geladi, *Analysis* **1992**, *20*, 81.
- [70] N. Bonnet, *Mikrochim. Acta* **1995**, *120*, 195.
- [71] N. Bonnet, M. Herbin, P. Vautrot, *Ultramicroscopy* **1995**, *60*, 349.
- [72] D. B. Williams, R. Levi-Setti, J. M. Chabala, D. E. Newbury, *J. Microsc.* **1987**, *148*, 241.
- [73] F. Adams, L. Butaye, G. Janssens, P. Van Espen, *Anal. Sci.* **1991**, *7*, 383.
- [74] J. F. Lovering, *NBS Special Publication 427*, US National Bureau of Standards, Gaithersburg, MD **1975**.
- [75] W. A. Ausserer, Y.-C. Ling, S. Chandra, G. H. Morrison, *Anal. Chem.* **1989**, *61*, 2690.
- [76] C. M. John, J. A. Chakel, R. W. Odom in *Secondary Ion Mass Spectrometry VIII* (Eds.: A. Benninghoven, K. T. F. Janssen, J. Tümpner, H. W. Werner), Wiley, Chichester **1992**, p. 657.
- [77] B. W. Schueler, R. W. Odom, J. A. Chakel in *Secondary Ion Mass Spectrometry VIII* (Eds.: A. Benninghoven, K. T. F. Janssen, J. Tümpner, H. W. Werner), Wiley, Chichester **1992**, p. 281.
- [78] P. K. Chu, W. C. Harris, G. H. Morrison, *Anal. Chem.* **1982**, *54*, 2208.
- [79] M. T. Bernius, S. Chandra, G. H. Morrison, *Rev. Sci. Instrum.* **1985**, *56*, 1347.
- [80] A. Lodding, H. Odelius, *Mikrochim. Acta* **1983**, *10*, 21.
- [81] A. Lodding, *Scanning Electron Microsc.* **1983**, 1229.

- [82] G. E. Cabaniss, R. W. Linton, *Environ. Sci. Technol.* **1984**, *18*, 319.
- [83] T. R. Keyser, D. F. S. Natusch, C. A. Evans, Jr., R. W. Linton, *Environ. Sci. Technol.* **1978**, *12*, 768.
- [84] I. Geuens, R. Gijbels, W. Jacob in *Secondary Ion Mass Spectrometry VIII* (Eds.: A. Benninghoven, K. T. F. Janssen, J. Tümpner, H. W. Werner), Wiley, Chichester **1992**, p. 479.

## 2.7 Scanning Electron Microscopy with Polarization Analysis (SEMPA)

### 2.7.1 Introduction

Scanning electron microscopy with polarization analysis (SEMPA) [1] is a technique that provides high resolution images of magnetic microstructure by measuring the spin polarization of low energy secondary electrons generated in a scanning electron microscope [2–5]. This is possible because the emitted secondary electrons retain the spin polarization present in the material; SEMPA therefore produces a direct image of the direction and magnitude of the magnetization in the region probed by the incident SEM electron beam. SEMPA determines all three components of the spin polarization, and hence of the magnetization. SEMPA records the magnetic and topographic images simultaneously, but independently. Polarization is normalized to the number of emitted electrons, that is, to the intensity, or the quantity measured in a secondary electron SEM topographic image. Thus, SEMPA measurements are intrinsically independent of topography. This feature allows the investigation of the correlation between magnetic and topographic structures. SEMPA can characterize ferromagnetic materials with a

sensitivity down to a fraction of an atomic layer and a lateral resolution of 20 nm. The surface sensitivity of SEMPA is particularly advantageous for studies of thin film and surface magnetism [6, 7] but puts requirements on the cleanliness of specimen surfaces. SEMPA also has other advantages common to scanning electron microscopes, such as long working distance, large depth of field, and large range of magnifications. The zoom capability is especially useful for imaging the magnetization distribution in ferromagnets where length scales vary over several orders of magnitude from relatively large ( $>10\ \mu\text{m}$ ) magnetic structures such as ferromagnet domains, to intermediate size (200 nm) structures found in Bloch, Néel, asymmetric Bloch or cross-tie domain walls, to the finest structures ( $<50\ \text{nm}$ ) found in magnetic singularities such as Bloch lines, Néel caps and magnetic swirls.

To put SEMPA in perspective, it is useful to compare it to other methods of imaging magnetic microstructure, some of which are discussed at length in other chapters of this handbook. Different magnetic imaging methods are distinguished by the quantity measured to obtain magnetic contrast, the resolution, the ease of

**Table 1.** Comparison of several magnetic imaging techniques

Method	Quantity measured	Resolution [nm]	Information depth [nm]	Reference
SEMPA	magnetization	20	1	[2-7]
magneto-optic Kerr	magnetization	500	10	[16]
MCD	magnetization	300	2	[17]
SPLEEM	magnetization	20	1	[18]
Bitter	field	1000	1000	[8]
MFM	field	10	1000	[15]
TEM Lorentz	field	2	average	[9, 10]
STEM differential phase contrast	field	2	through	[14]
Electron holography	field	2	sample	[12, 13]

interpretation of the measurement, the requirements on sample thickness and surface preparation, the cost, and so on. A summary of this information for the various imaging techniques mentioned below is displayed in Table 1. The values given for the resolution of each technique are estimates of the current state of the art; they should only be taken as a rough guide.

Most methods used for the observation of magnetic microstructure rely on the magnetic fields in and around a ferromagnet to produce magnetic contrast. For example, the oldest method for imaging magnetic microstructure is the Bitter method [8] where fine magnetic particles in solution are placed on the surface of a ferromagnet. The particles agglomerate in the fringe fields at domain walls thereby delineating the magnetic domains; the particles may be observed in an optical microscope or even an SEM. In Lorentz microscopy, the magnetic contrast is derived from the deflection of a focused electron probe as it traverses a ferromagnetic sample [9, 10]. In the transmission electron microscope (TEM), Lorentz microscopy can achieve a high lateral resolution of order 10 nm, but the

measurement represents an average over specimen thickness. Only thin samples (<300 nm) are suitable for high spatial resolution studies. Unfortunately, such thin samples may not have a magnetization distribution characteristic of the bulk. Lorentz microscopy in the reflection mode in an SEM has also been demonstrated [11]. It has the advantage that the near surface of bulk specimens can be examined, but the lateral resolution is seldom much better than 1  $\mu\text{m}$ . Electron holography [12, 13] is an electron interferometric method for obtaining absolute values of the magnetic flux in and around thin ferromagnetic samples. It is a high resolution (2 nm) method with contrast derived from the measurement of electron phase shifts that occur in electromagnetic fields. Differential phase contrast microscopy also measures electron phase shifts to give magnetic contrast at high lateral resolution (2 nm) in the scanning TEM [14]. Magnetic force microscopy (MFM) is an imaging technique suitable for thick (bulk) specimens. It achieves contrast through the magnetostatic interaction between a ferromagnetic tip and the fringe fields of the ferromagnet. MFM can be used to locate domain walls with a spatial resolution of

about 10 nm, but it is difficult to extract quantitative information from MFM images [15].

Direct methods for measuring micro-magnetic structure rely on contrast mechanisms which reveal the magnetization rather than the magnetic induction. The magneto-optic Kerr effect [16] uses the rotation of the plane of polarization of light upon reflection to map surface magnetization distributions. As an optical method, its spatial resolution is diffraction limited to optical wavelengths, but it has the advantage that a magnetic field can be applied and varied during measurement. Magnetic circular dichroism (MCD) images domains using a photoelectron microscope [17]. Photoelectron images are recorded for circularly polarized incident X-rays. Since the photoelectron yield is proportional to the spin-dependent density of states at the surface and the helicity of the X-rays which selectively excite atomic core levels, images of domain structure can be obtained with elemental specificity. Although the information depth within the magnetic material is about 2 nm, secondaries from a 10 nm carbon overcoat have been found to reflect the underlying magnetic structure [17]. Spin-polarized low energy electron microscopy (SPLEEM) is a very new high resolution (20 nm) method for resolving surface magnetic microstructure [18] which relies on the spin-dependence of the (quasi)elastic scattering cross section for polarized electrons from ferromagnets. A spin polarized electron source is required to modulate the spin of the incident beam. Magneto-optic Kerr, MCD, and SPLEEM are like SEMPA in that they measure quantities directly proportional to the sample magnetization.

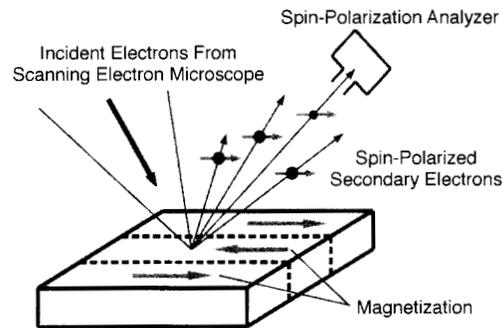
## 2.7.2 Principle of SEMPA

Scanning electron microscopy with polarization analysis (SEMPA), first demonstrated in the mid-1980s [19–23], is a micromagnetic imaging technique that derives magnetic contrast from the spin polarization of secondary electrons extracted from a ferromagnetic surface. The secondary electron magnetic moments are parallel, and consequently their spins antiparallel, to the magnetization vector at their point of origin in the sample [24].

The SEMPA method is schematically depicted in Fig. 1. As the electron beam is scanned across the sample, the secondary electrons are collected and their polarization analyzed. An electron spin analyzer measures each component of the polarization vector,  $\mathbf{P}$ , separately. For example the  $x$  component of polarization is given by

$$P_x = (N_{\uparrow} - N_{\downarrow}) / (N_{\uparrow} + N_{\downarrow}) \quad (1)$$

where  $N_{\uparrow}$  ( $N_{\downarrow}$ ) are the number of electrons detected with spins parallel (antiparallel)



**Figure 1.** The principle of SEMPA. A scanned beam of electrons incident on the surface of a ferromagnet creates spin-polarized secondary electrons which are subsequently spin-analyzed to yield a high resolution magnetization image.

to the  $+x$  direction. The degree of electron spin-polarization varies in the range  $-1 \leq P \leq 1$ . Note that  $P$  is normalized to the total number of electrons emitted,  $(N_{\uparrow} + N_{\downarrow})$ . The polarization and intensity are measured simultaneously, but independently. Thus, the magnetic and topographic images are determined separately.

Spin polarized secondary electrons emitted from a ferromagnet reflect the spin part of the magnetization,

$$M = -\mu_B(n_{\uparrow} - n_{\downarrow}) \quad (2)$$

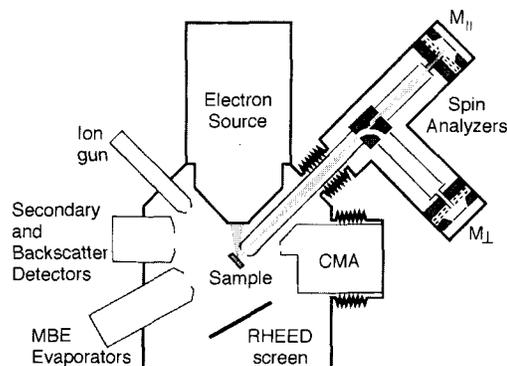
Here  $n_{\uparrow}$  ( $n_{\downarrow}$ ) are the number of spins per unit volume parallel (antiparallel) to a particular orientation, and  $\mu_B$  is the Bohr magneton. To the extent that the secondary electron cascade represents a uniform excitation of the valence electrons, the expected secondary electron polarization can be estimated as  $P = n_B/n_v$ , where  $n_B$  is the number of Bohr magnetons per atom and  $n_v$  the number of valence electrons per atom. In this way, one estimates a polarization of 0.28, 0.19, and 0.05 for Fe, Co, and Ni, respectively. These agree reasonably well with measurements of secondary electrons in the 10 to 20 eV range [25, 26]. At lower energies, spin dependent scattering [27] actually increases the polarization, improving the contrast of SEMPA measurements. SEMPA provides vector magnetization maps for conducting or semiconducting bulk specimens, and thin films and monolayer films where specimen charging is not a problem. In some cases charging can be avoided and the magnetization enhanced by evaporating a 1 nm thick film of Fe on the specimen; this has allowed the imaging of insulating Fe garnets [28]. The magnetization along the

measurement direction is proportional, but oppositely directed, to the electron polarization along that direction. In practice, the constant of proportionality is not precisely known; the detailed scattering dynamics for the production of polarized secondary electrons is dependent upon the surface band structure, which varies from material to material. Measurements using SEMPA reveal the spatial dependence of the relative value of the surface magnetization distribution rather than the absolute size of the surface moments.

Important features of SEMPA include its high spatial resolution and its surface sensitivity. The spatial resolution of SEMPA is primarily determined by the incident beam diameter focused on the sample surface. Even though the profile of the energy deposited in bulk samples expands greatly in the bulk due to multiple scattering [11], the escape depth of polarized secondary electrons is on the order of nanometers. The distance over which the secondary electron spin polarization is exponentially reduced by  $1/e$  is about 0.5 nm for a transition metal like Cr [29], and about 1.5 nm for a noble metal like Ag [30]. Although SEMPA measures only the near-surface micromagnetic structure, the underlying and bulk magnetic structure can be determined by solution of the micromagnetic equations using the surface magnetization measurements as boundary conditions [31]. The surface sensitivity of SEMPA is advantageous for studies of surface and thin film magnetism but can be a limitation in the sense that sample surfaces must be clean. Thick oxides or hydrocarbon layers will diminish the polarization and hence the magnetic contrast.

## 2.7.3 Instrumentation

The electron probe forming column, transport optics, and spin-polarization detectors comprise the essential electron optical components of the SEMPA system. A schematic of a SEMPA instrument is shown in Fig. 2. Because of the surface sensitivity of SEMPA, the specimen should be cleaned and maintained in ultrahigh vacuum ( $P < 5 \times 10^{-8}$  Pa). Conventional surface science preparation and analysis tools including an ion-beam sputtering gun, an electron beam evaporator, an Auger electron spectrometer and a reflection high energy electron diffraction screen greatly facilitate the preparation and characterization of the sample surface. The SEMPA system may be equipped with a single spin detector [4–7], or multiple spin detectors [2, 3] as shown in Fig. 2. Two detectors are used for the acquisition of all three orthogonal components of the vector polarization (magnetization) signal.



**Figure 2.** Schematic of a SEMPA apparatus. The electron source, polarization detectors, cylindrical mirror analyzer (CMA), and reflection high energy electron diffraction (RHEED) screen are shown in their actual relative positions; the rest of the instruments are not. The CMA and polarization analyzers are retractable.

The SEM probe forming column, the transport optics, the polarization analyzers, the electronics and signal processing will be considered in turn. A general description will be given of the generic components of SEMPA; we will use our apparatus as an example for detailed analysis.

### 2.7.3.1 Scanning Electron Microscopy Probe Forming Column

An SEM beam of 10 keV is a reasonable compromise among the constraints of secondary electron yield, spatial resolution, and beam stability in the secondary electron extraction field. The beam energy must be high enough to reduce the deleterious effects of electron lens aberrations, yet low enough to sustain reasonable secondary electron yields from the sample. Submicrometer beam diameters can be obtained for electron energies above 5 keV while the secondary electron yield, for example from Al, falls from 0.40 at an incident beam energy of 5 keV to 0.05 at 50 keV [11]. The incident electron beam must also be energetic enough such that the extraction optics which transport the polarized secondary electrons to the spin detector do not severely aberrate the focused spot on the sample. Extraction optics typically have fields on the order of 100 V/mm to achieve adequate collection efficiency. A 10 keV beam suffers minimal distortion in such an extraction field.

Two essential components of the electron optical column, the electron source and the probe forming objective lens, can be optimized for SEMPA. For reasonable

SEMPA acquisition times, electron sources must provide a 10 keV incident beam with a current of at least 1 nA to the specimen. It is this constraint that determines the SEMPA spatial resolution rather than the ultimate resolution of the column when used as a standard SEM. The selection of an electron source rests on the spatial resolution required for a specific micromagnetic measurement. Lanthanum hexaboride ( $\text{LaB}_6$ ) [2, 3], cold field emission cathodes [4–7], and thermal field emitters [32] have all been employed in SEMPA. The highest resolution, approximately 20 nm, has been achieved with field emitters. Thermal field emitters [33, 34] have somewhat larger source size than cold field emitters, but have greater stability (current variations  $<1\%$ ), high emission currents, and moderate energy width, making them well suited for use in high resolution SEMPA systems.

In SEM columns, the spherical aberration of the probe forming objective lens increases rapidly with increasing working distance, the distance between the lens exit pole face and the sample. For high resolution, one wants a short working distance. On the other hand, longer working distances are desirable to obtain a region at the sample surface free ( $<80 \text{ A/m}$ ) from the depolarizing effects of the stray magnetic field of the objective lens. Working distances between 5 and 15 mm provide a satisfactory trade off.

### 2.7.3.2 Transport Optics

The purpose of the transport optics is to efficiently collect and transfer the spin-polarized secondary electrons from the

specimen surface to the spin-polarization detectors without introducing instrumental asymmetries. Instrumental asymmetries are systematic errors which may be accounted for in a variety of ways. To reduce the effects of chromatic aberrations on the transported beam, the secondary electrons are first accelerated in a potential greater than about 500 V. In order to achieve the highest possible efficiency, the transport energy window of the electron optics should be about 8 eV wide and centered at 4 eV. The optical properties of low energy electron lenses used for transport can be computed from the numerical solution of Laplace's equation and subsequent ray tracing of the charged particle trajectories through the fields. An invaluable compendium of electron optical properties of common electron lens configuration has been compiled by Harting and Read [35]. At low magnification, the motion of the incident SEM beam on the specimen is translated into motion of the beam on the spin analyzer target leading to an instrumental asymmetry; a dynamic beam descanning scheme can be employed to remove scan related asymmetries [3]. The transport optics can be further optimized to reduce instrumental asymmetries and compensate for variations in the position of the beam [36].

### 2.7.3.3 Electron Spin Polarization Analyzers

Ideally a spin-polarimeter suitable for SEMPA should be efficient, small in size and compatible with the UHV ambient required for sample preparation. There has been considerable progress in reducing

the size and increasing the efficiency of electron spin polarimeters, yet spin detectors remain quite inefficient [37]. The polarization of a beam of  $N_0$  electrons is  $P_0 \pm \delta P$ , where the uncertainty is  $\delta P = 1/(N_0 F)^{1/2}$ . The figure of merit,  $F$ , for a spin polarization analyzer, rarely is much greater than  $10^{-4}$  even for the highest performance spin detectors. Thus, a polarization measurement with a relative uncertainty,  $\delta P/P_0 = 1/(P_0^2 N_0 F)^{1/2}$ , equivalent to the relative uncertainty in an intensity measurement,  $\delta N/N_0 = 1/(N_0)^{1/2}$ , can take over  $10^4$  times as long as the intensity measurement solely due to the inefficiency of the electron polarimeter.

Most spin polarimeters rely on a spin-orbit interaction for spin sensitivity. When an electron scatters from a central potential  $V(\mathbf{r})$ , the interaction of the electron spin  $s$  with its own orbital angular momentum  $\mathbf{L}$  [38] has the effect of making the cross section larger or smaller for electrons with spin parallel or antiparallel to  $\mathbf{n}$ , the unit vector normal to the scattering plane. The scattering plane is defined by the incident electron wave-vector  $\mathbf{k}_i$  and scattered electron wavevector  $\mathbf{k}_f$  such that  $\mathbf{n} = (\mathbf{k}_i \times \mathbf{k}_f)/|\mathbf{k}_i \times \mathbf{k}_f|$ . The cross section for the spin-dependent scattering can be written [38] as

$$\sigma(\Theta) = I(\Theta)[1 + S(\Theta)\mathbf{P} \cdot \mathbf{n}] \quad (3)$$

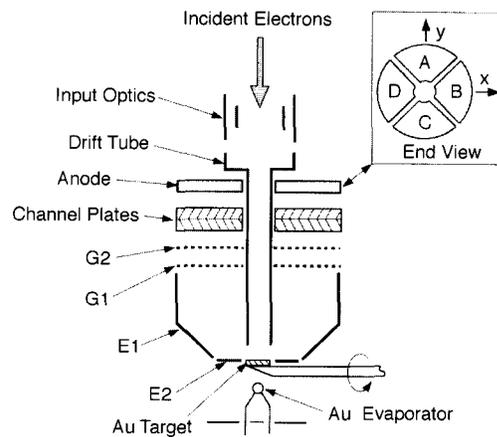
where  $I(\Theta)$  is the angular distribution of back scattered current in the detector and  $S(\Theta)$  is the Sherman function for the detector scattering material at the scattering angle,  $\Theta$ . The Sherman function is a measure of the strength of the spin-dependent scattering in the detector [39]. Typical values for  $S$  are  $|S| < 0.3$ . The polarization of the beam is determined from a spatial asymmetry  $A$  between the number of

electrons scattered to the left,  $N_L$ , and to the right,  $N_R$ , relative to the incident beam direction. The measured scattering asymmetry,  $A$ , is

$$A = (N_L - N_R)/(N_L + N_R) = PS \quad (4)$$

Differences in the left/right scattering can also arise from instrumental asymmetries and cause systematic errors that contribute to the uncertainty in the polarization measurement. These instrumental asymmetries result from: (i) unequal gains in the left and right channels of the signal processing electronics; (ii) unequal sensitivities of the electron multipliers; and (iii) mechanical imperfections which result in a detector geometry that is not symmetric.

As an example of a scattering type spin analyzer, we describe the low energy diffuse scattering (LEDS) detector [36, 40] used in our work. A schematic of this analyzer is shown in Fig. 3. The analyzer is quite compact since it operates at 150 eV; in our design it is about 10 cm



**Figure 3.** Schematic of the low energy diffuse scattering spin polarization analyzer. The divided anode assembly is shown in the inset as viewed from the Au target.

long. It employs an evaporated polycrystalline Au target. The efficiency of this spin analyzer is increased by collecting the scattered electrons over large solid angles. The large ratio of electrons collected to the number incident compensates for the moderate Sherman function of about 0.1, and leads to a high figure of merit,  $2 \times 10^{-4}$  [36]. The electrodes E1 and E2 in Fig. 3 focus the electrons that are diffusely backscattered from the Au target such that their trajectories are nearly normal to the retarding grids G1 and G2. The energy selectivity of the retarding grids enhances the Sherman function while the collection efficiency is increased by electrodes E1 and E2. The anode, which is also shown in the inset of Fig. 3, is divided into quadrants. Two orthogonal components of the polarization vector transverse to the electron beam direction ( $z$ ) may be measured simultaneously with this detector as

$$P_x = 1/S(N_C - N_A)/(N_C + N_A) \quad (5a)$$

$$P_y = 1/S(N_B - N_D)/(N_B + N_D) \quad (5b)$$

where  $N_i$  is the number of electrons counted by each quadrant ( $i = A, B, C, D$ ).

No single electron spin analyzer has all the features one might desire for highest performance in a SEMPA application. The traditional Mott analyzer which utilizes the asymmetry of the spin-dependent high energy (100–200 keV) electrons [38, 41, 42] has a Sherman function  $S$  that is larger than that of the LEDS detector and a comparable figure of merit. The high energy operation makes it less susceptible to apparatus asymmetries than low energy spin analyzers which therefore require more care in design of the transport optics. However, operation at the required high

voltage leads to large detector sizes making the Mott analyzer challenging to integrate with the SEM. Nevertheless, such analyzers have been used quite successfully for SEMPA [4, 7]. A low energy electron diffraction (LEED) electron-spin polarization analyzer [43, 44] has also been used very successfully for SEMPA [5, 6]. The collimating properties of diffraction by a single crystal, usually W(100) at about 100 eV, increase the efficiency of this spin analyzer leading to a relatively compact analyzer with a competitive figure of merit.

### 2.7.3.4 Electronics and Signal Processing

The electron signal is measured with surface barrier Si detectors, channeltrons, or stacked microchannel plates with a segmented anode, respectively, in the Mott, LEED, and LEDS spin analyzers. Each pair of detectors determines a component of the spin polarization vector transverse to the beam. For pulse counting, each signal channel consists of a preamplifier, amplifier/discriminator, and a scaler that is read by the computer. The signal processing electronics for the LEDS spin analyzer have been realized in both the pulse counting and analog modes; we describe aspects of each of these methods below.

When the electron probe beam is focused to very high spatial resolution in the SEM column, the beam current is reduced and pulse counting in the polarimeter is necessary. The short pulse widths (above 1 ns) in stacked microchannel plates facilitate high speed counting.

Dark currents are typically less than  $1 \text{ count cm}^{-2} \text{ s}^{-1}$ . In pulse counting, the quadrant anode structure must be designed to minimize cross talk between adjacent channels. Otherwise, pulses from adjacent channels will appear at reduced amplitude making pulse discrimination difficult. Reduction of inter-anode capacitance and capacitive coupling from each anode to common surrounding surfaces is effective in reducing this problem. Fast (20 ns) charge sensitive preamps can also be used. These are less affected by crosstalk than voltage preamps because they average over the characteristic ringing signal of capacitively coupled cross talk.

Fast, low resolution magnetization imaging with high incident current is very useful to survey a sample. At high incident beam currents, analog signal processing becomes necessary since the microchannel plates are count rate limited [45, 46]. With separate direct-coupled outputs from the anode quadrants, it is straightforward to switch over to an analog measurement of the current to the quadrants. When analog detection is used at high incident beam currents, the channel plate bias voltage is decreased to maintain linear gain. Each anode pair is connected to signal processing electronics including matched current-to-voltage converters, and sum and difference amplifiers. (Alternatively, the sum and difference can be performed later in the computer). The sum and difference signals are converted to pulse trains by separate voltage to frequency converters. Optocouplers provide the isolation necessary for the input stages to operate at the microchannel plate anode voltage and the signals are counted with a conventional scaler and timer system. Since the difference signal may change

sign, an offset voltage is applied to that voltage-to-frequency converter to prevent zero crossing and minimize digitization errors [3].

## 2.7.4 System Performance

The performance of a SEMPA system can be analyzed by examining the efficiency of the entire production, collection and processing chain. Although some of the considerations in the analysis are generally applicable to any SEMPA system, in order to provide specific numbers we will give parameters for our SEMPA system with the LEDS spin analyzer [2, 3]. The production efficiency of secondary electrons by a 10 keV electron beam at the surface of a ferromagnetic specimen tilted by  $45^\circ$ , is roughly 0.45 [11]. Only 37% of the secondary electrons produced at the sample are collected since the extraction optics only collect a narrow secondary electron energy window,  $4.0 \text{ eV} \pm 4.0 \text{ eV}$ . The efficiency of the transport optics between the extraction aperture and the spin analyzer may be as high as 1.00, but for normal operation the transport efficiency is closer to 0.88. The scattering efficiency, or ratio of the current incident on the detector channel plate input to that incident on the Au target, is 0.04 for nominal operating conditions [36, 40]. The channel plate itself has a finite collection efficiency of about 0.85 [46] due to final cell size. The product of all of these factors is the collection efficiency of the system,  $\epsilon = 0.005$ . For 1 nA incident beam current, only 4 pA ( $0.004 I_0$ ) of secondary electrons will be detected in the electron polarimeter, or approximately 1 pA to each quadrant.

Signal levels and integration times required to reach selected signal-to-noise ratios can also be estimated [3]. The simplest case to consider is the image of two adjacent domains with oppositely directed magnetization. Assume that the sample is oriented along a single detector direction such that the measured component of the polarization will be  $+P$  in one domain and  $-P$  in the other. The total change in that polarization component between the two domains (i.e., the signal) is  $2P$ . For a polarization measurement limited by counting statistics [38], one standard deviation statistical error in the polarization,  $P = (1/S)(N_C - N_A)/(N_C + N_A)$ , is given by  $\delta P = (1/(N_C + N_A)S^2)^{1/2}$ . The particle number reaching any pair of detector quadrants ( $N_A + N_C$ ) in a time interval  $\tau$  is  $\epsilon(I_0/2e)\tau$ . The signal-to-noise ratio is

$$SNR \equiv 2P/\delta P = 2PS(\epsilon I_0 \tau / 2e)^{1/2} \quad (6)$$

The dwell time required for each pixel in the image as a function of the desired signal-to-noise ratio and the experimental parameters is

$$\tau = (SNR)^2 e / (2P^2 S^2 \epsilon I_0) \quad (7)$$

The upper limit on the count rate will be set by the channel plate response. Assuming that the incident electron beam current in the electron microscope column is  $I_0 = 1 \text{ nA}$  and  $S = 0.11$ , the dwell time per pixel for various signal-to-noise ratios, and polarizations is given in Table 2. The elements in the table must be multiplied by the number of pixels in an image for the total data acquisition time. Thus, it takes about 54 s to acquire a  $256 \times 256$  pixel image with a signal-to-noise ratio of 5 and a mean polarization of 0.20. For analog signal detection, it is possible to reduce the noise introduced by the analog

**Table 2.** Pixel dwell time,  $\tau$  (ms), as a function of the SNR and polarization for SEM beam current,  $I_0 = 1 \text{ nA}$

$P$	$SNR = 2$	$SNR = 3$	$SNR = 5$	$SNR = 10$
0.01	53.185	119.666	332.407	1329.626
0.10	0.532	1.197	3.324	13.296
0.20	0.133	0.299	0.831	3.324
0.40	0.033	0.075	0.208	0.831

amplifier well below the shot noise of the incident beam for incident currents  $\geq 1 \text{ nA}$ . Hence the dwell times given in Table 2 also apply for analog signal acquisition.

## 2.7.5 Data Processing

Conventional image processing methods, such as filtering and contrast enhancement, can be used for processing SEMPA images. There are, however, some image processing steps that are unique to SEMPA since the contrast is derived from a vector magnetization and the spin detector sensitivity results from a scattering asymmetry. For SEMPA, common image processing steps include the subtraction of a zero offset and a background asymmetry. To do this, use is made of the fact that the magnitude of the magnetization  $|M|$  is constant. Consider the common case where the magnetization vector lies entirely in-plane. (This is expected for all but materials with particularly large magnetic anisotropy perpendicular to the surface.) In this case, it is possible to subtract a background and remove zero offsets by requiring that the in-plane magnetization,  $(M_x^2 + M_y^2)^{1/2}$ , have constant magnitude. In general, the

background subtracted may be nonlinear and fit with a polynomial. In special cases, the specimen topography can cause trajectories which couple with instrumental asymmetries to produce artifacts in the polarization measurements. In such cases the polarization detector's gold target can be replaced by a low atomic number target, such as graphite, for which there is no spin-dependent scattering asymmetry. The image acquired with the graphite target is then subtracted from that acquired with the Au target to remove instrumental asymmetries.

There are two basic formats to represent SEMPA data. The first uses the projection of the magnetization on orthogonal axes (i.e.,  $M_x$  and  $M_y$ ) and uses a gray map encoding scheme where white (black) represents the maximum value of the magnetization in the positive (negative) directions. The second format uses the magnitude  $|M|$  and the angle  $\vartheta$  of the magnetization vector projected onto some plane. Whether it is easier to identify a surface magnetic domain structure in  $M_x$  and  $M_y$  images or in  $|M|$  and  $\vartheta$  images depends largely upon the surface magnetic microstructure. The magnitude of the magnetization is determined as

$$|M| = [M_x^2 + M_y^2 + M_z^2]^{1/2} \quad (8)$$

and the direction, with respect to the positive  $x$ -axis of the in-plane magnetization, is (in the absence of any out-of-plane components)

$$\vartheta = \tan^{-1}(M_y/M_x) \quad (9)$$

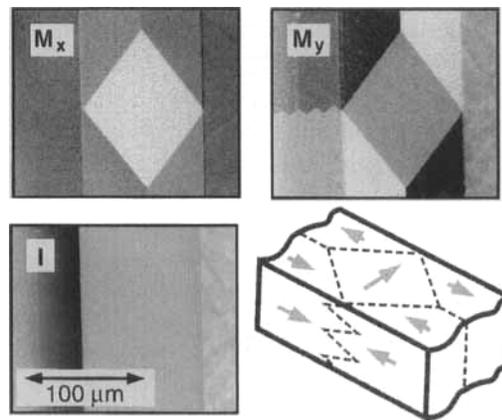
The map of the angle  $\vartheta$  can be displayed using color where the direction is read from an accompanying color wheel. Alternatively, it is sometimes helpful to

visualize the magnetization pattern by using small arrows to create a vector map.

## 2.7.6 Examples

### 2.7.6.1 Iron Single Crystals

The large magnetic moment per atom of Fe leads to a large intrinsic secondary electron polarization which makes Fe a favorable specimen to use for demonstrating SEMPA features. Figure 4 shows SEMPA images [2] of the (100) surface of a high quality Fe single crystal whisker. In the image labeled I, one observes the flat featureless upper surface of the whisker running vertically, centered in the frame. The image of the  $x$  component of the magnetization,  $M_x$ , shows a diamond-shaped domain with magnetization pointing to the right. The domain pattern is

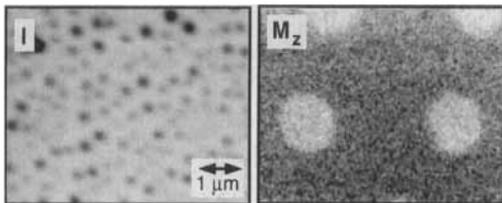


**Figure 4.** SEMPA measurements of the topography,  $I$ , and images of two components of the magnetization,  $M_x$  and  $M_y$ , from an Fe whisker. The depth of focus is demonstrated by domains clearly visible on the top and side of this slightly tilted sample of rectangular cross section.

shown schematically in the line drawing. The region to the right of the diamond in the figure is the nonmagnetic sample holder; to the left is the side of the crystal. The image of the vertical component of the magnetization,  $M_y$ , contains domains with magnetization pointing in the  $+y$  direction (white) and in the  $-y$  direction (black). The zig-zag domain wall visible in this image actually runs down the side of the whisker and is visible because the sample is slightly tilted. This large depth of focus is characteristic of SEMPA since a scanning electron microscope is used as the probe.

### 2.7.6.2 CoPt Magneto-optic Recording Media

In magneto-optic recording, information is stored by using a focused laser beam to read and write the bits. A bit is written when the laser locally heats the media in an applied magnetic field. The result is seen in Fig. 5 where the white dots correspond to magnetization out of plane,  $M_z$ , in the  $+z$  direction contrasted against a background previously magnetized in the  $-z$  direction. The corresponding intensity image shows the nonuniform topography of the Co-Pt

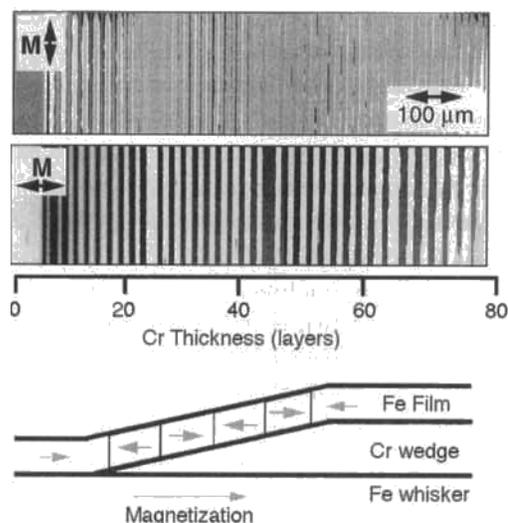


**Figure 5.** The topography,  $I$ , and perpendicular magnetization,  $M_z$ , are imaged from an CoPt magneto-optic recording medium. The round bits are approximately  $1.4 \mu\text{m}$  in diameter.

multilayer sample. In use, the bits are read by sensing the rotation of polarization of reflected light from the surface. The circular bits observed in Fig. 5 are about  $1.4 \mu\text{m}$  in diameter. A large perpendicular magnetic anisotropy is necessary to overcome the increased magnetostatic energy of out-of-plane magnetization. Domains with out-of-plane magnetization have also been observed on Co(0001) single crystals [47] and on TbFeCo magneto-optic storage media [48].

### 2.7.6.3 Exchange Coupling of Magnetic Layers

Two ferromagnetic layers separated by a nonferromagnetic layer may be exchange coupled such that the magnetic moments in the two ferromagnetic layers are parallel (ferromagnetic exchange coupling) or antiparallel (antiferromagnetic exchange coupling) depending on the spacer layer material and its thickness. An example is two Fe layers separated by a Cr layer to form a sandwich structure Fe/Cr/Fe(100). SEMPA is particularly well suited to determine the period (or periods) of oscillation of the exchange coupling between the magnetic layers as a function of spacer layer thickness [49, 50]. For example, consider the Fe/Cr/Fe(100) sandwich structure shown schematically in Fig. 6. A varying thickness Cr 'wedge' is deposited on the Fe(100) whisker substrate. This is covered with an Fe film approximately 10 layers thick. As shown in the schematic, the Fe layers are ferromagnetically coupled for small Cr thickness, and the sign of the coupling oscillates as the Cr thickness increases.



**Figure 6.** The magnetization images,  $M_x$  and  $M_y$ , of the top layer of a Fe/Cr/Fe(100) sandwich structure, shown schematically at the bottom, provide a precise measure of the oscillation of the sign of the magnetic exchange coupling as the thickness of the Cr spacer layer increases.

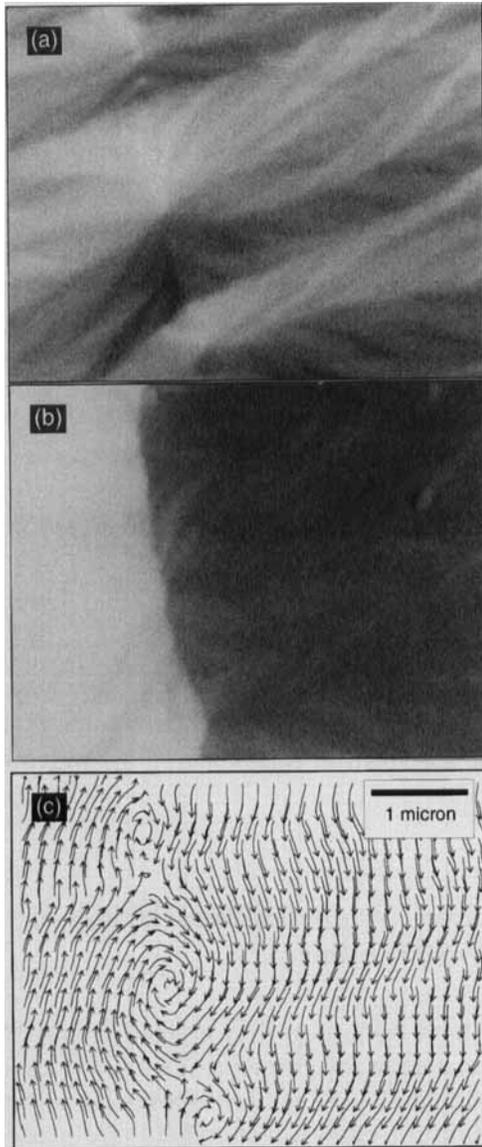
The SEMPA image of the magnetization in the direction of the wedge shows many changes in the magnetization as the Cr increases in thickness to nearly 80 layers (1 layer = 0.14 nm) over the length of the wedge shown in the image, about 0.8 mm. Such an image allows a very precise determination of the period of the exchange coupling. At Cr thicknesses where there is a reversal in the magnetic coupling, one observes a component of the magnetization transverse to the wedge direction as seen in the upper magnetization image. This provided early evidence for a different kind of coupling known as biquadratic exchange coupling.

Several features of SEMPA were successfully exploited in these studies. The high spatial resolution of the SEM allows the use of a small, nearly perfect Fe single crystal whisker substrate. The ultrahigh

vacuum allows deposition of a Cr wedge in situ. The clear advantage of the wedge structure is that it allows measurements at many different thicknesses with a reproducibility that could not be obtained by producing multiple films of uniform thickness. With a reflection high energy diffraction (RHEED) screen below the sample stage, and using the SEM column as a RHEED gun, it is possible to make spatially resolved RHEED measurements along the Cr wedge to determine the thickness of the interlayer material with single atomic layer resolution. The surface sensitivity of SEMPA allows the observation of the changes in the magnetization direction of the Fe overlayer without any interference from the lower Fe layer. Finally, since the SEMPA system used for these measurements was part of a scanning Auger microprobe, Auger spectroscopy could be utilized to monitor cleanliness of the specimen at each stage of its preparation.

#### 2.7.6.4 Magnetic Singularities in Fe-SiO<sub>2</sub> Films

A high resolution field emission SEMPA was used to image magnetic singularities in granular Fe-SiO<sub>2</sub> films above the percolation threshold [51]. These are nanocomposite materials with highly isotropic magnetic properties [52]. Figure 7 shows a segment of a 180° domain wall with two cross ties and a dramatic ripple pattern. Fig. 7a depicts the  $x$  component of the magnetization. This image clearly shows the fine structure of the ripple pattern; the cross ties appear as diamond shaped regions elongated in the direction



**Figure 7.** SEMPA images of the in-plane magnetization in a Fe-SiO<sub>2</sub> film near a domain wall: (a) *x*-component image; (b) *y*-component image; (c) vector map representation.

orthogonal to the wall. In the *y* component image, Fig. 7b, the wall itself is readily apparent. The vector map representation given in Fig. 7c shows the coarse structure

of the ripple and the vortices that occur between cross ties.

Because SEMPA is only sensitive to the magnetization at the surface, it is ideally suited to an investigation of surface features such as the singularities in Fig. 7. Line scans taken across the large central vortex indicate that the in-plane component of the magnetization decreases within the vortex showing that there must be a perpendicular component to the magnetization. The resolution of these images was determined to be 20 nm by analysis of line scans across several surface features.

## Acknowledgements

The iron whiskers were grown at Simon Fraser University under an operating grant from the National Science and Engineering Research Council of Canada. We thank all participating members of the Electron Physics Group for numerous contributions. This work was supported in part by the Office of Naval Research.

## 2.7.7 References

- [1] This technique is sometimes referred to as spin polarized SEM. We prefer the name SEMPA because the SEM beam is not spin polarized.
- [2] J. Unguris, M. R. Scheinfein, R. J. Celotta, D. T. Pierce in *Chemistry and Physics of Solid Surfaces VII* (Eds.: R. Vanslow, R. Howe), Springer Verlag, Germany **1990**, p. 239.
- [3] M. R. Scheinfein, J. Unguris, M. H. Kelley, D. T. Pierce, R. J. Celotta, *Rev. Sci. Instrum.* **1990**, *61*, 2501.
- [4] H. Matsuyama, K. Koike, *Rev. Sci. Instrum.* **1991**, *62*, 970; *J. Electron. Microsc.* **1994**, *43*, 157.
- [5] H. P. Oepen, J. Kirschner, *Scanning Microsc.* **1991**, *5*, 1.
- [6] H. P. Oepen, *J. Magn. Magn. Mat.* **1991**, *93*, 116.
- [7] R. Allenspach, *J. Magn. Magn. Mat.* **1994**, *129*, 160.
- [8] F. Bitter, *Phys. Rev.* **1931**, *38*, 1903.

- [9] K. Tsuno, *Rev. Solid State Science* **1988**, 2, 623.
- [10] J. P. Jacobovics, *Electron Microscopy in Materials Science Part IV* (Eds.: E. Ruedle, U. Valdre), Commission of European Communities, Brussels **1973**, p. 1305.
- [11] D. E. Newbury, D. E. Joy, P. Echlin, C. E. Fiori, J. I. Goldstein, *Advanced Scanning Electron Microscopy and X-ray Microanalysis*, Plenum Press, New York **1986**.
- [12] A. Tonomura, *Rev. Mod. Phys.* **1987**, 59, 639.
- [13] M. Mankos, M. R. Scheinfein, J. M. Cowley, *J. Appl. Phys.* **1994**, 75, 7418.
- [14] J. N. Chapman, *J. Phys.* **1984**, D17, 623; J. N. Chapman, R. Ploessl, D. M. Donnet, *Ultramicroscopy* **1992**, 47, 331; G. R. Morrison, H. Gong, J. N. Chapman, V. Hrnčiar, *J. Appl. Phys.* **1988**, 64, 1338.
- [15] P. Grutter, H. J. Mamin, D. Rugar, *Springer Ser. Surf. Sci.* **1992**, 28, 151.
- [16] J. Kranz, A. Z. Hubert, *Z. Angew. Phys.* **1963**, 15, 220.
- [17] J. Stohr, Y. Wu, B. D. Hermsmeier, M. G. Samant, G. R. Harp, S. Koranda, D. Dunham, B. P. Tonner, *Science* **1993**, 259, 658.
- [18] K. Grzelakowski, T. Duden, E. Bauer, H. Poppa, S. Chiang, *IEEE Trans. Magn.* **1994**, 30, 4500.
- [19] R. J. Celotta, D. T. Pierce, *Microbeam Analysis-1982* (Ed.: K. F. J. Heinrich), San Francisco Press, San Francisco **1982**, p. 469.
- [20] K. Koike, K. Hayakawa, *Jpn. J. Appl. Phys.* **1984**, 23, L187.
- [21] J. Unguris, G. G. Hembree, R. J. Celotta, D. T. Pierce, *J. Microscopy* **1985**, 139, RP 1.
- [22] K. Koike, H. Matsuyama, K. Hayakawa, *Scanning Microsc. Supp.* **1987**, 1, 241.
- [23] G. G. Hembree, J. Unguris, R. J. Celotta, D. T. Pierce, *Scanning Microsc. Supp.* **1987**, 1, 229.
- [24] J. Unguris, D. T. Pierce, A. Galejs, R. J. Celotta, *Phys. Rev. Lett.* **1982**, 49, 72.
- [25] H. Hopster, R. Raue, E. Kisker, G. Guntherodt, M. Campagna, *Phys. Rev. Lett.* **1983**, 50, 70.
- [26] E. Kisker, W. Gudat, K. Shroder, *Solid State. Commun.* **1982**, 44, 591.
- [27] D. R. Penn, S. P. Apell, S. M. Girvin, *Phys. Rev. Lett.* **1985**, 55, 518; *Phys. Rev.* **1985**, 32, 7753.
- [28] J. Unguris, M. W. Hart, R. J. Celotta, D. T. Pierce, Proc. 49th EMSA (Ed.: G. W. Bailey), San Francisco Press, San Francisco **1991**, p. 764.
- [29] J. Unguris, R. J. Celotta, D. T. Pierce, *Phys. Rev. Lett.* **1992**, 69, 1125.
- [30] J. Unguris, R. J. Celotta, D. T. Pierce, *J. Mag. Mag. Mat.* **1993**, 127, 205.
- [31] M. R. Scheinfein, J. Unguris, J. L. Blue, K. J. Coakley, D. T. Pierce, R. J. Celotta, P. J. Ryan, *Phys. Rev.* **1991**, B43, 3395.
- [32] A. Gavrin, M. H. Kelley, J. Q. Xiao, C. L. Chien, *Appl. Phys. Lett.* **1995**, 66, 1683.
- [33] D. W. Tuggle, J. Z. Li, L. W. Swanson, *J. Microsc.* **1985**, 140, 293.
- [34] D. W. Tuggle, L. W. Swanson, *J. Vac. Sci. Technol.* **1985**, B3, 193.
- [35] E. Harting, F. H. Read, *Electrostatic Lenses*, Elsevier, Amsterdam, **1976**.
- [36] M. R. Scheinfein, D. T. Pierce, J. Unguris, J. J. McClelland, R. J. Celotta, *Rev. Sci. Instr.* **1989**, 60, 1.
- [37] D. T. Pierce, R. J. Celotta, M. H. Kelley, J. Unguris, *Nuc. Inst. Meth.* **1988**, A266, 550.
- [38] J. Kessler, *Polarized Electrons, 2nd ed.*, Springer-Verlag, Berlin **1985**.
- [39] M. Fink, A. C. Yates, *At. Data* **1970**, 1, 385; M. Fink, J. Ingram, *At. Data* **1972**, 4, 129.
- [40] J. Unguris, D. T. Pierce, R. J. Celotta, *Rev. Sci. Instrum.* **1986**, 57, 1314.
- [41] E. Kisker, R. Clauberg, W. Gudat, *Rev. Sci. Instrum.* **1982**, 53, 1137.
- [42] G. D. Fletcher, T. J. Gay, M. S. Lubell, *Phys. Rev.* **1986**, A34, 911.
- [43] J. Kirschner, *Polarized Electrons at Surfaces*, Springer-Verlag, Berlin, **1985**.
- [44] J. Sawler, D. Venus, *Rev. Sci. Instrum.* **1991**, 62, 2409.
- [45] J. L. Wiza, *Nucl. Instrum. Meth.* **1979**, 162, 578.
- [46] E. H. Eberhardt, *IEEE Trans. Nucl. Sci.* **1981**, NS 28, 712.
- [47] J. Unguris, M. R. Scheinfein, R. J. Celotta, D. T. Pierce, *Appl. Phys. Lett.* **1989**, 55, 2553.
- [48] M. Aeschlimann, M. Scheinfein, J. Unguris, F. J. A. M. Greidanus, S. Klahn, *J. Appl. Phys.* **1990**, 68, 4710.
- [49] J. Unguris, R. J. Celotta, D. T. Pierce, *Phys. Rev. Lett.* **1991**, 67, 140.
- [50] D. T. Pierce, J. Unguris, R. J. Celotta in *Ultra-thin Magnetic Structure II* (Eds.: B. Heinrich, J. A. C. Bland), Springer-Verlag, Berlin **1994**, p. 117.
- [51] A. Gavrin, J. Q. Xiao, C. L. Chien, to be published.
- [52] C. L. Chien, *J. Appl. Phys.* **1991**, 69, 5267.

## Index

- Abbe theory 266
- Lorentz microscopy 266
- absorption 29, 31, 32, 35
- anomalous 32, 35
- dynamical equations 31
- transmission electron microscopy 29 ff
- adsorption layers 251
- low-energy electron microscopy 251
- advanced universal Auger electron microscope 443, 444
- experimental set-up 443
- Ag-Sn 96
- extended dislocations 96
- Aharonov–Bohm effect 272
- conventional transmission electron microscopy 272
- Al<sub>2</sub>O<sub>3</sub> (011) 174
- reflection electron microscopy 174
- AlN 89
- dislocation in 89
- a*-alumina 175
- reflection electron microscopy 175
- aluminum–zirconium 217
- crystalline–amorphous transition 217
- aluminum 94, 206
- high voltage electron microscopy 206
- unfaulted dislocation loops 94
- aluminum crystals 203
- cell formation 203
- dislocation density 203
- amplitude–phase diagram 13, 14, 61 ff
- dislocation contrast 61 ff
- transmission electron microscopy 13, 14
- analog detection 491
- scanning electron microscopy with polarization analysis 491
- analytical electron microscopes 439
- experimental set-up 440
- analytical scanning transmission electron microscopy 440
- experimental set-up 440
- analyzing crystals 445
- analysis range of 445
- annealing phenomena 215
- high voltage electron microscopy 215
- annular dark-field imaging 330
- heavy atoms 330, 334
- high angle 333
- platinum on zeolite 334
- uranium 330
- anomalous absorption 41
- in cubic-close-packed structures 41
- anorthite 99, 101
- antiphase boundaries 101
- domain boundaries 99
- aperture function 110
- area analysis 447
- energy-dispersive spectrometry 447
- artificial neural networks 200
- electron energy loss spectroscopy 200
- artifacts 288
- electron holography 288
- as-grown semiconductor crystals 323
- electron beam induced channeling image 323
- As<sub>2</sub>Te<sub>3</sub>(GeTe)<sub>*n*</sub> 143
- high-resolution electron microscopy 143
- atom implantation 219
- electron irradiation induced 219
- atomic number correction 449
- in quantitative X-ray scanning microanalysis 449
- atomic scattering factor 7
- transmission electron microscopy 7
- Au–Ni multilayer 198
- electron energy loss spectroscopy 198
- Au<sub>4</sub>Mn 138, 139, 158
- electron density 158
- high-resolution electron microscopy 138, 139
- Auger electron emission microscopy 247
- instrumentation 246
- Auger electron microscope 443
- experimental set-up 443
- Auger electron spectroscopy 391, 393
- beam effect 391
- quantitative analysis 393
- sample preparation 391
- Auger electrons 388
- Auger energies 390
- Auger notations 389
- Auger spectrum 390
- line shape 390
- Auger transitions 389, 390
- kinetic energy 389, 390
- Auger-electron microanalyzers 442
- β*-conductivity 320, 321
- electron beam induced conductivity 321
- Ba<sub>2</sub>NaNb<sub>5</sub>O<sub>15</sub> 136
- exit wave function 136
- background 396
- Auger electron spectroscopy 396
- X-ray photoelectron spectroscopy 396

- background slope effects 422
  - scanning Auger microscopy 422
- backscattered electron imaging 314
  - applications 315
  - atomic number contrast imaging 315
  - detectors 314
  - limitation 314 f
- backscattered electrons 314, 316
  - range 316
  - yield 314
- backscattering coefficient 316
- BaCuO<sub>2</sub> 236
  - convergent beam electron diffraction 236
- Ba<sub>x</sub>Mn<sub>8</sub>O<sub>16</sub> 144, 146
  - high-resolution electron microscopy 144, 146
- BC<sub>3</sub> thin foil 191
  - electron energy loss spectroscopy 191
- beam current fluctuation effects 423
  - scanning Auger microscopy 423
- beam effect 391, 393
  - Auger electron spectroscopy 391
  - X-ray photoelectron spectroscopy 392
- Beamson approach 420
- binary alloys 129
  - building block structure 129
- biological applications 479
  - imaging secondary ion mass spectrometry 479
- Bitter method 484
  - imaging magnetic microstructure 484
- Bloch wave formalism 233
  - higher order Laue zone reflections 233
- Bloch wave theory 227
- borosilicate glass 434
  - scanning microanalysis 434
- Borrmann effect 29, 32, 45
- Bravais lattice 225
  - convergent beam electron diffraction 225
- bremsstrahlung 435
  - to calibrate the film thickness 435
- bremsstrahlung intensity 435
- building block structures 128
  - image interpretation 128
- bulk magnetic structure 486
  - scanning electron microscopy with polarization analysis 486
- butylbenzotriamide 477
  - secondary ion mass spectrometry 477
- C<sub>60</sub> 146
  - high-resolution electron microscopy 146
- CaAl<sub>2</sub>Si<sub>2</sub>O<sub>8</sub> 101
  - antiphase boundaries 101
- CaF<sub>2</sub>/silicon interface 318
  - channeling pattern, selected area 318
- cage structures 146
  - high-resolution electron microscopy 146
- Cameca IMS secondary ion mass spectrometer 464 f
  - experimental set-up 464
  - mass range 465
  - ion microscope mode 465
- Cameca TOF secondary ion mass spectrometer 466, 467
  - detection system 467
  - experimental set-up 466
  - ion sources 467
  - mass resolution 467
- camera obscura 109
- carbon nanotubes 343, 357
  - nanodiffraction patterns 343
- scanning transmission electron microscopy 357
- carbon structures 146
  - high-resolution electron microscopy 146
- carrier diffusion length 322
  - electron beam induced channeling image 322
- catalysts 384
  - Z-contrast scanning transmission electron microscopy 384
- cathodoluminescence 320, 324 ff
  - II–VI compound semiconductors 324
  - modes of operation 326
  - signal collection 325
  - spatial resolution 326
- cathodoluminescence in scanning electron microscopy 325
  - experimental set-up 325
- CCD cameras 122, 282
  - electron holography 282
  - high-resolution electron microscopy 122
- CCD detectors 351
  - scanning transmission electron microscopy 351
- CdMnTe 240
  - convergent beam electron diffraction 240
- CdTe(001)–GaAs(001) interface 377, 378
  - edge dislocation in 377
  - Z-contrast scanning transmission electron microscopy 377, 378
- ceramics 379
  - Z-contrast scanning transmission electron microscopy 379
- channeling 128
  - high-resolution electron microscopy 128
- channeling parameters 207
  - high voltage electron microscopy (table) 207
- channeling patterns 316 ff
  - backscattered electron imaging 316
  - CaF<sub>2</sub>/silicon interface 318
  - information about crystal structure 317
  - limitations 318
  - quality of a crystal for determining 318

- selected area 318
- channeling theory 129
- charge collection mode 320
  - electron beam induced channeling 320
  - for examining semiconductor devices 321
- chemical composition determination, quantitative 396
  - Auger electron spectroscopy 396
  - X-ray photoelectron spectroscopy 396
- chemical shifts 391
  - Auger electron spectroscopy 391
- chemisorption 254
  - low-energy electron microscopy 254
- chonospectroscopy 188
  - electron energy loss spectroscopy 188
- chromatic contrast 472
  - secondary ion mass spectrometry 472
- chromium trihalides 89
  - dislocations 89
- Cliff-Lorimer sensitivity factor 453
  - analytical electron microscopy 453
- Co/Cu multilayers 264, 265
  - Lorentz microscopy 264, 265
- Co<sub>2</sub>Si 256
  - low-energy electron microscopy 256
- cobalt 294
  - quantum size effects in 294
- cobalt film 268
  - Lorentz microscopy 268
- cobalt layers 295, 298, 299
  - imaging by spin-polarized low-energy electron microscopy 295, 298, 299
  - magnetic domain imaging by SPLEEM 298
- coherence 276
  - electron holography 276
- coherent aberrations 273
  - electron holography 273
- coherent convergent beam electron diffraction 236 ff
  - scanning transmission electron microscopy 344
- column approximation 11, 12
  - transmission electron microscopy 11, 12
- columnar coherence 368 ff
  - Z-contrast scanning transmission electron microscopy 368
- composition-composition histogram 447
- concentric hemispherical analyzer 403, 404
  - focusing properties 404
- continuous beam reflective high-energy electron diffraction 353
  - scanning transmission electron microscopy 353
- conventional transmission electron microscopy 271 ff
  - Aharonov-Bohm effect 272
  - limitations 273
  - theory 271 f
- convergent beam electron diffraction 223 ff, 232, 234, 235, 238 ff
  - applications
    - BaCuO<sub>2</sub> 236
    - CdMnTe 240
    - InGaAs/InP superlattice 238
    - NbSe<sub>3</sub> 225
    - Si 237
    - Si/SiGe 239
    - silicon 229
  - coherent 236 ff
  - crystal structure determination 232
  - dislocations 240
  - large angle convergent beam electron diffraction 223
  - metastable alloys 234
  - quantitative 235 f
  - semiconductors 238
  - Si/SiGe 239
- convergent beam interferometry 284
- convergent-beam scanning transmission electron microscopy 343
- copper on Mo(110) 255
  - low-energy electron microscopy 255
- CoPt magneto-optic recording medium 494
  - scanning electron microscopy with polarization analysis 494
- CoSi<sub>2</sub>-Si (111) epitaxial interface 372
  - Z-contrast imaged 372
- CoSi<sub>2</sub> 256
  - layer growth studied by LEEM 256
  - low-energy electron microscopy imaging 256
- CoSi<sub>2</sub>/(111)Si 240
  - applications
    - convergent beam electron diffraction 240
- Coster-Kronig transitions 389
- CrCl<sub>3</sub> 89
  - dislocation in 89
- critical voltage effect 231
  - convergent beam electron diffraction 231
- cross-grating diffraction 229
  - Bloch wave theory 229
- crystal structure determination 232
  - convergent beam electron diffraction 232
- crystal symmetry determination 224
  - convergent beam electron diffraction 224
- crystalline-amorphous transition 217
  - electron irradiation induced 217 ff
- crystallographic contrast 472
  - secondary ion mass spectrometry 472
- crystallographic defects 323
  - electron beam induced channeling image 323

- Cu-Ga 90 ff
  - dislocation in 91
  - glide dislocations 90, 92
  - NbTe<sub>2</sub> 91
- Cu<sub>3</sub>Pd 99, 101
  - antiphase boundaries 99
  - diffraction contrast image 101
- Cu<sub>3</sub>Si 256
  - low-energy electron diffraction 256
  - low-energy electron microscopy 256
- cylindrical mirror analyzer 403, 415
  - double-pass 403
  - focusing properties 404
  - imaging properties as a function of energy 415
  - single-pass 403
  - use in X-ray photoelectron spectroscopy 405
- damping 112
- dark field image 125
  - definition 125
- Darwin-Howie-Whelan differential equations 27, 33
- datolite 434
  - scanning microanalysis 434
- deep level transient spectroscopy 323
- defects in a bulk wafer 323
  - scanning electron microscopy 323
- deflection type analyzers 400, 401
  - cylindrical sector fields 400
  - energy resolution 400
  - hemispherical analyzers 400
  - parallel plate analyzers 400
- defocus method of imaging 266
  - Lorentz microscopy 266
- deformation 212
  - high voltage electron microscopy 212
- deformed crystals 21, 24, 27, 33
  - diffraction models 21
  - dynamical diffraction 33
  - dynamical equations 27
  - lattice potential 24
- depth profiles 457
  - secondary ion mass spectrometry 457
- detection defined imaging 398
  - scanning Auger microscopy 398
- detection defined XPS imaging 415
- detection limit 393
  - Auger electron spectroscopy 394
  - X-ray photoelectron spectroscopy 393
- detector defined imaging 410, 414
  - 2D position analysis 415
  - detection efficiency 415
  - X-ray photoelectron spectroscopy 410, 414
- determination of order parameters 104
  - diffraction contrast transmission electron microscopy 104
  - determination of the type of stacking fault 104
  - diffraction contrast transmission electron microscopy 104
  - diamond 217
    - dislocation in 326
      - cathodoluminescence image 326
    - high voltage electron microscopy 217
    - planar defects in 356
  - scanning transmission electron microscopy 356
- diatoms 312
  - scanning electron microscopy 312
- differential phase contrast microscopy 484
  - imaging magnetic microstructure 484
- diffraction 7
  - by crystals 7
  - diffraction amplitude 7, 9
    - transmission electron microscopy 7 f
  - diffraction conditions 8
    - transmission electron microscopy 8 f
  - diffraction contrast 85
    - applications
      - dislocations, fine structure 85
  - diffraction contrast electron microscopy 98
    - applications
      - minerals 99
      - ordered alloys 99
      - defects in semiconductors 100
  - diffraction contrast imaging 112
  - diffraction contrast mode 5 ff
    - transmission electron microscopy 4
  - diffraction contrast transmission electron microscopy 3 ff, 103, 104
    - applications
      - determination of order parameters 104
      - determination of the type of stacking fault 104
      - in situ studies 103
      - phase transition 103
      - radiation damage 103
  - diffraction patterns 124, 263
    - effect of magnetic domains 263
    - schematic representation 124
- dislocation contrast 23, 53, 61 ff, 67
  - amplitude-phase diagrams 61 ff
  - dynamical theory 67
  - kinematical theory 61
  - weak-beam method 63 ff
- dislocation dipoles 68
  - diffraction contrast transmission electron microscopy 68
- dislocation loops 173
  - reflection electron microscopy 173

- dislocation ribbons 74 ff
  - diffraction contrast transmission electron microscopy 74
  - imaging of 74, 76
- dislocations 22, 68, 85, 124, 141, 172, 240
  - applications
  - – convergent beam electron diffraction 240
  - bright field image 54
  - characterization of 79
  - dark field image 54
  - extinction criterion 68
  - fine structure 85
  - fringe pattern 22
  - reflection electron microscopy 172 ff
  - schematized model 124
  - study by high-resolution electron microscopy 141
- domain boundaries 21, 23, 38, 44, 46, 178
  - displacement function 21
  - matrix formulation 38
  - modes 23
  - reflection electron microscopy 178
- domain boundary contrast 23
- domain boundary fringes 43
- domain contrast 50
  - imaging of microtextures 50
- domain images 264
  - transmission electron micrographs 264
- domain structures 96
  - applications
  - – diffraction contrast electron microscopy 96
- domain textures 47
  - diffraction patterns 47
- double exposure hologram 281
  - off-axis image plane holography 281
- dynamic secondary ion mass spectrometry 458
- dynamical equations 27, 28, 31
  - absorption 31
  - deformed crystals 27
  - perfect crystal 28
- dynamical theory 67
  - dislocation contrast 67
  - two-beam 26
- edge dislocation 25, 56, 58, 73, 141, 173
  - diffraction contrast transmission electron microscopy 56, 72
  - image profiles 25
  - imaging of 55
  - reflection electron microscopy 173
  - study by high-resolution electron microscopy 141
  - transmission electron microscopy 53, 57
  - two-beam image characteristics 58
- edge effects 423
  - scanning Auger microscopy 423
- edge fringes 42, 44, 45
  - fringe profile 45
- EELS and incoherent structure imaging 378
- elastic scattering 364
  - incoherent imaging 364
- electric structures 283
  - electron holography 283
- electron–solid interactions 429
  - mechanism of 429
- electron–specimen interactions 439
  - electron probe X-ray microanalysis 435
- electron backscattering factor 450
  - in quantitative X-ray scanning microanalysis 449
- electron beam induced conductivity 320, 321
  - $\beta$ -conductivity 321
  - p–n junction 321
  - Schottky surface barrier 321
- electron beam induced current 321
- electron beam interactions 428
  - analytical electron microscopy 428
- electron beam shaping 6
  - transmission electron microscopy 6
- electron channeling 137, 156, 205 ff
  - diffraction pattern 159
  - isolated columns 157
  - parameterization 160
  - principle 156
- electron channeling pattern 317
  - InP 317
- electron diffraction 7, 9
  - convergent beam 223 ff
  - scattered amplitude 9
- electron energy loss spectroscopy 181 ff, 190 ff
  - applications
  - – artificial neural networks 200
  - – Au–Ni multilayer 198
  - – BC<sub>3</sub> thin foil 191
  - – elemental mapping 195
  - – Fe/Fe<sub>x</sub>Ge<sub>1–x</sub> multilayer 195
  - – LaBaCo<sub>2</sub>O<sub>5+d</sub> 188
  - – mapping of fine structures 198
  - – Nb mapping 196
  - – Si–SiO<sub>2</sub> interfaces 200
  - – Si–TiO<sub>2</sub> interface 199
  - data processing 190, 196
  - detection limit 193
  - energy-analyzing mode 182
  - energy-filtered images 184
  - energy-selecting mode 183
  - experimental set-up 183
  - information accessible through a detailed analysis (table) 189
  - instrumentation 182 ff
  - multiple least-squares fitting techniques 192

- parameters currently used for the acquisition of spectrum lines (table) 187
- point analysis 184
- quantitative analysis technique 190, 193
- spectrum imaging 184
- electron energy loss spectrum 188
  - characteristic features 188
- electron holography 178, 271 ff, 276, 283, 284, 287, 289, 346, 484
  - applications
    - electric structures 283
    - magnetic domain walls 283
    - magnetic structures 283
    - superconductors 283
    - surface analysis 284
    - thin films 283
  - artifacts 288
  - at high and medium resolution 283
  - coherence 276
  - comparison with SEMPA 484
  - corrections of aberrations 287
  - geometries 277 ff
  - high resolution 284 ff
  - image recording 286
  - in the reflection electron microscopy mode 178
  - object wave reconstruction 286, 289
  - operating principle 276
  - signal-to-noise properties 288
- electron interferometry, as a basis for off-axis electron holography 279
- electron irradiation induced methods 217 ff
- electron microanalysis 427 ff
- electron microscope 122, 204
  - resolution capability 122 ff
  - voltage dependence 204
- electron microscopy 109
  - high resolution 109 ff
  - high voltage 203 ff
  - reflection 163 ff
  - transmission 3 ff
- electron optical column 487
  - optimization for SEMPA analysis 487 f
- electron probe microanalysis 427 ff
- electron probe X-ray microanalysis 435
  - at low voltage 439
  - instrumentation 435
- electron probe X-ray microanalyzers 435 f
- electron scattering 118
  - in a thick object
    - derivation of mathematical expressions 152 ff
    - quantum mechanical approach 153
  - in a thin object
    - derivation of mathematical expressions 151 ff
  - schematic representation 118
- electron sources 5
  - transmission electron microscopy 5
- electron wave 271, 272
  - modulation of 271
  - propagation of 272
- electron-beam-induced current 428
- electrostatic transfer lens 408
  - retarding effect 408
- element mapping 195, 411
  - electron energy loss spectroscopy 195
  - scanning Auger microscopy 411
- elemental distributions 447
  - X-ray mapping 447
- energy filtering microscope 184
- energy loss distribution 341
  - scanning transmission electron microscopy 341
- energy resolution 401
  - deflection type analyzers 401
- energy selected imaging 184
  - electron energy loss spectroscopy 184
- energy-dispersive spectrometry 446
  - scanning emission electron microscopy 446
- energy-dispersive spectroscopy 432
  - comparison with wavelength dispersive spectroscopy 445
  - $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$  432
- energy-filtered images 184
- epitaxial growth 178
  - reflection electron microscopy 178
- Everhart-Thornley detector 311
- exit wave reconstruction 134
  - high-resolution electron microscopy 134
- extinction criteria 45, 55
  - dislocations 55
  - fringe profile 45
- extrinsic luminescence 324
- extrinsic stacking fault 104
  - diffraction contrast transmission electron microscopy 104
- $f(pz)$  methods 451
  - quantitative X-ray scanning microanalysis 451
- f.c.c. metals 174
  - deformation fault lines 174
- face-centered-cubic structures 105
  - determination of stacking faults 105
- Fe-SiO<sub>2</sub> film 495, 496
  - scanning electron microscopy with polarization analysis 495, 496
- Fe-Cr-Mo-Ti 47, 51
  - inversion boundaries 47, 51
- Fe/Cr/Fe(100) 494
  - sandwich structure 495

- scanning electron microscopy with polarization analysis 495
- scanning electron microscopy with polarization analysis 494
- Fe/Fe<sub>x</sub>Ge<sub>1-x</sub> multilayer 195
- electron energy loss spectroscopy 195
- ferromagnetic films 349
- off-axis holography 349
- ferromagnetic materials 294, 348, 483
- magnetic fields in 348
- scanning electron microscopy with polarization analysis 483
- spin-polarized low-energy electron microscopy 294
- field-effect device 102
- diffraction contrast transmission electron microscopy 102
- transmission electron microscopy 102
- field-emission gun 352
- scanning transmission electron microscopy 352
- first order Laue zone reflections 225
- convergent beam electron diffraction 225
- fitness function 131
- quantitative imaging in high-resolution electron microscopy 131
- fluorescence correction 450
- quantitative X-ray scanning microanalysis 450
- focus variation method 133
- phase retrieval 133
- foreign atom implantation 219 ff
- electron irradiation induced 219 ff
- Foucault method 264, 267
- Lorentz microscopy 264 f, 267
- Fourier holograms 278
- Fourier lenses 417
- energy selection 417
- imaging 417
- Fraunhofer holograms 278
- Fresnel holograms 278
- Fresnel method 266, 267
- Lorentz microscopy 266, 267
- fringe profile 39, 40, 42, 44, 45
- edge fringes 45
- extinction criteria 45
- planar interfaces 39
- translation interfaces 40, 45
  
- GaAs 166, 376
- dislocation in 326
- cathodoluminescence image 326
- doping variation in 323
- electron beam induced channeling image 323
- reflection electron microscopy 166
- Z-contrast scanning transmission electron microscopy 376
  
- Gabor holography 346
- limitation 346
- Gatan imaging filter 183
- germanium 141
- grain boundary 141
- Glaser unit 117
- glide dislocations 90
- observation by high voltage electron microscopy 90
- gold 94, 95
- fault 94
- stacking fault tetrahedra in 95
- gold on Mo(110) 253
- low-energy electron microscopy 253
- grain boundaries 139, 381, 382
- structure-property relationships 382
- study by high-resolution electron microscopy 139
- superconducting properties 381
- Z-contrast scanning transmission electron microscopy 381
- grain boundary engineering 382
- graphite 23, 77, 79, 86, 88, 93
- dislocations 23, 93
- dislocation ribbons 77, 79, 86, 88
- stacking faults 23
- graphite foil 13
- transmission electron microscopy 13
  
- heavy atoms 330, 333, 334, 353
- annular dark-field (ADF) imaging 330, 333, 334
- on zeolites 355
- scanning transmission electron microscopy 355
- scanning transmission electron microscopy 353
- heteroepitaxy 256
- low-energy electron microscopy 256
- on Si 254
- low-energy electron microscopy 254
- high resolution mode 4
- transmission electron microscopy 4
- high voltage electron microscopy 203 ff
- applications
- aluminum (111) 206
- annealing phenomena 215
- foreign atom implantation 219
- magnetic domain measurements 216
- martensitic transformations 215
- radiation effects 216
- study of deformation 212
- in situ experiments 208
- instrumentation 209
- quantitative studies 216 ff
- specimen treatment devices 208 ff

- high-angle annular dark-field imaging, heavy atoms 333 fff
- high-angle detector 368
  - incoherent imaging 368
- high-resolution electron microscopy 109 ff
- high- $T_c$  superconductors 144
  - high-resolution electron microscopy 144
- higher order Laue zone 224, 226
  - convergent beam electron diffraction 224
  - lattice parameter measurement 226
  - strain measurement 226
- higher order Laue zone lines 232
  - quantitative analysis 232
- higher order Laue zone reflections 231, 233
  - Bloch wave formalism 233
  - convergent beam electron diffraction 231 ff
  - kinematic approximation 233
- holographic reconstruction methods 132
  - quantitative imaging 132
- holography 178, 275 ff, 347
  - basis principle 275
  - in the reflection electron microscopy mode 178
  - off-axis technique 279
  - scanning transmission electron microscopy 347
- homoepitaxy 255
  - low-energy electron microscopy 255
- II–VI compound semiconductors 324
  - cathodoluminescence 324
- image artefacts 420
  - scanning Auger microscopy 420
- image contrast 169 ff, 336
  - reflection electron microscopy 169, 170
  - scanning transmission electron microscopy 336
- image contrasting, theory 266
  - Lorentz microscopy 266
- image density distribution 337
  - scanning transmission electron microscopy 337
- image depth sequences 473 f
  - secondary ion mass spectrometry 473
- image formation 109 ff
  - derivation of mathematical expression 146 ff
  - high-resolution electron microscopy 109 ff
  - pinhole camera 113
  - transmission electron microscope 114
- image intensity 338
  - scanning transmission electron microscopy 338
- image intensity distribution 339
  - scanning transmission electron microscopy 339
- image interpretation 125, 312
  - high-resolution electron microscopy 125
  - secondary electron imaging 312
- image plane holograms 278
- image processing 492
  - scanning electron microscopy with polarization analysis 492
- image reconstruction 281
  - off-axis image plane holography 281
- image recording 286, 462
  - electron holography 286
  - secondary ion mass spectrometry 462
- image simulation 67, 126, 130, 131, 152
  - commercial software packages 131
  - dislocation contrast 67
  - high-resolution electron microscopy 126, 130
  - multislice method 152
- image-processing 473
  - secondary ion mass spectrometry 473
- imaging modes 332
  - scanning transmission electron microscopy 332 ff
- imaging secondary ion mass spectrometry 457 ff
  - comparison of SIMS with other imaging microanalytical and surface analytical techniques (table) 459
  - data acquisition modes 458
  - experimental set-up 458
- immersion lens 307
  - scanning electron microscopy 307
- imperfect crystals 238
  - large angle convergent beam electron diffraction 238
- imperfections in crystals 344
  - scanning transmission electron microscopy 344
- impulse response function 110, 114, 115
- in situ studies 103
  - diffraction contrast transmission electron microscopy 103
- in-line holograms 277
- in-line holography 277 ff
  - resolution 278
- incoherent aberrations 273
  - electron holography 273
- incoherent imaging 361 ff
  - atomic number ( $Z$ ) sensitivity 363
  - direct structure inversion 364
  - dynamic diffraction conditions 363
  - elastically scattered electrons 364
  - experimental set-up 364
  - high resolution 365
  - image intensity 367
  - residual correlations 368
  - resolution 365
  - resolution criteria 361
  - STEM geometry 364
  - theory 364
  - theory using elastically scattered electrons 366
  - using inelastically scattered electrons 370
  - with thermally scattered electrons 367

- incoherent structure imaging and EELS 377
- inelastic scattering 340
  - scanning transmission electron microscopy 340
- infinite-magnification circle 345
- information depth 245, 295, 395
  - Auger electron spectroscopy 394
  - low-energy electron microscopy 245
  - spin-polarized low-energy electron microscopy 295
- information limit 116, 117, 274
  - transmission electron microscopy 274
- InGaAs/InP superlattice 238
  - convergent beam electron diffraction 238
- inner-shell energy losses 341
  - scanning transmission electron microscopy 341
- InP 317
  - electron channeling pattern 317
- input lenses 407
  - cylindrical mirror analyzer 407
- interface displacements 240
  - applications
    - – convergent beam electron diffraction 240
- interference effects 267
  - Lorentz microscopy 267
- intrinsic luminescence 324
- intrinsic stacking fault 104
  - diffraction contrast transmission electron microscopy 104
- inversion boundaries 51
  - dislocation contrast transmission electron microscopy 51
- ion detection 462
  - secondary ion mass spectrometry 462
- ion implantation 461
  - secondary ion mass spectrometry 461
- ion microprobe mode 458
  - secondary ion mass spectrometry 458
- ion microprobes 468
  - comparison with ion microscopes (table) 468
- ion microscope mode 458
  - secondary ion mass spectrometry 458
- ion microscopes 468
  - comparison with ion microprobes (table) 468
- iron 214
  - dislocations studied by HVEM 214
- iron single crystals 493
  - scanning electron microscopy with polarization analysis 493
- Kikuchi cones 15
- Kikuchi lines 14 ff, 166
- Kikuchi pattern 64
- kinematic approximation 233
  - higher order Laue zone reflections 233
- kinematical diffraction theory 18 ff
  - Born approximation 18
  - scattered amplitude 22
- kinematical theory 61
  - dislocation contrast 61
- Kirchhoff diffraction integral 272
- knife-edge test 350
  - scanning transmission electron microscopy 350
- Kramers' equation 435
- LaB<sub>6</sub> emitters 411
  - scanning Auger microscopy 411
- LaBaCo<sub>2</sub>O<sub>5+d</sub> 188
  - electron energy loss spectroscopy 188
- large angle convergent beam electron diffraction 223, 224, 228
  - large-area phase contrast 275
    - transmission electron microscopy 275
  - (large-area) surface analysis 457
    - secondary ion mass spectrometry 457
- lateral resolution 441, 468, 469
  - ion microprobe 468
  - ion microscope 468
    - scanning transmission electron microscopy 441
    - secondary ion mass spectrometry 469
- lateral resolution limit 461
  - secondary ion mass spectrometry 461
- lattice fringes 29, 174
  - reflection electron microscopy 174
  - transmission electron microscopy 29
- lattice parameter measurement 226
  - higher order Laue zone 226
- Laue circles 166
- lock-in detection techniques 402
  - deflection type analyzers 402
- Lorentz imaging 269
  - in situ experiments 269
- Lorentz microscopy 261 ff, 266 ff
  - application
    - – Co/Cu multilayers 265
    - – cobalt film 268
    - – magnetic domains 263
  - control the magnetic field strength 269
  - image contrasting, theory 266
  - imaging modes 267
    - reflection mode 484
      - – comparison with SEMPA 484
- low energy diffuse scattering detector 489
  - scanning electron microscopy with polarization analysis 489
- low energy diffuse scattering spin polarization analyzer 489
  - scanning electron microscopy with polarization analysis 489
- low energy electron diffraction electron-spin polarization analyzer 490

- scanning electron microscopy with polarization analysis 490
- low-energy electron microscopy 243 ff, 257
  - applications
  - adsorption layers 251
  - chemisorption 254
  - copper on Mo(110) 255
  - CoSi<sub>2</sub> 256
  - Co<sub>2</sub>Si 256
  - Cu<sub>3</sub>Si 256, 256
  - gold on Mo(110) 253
  - heteroepitaxy on Si 254
  - Mo(110) surface 253
  - Mo<sub>2</sub>C 253
  - Mo<sub>2</sub>C formation on Mo(110) 253
  - order-disorder transitions 251
  - oxidation processes 253
  - Pb on Si(111) 250
  - Si(110) 256
  - Si(111) 254
  - Si(111) phase transitions 252
  - Si homoepitaxy 255
  - surface topography studies 250
  - thin film growth 254
  - W(100) 254
  - W(100) surface 254
  - comparison with other techniques 257
  - instrumentation 246 ff
  - limitations 257
  - surface topography 245
  - theory 243
- magnetic circular dichroism 485
  - comparison with SEMPA 485
- magnetic domain measurements 216
  - high voltage electron microscopy 216
- magnetic domain structure 297, 316
  - backscattered electron imaging 316
  - spin-polarized low-energy electron microscopy 297
- magnetic domain walls 283
  - electron holography 283
- magnetic domains 261 ff, 297
  - imaging by spin-polarized low-energy electron microscopy 297
  - Lorentz microscopy 261, 263
  - perturbation to electron diffraction patterns 262
  - transmission electron microscopy 262
- magnetic fields 418
  - in imaging XPS 418
- magnetic force microscopy 484
  - comparison with SEMPA 484
- magnetic imaging 484
  - comparison of different techniques 484
- magnetic layers, exchange coupling 494
  - scanning electron microscopy with polarization analysis 494
- magnetic media 312, 313
  - secondary electron imaging 312, 313
- magnetic microstructure 483
  - scanning electron microscopy with polarization analysis 483
- magnetic objective lens 420
  - X-ray photoelectron imaging 420
- magnetic phase transitions 300
  - spin-polarized low-energy electron microscopy 300
- magnetic sector mass spectrometer 462
- magnetic singularities 495
  - scanning electron microscopy with polarization analysis 495
- magnetic structures 283
  - electron holography 283
- magnetization distribution 267
  - determination by Lorentz microscopy 267
- magneto-optic Kerr effect 485
  - comparison with SEMPA 485
- magneto-optic recording media 494
  - scanning electron microscopy with polarization analysis 494
- mapping of fine structures 198
  - electron energy loss spectroscopy 198
- martensitic transformations 215
  - high voltage electron microscopy 215
- mass determinations of macromolecules 354
  - scanning transmission electron microscopy 353
- mass resolution 470, 471
  - ion microprobe 469
  - ion microscope 469
  - secondary ion mass spectrometry 471
- mass spectrometers 462, 463
  - characteristics of for use in secondary ion mass spectrometry (table) 463
  - secondary ion mass spectrometry 462
- matrix contrast 472
  - secondary ion mass spectrometry 472
- maximum likelihood common factor analysis 395 f
  - Auger electron spectroscopy 396
- metastable alloys 234
  - applications
  - convergent beam electron diffraction 234
- MgO 177, 331, 344, 345
  - electron intensity distributions 177
  - in-line holography 345
  - nanodiffraction patterns 344
  - scanning transmission electron microscopy 331
- micro-analysis 341
  - scanning transmission electron microscopy 341

- micromagnetic imaging 485
  - scanning electron microscopy with polarization analysis 485
- microtextures 46, 50, 51
  - imaging of, by interface contrast 51
  - imaging of, by domain contrast 50
  - of domain fragmented crystals 46
- MIDAS system 352
- mixed dislocations 71, 72
  - diffraction contrast transmission electron microscopy 71
- mixed-layer compounds 143
  - high-resolution electron microscopy 143
- Mo(110) 245
  - low-energy electron microscopy 245
- Mo(110) surface 253
  - low-energy electron microscopy 253
- Mo<sub>2</sub>C 253
  - low-energy electron microscopy 253
- Mo<sub>2</sub>C formation on Mo(110) 253
  - low-energy electron microscopy 253
- modulation transfer function 110
- moiré patterns 82ff
  - in the determination of lattice parameter 84
  - in the study of dislocations 84
  - theoretical considerations 82
- molybdenum–sulphide fullerene 442
  - high-resolution image 442
  - X-ray energy-dispersive spectra 442
- Moseley's law 432
- Mott analyzer 490
  - scanning electron microscopy with polarization analysis 490
- MTS<sub>3</sub>(SnTaS<sub>3</sub>) 143
  - high-resolution electron microscopy 143
- multielement characterization 459
  - secondary ion mass spectrometry 459
- multilayer structure 474
  - secondary ion mass spectrometry 474
- multiple detector systems 428
  - analytical electron microscopy 428
- multiple least squares fitting techniques 194
  - electron energy loss spectroscopy 194
- multiribbons, dislocations in 88f
- multislice calculation 340
  - scanning transmission electron microscopy 340
- multislice method 152
  - image simulation 152
- multispectral Auger microscope 421
  - reduction of image artefacts 421
- multivariate data analysis 476
  - secondary ion mass spectrometry 476
- Na<sub>8</sub>Si<sub>46</sub> 146
  - high-resolution electron microscopy 146
- nanodiffraction 343
  - scanning transmission electron microscopy 343ff
- Na<sub>x</sub>Si<sub>136</sub> 146
  - high-resolution electron microscopy 146
- Nb mapping 196
  - electron energy loss spectroscopy 196
- Nb<sub>3</sub>Ga<sub>13</sub> 139
  - high-resolution electron microscopy 139
- Nb<sub>5</sub>Ga<sub>13</sub> 139, 140
  - high-resolution electron microscopy 139, 140
- NbSe<sub>3</sub> 225
  - convergent beam electron diffraction 225
- NbTe<sub>2</sub> 89f, 92
  - dislocation in 89f
  - glide dislocations 92
  - multiribbons in 90
- near-ultraviolet electron emission microscopy 249
- Ni-50 at % Ti alloy 218
  - electron irradiation induced crystalline–amorphous transition 218
- Ni-50 at % Ti shape-memory alloy 218
  - electron irradiation induced crystalline–amorphous transition 218
- Ni<sup>3+</sup>Mo 99
  - discommensurations 99
- Ni<sub>3</sub>Mo 50, 100
  - composite diffraction pattern 50
  - diffraction contrast image 100
- Ni<sub>4</sub>Mo 48, 52
  - diffraction pattern 48
  - imaging of interfaces in 52
- nickel oxide 100
  - domain walls 100
- NiO 98
  - antiferromagnetic domain structure 98
- niobium 92, 93, 284
  - electron holography 284
  - tilt boundaries 92, 93
- non-equilibrium phases 217
  - formation of high voltage electron microscopy 217
- numerical reconstruction 282
  - off-axis electron holography 282
- off-axis electron holography 279
  - beam splitting 279
- off-axis Fresnel holography 279
- off-axis holograms 277
- off-axis holography 133
  - phase retrieval 133
- off-axis image plane holography 280, 281
  - image reconstruction 281
  - image recording 281

- instrumentation 281
- theory 280
- off-axis STEM holography 348
  - instrumentation 348
- omega filter 183
- order-disorder transitions 251
  - low-energy electron microscopy 251
- organic coatings 479
  - secondary ion mass spectrometry 479
- oxidation processes 253
  - low-energy electron microscopy 253
  
- p-n junction 321, 322
  - electron beam induced channeling image 322
  - electron beam induced conductivity 321
- Packwood-Brown model 451
  - quantitative X-ray scanning microanalysis 451
- parallel detection 406
  - implementation in XPS and AES 406
- partial dislocations 73
  - diffraction contrast transmission electron microscopy 73
- particles 453
  - analytical electron microscopy 453
- passivation layers 476
  - secondary ion mass spectrometry 476
- Pb on Si(111) 250
  - low-energy electron microscopy 252
- Pb<sub>3</sub>(VO<sub>4</sub>)<sub>2</sub> 97, 98
  - domain structures 97, 98
- Pd particles on a carbon support 356
  - bright-field STEM image 356
- perfect crystals 28, 32, 35
  - anomalous absorption 32
  - dynamical equations 28
  - matrix formulation for the amplitudes of transmitted and scattered beams 35
- performance 308
  - scanning electron microscopy 308 ff
- phase contrast imaging 116
  - electron microscopy 116
- phase contrast microscopy 115
  - high-resolution electron microscopy 115
- phase contrast transfer function 274
  - transmission electron microscopy 274
- phase retrieval 133, 160
  - focus variation method 133
  - off-axis holography 133
- phase retrieval procedure 133
  - high-resolution electron microscopy 133
- phase transformations 46
  - diffraction contrast transmission electron microscopy 46
- phase transitions 103, 252
  - diffraction contrast transmission electron microscopy 103
  - low-energy electron microscopy imaging 252
- pinhole camera 113
- planar interfaces 39, 95
  - applications
    - diffraction contrast electron microscopy 95
  - fringe profile 39
  - plastic deformation 90
  - observation by high voltage electron microscopy 90
- platinum 92, 93
  - twist boundaries 92, 93
- platinum on zeolite 334
  - annular dark-field imaging 334
- point analysis 184
  - electron energy loss spectroscopy 184
- point defect clusters 93
  - applications
    - diffraction contrast electron microscopy 93
- point resolution 116, 117, 119
  - electron microscope 119
- point spread function 272
- polarization analyzers 488 f
  - scanning electron microscopy with polarization analysis 488
- polished surfaces 435
  - electron probe X-ray microanalysis 435
- primary ion sources 461
  - secondary ion mass spectrometry 461
- probe channeling 372 ff
  - Z-contrast scanning transmission electron microscopy 372
- profile analysis 447
  - energy-dispersive spectrometry 447
- projected charge density approximation 127
- projection box 110 f
- Pt (111), surface steps 164, 172, 173
  - reflection electron microscopy 164, 172, 173
- ptychology 346
  
- quadrupole mass spectrometer 462
- quantitative image interpretation, high-resolution electron microscopy 131 ff
- quantitative X-ray scanning microanalysis 448
- quantum size effect 294
  - ultrathin magnetic layers 294
- quantum size effect contrast 245
  - low-energy electron microscopy 245
- quartz 47, 96, 97
  - $\alpha$ - $\beta$ -transition 47, 96
  - domain fragmented a phase 96, 97
- $\alpha$ -quartz 46
  - Dauphiné twins 47
  - domain structure 46

- $\beta$ -quartz 46
  - domain structure 46
- quenched aluminum 94
  - unfaulted dislocation loops 94
- radiation damage 103, 207
  - diffraction contrast transmission electron microscopy 103
  - high voltage electron microscopy 206
- radiation effects 216
  - high voltage electron microscopy 216
- Rb<sub>6</sub>C<sub>60</sub> 146
  - high-resolution electron microscopy 146
- reduced aberration 405
- reflection electron microscopy 163 ff, 176, 178
  - applications 178
  - domain boundaries 178
  - epitaxial growth 178
  - surface steps 178
  - theory 176 ff
- reflection high energy electron diffraction 165 ff, 176
  - formation of Kikuchi lines 168
  - theory 176 ff
- reflection high energy electron microscopy 167
  - Kikuchi lines 167
- reflection high-energy electron diffraction patterns 353
  - scanning transmission electron microscopy 353
- refractive bending 169
  - reflection high energy electron diffraction 169
- resistive anode encoder 463, 470
  - experimental set-up 463
  - secondary ion mass spectrometry 463, 470
- resolution 110, 116, 119, 308, 333, 339, 349, 410, 488
  - high-resolution electron microscopy 110, 116, 119
  - improvement by holography 349
  - improvement of in the electron microscope 119
  - scanning Auger microscopy 410
  - scanning electron microscopy 308 ff
  - with polarization analysis 488
  - scanning transmission electron microscopy 333, 339 349
- resolution limit 117, 118
  - coherent imaging 119
  - high-resolution electron microscopy 117, 118
- resolution number 116
- retarding field 407
  - concentric hemispherical analyzer 407
  - cylindrical mirror analyzer 407
- retarding field analyzers 398
- rhodium catalyst on g-Al<sub>2</sub>O<sub>3</sub> 384
  - Z-contrast images 384
- RuSe<sub>2</sub> 66
  - weak-beam image 66
- rutile (100) 170
  - reflection electron microscopy 170
- sample damage 412
  - scanning Auger microscopy 412
- sample preparation 391
  - Auger electron spectroscopy 391
- scanning Auger electron microscopy 342
- scanning Auger microscopy 387 ff, 396 ff, 420
  - background slope effects 422
  - beam current fluctuation effects 423
  - edge effects 423
  - image artefacts 420
  - imaging technique 396
  - substrate backscattering effects 422
  - topographic effects 422
- scanning beam methods 303 ff
- scanning electron beam analysis 435
  - instrumentation 435 ff
- scanning electron microscope 305, 306
  - operating principle 306
- scanning electron microscopy 305 ff, 312, 483 ff
  - capabilities 305
  - comparison with scanning transmission microscopy 312 f
  - instrumentation 306, 307
  - operating principle 305
  - with polarization analysis 483 ff, 487, 488, 491, 493 ff
  - applications
    - Fe/Cr/Fe(100) 494
    - Fe–SiO<sub>2</sub> films 495
    - iron single crystals 493
    - magnetic layers, exchange coupling 494
    - magnetic singularities 495
    - magneto-optic recording media 494
  - comparison with other magnetic imaging techniques 484
  - instrumentation 487
  - operating principle 485
  - resolution 488
  - system performance 491
  - theory 485 f
- scanning electron probe methods 436
  - basic characteristics (table) 436 f
- scanning microanalysis 427 ff
  - resolution 430
- scanning transmission electron microscopy 329 ff, 336, 343, 350, 353, 361, 484
  - applications
    - carbon nanotubes 357
    - diamond, planar defects in 356

- heavy atoms on light atom supports 353
- heavy atoms on zeolites 355
- mass determinations of macromolecules 354
- small particles 354
- surface superstructures 357
- bright field imaging mode 332
- comparison with conventional transmission electron microscopy 331
- comparison with SEMPA 484
- 3 contrast ((?)) 361 ff
- convergent-beam 343
- differential imaging mode 336
- high-angle annular dark-field imaging 333 f
- instrumentation 329 ff
- instrumentation 350
- off-axis holography 347
- resolution 333, 339
- serial imaging mode 332
- thin-annular bright-field imaging 339
- scattering factor 117
- scattering theory 9 ff
- Scherzer defocus 116
- Scherzer point-resolution 274
- transmission electron microscopy 274
- Schottky barrier 323
- depletion depth 323
- electron beam induced channeling image 323
- Schottky field emitters 411
- scanning Auger microscopy 411
- Schottky surface barrier 321
- electron beam induced conductivity 321
- screw dislocations 24, 25, 54, 61, 71
- image profiles 24, 71
- imaging of 55, 56, 69
- kinematical theory 61
- transmission electron microscopy 55
- secondary electron detector 410
- scanning Auger microscopy 410
- secondary electron emission microscopy 247
- instrumentation 246
- secondary electron imaging 163, 310, 311, 313
- Everhart-Thornley detector 311
- voltage contrast 313
- secondary electrons 310
- generation 310
- yield 310
- secondary ion formation 460
- secondary ion mass spectrometry 457, 458, 461, 462, 478
- applications 478 f
- dynamic mode 458
- experimental set-up 458
- image recording 462
- instrumentation 461
- ion detection 462
- ion microprobe mode 458
- ion microscope mode 458
- sample requirements 478
- sensitivity 459
- static mode 458
- selective beam imaging 129
- high-resolution electron microscopy 129
- semiconductor interfaces 377
- Z-contrast scanning transmission electron microscopy 377
- semiconductors 19, 100, 238, 376
- applications
- convergent beam electron diffraction 238
- fabrication-induced defects 100
- scanning electron microscopy 319
- transmission electron microscopy 100
- Z-contrast scanning transmission electron microscopy 376
- sensitivity 459
- secondary ion mass spectrometry 459
- sequential energy analysis 397
- scanning Auger microscopy 397
- Sherman function 489
- scanning electron microscopy with polarization analysis 489
- Si 237, 377
- convergent beam electron diffraction 237
- tilt grain boundaries in 377
- Si homoeopitaxy 255
- low-energy electron microscopy 255
- Si (110) 256
- low-energy electron microscopy 256
- Si (111) 165, 170, 174, 175, 254
- lattice fringes 175
- low-energy electron microscopy 254
- reflection electron microscopy 165, 170, 174, 175
- Si/SiGe 239
- applications
- convergent beam electron diffraction 239
- convergent beam electron diffraction 239
- layer strains 239
- Si-SiO<sub>2</sub> interfaces 200
- electron energy loss spectroscopy 200
- Si-TiO<sub>2</sub> interface 199
- electron energy loss spectroscopy 199
- signal collection 325
- cathodoluminescence 325
- signal processing 490 f
- scanning electron microscopy with polarization analysis 490
- signal-to-noise ratio 119, 492
- high-resolution electron microscopy 119
- scanning electron microscopy with polarization analysis 492

- silicon 17, 94 ff, 142, 217, 228, 229, 255, 285
  - convergent beam electron diffraction 229
  - dissociated dislocation 142
  - electron holography 285
  - extended dislocations 96
  - faulted Frank loops 94, 95
  - faulted loop 95
  - high voltage electron microscopy 217
  - homoepitaxy 255
  - study of by LEEM 255
  - Kikuchi bands 17
  - large angle convergent beam electron diffraction pattern 228
  - stacking fault tetrahedron 142
- silicon nanocrystals 383
  - Z-contrast images 383
- SiO<sub>2</sub> gate oxide layers 411
  - scanning Auger microscopy 411
- small particles 354
  - scanning transmission electron microscopy 354
- snorkel lens 307
  - scanning electron microscopy 307
- SnS<sub>2</sub> 70, 87
  - dislocation ribbon in 87
  - stacking fault energy 87
- soft X-ray spectra 434
  - scanning microanalysis 434
- source defined imaging 398, 410, 413
  - Auger electron spectroscopy 410
  - scanning Auger microscopy 398
  - X-ray photoelectron spectroscopy 413
- spatial displacement 397
  - scanning Auger microscopy 397
- spatial incoherence 117
- spatial resolution 308, 316, 322, 326, 442, 486
  - backscattered electron imaging 316
  - cathodoluminescence 326
  - electron beam induced channeling image 322
  - scanning electron microscopy 308 ff
  - with polarization analysis 486
  - X-ray microanalysis 442
- specimen preparation 123
  - high-resolution electron microscopy 123 ff
- spectrum imaging 184, 186
  - electron energy loss spectroscopy 184, 186
- spherical inclusion 60
  - two-beam image characteristics 60
- spin dependent scattering 486
  - scanning electron microscopy with polarization analysis 486
- spin-polarized low-energy electron diffraction 293
- spin-polarized low-energy electron microscopy 293 ff, 485
  - applications
    - in situ studies 298
    - magnetic domain structure 297
    - magnetic phase transitions 300
    - superlattices 297
    - thin films 297
  - comparison with SEMPA 485
  - information depth 295
  - instrumentation 295, 296
  - resolution 297
  - theory 293
- sputter yield 460
  - secondary ion mass spectrometry 460
- sputtering process 460
  - secondary ion mass spectrometry 460
- SrTiO<sub>3</sub> 379, 380
  - tilt boundaries in 379
  - [001] tilt grain boundaries 379
  - tilt grain boundaries in 380
- stacking fault contrast 22
- stacking fault energy, measurement of 85
- stacking fault tetrahedron 140
- stacking faults 21, 22, 42, 43, 95, 105, 140, 172 f
  - determination of in fcc structures 105
  - displacement function 21
  - extrinsic faults 96
  - intrinsic faults 95
  - reflection electron microscopy 172 ff
  - study by high-resolution electron microscopy 140
- stainless steel 90
  - glide dislocations 90
- standardless analysis 452
  - quantitative X-ray scanning microanalysis 452
- static secondary ion mass spectrometry 458
- strain measurement 226
  - higher order Laue zone 226
- space group determination 227
  - convergent beam electron diffraction 227
- structural resolution 116
- structure retrieval 136
  - high-resolution electron microscopy 136
- subgrain boundaries 91, 92
  - applications
    - diffraction contrast electron microscopy 91, 92
- substitutional alloys 138
  - high-resolution electron microscopy 138
- substrate backscattering effects 422
  - scanning Auger microscopy 422
- superconductors 283
  - electron holography 283
- superlattices 174, 297
  - reflection electron microscopy 174
  - spin-polarized low-energy electron microscopy 297

- supported metal catalysts 344
  - scanning transmission electron microscopy 344
- surface analysis 284
  - electron holography 284
- surface lattice fringes 176
  - reflection electron microscopy 176
- surface layers 174
  - reflection electron microscopy 174
- surface projections 169
  - reflection electron microscopy 169
- surface steps 169, 178
  - reflection electron microscopy 169, 178
- surface superstructures 357
  - scanning transmission electron microscopy 357
- surface topography 245
  - low-energy electron microscopy 245
- surface topography studies 250
  - low-energy electron microscopy 250
  
- temporal incoherence 117
- thermal vibrations 367
  - incoherent imaging 367
- thermionic electron emission microscopy 249
- thin-annular bright-field imaging 339
- thin-film criterion 453
- thin film growth 254
  - low-energy electron microscopy 254
- thin films 283, 297, 348, 453
  - analytical electron microscopy 453
  - electron holography 283
  - magnetic fields in 348
  - spin-polarized low-energy electron microscopy 297
- thin phase objects 125
  - electron scattering 125
- through-focus series 125
- Ti<sub>2</sub>Nb<sub>10</sub>O<sub>29</sub> 126, 130
  - image simulation 126, 130
- time of flight analyzers 400
- time-of-flight (TOF) mass spectrometer 462
- time-resolved spectroscopy 188
  - electron energy loss spectroscopy 188
- TiO<sub>2</sub> 135
  - exit wave entropy 135
- topographic effects 422
  - scanning Auger microscopy 422
- transfer function 110 ff, 115, 118
  - electron microscope 118
- transfer lens system for X-ray photoelectron imaging 417
  - instrumentation 417
- transfer lenses 407 f
  - cylindrical mirror analyzer 407
  - retarding effect 408
- translation interfaces 36, 40, 45
  - diffraction effects 36
  - fringe profile 40, 45
  - transmission cross-coefficient 112
  - transmission electron microscope 363
    - coherent and incoherent imaging optics 363
    - transmission electron microscopy 362
      - bright-field coherent imaging 363
      - conventional 271 ff
      - diffraction contrast 3 ff
      - incoherent dark field imaging 362
      - instrumentation 3
    - transmission microscopes 3 ff
    - transport optics 488
      - scanning electron microscopy with polarization analysis 488
  - TRIFT mass spectrometer 468
  - tungsten 214
    - dislocations studied by HVEM 214
  - twins 46, 48, 49
    - diffraction patterns 48, 49
  - two-beam dynamical diffraction theory 227
  - two-beam dynamical theory 26 ff, 68
  - two-beam lattice fringes 29
    - transmission electron microscopy 29
  
  - ultrathin magnetic layers 294
    - spin-polarized low-energy electron microscopy 294
  - universal escape depth 394
    - Auger electron spectroscopy 394
  - University of Chicago UC-HRL scanning ion microprobe 468
  - uranium 330
    - annular dark-field imaging 330
  
  - V<sub>2</sub>O<sub>5</sub> 393, 399
    - Auger spectrum 399
    - X-ray photoelectron spectroscopy 393
  - vacuum matrix 38
    - matrix formulation 38
  - van Cittert–Zernike theorem 277
    - electron holography 277
  - vector magnetization maps 486, 493
    - scanning electron microscopy with polarization analysis 486, 493
  - video camera based acquisition 470
    - secondary ion mass spectrometry 470
  - voltage contrast 313
    - secondary electron imaging 313
  
  - W(100) 254
    - low-energy electron microscopy 254
  - W(100) surface 254
    - low-energy electron microscopy 254
  - Warren approximation 369

- Z-contrast scanning transmission electron microscopy 369
- wave reconstruction 134
- high-resolution electron microscopy 134
- wavelength dispersive spectroscopy
- comparison with energy-dispersive spectroscopy 445
- wavelength-dispersive spectrometry 445
- scanning emission electron microscopy 445
- wavelength-dispersive spectroscopy 432
- $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$  432
- weak phase object 115
- optimum focus imaging 115
- weak-beam method 63 ff
- transmission electron microscopy 63
- wedge fringes 22
- fringe pattern 22
- window function 114
  
- X-ray absorption correction 450
- in quantitative X-ray scanning microanalysis 449
- X-ray analysis 439
- X-ray detector 352
- scanning transmission electron microscopy 352
- X-ray distribution function 429, 452
- analytical electron microscopy 429, 452
- X-ray distributions 451
- integration of for quantitative X-ray scanning microanalysis 451
- X-ray emission spectra 430 f
- X-ray emission spectrometry 445
- X-ray mapping 447, 448
- ceramics 448
- energy-dispersive spectrometry 447
- wavelength-dispersive spectrometry 447
- X-ray microanalysis 441
- in scanning transmission electron microscopy 441
- X-ray photoelectron spectroscopy 387 ff, 391, 392
- X-ray photoelectron spectroscopy 413
- beam effects 393
- detector defined imaging 414
- quantitative analysis 393
- kinetic energy 392
- source defined imaging 413
- X-ray scanning microanalysis 448
- quantitative 448
- X-ray spectra 432
- scanning microanalysis 432
- X-rays 412
- generation 412
- monochromatizing 412
  
- $\text{YBa}_2\text{Cu}_3\text{O}_4$  146
- high-resolution electron microscopy 146
- $\text{YBa}_2\text{Cu}_3\text{O}_{7-d}$  99, 144
- high-resolution electron microscopy 144
- orthorhombic twins 99
- $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$  383, 430, 431
- energy-dispersive spectroscopy 432
- grain boundaries 381
- - Z-contrast scanning transmission electron microscopy 381
- wavelength-dispersive spectroscopy 432
- X-ray emission spectra 430, 431
- Z-contrast image of 383
- $\text{YBa}_2\text{Cu}_3\text{O}_7$  98
- disorder-order transition 98
- $\text{YBa}_2\text{Cu}_4\text{O}_8$  134
- exit wave function 134 f
  
- Z-contrast scanning transmission electron microscopy 361 ff
- applications
- - catalysts 384
- - CdTe(001)-GaAs(001) interface 377, 378
- - ceramics 379
- - GaAs 376
- - nanocrystalline materials 382
- - semiconductor interfaces 377
- - semiconductors 376
- - tilt boundaries in silicon 377
- -  $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$  grain boundaries 381
- ZAF method 449
- quantitative X-ray scanning microanalysis 449
- Zernike-van Cittert theorem 337
- scanning transmission electron microscopy 337
- zero order Laue zone 224
- convergent beam electron diffraction 224
- zinc 92, 94
- undissociated dislocations 92, 94
- zone axis diffraction 229
- Bloch wave theory 229