

# Handbook of Microscopy

Applications in Materials Science,  
Solid-State Physics and Chemistry

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

Methods II



**VCH** 

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

# **Handbook of Microscopy**

## **Methods II**



# **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. Magonor, 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

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

#### **Distribution:**

VCH, P.O. Box 10 11 61, D-69451 Weinheim (Federal Republic of Germany)

Switzerland: VCH, P.O. Box, CH-4020 Basel (Switzerland)

United Kingdom and Ireland: VCH (UK) Ltd., 8 Wellington Court, Cambridge CB1 1HZ (England)

USA and Canada: VCH, 333 7th Avenue, New York, NY 10001 (USA)

Japan: VCH, Eikow Building, 10-9 Hongo 1-chome, Bunkyo-ku, Tokyo 113 (Japan)

ISBN 3-527-29473-2

# **Handbook of Microscopy**

Applications in Materials Science,  
Solid-State Physics and Chemistry

Edited by

S. Amelinckx, D. van Dyck,

J. van Landuyt, G. van Tendeloo

Methods II



Weinheim · New York · Basel · Cambridge · Tokyo



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. Ute Anton, Dr. Jörn Ritterbusch  
Production Manager: Dipl.-Wirt.-Ing. (FH) 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  
**Handbook of microscopy** : applications in materials science,  
solid state physics and chemistry / ed. by S. Amelinckx ... -  
Weinheim ; New York ; Basel ; Cambridge ; Tokyo : VCH.  
NE: Amelinckx, Severin [Hrsg.]  
Methods 2 (1997)  
ISBN 3-527-29473-2

© 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

## Short biography of the editors



*Severin Amelinckx* was born in Willebroek, Belgium, in 1922. He studied at the University of Ghent, receiving his first degree (licence) in mathematics in 1944, his doctorate in physics in 1952, and his aggregation in physics in 1955. Currently he is Emeritus Professor of General Physics and Crystallography associated with the EMAT laboratory of the University of Antwerp (RUCA). Until 1987 he was Director General of the Belgian Nuclear Research Establishment at Mol. He is a member of the European Academy and of the Koninklijke Academie voor Wetenschappen, Letteren en Schone Kunsten van België and former chairman of the division of sciences of this academy.

His research interests include electron diffraction contrast imaging, defects in solids, phase transformations and their resulting domain structures, crystal growth, dislocations, fullerenes and nanotubes, the structure of high- $T_c$  superconductors, modulated structures, and order-disorder in alloys.



*Joseph Van Landuyt*, who was born in St. Amandsberg, Belgium, in 1938, obtained both his licence (1960) and doctorate in physics (1965) from the University of Ghent. At present he is Professor of General Physics and Crystallography at the University of Antwerp (RUCA and UIA) and of Electron Microscopy at UIA and the Flemish University of Brussels (VUB). He is a member of the Koninklijke Academic voor Wetenschappen, Letteren en Schone Kunsten van België.

His research interests are centered on the study of nanostructural features in alloys, ceramics, and minerals (in particular gems), with special interest in defects in semiconductors and their relation to device performance. More general subjects of interest are structural variants, defects, and phase transitions in various solids.



*Gustaaf Van Tendeloo*, born in Lier, Belgium, in 1950, received his licence in physics from the University of Brussels (VUB) in 1972, his doctorate from the University of Antwerp (UIA) in 1974, and his aggregation from the University of Brussels (VUB) in 1981. He has been associated with the University of Antwerp (RUCA) since 1972, but has spent extended periods of time as a researcher in Berkeley (USA), Caen (France), and elsewhere. He is currently Professor of Solid-State Physics at the University of Brussels (VUB) and of the Physics of Materials at the University of Antwerp (RUCA and UIA).

His research interests include the electron microscopy of inorganic solids (in particular high- $T_c$  superconductors), solid-state phase transitions, modulated structures, fullerenes, defects in crystals order-disorder in alloys, and nanostructural features in general.



*Dirk Van Dyck* was born in Wilrijk, Belgium, in 1948. He studied physics, receiving his licence from the University of Brussels (VUB) in 1971 before moving to the University of Antwerp (UIA) for his doctorate (1977) and aggregation (1987). He has been associated with the University of Antwerp since 1971, and is at present Professor of Theoretical Mechanics, Digital Techniques and Image Processing at the University of Antwerp.

Among his research interests are theoretical aspects of dynamic electron diffraction and imaging, holographic reconstruction and structural retrieval, image processing and pattern recognition, and artificial intelligence. In particular, he is involved in the development of a 1 Å resolution microscope in the framework of the Brite/Euram program of the European Union.

The four editors belong to the Electron Microscopy for Materials Science (EMAT) laboratory, University of Antwerp (RUCA), which was founded in 1965. All four have contributed significantly to the development of electron microscopy and its application by numerous publications in journals and books and are members of editorial boards of several international journals in the field of materials science. They have also delivered numerous invited lectures at international conferences in their respective areas of research.

# List of Contributors

Bauer, Ernst (V:3)  
Physikal. Institut  
Technische Universität Clausthal  
38678 Clausthal  
Germany

Bonnet, Noël (VIII:2)  
INSERM Unit 314  
University of Reims  
21, rue Clément Ader  
51100 Reims  
France

Cerezo, Alfred; Smith, George D. W.  
(VI:2)  
Department of Materials  
University of Oxford  
Parks Road  
Oxford OX1 3PH  
United Kingdom

Cory, David G.; Choi, Sungmin (V:1)  
Dept. of Nuclear Engineering,  
NW 14-4111  
Massachusetts Institute of Technology  
Cambridge, MA 02139  
USA

Cowley, John M. (IV:2.2)  
Arizona State University  
Dept. of Physics & Astronomy  
Box 87 15 04  
Tempe, AZ 8528 -1504  
USA

DiNardo, N. John (VII:4)  
Department of Physics and  
Atmospheric Science  
Drexel University  
Philadelphia, PA 19104  
USA

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

Herrmann, Karl Heinz (VIII:1)  
Institute of Applied Physics  
University of Tübingen  
Auf der Morgenstelle 10  
72076 Tübingen  
Germany

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

Kruit, Peter (IX:1)  
Delft University of Technology  
Dept. of Applied Physics  
Lorentzweg 1  
2628 CS Delft  
The Netherlands

Mundschau, Michael V. (VI:1)  
Bowling Green State University  
Dept. of Chemistry  
Bowling Green, Ohio 43403-0213  
USA

Oleshko\*, Vladimir; Gijbels, Renaat  
(IV: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. (IV:2.3)  
Solid State Division  
Oak Ridge National Laboratory  
P. O. Box 2008  
Oak Ridge, TN 37831-6030  
USA

Scheinfein, Mike R. (V:2)  
Dept. of Physics and Astronomy  
Arizona State University  
Tempe, AZ 85287-1504  
USA

Unguris, John; Kelley, Michael H.;  
Gavrin, Andrew; Celotta, Robert J.;  
Pierce, Daniel T. (V:2)  
Electron Physics Group  
Nat. Institute of Standards and  
Technology  
Gaithersburg, MD 20899  
USA

Schwarz, Udo D. (VII:2)  
Institute of Applied Physics  
University of Hamburg  
Jungiusstr. 11  
20355 Hamburg

Spence, John C. H. (IX:2)  
Dept. of Physics  
A. S. U.  
Tempe, AZ 85287  
USA

Ventrice, Carl A., Jr. (VII:4)  
Department of Physics  
Rensselaer Polytechnic Institute  
Troy, NY 12180  
USA

Wadas, Andrzej (VII:3)  
Institute of Applied Physics  
University of Hamburg  
Jungiusstr. 11  
20355 Hamburg  
Germany

Wiesendanger, Roland (VII:1)  
University of Hamburg  
Institute of Applied Physics and  
Microstructure Research Center  
Jungiusstr. 11  
20355 Hamburg  
Germany

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

# Outline

## Volume 1: Methods I

### **I Light Microscopy**

- 1 Fundamentals of Light Microscopy**  
*F. Mücklich*
- 2 Optical Contrasting of Microstructures**  
*F. Mücklich*
- 3 Raman Microscopy**  
*P. Dhamelincourt, J. Barbillat*
- 4 Three-Dimensional Light Microscopy**  
*E. H. K. Stelzer*
- 5 Near Field Optical Microscopy**  
*D. Courjon, M. Spajer*
- 6 Infrared Microscopy**  
*J. P. Huvenne, B. Sombret*

### **II X-Ray Microscopy**

- 1 Soft X-Ray Imaging**  
*G. Schmahl*
- 2 X-Ray Microradiography**  
*D. Mouze*
- 3 X-Ray Microtomography**  
*J. Cazaux*
- 4 Soft X-Ray Microscopy by Holography**  
*D. Joyeux*
- 5 X-Ray Diffraction Topography**  
*M. Schlenker, J. Baruchel*

### **III Acoustic Microscopy**

- 1 Acoustic Microscopy**  
*A. Briggs*

**IV Electron Microscopy****1 Stationary Beam Methods**

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

**Volume 2: Methods II****IV Electron Microscopy****2 Scanning Beam Methods**

- 2.1 Scanning Reflection Electron Microscopy  
*D. C. Joy*
- 2.2 Scanning Transmission Electron Microscopy  
*J. M. Cowley*
- 2.3 Scanning Transmission Electron Microscopy: Z Contrast  
*S. J. Pennycook*
- 2.4 Scanning Auger Microscopy (SAM) and Imaging X-Ray Photoelectron Microscopy (XPS)  
*R. De Gryse, L. Fiermans*
- 2.5 Scanning Microanalysis  
*R. Gijbels*
- 2.6 Imaging Secondary Ion Mass Spectrometry  
*P. van Espen, G. Janssens*

## **V Magnetic Methods**

- 1 Nuclear Magnetic Resonance**  
*D. G. Cory, S. Choi*
- 2 Scanning Electron Microscopy with Polarization Analysis (SEMPA)**  
*J. Unguris, M. H. Kelley, A. Gavrin, R. J. Celotta,  
D. T. Pierce, M. R. Scheinfein*
- 3 Spin-Polarized Low-Energy Electron Microscopy**  
*E. Bauer*

## **VI Emission Methods**

- 1 Photoelectron Emission Microscopy**  
*M. Mundschau*
- 2 Field Emission and Field Ion Microscopy (Including Atom Probe FIM)**  
*A. Cerezo, G. D. W. Smith*

## **VII Scanning Point Probe Techniques**

- General Introduction
- 1 Scanning Tunneling Microscopy**  
*R. Wiesendanger*
- 2 Scanning Force Microscopy**  
*U. D. Schwarz*
- 3 Magnetic Force Microscopy**  
*A. Wadas*
- 4 Ballistic Electron Emission Microscopy**  
*J. DiNardo*

## **VIII Image Recording, Handling and Processing**

- 1 Image Recording in Microscopy**  
*K.-H. Herrmann*
- 2 Image Processing**  
*N. Bonnet*

## **IX Special Topics**

- 1 Coincidence Microscopy**  
*P. Kruit*



- 2 Low Energy Electron Holography and Point-Projection Microscopy**  
*J. C. H. Spence*

## **Volume 3: Applications**

### **I Classes of Materials**

- 1 Metals and Alloys**  
*J. Th. M. De Hosson*  
*G. van Tendeloo*
- 2 Microscopy of Rocks and Minerals**  
*D. J. Barber*
- 3 Semiconductors and Semiconducting Devices**  
*H. Oppolzer*
- 4 Optoelectronic Materials**  
*I. Berbezier, J. Derrien*
- 5 Domain Structures in Ferroic Materials**  
*E. K. H. Salje*
- 6 Structural Ceramics**  
*M. Rühle*
- 7 Microscopy of Gemmological Materials**  
*J. van Landuyt, M. H. G. van Bockstael, J. van Royen*
- 8 Superconducting Ceramics**  
*G. van Tendeloo*
- 9 Non-Periodic Structures**
- 9.1 High-Resolution Imaging of Amorphous Materials  
*P. H. Gaskell*
- 9.2 Quasi-Crystalline Structures  
*K. H. Kuo*
- 10 Medical and Dental Materials**  
*K. Yasuda, K. Hisatsune, H. Takahashi, K.-I. Udoh, Y. Tanaka*
- 11 Carbon**  
*D. Bernaerts and S. Amelinckx*
- 12 Composite Structural Materials**  
*O. Van der Biest, P. Lust, K. Lambrinou, J. Ivens, I. Verpoest, L. Froyen*
- 13 The Structure of Polymers and Their Monomeric Analogs**  
*I. G. Voigt-Martin*

- 14      **Nuclear Materials**  
         *H. Blank, Hj. Matzke, H. Maußner, I. L. F. Ray*
- 15      **Magnetic Microscopy**  
         *A. Hubert*

## **II      Special Topics**

- 1      **Small Particles**  
         **(Catalysis, Photography, Magnetic Recording)**  
         *H. W. Zandbergen, C. Træholt*
- 2      **Structural Phase Transformations**  
         *H. Warlimont*
- 3      **Preparation Techniques**  
         **for Transmission Electron Microscopy**  
         *A. Barna, G. Radnóczy, B. Pécz*
- 4      **Environmental Problems**  
         *W. Jambers, R. E. Van Grieken*
- 5      **Quantitative Hyleography:**  
         **The Determination of Quantitative Data From Micrographs**  
         *P. J. Goodhew*



# Contents

## Volume 2: Methods II

### IV Electron Microscopy

- 2 Scanning Beam Methods 537**
- 2.1 Scanning Reflection Electron Microscopy 539
  - D. C. Joy*
  - 2.1.1 Introduction 539
  - 2.1.2 Instrumentation 540
  - 2.1.3 Performance 542
  - 2.1.4 Modes of Operation 544
    - 2.1.4.1 Secondary Electron Imaging 544
    - 2.1.4.2 Backscattered Electrons 548
    - 2.1.4.3 Special Techniques 553
  - 2.1.5 Conclusions 561
  - 2.1.6 References 561
- 2.2 Scanning Transmission Electron Microscopy 563
  - J. M. Cowley*
  - 2.2.1 Introduction 563
  - 2.2.2 Scanning Transmission Electron Microscopy Imaging Modes 566
  - 2.2.3 Scanning Transmission Electron Microscopy Theory 570
  - 2.2.4 Inelastic Scattering and Secondary Radiations 574
  - 2.2.5 Convergent-Beam and Nanodiffraction 577
  - 2.2.6 Coherent Nanodiffraction, Electron Holography, Ptychology 578
  - 2.2.7 Holography 581
  - 2.2.8 STEM Instrumentation 584
  - 2.2.9 Applications of Scanning Transmission Electron Microscopy 587
  - 2.2.10 References 592
- 2.3 Scanning Transmission Electron Microscopy: Z Contrast 595
  - S. J. Pennycook*
  - 2.3.1 Introduction 595

2.3.2	Incoherent Imaging with Elastically Scattered Electrons	598
2.3.3	Incoherent Imaging with Thermally Scattered Electrons	601
2.3.4	Incoherent Imaging using Inelastically Scattered Electrons	604
2.3.5	Probe Channeling	606
2.3.6	Applications to Materials Research	610
2.3.6.1	Semiconductors	610
2.3.6.2	Ceramics	613
2.3.6.3	Nanocrystalline Materials	616
2.3.7	References	619
2.4	Scanning Auger Microscopy (SAM) and Imaging X-Ray Photoelectron Microscopy (XPS)	621
	<i>R. De Gryse, L. Fiermans</i>	
2.4.1	Introduction	621
2.4.2	Basic Principles of Auger Electron Spectroscopy (AES) and X-Ray Photoelectron Spectroscopy (XPS)	622
2.4.2.1	Auger Electron Spectroscopy (AES)	622
2.4.2.2	X-Ray Photoelectron Spectroscopy (XPS)	625
2.4.2.3	Quantitative Analysis in AES and XPS	627
2.4.3	Scanning Auger Microscopy (SAM) and Imaging XPS	630
2.4.3.1	Basic Principles of Imaging	630
2.4.3.2	General Aspects of Analyzers	632
2.4.3.3	Energy Resolution of Deflecting Electrostatic Analyzers	635
2.4.3.4	Cylindrical Mirror Analyzer (CMA) versus the Concentric Hemispherical Analyzer (CHA)	637
2.4.3.5	Imaging Techniques	644
2.4.3.6	Magnetic Fields in Imaging XPS	652
2.4.4	Characteristics of Scanning Auger Microscopy Images	654
2.4.4.1	General Aspects	654
2.4.4.2	Background Slope Effects	656
2.4.4.3	Substrate Backscattering Effects	656
2.4.4.4	Topographic Effects	656
2.4.4.5	Beam Current Fluctuation Effects	657
2.4.4.6	Edge Effects	657
2.4.5	Conclusion	658
2.4.6	References	658
2.5	Scanning Microanalysis	661
	<i>R. Gijbels</i>	
2.5.1	Physical Basis of Electron Probe Microanalysis	661
2.5.1.1	Electron Interactions with Solids	661
2.5.1.2	X-Ray Emission Spectra	664
2.5.1.3	Characteristic X-Ray Spectra	666

2.5.1.4	Soft X-Ray Spectra	668
2.5.1.5	X-Ray Continuum	669
2.5.1.6	Overview of Methods of Scanning Electron Beam Analysis	669
2.5.1.7	Electron Probe X-Ray Microanalyzers	669
2.5.1.8	Analytical Electron Microscopes	673
2.5.1.9	Multipurpose Electron Probe Analytical Systems	675
2.5.1.10	X-Ray Emission Spectrometry	679
2.5.1.11	Wavelength-Dispersive Spectrometry	679
2.5.1.12	Energy-Dispersive Spectrometry	680
2.5.1.13	X-Ray Mapping	681
2.5.2	Introduction to Quantitative X-Ray Scanning Microanalysis	682
2.5.2.1	ZAF Method	683
2.5.2.2	Atomic Number Correction	683
2.5.2.3	X-Ray Absorption Correction	684
2.5.2.4	Fluorescence Corrections	684
2.5.2.5	$f(pz)$ Methods	685
2.5.2.6	Standardless Analysis	686
2.5.2.7	Analysis of Thin Films and Particles	687
2.5.3	Conclusions	688
2.5.4	References	689
2.6	Imaging Secondary Ion Mass Spectrometry	691
	<i>P. Van Espen, G. Janssens</i>	
2.6.1	Introduction	691
2.6.1.1	Types of Secondary Ion Mass Spectrometry Measurements	691
2.6.1.2	Dynamic and Static Secondary Ion Mass Spectrometry	692
2.6.1.3	Ion Microscope and Ion Microprobe	692
2.6.1.4	Characteristics of Secondary Ion Mass Spectrometry	693
2.6.2	Secondary Ion Formation	694
2.6.3	Instrumentation	695
2.6.3.1	Primary Ion Sources	695
2.6.3.2	Sample Chamber	696
2.6.3.3	Mass Spectrometer	696
2.6.3.4	Ion Detection and Image Registration	696
2.6.3.5	Typical Configurations	698
2.6.4	Comparison of Ion Microprobe and Ion Microscope Mode	702
2.6.5	Ion Image Acquisition and Processing	704
2.6.5.1	Dynamic Range of Secondary Ion Mass Spectrometry Ion Images	704
2.6.5.2	Influence of Mass Resolution	705
2.6.5.3	Image Sequences	705
2.6.5.4	Interpretation and Processing of Ion Images	706

2.6.5.5	Analysis of Image Depth Sequences	707
2.6.5.6	Analysis of Multivariate Ion Images	709
2.6.6	Sample Requirements	711
2.6.7	Application Domain	712
2.6.8	References	714

## **V Magnetic Methods**

<b>1</b>	<b>Nuclear Magnetic Resonance</b>	<b>719</b>
	<i>D. G. Cory, S. Choi</i>	
1.1	Introduction	719
1.2	Background	720
1.3	Magnetic Field Gradients, Magnetization Gratings, and $k$ -Space	723
1.4	Nuclear Magnetic Resonance	727
1.5	Echoes and Multiple-Pulse Experiments	727
1.6	Two-Dimensional Imaging	730
1.7	Slice Selection	731
1.8	Gratings and Molecular Motions	732
1.9	Solid State Imaging	733
1.10	References	734
 <b>2</b>	 <b>Scanning Electron Microscopy with Polarization Analysis (SEMPA)</b>	 <b>735</b>
	<i>J. Unguris, M. H. Kelley, A. Gavrin, R. J. Celotta, D. T. Pierce, M. R. Scheinfein</i>	
2.1	Introduction	735
2.2	Principle of SEMPA	737
2.3	Instrumentation	739
2.3.1	Scanning Electron Microscopy Probe Forming Column	739
2.3.2	Transport Optics	740
2.3.3	Electron Spin Polarization Analyzers	740
2.3.4	Electronics and Signal Processing	742
2.4	System Performance	743
2.5	Data Processing	744
2.6	Examples	745
2.6.1	Iron Single Crystals	745
2.6.2	CoPt Magneto optic Recording Media	746
2.6.3	Exchange Coupling of Magnetic Layers	746

2.6.4	Magnetic Singularities in Fe-SiO <sub>2</sub> Films	747
2.7	References	748
<b>3</b>	<b>Spin-Polarized Low-Energy Electron Microscopy</b>	<b>751</b>
	<i>E. Bauer</i>	
3.1	Introduction	751
3.2	Theoretical Foundations	751
3.3	Instrumentation	753
3.4	Areas of Application	755
3.5	Discussion	758
3.6	Concluding Remarks	758
3.7	References	758

## **VI Emission Methods**

<b>1</b>	<b>Photoelectron Emission Microscopy</b>	<b>763</b>
	<i>M. Mundschau</i>	
1.1	Introduction	763
1.2	Photoelectron Emission	763
1.3	Microscopy with Photoelectrons	765
1.4	Applications	768
1.4.1	Monolayer Epitaxial Growth	768
1.4.2	Chemical Kinetic Reaction-Diffusion Fronts in Monolayers	769
1.4.3	Magnetic Materials	769
1.5	Choice and Preparation of Samples	771
1.6	References	771
<b>2</b>	<b>Field Emission and Field Ion Microscopy (Including Atom Probe FIM)</b>	<b>775</b>
	<i>A. Cerezo, G. D. W. Smith</i>	
2.1	Field Emission Microscopy	775
2.2	Field Ion Microscopy	777
2.2.1	Principle of the Field Ion Microscope	778
2.2.2	Field Ionization	779
2.2.3	Field Evaporation	780
2.2.4	Image Formation, Magnification, and Resolution	781



2.2.5	Contrast from Lattice Defects and Alloys, and Analysis of Field Ion Microscope Images	784
2.2.6	Specimen-Preparation Techniques	787
2.3	Atom Probe Microanalysis	788
2.3.1	Principles of the Atom Probe Field Ion Microscope	788
2.3.2	Energy Deficits and Energy Compensation	791
2.3.3	Accuracy and Precision of Atom Probe Analysis	792
2.3.4	Atomic Plane Depth Profiling	793
2.3.5	Analysis of Semiconductor Materials	795
2.4	Three-Dimensional Atom Probes	795
2.4.1	Position-Sensing Schemes	797
2.4.2	Mass Resolution in the Three-Dimensional Atom Probe	798
2.4.3	Three-Dimensional Reconstruction of Atomic Chemistry	798
2.5	Survey of Commercially Available Instrumentation	799
2.6	References	800

## VII Scanning Point Probe Techniques

General Introduction 805

<b>1</b>	<b>Scanning Tunneling Microscopy</b>	<b>807</b>
	<i>R. Wiesendanger</i>	
1.1	Introduction	807
1.2	Topographic Imaging in the Constant-Current Mode	807
1.2.1	Effects of Finite Bias	809
1.2.2	Effects of Tip Wave Functions with Angular Dependence	810
1.2.3	Imaging of Adsorbates	811
1.2.4	Spatial Resolution in Constant-Current Topographs	812
1.3	Local Tunneling Barrier Height	815
1.3.1	Local Tunneling Barrier Height Measurements at Fixed Surface Locations	816
1.3.2	Spatially Resolved Local Tunneling Barrier Height Measurements	816
1.4	Tunneling Spectroscopy	817
1.4.1	Scanning Tunneling Spectroscopy at Constant Current	819
1.4.2	Local Spectroscopic Measurements at Constant Separation	820
1.4.3	Current Imaging Tunneling Spectroscopy	820
1.5	Spin-Polarized Scanning Tunneling Spectroscopy	821
1.6	Inelastic Tunneling Spectroscopy	823
1.6.1	Phonon Spectroscopy	824

- 1.6.2      Molecular Vibrational Spectroscopy 824
- 1.7        References 825

## **2            Scanning Force Microscopy 827**

*U. D. Schwarz*

- 2.1        Introduction 827
- 2.2        Experimental Aspects 829
  - 2.2.1      The Force Sensor 829
  - 2.2.2      Deflection Sensors 831
  - 2.2.3      Imaging Modes 832
    - 2.2.3.1    Constant Force Mode 832
    - 2.2.3.2    Variable Deflection Mode 833
    - 2.2.3.3    Noncontact Dynamic Modes 834
    - 2.2.3.4    Imaging Friction, Elasticity, and Viscosity 835
    - 2.2.3.5    Other Imaging Modes 836
  - 2.2.4      Force–Distance Curves 836
  - 2.2.5      Tip Artefacts 837
  - 2.2.6      Scanning Force Microscopy as a Tool for Nanomodifications 838
- 2.3        Theoretical Aspects 838
  - 2.3.1      Forces in Force Microscopy 838
    - 2.3.1.1    Pauli Repulsion and Ionic Repulsion 838
    - 2.3.1.2    Van der Waals Forces 839
    - 2.3.1.3    Adhesion 839
    - 2.3.1.4    Capillary Forces 840
    - 2.3.1.5    Interatomic and Intermolecular Bonds 840
    - 2.3.1.6    Frictional and Elastic Forces 840
    - 2.3.1.7    Magnetic and Electrostatic Forces 841
  - 2.3.2      Contrast Mechanism and Computer Simulations 841
- 2.4        References 842

## **3            Magnetic Force Microscopy 845**

*A. Wadas*

- 3.1        Introduction 845
- 3.2        Force Measurement 845
- 3.3        Force Gradient Measurement 849
- 3.4        References 852

<b>4</b>	<b>Ballistic Electron Emission Microscopy</b>	<b>855</b>
	<i>J. DiNardo</i>	
4.1	Introduction	855
4.2	Experimental Considerations	857
4.3	First Demonstrations of Ballistic Electron Emission Microscopy	858
4.4	Theoretical Considerations	860
4.5	Ballistic Electron Emission Microscopy Analysis of Schottky Barrier Interfaces	864
4.5.1	Epitaxial Interfaces	865
4.5.2	Nonepitaxial Interfaces	867
4.5.3	Au/Si Interfaces	868
4.5.4	Metal-Film Dependence	869
4.5.5	Surface Gradients	869
4.5.6	Interfacial Nanostructures	870
4.5.7	Local Electron Tunneling Effects	872
4.5.8	Impact Ionization	873
4.6	Probing Beneath the Schottky Barrier	874
4.7	Ballistic Hole Transport and Ballistic Carrier Spectroscopy	879
4.8	Summary	881
4.9	References	881

## **VIII Image Recording, Handling and Processing**

<b>1</b>	<b>Image Recording in Microscopy</b>	<b>885</b>
	<i>K.-H. Herrmann</i>	
1.1	Introduction	885
1.2	Fundamentals	885
1.2.1	The Primary Image	885
1.2.2	The General Recorder	886
1.2.3	Quantum Efficiency of Conversion Processes	889
1.2.3.1	Photographic Recording	890
1.2.3.2	Photoeffect	890
1.2.3.3	Scintillators	892
1.2.3.4	Light Optical Elements	893
1.2.3.5	Secondary Emission	895
1.2.3.6	Electron Beam-Induced Conduction	896
1.2.3.7	Imaging Plate	896
1.2.4	Composed Systems and Optoelectronic Components	896

1.2.4.1	Scintillator-Photosensor Combination	897
1.2.4.2	Image Intensifiers	897
1.2.4.3	Microchannel Plates	899
1.2.4.4	Television Camera Tubes	899
1.2.4.5	Charge-Coupled Devices	901
1.2.5	Resolution and Sampling	904
1.3	Light Microscopy	905
1.3.1	Video Recording	905
1.3.2	Low-Light-Level Detection	905
1.4	Electron Microscopy	906
1.4.1	Photographic Recording	906
1.4.2	Imaging Plate	907
1.4.3	Electronic Recording	908
1.4.3.1	Television Chains	908
1.4.3.2	Slow-Scan Charge-Coupled Device Converters with a Scintillator	909
1.4.3.3	Directly Back-Illuminated Charge-Coupled Devices	912
1.5	X-Ray Microscopy	912
1.5.1	Photographic Film and Imaging Plate	913
1.5.2	Resist	914
1.5.3	Transmission Photocathodes	915
1.5.4	Microchannel Plates	916
1.5.5	Television Chains	916
1.5.5.1	X-ray-Sensitive Vidicons	917
1.5.5.2	Conversion to Visible Radiation	917
1.5.6	Slow-Scan Charge-Coupled Device Chains	917
1.5.7	Directly Illuminated Charge-Coupled Device Sensors	918
1.6	References	919

## **2 Image Processing 923**

*N. Bonnet*

2.1	Introduction	923
2.2	Image Preprocessing	924
2.2.1	Global Methods for Image Preprocessing	925
2.2.1.1	Examples of Global Image Preprocessing in Image Space or Image Frequency Space	925
2.2.1.2	Examples of Global Image Preprocessing in Parameter Space	926
2.2.2	Local Methods for Image Preprocessing	927
2.2.2.1	Example of Algorithm for Local Contrast Enhancement	928
2.2.2.2	Example of Algorithm for Improving the Signal-to-Noise-Ratio	929

2.3	Processing of Single Images	931
2.3.1	Image Restoration	931
2.3.1.1	Restoration of Linear Degradations	931
2.3.1.2	Restoration of Partly Linear Degradations: Very High Resolution Electron Microscopy	932
2.3.1.3	Example of a Completely Nonlinear Restoration: Near-Field Microscopies	933
2.3.2	Image Segmentation	934
2.3.2.1	Segmentation on the Basis of Grey Levels Only	934
2.3.2.2	Segmentation on the Basis of Grey-Level Gradients	935
2.3.2.3	Segmentation on the Basis of Grey-Level Homogeneity and the Concept of Connectivity	935
2.3.2.4	Segmentation on the Basis of Grey Levels, Gradients, and Connectivity: Functional Minimization	936
2.3.2.5	Mathematical Morphology	936
2.4	Analysis of Single Images	937
2.4.1	Object Features	937
2.4.2	Pattern Recognition	937
2.4.3	Image Analysis without Image Segmentation	938
2.4.3.1	Texture Analysis	938
2.4.3.2	Fractal Analysis	939
2.4.3.3	Stereology	939
2.4.4	Mathematical-Morphology Approach to Image Analysis	940
2.4.4.1	Granulometry	940
2.4.4.2	Distance Function	940
2.4.4.3	Skeleton	942
2.5	Processing/Analysis of Image Series	942
2.5.1	Three-Dimensional Reconstruction	942
2.5.1.1	Serial Sections	943
2.5.1.2	Stereoscopy	943
2.5.1.3	Microtomography	944
2.5.1.4	Three-Dimensional Display	945
2.5.2	Processing and Analysis of Spectral, Temporal and Spatial Image Series	945
2.6	Conclusion	950
2.7	References	950

## IX Special Topics

### 1 Coincidence Microscopy 955

*P. Kruit*

- 1.1 Introduction 955
- 1.2 Instrumentation 956
- 1.3 Coincidence Count Rates 958
- 1.4 Signal Combinations 958
  - 1.4.1 EELS–Emitted Electron 958
  - 1.4.2 EELS–X-ray 961
  - 1.4.3 EELS–Cathodoluminescence 961
  - 1.4.4 Backscattered Electron–Secondary Electron 962
  - 1.4.5 Other Combinations 962
- 1.5 References 962

### 2 Low Energy Electron Holography and Point-Projection Microscopy 963

*J. C. H. Spence*

- 2.1 Introduction and History 963
- 2.2 Electron Ranges in Matter: Image Formation 966
- 2.3 Holographic Reconstruction Algorithms 970
- 2.4 Nanotips, Tip Aberrations, Coherence, Brightness, Resolution Limits, and Stray Fields 973
- 2.5 Instrumentation 978
- 2.6 Relationship to Other Techniques 981
- 2.7 Future Prospects, Radiation Damage, and Point Reflection Microscopy 982
- 2.8 References 985

## General Reading

## List of Symbols and Abbreviations

## List of Techniques

## Index

## Part IV

---

# Electron Microscopy

## **2 Scanning Beam Methods**





## 2.1 Scanning Electron Microscopy

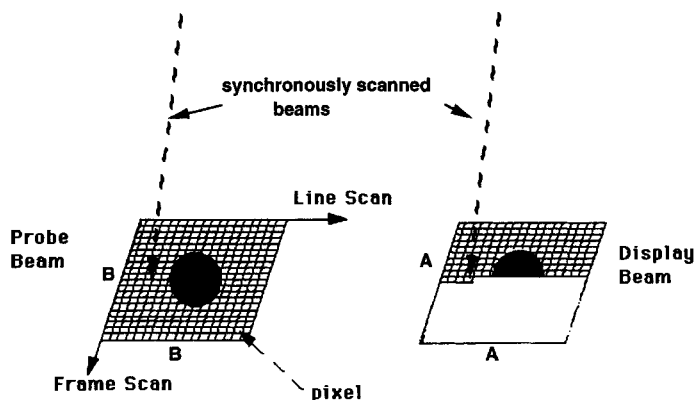
### 2.1.1 Introduction

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

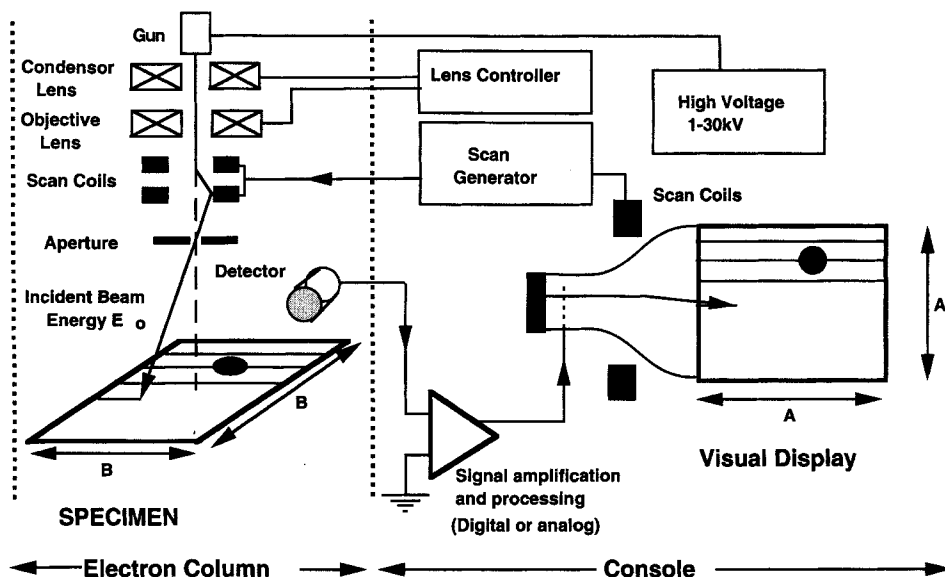


Figure 2. Basic components of the scanning electron microscope.

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

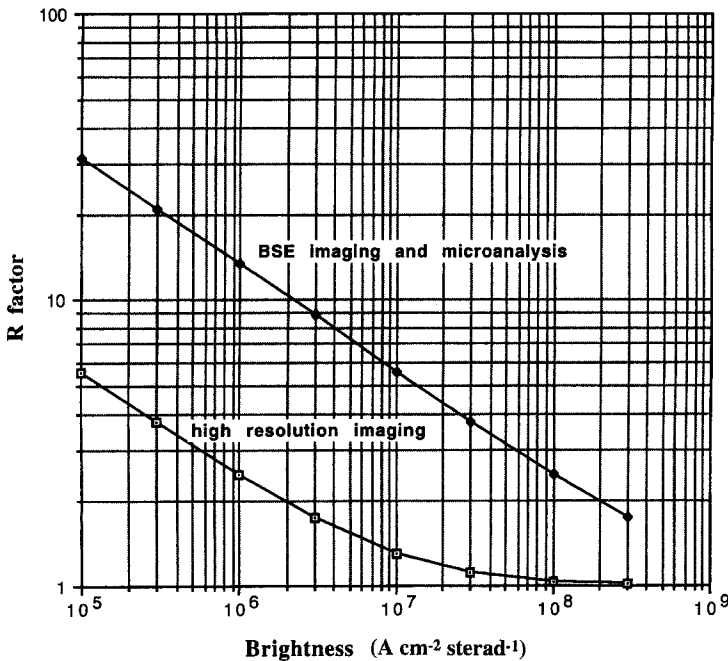
through an analog to digital converter (ADC) and then handled digitally rather than as an analog video signal. This permits images to be stored, enhanced, combined, and analyzed using either an internal or an external computer. Although the majority of images are still recorded on to photographic film, 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  $10000\times$  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

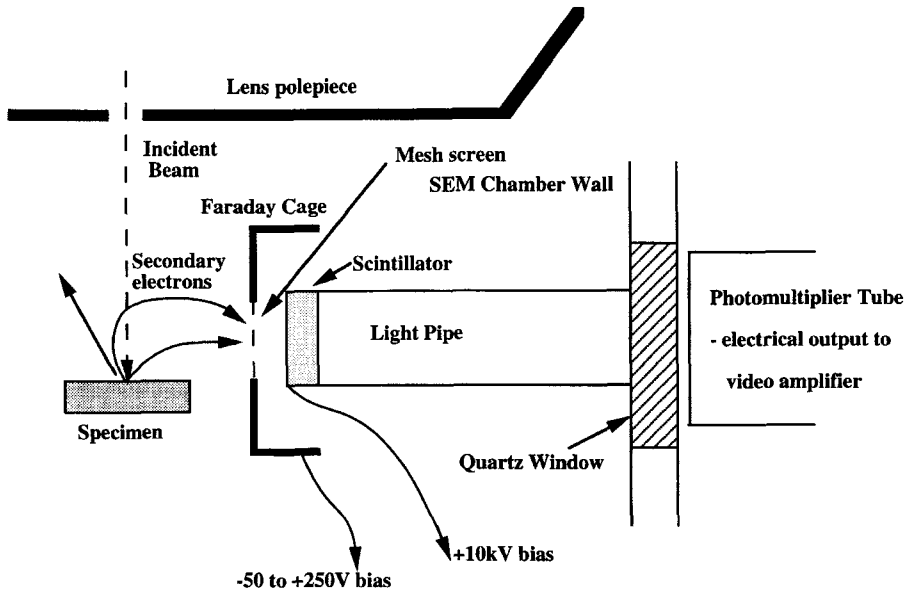


Figure 4. Everhart-Thornley SE detector and specimen chamber geometry.

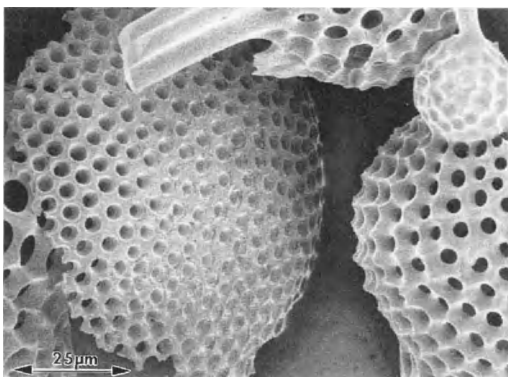


to produce, and the detector may be placed anywhere within the specimen chamber of the SEM and still collect 50 to 60% of the SE leaving the specimen. On some advanced instruments the SE detector may be placed above the objective lens and out of line of sight of the sample [10]. The magnetic field of 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.

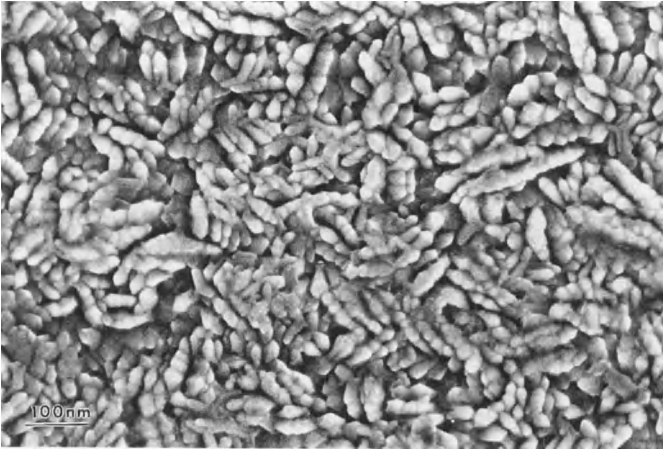
The general characteristics of SE signals are illustrated in Fig. 5 which shows a collection of diatoms. The image has a marked three-dimensional effect caused by the lighting and shadow effects which decorate each object, and each edge is outlined by a fine bright line which adds definition to the detail. 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

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 5.** Secondary electron images of Radiolarium. Recorded in Hitachi S-4500 field emission SEM at 5 keV beam energy. Magnification: 800 $\times$ .

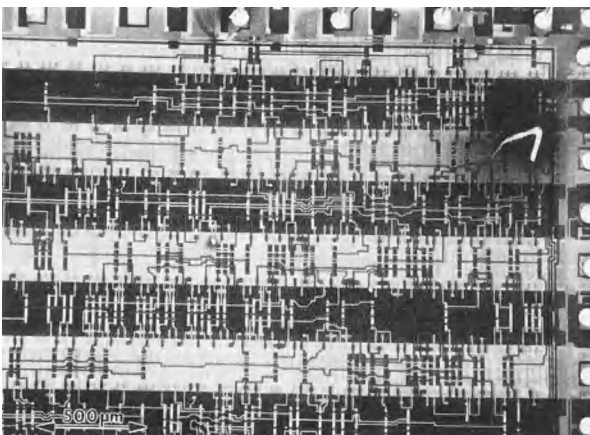


**Figure 6.** High resolution image of magnetic disc media surface recorded at 30 keV in Jeol JSM890 field-emission SEM.

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  $0.1\text{ }\mu\text{m}$ . 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

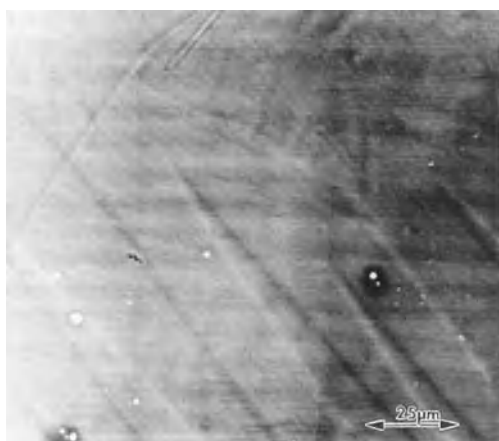


**Figure 8.** BSE Z-contrast image of Assyrian glass showing  $\text{Cu}_2\text{O}$  dendrites. Recorded on Cambridge S250 SEM at 10 keV.

which were later determined to be  $\text{Cu}_2\text{O}$ , probably dissolved from the vessel in which the glass was melted. The phases are readily distinguished here because the glass, essentially  $\text{SiO}_2$ , 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.

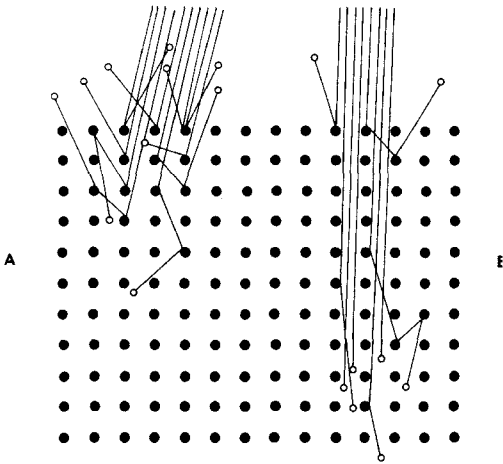
Unlike secondary electrons which come from the surface region of the specimen, backscattered electrons emerge from a volume which is of the order of one-third

of the incident beam range in depth and radius. To a first approximation, the range  $R$  (in nm) is given by

$$R = \frac{75E_0^{5/3}}{\rho} \quad (2)$$

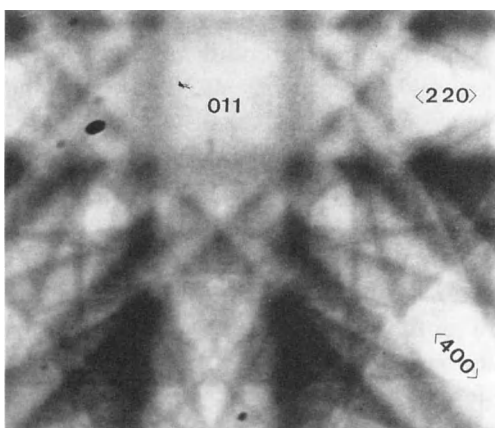
where  $E_0$  is the beam energy in keV and  $\rho$  is the density of the specimen in  $\text{g cm}^{-3}$ . Thus for beam energies of 10 keV, and typical densities, the backscattered signal comes from a region of the order of a fraction of a micrometer or more in breadth and depth. Consequently, the 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.

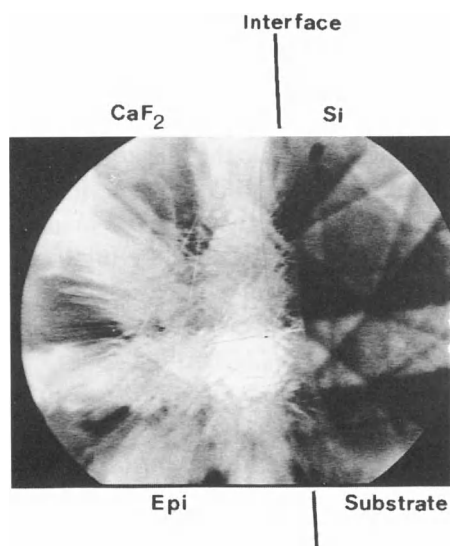
of the crystal. The image, or electron channeling pattern, produced in such a case is illustrated in Fig. 11 for a single crystal of InP. Here the variation in the angle of incidence has been achieved simply by scanning the large crystal at low magnification, so that between extremes of the field of 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



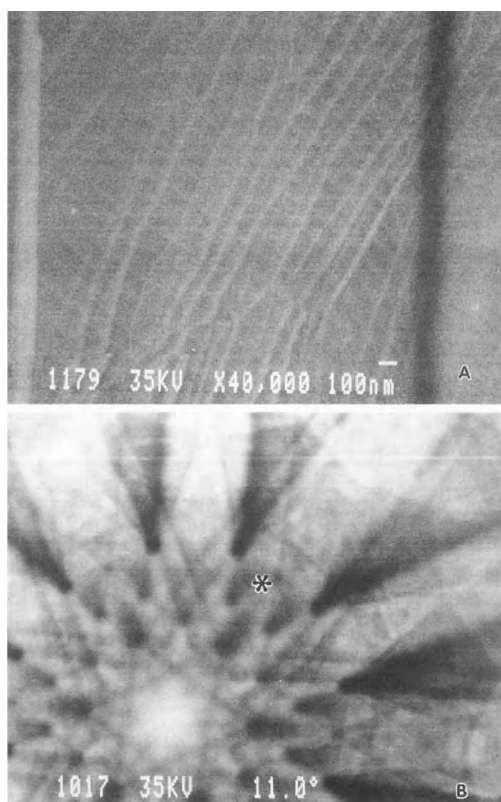
**Figure 12.** Selected area channeling pattern across the epitaxial interface between  $\text{CaF}_2$  and silicon. Recorded on Cambridge S250 at 25 keV.

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  $1\text{ }\mu\text{m}$  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  $\text{MoS}_2$  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

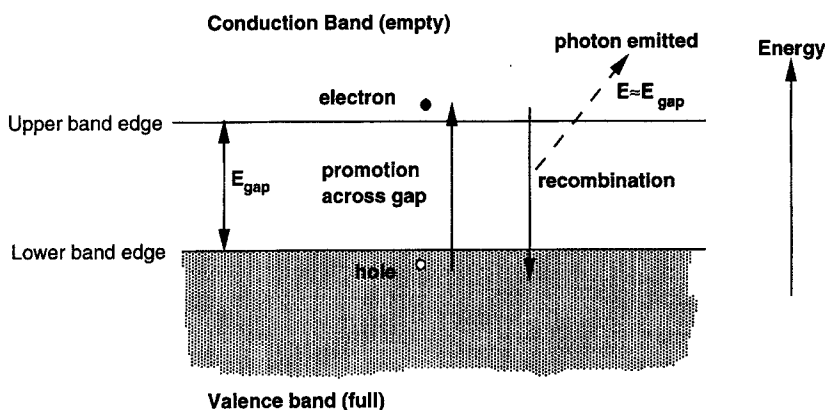
Unlike the techniques already discussed the imaging modes that will now be

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  $E_{\text{eh}}$  required to form one electron-hole pair is about three times the band gap, for example, in silicon  $E_{\text{eh}}$  is 3.6 eV. A single incident electron of energy  $E_0$  can then create about  $E_0/E_{\text{eh}}$  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





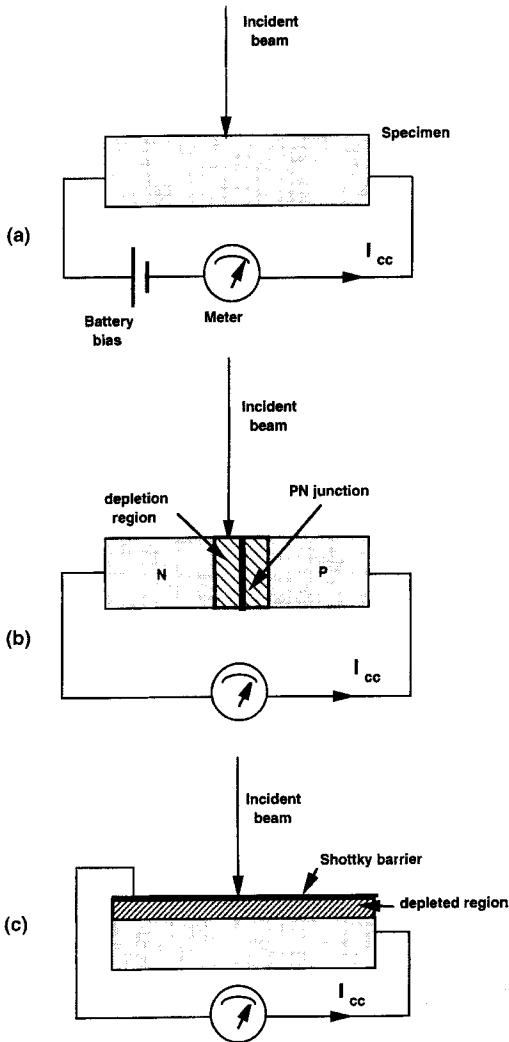
**Figure 14.** Band gap diagram for a semiconductor.

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.

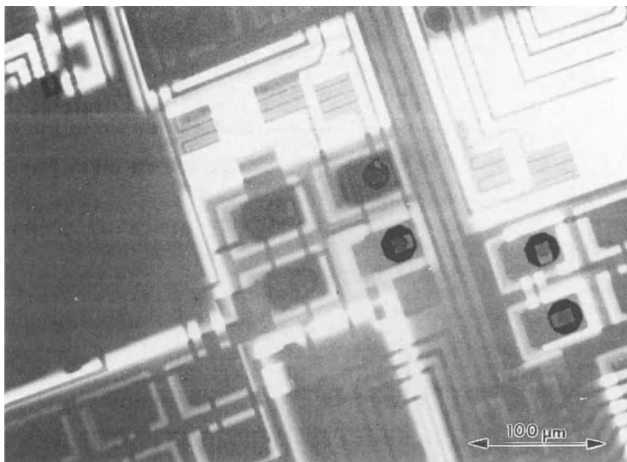
the external circuit. This current will be  $e(E_0/E_{ch})$  where  $e$  is the electronic charge and is referred to as the electron beam induced current (EBIC). (To avoid confusion it is better to describe all of these modes as being 'charge collected microscopies'.) If the beam is placed outside of 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

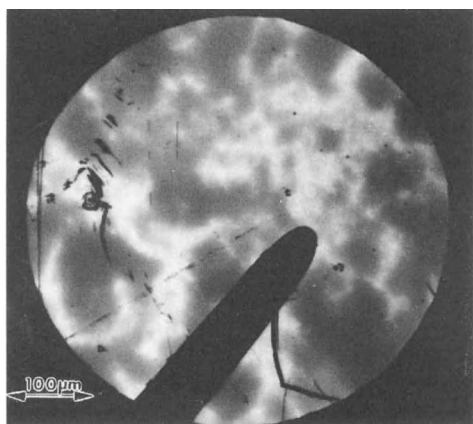


**Figure 16.** EBIC image of p-n junctions in an integrated circuit. Recorded at 15 keV in Hitachi S-800 FE SEM.

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

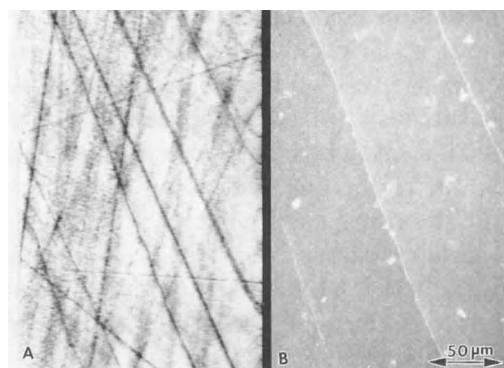
the junction and its depletion region. The technique is made more useful, as shown in Fig. 15(c), when a Schottky barrier is deposited on to the surface of the semi-conducting material. The depleted region now extends beneath the whole area of the barrier, and downwards from the surface to a depth dependent 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 electronics 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

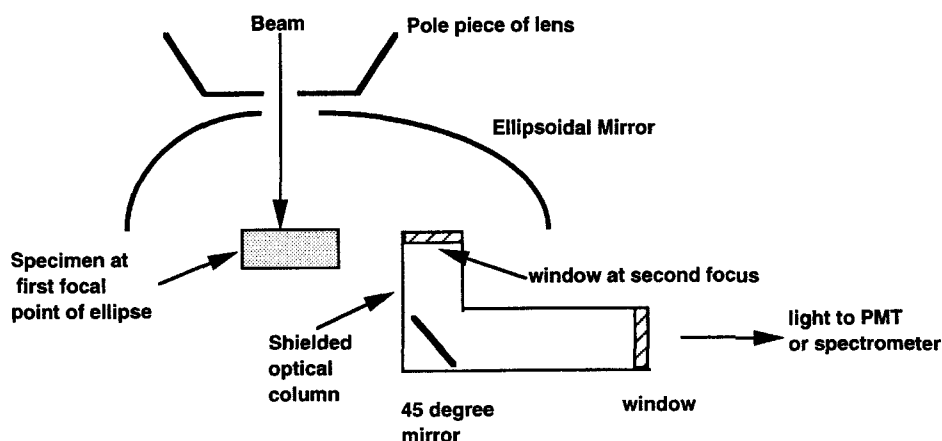
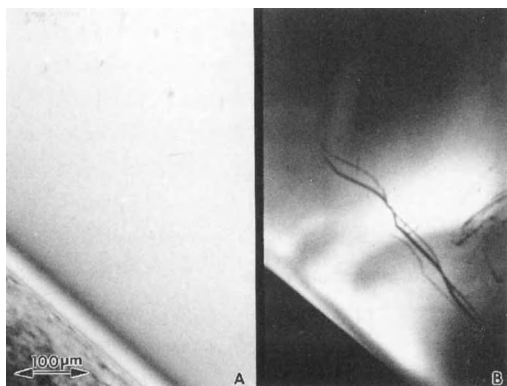


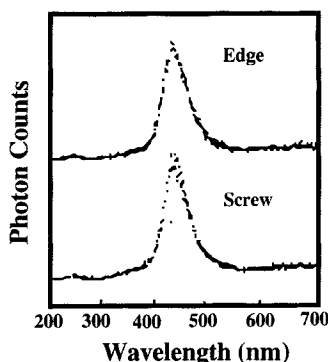
Figure 19. Schematic layout for doing CL in SEM.

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. Serce, 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

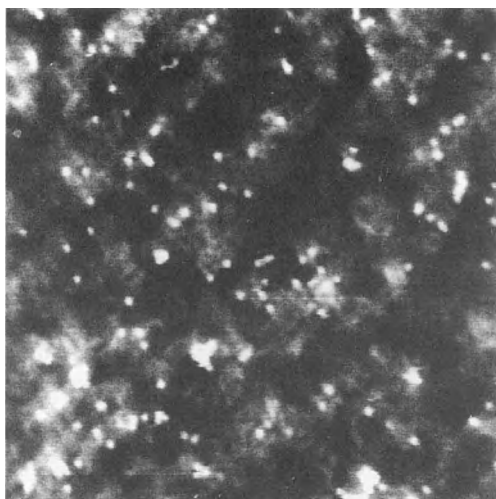
Although the concept of scanning transmission electron microscopy (STEM) was not new, the possibility of employing the scanning principle for high-resolution imaging of thin specimens in transmission was first realized by Crewe, who introduced the use of field-emission guns for this purpose [1, 2]. Dark-field 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,

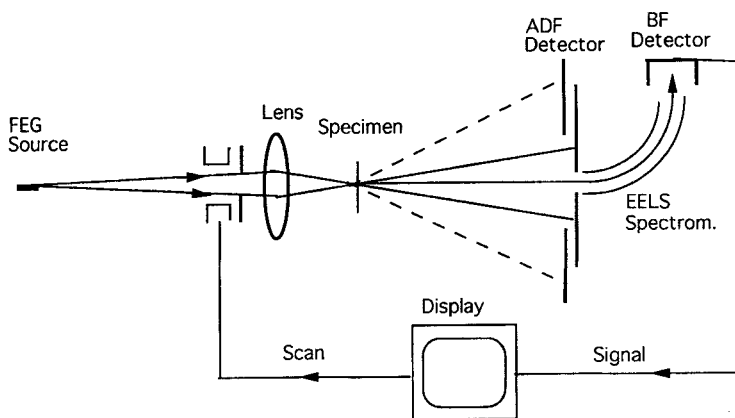


**Figure 1.** Annular dark-field image of uranium atoms on an amorphous carbon film. Full width of image: 31.5 nm. (Courtesy of J. Wall, Brookhaven National Lab.).

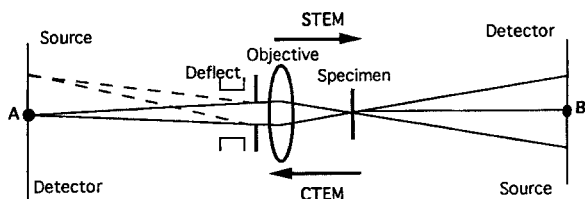
but with the electrons traveling in the opposite direction. In this diagram condenser and projector lenses have been omitted and only the essential objective

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 CTM: 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 CTM, 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 CTM, 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 CTM, 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 CTM. 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 CTM, the chromatic aberration of the objective

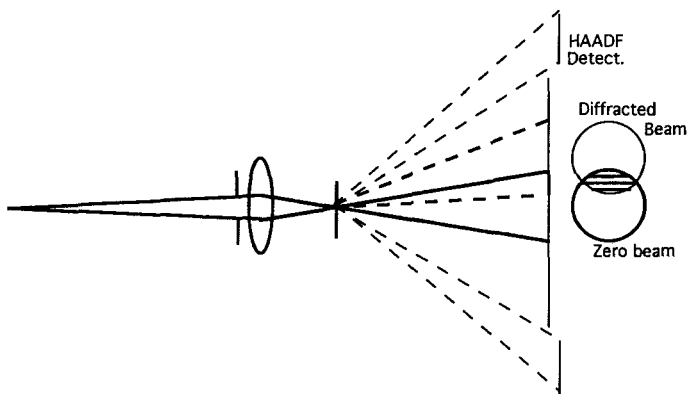
lens then degrades the resolution, but in STEM, because the objective lens comes before the specimen, the effect of such an energy spread is negligible. Hence, in general, the resolution and/or contrast is better for STEM than for CTEM for a given specimen thickness: or, for a given resolution level, 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

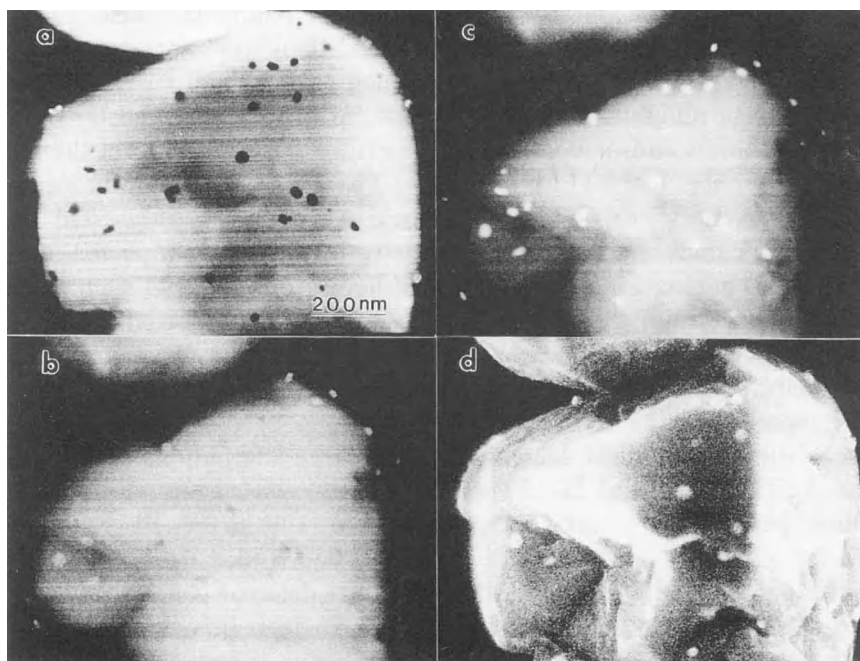
collection of the dark-field signal is much better than for dark-field CTEM, for which only a small part of the diffraction pattern can be collected by the objective aperture and used to form the image.

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.



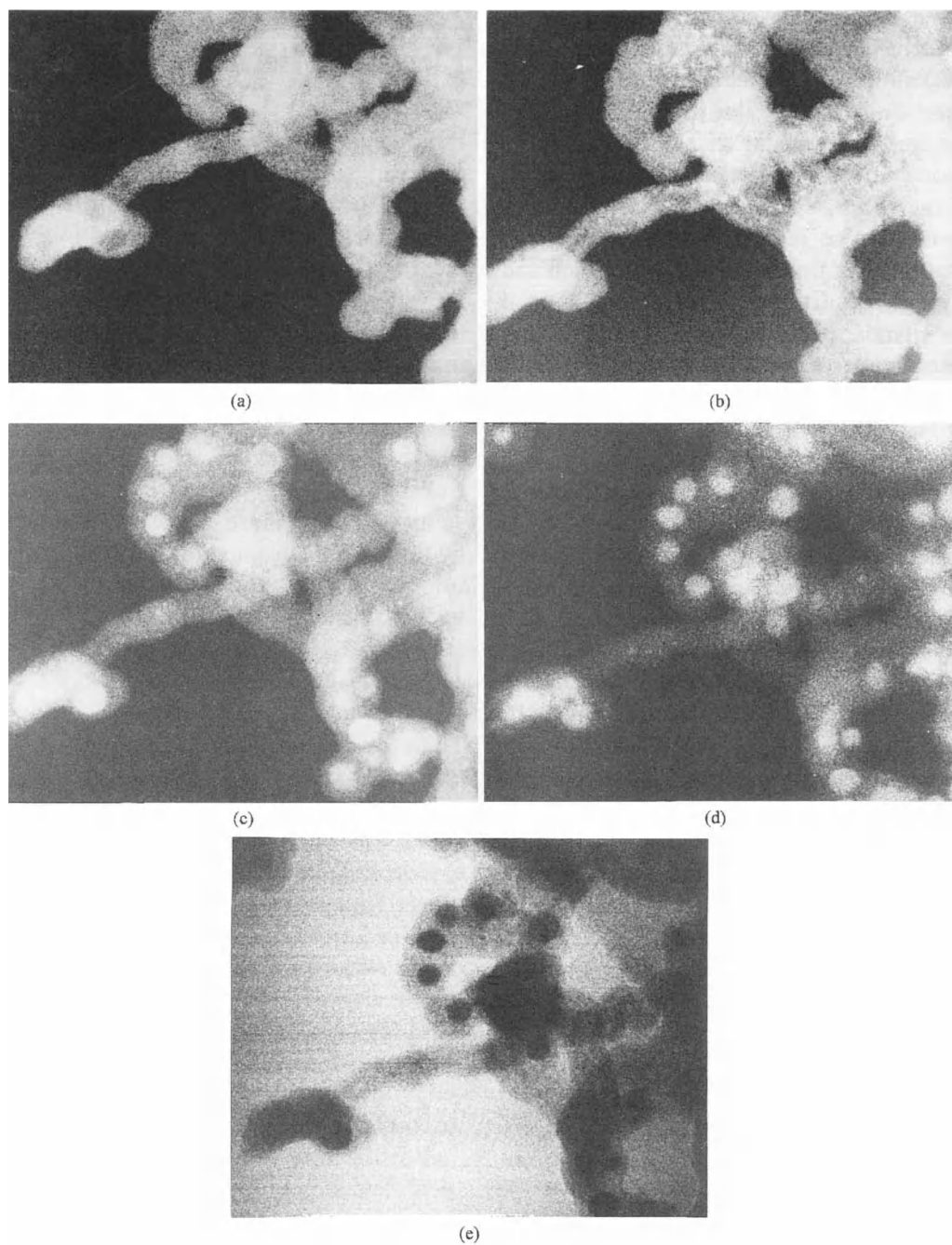
**Figure 5.** (a)–(c) ADF STEM images of Pt particles on a thick NaY–zeolite crystal with inner collector angle 20 mrad for (a) and increasing for (b) and (c), plus a SEM image (d), of the same specimen showing the Pt particles on one surface. Note the reversals of contrast of the Pt particles as the 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  $1\text{ }\mu\text{m}$ , so that it is a good assumption that electron waves arriving at the specimen from all parts of a  $50$  or  $100\text{ }\mu\text{m}$  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 under-focus, 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 CTM 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

$$\begin{aligned} |\Psi(u + \tfrac{1}{2})|^2 &= |F_h \exp[i\chi(\tfrac{1}{2}a) + \pi i X/a] \\ &\quad + F_{h+1} \exp[i\chi(-\tfrac{1}{2}a) - \pi i X/a]|^2 \\ &= |F_h|^2 + |F_{h+1}|^2 + 2|F_h| \\ &\quad \times |F_{h+1}| \cos(2\pi X/a + \alpha) \quad (9) \end{aligned}$$

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].

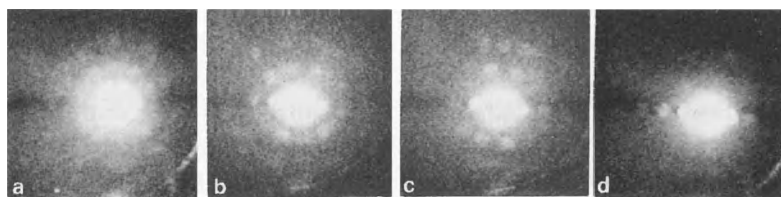
## 2.2.5 Convergent-Beam and Nanodiffraction

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.

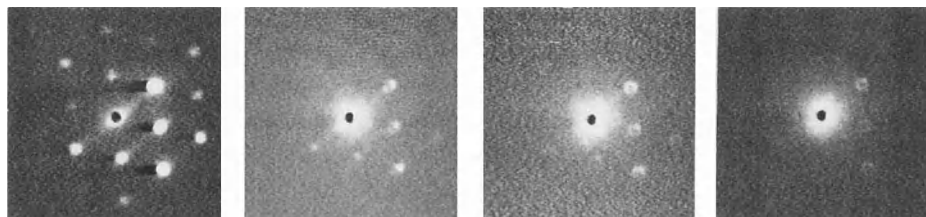
If they do overlap, interference fringes appear in the region of overlap if the incident beam is coherent [45, 46]. The coherent interference effects in diffraction patterns will be considered in detail in the next part of this Section. For now we concentrate on the possibilities for obtaining diffraction 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.

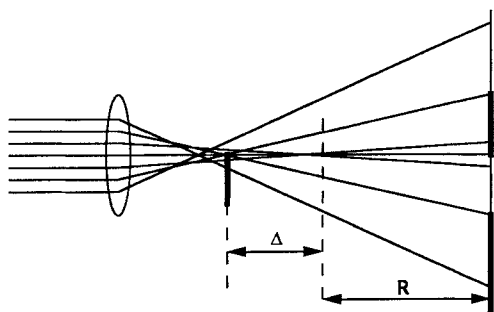
Imperfections in crystals, including crystal edges, faults or other defects, give scattering in between the Bragg diffraction spots of the pattern. Then coherent interference effects, characteristic of the type of defect, are seen. When the region of the specimen illuminated by the incident beam includes 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].

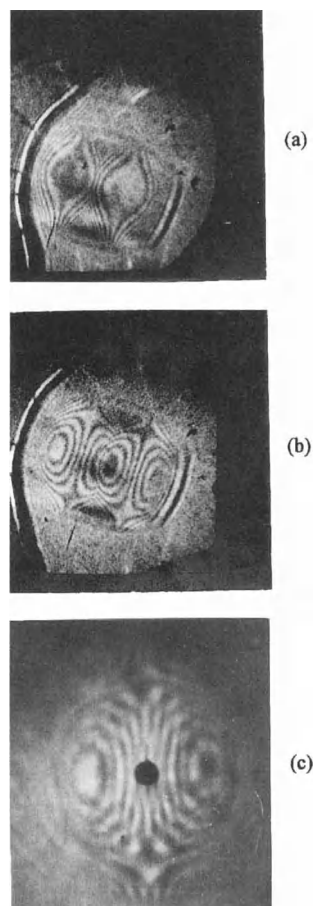


**Figure 9.** Geometrical optics ray diagram of the shadow-image formation with strong spherical aberration. The infinite-magnification circle is formed when adjacent rays cross over at the specimen plane. Within this circle, the image of the straight-edge is inverted.

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 CTM 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. Konnert 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

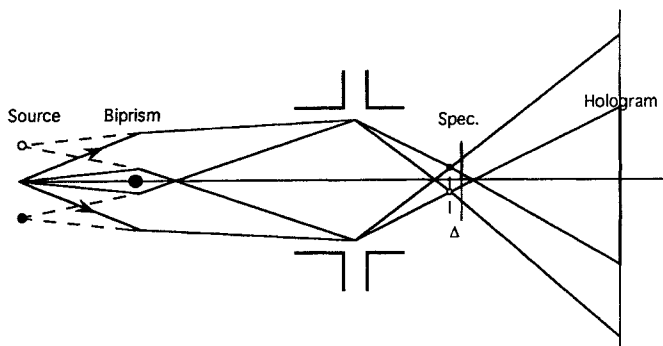
light-optical analogue experiments and STEM experiments with moderate resolution. Complications arise, for example, from the fact that the aperture limitation cannot be ignored, but evidence that the method may provide considerable enhancement of resolution has been given.

## 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.

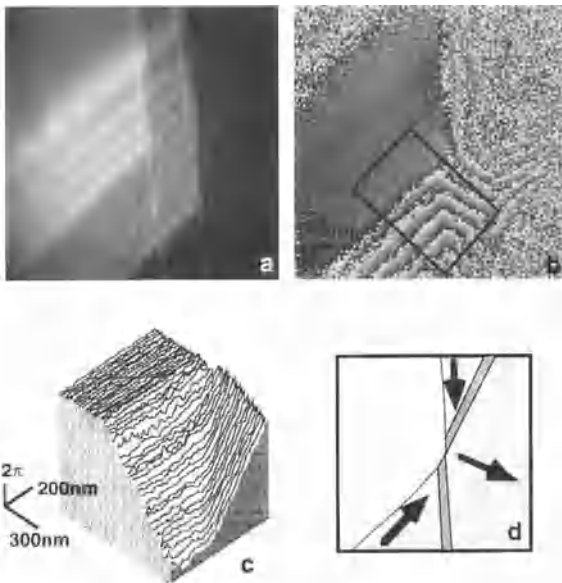


patterns can be used to reduce, and effectively eliminate, the effects of the conjugate images [55]. An alternative approach is that of Veneklasen [65], who proposed that by making a detector with a detector function  $D(u) = T(u)$ , an on-line reconstruction of the wave-function may be made, giving 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.

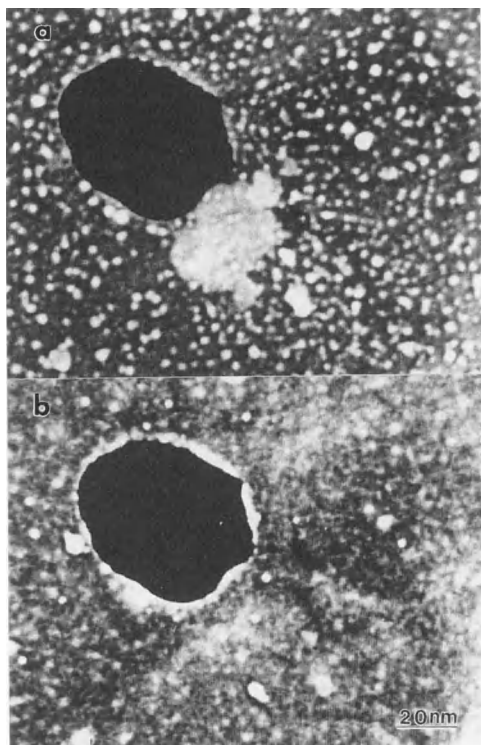
In observations of heavy atoms sitting on thin films of amorphous carbon, it appeared that the heavy atom positions tended to change from one image scan to the next [78]. Experiments were made to determine whether the atom movements resulted from energy pulses given to the atoms by the incident electron 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.

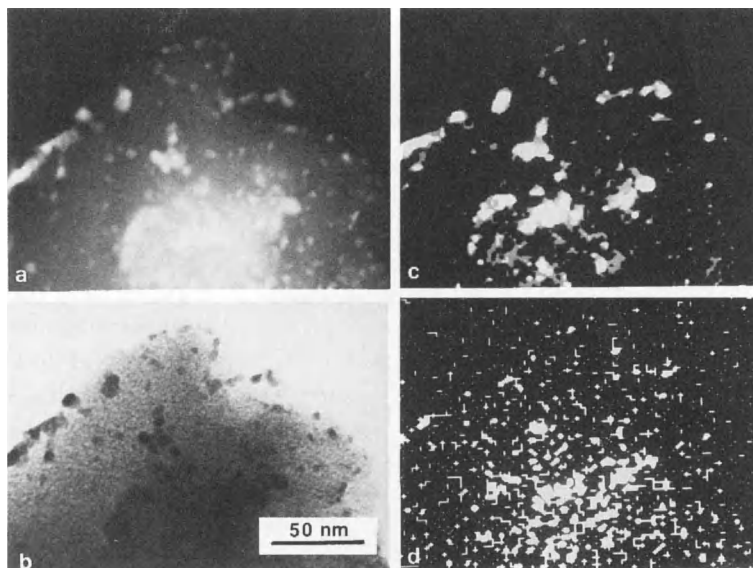
The problem of locating heavy-metal atoms or small clusters within the channels in the structures of zeolites becomes more significant as the use of zeolites in industrial catalysts becomes more extensive. A major difficulty for the application of electron microscopy techniques to such materials is that 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

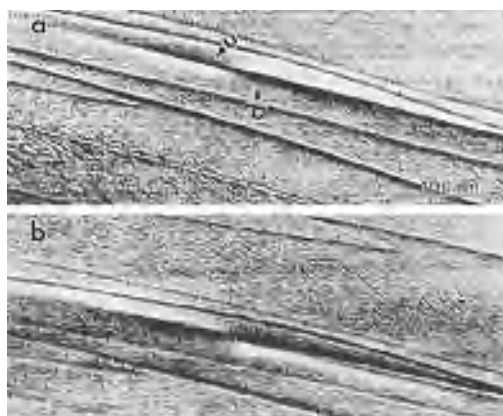
disagreement with the theoretical studies which referred only to isolated particles with no perturbation of surface energies by support material.

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. XIIth 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

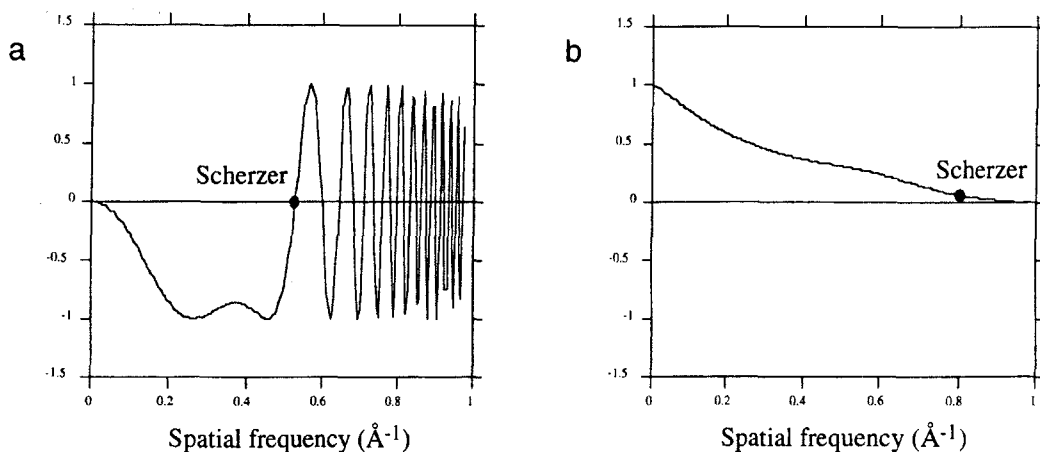
### 2.3.1 Introduction

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].

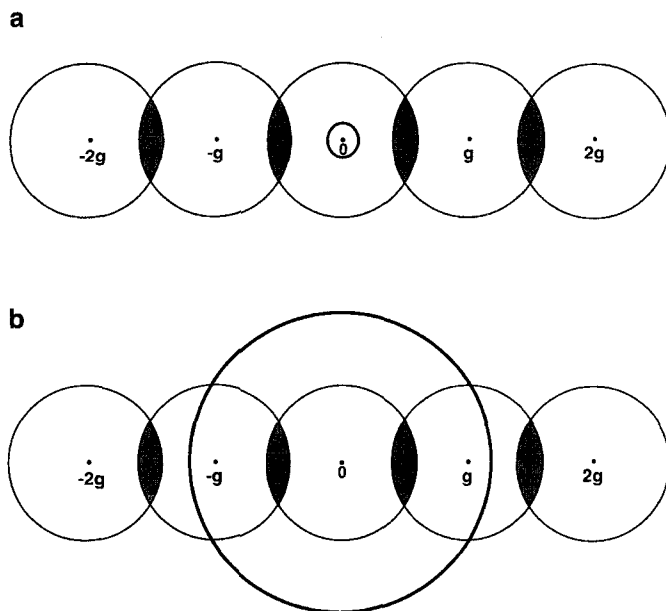
These resolution criteria are deduced assuming the objective aperture can be set to the optimum size appropriate to the imaging conditions used. For a



**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.

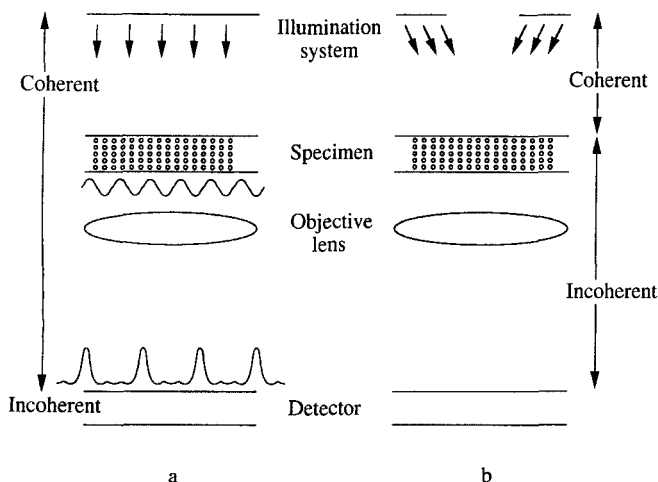
spacings up to the aperture diameter. The dark field image therefore shows double the resolution of the bright field image.

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

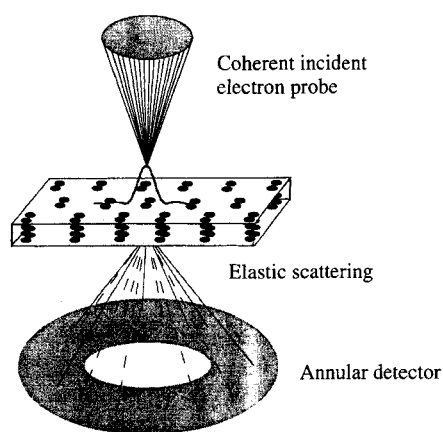


**Figure 3.** Schematic comparing (a) coherent and (b) incoherent imaging optics for a transmission electron microscope.

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-

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.



**Figure 4.** Schematic of the STEM optics for incoherent imaging.

## 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,

$$P(\mathbf{R} - \mathbf{R}_0) = \int_{\text{objective aperture}} \exp\{i[\mathbf{K} \cdot (\mathbf{R} - \mathbf{R}_0) + \gamma(\mathbf{K})]\} d\mathbf{K} \quad (1)$$

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,

$$\psi_s(\mathbf{K}_f) = \frac{m}{2\pi\hbar^2} \int \exp[i\mathbf{K}_f \cdot \mathbf{R}] \times V(\mathbf{R}) P(\mathbf{R} - \mathbf{R}_0) d\mathbf{R} \quad (3)$$

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

$$\int \exp[-i\mathbf{K}_f \cdot (\mathbf{R} - \mathbf{R}')] d\mathbf{K}_f = (2\pi)^2 \delta(\mathbf{R} - \mathbf{R}') \quad (4)$$

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

$$\tilde{O}(\mathbf{K}) = \int O(\mathbf{R}) \exp(-i\mathbf{K} \cdot \mathbf{R}) d\mathbf{R} \quad (7)$$

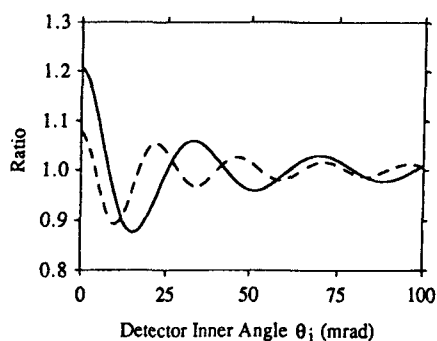
which can be written as a self-convolution of elastic scattering form factors  $f(\mathbf{K})$ ,

$$\tilde{O}(\mathbf{K}) = \frac{1}{\chi^2} \int f(\mathbf{K} + \mathbf{K}') f(\mathbf{K}') d\mathbf{K}' \quad (8)$$

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,

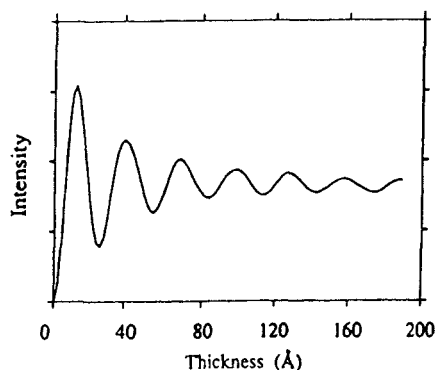
$$O(\mathbf{R}) = \sigma^2 |V(\mathbf{R}) * d(\mathbf{R})|^2 \quad (10)$$

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,

$$\begin{aligned}\tilde{O}^{\text{TDS}}(\mathbf{K}) &= \frac{1}{\chi^2} \int f(|\mathbf{s}'|) f(|\mathbf{s} - \mathbf{s}'|) \\ &\times [\exp(-M s^2) - \exp(-M s'^2)] \\ &\times \exp[-M(\mathbf{s} - \mathbf{s}')^2] d\mathbf{K}'\end{aligned}\quad (12)$$

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

$$\begin{aligned}I(\mathbf{R}_0) &\approx P^2(\mathbf{R} - \mathbf{R}_0) \int O^{\text{TDS}}(\mathbf{R}) d\mathbf{R} \\ &= P^2(\mathbf{R} - \mathbf{R}_0) \tilde{O}^{\text{TDS}}(0)\end{aligned}\quad (13)$$

where  $\tilde{O}^{\text{TDS}}(0)$  is simply the atomic diffuse scattering distribution integrated over the high-angle detector,

$$\begin{aligned}\tilde{O}^{\text{TDS}}(0) &= \frac{1}{\chi^2} \int_{\text{detector}} f^2(|\mathbf{s}'|) \\ &\times [1 - \exp(-2M s'^2)] d\mathbf{K}' \\ &= \sigma_{\text{at}}\end{aligned}\quad (14)$$

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,

$$W_{mn} = \langle \exp\{i\mathbf{k} \cdot (\mathbf{u}_m - \mathbf{u}_n)\} \rangle = \exp \left[ 2M \left( \frac{\text{Si}(2\pi q_B r_{mn})}{2\pi q_B r_{mn}} - 1 \right) \right] \quad (17)$$

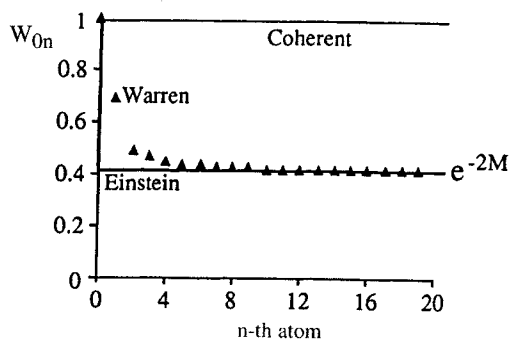
This factor expresses the degree of correlation between the  $m$ th and  $n$ th atoms in a column, where  $r_{mn} = |\mathbf{r}_m - \mathbf{r}_n|$  and  $\text{Si}(x)$  is the sine integral function

$$\text{Si}(x) = \int_0^x \frac{\sin u}{u} du \quad (18)$$

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

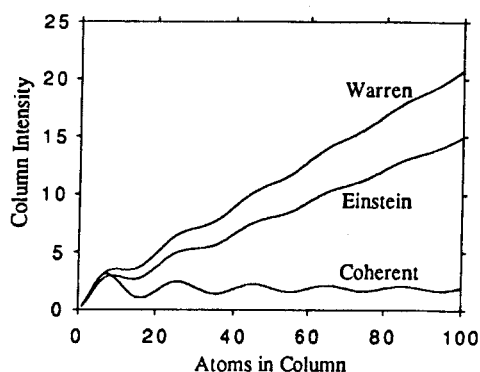
$$W_{0n} = \exp[2M(\delta_{0n} - 1)] \quad (19)$$

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

$$O'(\mathbf{R}) = \left( \frac{e^2}{\pi \hbar v} \right)^2 \times \left| \int \frac{\rho_{n0}}{q^2}(\mathbf{q}) \exp(-i\mathbf{K} \cdot \mathbf{R}) d\mathbf{K} \right|^2 \quad (21)$$

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],

$$\tilde{O}'(\mathbf{K}) = \left( \frac{2e^2}{\hbar v} \right)^2 \int \frac{\rho_{n0}(\mathbf{q}')}{q'^2} \frac{\rho_{n0}^*(\mathbf{q}' + \mathbf{K})}{|\mathbf{q}' + \mathbf{K}|^2} d\mathbf{K}' \quad (22)$$

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

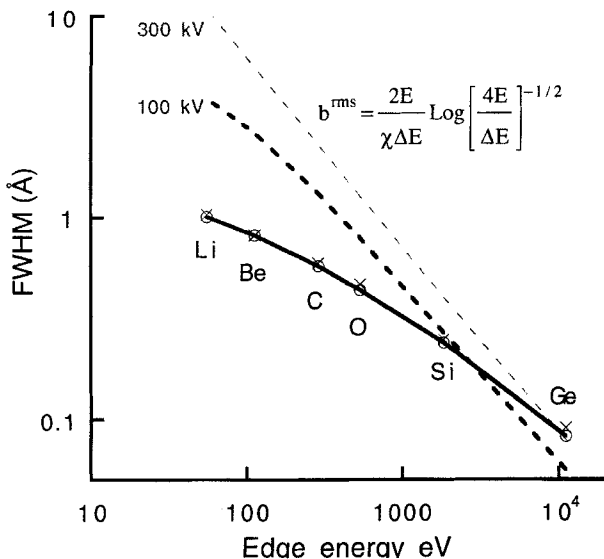
$$\int \rho_{n0}(\mathbf{q}', \kappa) \rho_{n0}^*(\mathbf{q}' + \mathbf{K}, \kappa) d\kappa \quad (24)$$

Analytical expressions have been given for these integrals [25, 26], and their general properties have been considered by Kohl and Rose [27].

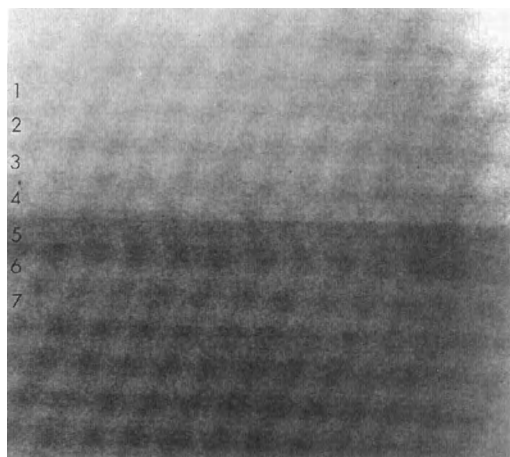
Figure 9 plots the full-width half-maxima of object functions calculated for a number of elements, assuming that all the scattered electrons are collected. Somewhat surprisingly it would appear that atomic resolution is possible even for the lightest elements. This is a significant difference from impact 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  $\hat{O}'(\mathbf{K})$ . 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].

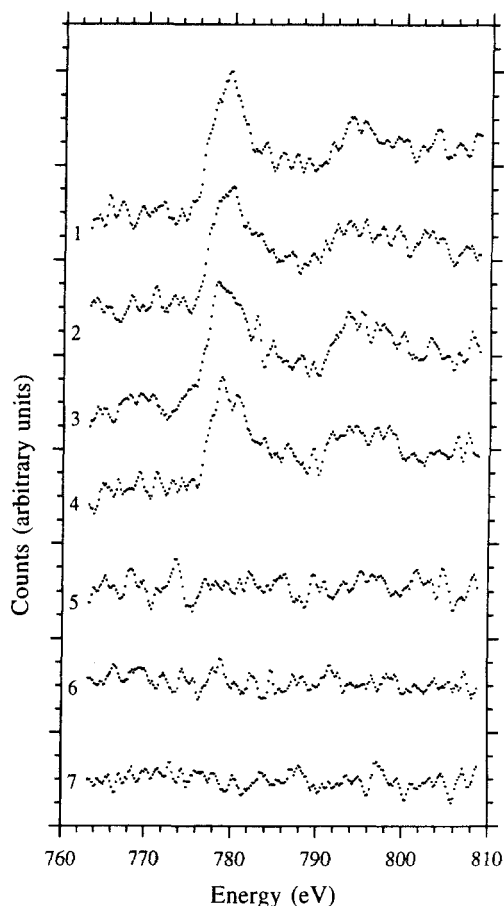


**Figure 10.** The Z-contrast image of a  $\text{CoSi}_2$ -Si (111) epitaxial interface used as an edge resolution test specimens for EELS. Cobalt atoms image bright; planes numbered correspond to the spectra in Fig. 11.

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,

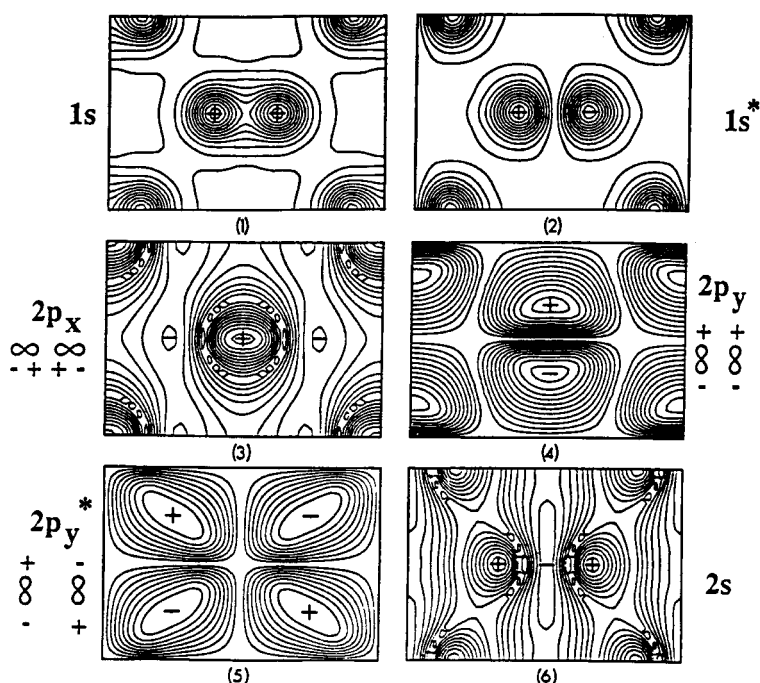
$$\begin{aligned}\psi(\mathbf{R}, z) = & \sum_j \epsilon^j(\mathbf{K}) \tau^j(\mathbf{K}, \mathbf{R}) \\ & \times \exp(-is^j(\mathbf{K})z/2\chi) \\ & \times \exp(-\mu^j(\mathbf{K})z) \quad (25)\end{aligned}$$

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

$$\begin{aligned}\psi(\mathbf{R} - \mathbf{R}_0, z) = & \sum_j \int_{\text{objective aperture}} \epsilon^j(\mathbf{K}) b^j(\mathbf{K}, \mathbf{R}) \\ & \times \exp(-is^j(\mathbf{K})z/2\chi) \exp(-\mu^j(\mathbf{K})z) \\ & \times \exp\{i[\mathbf{K} \cdot (\mathbf{R} - \mathbf{R}_0) + \gamma(\mathbf{K})]\} d\mathbf{K} \quad (26)\end{aligned}$$

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,

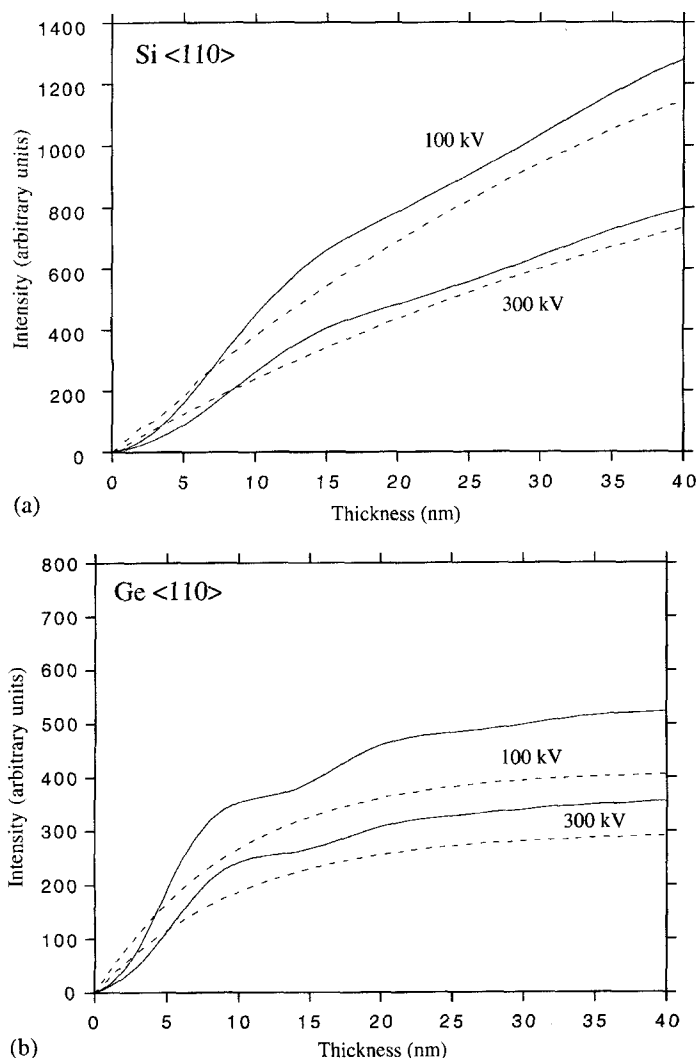
$$\begin{aligned} O(\mathbf{R}, t) &= \epsilon^{1s^2}(0) b^{1s^2}(\mathbf{R}, 0) \\ &\times \left( \frac{1 - \exp(-2\mu^{1s}(0)t)}{2\mu^{1s}(0)} \right) O^B(\mathbf{R}) \end{aligned} \quad (29)$$

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 <110> and (b) germanium <110> 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

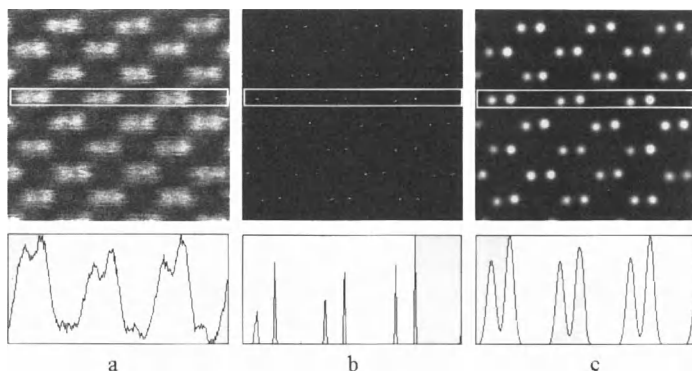
One of the alluring properties of a 300 kV scanning transmission electron microscope is a demonstrated probe size of 1.26 Å, less than the separation of the ‘dumbbells’ seen in the  $\langle 110 \rangle$  projection of diamond cubic semiconductors. In compound semiconductors, the two columns of the dumbbell 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

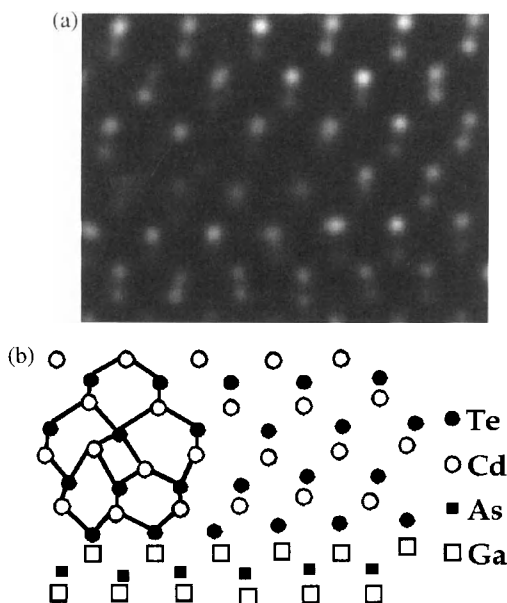
**Figure 14.** Images of GaAs  $\langle 110 \rangle$  in the HB603U 300 kV scanning transmission electron microscope. (a) Raw image, with arsenic columns showing the expected  $\sim 10\%$  higher scattering power, (b) a maximum entropy reconstruction of the object, and (c) a reconstructed image. Line 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.

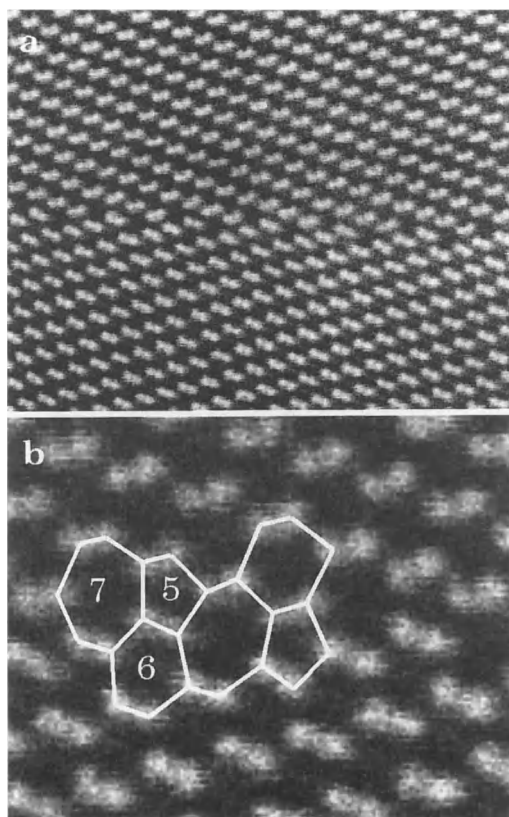
The greatest value of this direct inversion capability is that it can reveal unexpected features of the object that would not have been included in any model structure. Figure 15 shows an image of a perfect edge dislocation core at a CdTe(001)–GaAs(001) interface reconstructed in this way. It is clearly 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.

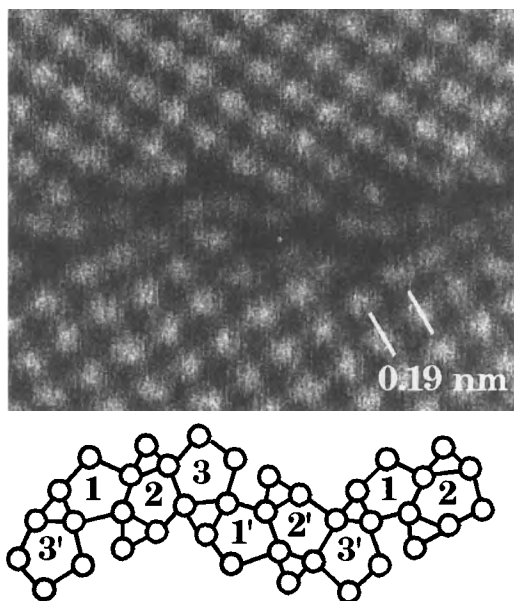
parallel to the  $\langle 110 \rangle$  tilt axis. This is seen clearly to comprise the five- and seven-membered ring structure first shown by Krivanek et al. [45]. In contrast, the symmetric  $\Sigma = 13, \{150\}\langle 001 \rangle$ , tilt boundary in silicon (Fig. 17) does show an unexpected arrangement 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.

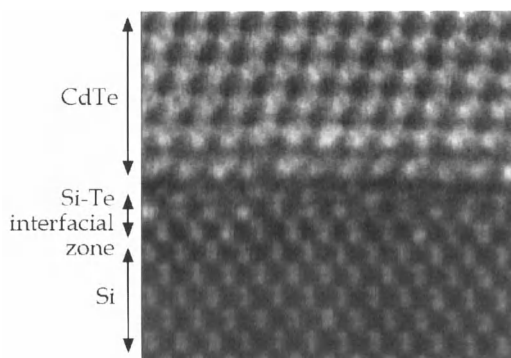
boundary containing multiple structures reported previously [47], and is not the structure predicted from the principle of continuity of boundary structure [48].

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/Si(100) interface in Fig. 18. Growth by MBE, the image has been high-pass filtered to remove the high background

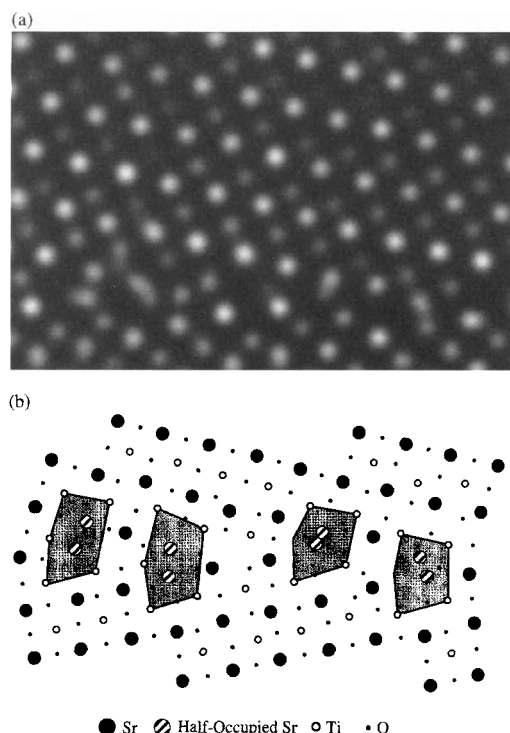


**Figure 18.** Complex atomic structure at  $(111)\text{-Si}(100)$  interface grown by MBE. The interface ends on the tellurium sublattice, but EELS shows the presence of tellurium diffused in the monolayers of the silicon substrate. Tellurium columns are seen bright.

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  $\Sigma = 85, \{920\}\{001\}$  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.

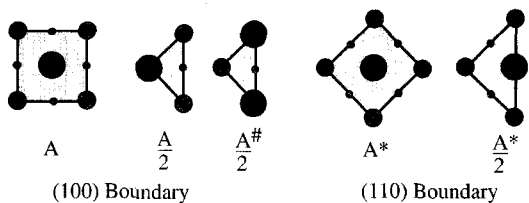
Such sites give insight into how impurities may segregate into the boundary plane; they have significant excess volume and could be attractive sites for high atomic radius impurity species.

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].

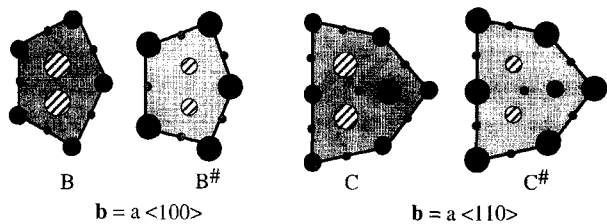
Examining  $36^\circ$  and  $67^\circ$  symmetric tilt boundaries revealed further structural units, again all characterized by the presence of half-filled columns. For these

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

Unit Cells and Fragments: Spacer Units



Pentagons: Dislocation Cores



**Figure 20.** Structural units for all symmetric and asymmetric  $\langle 001 \rangle$  tilt grain boundaries in  $\text{SrTiO}_3$ .

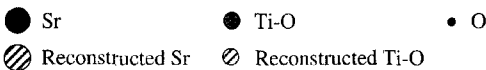


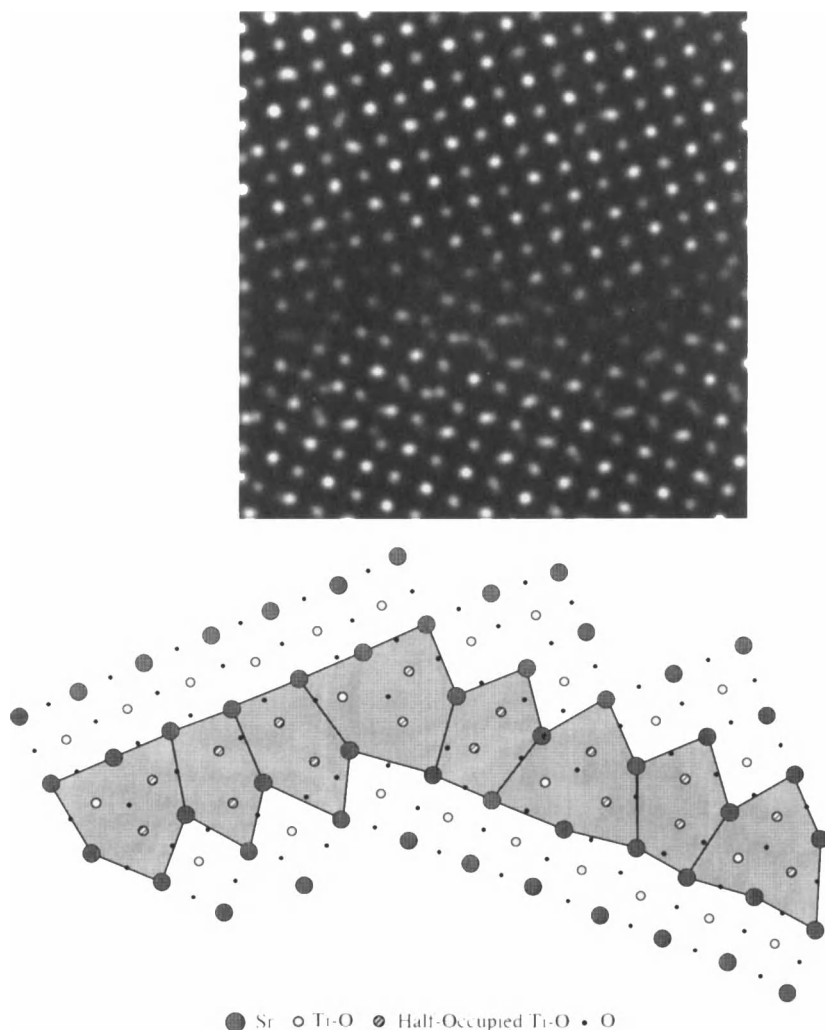
Fig. 20). Thus, only a single type of lattice dislocation, in these two variants, is needed to account for all symmetric boundaries. In this description, continuity of boundary structure is no longer maintained but the number of units required is reduced.

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

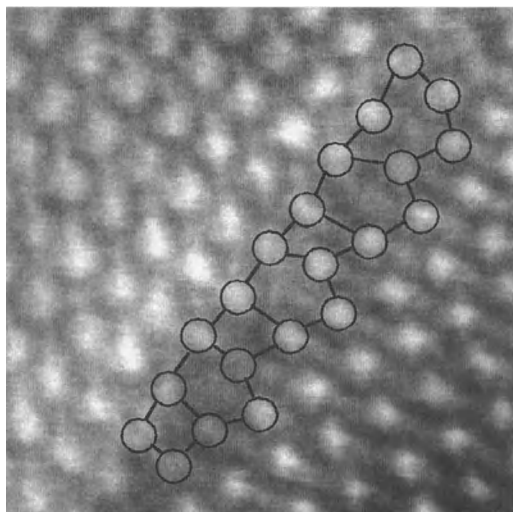


**Figure 21.** Reconstructed 100 kV image of a macroscopically symmetric  $45^\circ$  tilt boundary in  $\text{SrTiO}_3$ : decomposition into facets, with deduced atomic structure.

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.

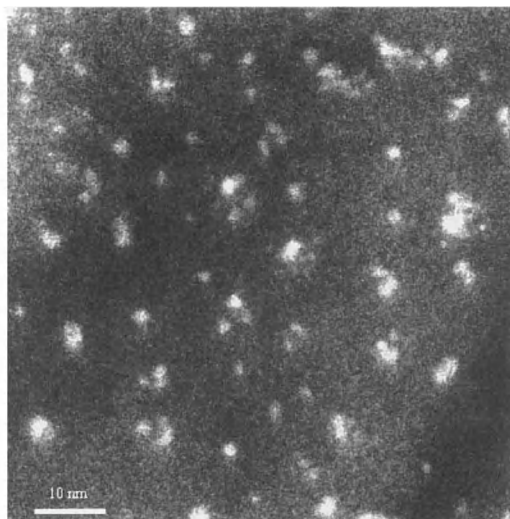
### 2.3.6.3 Nanocrystalline Materials

Of increasing technological importance, nanocrystalline materials represent another area where Z-contrast microscopy offers major advantages. It is the combination of freedom from coherent interference effects, such as Fresnel fringes and speckle pattern from an amorphous



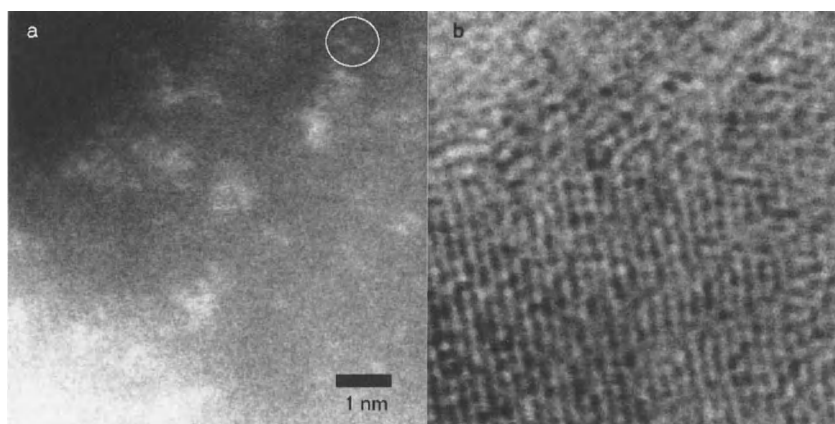
**Figure 22.** Z-contrast image (300 kV) of a predominantly asymmetric tilt boundary in  $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ , grown on a symmetric  $30^\circ$   $\text{SrTiO}_3$  bicrystal. Structural units are equivalent to units  $\text{B}^\#$  and  $\text{C}^\#$  for  $\text{SrTiO}_3$  (see Fig. 20).

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



**Figure 23.** Z-contrast images (300 kV) of silicon nanocrystals formed by implantation into  $\text{SiO}_2$  revealing nanometer-sized clusters.

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.



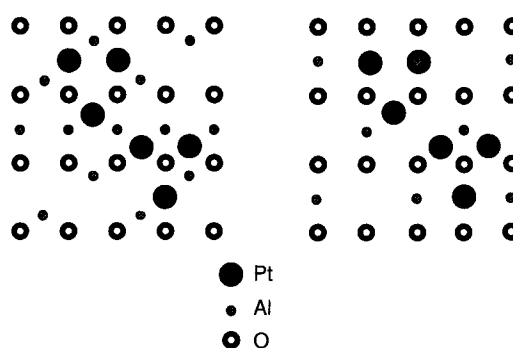
**Figure 24.** Images (300 kV) of a platinum catalyst on a  $\gamma\text{-Al}_2\text{O}_3$  support (a) phase contrast and (b) Z-contrast, obtained simultaneously. The catalyst is seen to comprise of platinum dimers and trimers with some unresolved three-dimensional clusters.



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

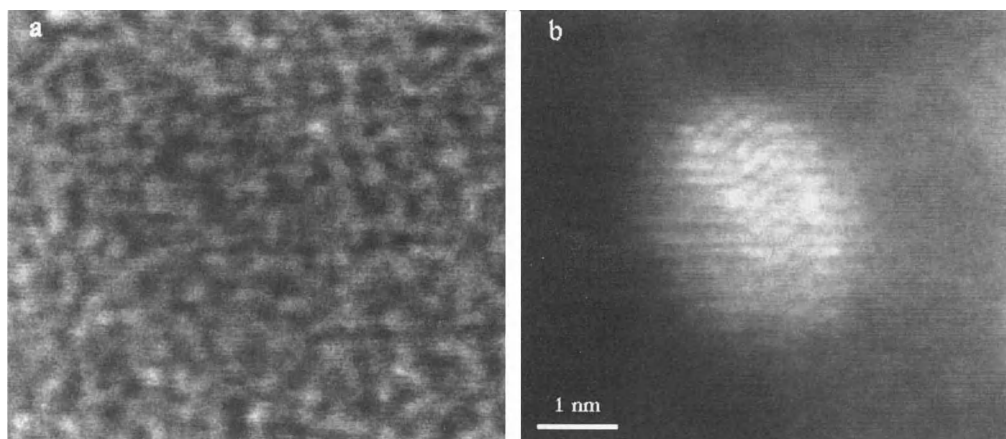
The ultimate example of Z-contrast sensitivity to date is the ability to image single platinum atoms on a  $\gamma\text{-Al}_2\text{O}_3$  support, as seen in Fig. 24. The small bright dots are visible well above the noise level of the substrate, and are arranged as dimers, trimers, and other configurations. 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  $\gamma\text{-Al}_2\text{O}_3$ . 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. Jebasinski, 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):

$$E_{\text{kin}}(jkl; X) = E_b(j) - E_b(k) - E_b(l) - F(kl; X) + R_c \quad (1)$$

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]:

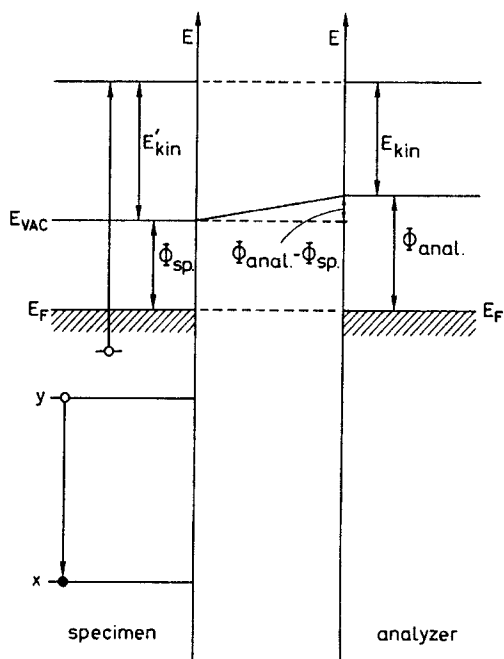
$$E_{\text{kin}} = E_K(Z) - \frac{1}{2} [E_{L_2}(Z) + E_{L_2}(Z+1)] - \frac{1}{2} [E_{L_3}(Z) + E_{L_3}(Z+1)] \quad (3)$$

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  $\text{Ar}^+$ -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  $\text{Ar}^+$  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

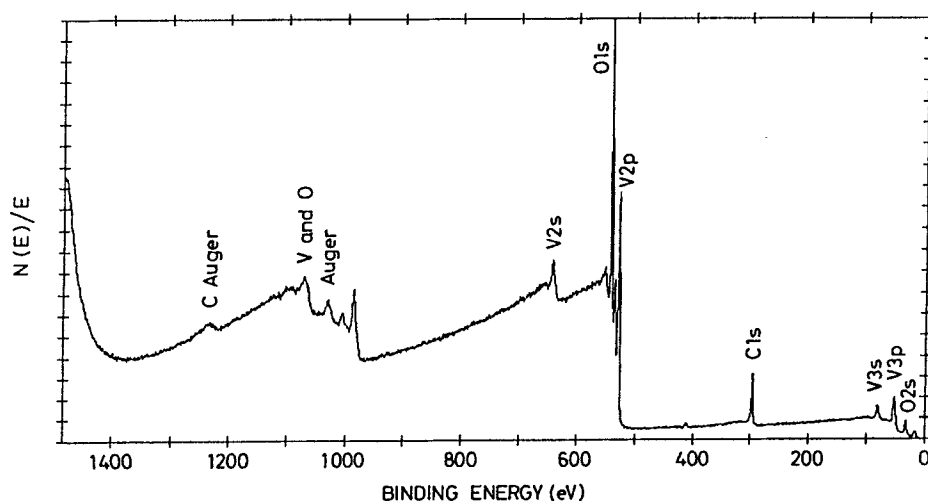
characterized by completely filled shells, while after photoemission one electron hole is present. One consequently only expects the corresponding spin-orbit splitting spectroscopic terms, explaining the notation used above. If, however, the electron hole thus created can couple with other holes already 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 eV) 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

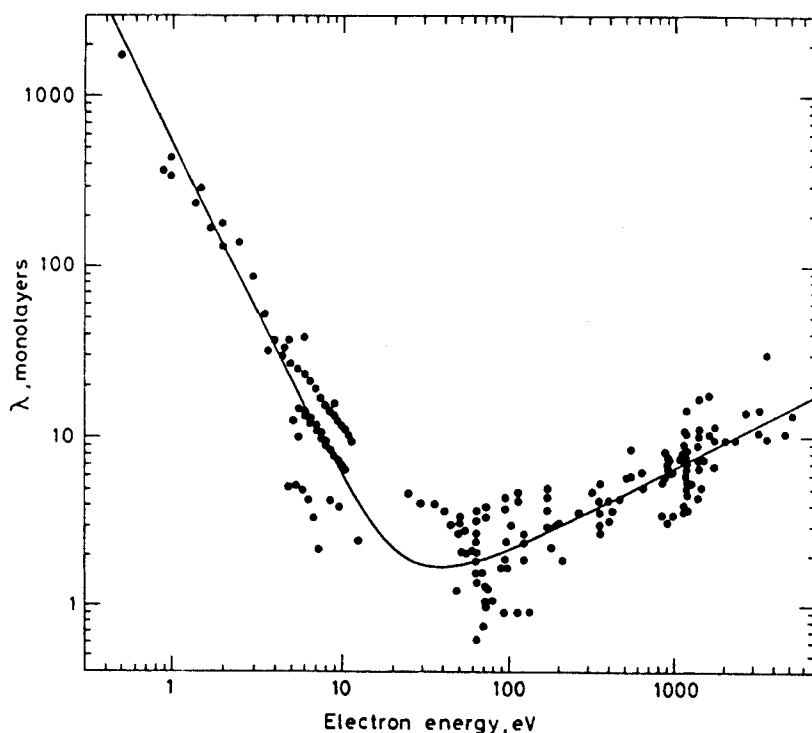
[7] (and using the notation of Nebesny et al. [6]):

$$I_A = I_0 \times \sigma_{A(E_p)} \times (1 + r_{M(E_A, \alpha)}) \times T_{(E_A)} D_{(E_A)} \times \int_0^\infty N_A(z) \times \exp[-z/\lambda_{M,E} \cos \theta] dz \quad (4)$$

where  $I_0$  is the primary electron beam current ( $A\,cm^{-2}$ ),  $\sigma_{A(E_p)}$  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:

$$C_X = \frac{I_X}{I_{Ag} S_X D_X} \quad (8)$$

where  $C_X$  is the atomic concentration of element X,  $I_X$  is normally the peak-to-peak amplitude of the element X,  $I_{Ag}$  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

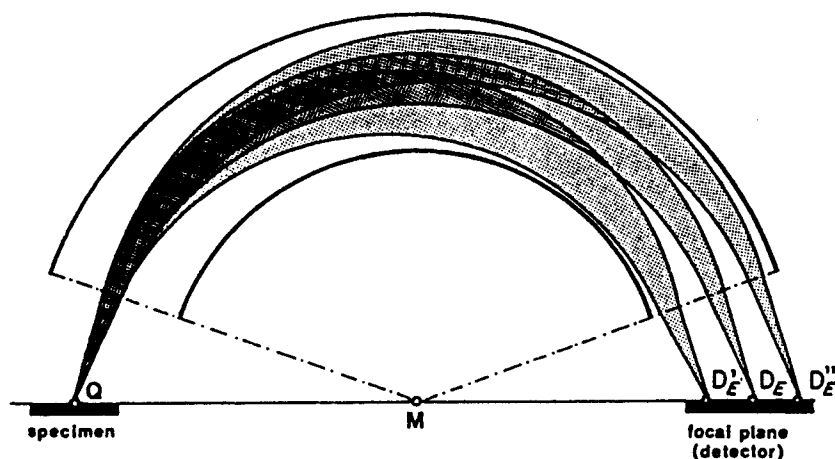
### 2.4.3.1 Basic Principles of Imaging

In microscopic imaging techniques, just as in classical optical systems, the object or

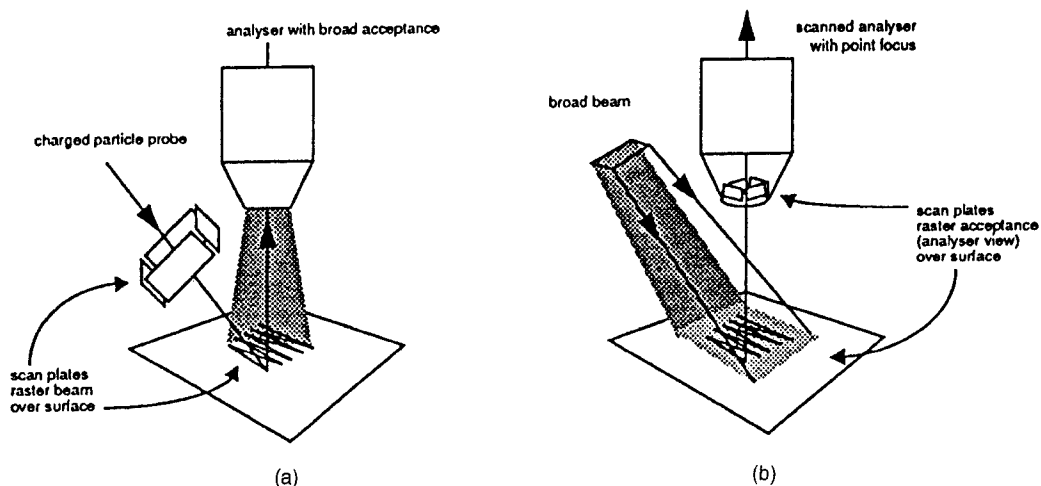
sample to be viewed is illuminated by a source, broad enough to illuminate the whole sample at once. The reflected light or emitted beam from the sample is then passed through a magnifying imaging device (cf. an optical lens) and hence a magnified image of an object or sample is obtained in which all 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.

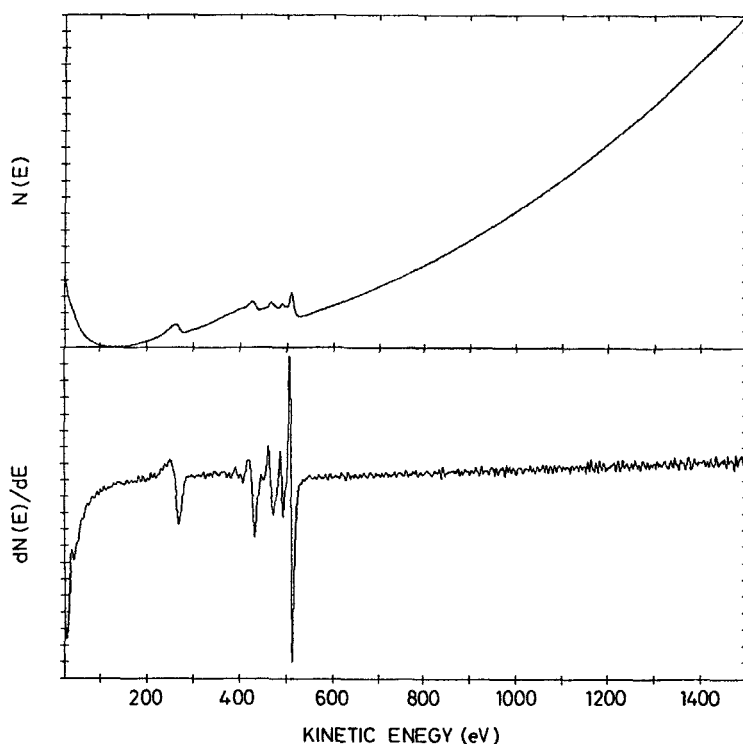
By modulating the retarding potential  $V_0$  with an alternating signal  $\Delta V_0$  and by synchronously detecting the AC part of the collected current, one effectively measures the first derivative of  $I$  with respect to  $E_0$  [12], that is,

$$\frac{dI(E_0)}{dE_0} \sim N(E_0) \quad (10)$$

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

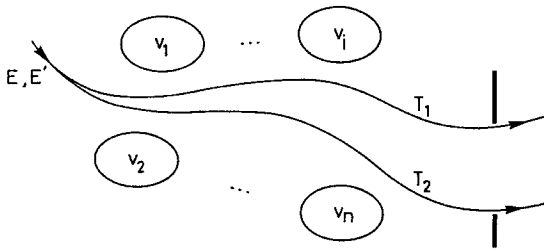
only intervene in the transit time of the particle along the followed ray path. Clearly, this property is of value in the energy analysis of ions having different masses and (multiple) charges; only the ratio of electrode potentials to the accelerating potential of the particles remains constant [18].

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

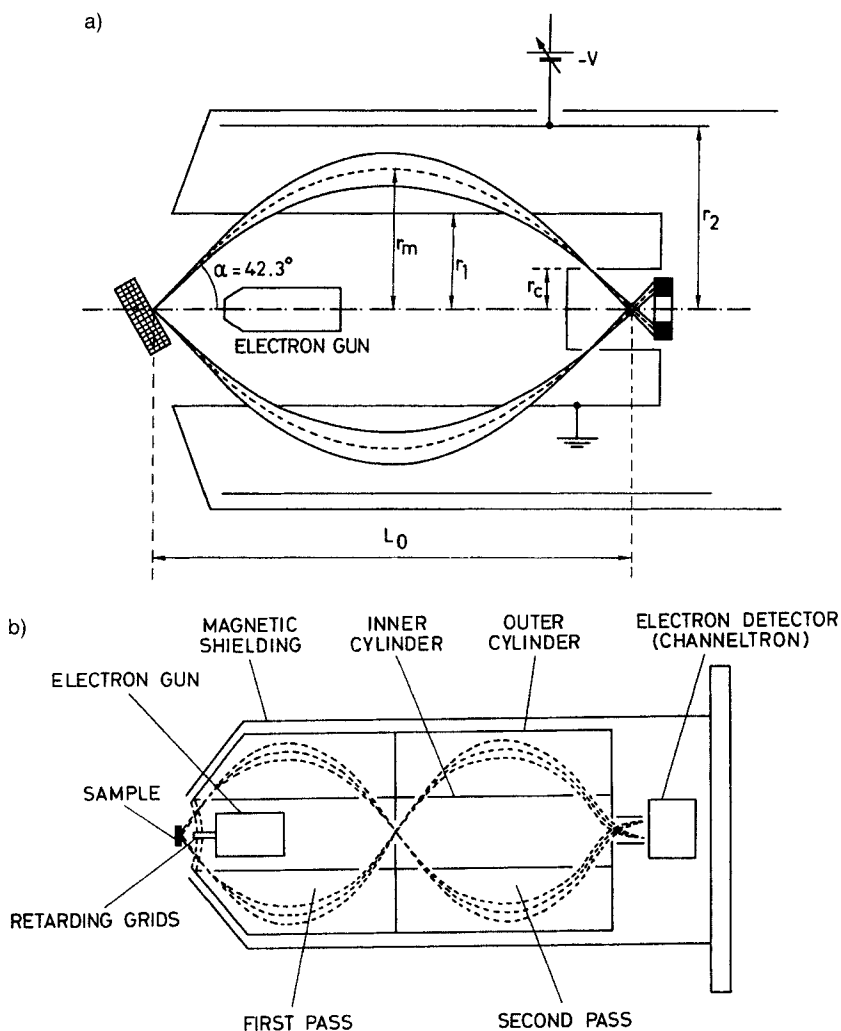
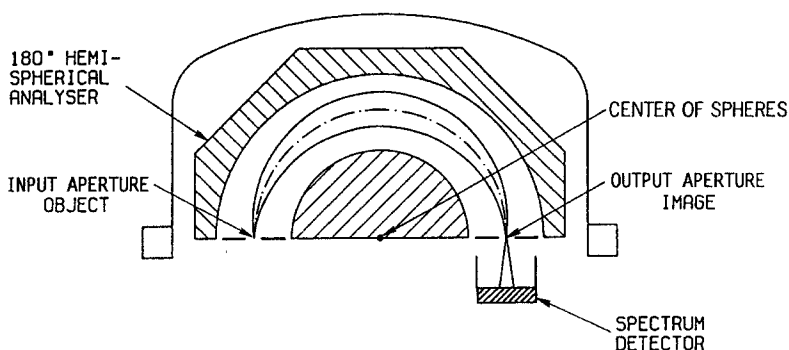


Figure 8. (a) Single-pass CMA; (b) double-pass CMA.



**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.

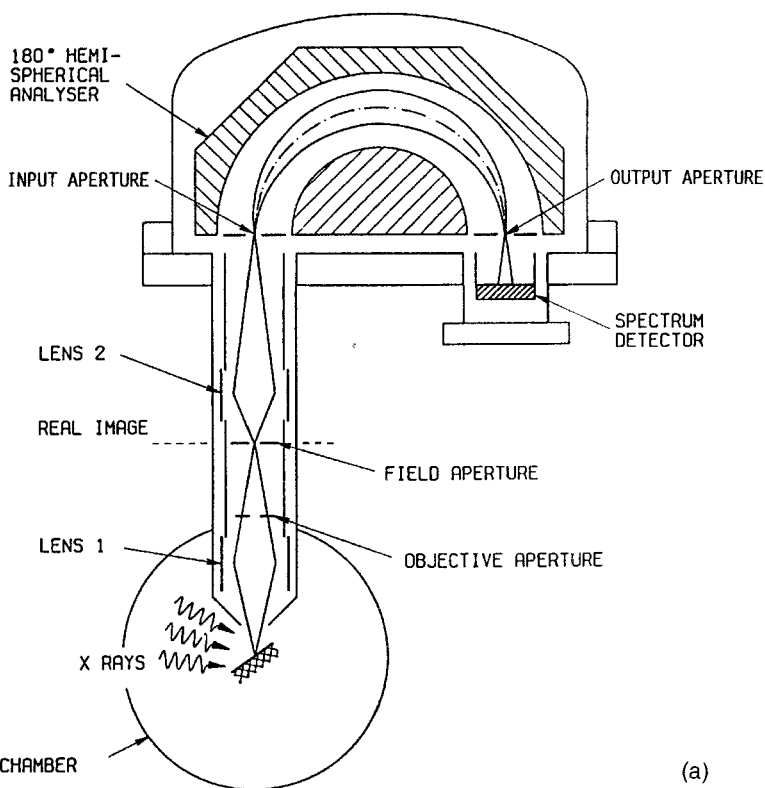
In modern equipment for AES and XPS the concept of parallel detection is fully implemented for many applications. As in a CHA, the energy dispersive direction coincides with the produced direction of the input aperture and the center of the deflecting spheres. There is also good accessibility for mounting 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

**PAGE MISSING OR DUPLPLICATE**





**Figure 10.** (a) Combination of transfer lens and CHA defining the field of view and CHAMBER the acceptance angle.

(a)

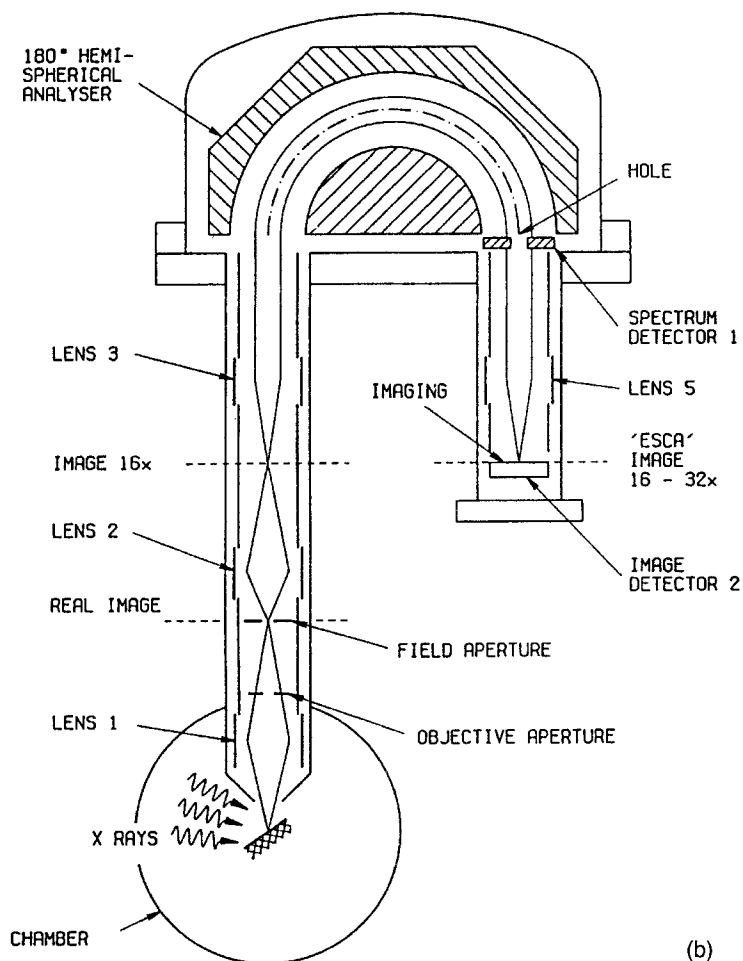
kind of magnetic lens close to the sample with the  $\mathbf{B}$  vector pointing towards the transfer lens.

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 CHA. This lens provides extra magnification and also performs the decelerating function (without the use of grids) if the CHA 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 CHA vastly enhances the performance of the CHA 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.

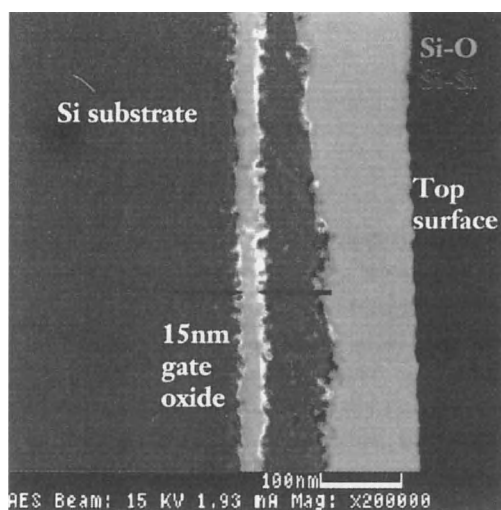
In most commercial SAM instruments a secondary electron detector (SED) is included to produce SEM images of the surface under study. Furthermore, absorption current images can also be obtained. Both are very useful in selecting the surface area which has to be studied with the SAM.

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  $\text{LaB}_6$  emitters has proven to be a leap forward in brightness as compared with the classic thermionic emitters. However,  $\text{LaB}_6$  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  $\text{LaB}_6$  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  $\text{LaB}_6$  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  $\text{LaB}_6$  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  $\text{LaB}_6$  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  $\text{SiO}_2$  gate oxide layer, 200 nm below



**Figure 11.** A 15-nm wide  $\text{SiO}_2$  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.

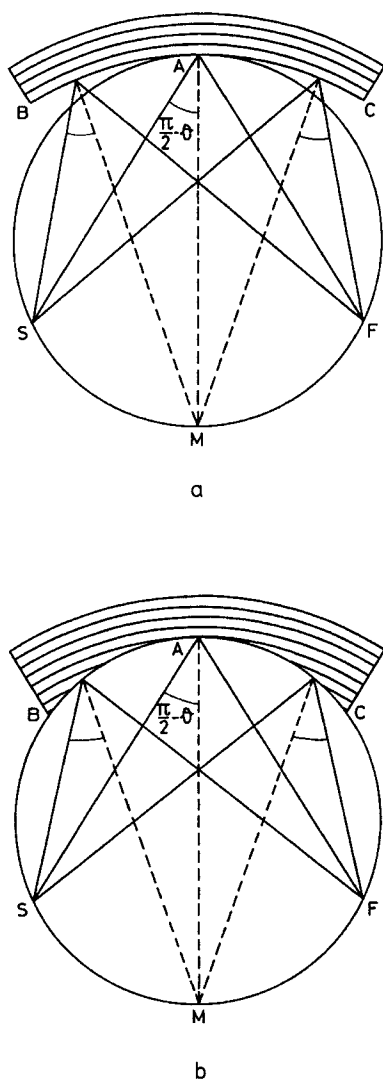
With the simple cathode–anticathode arrangement the emerging X-ray beam is usually very broad and so spatial resolution can only be obtained by collimating the beam. This is done at the expense of intensity. Besides intensity, there remain several other drawbacks and limitations for XPS work using a 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



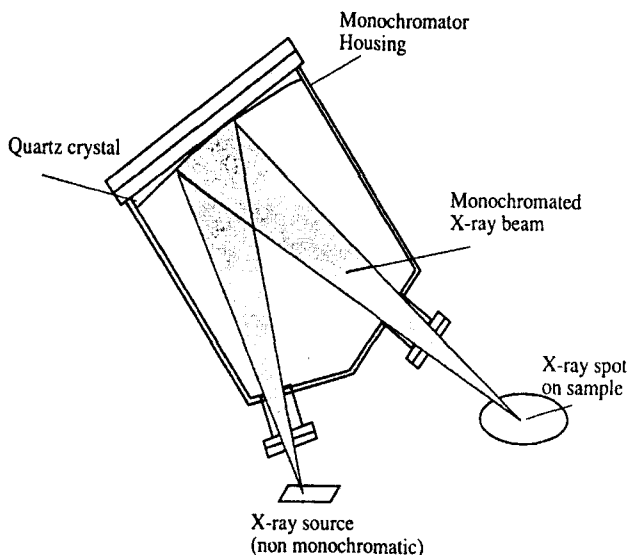
**Figure 12.** (a) The Johann solution: the bent crystal has a radius equal to the diameter of the Rowland circle; (b) the Johansson or Guinier solution.

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.

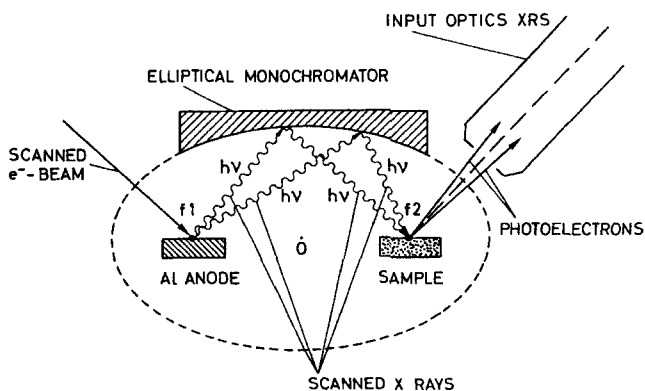
A breakthrough in the pixeling process has been made by Physical Electronics Inc. using elliptically bent crystals. The principle underlying this approach is shown in Fig. 14. At the foci of the ellipse, which defines the bent crystals lattice surface, are the anticathode and sample. By scanning the 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



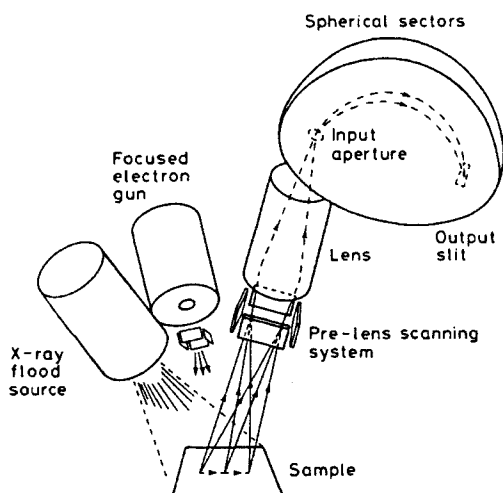
**Figure 13.** Practical focusing X-ray monochromator of the Johansson or Guinier configuration, ultrahigh vacuum adapted for XPS work. Adapted from Kratos Analytical, Axis product brief. Reproduced by permission of Kratos Analytical Ltd.

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.

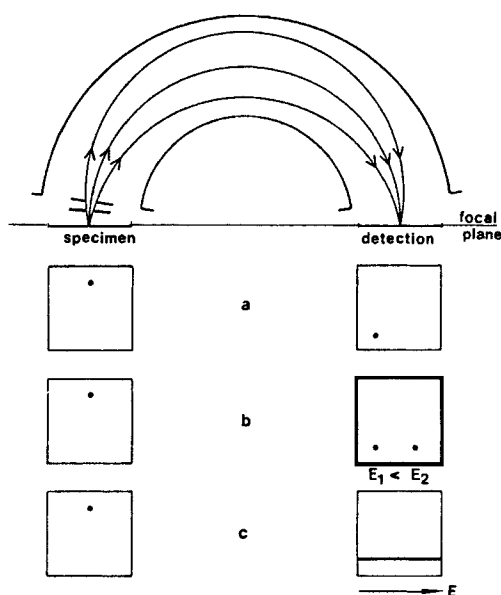


**Figure 15.** In the detection defined XPS imaging, pixeling can be obtained by the use of a pre-lens scanning system. Adapted from M. P. Seah and G. C. Smith, *Surf. Interface Anal.* **1988**, *11*, 69 (Fig. 3). Reproduced by permission of M. P. Seah and John Wiley & Sons Ltd.

moving the sample stage, a time-consuming and trivial solution to detection defined XPS imaging can be attained.

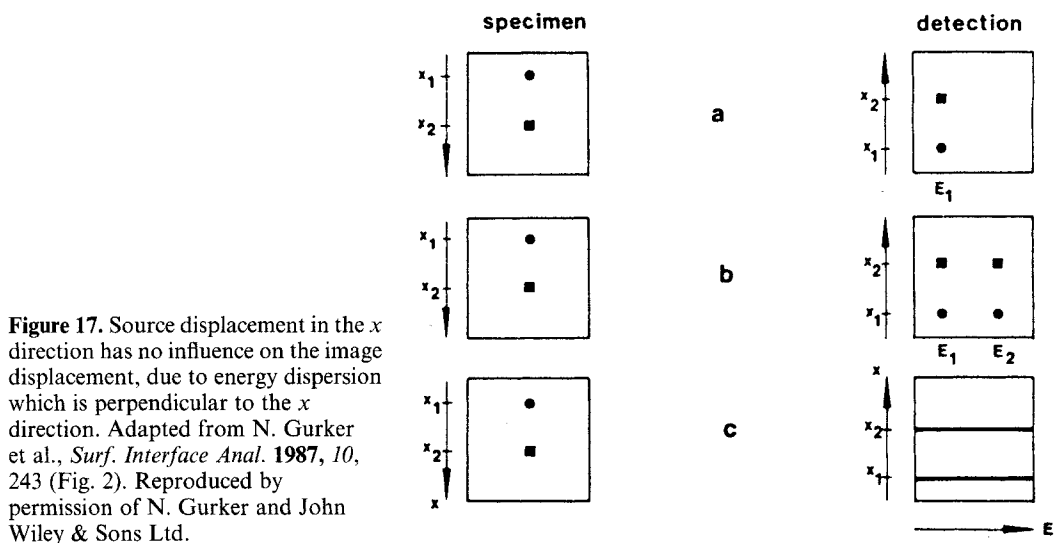
A second logical step in DD XPS imaging has already been mentioned. Namely, this approach entails pixeling by using scanning deflection plates (Fig. 15) in front of the transfer lens of the CHA [55]. The transfer lens defines the field of view or the lateral resolution, whereas the scan plates provide 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.

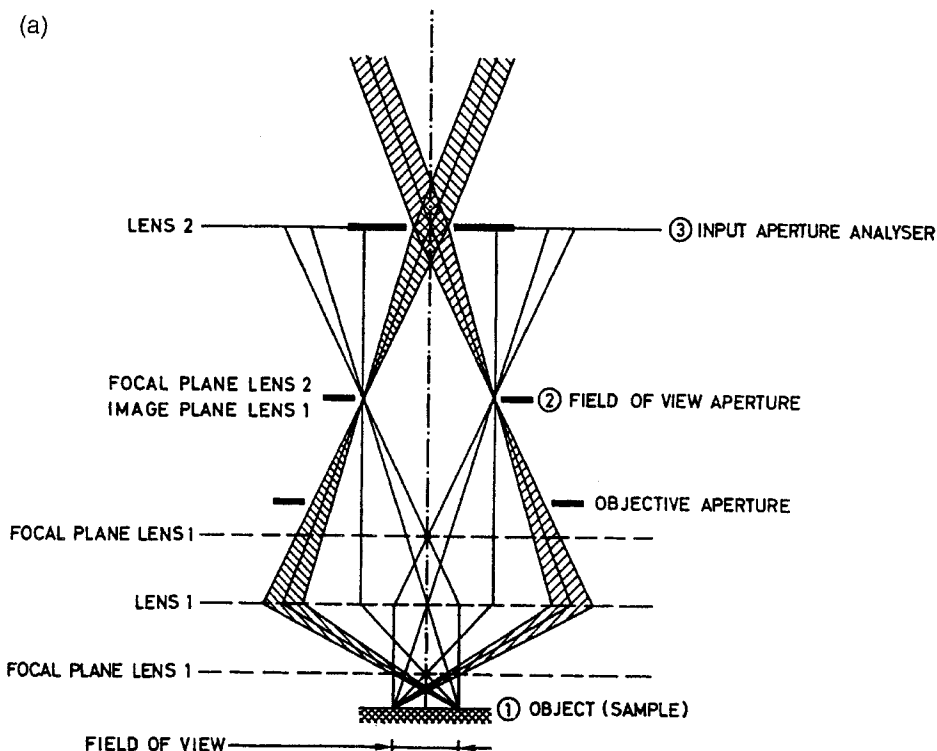
Recently, VG Scientific Ltd have pioneered a breakthrough on the side of DD

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

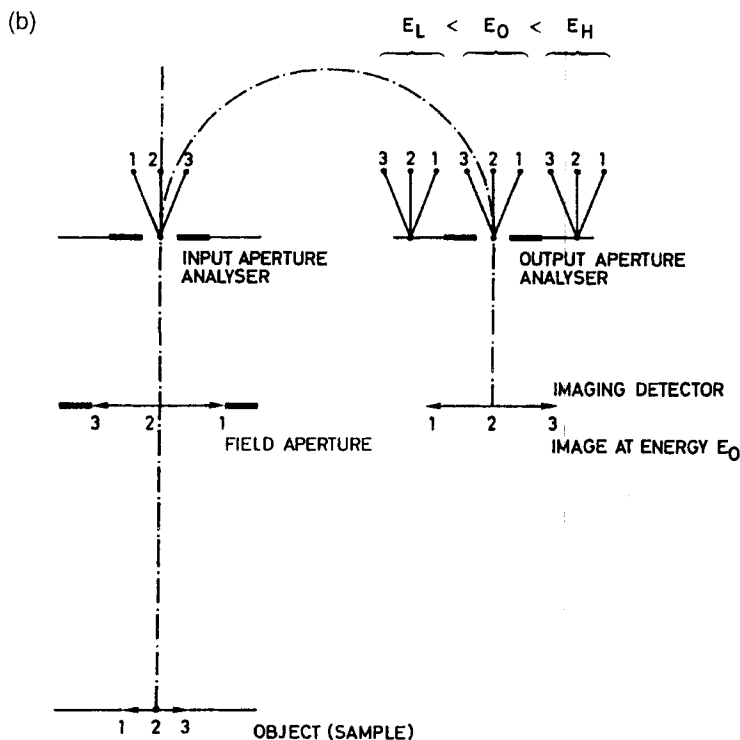
$E_0$ , giving the required energy dispersion. The energy resolution can be changed by changing the size of output and input aperture but, as discussed above, this has no influence on the image information, only on the intensity which translates itself into acquisition time of the image.

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 CHA.

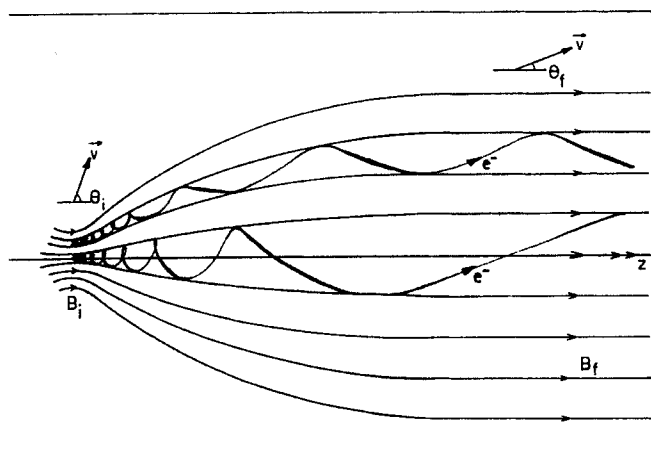
the influence of a Lorentz force. If the change in magnetic field strength during one such spiraling motion is compared with the average magnetic field, the total energy and angular momentum are



**Figure 19.** 5 μm wide W bars, with 5 μm 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:

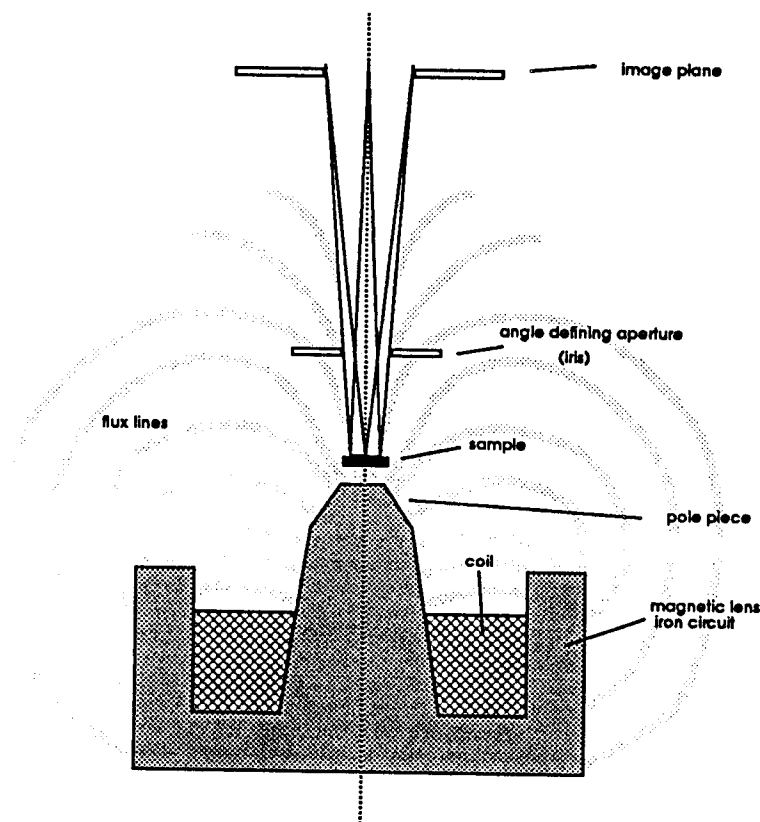


Figure 21. Magnetic objective lens as used in XPS. Adapted from Kratos Axis 165 System Overview. Reproduced by permission of Kratos Analytical Ltd.

- (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

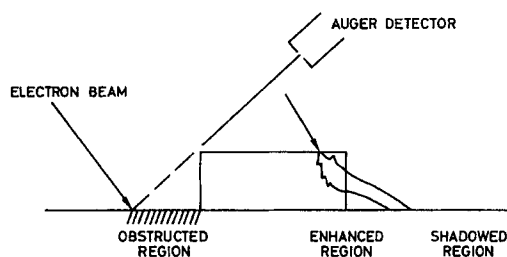
Again, Prutton and his colleagues describe how the acquisition of both a secondary electron image and a sample absorption current image is necessary to correct for beam current fluctuations. Both images are anticorrelated, a feature exploited to form a beam current image, which is then divided into any 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. 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*, 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

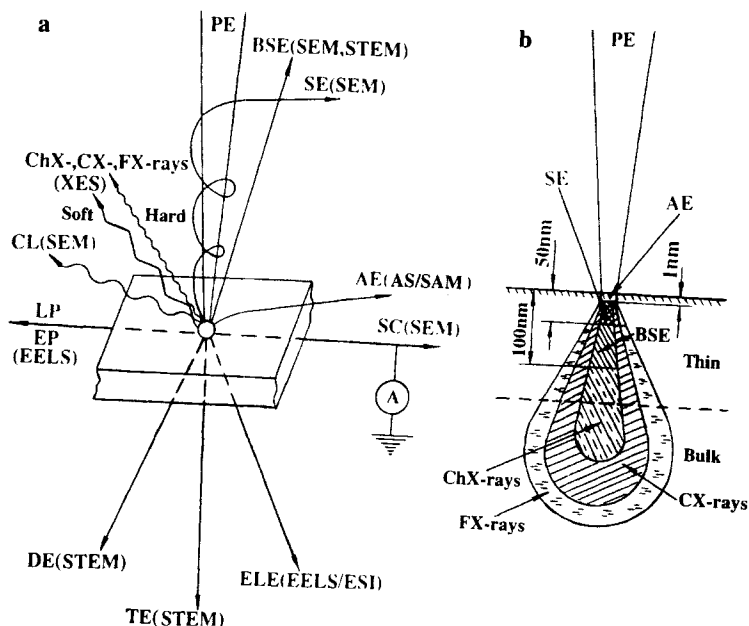
Classical electron probe microanalysis (EPMA) is carried out by determining the intensity of characteristic X-ray emission from an element of interest in the specimen and comparing it with that from a standard of known composition. Maintaining the same analysis conditions enables an X-ray intensity ratio 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 ( $0 < h\nu < 1\text{--}3 \text{ eV}$ ); 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 ( $10^{-3}\text{--}10^{-1} \text{ sr}$ ). Moreover, inner-shell ionizations result in the emission not only of X-rays but also of Auger electrons, and the X-ray fluorescence yield

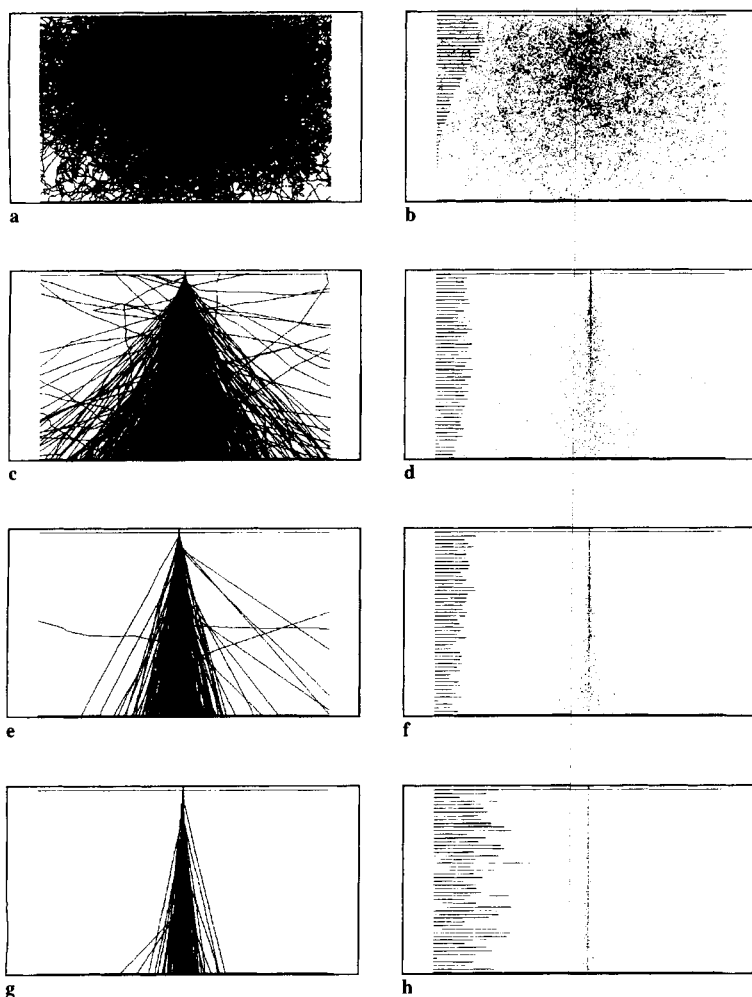


**Figure 1.** Main processes of electron beam interactions and signals detected in different operating and analytical modes of an analytical electron microscope (a) and schematic arrangement of their information volumes (b); PE, primary electrons; TE, transmitted electrons; DE, diffracted electrons; 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

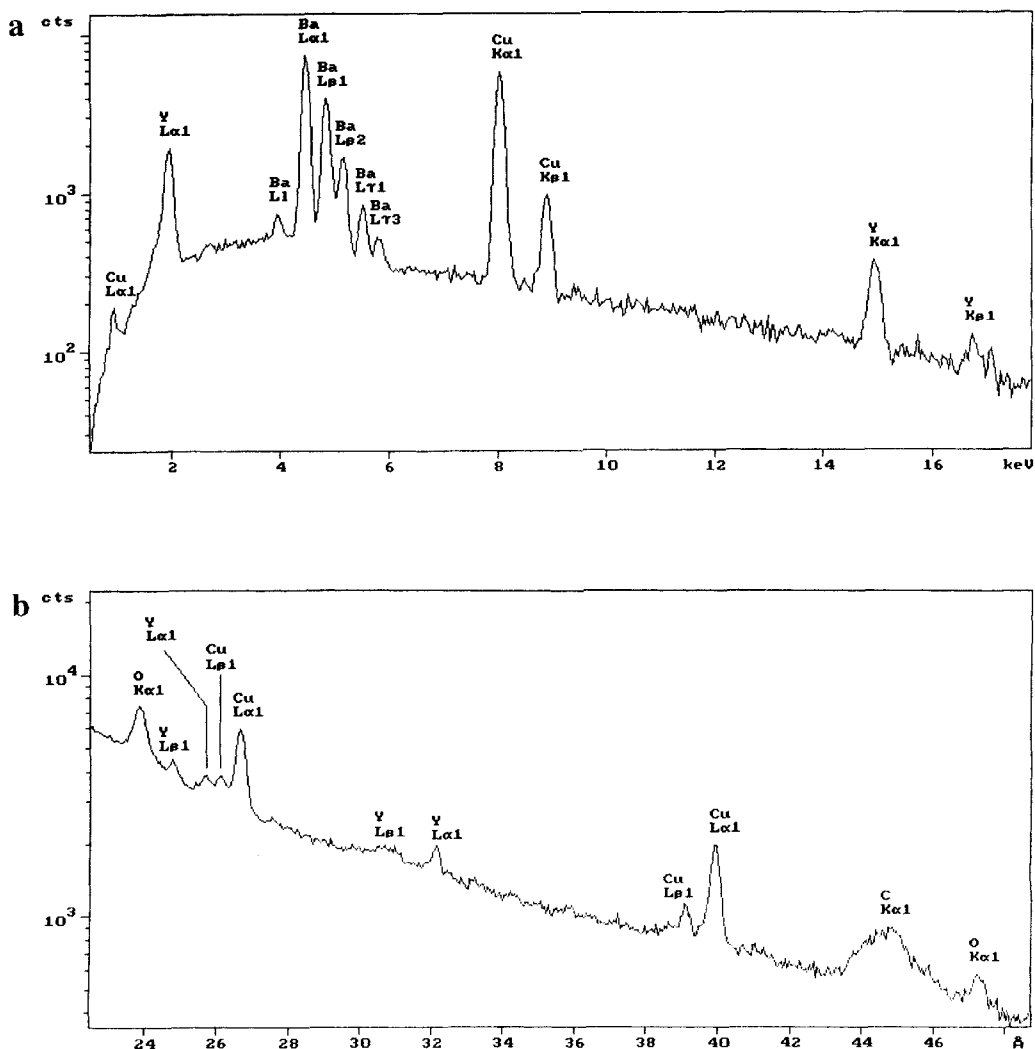


**Figure 2.** Single scattering Monte Carlo simulations of a lateral electron distribution (a, c, e, g) and  $\text{Ag } L_{\alpha}$  X-ray generation (b, d, f, h) in an AgBr tabular crystal of 100-nm thickness with a point electron source at 5 keV (a, b), 30 keV (c, d), 100 keV (e, f), and 300 keV (g, h) for 1000 trajectories. A histogram of the normalized yield of  $\text{Ag } L_{\alpha}$  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

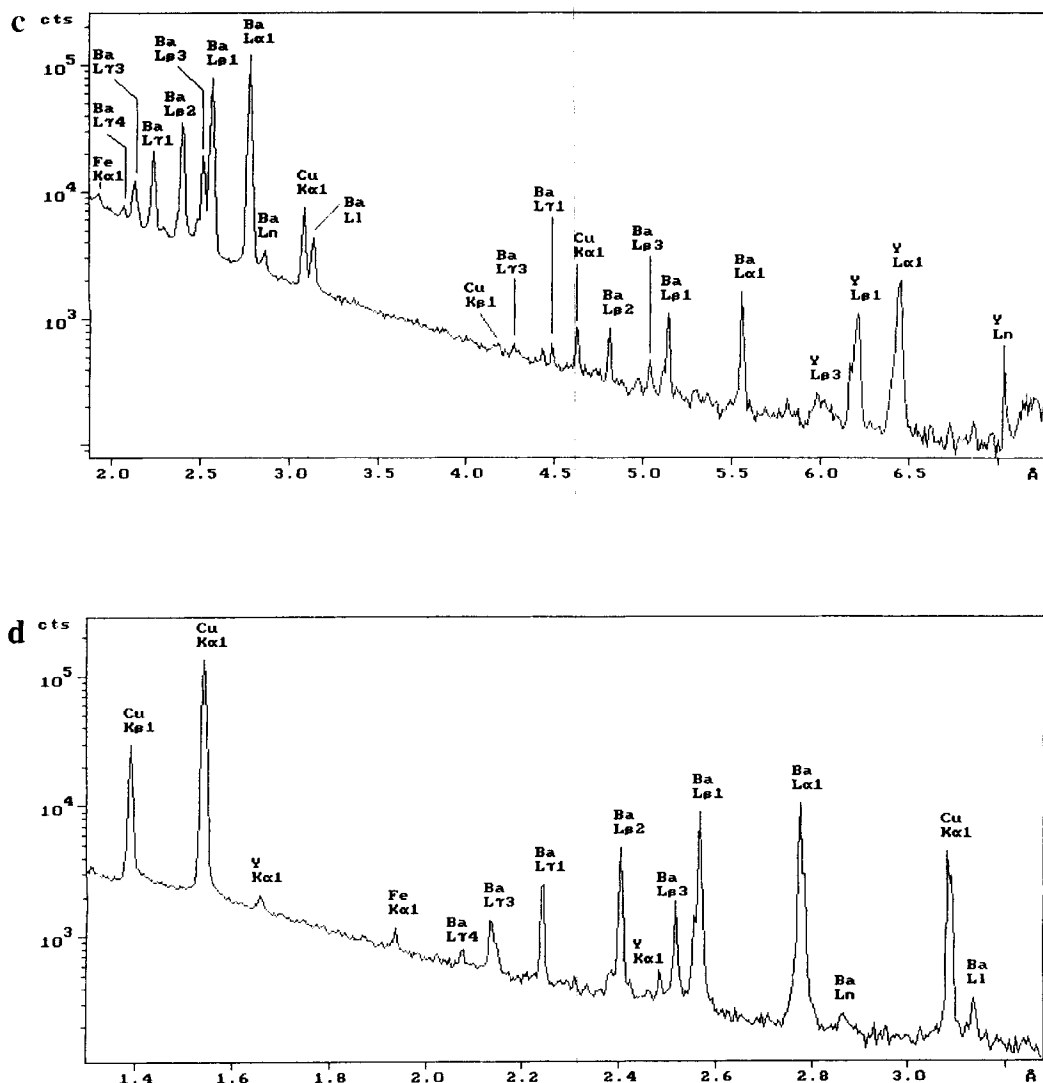


Figure 3 Continued

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  $\text{CK}_\alpha$  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 ( $\text{FeK}_\alpha$  at 0.194 nm; 6.398 keV) and Pt ( $\text{PtL}_\alpha$  at 0.131 nm; 9.441 keV), probably from crucible material.

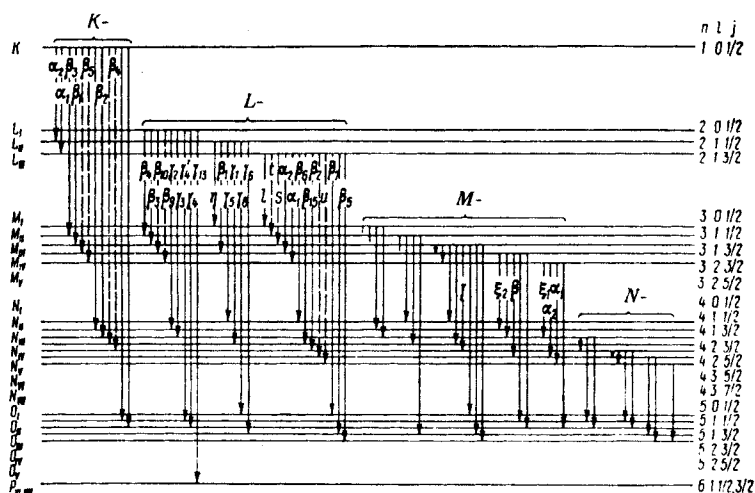
### 2.5.1.3 Characteristic X-Ray Spectra

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:

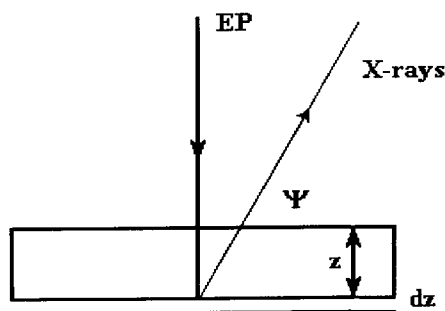
$$E = A(Z - 1)^2 \quad (1)$$

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

$$I = \phi(\Delta\rho z) \int_0^\infty \phi(\rho z) \times \exp(-\chi\rho z) d\rho z f(\chi)(1 + \gamma + \delta) \quad (2)$$

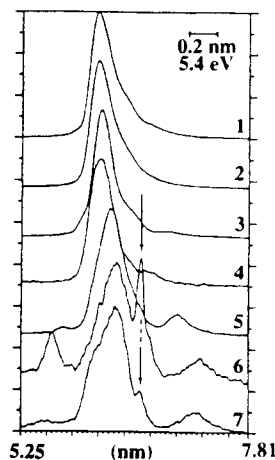
where  $\phi(\Delta\rho z)$  is the emission from an isolated thin film of mass thickness  $\Delta\rho z$ ,  $\chi = (\mu/\rho) \csc \psi$ ,  $\rho z$  is the specimen mass thickness, and  $\mu$  is the linear absorption coefficient. The absorption factor is defined as

$$f(\chi) = \frac{\int_0^\infty \phi(\rho z) \exp(-\chi\rho z) d\rho z}{\int_0^\infty \phi(\rho z) d\rho z} \quad (3)$$

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  $BK_\alpha$  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

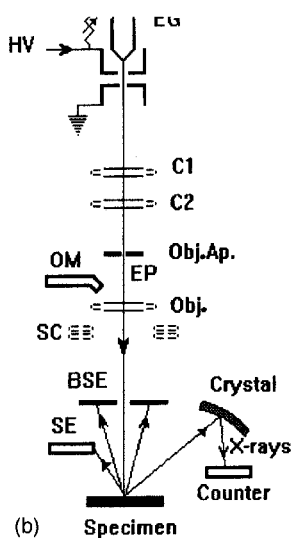
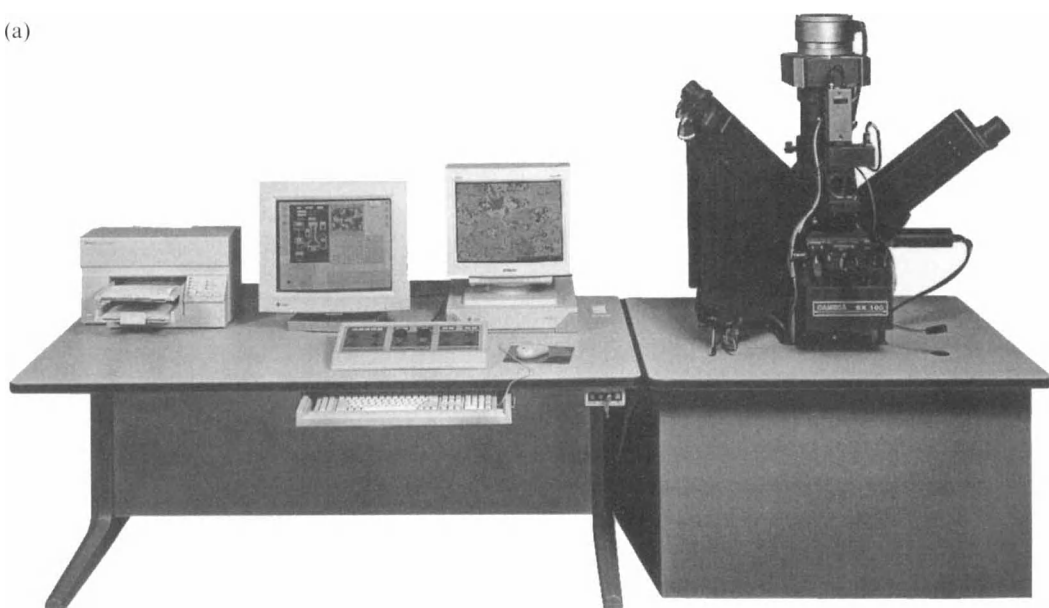
Table 1. Basic characteristics of scanning electron probe methods<sup>a</sup>

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 'compo', topography 'topo', orientation	1.5–15 ( $\times 10-10^4$ )	Elemental (composition images), structural (surface topography images)	100	$\Delta Z < 0.1$	–	
STEM	SC	Atomic number-orientation	3–100	Elemental (composition images), structural images	1000	–	–	
	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)				
	TE+DE Z contrast			Structural (dark-field images)				
SAED/STEM	BSE	Atomic number	3–10	Elemental	$t$ (film)	$\Delta Z < 0.1$	2–3	
	DE	–	1000	Structural (point, ring diffraction patterns)	–	–	(10)	2
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 spectral profiles, 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)	1	3–92	0.1 (10)	2, 21
CL, UV-VIS, SEM IR photons $0 < h\nu < 1\text{--}3 \text{ eV}$	—	—	30–500	Chemical (spectra), structural (spectroscopic images)				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  $1024 \times 768$  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.

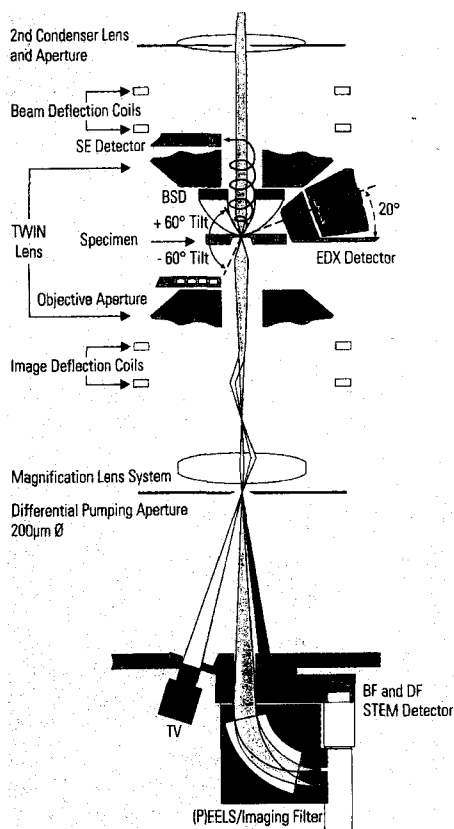
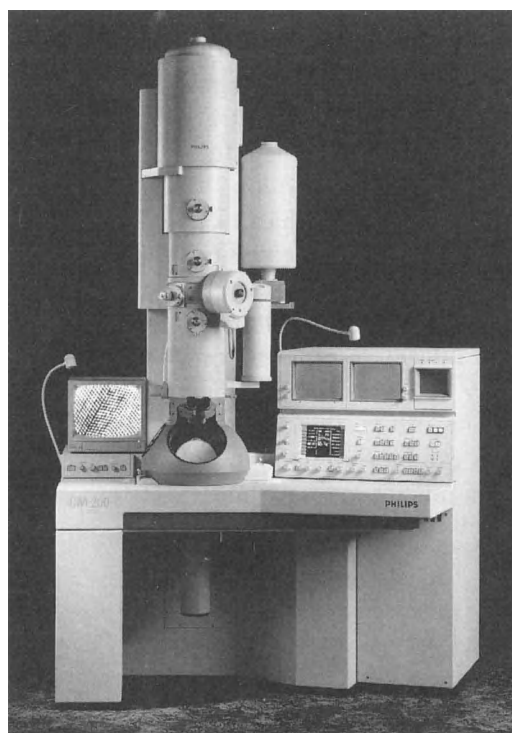
Traditionally, the energy- and wavelength-dispersive systems in an electron probe X-ray microanalyzer or in an analytical scanning electron microscope are operated at a high voltage (15–40 kV). These conditions are sufficient to allow

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



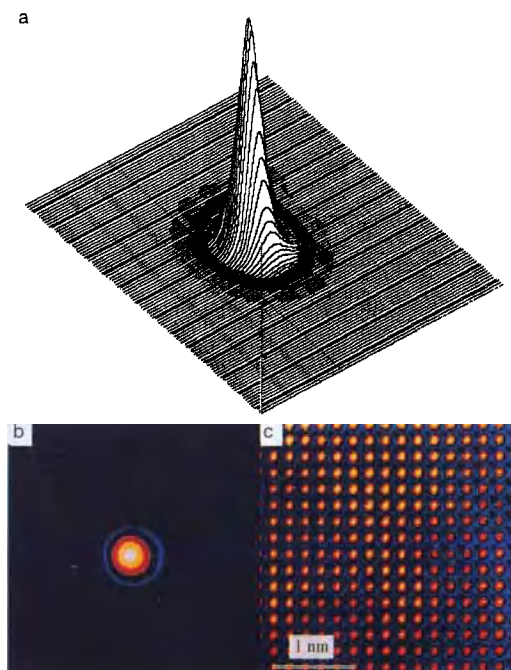


**Figure 8.** General view (a) and schematic diagram of the electron optics (b) of a dedicated analytical scanning transmission electron microscope. On the left-hand side in (b) is the high-resolution TV monitor and the electron optical column with side-entry airlock and windowless electron detector. 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.

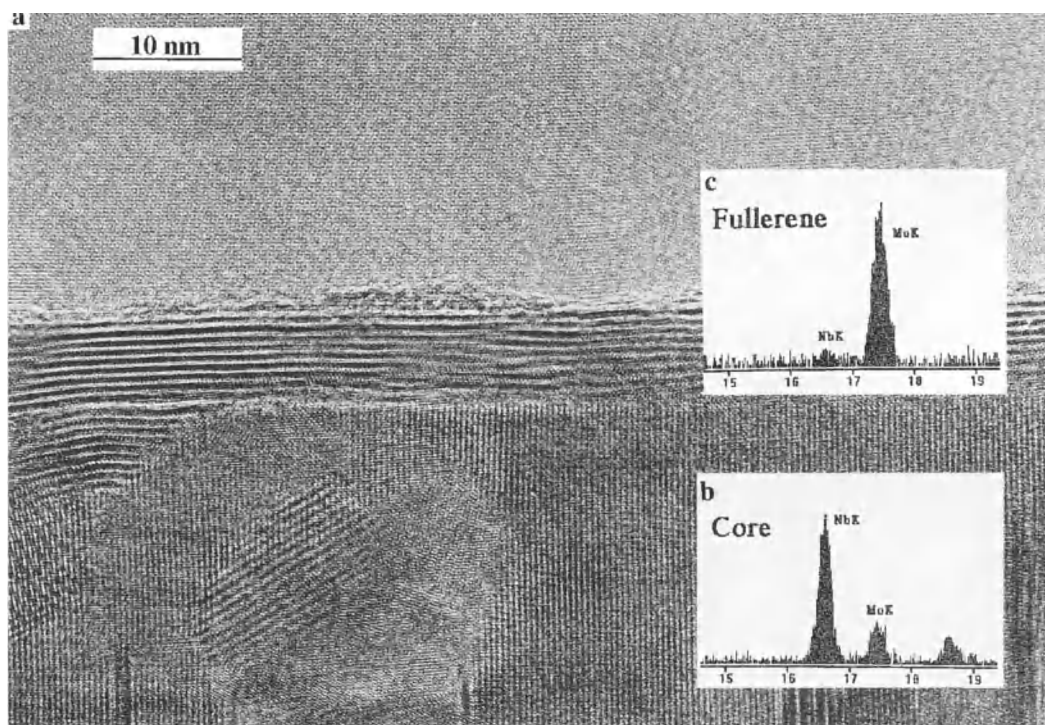
imaged using Z-contrast with a wide-angle annular dark-field semiconductor detector [19]. A higher contrast than in conventional TEM bright- or dark-field modes is achieved in this case.

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.

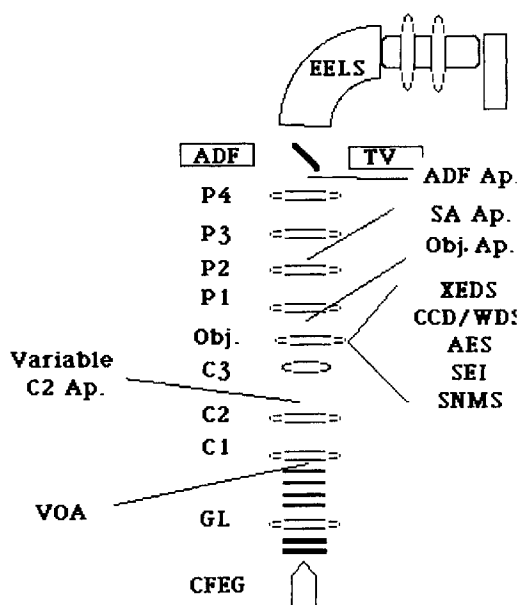
instrument supplied with an energy-dispersive spectrometer [23]. A microanalyzer may also be equipped with a cylindrical mirror Auger-electron spectrometer [21]; this, however, needs an ultrahigh vacuum. Auger-electron microanalyzers, in which the 1–10 keV electron gun is incorporated in the inner cylinder 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].

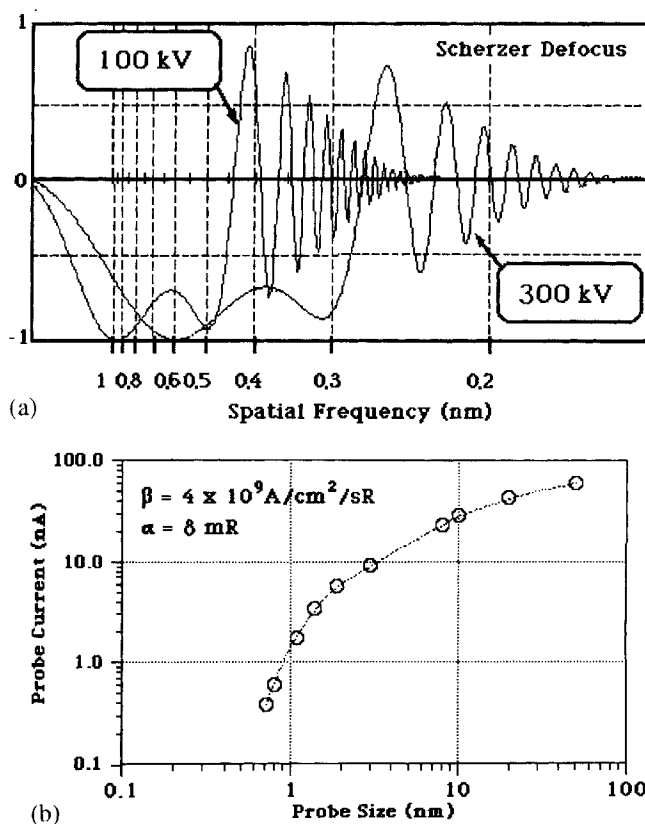
Recently an advanced 300 kV field-emission analytical electron microscope (Argonne National Laboratory) has been designed to attain the best possible analytical sensitivity, resolution, and versatility for EDS, EELS, Auger electron spectroscopy (AES), selected area electron diffraction (SAED), convergent 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  $4 \times 10^9 \text{ Å cm}^{-2} \text{ sr}^{-1}$  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



**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  $60\text{ l s}^{-1}$  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

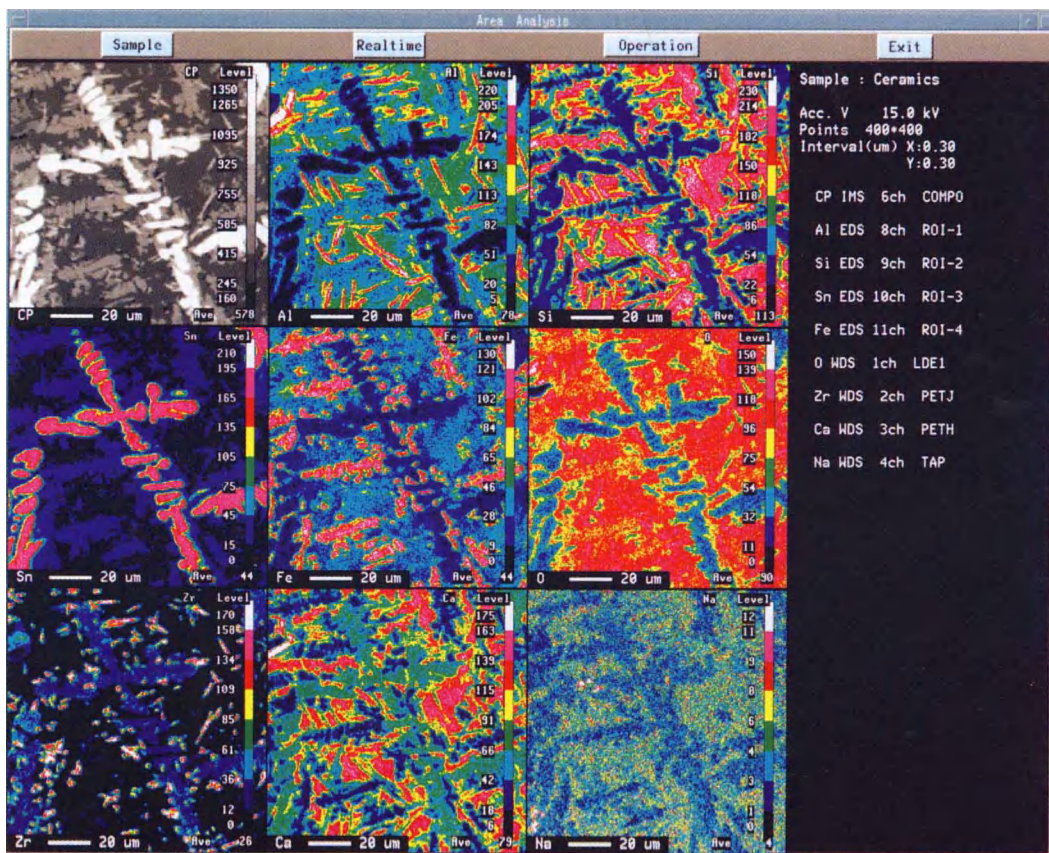
X-ray maps (Fig. 13) provide valuable information on the two-dimensional elemental distributions over bulk or thin samples. X-ray mapping can be undertaken with both EDS and WDS, where

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 a 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:

$$k_z^a = \frac{R_{sp}^a}{R_{st}^a} \frac{\int_{E_c}^{E_0} (Q/S)_{sp}^a dE}{\int_{E_c}^{E_0} (Q/S)_{st}^a dE} \quad (9)$$

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

$$F_i = \frac{\left(1 + \sum_j I_{i,j}^f / I_i\right)_{sp}}{\left(1 + \sum_j I_{i,j}^f / I_i\right)_{st}} \quad (14)$$

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

$$F_i = \left(1 + \frac{\sum_j I_{i,j}^f}{I_i}\right)_{sp} \quad (15)$$

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  $\bar{Z}_{sp} \neq \bar{Z}_{st}$ , 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

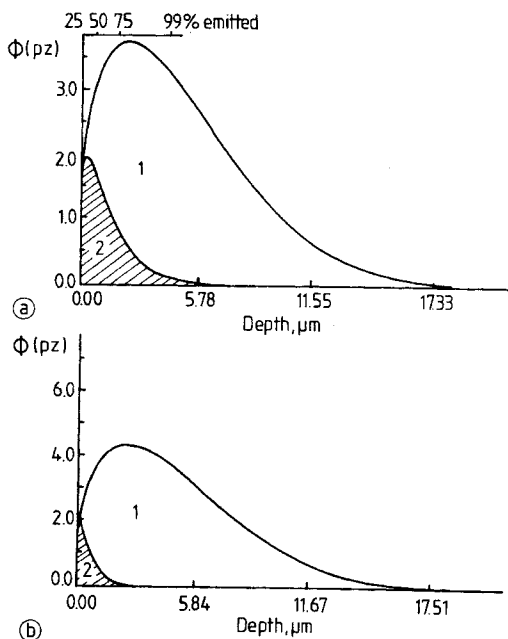
Quantitative microanalysis can be carried out more suitably by treating the atomic number and absorption factors together rather than as separate entities. The correction procedure based on the integration of X-ray distributions,  $\phi(\rho z)$  curves, is attractive for the analysis of low-energy X-ray 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<sub>act</sub> and Si<sub>dead</sub> 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

$$k_{ij} = \frac{k_{i, \text{is}}}{k_{j, \text{is}}} = \frac{k_i}{k_j} \quad (21)$$

where  $k_{i(j), \text{is}} = k_{i(j)}$  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]

$$k_a = \frac{(\mu/\rho)_{\text{sp}}^i}{(\mu/\rho)_{\text{sp}}^j} \cdot \frac{1 - \exp[-\chi^i(\rho t)]}{1 - \exp[-\chi^j(\rho t)]} \quad (23)$$

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. Utlaut 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

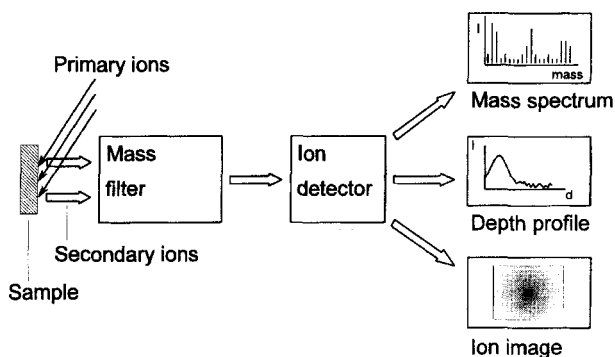
Secondary ion mass spectrometry (SIMS) is a surface and microanalytical technique for the investigation of the chemical composition of solids. The specimen to be analyzed is brought under high vacuum and bombarded with energetic ions (e.g., 10 KeV  $\text{Ar}^+$ ). The sputtering process results in the 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)	$\geq \text{H}$	ppm	0.3–2 nm	$\sim 100$ nm	Elemental, isotopic, organic
Scanning Auger microscopy (SAM)	$\geq \text{Li}$	0.1%	0.3–2.5 nm	50 nm	Elemental, some chemical states
X-ray mapping in electron probe X-ray microanalysis (EPXMA)	$\geq \text{B}$	0.1%	0.5–2 $\mu\text{m}$	1 $\mu\text{m}$	Elemental
X-ray photoelectron microscopy (XPS)	$\geq \text{He}$	1%	1–3 nm	$\sim 30$ $\mu\text{m}$	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

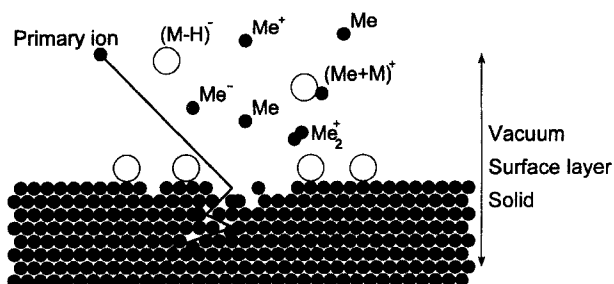
When a primary ion with an energy in the kiloelectronvolt range impacts on the surface of a solid it will collide with the sample atoms and gradually lose its energy until it comes to rest a few tenths of nanometers below the sample surface. The energy released in the collision cascade results in a disruption 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  $\pm$  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 m<sup>2</sup> 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  $M = 30$ .

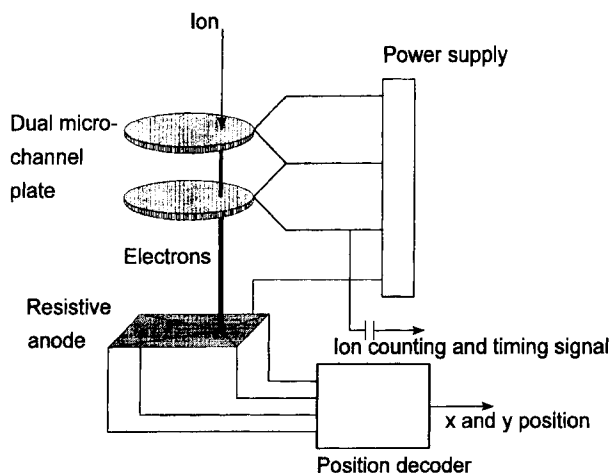
consists of a single or dual microchannel plate (MCP), which acts as an ion-to-electron converter and amplifier. The gain of the MCP depends on the applied voltage, and is limited to about 1000 for a single MCP and can be up to  $10^5$  (one ion producing  $10^5$  electrons) for a dual MCP. The localized 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.

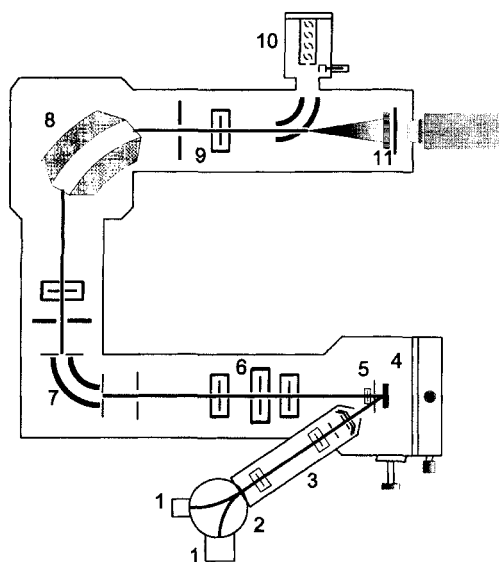


and the CCD camera [37–39]. The lateral resolution of both the RAE and camera-based system is sufficient not to cause any additional degradation of the ion image.

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 originally 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).

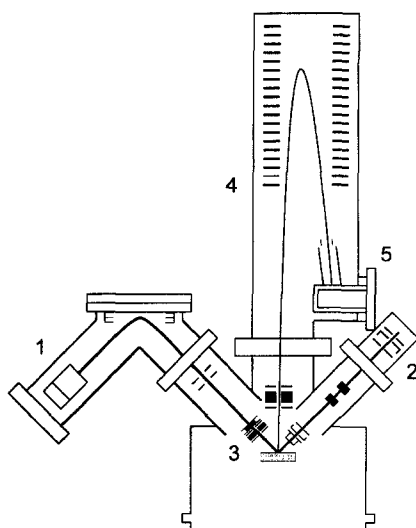
The projection lenses, situated behind the exit slit of the mass spectrometer, are used to postaccelerate the energy- and mass-filtered secondary ions and to direct them into the detection system. The beam can be focused and deflected into an electron multiplier or a Faraday cage to measure the ion current 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  $M/\Delta M$  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  $\text{Ga}^+$  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  $\sim 20\%$  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  $\text{Cs}^+$  ion source for ion microscope analysis and

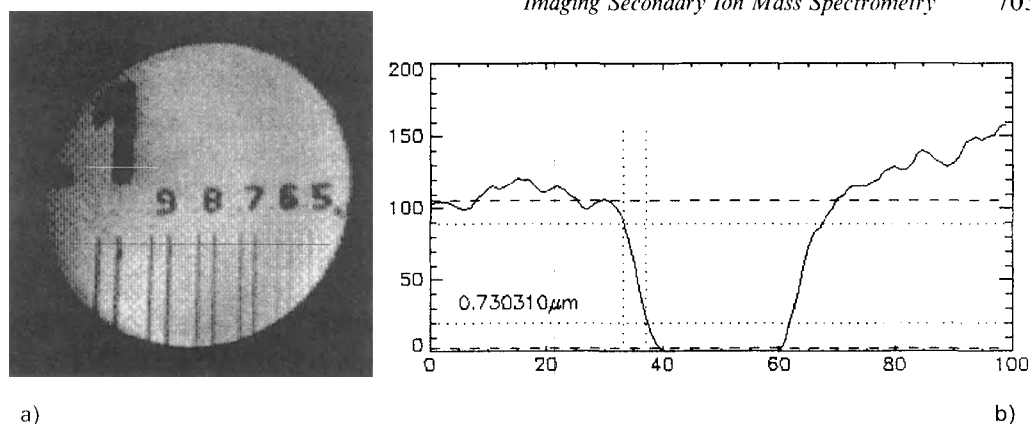
with a  $\text{Ga}^+$  liquid metal ion source for imaging in the microprobe mode. The TRIFT mass spectrometer has stigmatic optics consisting of three electrostatic energy analyzers [47]. In the ion microscope mode, a mass resolution  $>5000$  can be obtained, while the lateral resolution is limited to  $\sim 3\ \mu\text{m}$ . 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  $1\ \mu\text{m}$  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 $\mu\text{A}$
Beam diameter	$<1\ \mu\text{m}$	100 $\mu\text{m}$
Analyzed area	10 $\mu\text{m}$	150 $\mu\text{m}$
Lateral resolution	$\sim 100\ \text{nm}$	$\sim 1\ \mu\text{m}$
Detection limits	$\sim 0.1\%$	$\sim 1\ \text{ppm}$

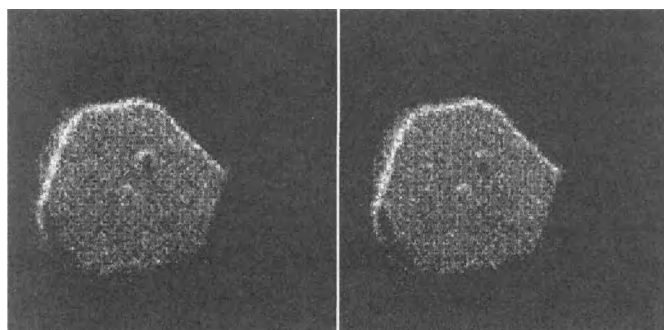


**Figure 6.** (a)  $\text{Si}^+$  ion image obtained from a silicon substrate with a gold test pattern (Cameca 4f in ion microscopy mode) and (b) line scan over the digit '1'. The image has a lateral resolution of  $0.7\text{ }\mu\text{m}$  close to the maximum attainable resolution in the ion microscope 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  $1\text{ }\mu\text{m}$  down to  $0.5\text{ }\mu\text{m}$ . 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  $0.7\text{ }\mu\text{m}$ , 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  $100\text{ ns}$  at a mass resolution of 300. The image field is  $6 \times 6\text{ }\mu\text{m}^2$ , and the lateral resolution is about  $75\text{ nm}$ , 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  $6 \times 6\text{ }\mu\text{m}$  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].

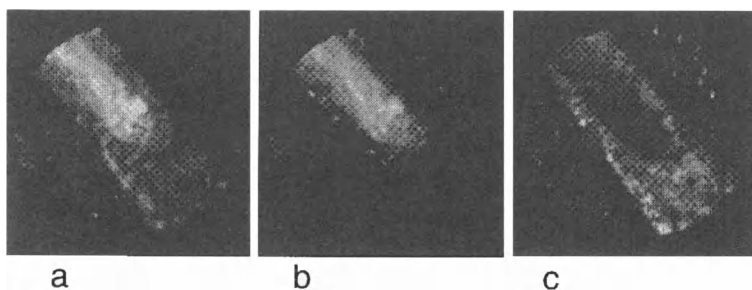
### 2.6.5.2 Influence of Mass Resolution

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 predefined 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.

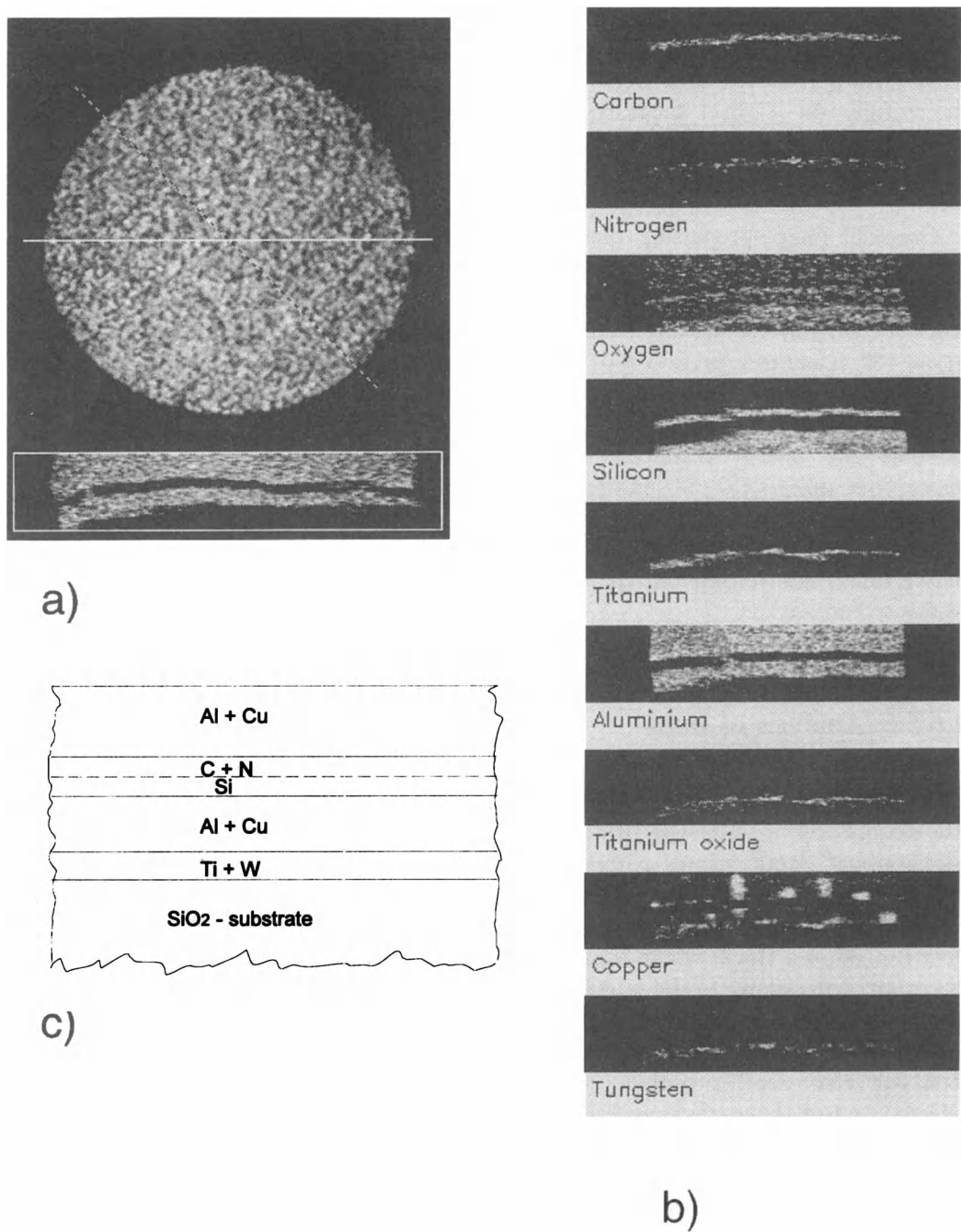
### 2.6.5.5 Analysis of Image Depth Sequences

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

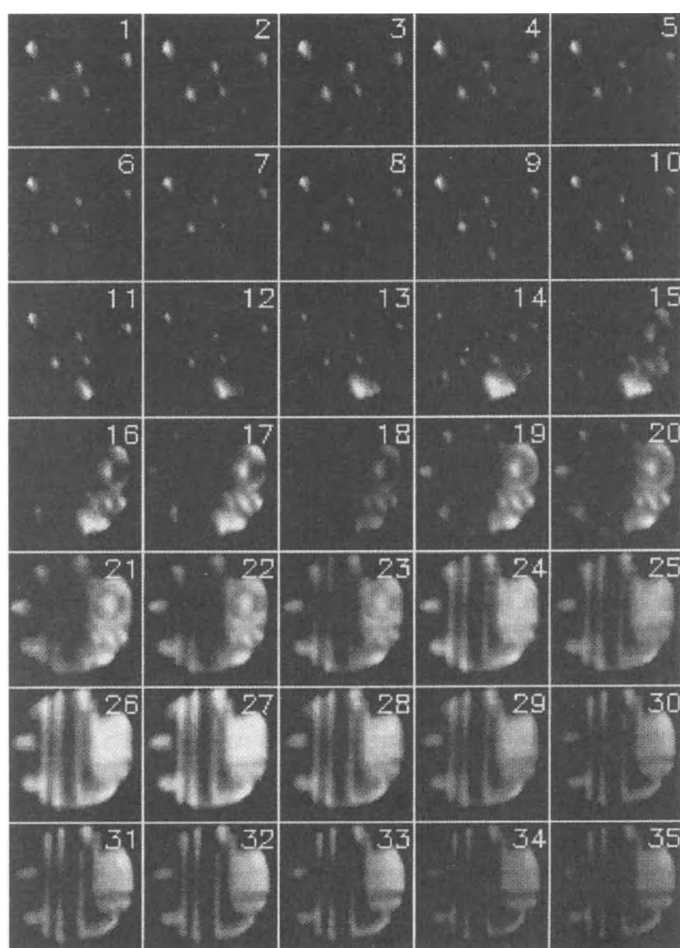
for the investigation of complex samples containing certain features buried below the surface. Cross-sections of the image stack allow the depth distribution of the various species to be viewed under different angles. Finally, a pseudo three-dimensional reconstruction of the interior structure can be 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.

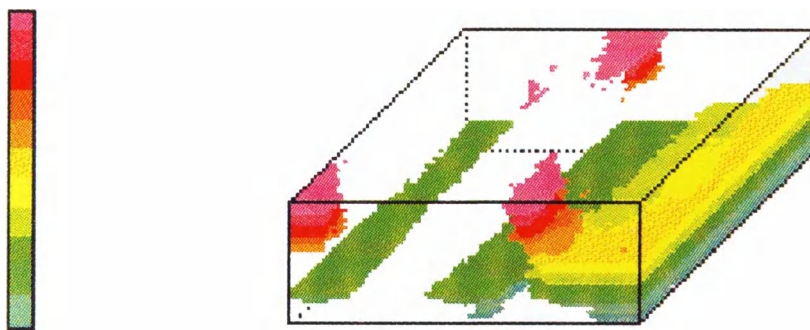


**Figure 11.** Lateral distribution of aluminum as a function of depth in a sample containing aluminum structures buried under an  $\text{SiO}_2$  passivation layer. Cycle 1 corresponds to the surface, and cycle 35 to  $\sim 1 \mu\text{m}$  below the surface.

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.



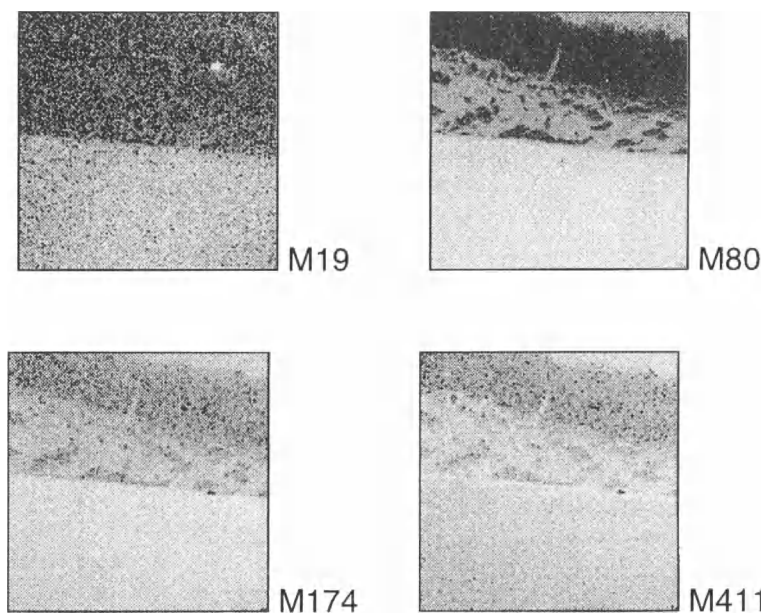
**Figure 12.** Reconstructed pseudo three-dimensional view of the aluminum distribution in the interior of a sample. In the transparent view the grey scale corresponds to depth in the sample.

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  $1\text{ }\mu\text{m}$  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  $\text{SO}_3^{2-}$ , 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



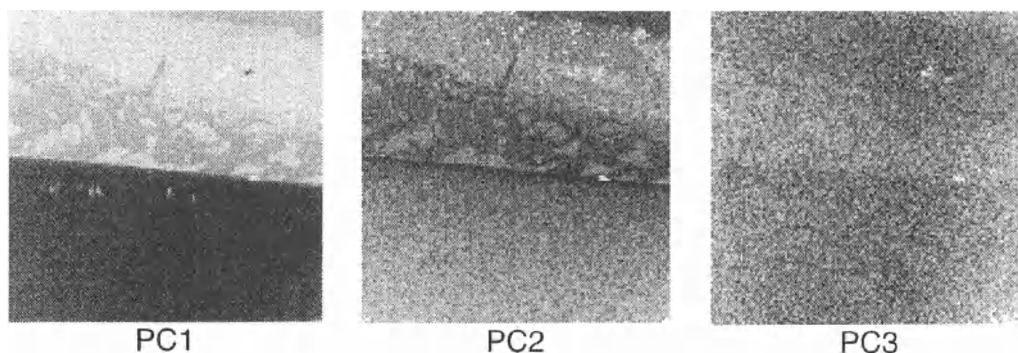
**Figure 13.** TOF-SIMS negative ion images from a copper surface treated with BBT: mass 19,  $F^-$ ; mass 80,  $SO_3^{2-}$ ; mass 174,  $(BBT-H)^-$ ; mass 411,  $(Cu(BBT-H)_2)^-$ .

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.

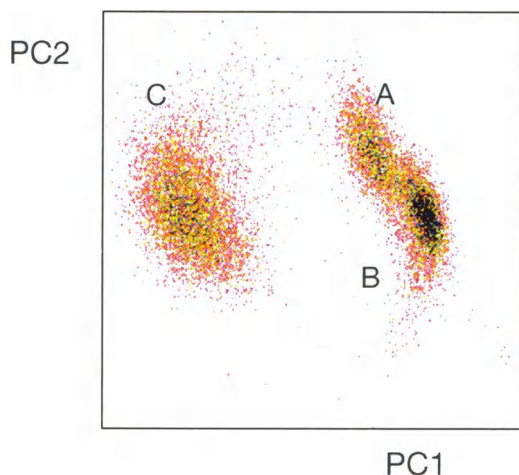
In this way it is possible to identify locations on the surface where specific types of chemical reactions have occurred.

## 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.

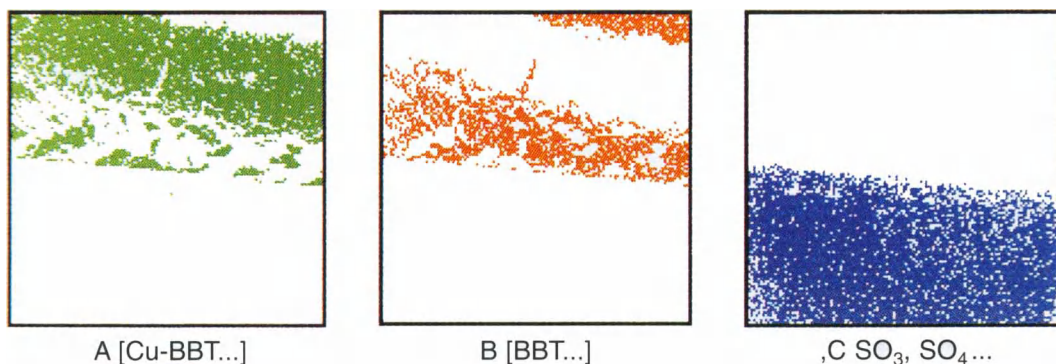
particles the samples should be conducting. Charge build-up will strongly affect the secondary ion extraction and cause considerable image degradation. Coating the sample with a conducting layer (gold), as is done in scanning electron microscopy (SEM), is less advantageous for SIMS. One has to sputter 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. Rüdenauer, *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. Rüdenauer, *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. Griffiths, *J. Vac. Sci. Technol.* **1985**, *A3*, 2102.
- [14] H. Hutter, M. Grasserbauer, *Mikrochim. Acta* **1992**, *107*, 137.
- [15] A. Benninghoven, F. G. Rüdenauer, 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. Griffiths 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. Zehnpenning, 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. Degrève, *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. Rüdénauer, 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.

# Magnetic Methods



# 1 Nuclear Magnetic Resonance

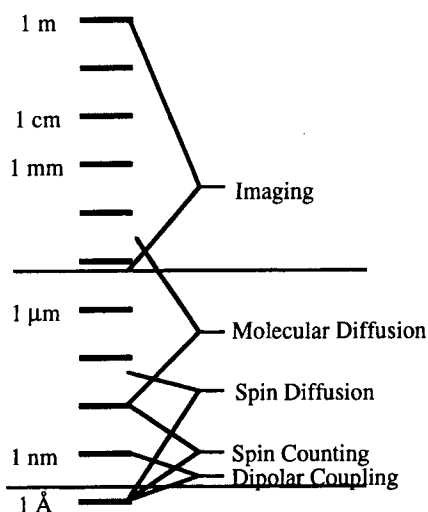
## 1.1 Introduction

The aim of this Chapter is to present a physical picture of the basic principles of nuclear magnetic resonance (NMR) microscopy. Since the focus is on spatial information, we take the unusual approach of discussing the NMR experiment in terms of magnetization gratings, how they are created and how they evolve. For length scales that are longer than about 100 Å, spatial NMR measurements rely on the use of magnetic field gradients to create a spatial grating in the spin magnetization. The spatial resolution of the measurement depends on the pitch of this grating and the period for which it can be maintained. To acquire an image, however, also requires that each individual volume element, a voxel, contains a detectable number of nuclear spins. Since NMR is a relatively insensitive technique, voxels are typically much larger than the finest grating pitch that can be created. In certain cases spatial information is available at these higher resolutions by performing a form of scattering experiment. Figure 1 shows the spatial scales of NMR imaging and scattering.

In the imaging regime, each set of spins in the sample is correlated with a particular volume element, and the desired

information is either the density of spins, or the density of a subset of the sample. In the scattering case the information is a local correlation, such as the mean distance that spins can move over a short period. For instance, if the displacement is molecular diffusion of a liquid confined to small regions by boundaries, then the scattering experiment will return the size distribution of the confining cells, but will include no information on the location of an individual cell. Since each set of spins in the sample contributes coherently to the overall signal, then the sensitivity of the measurement is independent of the length scale, and arbitrarily small features can be explored. In imaging where each feature is observed as an isolated response, this is no longer the case. In Fig. 1 the horizontal line at about 7 µm separates the area above which imaging is viable and below which scattering methods must be employed.

Image contrast is an important issue which we will not have the space to explore. An NMR spectrum is rich in chemical and morphological detail, and any feature (chemical shift, coupling, relaxation time) may be used as the basis for creating contrast in NMR imaging. Any imaging experiment may be preceded by a period of contrast creation where a select spectroscopic feature is enhanced relative to another. In such fashion a single



**Figure 1.** Spatial scales of NMR experiments. The scale is split into three regions, at the longest length scales absolute spatial information is recorded where the signal is correlated to a specific area of space. In the intermediate length scales a resolved volume element no longer contains sufficient number of spins to be individually resolved, but one can still measure average properties of the system, often associated with some form of transport, the scattering regime. In both of the above, linear magnetic field gradients are employed to create a well-defined spatial grating. At the highest resolutions, a grating can no longer be created via the gradient, and methods rely instead on spatially heterogeneous chemistry. Here, detailed average information such as internuclear distances is still available, but the information is not obtained in a linear fashion.

chemical component in a mixture may be selected as the basis for image formation, and in the solid state the selectivity may be extended to a specific molecular orientation. The details of such experiments have been reviewed, and a list of general references is included at the end of the chapter.

## 1.2 Background

In its most general form, the NMR experiment requires a quantum mechanical

description which is usually in terms of the density matrix approach since the measured bulk magnetization is from a collection of nuclear spins. For non-interacting spin =  $\frac{1}{2}$  nuclei, however, a semiclassical description introduced by Bloch provides a complete picture of the dynamics, and is adequate for most aspects of the imaging experiment. Only the Bloch picture will be introduced here.

Most elements have at least one isotope with a nonzero spin angular momentum,  $I$ , and an associated magnetic moment,  $\mu$ , which are related by the gyromagnetic ratio,  $\gamma$ ,

$$\mu = \hbar\gamma I \quad (1)$$

Both the gyromagnetic ratio and the magnitude of the nuclear spin  $I$  are characteristic of the nuclear state and are invariant. Protons have the highest gyromagnetic ratio of stable nuclei with  $\gamma = 2\pi \cdot 4250 \text{ Hz G}^{-1}$ ,  $I = \frac{1}{2}$ , and are nearly 100% naturally abundant. For sensitivity reasons the majority of imaging experiments are measures of proton spin density.

The NMR experiment involves the interaction of the bulk magnetic moment,  $M$ , of a collection of nuclear spins with a strong externally applied magnetic field. The interaction follows the Zeeman interaction, with the energy being proportional to the field,  $B_0$ ,

$$E = -M \cdot B_0 \quad (2)$$

The bulk magnetization is a superposition of the magnetic moments of a large number of individual spins. For spin =  $\frac{1}{2}$  nuclei, the nuclear spin is a magnetic dipole with two allowed states, parallel and antiparallel to the applied field, with an energy difference of

$$\Delta E = \hbar\gamma|B_0| = \hbar\omega_0 \quad (3)$$

This energy difference is quite small, and the associated frequency of the electromagnetic photons,  $\omega_0$ , is in the radio-frequency (RF) portion of the spectrum. A collection of spins in thermal equilibrium will populate these two states according to Boltzmann statistics,

$$\frac{N_-}{N_+} = e^{-\Delta E/kT} \quad (4)$$

At field strengths that can be created in superconducting magnets (1–10 T), and at room temperatures, the resulting population difference is about 1 spin in  $10^6$ . The detected magnetization is this small excess, and so the NMR experiment is relatively insensitive; normally greater than  $10^{15}$  spins are required for detection.

Since the individual spins are quantized along the direction of the external magnetic field (along the  $z$ -axis), then only this component of the nuclear spin has a definite value at equilibrium, the two transverse components ( $x$  and  $y$ ) are in superposition states. The bulk nuclear magnetization at equilibrium is then a stationary magnetic moment aligned along the  $z$ -axis.

The dynamics of the bulk magnetization away from equilibrium can be broken down into two simple motions: (1) a precession about the applied magnetic field associated with the torque originating from Eq. (2).

$$\frac{d}{dt}\mathbf{M}(t) = \gamma\mathbf{M}(t) \times \mathbf{B} \quad (5)$$

and (2) a relaxation that carries the magnetization back to equilibrium.

Since the external field is about the  $z$ -axis, the precession is around the  $z$ -axis at a frequency of

$$\omega_0 = \gamma|\mathbf{B}_0| \quad (6)$$

the so-called Larmor frequency. This linear dependence of the precession frequency to the applied magnetic field will form the basis of NMR imaging methods.

The relaxation process has two components, the magnetization along the  $z$ -axis relaxes towards its equilibrium value,  $M_0$ , and the transverse components relax towards zero. Since the bulk magnetization is composed of the magnetic moments of many individual spins, the length of the magnetization vector has a maximum but no minimum value (two individual magnetic moments can be antiparallel and thereby reduce the length of the bulk magnetization vector). Therefore, the transverse component may relax at a faster rate than the  $z$ -component does in its return to equilibrium, and two time constants are required to describe spin relaxation. These two relaxation times are called the spin–lattice relaxation time,  $T_1$  (along the  $z$ -axis) and the spin–spin relaxation time,  $T_2$  (governing the decay of transverse magnetization).

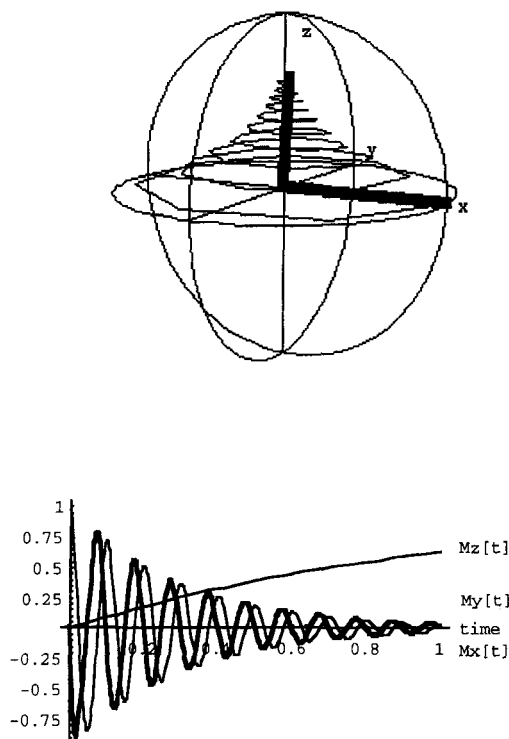
The Bloch equations describing the complete dynamics of a collection of isolated spins are

$$\begin{aligned} \frac{d}{dt}\mathbf{M}(t) = & \gamma\mathbf{M}(t) \times \mathbf{B} - \frac{1}{T_1}(M_z - M_0)\hat{z} \\ & - \frac{1}{T_2}(M_x\hat{x} + M_y\hat{y}) \end{aligned} \quad (7)$$

Notice that  $T_2$  must be less than or equal to  $T_1$  to conserve the maximum length of the magnetization vector. Two useful pictorial representations of the evolution of the spin magnetization back to equilibrium, after having first been placed along the  $x$ -axis, are shown in Fig. 2.

In terms of the Bloch equations, Eq. (7), notice that two types of magnetic fields can





**Figure 2.** Diagrams of the return of the spin magnetization to equilibrium after being aligned along the  $x$ -axis. In both figures the evolution of a single bulk magnetization vector is being followed. The initial position is shown as the transverse vector at the top, which spirals in to the  $z$ -axis, the vertical vector. In the lower figure the three individual components of the magnetization are shown as a function of time. The NMR experiment measures the two transverse components,  $M_x$  and  $M_y$ . There are three motions, a precession about the  $z$ -axis, a decay of the transverse components and a slower growth along  $z$  toward the static equilibrium value.

have a pronounced influence on the dynamics: static fields along the  $z$ -axis, and transverse fields rotating at the Larmor frequency. Changes in the strength of the field along the  $z$ -axis result in changes in the Larmor frequency. Since the magnetization is precessing at the Larmor frequency, transverse magnetic fields that rotate about  $z$  at the same frequency (resonant fields) will create a constant torque

on the spins and rotate them away from the  $z$ -axis. This second motion is most readily seen in a reference frame where the magnetization is stationary, the 'rotating' frame defined by the transformation

$$\frac{d}{dt} M_{\text{rotating}} = \frac{d}{dt} M_{\text{lab}} - \omega_0 M_x \hat{y} + \omega_0 M_y \hat{x} \quad (8)$$

If the applied magnetic fields are separated into two components, a static field along the  $z$ -axis,  $B_0$ , and a rotating field along the  $x$ -axis of the rotating frame,  $B_1$ , then the Bloch equations simplify to

$$\begin{aligned} \frac{d}{dt} M_x &= \Delta\omega M_y - \frac{M_x}{T_2} \\ \frac{d}{dt} M_y &= -\Delta\omega M_x + \omega_1(t) M_z - \frac{M_y}{T_2} \\ \frac{d}{dt} M_z &= -\omega_1(t) M_y - \frac{M_z - M_0}{T_1} \end{aligned} \quad (9)$$

where  $\Delta\omega$  is a small off-resonance term (the difference between the rotating frame frequency and the Larmor frequency), and  $\omega_1$  is the strength of the rotating field (the RF field),

$$\omega_1 = \gamma |B_1| \quad (10)$$

For simplicity, the NMR experiment is usually referenced to a frame that moves with an externally applied RF field.

Notice that in transforming to the rotating frame, only those components that will influence the spin dynamics were kept, the fields that are static in the rotating frame. All other fields will be purely oscillatory in this reference frame, and the long time behavior of the spins will not be influenced by them.

The overall NMR experiment can be described with reference to Eq. (9). The spins are placed in a homogeneous

magnetic field, and after the populations have reached equilibrium, an RF field is applied for a short time. This RF pulse rotates the bulk magnetization away from the  $z$ -axis at an angular velocity of  $\omega_1$ . By varying the length of the RF pulse, the angle can be made equal to  $90^\circ$ ,  $\omega_1 t = \pi/2$  (a so-called  $\pi/2$  pulse), and the magnetization will be aligned along the  $y$ -axis. The RF field is then turned off and the spins return to equilibrium. The transverse components of the magnetization vector ( $M_x$  and  $M_y$  in Fig. 2) are captured by monitoring the voltage induced in a turned coil wrapped around the sample.

The measured signal, the free induction decay (FID), is in the time domain, and according to Eq. (9) has the form

$$M_{x,y}(t) = M_0 e^{-i\Delta\omega t} e^{-t/T_2} \quad (11)$$

The Fourier transformation of the FID is a Lorentzian with a width of

$$\nu_{\text{fwhm}} = \frac{1}{\pi T_2} \quad (12)$$

where 'fwhm' denotes full width at half-maximum. Since both the Fourier transformation and the Bloch equations in the absence of an RF field are linear, we are free to break the NMR experiment into a superposition of identical experiments on each NMR distinct set of spins. So a general result is a frequency spectrum showing many resonances each corresponding to an identifiable spin system. The frequency differences may arise from small chemical shifts associated with the partial screening of the magnetic field by the electrons, from spin-spin couplings, or from local variations in the magnetic field strength due to the heterogeneity of the magnetic susceptibility throughout the sample.

### 1.3 Magnetic Field Gradients, Magnetization Gratings, and $k$ -Space

In order to record an image of a sample, there must be a measurable difference introduced between two locations in the sample. The most straightforward approach to this is to apply a magnetic field gradient,

$$B = B_0 + \frac{\partial}{\partial u} B_z \quad (13)$$

where  $u$  is  $x$ ,  $y$ , or  $z$ , so that the resonance frequency varies across the sample. Only the three partial derivatives of the  $z$ -component of the magnetic field are of interest, since the others correspond to static fields in the transverse direction of the laboratory frame, and thus rotating fields in the rotating frame that do not influence the dynamics.

The presence of a magnetic field gradient introduces a spatial heterogeneity into the experiment, and the observed bulk magnetization is that integrated over the entire sample. If a  $z$ -gradient ( $\partial B_z / \partial z$ ) is applied to a sample whose spin density is described by  $\rho(x, y, z)$  then the FID, Eq. (11), becomes

$$\begin{aligned} M_{x,y}(t) = & M_0 e^{i\Delta\omega t} e^{-t/T_2} \\ & \times \int \rho(x, y, z) \\ & \times e^{-i\gamma(\partial B_z / \partial z)zt} dx dy dz \end{aligned} \quad (14)$$

This is put in a more recognizable form by introducing the reciprocal space vector,

$$k_i = \int \gamma \frac{\partial}{\partial i} B_z(t) dt \quad (15)$$

and by defining the projection of the spin density along the  $z$ -axis,

$$P(z) = \int \rho(x, y, z) dx dy \quad (16)$$

so that

$$M_{x,y}(t, k_z) = M_{x,y}(t, 0) \int P(z) e^{-ik_z z} dz \quad (17)$$

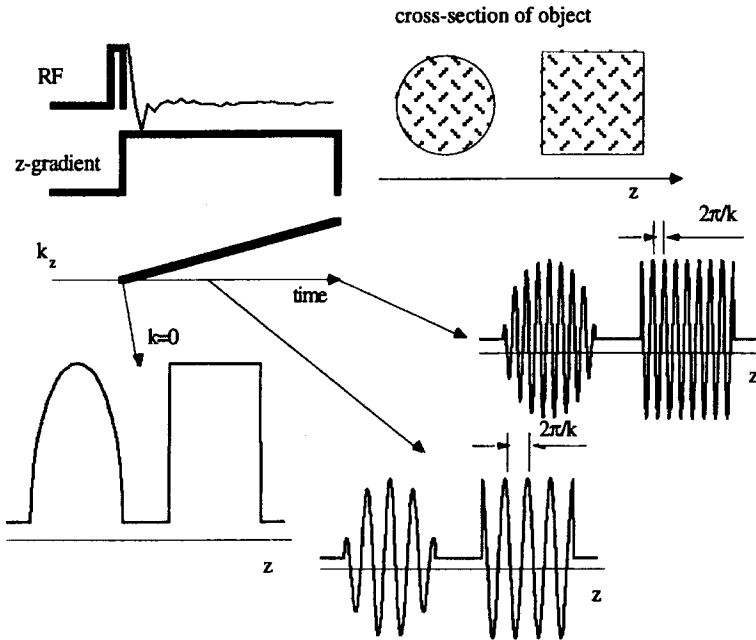
where  $M_{x,y}(t, 0)$  is the NMR signal in the absence of the magnetic field gradient. The measured NMR signal in the presence of a magnetic field gradient is the Fourier transform of the projection of the spin density along the direction of the gradient

convoluted with the NMR lineshape in the absence of the gradient,

$$\text{Image} = \text{ObjectFunction} \otimes \text{NMR lineshape} \quad (18)$$

It is helpful to understand how it is that the NMR signal is a measure of the Fourier components of the spin density. As shown in Fig. 3, spin evolution in the presence of a magnetic field gradient introduces a sinusoidal magnetization grating across the sample, and the  $k$ -vector describes the spacing of the grating,

$$k_z = \frac{2\pi}{\lambda_z} \quad (19)$$



**Figure 3.** A schematic description of a one-dimensional NMR imaging experiment to measure the projection of the object spin density along a given axis. The object is shown and consists of two regions of spins with circular and square cross-sections. The  $z$ -axis is identified and the applied magnetic field gradient is a  $\partial B_z / \partial z$  gradient, so the Larmor frequency of the spins in the object increases from left to right. The measurement is proceeded by an RF pulse that excites the spins and places them in the transverse plane. The three profiles of the object show one axis of the transverse magnetization at three different values of  $k$ . Higher  $k$ -values correspond to finer magnetization gratings. The NMR signal is the integrated spin magnetization across the sample, which for this sample has the shape shown following the RF pulse.

where  $\lambda_z$  is the period of the grating. The essence of the NMR imaging experiment is that by a suitable combination of magnetic field gradients oriented along the  $x$ ,  $y$ , and  $z$  laboratory frame directions any Fourier component of the spin density can be measured. Since these gradients are under experimental control and can be switched on in a very short time compared to the NMR relaxation times, then there is a great deal of flexibility in how one proceeds to record all of the necessary Fourier components needed to reconstruct an image. We will return to this after discussing the limits to resolution, and multiple-pulse experiments.

In Eqs. (17) and (18), the NMR image is described as a convolution of a spatial function describing the density of the spins, and a frequency function describing the NMR spectrum. To make sense of this, the NMR spectrum must be converted into spatial units. Notice in Fig. 3 that in the presence of a constant magnetic field gradient, the two reciprocal space vectors, time and  $k$ , are directly proportional to each other,

$$\frac{k_z}{t} = \gamma \frac{\partial}{\partial z} B_z = \frac{\Delta\omega}{\Delta z} \quad (20)$$

So if the NMR spectrum corresponds to a single resonance line (such is the case for water) then the image resolution can be described in terms of the linewidth,  $\nu_{\text{fwhm}}$  (or the spin-spin relaxation time), and the gradient strength,

$$\text{resolution} \geq \nu_{\text{fwhm}} \left( \gamma \frac{\partial}{\partial z} B_z \right)^{-1} \quad (21)$$

Sharper lines and stronger gradients lead to higher resolution. This is what makes

NMR imaging of solids challenging since the  $T_2$  is approximately 1000 times shorter for solids than for liquids.

There are other limitations to the obtainable resolution, the most important in microscopy of liquids being the random molecular motions of the spins. Water has a diffusion coefficient of about  $3 \mu\text{m}^2 \text{ms}^{-1}$  at room temperature, and as the molecules move they carry their magnetic moment with them. This leads to an irreversible blurring of the magnetization grating and hence a loss in high spatial frequency information (Fig. 4). The influence of the spin motion on the magnetization grating is included in the Bloch equation by adding a diffusion term,

$$\frac{dM_i(t)}{dt} = D \nabla^2 M_i(t) \quad (22)$$

Equation (22) has solutions of

$$M_i(t) = M_i(0) e^{-Dk^2 t} \quad (23)$$

in a constant grating, and

$$M_{x,y}(k, t) = M_{x,y}(0) e^{-ikz} e^{-Dk^2 t/3} \quad (24)$$

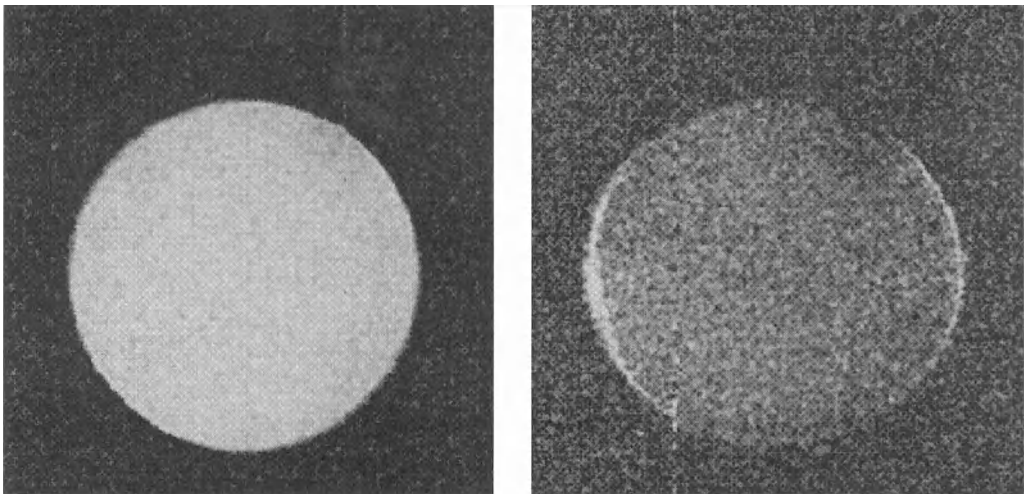
in a constant magnetic field gradient. The factor of three in the exponential of Eq. (24) arises since the spacing of the grating is varying with time, and the attenuation is most pronounced for finer gratings.

Recognizing that there are potentially many contributions to the overall resolution of the imaging experiment, the convolution of Eq. (18) may be generalized as an ideal linear model,

Image = ObjectFunction

$\otimes$  PointSpreadFunction + noise

(25)



**Figure 4.** Two NMR images of a small tube of water. The image on the left was acquired with strong gradients so that molecular diffusion does not influence the image intensity, that on the right was acquired with much lower gradients strength and molecular diffusion destroyed the grating prior to acquisition. Notice the sharp half moons on the edges of the image; these arise from molecules that have an effective lower diffusion coefficient since they are near boundaries. The images are most sensitive to diffusion along the horizontal axis where a gradient echo was employed. The image field of view is  $1.7 \times 1.7$  mm and the resolution is  $6.6 \times 6.6$   $\mu$ m.

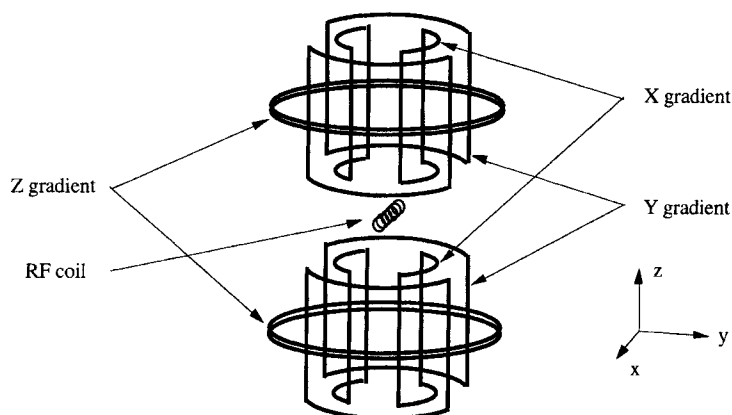
The total point spread function (PSF) is then the convolution of the individual contributions, some of which are listed in Table 1.

In Table 1, interrupted molecular diffusion, which occurs where boundaries effectively reduce the diffusion coefficient, leads to a spatially dependent diffusion constant. Any contribution to the NMR linewidth that is spatially varying is nonlinear.

One of the reasons that NMR imaging is so successful is a result of the truncation of Hamiltonians to their secular components, that is, all of the interactions (except RF pulses) commute and hence do not interfere with one another. Since each interaction is itself linear, then the entire spin dynamics are linear and, to a good approximation, the linear model is accurate to the extent that the instrument is correctly engineered.

**Table 1.** Contributions to the NMR resolution

Contribution	Origin	Linear/nonlinear
NMR linewidth	$T_2$ relaxation	Linear
Bulk susceptibility	Variation of the local magnetic field due to variations in the local bulk susceptibility	Nonlinear
Free molecular diffusion	Random motion of spins	Linear
Interrupted molecular diffusion	Reduction of the effective diffusion coefficient by boundaries	Nonlinear
Instrumental imperfections	Nonlinearity in the gradients, inhomogeneity of the RF field	Nonlinear



**Figure 5.** A schematic representation of the NMR microscopy probe; the magnetic field is along the vertical axis. The probe has a small RF coil wrapped tightly about the sample; the good filling of the coil is necessary to reduce resistive losses in the coil. This is surrounded by three gradient coils that are connected to audio-frequency amplifiers. The z-gradient is a Maxwell pair through which currents flow in opposite directions. Moving inwards, the *x*- and *y*-gradient coils are each composed of four semicircular current paths through which currents flow in parallel.

## 1.4 Nuclear Magnetic Resonance

There are many excellent descriptions of NMR instrumentation; here the focus is on the changes necessary to implement microscopy on a high-resolution, high-field spectrometer. Fortunately these are quite modest, since many spectrometers have the necessary gradient amplifiers and controllers, and all modern instruments are capable of performing the experiments. The fundamental difference is the NMR probe, shown schematically in Fig. 5.

Notice that the field of view is normally limited physically by the size of sample

that can be accommodated in the RF coil. This is advantageous for sensitivity reasons, and quite necessary for reasonable experimental times. Some typical values are given in Table 2.

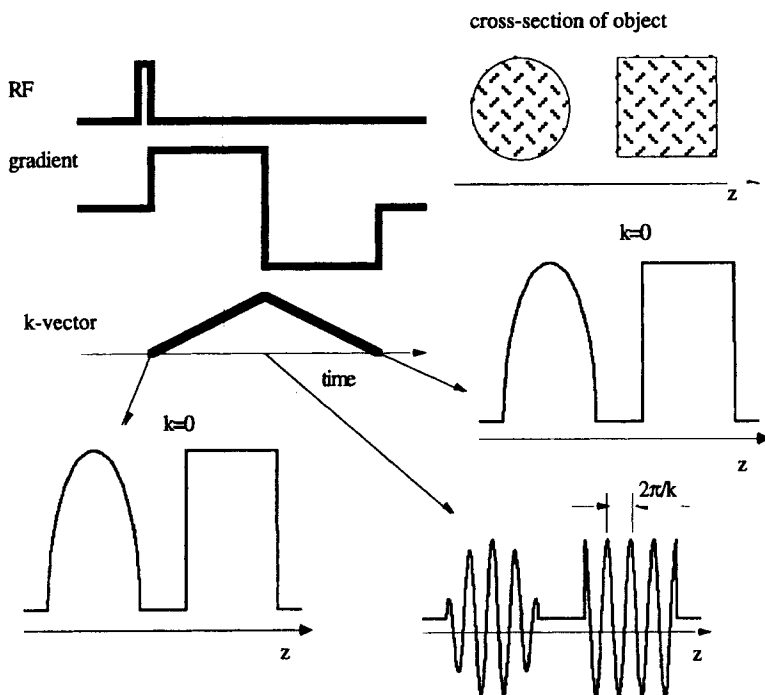
## 1.5 Echoes and Multiple-Pulse Experiments

That the various interactions, the gradient, chemical shift, susceptibility, and couplings, each drive the spin dynamics individually is useful and permits a simple linear model of the sensitivity and resolution of the imaging system, but equally important

**Table 2** Some typical NMR microscopy probe configurations

Field of view	2.5 cm	1.0 cm	2.5 mm
Resolution	100 $\mu\text{m}$	20 $\mu\text{m}$	5 $\mu\text{m}$
Gradient strength	50 $\text{G cm}^{-1}$	100 $\text{G cm}^{-1}$	1000 $\text{G cm}^{-1}$
Magnet	9.6 T/89 mm	9.6 T/89 mm	14 T/54 mm

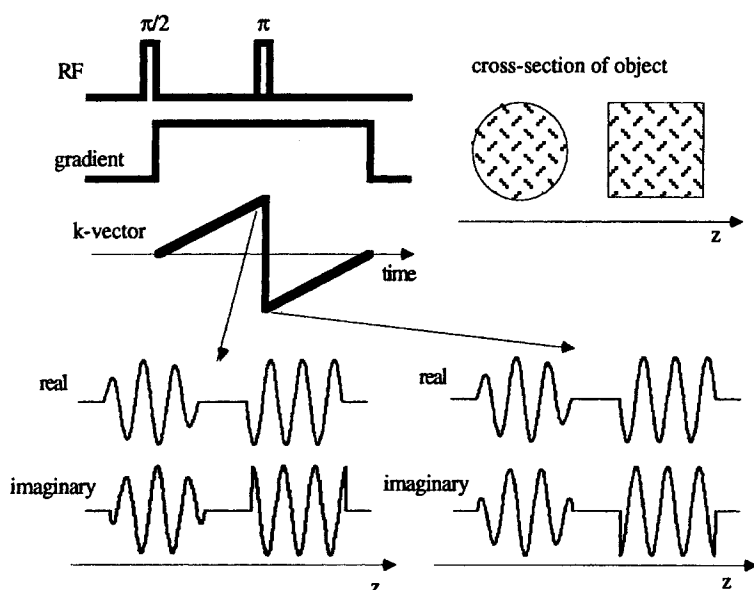
**Figure 6.** A gradient echo is generated by a bipolar gradient waveform. Since the  $k$ -vector is the integrated area under the gradient waveform, Eq. (15), at the end of the bipolar gradient  $k$  returns to zero, and there is no magnetization grating across the sample. At the mid-point of the gradient waveform  $k$  is at a maximum and the grating is at its finest. Since the NMR signal is the integration of the spin magnetization across the sample, the signal maxima correspond to the two points where  $k$  is zero, and hence an echo is observed.



(particularly in creating contrast) is the experimenter's ability to separate these by selectively refocusing one or more interactions. The gradient echo is perhaps the simplest example of this. Recall that the periodicity of the magnetization grating is the zero moment of the time-dependent gradient waveform (the integrated area of the waveform), Eq. (15). By applying the gradient as a bipolar pulse pair (Fig. 6) the grating is removed from the object at the end of the gradient waveform and the spins are refocused (back in phase). Since the only interaction that is influenced by the presence of a gradient is the gradient evolution itself, the bipolar gradient waveform shown in Fig. 6 will not refocus the chemical shift, or any other internal interaction. We will leave the applications of gradient echoes to later and turn to the question of how other interactions can be refocused.

In select cases the sign of the interaction can be changed, as it was in the gradient echo, but, in general, refocusing is accomplished through the actions of RF pulses that strongly modulate the spin dynamics. RF pulses are normally arranged such that the RF field strength is stronger than any internal interaction and so they can be thought of as instantaneously changing the spin state of the system.

The simplest RF pulse-created echo is the two-pulse  $(\pi/2, \pi)$  sequence shown in Fig. 7. The first RF pulse rotates the magnetization from the  $z$ -axis into the transverse plane, where in the presence of a gradient it creates a magnetization grating. The  $\pi$ -pulse inverts the sense (directionality) of this grating so that continued evolution in the magnetic field gradient leads to the formation of an echo.



**Figure 7.** The figure shows the action of a  $\pi$ -pulse when applied to a magnetization grating. Here both quadrature components (real and imaginary) of the transverse magnetization must be kept track of. The two gratings are shown just before and just after the  $\pi$ -pulse which inverts the sign of one of the two components (the real component in the figure). This is consistent with the RF pulse corresponding to a rotation of  $180^\circ$  about an axis in the transverse plane. Before and after the  $\pi$ -pulse the magnetization grating has the same pitch, but the sign of the grating is changed by the pulse. Following the  $\pi$ -pulse the magnetic field gradient is still in the same direction and so the  $k$ -vector continues to increase, but now towards zero, and following an equal length of time  $k$  returns to zero and all of the magnetization is again coherent. A spin echo is formed.

The spin echo depicted in Fig. 7 will refocus any time-independent variations in the resonance frequency, including chemical shifts, susceptibility shifts, and hetero-nuclear spin couplings. Earlier the spin-spin relaxation time was identified with the decay of the bulk magnetization in the transverse plan, and we now see that this is not a fundamental relaxation phenomenon since, for example, the application of a magnetic field gradient can increase its rate, and since the magnetization can be refocused. The fundamental spin-spin relaxation rate is associated with those decay processes that cannot be refocused by methods such as the spin echo, and the observed decay of the transverse magnetization is the rate sum of this and other

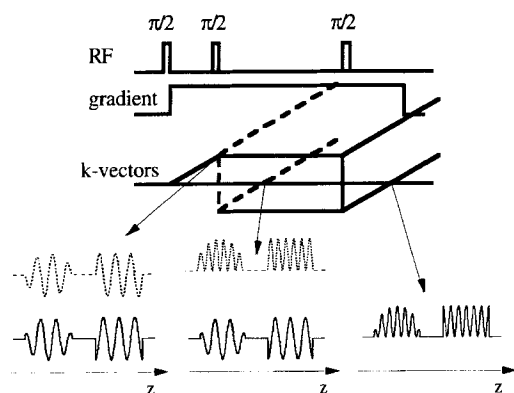
contributions,

$$\frac{1}{T_2^{\text{observed}}} = \frac{1}{T_2} + \frac{1}{T_2^{\text{other}}} \quad (26)$$

One very useful imaging application of the spin echo is to change the image contrast from being a measure of the pure spin density (by using a spin echo to refocus spatial variations in the observed  $T_2$ ) to being dependent on the local magnetic field variations (by using a gradient echo).

Any train of RF pulse will create a pattern of echoes that can refocus line-broadening terms. A second application of spin echoes is to store information that can then be read out later. The stimulated echo shown in Fig. 8 is a classic example of this.





**Figure 8.** Schematic view of a  $\pi/2$ ,  $\pi/2$  echo and the stimulated echo. The second  $\pi/2$  RF pulse rotates one component of the transverse spin magnetization to the  $z$ -axis (along the field), where it does not evolve (recall from the Bloch equation, Eq. (9), that the  $z$ -magnetization only slowly decays). The remaining transverse magnetization is now an oscillator function of the spatial coordinate and so is a combination of  $+k$  and  $-k$  states (the dotted trajectories in the figure). Likewise, the  $z$ -magnetization is also a combination of  $+k$  and  $-k$  (the solid lines). Between the second and third RF pulses, only the transverse components evolve, and only one-half of the signal forms an echo. The third RF pulse returns the magnetization that was stored along the  $z$ -axis back into the transverse plane, where it too evolves and eventually yields the 'stimulated' echo. This echo has special importance, since between the second and third pulses the corresponding magnetization was stored as a special grating.

One of the most useful features of the stimulated echo is that a magnetization grating can be stored for a time comparable to the spin-lattice relaxation time (on the order of a few seconds), which is an ideal approach to measuring extremely slow processes—for instance molecular transportation over macroscopic distances.

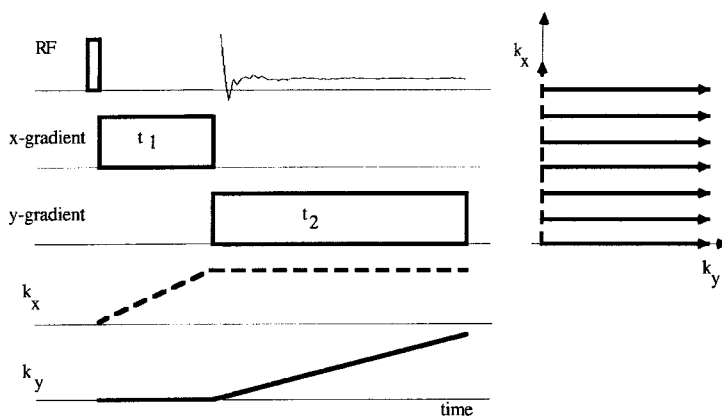
## 1.6 Two-Dimensional Imaging

Having seen that by applying a magnetic field gradient one can encode a

one-dimensional image and collect a projection of the spin density of the sample, one clear path to two-dimensional imaging is to collect a series of such projections with the gradient at various orientations and then to use Radon-filtered back projection to reconstruct the image. Early NMR images were acquired in this fashion, and occasionally solid state images still are. However, most imaging is performed via Fourier imaging where a two- or three-dimensional region of reciprocal space is sampled corresponding to the desired resolution and field of view. Such data are easily measured since the gradients are under experimental control and virtually any gradient waveform can be generated. Since the various sequences can be rather complex, often a pictorial representation of reciprocal space is employed where the trajectory of the experiment is mapped out. It is then possible to focus on the manner in which two-dimensional  $k$ -space is sampled rather than to get caught up in the details of the NMR experiment.

Notice that the imaging experiment contains two fundamentally different times, a phase-encoding interval ( $t_1$  in Fig. 9 during which no data are acquired) and a frequency-encoding interval (during data acquisition). These two times permit the separate encoding of the two interactions—the gradients in the two orthogonal directions—and thus allow all of  $k$ -space to be sampled. The image is the two-dimensional Fourier transform of the collected data.

The gradient evolution during the phase-encoding interval can be carried out in two fashions, the gradient strength can be kept constant and the time incremented, or the time can be kept constant



**Figure 9.** A generic two-dimensional Fourier imaging experiment. The data are collected on a Cartesian raster during the time period labeled  $t_2$ , and in the presence of a  $y$ -gradient. Prior to this, a brief  $x$ -gradient pulse has been applied to create a magnetization grating in the  $x$ -direction. The collection period therefore corresponds to the rays in the reciprocal space picture on the right, and each individual ray is collected during a separate experiment. The length of  $t_1$  is systematically varied to achieve the desired offset along the  $k_x$  axis.

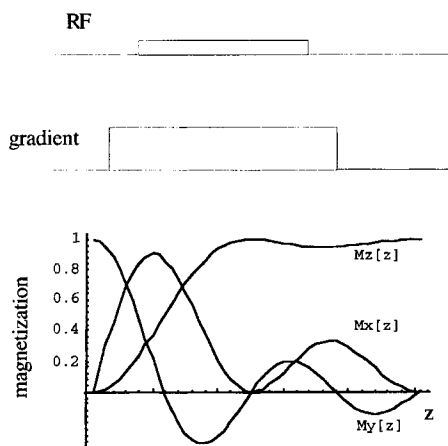
and the gradient strength incremented. The second, called spin warp imaging, has advantages since the extent of the evolution due to chemical shift or other non-gradient interactions is kept constant and these then appear solely as a signal attenuation factor. The resultant point spread function for a constant encoding time is an attenuated delta function (neglecting the contribution from sampling).

There is a wide range of imaging sequences, and we will not attempt to review these here; they all include the general features shown in Fig. 9. The two-dimensional experiments can be extended to three dimensions by encoding the third direction as a second phase-encoded axis.

## 1.7 Slice Selection

Recording a full three-dimensional image is often not the most economical approach to imaging, and slice selection can be

achieved by taking advantage of the frequency offset dependence of RF excitation. Looking back at the Bloch equations, Eq. (9), for an on-resonance RF pulse ( $\Delta\omega = 0$ ) the RF field is along the  $x$ -axis and hence the evolution of the spins is a simple rotation about the  $x$ -axis. However, as  $\Delta\omega$  increases, then the dynamics become more complex and are most easily visualized by considering an 'effective' field that is the vector sum of the RF field along the  $x$ -axis and the off-resonance field along the  $z$ -axis. The dynamics are still a simple rotation about this effective field, but the motion of the magnetization vector now describes a cone rather than a plane. The result is that as the frequency offset is increased, the angle of the effective field to the  $z$ -axis decreases, and eventually the RF pulse has very little influence. The key to slice selection then is to apply a relatively weak RF pulse in the presence of a strong gradient, so that the frequency offset is spatially dependent (Fig. 10).



**Figure 10.** Calculated selective excitation profiles for a weak RF pulse in the presence of a magnetic field gradient. The pulse length is set to rotate the on-resonance spins (at the origin) by  $90^\circ$ . Notice that as the resonance offset increases (with increasing  $z$ ), the effective rotation angle becomes smaller and most of the magnetization remains along the  $z$ -axis. For most images, a shaped RF pulse is employed that creates a square magnetization profile.

There is not a simple linear picture of the dynamics, but various shaped (amplitude-modulated) RF pulses have been developed that give well-defined square slice selection profiles.

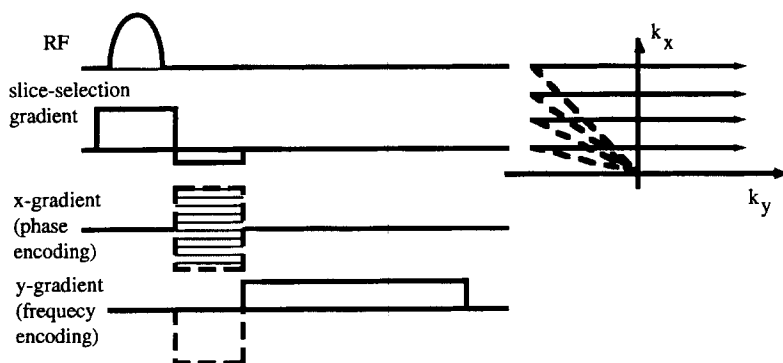
So the overall slice-selected two-dimensional imaging experiment might look like that shown in Fig. 11.

Figure 12 shows images of rat arteries acquired using a multislice two-dimensional spin echo technique. Such images might form the basis of rapid, nondestructive, three-dimensional imaging.

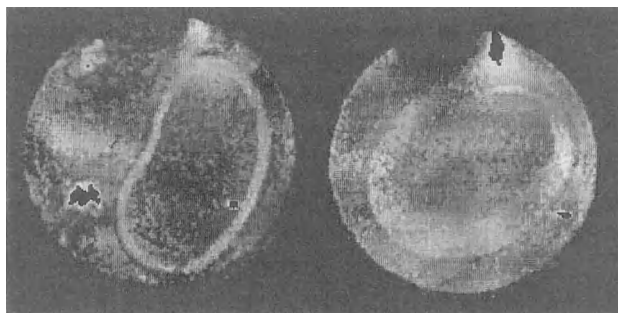
## 1.8 Gratings and Molecular Motions

We have already seen in the case of molecular diffusion that molecular motions vary the magnetization grating. Random motions, such as diffusion, lead to an attenuation of the grating, and coherent processes such as flow displace the grating. The precession due to a  $z$ -gradient at a position  $z_0$  has an instantaneous phase angle of  $k_z z_0$ , if the spins are not moving. In the presence of motion this angle (for a single spin packet) has the general form

$$\Theta_{\text{gradient}} = \gamma \int \frac{d}{dz} B_z(t) z(t) dt \quad (27)$$



**Figure 11.** A slice-selected two-dimensional spin warp, gradient echo sequence. Notice that during the phase-encoding time the  $k_y$  vector is offset so that both positive and negative values can be sampled. Since NMR is a coherent spectroscopic technique, this has the advantage of measuring the phase. The gradient echo in the slice selection direction refocuses the evolution of the spins during the selective RF pulse.



**Figure 12.** Images of a healthy rat artery (left), and a rat artery with atherosclerosis (right). The images were acquired using multislice two-dimensional spin echo in 25 min with eight averages, and have an in-plane resolution of  $13 \times 13 \mu\text{m}$  with a slice thickness of  $200 \mu\text{m}$ .

The time-dependent position of the spin can be expanded in the usual power series,

$$z(t) = z_0 + vt + at^2 + \dots \quad (28)$$

and the angle of precession is then a moment expansion of the gradient waveform,

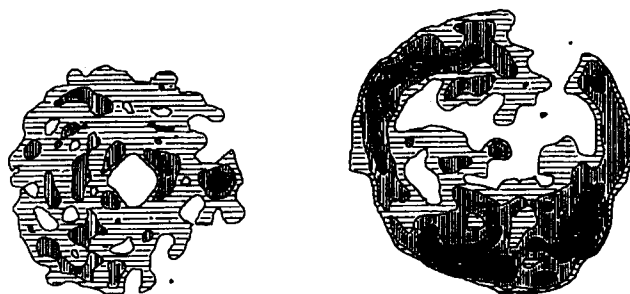
$$\begin{aligned} \Theta_{\text{gradient}} = \gamma \left\{ z_0 \int \frac{d}{dz} B_z(t) dt \right. \\ \left. + v \int \frac{d}{dz} B_z(t) t dt \right. \\ \left. + a \int \frac{d}{dz} B_z(t) t^2 dt \right\} \quad (29) \end{aligned}$$

Based on this expansion it is straightforward to generate a gradient waveform that

will selectively encode the position, velocity, or acceleration of the spin. For example, the gradient echo introduced earlier (see Fig. 6) does not encode the position (this is what is meant by an echo), but it does encode the velocity of the spin. Using gradient modulation methods it is possible to record images with contrast based on spin density, the diffusion coefficient, the velocity, or the acceleration.

## 1.9 Solid State Imaging

There is a special challenge when exploring solid samples in that the NMR linewidth is



**Figure 13.** Two-dimensional solid state images of the polybutadiene distribution in a solid, circular disk composed of a polystyrene/polybutadiene blend. Darker intensities correspond to higher butadiene concentrations. The contrast mechanism depended on the strength of the residual homonuclear dipolar coupling, and the diameter of the disk was 4 mm. The in-plane resolution is slightly finer than  $50 \mu\text{m}$ . (Reproduced with permission from D. G. Cory, J. C. de Boer and W. S. Veeman, *Macromolecules* **1989**, 22, 1618.)

normally quite broad (50 000 Hz), and straightforward applications of the above experiments require strong gradients, and result in a low signal-to-noise ratio. Static gradients of up to  $10\,000\text{ G cm}^{-1}$  have been achieved and used with back-projection methods, but more powerful methods are often based on artificially narrowing the NMR resonance. Much like a spin echo can be used to reduce the linewidth associated with static variations in the resonance frequency, there are a variety of coherent averaging schemes based on mechanical motions or RF pulse trains that greatly reduce the solid state NMR linewidths. Reductions in the linewidth of factors of 1000 are achievable for plastics and other organic solids, and these schemes have been combined with imaging methods.

Unfortunately, one of the results of most coherent averaging schemes is that the dynamics become highly nonlinear and therefore the experiments themselves become rather complex, in addition to requiring novel instrumentation. Such methods have been reviewed, and the interested reader is referred to these. An example is shown in Fig. 13. The fundamental resolution limit for solid state imaging is not limited by diffusion, as is

the case for liquids, but by the sensitivity of the measurement. To date, images have rarely been obtained with finer than  $100\text{ }\mu\text{m}$  resolution, but it is thought that the ultimate resolution will be on the order of a few micrometers for small samples.

## Acknowledgments

The authors thank the Whitaker Foundation, the National Science Foundation (DMR-9357603), and the National Institutes of Health (RR-00995) for supporting our imaging research, and S. C. acknowledges the receipt of a fellowship from the Korean government. We also thank Dr. S. Gravina for assistance with the experiments, and Dr. R. Kapadia for supplying the samples of rat arteries.

## 1.10 References

- B. Blümich, W. Kuhn, *Magnetic Resonance Microscopy*, VCH, Weinheim **1992**.
- P. Blümli, B. Blümich, *NMR—Basic Principles and Progress*, Springer-Verlag, Berlin **1993**.
- P. T. Callaghan, *Principles of Nuclear Magnetic Resonance Microscopy*, Oxford University Press, Oxford **1991**.
- D. G. Cory, *Annu. Rep. NMR Spectrosc.* **1992**, 24, 88.
- P. Jezzard, J. J. Attard, T. A. Carpenter, L. D. Hall, *Prog. NMR Spectrosc.* **1992**, 23.
- P. Mansfield, E. L. Hahn, *NMR Imaging*, The Royal Society, London **1990**.

## 2 Scanning Electron Microscopy with Polarization Analysis (SEMPA)

### 2.1 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.

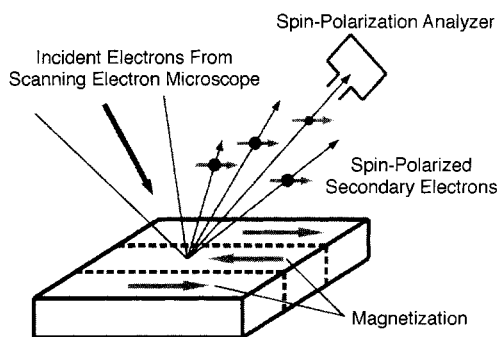
## 2.2 Principle of SEMPA

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,

$$M = -\mu_B(n_{\uparrow} - n_{\downarrow}) \quad (2)$$

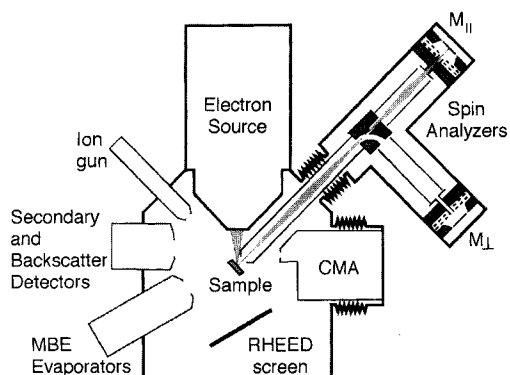
Here  $n_{\uparrow}$  ( $n_{\downarrow}$ ) are the number of spins per unit volume parallel (antiparallel) to a particular orientation, and  $\mu_B$  is the Bohr magneton. To the extent that the secondary electron cascade represents a uniform excitation of the valence electrons, the expected secondary 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.

## 2.3 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.

### 2.3.1 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 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.

### 2.3.2 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].

### 2.3.3 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  $\mathbf{s}$  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

$$\sigma(\Theta) = I(\Theta)[1 + S(\Theta)\mathbf{P} \cdot \mathbf{n}] \quad (3)$$

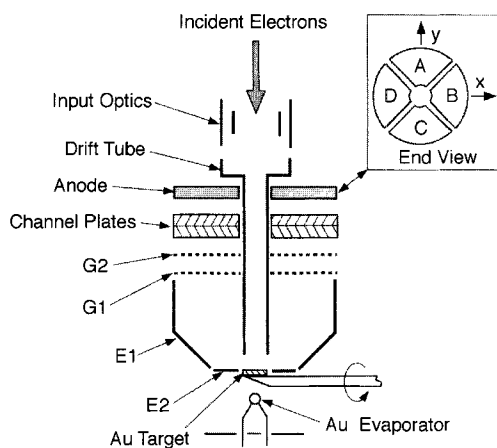
where  $I(\Theta)$  is the angular distribution of back scattered current in the detector and  $S(\Theta)$  is the Sherman function for the detector scattering material at the scattering angle,  $\Theta$ . The Sherman function is a measure of the strength of the spin-dependent scattering in the detector [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.

### 2.3.4 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].

## 2.4 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.

## 2.5 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)

$$\vartheta = \tan^{-1}(M_y/M_x) \quad (9)$$

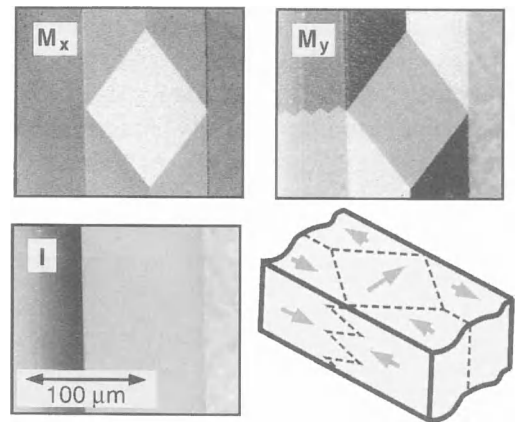
The map of the angle  $\vartheta$  can be displayed using color where the direction is read from an accompanying color wheel. Alternatively, it is sometimes helpful to

visualize the magnetization pattern by using small arrows to create a vector map.

## 2.6 Examples

### 2.6.1 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 (1 0 0) 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



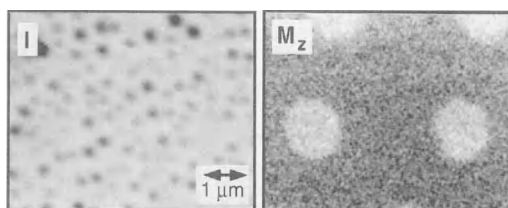
**Figure 4.** SEMPA measurements of the topography,  $I$ , and images of two components of the magnetization,  $M_x$  and  $M_y$ , from an Fe whisker. The depth of focus is demonstrated by domains clearly visible on the top and side of this slightly tilted sample of rectangular cross section.



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.

## 2.6.2 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

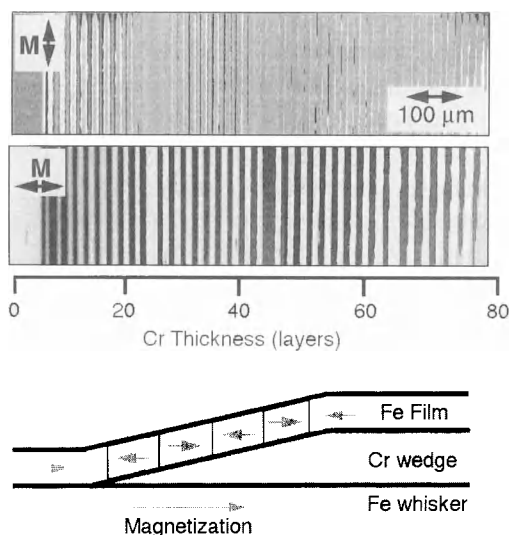


**Figure 5.** The topography,  $I$ , and perpendicular magnetization,  $M_z$ , are imaged from an CoPt magneto-optic recording medium. The round bits are approximately  $1.4\ \mu\text{m}$  in diameter.

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].

## 2.6.3 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.

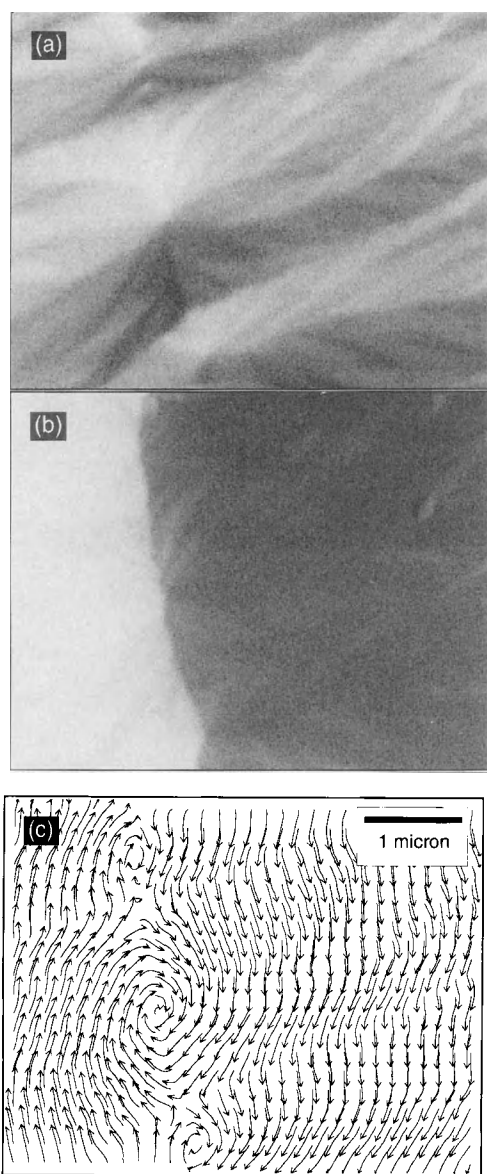
The SEMPA image of the magnetization in the direction of the wedge shows many changes in the magnetization as the Cr increases in thickness to nearly 80 layers (1 layer = 0.14 nm) over the length of the wedge shown in the image, about 0.8 mm. Such an image allows a very precise determination of the period 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.

## 2.6.4 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 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. Vanselow, 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. Hrniciar, *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. Hermseier, 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. Clauber, 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.



## 3 Spin-Polarized Low-Energy Electron Microscopy

### 3.1 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,

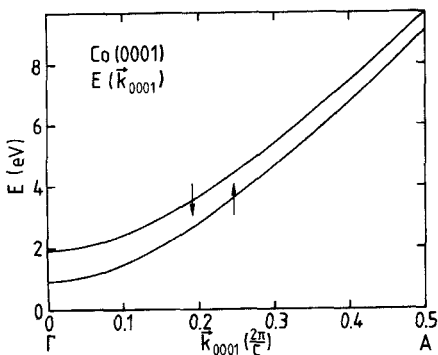
SPLEEM is an excellent method for the study of the correlation between magnetic structure, microstructure, and crystal structure.

### 3.2 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

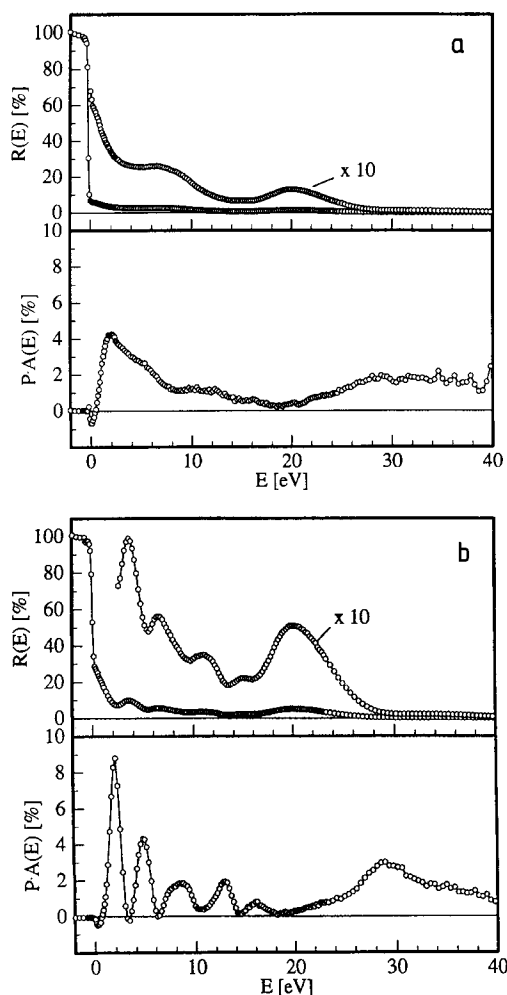
magnetization resulting from the preferred alignment of the target spins  $s_i$ , then  $V_{\text{ex}} \sim \mathbf{P} \cdot \mathbf{M}$ , which causes a  $\mathbf{P} \cdot \mathbf{M}$ -dependent contribution to the reflected intensity  $I = I_0 + I_{\text{ex}}$ . If  $\mathbf{P}$  is reversed,  $V_{\text{ex}}$  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  $\mathbf{k} \equiv \mathbf{k}_\perp$ . Although  $\mathbf{k}_\perp$  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  $A_{\text{ex}}$  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

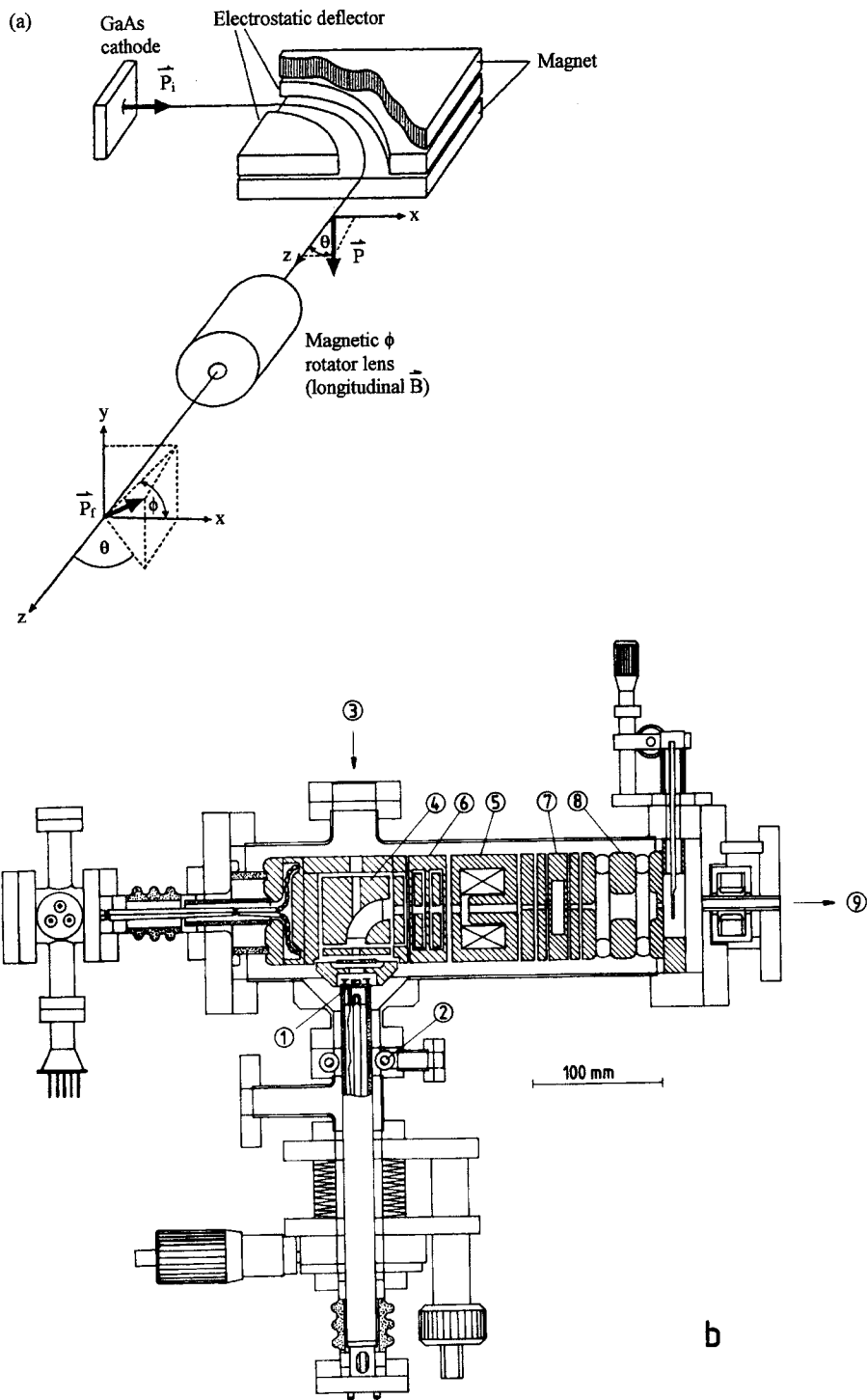
of occupied d orbitals of one spin state [8]. Thus, there is no universal behavior but rather a pronounced material dependence. For cobalt ( $n = 4.1$ )  $\lambda \approx 6 \text{ \AA}$ , for iron ( $n = 3.5$ )  $\lambda = 4.5 \text{ \AA}$ , while for gold ( $n = 5$ )  $\lambda = 10 \text{ \AA}$ . In addition to the limitation by inelastic scattering, the information depth is strongly limited in bandgaps, for example on Co(0001) below 1 eV.

### 3.3 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  $\text{LaB}_6$  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.

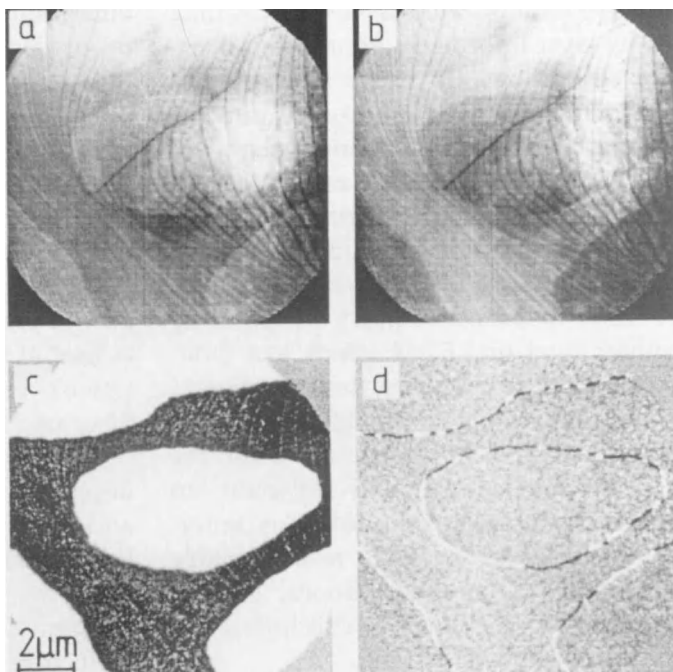
The resolution in SPLEEM is comparable to that in LEEM, but the signal-to-noise ratio is poorer due to the need for image subtraction if purely magnetic images are desired. Inasmuch as  $I_{\text{ex}} \ll I_0$ , longer exposures (in the 1 s range) are needed for a good signal-to-noise ratio. Further 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.

### 3.4 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  $\text{Co}(0001)$  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  $\text{W}(110)$  illustrate this influence. On clean  $\text{W}(110)$ , 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  $\text{W}_2\text{C}$ -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  $\text{W}[110]$  [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  $\text{W}(110)$ , which is in-plane magnetized from the very beginning, cobalt on  $\text{Au}(111)$  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  $\text{Au}(111)$  surface. In particular, when cobalt is grown on an  $\text{Au}(111)$  double layer on  $\text{W}(110)$ , the step distribution can have a strong influence on domain size and shape [16a].

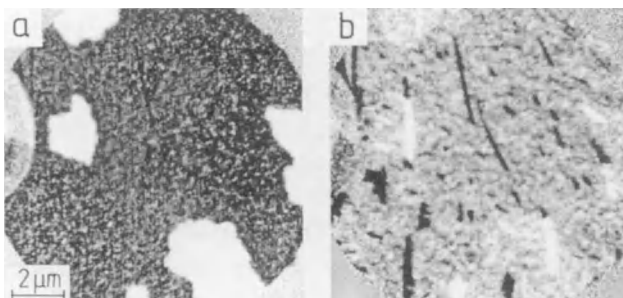


**Figure 4.** Magnetic domain images of a six monolayer thick cobalt layer on W(110). Images with  $P$  (a) parallel and (b) antiparallel to  $M$ , respectively. (c) Difference image between (a) and (b). (d) Difference image between two images with  $P$  perpendicular to  $M$ . Electron energy 2 eV [15].

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

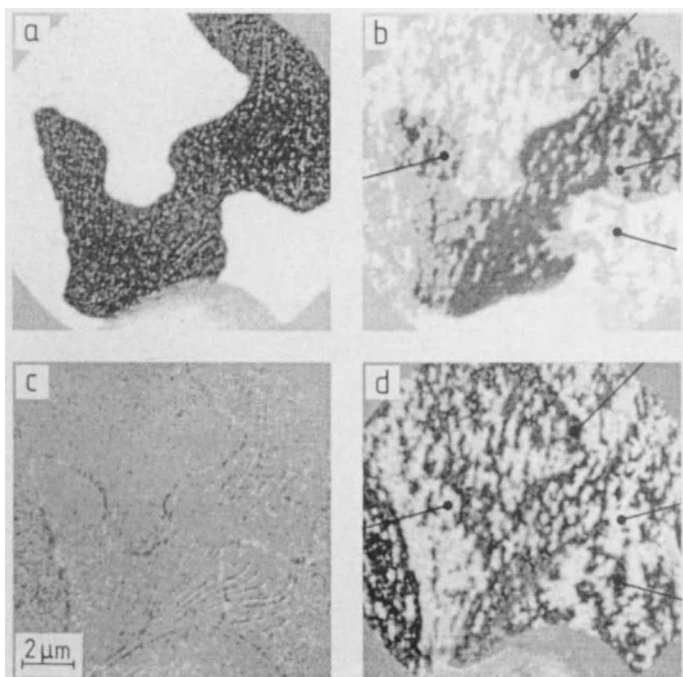


**Figure 5.** Influence of a gold overlayer on the magnetic structure of a 10 monolayer thick cobalt layer on W(110). (a) SPLEEM image of the cobalt layer and (b) SPLEEM image of the same area covered with two monolayers of gold. Electron energy 1.5 eV [16b].

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  $\mathbf{M} \perp \mathbf{P}$  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.  $\mathbf{P} \parallel \mathbf{M}$  in (a) and  $\mathbf{P} \perp \mathbf{M}$  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.

### 3.5 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.

### 3.6 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.

### 3.7 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.



# Emission Methods





# 1 Photoelectron Emission Microscopy

## 1.1 Introduction

Photoelectron emission is the ejection of electrons from matter by the action of light. Once released from matter, photoelectrons become indistinguishable from electrons produced by other means. In particular, photoelectrons emitted into vacuum are deflected by electric and magnetic fields and a beam of photoelectrons can be focused by electrostatic and electromagnetic lenses to form a photoelectron image of a sample. This is the basis of photoelectron microscopy.

## 1.2 Photoelectron Emission

The origins of research into photoelectron emission are traditionally traced to a report by Heinrich Hertz that ultraviolet light increased the length of an electric spark in the spark gap of an electric oscillator [1, 2]. Hertz was in the process of discovering radio waves, which he detected by observing sparks in a secondary electric oscillator which was in resonance with a primary oscillator powered by an induction coil. He noticed that the sparks of the secondary were more intense if the spark gap had direct line-of-sight to the spark

gap of the primary. Using ultraviolet filters, he showed that ultraviolet light produced by the primary spark was responsible for the effect. Furthermore, he showed that the effect was produced by using other sources of ultraviolet light such as a carbon arc lamp.

The next advance was initiated by Hallwachs, who used a gold-leaf electroscope in air to show that negative charge was emitted from zinc exposed to ultraviolet light [3–5]. An insulated, uncharged sample acquired a positive charge, and a negatively charged sample lost its negative charge and became positively charged. However, a strongly positively charged sample retained its charge. Hallwachs suggested that the ultraviolet light caused the emission of negatively electrified particles [5]. A high positive potential on the sample prevented emission of negative particles by electrostatic attraction.

Elster and Geitel showed that photoelectric rays in a cathode ray tube were deflected by a magnet and behaved like other cathode rays [6, 7]. Using crossed electric and magnetic fields, Thomson [8, 9] and Lenard [9, 10] showed that the charge-to-mass ratio [11] of the negative charge carriers of photoemission was identical to that of electrons. Lenard showed that the photoemitted current (i.e., the number of electrons released)

increased with the intensity of light. However, the kinetic energy of the photoelectrons was independent of the light intensity. Moreover, the kinetic energy of the photoelectrons increased with the increase in the frequency of light. Furthermore, photoelectrons were not ejected from a sample unless the frequency of the incident light was greater than a threshold value which was characteristic for each substance.

Einstein [12,13] explained Lenard's results by making the following assumption: matter contains bound electron resonators which emit and absorb electromagnetic waves of a definite period. The resonators emit light with a specific quantum of energy,  $E = h\nu$  (in modern terms) where  $h$  is Planck's constant and  $\nu$  is frequency of light. Unlike water waves which have a decreasing wave amplitude and eventually become imperceptible, electromagnetic waves do not dissipate as they propagate. Rather, the energy emitted,  $h\nu$ , remains intact in a light quantum, which can deliver its entire energy to a single electron. This energy can be transformed into the kinetic energy of an electron. Further, if it is assumed that the electron is bound to matter by some potential energy,  $P$ , the electron cannot be emitted unless the quantum of light possesses at least this binding energy. Perhaps the most convincing evidence evoked by Einstein to show that light does not spread as does a water wave is the fact that far ultraviolet light photoionizes single atoms and molecules in the gas phase [14]. Because gas molecules have cross sections of only a fraction of a few square nanometers, the quantum of light cannot extend over an area much greater than atomic dimensions.

It does not follow from Einstein's theory of photoemission that light is a billiard ball like corpuscle. Light must be a form of motion for the same reason that heat is considered to be a form of motion. In his classic studies of heat, Rumford showed that water could be boiled and heat could be produced indefinitely by friction. Because the heat produced was inexhaustible, Rumford concluded that heat could not be a material substance but must be a form of motion [15]. In taking Rumford's experiment a step further, a sample can be warmed by friction until it glows with a red heat. Like heat, an inexhaustible supply of photons is produced from matter by a hot glowing body. Photons, therefore, cannot be a material substance released from matter, but must be a form of motion. This motion can be transformed into the kinetic energy of a photoelectron.

Our most detailed knowledge of the events involved in photoelectron emission is revealed by photoelectron spectroscopy [16–19]. In this technique, the kinetic energy of photoelectrons can be analyzed at meV resolution. The following fact has emerged: photoelectrons can be ejected from all atoms, molecules, liquids, and solids if photons of sufficient energy to overcome the electron binding energy are absorbed. Electrons within isolated atoms and molecules in the gas phase are bound in atomic and molecular orbitals, each with a characteristic binding energy [18]. For most gases, including organic molecules, first ionization potentials are rarely below 5 eV and more generally start between 9 and 13 eV and require vacuum ultraviolet radiation for photoionization [18]. More strongly bound core electrons need X-ray energies for emission [17]. In

addition to overcoming the potential energy which binds the photoelectron, photoemission may be accompanied by the excitation of the photoionized molecule to higher translational, vibrational, rotational and electronic states [18]. Thus the kinetic energy of the photoelectron can be reduced by these loss mechanisms, and a spread of kinetic energies is expected even if using a monochromatic light source. Conversely, if the molecule initially is in an excited state, the translational, vibrational, rotational, or electronic energy can be added to the kinetic energy of the photoelectron, albeit usually with a lower probability than in the case of energy loss.

Solid metals and semiconductors generally have threshold energies below 6.5 eV, which is much lower than typical isolated molecules. For these materials, near ultraviolet light is sufficient to induce photoemission. Films of strontium, barium, sodium, potassium, rubidium, cesium and lithium, distilled onto silver coatings on glass will emit with visible light, and were used in early photocells [20]. The remarkable material Ag-O-Cs, has a photoelectric threshold near 1400 nm [7] (p.93). It was commercialized in the 1930s, but remains one of the very few materials which photoemits in the near infrared [21]. A roughened surface of silver is essential for emission [21].

Once ejected from an atom within a solid, the photoelectron becomes subject to the same forces which dictate motion of all electrons in matter. Photoelectrons undergo electron diffraction [22] and can propagate only with energies and in crystallographic directions allowed by band theory. Photoelectrons are subject to the same energy loss mechanisms as other

electrons and can lose kinetic energy to lattice vibrations (phonons), plasmons, or electronic excitations. Electrons in semiconductors can be excited across the band gap, and very energetic photoelectrons can induce a secondary electron cascade. The mean-free-path of a photoelectron in matter depends upon these loss mechanisms. For photoelectrons of energies from 1–2 eV, the mean-free-path can be large (100 nm) [19]. This is because electrons at very low energies are not energetic enough to excite loss mechanisms, which require a definite minimum quantum of energy. Electrons of 100 eV energy have much shorter paths of only a few nanometers [19].

To be emitted, a photoelectron originating in the bulk must reach the surface with sufficient energy to overcome the surface potential barrier, which in metals is the photoelectric work function. The work function varies between materials and crystallographic planes of a material. The work function can be drastically changed by absorption of a single monolayer [19]. Photoelectric threshold and photoelectric intensity depend critically on the state of the surface and the nature of the adsorbate. For this reason, it is possible to image single adsorbed monolayers, including submonolayer coverage, by photoelectron microscopy.

### 1.3 Microscopy with Photoelectrons

Photoelectron emission microscopy dates to the very beginnings of electron microscopy. Brüche [23,24] published a photoelectron image of zinc in 1933, just one

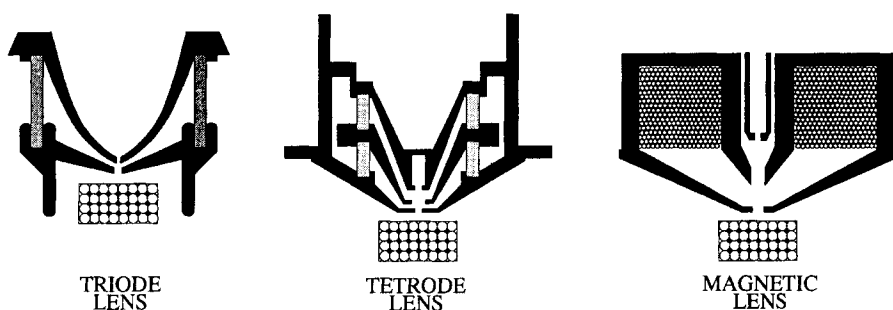
year after publication of the first thermionic electron emission images by Brüche and Johansson [25] and transmission electron images by Knoll and Ruska [26]. Other early papers showed that the technique was sensitive to the adsorption of gases on platinum [27, 28] and nickel [29]. It was proved that metals such as potassium and barium could be deposited in situ onto metallic substrates, and that their adsorption (presumably of a single monolayer) caused a great increase in photoelectron emission by reduction of the work function. Several reviews are written on photoelectron microscopy [30–37]. In traditional photoelectron emission microscopy, flat solid samples are flooded with light of sufficient energy to overcome the work function threshold of the sample. Commonly used is the low pressure mercury lamp (Osram HBO 100W/2). Spectra of Hg [38] show sharp peaks of various intensity at 234.54 nm (5.29 eV, very faint), 239.97 nm (5.17 eV, faint), 244.69 nm (5.07 eV, faint), 246.41 nm (5.03 eV, faint), 248.27 nm (5.00 eV, moderate intensity) and 253.65 nm (4.89 eV, the strongest). Many other peaks exist at lower energies [38–41]. Quartz optics and windows transparent to UV light are necessary. The user needs to be aware that samples with work functions above 5.3 eV will not be imaged using a mercury lamp. Oxide and carbon monoxide contamination often raises the work function of samples to above 5.3 eV. Conversely, ubiquitous carbon contamination can lower the work function to below 4 eV.

Deuterium lamps, which have a maximum energy of 6.9 eV, are used to eject photoelectrons from samples with higher work function such as platinum or oxygen and carbon monoxide adsorbed on

platinum. Sapphire windows transparent to UV are used, and oxygen must be purged from the beam line to avoid production of ozone and adsorption of UV light by oxygen. Synchrotron radiation [42] has been used successfully in photoelectron imaging [36, 43–46]. Attempts have been made to use conventional X-ray sources [47], and lasers [48], but these are not commonly used. Xenon lamps have also worked. In conventional nonscanning photoelectron microscopy, light need not be focused to a small spot size. However, some focusing is desirable to increase intensity. For Hg lamps, spots sizes are typically greater than 3 mm in diameter.

Photoelectrons in vacuum are accelerated and focused with classical electron optics, such as the electrostatic triode and tetrode objective lenses and the magnetic objective lens drawn in Fig. 1. Most instruments employ additional intermediate and projector lenses similar in principle to those used in transmission electron microscopes. The electron optics of such lenses is discussed elsewhere [30, 33, 34, 49–60]. Channel plate image intensifiers and ultra-high vacuum conditions are used in most advanced instruments.

The early triode lens of Brüche and Johansson was made simply of two flat parallel metal plates spaced 1.0 mm apart with a 1.2 mm hole bored into the plate nearest the sample and a 1.0 mm hole in the second [55]. A flat sample, negative with respect to the plates, formed the third electrode of the triode. Focusing was achieved by varying the potential of the plate nearest the sample. In the advanced version (that of E. Bauer [34]) of the triode lens of Fig. 1, the metal plates are sloped to allow light to reach the sample ( $\approx 75^\circ$  off normal) and to allow



**Figure 1.** The three main types of objective lenses used in photoelectron emission microscopy. In the electrostatic triode and tetraode lenses, metallic regions (dark) are separated by ceramic insulators (grey). Flat photoemitting solids are placed  $\sim 4$  mm beneath the lenses and are illuminated with light of sufficient energy to overcome the photoelectric threshold of the sample. Samples can be thick (many centimeters) and can be heated or cooled from beneath.

insertion of a wider ceramic insulator (shaded) to avoid voltage breakdown. The triode lens has a proven resolution of 15 nm [61]. The tetraode electrostatic objective lens developed by Johansson in the early 1930s [55], consisted of three flat parallel plates. The flat sample formed the fourth electrode of the tetraode. In the advanced version (that of W. Engel [88]) of Fig. 1, the plates are again sloped away from the optical axis to allow the beam of light to reach the sample and to allow insertion of wider insulators. Focusing is achieved by varying the potential of the central electrode. Proven resolution is better than 10 nm [62]. A disadvantage of the tetraode is that the entire accelerating field (up to 30 keV over 1–4 mm) is between the sample and nearest electrode. Voltage breakdown between sample and first electrode can occur if the sample is not smooth. In the triode lens, the voltage between sample and focusing electrode is only a few hundred volts over a few millimeters. Resolution is lost, but voltage breakdown is less likely. The early photoelectron microscopes used simple coils of copper as electromagnetic lenses [23, 27].

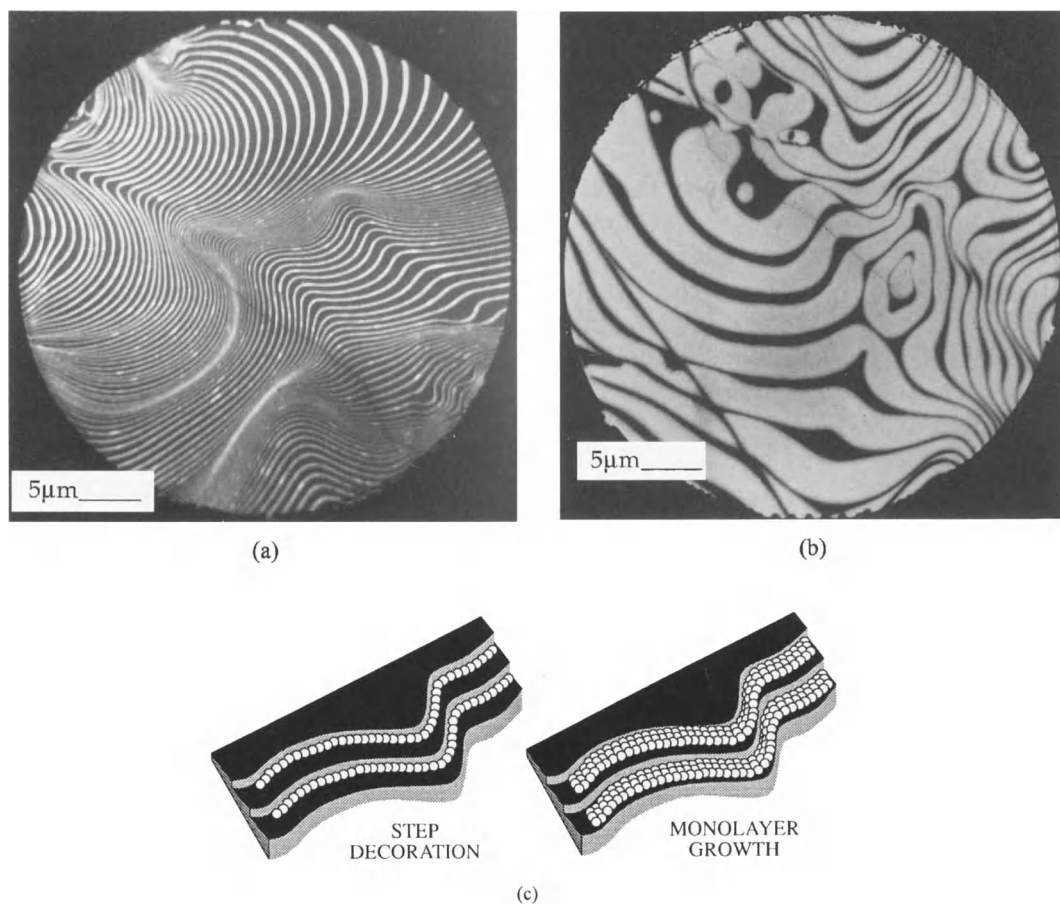
The version of the magnetic lens of Fig. 1, developed by W. Engel, has a resolution of better than 10 nm [34]. It is more robust than the electrostatic lenses, which must best be used in ultrahigh vacuum to avoid deposition of conducting carbonaceous material onto the insulators. The magnetic field, however, can interfere with the study of magnetic materials. The coils of the lens need to be cooled to avoid heating and degassing of the lens. Resolution is limited in practice to chromatic aberrations due to the kinetic energy spread of the emitted photoelectrons [50]. Optimum resolution is achieved only if the energy spread is less than 0.5 eV. Thus best resolution is obtained if the light energy is only a few tenths of an electron volt above threshold energy. For a given instrument, resolution will be highly sample dependent. Although it may seem trivial, mechanical vibrations from pedestrian and motor traffic can easily degrade resolution into the micron range. Stray electromagnetic fields, for example, from fluorescent lights in the same building (even a few floors away) or from ionization gauges, etc., can also degrade resolution into the micron range.

## 1.4 Applications

### 1.4.1 Monolayer Epitaxial Growth

Figure 2a shows the initial stages of the epitaxial growth of the first monolayer of copper evaporated onto an Mo {0 1 1} single crystal surface. To avoid contamination

by residual gas, the system had a base pressure of  $2 \times 10^{-10}$  torr, and the film was grown in situ at a rate of one full monolayer in 2.5 min [61]. The substrate was at 700 K which allowed the copper to diffuse to atomic steps which were present. Figure 2b shows continued growth of the first monolayer in another region having wider terrace widths. The curved step



**Figure 2.** (a) Initial growth of copper monolayer at atomic steps of an Mo {0 1 1} surface at 700 K. Curved steps of Mo arise by step migration during substrate preparation. (b) Continued first monolayer growth of Cu/Mo {0 1 1} under step flow growth conditions. Near perfect contrast in monolayer growth is achieved by choosing a substrate with threshold greater than the light source and an adsorbate with a lower work function. (c) Schematic model of adsorption and growth at atomic steps. In Fig. 2a the white bands are one monolayer thick and  $\sim 250$  copper atoms wide.

structure was due to step migration during preparation and cleaning of the molybdenum. Details are explained elsewhere [61,63–69]. Atomic steps of copper form at the steps of molybdenum, as shown schematically in Fig. 2c, and monolayer growth follows the classical step-flow model of crystal growth [70]. An Hg lamp and triode lens were used. The clean Mo {011} face has a work function above 5.3 eV. Its photoelectric threshold is, therefore, above the maximum transmitted energy of the Hg lamp. No photoelectrons are emitted from the uncovered molybdenum, which appears absolutely black. Adsorbed copper lowers the work function to 4.8 eV [61]. Photoelectrons are emitted from the copper covered areas and appear bright. By choosing a substrate with work function above the maximum energy of the light source, and by choosing an adsorbate with a threshold within the range of the lamp, single monolayers are imaged and with perfect contrast.

The photoelectrons emitted from the copper covered areas can originate from the underlying Mo as well as from the copper atoms. By reducing the surface potential, photoelectrons from Mo can escape if they have kinetic energy, greater than 4.8 eV. Because the energy spread does not exceed 0.5 eV, optimal resolution of 15 nm is achieved with a triode lens. Use of photoelectron microscopy to study epitaxial growth is described elsewhere [61,63–73]. It should be pointed out that the crystal face of Fig. 2 was especially well oriented ( $0.03^\circ$  off axis). This gave exceptionally wide terraces. Samples with terrace widths smaller than the resolution limits will not give such results.

### 1.4.2 Chemical Kinetic Reaction-Diffusion Fronts in Monolayers

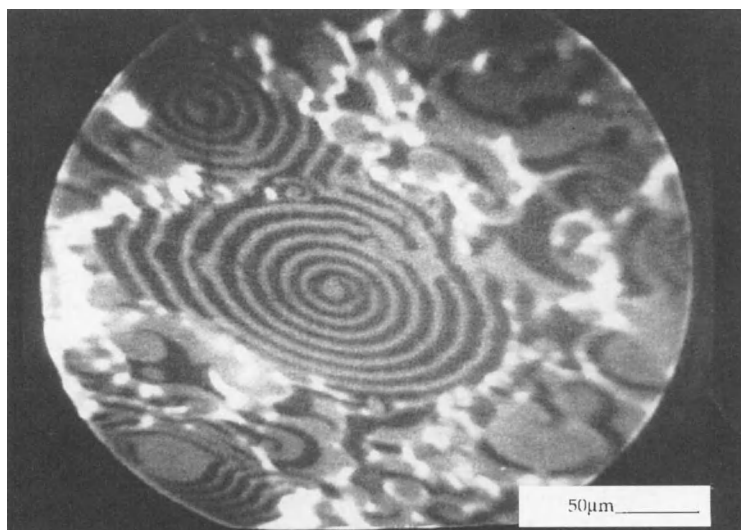
Study of spatio-temporal pattern formation and temporal oscillations in chemical systems is now at the forefront of chemical kinetics and chemical physics [74–78]. Chemical waves, target patterns, spirals, and complex ‘chemical chaos’ are observed in reactive systems and can be mathematically modeled by solving nonlinear differential equations [74–78]. Such phenomena occur in reactive monolayer films on catalyst surfaces and can be observed directly by photoelectron microscopy [79–87]. In fact, no other microscopic technique has revealed these phenomena in monolayers in such detail as has photoelectron microscopy. Figure 3 shows spiral formation seen during the catalytic reaction between carbon monoxide and oxygen on Pt {110} at 435 K. The darker areas are covered with an adsorbed monolayer of oxygen and the gray areas are covered with a monolayer of adsorbed carbon monoxide. Such spirals are also seen in liquid systems and arise from the chemical kinetics and *not*, for example, from screw dislocations on a surface. A still photograph cannot do justice to the myriad of dynamic events recorded at video rates. Slight changes in reaction conditions can completely alter the spatio-temporal patterns. The image was taken using a deuterium lamp and electrostatic tetrode lens [88].

### 1.4.3 Magnetic Materials

Ferromagnetic domains have been imaged by photoelectron microscopy [89–92]. The



**Figure 3.** Distribution image of the reactants CO (grey) and O (dark) on Pt {110} during catalytic reaction. Sample temperature of 435 K and partial pressures of  $3 \times 10^{-4}$  mbar  $O_2$  and  $3.2 \times 10^{-5}$  mbar CO in the reaction chamber. (Courtesy of W. Engel.)

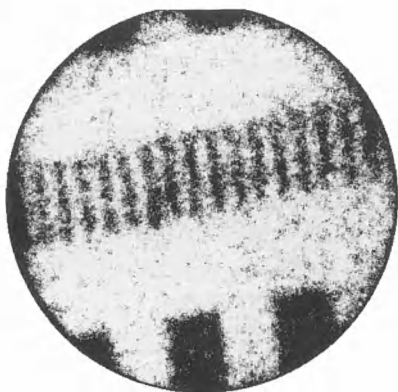


magnetic field of oppositely magnetized domains can split the photoelectron beam into two. As in Lorentz microscopy, one beam can be stopped by an aperture, allowing domains of one orientation to

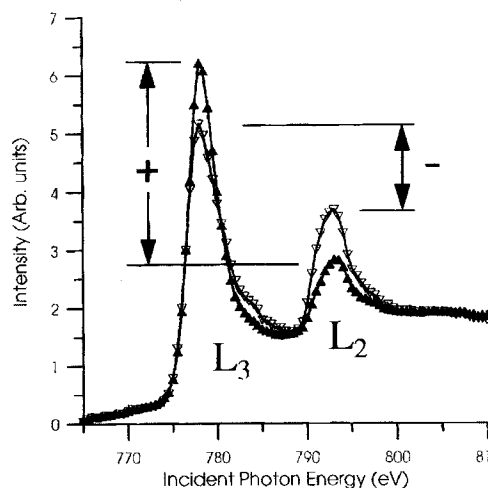
appear bright and those of opposite orientation to be dark.

Using circularly polarized X-rays from a synchrotron source to excite electrons of a specific spin orientation, Tonner et al.

a)



b)



**Figure 4.** (a) High resolution MCXD image of magnetic domains on a Co-Pt magnetic recording disk obtained using circularly polarized X-rays. The image was recorded by subtracting images obtained at the cobalt  $L_3$  and  $L_2$  edges for enhanced contrast. Image field-of-view:  $70 \mu\text{m}$ . (b) X-ray absorption spectra near the Cobalt  $L$  edges of the same sample recorded with circularly polarized X-rays and parallel (filled triangles pointing up) and antiparallel (open triangles pointing down) alignment of photon spin and magnetization vectors, respectively. The contrast in the image is indicated by the difference in height of the + and - arrows. (Courtesy of D. Dunham and B. P. Tonner.)

[93,94] have imaged magnetized regions of a magnetic recording disc as shown in Fig. 4a. The image was correlated with photoelectron spectra (Fig. 4b). The use of electrostatic objective lenses is one great advantage of photoelectron microscopy in the study of magnetic materials because the field of a magnetic objective lens does not interfere with domain structure of the magnetic material.

## 1.5 Choice and Preparation of Samples

Because the sample is actually part of the objective lens, samples should be flat to avoid field distortion. Single crystal surfaces, polished metallurgical specimens, silicon wafers and integrated circuits work well. Samples need to conduct well enough to replace the electrons emitted in order to avoid charging. Metals and semiconductors work well. Samples which are normally classified as insulators have been imaged if they become photoconducting or can be heated until they conduct [95]. Insulators can, of course, be coated with a conducting layer if surface sensitivity is not necessary. For controlled experiments and monolayer studies, ultra high vacuum compatibility is essential. Sample surfaces need to be cleaned in situ by heating, chemical reaction or sputtering. Although photoelectron emission microscopy can be used for routine postmortem examination of samples, best results are obtained by performing experiments in situ with the sample in the observational position. These include epitaxial growth and desorption of monolayers, chemical vapor

deposition [96], reaction-diffusion phenomena, etc. Samples can be heated and cooled. Thick samples (1–3 cm) and even entire (10 cm) silicon wafers can be used.

## 1.6 References

- [1] H. Hertz, *Ann. Phys.* **1887**, 31, 983.
- [2] H. Hertz, *Electric Waves*, Dover, New York **1962**, p. 63.
- [3] W. Hallwachs, *Ann. Phys.* **1888**, 33, 301.
- [4] S. P. Thompson, *Light Visible and Invisible*, MacMillan, London **1928**, p. 181.
- [5] T. H. Osgood, A. E. Ruark, E. Hutchisson, *Atoms, Radiation and Nuclei*, Wiley & Sons, New York **1964**, p. 78.
- [6] J. Elster, H. Geitel, *Ann. Phys.* **1890**, 41, 166.
- [7] V. K. Zworykin, E. D. Wilson, *Photocells and Their Application*, Wiley, New York, **1932**, p. 9.
- [8] J. J. Thomson, *Phil. Mag.* **1899**, 48, 547.
- [9] A. L. Hughes, L. A. DuBridge, *Photoelectric Phenomena*, McGraw-Hill, New York **1932**, p. 4.
- [10] P. Lenard, *Ann. Phys.* **1902**, 8, 169.
- [11] O. Klemperer, *Electron Physics, The Physics of the Free Electron*, Academic Press, New York **1959**, p. 121.
- [12] A. Einstein, *Ann. Phys.* **1905**, 17, 132.
- [13] A. B. Arons, M. B. Peppard, *Am. J. Phys.* **1965**, 33, 367.
- [14] A. von Engel, *Ionized Gases*, Oxford University Press, Oxford **1955**, p. 60.
- [15] B. Thompson, Count Rumford in *A Source Book in Physics* (Ed.: W. F. Magie), Harvard University Press, Cambridge, MA **1969**, p. 161.
- [16] K. Siegbahn, C. Nordling, A. Fahlman, R. Nordberg, K. Hamrin, J. Hedman, G. Johanson, T. Bergmark, S. Karlsson, I. Lindgren, B. Lindberg, *Nova Acta Regie Soc. Sci. Upsal.* **1967**, 20, 4.
- [17] K. Siegbahn in *Electron Spectroscopy* (Ed.: D. A. Shirley), North-Holland, Amsterdam **1972**, p. 15.
- [18] C. R. Brundle, M. B. Robin in *Determination of Organic Structures by Physical Methods* (Eds.: F. C. Nachod, J. J. Zuckerman), Academic Press, New York **1971**, p. 1.
- [19] D. P. Woodruff, T. A. Delchar, *Modern Techniques of Surface Science*, 2nd ed., Cambridge University Press, Cambridge **1994**, Chapter 3.
- [20] N. R. Campbell, D. Ritchie, *Photoelectric Cells*, Pitman, London **1929**, p. 4.

- [21] A. H. Sommer, *Photoemissive Materials*, Krieger, Huntington, New York **1980**, Chapter 10.
- [22] D. P. Woodruff, A. M. Bradshaw, *Rep. Prog. Phys.* **1994**, *57*, 1029.
- [23] E. Brüche, *Z. Phys.* **1933**, *86*, 448.
- [24] C. Ramsauer, *Elektronenmikroskopie*, Springer, Berlin **1943**, p. 32.
- [25] E. Brüche, H. Johannson, *Naturwissenschaften* **1932**, *20*, 353.
- [26] M. Knoll, E. Ruska, *Z. Phys.* **1932**, *78*, 318.
- [27] J. Pohl, *Physik. Z.* **1934**, *35*, 1003.
- [28] H. Mahl, J. Pohl, *Z. Tech. Phys.* **1935**, *16*, 219.
- [29] H. Gross, G. Seitz, *Z. Physik* **1937**, *105*, 734.
- [30] G. Möllenstedt, F. Lenz, *Advances in Electronics*, Vol. 13 **1963**, p. 251.
- [31] L. Wegmann, *J. Microsc.* **1972**, *96*, 1.
- [32] R. A. Schwarzer, *Microsc. Acta* **1981**, *84*, 51.
- [33] O. H. Griffith, G. F. Rempfer, *Advances in Optical and Electron Microscopy*, Vol. 10 **1987**, p. 269.
- [34] E. Bauer, W. Telieps in *Surface and Interface Characterization by Electron Optical Methods* (Eds.: A. Howie and U. Valdré), Plenum, New York **1988**, p. 195.
- [35] O. H. Griffith, W. Engel, *Ultramicroscopy* **1991**, *36*, 1.
- [36] M. Munschau, *Synchrotron Radiation News* **1991**, *4*, 29.
- [37] M. E. Kordes in *CRC Handbook of Surface Imaging and Visualization* (Ed.: A. T. Hubbard) CRC Press, Boca Raton **1995**, Chapter 56.
- [38] G. P. Harnwell, J. J. Livingood, *Experimental Atomic Physics*, McGraw-Hill, New York **1933**, 279.
- [39] W. Koch, *Z. Phys.* **1958**, *152*, 1.
- [40] L. Wegman, *Z. Angew. Phys.* **1969**, *27*, 199.
- [41] J. G. Eden, *Photochemical Vapor Deposition*, Wiley, New York **1992**, p. 59.
- [42] G. Margaritondo, *Introduction to Synchrotron Radiation*, Oxford University Press, Oxford **1988**.
- [43] B. P. Tonner, G. R. Harp, *Rev. Sci. Instrum.* **1988**, *59*, 853.
- [44] B. P. Tonner, G. R. Harp, *J. Vac. Sci. Technol. A* **1989**, *7*, 1.
- [45] B. P. Tonner, *Nucl. Instr. & Methods A* **1990**, *291*, 60.
- [46] F. Polack, S. Lowenthal in *X-ray Microscopy* (Eds.: G. Schmahl, D. Rudolph), Springer, Berlin **1984**, p. 251.
- [47] L. Y. Huang, *Z. Phys.* **1957**, *149*, 225.
- [48] M. D. Jones, G. A. Massey, D. L. Habliston, O. H. Griffith, *Beitr. elektronenmikroskop. Direktabb. Oberfl.* **1979**, *12*(2), 177.
- [49] E. Bauer, *Rep. Prog. Phys.* **1994**, *57*, 895.
- [50] L. H. Veneklasen, *Rev. Sci. Instrum.* **1992**, *63*, 5513.
- [51] C. E. Hall, *Introduction to Electron Microscopy*, McGraw-Hill, New York **1953**.
- [52] V. E. Cosslet, *Introduction to Electron Optics*, 2nd ed., Oxford University Press, Oxford **1950**.
- [53] O. Klemperer, *Electron Optics*, 2nd ed., Cambridge University Press, Cambridge **1953**.
- [54] B. Paszkowski, *Electron Optics*, American Elsevier, New York **1968**, 199.
- [55] E. F. Burton, W. H. Kohl, *The Electron Microscope*, 2nd ed., Reinhold, New York **1948**.
- [56] R. A. Schwarzer, *Beitr. elektronenmikroskop. Direktabb. Oberfl.* **1979**, *12*(2), 3.
- [57] Y. Uchikawa, S. Maruse, *J. Electron Microscopy* **1970**, *19*, 12.
- [58] G. N. Plass, *J. Appl. Phys.* **1942**, *13*, 49.
- [59] H. Liebl, *Optik* **1979**, *53*, 1.
- [60] V. K. Zworykin, G. A. Morton, E. G. Ramberg, J. Hillier, A. W. Vance, *Electron Optics and the Electron Microscope*, Wiley, New York **1945**.
- [61] M. Munschau, E. Bauer, W. Swiech, *Surf. Sci.* **1988**, *203*, 412.
- [62] G. F. Rempfer, W. P. Skoczylas, O. H. Griffith, *Ultramicroscopy* **1991**, *36*, 196.
- [63] M. Munschau, E. Bauer, W. Telieps, W. Swiech, *Surf. Sci.* **1989**, *213*, 381.
- [64] E. Bauer, M. Munschau, W. Swiech, W. Telieps, *Ultramicroscopy* **1989**, *31*, 49.
- [65] M. Munschau, E. Bauer, W. Swiech, *J. Appl. Phys.* **1989**, *65*, 581.
- [66] M. Munschau, E. Bauer, W. Swiech, *Phil. Mag. A* **1989**, *59*, 217.
- [67] M. Munschau, *Eur. Microsc. & Analysis* **1990**, (Sept.) 17.
- [68] M. Munschau, E. Bauer, W. Swiech, *Metall. Trans. A* **1991**, *22A*, 1311.
- [69] M. Munschau, E. Bauer, W. Swiech, *Mat. Res. Soc. Symp. Proc.* **1993**, *280*, 333.
- [70] W. K. Burton, N. Cabrera, F. C. Frank, *Phil. Trans. R. Soc. London* **1951**, *A243*, 299.
- [71] E. Bauer, M. Munschau, W. Swiech, W. Telieps, *Inst. Phys. Conf. Ser.* **1988**, Vol. 1, No. 93, p. 213.
- [72] E. Bauer, M. Munschau, W. Swiech, W. Telieps in *Evaluation of Advanced Semiconductor Materials by Electron Microscopy* (Ed.: D. Cherns), Plenum, New York **1990**, p. 283.
- [73] M. Munschau, E. Bauer, W. Telieps, W. Swiech, *J. Appl. Phys.* **1989**, *65*, 4747.
- [74] S. K. Scott, *Oscillations, Waves and Chaos in Chemical Kinetics*, Oxford University Press, Oxford **1994**.
- [75] S. K. Scott, *Chemical Chaos*, Oxford University Press, Oxford **1993**.
- [76] P. Gray, S. K. Scott, *Chemical Oscillations and Instabilities*, Oxford University Press, Oxford **1990**.

- [77] P. Grindrod, *Patterns and Waves*, Oxford University Press, Oxford **1991**.
- [78] V. A. Vasilev, Yu. M. Romanovski, D. S. Chernavskii, V. G. Yakhno, *Autowave Processes in Kinetic Systems*, Deutscher Verlag der Wissenschaften, Berlin **1987**.
- [79] H. H. Rotermund, W. Engel, S. Jakubith, A. von Oertzen, G. Ertl, *Ultramicroscopy* **1991**, 36, 164.
- [80] H. H. Rotermund, W. Engel, M. Kordesch, G. Ertl, *Nature* **1990**, 343, 355.
- [81] G. Ertl, *Science* **1991**, 254, 1750.
- [82] S. Jakubith, H. H. Rotermund, W. Engel, A. von Oertzen, G. Ertl, *Phys. Rev. Lett.* **1990**, 65, 3013.
- [83] H. H. Rotermund, S. Jakubith, A. von Oertzen, G. Ertl, *Phys. Rev. Lett.* **1991**, 66, 3083.
- [84] M. Bär, S. Nettesheim, H. H. Rotermund, M. Eiswirth, G. Ertl, *Phys. Rev. Lett.* **1995**, 74, 1246.
- [85] M. D. Graham, I. G. Kevrekidis, K. Asakura, J. Lauterbach, K. Krischer, H. H. Rotermund, G. Ertl, *Science* **1994**, 264, 80.
- [86] M. Munschau, M. E. Kordesch, B. Rausenberger, W. Engel, A. M. Bradshaw, E. Zeitler, *Surface Sci.* **1990**, 227, 246.
- [87] M. Munschau, B. Rausenberger, *Platinum Metals Rev.* **1991**, 35, 188.
- [88] W. Engel, M. E. Kordesch, H. H. Rotermund, S. Kubala, A. von Oertzen, *Ultramicroscopy* **1991**, 36, 148.
- [89] G. V. Spivak, T. N. Dombrovskaya, *Soviet Physics-Doklady* **1956**, 106, 9.
- [90] G. V. Spivak, T. N. Dombrovskaya, N. N. Sedov, *Soviet Physics-Doklady* **1957**, 107, 120.
- [91] K. Schur, G. Pfefferkorn, *Beitr. elektronenmikroskop. Direktabb. Oberfl.* **1971**, 4(2), 235.
- [92] M. Munschau, J. Romanowicz, *J. Vac. Sci. Technol. A*, to be published.
- [93] B. P. Tonner, D. Dunham, J. Zhang, W. L. O'Brien, M. Samant, D. Weller, B. D. Hermsmeier, J. Stöhr, *Nucl. Instr. & Methods A* **1994**, 347, 142.
- [94] J. Stöhr, Y. Wu, B. D. Hermsmeier, G. R. Harp, B. P. Tonner, *Science* **1993**, 259, 658.
- [95] L. Weber, H. R. Oswald, E. Grauer-Carstensen, *Micron* **1975**, 6, 129.
- [96] M. E. Kordesch, *J. Vac. Sci. Technol. A*, **1995**, 13, 1517.



## 2 Field Emission and Field Ion Microscopy (Including Atom Probe FIM)

### 2.1 Field Emission Microscopy

The field emission microscope (FEM) is a projection microscope, in which an enlarged image of a sharply pointed object is produced on a fluorescent screen (Fig. 1). The image is produced in vacuum, by applying a strong negative potential to the specimen, which must be an electrical conductor. At sufficiently high local field strengths, electrons can escape from the specimen by tunneling through the deformed surface potential barrier (Fig. 2). The intensity of the field emission current  $I$  is a function of the local field strength  $F$  and the work function of the surface  $\phi$ , and is expressed mathematically by the Fowler–Nordheim equation:

$$\frac{I}{F^2} = a \exp\left(-\frac{b\phi^{\frac{3}{2}}}{F}\right) \quad (1)$$

where  $a$  and  $b$  are effectively constants.

For electron tunneling to occur at a measurable rate, field strengths of  $1\text{--}5\text{ V nm}^{-1}$  are necessary. In order to obtain such fields with the application of only a few thousand volts, a sharp pointed geometry for the specimen is essential. For the case of an isolated charged sphere, the field

strength at the surface is given by:

$$F = \frac{V_0}{r} \quad (2)$$

where  $V_0$  is the applied voltage, and  $r$  is the radius of curvature. For a needle-like specimen, which can be approximated as a sphere on a cone, the apex field strength may be expressed as:

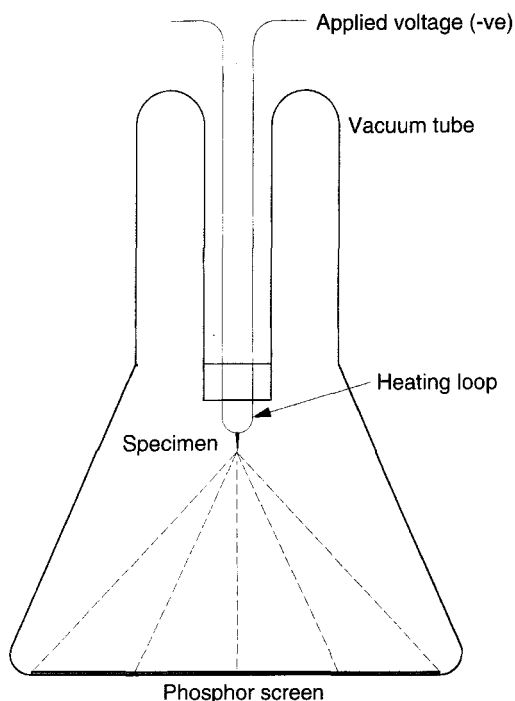
$$F = \frac{V_0}{kr} \quad (3)$$

where  $k$  is a constant depending on the half-angle of the conical shank, and is approximately equal to 5 for typical field emission specimens.

The trajectories of the electrons leaving the surface are approximately radial. Hence the magnification of the image is given by:

$$M = \frac{D}{\beta r} \quad (4)$$

where  $D$  is the specimen-to-screen distance, and  $\beta$  is an image compression factor (approximately 1.5) which takes into account the effects of the specimen shank. Thus, for specimens of end radii below  $0.5\text{ }\mu\text{m}$ , and specimen-to-screen distances of the order of  $100\text{ mm}$ , magnifications of the order of  $100\,000$  are obtained.

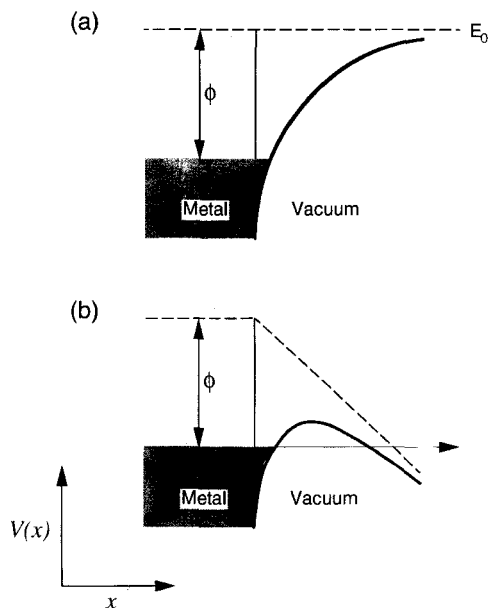


**Figure 1.** Schematic diagram of an FEM. A needle-like specimen is mounted along the axis of a vacuum tube, across the end of which is situated a fluorescent screen. Application of a high negative voltage to the specimen results in electron emission, and the production of a greatly enlarged image of the apex region on the microscope screen.

The resolution of the FEM image is limited by two main factors, the finite wavelengths of the emitted electrons, and the components of their momentum transverse to the emitting surface. The resolutions can be expressed by an equation of the form [1]:

$$\delta = C_1 \sqrt{r} \sqrt{\left( \frac{C_2}{\sqrt{V}} + \frac{C_3}{\sqrt{\phi}} \right)} \quad (5)$$

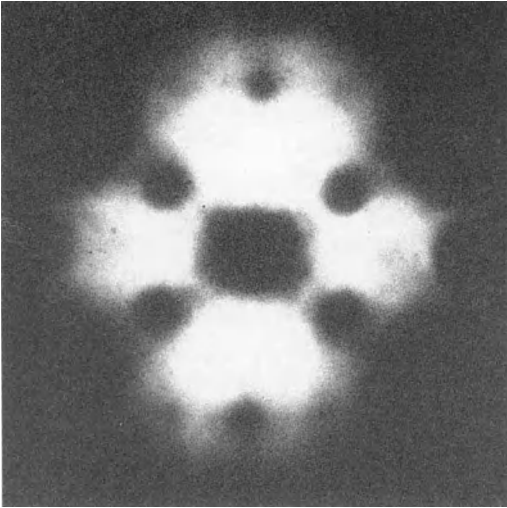
where  $C_1$ ,  $C_2$ , and  $C_3$  are constants. The first term in the brackets is related to the electron wavelength, and the second term arises from the distribution of transverse momentum components of the emitted



**Figure 2.** One-dimensional potential energy diagram for an electron on a metal surface: (a) under field-free conditions; (b) in the presence of a strong negative field. The application of the field produces a narrow energy barrier, through which tunneling of electrons can occur, resulting in field emission.

electrons. Equation (5) shows that  $\delta$  is independent of the specimen-to-screen distance, and varies with the square root of tip radius.

It should be noted that the electrons which are emitted are predominantly from energy levels at or near the Fermi surface. This is because these particles face the smallest barrier to escape (Fig. 2). Because of the quantized nature of the electron energy levels in solids, these electrons have inherently rather high energies (corresponding, in Maxwell-Boltzmann terms, to temperatures of many thousands of degrees kelvin). Hence, transverse momentum components tend to be high, and the point-to-point resolution of the FEM is limited to around 2 nm. It is not possible to improve on this resolution limit by, for



**Figure 3.** FEM image of a clean, thermally annealed tungsten specimen. The (011) plane is at the center.

example, cooling the specimen, because the Pauli exclusion principle prevents electron transitions to lower energy states.

Contrast in FEM images may arise from local variations in work function, field strength, or (very commonly) a combination of both. For example, Fig. 3 shows an FEM image of a single crystal tungsten specimen, prepared from tungsten wire by an electropolishing process, and then thermally cleaned by 'flashing' to a temperature in excess of 2000 K. The heating operation produces a rounded, almost hemispherical end-form on the specimen, with flat facets corresponding to low-index crystallographic planes such as {011}, {002}, and {112}. Atomically flat, low-index planes tend to have high work functions, and thus appear dark in the image, whereas the atomically rough, high-index planes have lower work functions and generally appear brighter.

Because of the radial nature of electron trajectories in the FEM, the symmetry of

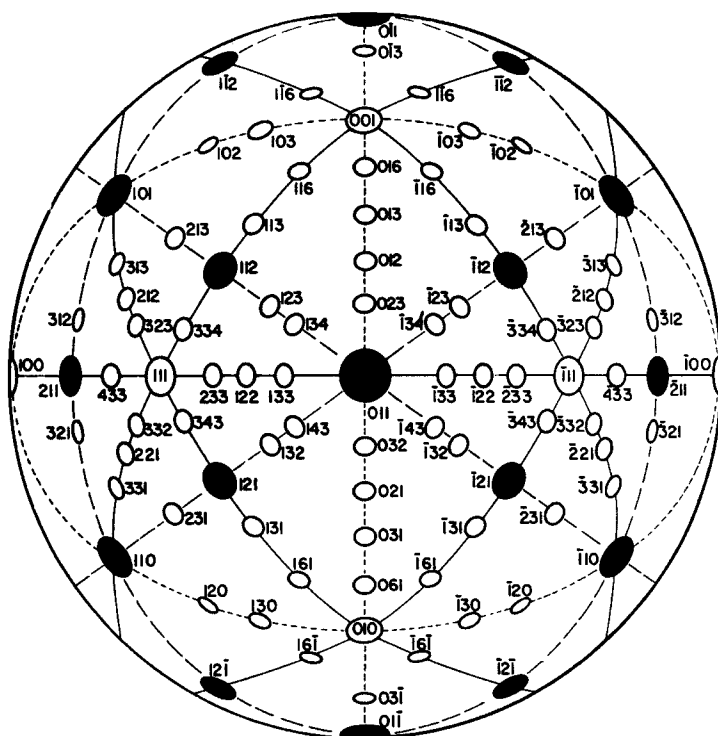
the specimen is preserved in the image, which resembles a stereographic projection of the original surface. Hence crystallographic indexing can usually be carried out very straightforwardly, by reference to standard projection maps. Figure 4 shows a [011] oriented projection corresponding to the orientation of the specimen in Fig. 3.

The main applications of the FEM have been to the study of adsorption and surface diffusion on metal and semiconductor surfaces, the investigation of thin film nucleation and growth processes, and the study of surface reactions such as oxidation and carburization. Field emitters are also used as electron sources of high intensity and good coherence. These have found wide application in other forms of microscopy, such as scanning electron (SEM), transmission electron (TEM), and scanning transmission electron microscopies (STEM), and electron holography. The main limitations of the field emission microscopy technique are its limited spatial resolution, and the absence of any simple means of chemical analysis of the specimen surface. These limitations were largely overcome by the introduction of the field ion microscope and the atom probe mass spectrometer, which are described below.

## 2.2 Field Ion Microscopy

The field ion microscope (FIM) allows the individual atoms on the surface of a solid to be imaged. It also permits the study of the three-dimensional structure of a material. Successive atom layers can be stripped from the surface in the form of ions, by the process of field evaporation. The material





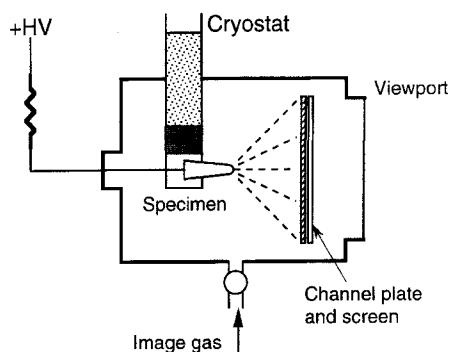
**Figure 4.** [011] centered projection of a cubic crystal structure. The orientation corresponds to the micrograph shown in Fig. 3. (Courtesy of Academic Press.)

removed from the surface by field evaporation can be chemically identified by linking the microscope to a time-of-flight mass spectrometer of single-particle sensitivity. This is known as an atom probe (AP). Nearly all metals and semiconductors can be examined by FIM/AP methods, but insulators and mechanically soft substances such as organic and biological materials are not amenable to study with this technique.

### 2.2.1 Principle of the Field Ion Microscope

The basic design of the FIM (Fig. 5) is similar to that of the FEM. There are some important differences, however. The

needle-like specimen is usually made sharper (50–100 nm end radius), and is kept at low temperature (about 20–70 K). A high positive voltage (5–20 kV) is applied, rather than a negative one. Also, a small

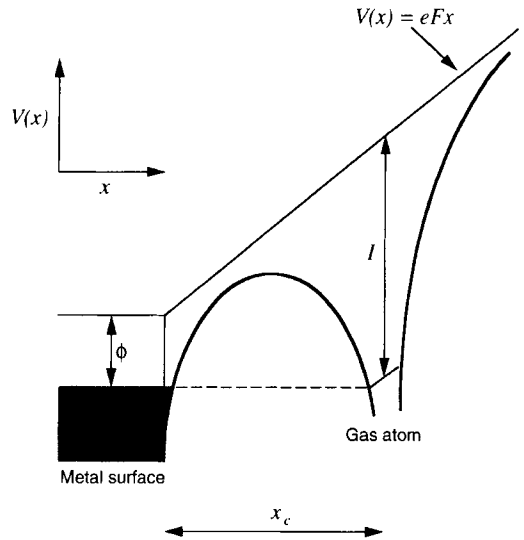


**Figure 5.** Schematic diagram of an FIM. Note that the specimen, which is typically 10 mm long and has an end radius of 100 nm, is not shown to scale.

amount of inert gas, such as helium or neon, is admitted to the microscope vacuum chamber. If the field is high enough, gas atoms which pass close to the specimen apex can become ionized. The resulting positive ions are accelerated away from the specimen along approximately radial paths, and strike a micro-channel plate (MCP) image converter assembly. The MCP converts the incident ion beams to intensified cascades of secondary electrons, which are then accelerated towards a phosphor screen, where the final image is produced. As in the FEM, the radial projection of the ions produces a very high magnification of the image. The use of ions, rather than electrons, to form the image effectively removes the diffraction limit on the spatial resolution. Also, the energies of gas atoms are not subject to the same quantization constraints as for electrons in solids, so cooling the specimen is beneficial in reducing the transverse momentum components of the imaging particles. The result is that the FIM has a point-to-point spatial resolution approaching 0.2 nm, which is sufficient to image individual atoms on the specimen surface.

### 2.2.2 Field Ionization

The field ionization process involves quantum mechanical tunneling of an electron from the upper level of the image gas atom, into a vacant energy level in the specimen. Ground state ionization energies for the inert gases typically used for field ion imaging are high (10–25 eV). The applied electric field distorts the potential well around the atom, as shown in Fig. 6.



**Figure 6.** Electron energy diagram for a gas atom above a metal surface under the influence of a large electric field. The field distortion of the potential well associated with the gas atom produces a barrier through which the uppermost electron in the gas atom can tunnel into an empty state in the metal. At sufficiently high fields, the barrier is thin enough for a significant ionization probability.

A narrow (approximately 0.5 nm wide) potential barrier is thus formed, through which electron tunneling can occur, resulting in ionization of the gas atom. Two conditions must be satisfied for there to be a significant probability of electron tunneling. The first condition is that the gas atom must be sufficiently distant from the specimen surface for the energy level in the gas atom to be above the Fermi level in the specimen, so that an empty final state exists. The critical distance  $x_c$  at which ionization can take place is given approximately by:

$$x_c = \frac{I - \phi}{eF} \quad (6)$$

where  $I$  is the ionization energy of the gas atom,  $\phi$  is the work function of the metallic specimen (more generally the Fermi

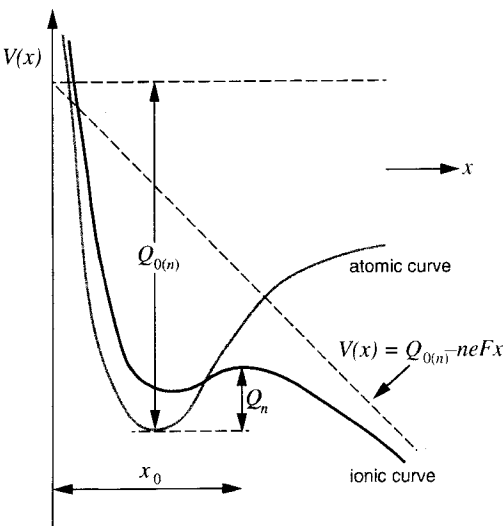
**Table 1.** Ionization potentials and imaging fields for selected gases (after Müller and Tsong [2])

Gas	Ionization potential (eV)	Best imaging field (V nm <sup>-1</sup> )
Helium	24.6	44
Neon	21.6	35
Hydrogen	15.4	22
Argon	15.8	22
Krypton	14.0	15
Xenon	12.1	12

energy),  $F$  is the applied field, and  $e$  is the electronic charge. The second condition is that the potential barrier must be narrow enough to permit tunneling. In order for this to occur, the gas atom must be close to the specimen surface; in general, the probability for tunneling falls rapidly at distances beyond about 0.5 nm. These two conditions can only be met with extremely strong fields, above  $10^{10}$  V m<sup>-1</sup>, that is at fields approaching an order of magnitude higher than those needed for field electron emission. A specimen with an apex radius of 50 nm will require voltages in the range 5–10 kV to attain the required ionization fields of common image gases such as neon or helium. Table 1 shows the ionization fields required for a number of gases used in field ion microscopy.

**2.2.3 Field Evaporation**

Under the intense fields produced in the FIM, atoms of the specimen material can also become ionized, and this process is termed field evaporation. It is usually a thermally activated process, except at the temperatures below approximately 50 K, where it is believed that ion tunneling may occur. The field reduces the energy barrier for desorption to the point that removal of



**Figure 7.** Schematic potential energy diagrams for an atom and an ion above the surface of a metal specimen in the presence of a large electric field. The field lowers the ion energy, so that at some distance from the specimen surface it is energetically favorable for ionization to occur. This crossover leads to a reduced activation barrier (energy  $Q_n$ ) for evaporation of a surface atom.

material occurs even at cryogenic temperatures. Potential energy diagrams for a neutral atom and a positive ion at the surface of an FIM specimen are shown schematically in Fig. 7. Under zero-field conditions, the energy  $Q_{0(n)}$  required to remove a surface atom as an  $n$ -fold charged ion is given by:

$$Q_{0(n)} = \Lambda + \sum_{k=1}^n I_k - n\phi \tag{7}$$

where  $\Lambda$  is the sublimation energy,  $I_k$  is the  $k$ th ionization energy of the atom, and  $\phi$  is the work function of the material. The presence of the electric field  $F$  at the specimen surface reduces the potential energy of the ion at a distance  $x$  from the surface by  $neFx$ , so that the ionic and atomic curves intersect, as shown in Fig. 7. In

the simplest model of field evaporation, repulsion terms in the ionic potential are ignored, and the potential energy of the ion in the region of the crossover is expressed approximately by the sum of an image potential and a field term:

$$V(x) = Q_{0(n)} - \frac{n^2 e^2}{16\pi\epsilon_0 x} - neFx \quad (8)$$

where  $\epsilon_0$  is the permittivity of free space. This potential curve has a maximum defined by  $dV/dx = 0$  which occurs at a distance  $x_0$  from the specimen surface

$$x_0 = \frac{1}{4} \sqrt{\frac{ne}{\pi\epsilon_0 F}} \quad (9)$$

Assuming that the neutral and ionic potential energy curves cross closer to the specimen than  $x_0$ , the maximum in the ionic potential (the 'image hump') represents the energy barrier to the removal of a surface atom as an  $n$ -fold charged positive ion. The activation energy  $Q_n$  is therefore given by:

$$Q_n = Q_{0(n)} - \sqrt{\frac{n^3 e^3 F}{4\pi\epsilon_0}} \quad (10)$$

Approximate fields required for evaporation at cryogenic temperatures are readily calculated by setting  $Q_n = 0$ . For any given element, the fields calculated for the evaporation of singly, doubly, and triply charged ions will generally be different. The evaporation field of that material is taken to be the lowest of these values. Evaporation fields calculated using this method are shown in Table 2 for a number of elements, together with the predicted charge states of the ions generated. Despite the simplicity of the model, the agreement with the experiment is remarkably good. For a fuller treatment of this and other

models of field evaporation, the reader is referred to Tsong [3] or Miller et al. [4].

## 2.2.4 Image Formation, Magnification, and Resolution

The field ion image is produced by a combination of several high-field effects, as indicated schematically in Fig. 8. Atoms of the image gas which arrive in the apex region are polarized and attracted towards the specimen. Their initial collision with the surface results in partial thermal accommodation to the temperature of the specimen, so that the atom is trapped at the surface. The atom executes a form of hopping motion across the surface, and multiple impacts lead to further thermal accommodation. The supply of gas atoms directly to the specimen apex, as shown in Fig. 8, is enhanced by gas collected on the shank, which diffuses preferentially towards the apex under the influence of the applied field. Some of the gas atoms which diffuse over the apex of the specimen become bound at specific locations on the surface, principally those at which the field is highest, such as kink site atoms. This field adsorbed layer increases the efficiency of thermal accommodation for other gas atoms incident on the surface. Gas atoms which pass over one of these high field sites may be field ionized, and then travel away from the specimen towards the image screen, to build up the field ion image.

The FIM image (Fig. 9) shows individual atoms on the specimen surface as distinct, well-resolved spots, arranged in patterns of concentric circles. As the process of field ionization is strongly field

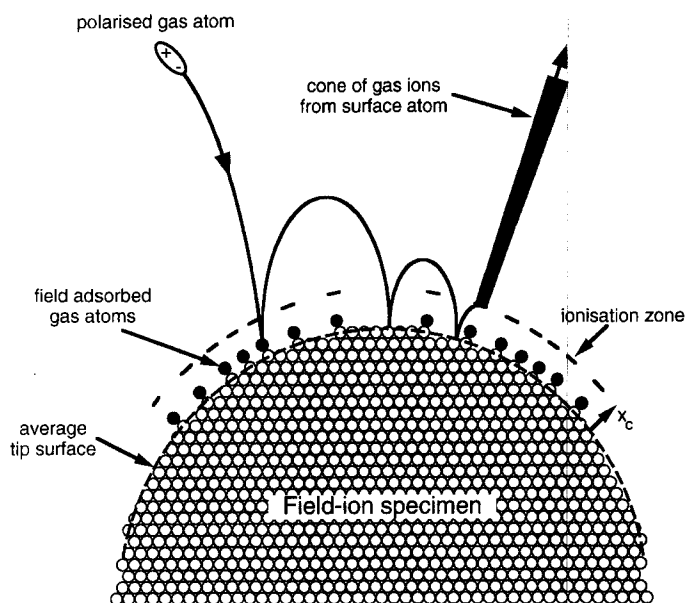
**Table 2.** Calculated and experimental evaporation fields (after Müller and Tsong [2] and Tsong [5])

Material	Calculated field for most easily evaporated species ( $\text{V nm}^{-1}$ )			Observed field ( $\text{V nm}^{-1}$ )	Main observed species
	$\text{M}^+$	$\text{M}^{2+}$	$\text{M}^{3+}$		
Be		46		34	$\text{Be}^{2+}$ , $\text{Be}^+$
Al	19			33	$\text{Al}^+$ , $\text{Al}^{2+}$
Si		32		30	$\text{Si}^{2+}$
Ti		26		25	$\text{Ti}^{2+}$
V		30			$\text{V}^{2+}$
Cr	27	29			$\text{Cr}^{2+}$
Mn	30	30			$\text{Mn}^{2+}$
Fe		33		35	$\text{Fe}^{2+}$
Co		37		36	$\text{Co}^{2+}$
Ni	35	36		35	$\text{Ni}^{2+}$
Cu	30			30	$\text{Cu}^+$
Zn	33				$\text{Zn}^{2+}$
Ga	15				
As		42			
Zr		28		35	$\text{Zr}^{2+}$
Nb		37		35	$\text{Nb}^{2+}$
Mo		41		46	$\text{Mo}^{2+}$ , $\text{Mo}^{3+}$
Ru		41		45	
Rh		41		46	$\text{Rh}^{2+}$ , $\text{Rh}^{3+}$
Pd	37				$\text{Pd}^+$
Ag	24				$\text{Ag}^+$
In	13				
Sn		23			
Hf		39			$\text{Hf}^{3+}$
Ta			44		$\text{Ta}^{3+}$
W			52	57	$\text{W}^{3+}$
Re		45		48	$\text{Re}^{3+}$
Ir		44		53	$\text{Ir}^{3+}$ , $\text{Ir}^{2+}$
Pt		45		48	$\text{Pt}^{2+}$
Au	53	54		35	$\text{Au}^+$ , $\text{Au}^{2+}$

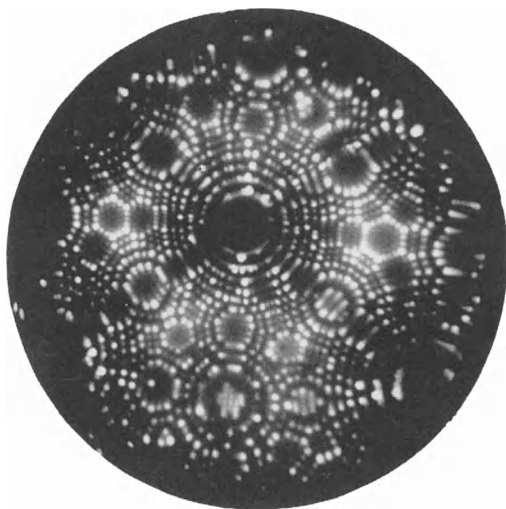
dependent, gas atoms are ionized preferentially above the highest field sites, which tend to be corner or kink site atoms. Each set of concentric rings reveals the prominent atoms surrounding an individual crystallographic pole, with the directions of largest interplanar spacing being the most prominent. It should be noted that the image rings are not due to a diffraction effect, but are a direct, real-space image, arising from the hemispherical shape of the specimen apex. The ball model shown in Fig. 10 has a large number of spheres,

arranged on a face-centered cubic lattice and shaped into an approximately hemispherical end-form. The spheres on kink sites, shown in white, form a pattern similar to the field ion image.

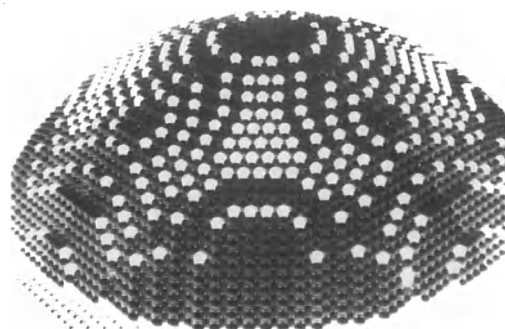
An important concept in field ion microscopy is the best image field (BIF). This refers to the field strength at which the best overall image contrast is obtained for a given specimen, in a given image gas. If the applied voltage is too low, the field is inadequate to produce an ion current from the specimen, while if the field is too high



**Figure 8.** Schematic diagram showing the processes which occur in the region of the specimen apex which lead to the formation of a field-ion image. Polarization forces attract gas atoms towards the specimen surface where they are thermally accommodated to the specimen temperature through a series of collisions. Some of these atoms are trapped at specific sites on the specimen surface, while others become field ionized above prominent atom sites and contribute to the image.



**Figure 9.** Field ion micrograph from a tungsten single crystal. The image was recorded at a specimen temperature of 28 K, using helium as the image gas, and with an applied voltage of 4.5 kV. The orientation of the sample was the same as that for the field emission micrograph shown in Fig. 3 and the crystallographic projection in Fig. 4.



**Figure 10.** Ball model of the specimen apex, constructed from a number of spheres arranged in a cubic lattice. The prominent sites, representing kink site atoms in the surface of the specimen, are shown white. These form intersecting sets of concentric rings similar to those seen in the field ion image.

ionization occurs uniformly above the tip surface and the image becomes fogged out. Image contrast is due to fluctuations in the overall rate of field ionization between different points on the specimen surface. There has been much debate about whether this is mainly associated with variations in the local concentration of image gas atoms

on the surface, or to variations in tunneling probability. Both effects would be expected to vary with local field strength and surface charge, and it now seems that both effects are important.

Additional contrast effects arise due to the phenomenon of field adsorption. Field ionization may take place by tunneling through an adsorbed gas atom, rather than through a vacuum barrier. There is no complete theory of this process, but field adsorption binding energies are of the order of  $10\text{--}20\text{ kJ mol}^{-1}$ , and so the surface will be almost completely covered with adsorbed gas atoms at the lowest imaging temperatures. The presence of adsorbed gas atoms does not seem to affect the determination of atom positions in the FIM, since the adsorbed atoms are located directly above the most prominent substrate atoms. However, if neon or argon atoms are adsorbed on the surface, the brightness of the image is increased, and the BIF is decreased, since the rate of electron tunneling through such atoms is greater than through helium atoms or a vacuum barrier. Small additions of hydrogen are also effective in lowering the BIF when helium or neon are used as the primary image gas, and this effect is also believed to be due to field adsorption. However, the addition of hydrogen commonly leads to the formation of metal hydride ions during field evaporation, and this leads to difficulties when carrying out atom probe microanalysis (see below).

As in the case of the FEM, the high magnification obtained in the FIM is due to the near-radial projection of ions from the specimen. With an apex radius of 50 nm, and a screen placed 10 cm away, a magnification of the order of  $10^6$  is achieved. The inverse dependence of magnification on

apex radius means there will be significant variations between specimens, and it is therefore important to record the applied voltage at which each FIM image is obtained. The spatial resolution in the FIM is limited at the lowest temperatures by the size of the ionization 'disk' above high field sites on the specimen. At higher operating temperatures ( $>50\text{ K}$ ), the resolution is controlled by the thermal velocities of the gas atoms at the moment of ionization. The lateral component of velocity leads to an angular spread in the trajectories of ions emitted from any given site and a corresponding increase in the size of the image spot. For a specimen temperature  $T$  this gives an approximate resolution for the microscope of:

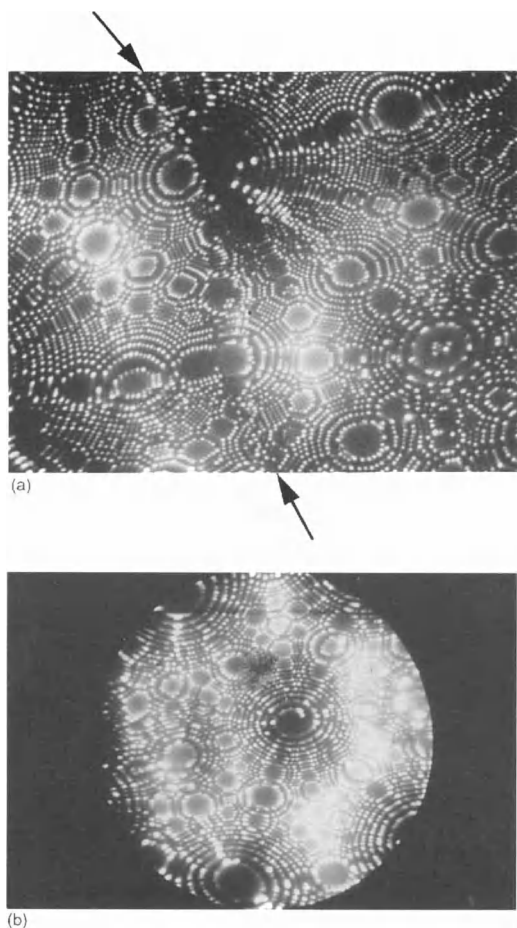
$$\delta = 4r\beta\sqrt{\frac{k_{\text{B}}T}{eV_0}} \quad (11)$$

For a specimen radius of 100 nm and an applied voltage of 10 kV, and taking  $\beta$  to be 1.5 gives a resolution of 1 nm at 300 K but 0.24 nm at 20 K.

A distinction should be made between the resolution of the image as defined by the ability to separate two atoms, and the precision with which the position of a single atom may be located. By determining the center of each spot in a well-resolved region of a field ion image, it is sometimes possible to measure atomic positions to within 0.02 nm.

### 2.2.5 Contrast from Lattice Defects and Alloys, and Analysis of Field Ion Microscope Images

If the apex of the specimen contains a grain boundary, the corresponding field



**Figure 11.** (a) Field ion micrograph of a tungsten specimen with a grain boundary. The presence of the boundary is seen as a discontinuity in the regular crystallographic contrast of the individual grains. (b) Field ion micrograph of a tungsten specimen containing a dislocation. The intersection of the dislocation with the surface of the specimen in the central (011) region converts the usual ring-like structure into a spiral. In order to produce this form of contrast, the dislocation must have a Burgers vector component normal to the specimen surface at the point of intersection. (Courtesy of T. J. Godfrey.)

ion image exhibits rings which are discontinuous at the boundary (Fig. 11a). Other types of lattice defect can also be seen in the FIM, but are more difficult to observe. If a dislocation intersects the surface of the

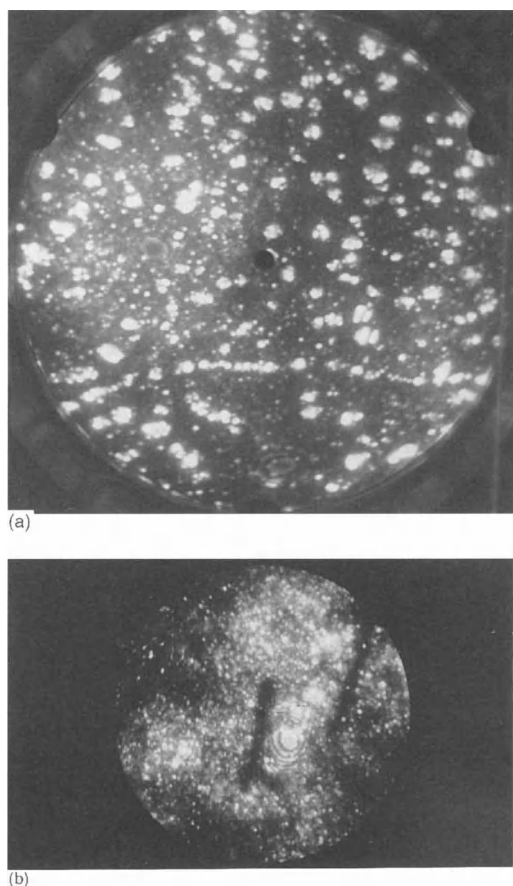
specimen close to a major crystallographic pole, and there is a component of the Burgers vector perpendicular to the surface, the image of the pole takes the form of a spiral (Fig. 11b). Point defects such as vacancies can also be observed, but because the field ion image does not show all the atoms on the surface, great care must be taken in the correct interpretation of such images.

When single-phase alloys are studied, there are often differences in the brightness of the field ion image of different atoms. This element-specific contrast may arise from preferential field ionization above different atoms, or from differences in the field evaporation behavior of the elements, or a combination of both effects. Species with high evaporation fields will be preferentially retained on the surface of the specimen, and will be associated with higher local fields and thus image more brightly. Elements with low evaporation fields will image darkly. In the case of concentrated solid solutions, the strong crystallographic contrast present in single-element materials is usually lost, because the image of the regular crystal structure is perturbed by the differences in image brightness from the constituent elements. In ordered materials in which there are large differences in the evaporation fields of the two species, exaggerated crystallographic contrast effects arise, in which only one of the elements is visible. Even when the difference in evaporation fields is not great, ordering on a superlattice pole can usually be observed as alternate bright and dim planes. Antiphase boundaries (APBs) are also visible in the field ion image as a mismatch between the bright/dark contrast of the individual atomic planes on either side of the boundary.



In the case of multiphase materials, there is often contrast which is very useful from a metallurgical point of view. This contrast results from differences in the local radius of curvature over the sample, due to differences in binding energy, and thus evaporation field, between the phases. A phase which has a lower evaporation field than the matrix must have a larger local radius of curvature than the surrounding material or it will field evaporate at a higher rate. Similarly, a region of high evaporation field must have a lower radius of curvature; that is, it must protrude slightly above the mean tip profile, or it will not evaporate sufficiently. Thus the shape which is produced by field evaporation is that which gives the same rate of removal from the whole imaged region. Less refractory regions will appear darker in the FIM image, because of reduced local fields, whereas more refractory phases will appear brighter. Examples of the images obtained from two-phase materials are shown in Fig. 12.

Once an FIM image has been established, field evaporation allows the internal structure of the material to be examined. The specimen is sectioned by removing the uppermost layers in a controllable manner. As the voltage is raised, atoms which occupy the most prominent positions on the surface (at edges and kink sites) are the first to field evaporate. The removal of these atoms exposes others to similarly high fields. The overall result is that every atomic terrace progressively shrinks in size as field evaporation progresses. The removal of single atomic layers can be observed from the collapse of individual rings in the images of low index poles, and thus the absolute depth traversed through the specimen can be



**Figure 12.** Examples of bright and dark contrast in field ion images, resulting from second-phase particles. (a) Image of a model maraging steel (Fe-18.2 at.% Ni-8.8% Co-2.9% Mo) specimen aged at 783 K for 1 h, showing brightly imaging spheroidal molybdenum rich precipitates. The distance across the image is approximately 165 nm. (Courtesy W. Sha and ASM International.) (b) Darkly imaging plates of cementite ( $\text{Fe}_3\text{C}$ ) in a pearlitic steel (Fe-3.47 at.% C-0.65% Mn-0.6% Si). (Courtesy of M. K. Miller.)

measured. In this way, the microstructure can be observed in three dimensions. The material removed by the field evaporation process can be chemically identified by mass spectrometry, thus allowing the composition of the imaged regions to be determined (see below).

Crystallographic parameters may be measured from FIM micrographs. As explained earlier, the image is a projection of the three-dimensional shape of the specimen surface. To a first approximation, this projection may be taken as stereographic. Major crystallographic poles may be indexed from the symmetry of the pattern, the main zone lines can then be constructed, and the indices of minor poles found from the zone addition rule. Once indexing has been carried out, the magnification of any given region of the image can be calculated. If the two-dimensional array of atoms on a given plane is fully resolved, and the lattice parameters and crystal structure of the material are known, then the magnification can be determined from direct measurement of the micrograph. More generally, if  $N$  is the number of image rings between a reference crystallographic pole ( $hkl$ ) and another pole separated from it by an angle  $\theta$ , then the radius of the specimen in this region is given by:

$$r(1 - \cos \theta) = Nd_{hkl} \quad (12)$$

where  $d_{hkl}$  is the interplanar spacing for the ( $hkl$ ) pole. The magnification can then be obtained from a calculation of the distance  $r\theta$  between the two poles, and comparison with the equivalent distance on the micrograph.

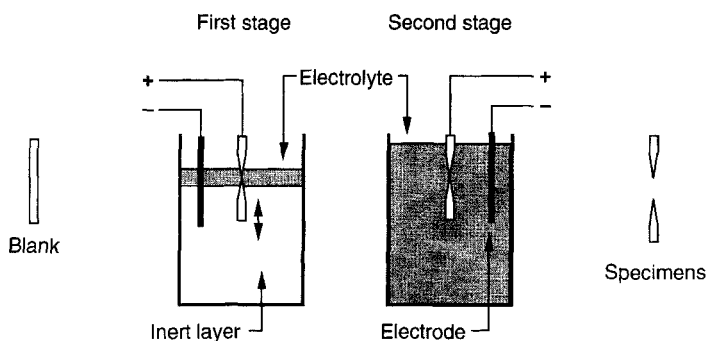
Three-dimensional reconstruction of the specimen geometry is possible if an extended sequence of micrographs is recorded at intervals during field evaporation. However, full analysis is hindered by the fact that the specimen end-form is never truly hemispherical. One reason is that the work function (and hence the evaporation field) varies with crystallographic orientation. Accurate measurement of curvilinear

distances across nonuniform specimen surfaces is difficult, except in very localized areas, and extensive computer calculation is often required. A simple measurement of distance may be made along the axis of the specimen if this coincides with a major crystallographic pole, since the step height of the atomic ledges of the axial pole will automatically provide a unit of distance. Long field evaporation sequences may be analyzed in this way, by recording the number of atom layers removed between successive micrographs. One application is to find the size of second-phase precipitate particles from their 'persistence depth', which is a more reliable estimate than lateral measurement based on an individual micrograph. Another important application is in the reconstruction of the atomic-scale topography of grain boundaries and interphase interfaces.

### 2.2.6 Specimen Preparation Techniques

The specimen needles required for the FIM (and FEM) may be prepared by chemical or electrolytic polishing, or by physical methods such as ion milling, depending on the nature and form of the starting material. It is convenient to start with materials of round cross-section and small diameter, such as wires. However, it is also possible to make samples from much larger objects (e.g. an aero engine turbine blade, a steel plate, or a semiconductor single crystal) by first slicing out a specimen blank of square cross-section, about 0.5 mm diameter and 10–15 mm long. Specimen blanks can also be prepared from very small objects, such as

**Figure 13.** Method typically used for the preparation of field ion specimens from bulk metallic alloys. Specimen blanks are usually 10–15 mm long and with a width of approximately 0.5 mm.



evaporated thin films, using microlithographic techniques.

Electropolishing is the most widely used preparation technique for metals and alloys. A two-layer polishing method (Fig. 13) is often employed. A layer of electrolyte, about 5 mm deep, is floated on top of a heavier inert liquid, such as a chlorofluorocarbon fluid. As material is removed in the electrolyte layer, a neck develops on the specimen, until it fractures under its own weight. A small amount of further polishing usually results in a sharp specimen. Blunted samples can be resharpened by a micropolishing technique. The end of the needle is immersed in a single drop of electrolyte, suspended on a small platinum loop, 1–2 mm in diameter, which acts as the counter electrode. The micropolishing is carried out under an optical microscope, so that progress can be monitored continuously.

Chemical polishing is used mainly for poorly conducting materials such as semiconductors, for which electropolishing cannot be used. The method usually involves a repeated dipping technique, and great care is needed in order to produce a smooth specimen surface by this approach, as etching commonly occurs.

Ion milling is a slow but controllable method of FIM specimen preparation, and

can be carried out in a conventional TEM specimen preparation unit, with some modification to the holder to accommodate the differently shaped specimen. The advantage of this method is that it generally produces a very smooth surface, without preferential etching at grain boundaries or second-phase particles, which can be a problem with the other techniques mentioned above. It is also suitable for specimens of low electrical conductivity.

A hybrid approach to specimen preparation which is useful in some circumstances is to begin the polishing using a chemical or electrolytic method, and then to use ion milling to carry out the final sharpening. A selection of recommended preparation techniques for different materials is given in Table 3.

## 2.3 Atom Probe Microanalysis

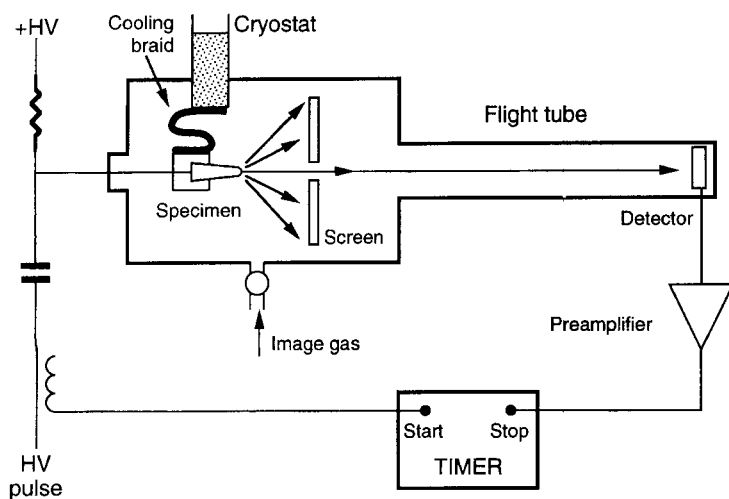
### 2.3.1 Principles of the Atom Probe Field Ion Microscope

The principle of the atom probe is outlined in Fig. 14. The simplest instrument consists

**Table 3.** Sample preparation techniques for selected materials

Material	Technique
Steels	First stage electropolish by two-layer method to form thin waist on specimen: 25% perchloric acid in glacial acetic acid, 20–25 V d.c., temperature 0–10°C. Final stage electropolish in a single bath of electrolyte: 2% perchloric acid in 2-butoxyethanol, 20–25 V d.c., temperature 0–10°C
Nickel alloys	First stage electropolish by two-layer method to form thin waist on specimen: 10% perchloric acid + 20% glycerol + 70% ethanol, 22 V d.c. Final stage electropolish in a single bath of electrolyte: 2% perchloric acid in 2-butoxyethanol, 25 V d.c.
Copper alloys	Electropolish in concentrated orthophosphoric acid, 16 V d.c.
Aluminum alloys	Procedures vary from alloy to alloy. One versatile method is to electropolish with 2–10% perchloric acid in methanol, 5–10 V a.c., at –10°C
Titanium alloys	Procedures vary from alloy to alloy. One method is to electropolish in 6% perchloric acid + 34% n-butyl alcohol + 60% methanol
Molybdenum alloys	First stage electropolish by two-layer method to form thin waist on specimen: 5 N aqueous sodium hydroxide solution, 6 V a.c. Final stage electropolish in a single bath of 12% concentrated sulphuric acid in methanol, 6 V d.c.
Platinum alloys	Electropolish in molten salt mixture of 80% sodium nitrate + 20% sodium chloride at 440–460°C, using a repeated dipping technique. Start at 5 V d.c., reducing to 3 V.
Tungsten	Electropolish using two-layer method in 5 wt.% aqueous sodium hydroxide solution, 5–6 V a.c.
Silicon	Chemically polish in a solution of 15% concentrated nitric acid, 80% hydrofluoric acid (40% solution) and 5% glacial acetic acid. Finish by ion milling if required
Gallium arsenide	Chemically polish in a solution of 44% concentrated sulphuric acid, 28% hydrogen peroxide (30% w/v solution) and 28% water, at about 60°C

Reproduced from Metals Handbook, 9th Edition, Volume 10, Materials Characterisation (1986), courtesy of ASM International.



**Figure 14.** Schematic diagram of a straight flight path atom probe. The flight path is typically 1–2 m but can be as long as 8 m.

of a combination of an FIM and a linear time-of-flight mass spectrometer, connected via an aperture hole, about 2 mm diameter, located in the center of the microscope screen. The spectrometer consists of an evacuated flight tube, at the end of which is situated a single-ion sensitivity particle detector. The FIM specimen is mounted on a goniometer stage and can be tilted to align any part of the image over the aperture hole. Ions are field evaporated from the surface at well-defined moments in time by means of a succession of short high-voltage pulses, superimposed on the d.c. imaging voltage applied to the specimen. Typically, these pulses are of subnanosecond rise time, 10 ns duration, and have amplitudes of 15–20% of the d.c. voltage. The field evaporated ions are accelerated towards the microscope screen, following trajectories which are very closely similar to those of the gas ions which formed the original FIM image. Most of the ions are stopped by the screen, but those from the selected region of the specimen will pass through the aperture hole into the spectrometer. Because of the high magnification of the FIM, the 2 mm diameter aperture in the screen defines a region 1–2 nm in diameter on the specimen surface.

The flight time of the ions from the specimen to the detector is measured by equating the applied potential  $V_0$  (the sum of the d.c. and the pulse voltages) with the final kinetic energy of the ion:

$$neV_0 = \frac{mv^2}{2} \quad (13)$$

where  $m$  is the mass of the ion and  $v$  its final velocity. The mass-to-charge ratio of the ion can be obtained from the flight

time  $t$  over a length  $d$ :

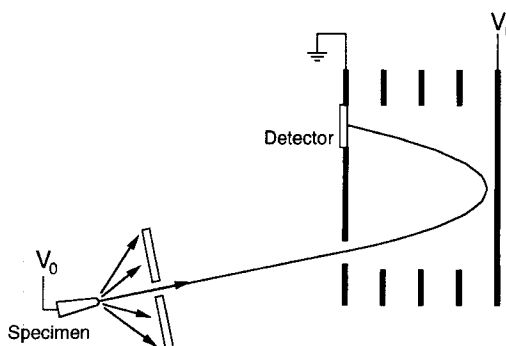
$$\frac{m}{n} = \frac{2eV}{d^2} t^2 \quad (14)$$

This method does not allow the mass of an ion to be determined independently of its charge state, but fortunately the number of charge states produced by any given element is usually small. Under particular conditions of field and temperature, most elements will evaporate in only one or two states, as indicated in Table 2. The identification of mass peaks in the spectrum is therefore relatively straightforward. In certain cases, the analysis can be complicated by the overlap of elements with a common isobar (e.g.  $^{58}\text{Fe}$  and  $^{58}\text{Ni}$ ) or by elements with different charge states (e.g.  $^{56}\text{Fe}^{2+}$  and  $^{28}\text{Si}^+$ ). When this occurs, reference data on natural isotopic abundances must be used to try and resolve the difficulty. These cases are few, however, and seldom pose a serious problem.

The lack of any spatially dispersive elements in the atom probe mass spectrometer design means it has a very high sensitivity, up to 100%, determined by the efficiency of the single-ion detector used in the instrument. The most common detectors are dual channel-plate assemblies (about 65% efficient). Channeltrons and multistage discrete dynode Cu–Be electron multipliers have also been used and, although less convenient to operate, give close to 100% efficiencies. However, the effectiveness of an atom probe is dependent not only on its sensitivity, but also on the accuracy with which it can identify a given ion. This is controlled by the spread of flight times for any given species, which in turn depends on the energy distribution of the field evaporated ions.

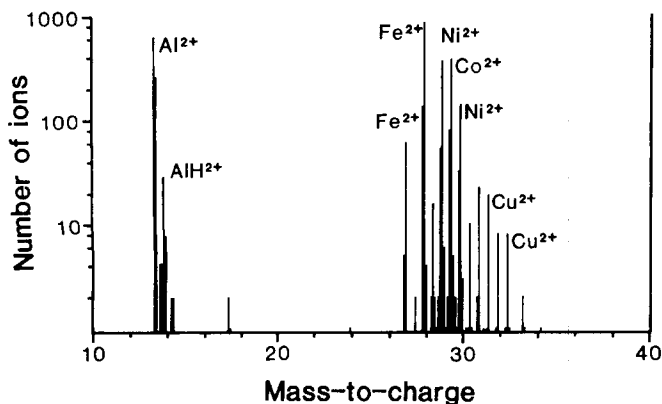
### 2.3.2 Energy Deficits and Energy Compensation

Because of the time varying electric field present at the specimen at the moment of evaporation, the ions are emitted with a small range of energies. This imposes a limit on the mass resolution of the simple linear atom probe of around  $m/\Delta m = 200$  to 300 (full-width at half maximum). For the study of complex engineering alloys, this resolution limit can impose a serious restriction, especially in the case of steels where the mass-to-charge ratios of some of the important alloying elements (Mn, Co, Ni, Cu) lie in the 'tails' of the main iron peaks. To improve the mass resolution of the instrument, some form of energy compensation must be incorporated into the design. Modern atom probes now include this feature as standard. Two main approaches are used, either a toroidal sector lens in a configuration first published by Poschenrieder [6] or a reflectron lens, as developed by Karataev et al. [7]. An atom probe design incorporating a reflectron, which is an increasingly favored design, is shown in Fig. 15. Ions entering the reflectron will be turned back at different points, depending on their total energy.



**Figure 15.** Schematic diagram of an energy compensated atom probe instrument using a reflectron. Ions of higher initial energy penetrate further into the reflectron. This increases their flight time inside the reflectron, so that they arrive simultaneously with ions of the same mass-to-charge ratio, but lower initial energy.

The highest energy ions will penetrate furthest into the reflectron, and will therefore have longer overall flight paths through the system. Thus, faster ions are effectively delayed, and the slower ions have an opportunity to 'catch up'. With suitable operating voltages, ions of the same mass-to-charge ratio can be time focused at the detector. Energy compensation is capable of producing an at least tenfold improvement in the mass resolution of the simple linear time-of-flight atom probe, making it possible to resolve individual isotopes fully (Fig. 16).



**Figure 16.** Energy compensated atom probe mass spectrum from the permanent magnet alloy, Alnico 5 (Fe-15.6 at.%Al-12.5%Ni-21.4%Co-2.5%Cu). At the high resolution available with this system, all the individual isotopes of the elements are well resolved.

### 2.3.3 Accuracy and Precision of Atom Probe Analysis

An essential factor in the operation of the atom probe is that the trajectories of field evaporated ions closely follow those of the corresponding image gas ions. Thus, when field evaporation occurs, only atoms whose images lie over the probe aperture pass into the time-of-flight mass spectrometer. The analysis area on the specimen surface is then defined by the region which corresponds to the physical aperture in the image screen. This limits the lateral resolution of conventional atom probe analysis. Some instruments allow the tip-to-screen distance to be altered, which allows the magnification of the FIM to be varied by a factor of 2–3. This has the secondary effect that the analysis area is also varied by the same amount, so that a smaller effective probe aperture can be selected if higher lateral resolution is required. However, as the analysis area becomes more restricted, the number of ions collected per atom plane field evaporated decreases. Therefore a compromise has to be achieved between spatial resolution and ion counting statistics.

In principle, the atom probe is a fully quantitative microanalytical tool. The method used to obtain an analysis is very straightforward. The absence of any need for external excitation to remove material from the specimen (e.g. by an electron beam or ion sputtering) means that interpretation of atom probe data is not affected by excitation cross-section, absorption, or fluorescence corrections. The specimen is held at cryogenic temperatures, and so atoms on the surface are only removed by field evaporation, and always in the

form of positive ions. Accurate quantitative analysis only requires detection of the resulting ions with an efficiency which is independent of the atomic species. The efficiency of the channel plate detectors used in most atom probes is effectively independent of mass for ion energies of above 5 keV [8], and therefore creates no difficulties. The main limitation of the use of time-of-flight analysis is that ions are detected only if they are produced during the field evaporation pulse. Since the evaporation fields of the various elements present will generally be different, elements with lower evaporation fields may be evaporated preferentially at the d.c. voltage, leading to loss in sensitivity and also inaccuracy in the analysis. The ratio of pulse voltage to d.c. voltage must be kept sufficiently high that evaporation during the d.c. part of the cycle is negligible. Fortunately, the kinetics of field evaporation are strongly field dependent, with evaporation rates changing by up to an order of magnitude for only a 1% change in applied field, so pulse amplitudes of 15% of the d.c. standing voltage are usually sufficient. However, in some materials for which the evaporation fields of the constituent elements differ widely, higher pulse ratios (up to 20%) may be required. It is usually advisable to determine the range of conditions for quantitative analysis of a given type of alloy by performing reference calibration experiments on homogeneous, solution treated material of accurately known bulk composition, over a range of pulse ratios. As the field dependence of evaporation rate is highest at low temperatures, analysis should be performed at the lowest convenient temperature, in order to reduce the pulse amplitude required.

Once appropriate experimental conditions have been established, composition determination in the atom probe simply involves atom counting. However, since the volume of material analyzed in the atom probe is so small, the total number of atoms available for composition measurement is limited. The precision of the measurement is therefore limited by counting statistics. If there are  $n_s$  atoms of the species of interest in a total number of atoms  $N$ , the estimate for the concentration  $c_0 = n_s/N$  has a standard deviation  $\sigma$  of:

$$\sigma = \sqrt{\frac{n_s(N - n_s)}{N^3}} \quad (15)$$

For example, in an analysis of a steel in which a total of 10 000 atoms were detected, 2388 of these were Cr atoms. The composition from the analyzed volume is therefore estimated to be  $23.8 \pm 0.4$  at.%.

Precise measurement of low atomic fractions in the atom probe is challenging, as it requires detection of very large numbers of atoms. When a trace element or impurity is localized within the material, for example at an interface, then the high spatial resolution of the atom probe allows precise measurement of the local concentration. The sensitivity of the ion detectors to light elements is a further advantage of the technique, allowing quantitative measurement of local concentrations of boron, carbon, or nitrogen, even in high atomic number matrix materials such as steels or nickel-based superalloys.

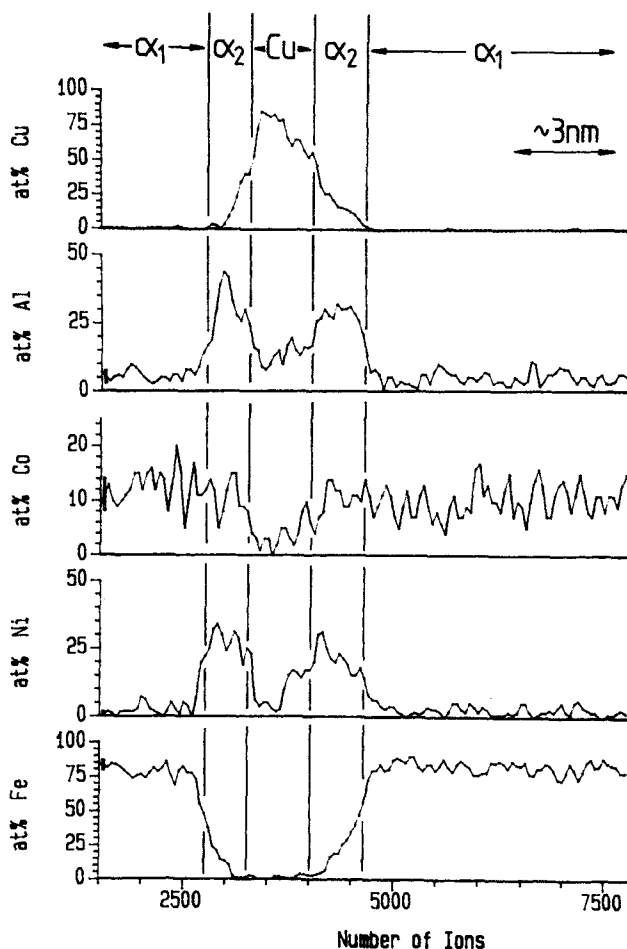
### 2.3.4 Atomic Plane Depth Profiling

The process of field evaporation provides depth profiling through the specimen. As

analysis proceeds, data are collected from a cylinder of material approximately 2 nm in diameter, over a depth which depends on the total amount of material evaporated. The composition of this volume can be estimated directly from the number of atoms of the various elements which are collected. In some cases this information is all that is needed, as the ability to sample volumes  $1\text{--}10\text{ nm}^3$  permits the identification of phase chemistry on a very fine scale. Often, however, the variation of composition along the cylinder of analysis is investigated. In this case it is not only the number of atoms which is important, but also the order in which they are detected. The data chain representing the sequence of detected atoms can be investigated by subdivision into smaller blocks. The composition of each block can then be calculated. Changes in composition from block to block will then give direct information about the variation of composition with depth in the specimen (Fig. 17).

The volume represented by each block of data will depend on the way in which the subdivision has been performed. Typically, the compositions are calculated using a fixed number of atoms per block, say 50–100 atoms. While this method is the simplest, there is not necessarily a linear relationship between the position of a given block in the data sequence and the absolute depth in the specimen. The number of atoms detected per atomic layer varies from specimen to specimen, even for the same material, since the magnification varies with the end radius, and is also dependent on tip shape. There may also be significant variations from one phase to another within a given specimen of a multi-phase material, due to local magnification changes produced by the variation





**Figure 17.** Atom probe composition-depth profile through the permanent magnet alloy Alnico 2 (Fe-20.8 at.%Al-14.8%Ni-11.3 Co-5.6%Cu). This material has a three-phase microstructure, consisting mostly of  $\alpha_1$  (Fe-Co rich) and  $\alpha_2$  (Al-Ni-Cu rich) phases, with a small volume fraction of copper precipitates formed within the  $\alpha_2$  phase. Reproduced from *Journal of Microscopy*, Volume 154, p. 215 (1989) by courtesy of the publishers.

in local radius discussed above. To obtain an absolute depth scale, the analysis is best performed with a small quantity of image gas present. The progress of field evaporation can then be monitored by following the collapse of rings on a known crystallographic pole. The size of each block in the composition profile can now be defined on the basis of the atoms detected during the removal of a fixed number of atomic planes (typically 2–5). Atomic layer depth resolution is obtained for composition variations perpendicular to the direction of analysis, but for changes of composition

in directions away from the local surface normal the resolution will be reduced. For example, if an interface lies at an angle to the direction of analysis, the finite diameter of the analysis cylinder will lead to a depth profile which shows an apparently diffuse interface.

One important application of the single atom layer resolution of the atom probe is in studies of site occupancy in ordered materials. In this type of experiment, the analysis region is selected to lie just within the top layer of the superlattice pole in the FIM image. During the analysis, atoms are

detected from only one superlattice plane at a time, and the ordering can be studied directly. The conditions for such an experiment need to be selected carefully, if the measurement of ordering is to be accurate. Preferential field evaporation or retention of individual species must be minimized, and this requires the specimen temperature, pulse fraction, and field evaporation rate to be set more accurately than is usually necessary for routine analysis. Figure 18 shows an example from the analysis of  $\text{Ni}_3\text{Al}$  with a variety of additions, where the data are represented in the form of a 'ladder' diagram. This is a vector plot, in which the sequence of detection of the major species is displayed by a movement of the plot horizontally for one species, and vertically for the other. The composition of individual atomic planes in the material is given directly by the local slope of the resultant line. In Fig. 18, the ladder diagram is divided into regions of  $45^\circ$  slope with horizontal lines in between, corresponding to the analysis of mixed  $\text{Ni} + \text{Al}$  planes and pure  $\text{Ni}$  planes, respectively. Atoms of additional elements ( $\text{Co}$ ,  $\text{Fe}$ ,  $\text{Hf}$ ) are represented by arrows, and show whether the solutes preferentially occupy sites in one plane or the other. Figure 18 indicates that both  $\text{Fe}$  and  $\text{Hf}$  are found predominantly in the mixed planes.

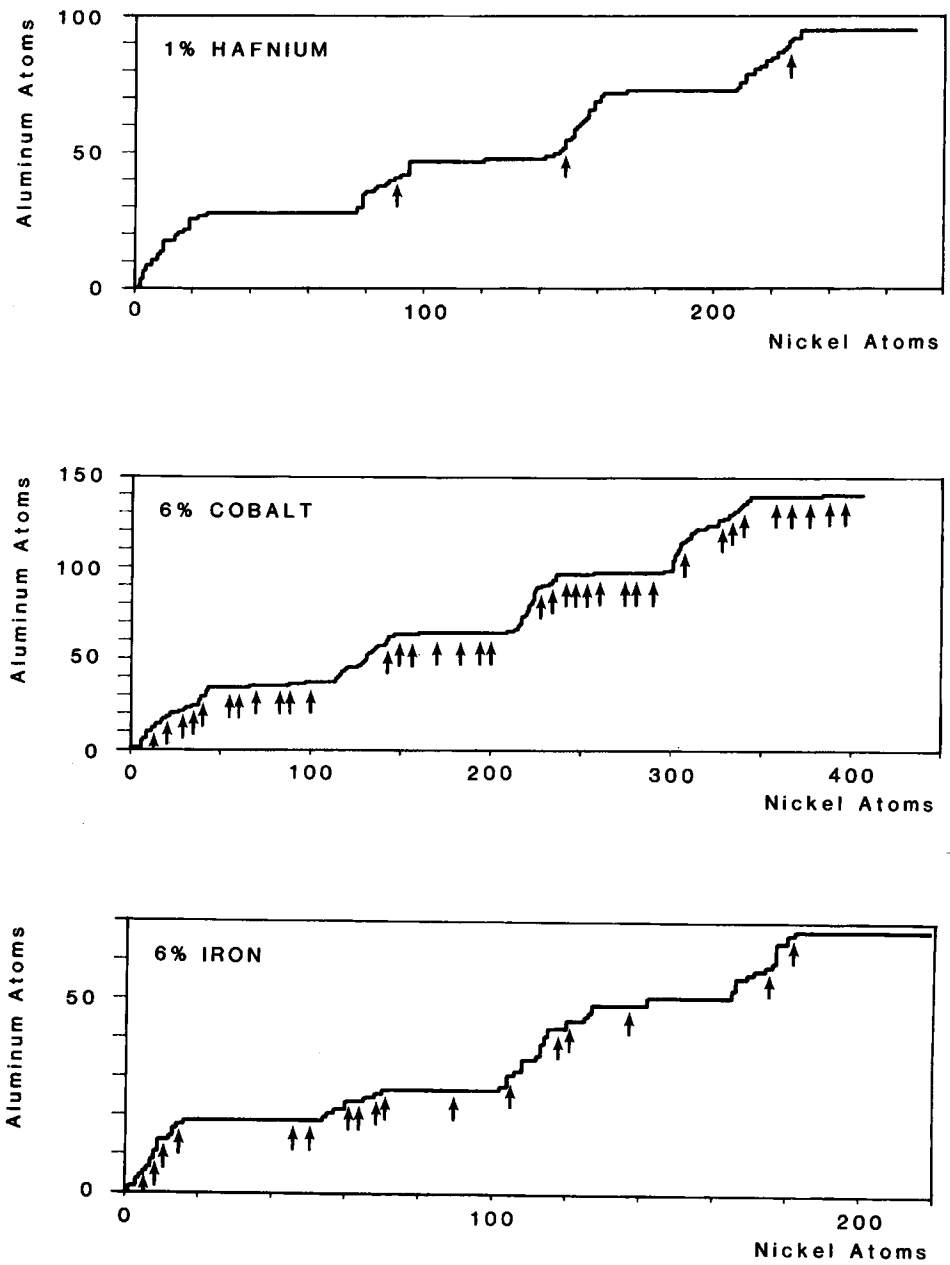
### 2.3.5 Analysis of Semiconductor Materials

For highly conductive materials such as metals, the use of voltage pulses to generate field evaporation is very effective, and most atom probe analysis is carried out

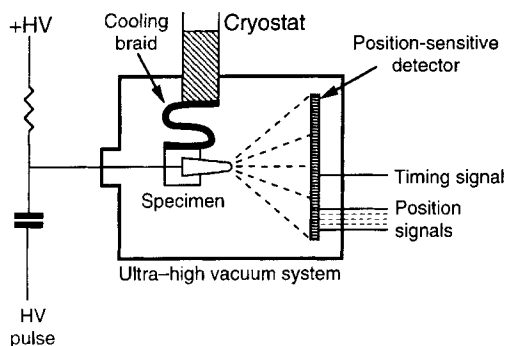
in this way. However, in propagation through less conductive materials, such as semiconductors, the pulse can be severely attenuated. Even where sufficient amplitude does reach the apex of the specimen, the mechanical shock associated with the pulse can lead to brittle fracture. In the analysis of semiconductors, therefore, voltage pulses are usually replaced by nanosecond laser pulses, focused on the apex of the specimen. This version of the atom probe is termed the pulsed laser atom probe (PLAP) [9]. The laser heating momentarily raises the temperature of the specimen from the base temperature to around 300 K, producing field evaporation under the influence of the applied d.c. voltage. This method has proved very useful in the analysis of a range of semiconducting materials and also for analysis of conducting ceramics. It should be emphasized that removal of material in the PLAP is by field evaporation, as in the atom probe, and not by thermal evaporation. Hence, the spatial resolution obtained in PLAP analysis is similar to that for the voltage pulsed atom probe. The technique should not be confused with laser ionization mass analysis (LIMA) in which material is removed from the specimen by thermal desorption at higher laser powers and much higher surface temperatures.

## 2.4 Three-Dimensional Atom Probes

The use of a probe aperture to define the lateral resolution in the atom probe severely restricts the total volume of



**Figure 18.** Ladder diagram showing the atom probe analysis along the (100) direction in  $L_{12}$ -ordered nickel aluminum alloys (based on  $Ni_3Al$ ) with additions of 1 at.%Hf, 6 at.%Co or 6 at.%Fe, showing the measurement of ordering along the superlattice direction, and the site occupation of the impurity atoms. The individual planes in the material are represented by regions of different slope on the plot, the horizontal lines being due to pure Ni planes, and a  $45^\circ$  slope representing a mixed Ni + Al plane. The position of the impurity atoms in the detection sequence is shown by arrows. Reproduced courtesy of Dr M. K. Miller.



**Figure 19.** Schematic diagram of a three-dimensional atom probe. Flight times are measured over the short (100–600 mm) flight path between specimen and detector, and the position of impact of each ion is determined from the position-sensitive detector.

material which can be analyzed in any sample. This design also requires the accurate alignment of the probe aperture when a selected-area analysis is being performed. A recent development in the field of atom probe microanalysis has essentially removed these major limitations, by combining the time-of-flight mass spectrometry of the atom probe with position-sensing (Fig. 19). In this arrangement, the original position of the evaporated atom is inferred from the position of detection, in the same way that atom positions are determined from the image spots in a field ion micrograph. This eliminates the need for the selection aperture, and a larger area may be analyzed without sacrificing lateral resolution. Indeed, the spatial resolution obtained with position-sensitive detectors is better than that achieved with the probe hole aperture. The lateral resolution obtained with this instrument is typically limited to  $\pm 0.5$  nm by trajectory aberrations during field evaporation. The analysis proceeds by continued removal and detection of individual ions, atomic layer by atomic

layer through the material. By reconstructing the position of the detected atoms within each layer, as well as the sequence of atomic layers, the original three-dimensional distribution of the elements can be displayed. This novel instrument design has therefore been termed the three-dimensional atom probe (3DAP).

## 2.4.1 Position-Sensing Schemes

There are a number of 3DAP designs, differing principally by the type of position-sensitive detector used. The first 3DAP was the position-sensitive atom probe (PoSAP), developed by Cerezo et al. [10]. In this instrument, the position-sensitive detector was based around a double MCP assembly with a wedge-and-strip anode. The wedge-and-strip anode was originally developed by Anger and coworkers [11] and works by dividing the charge incident from the MCPs between three electrodes in such a way as to encode the centroid of the charge cloud. This provides a simple method of position sensing, but one which is only effective for single ions. If more than one ion arrives within the time window for charge measurement (about 250 ns) then the position information is lost. This reduces the overall detection efficiency, and imposes analysis conditions which minimize the number of field evaporation pulses producing more than one detected ion, limiting the overall speed of analysis. Subsequently, Blavette et al. [12] developed their tomographic atom probe (TAP) which employs a multistrike detection system. In this detector, the electron cloud generated by ions striking the

MCPs is incident on a square anode array of  $10 \times 10$  elements. A cloud will typically cover four elements in the array, and the ratio of the charge collected on each of the elements is used to calculate the centroid of the cloud. If two ions with sufficient spatial separation are incident on the detector, the position of each can be calculated in a simple manner. Kelly et al. [13] have proposed a similar multistrike detection system, using the principle of charge division in a conductive anode.

Another approach to position determination is to use a detector consisting of MCPs and a phosphor screen, and to couple the detector optically to an external position sensing detector. This makes the detector within the vacuum system much simpler, and the basic instrument configuration essentially that of an imaging atom probe, first developed by Panitz [14]. An additional advantage to this arrangement is that the detector can be used to obtain a direct field ion image of the analysis area. In the optical atom probe constructed by Miller [15, 16], the detector is fiber optic coupled to an image-intensified diode array camera. From the image acquired by the camera, individual ion positions are indicated by single spots of light. Again, this is essentially a single-event detection system, since if two different ions are detected on a single field evaporation pulse, it is not possible to ascertain which ion corresponds to which position. Cerezo et al. [17] have developed an optically-coupled multistrike detector system, called the optical PoSAP, which combines an image-intensified camera with an anode array photomultiplier tube. The image from the detector is split by a partially reflective mirror, and focused onto both the photomultiplier tube and the camera.

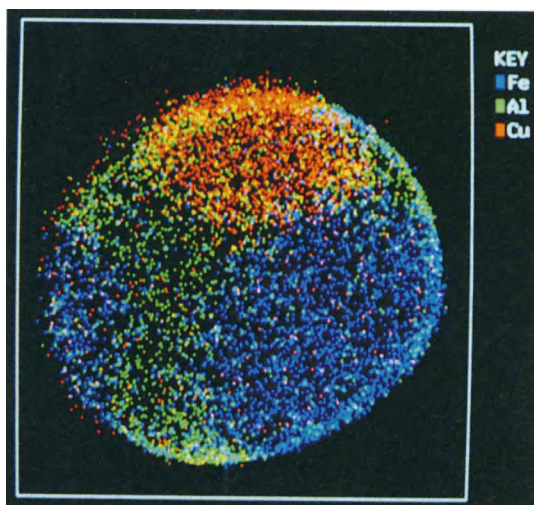
A fast phosphor is used in the detector, so that the signals from the elements of the anode array can be used for flight-time measurements. Each element in the array corresponds to a certain area of the detector, so that a timing signal from one of the elements can easily be correlated with the exact impact position determined from the image acquired by the camera.

### **2.4.2 Mass Resolution in the Three-Dimensional Atom Probe**

The first 3DAP designs used a straight flight path, and so suffered from the same limitations in mass resolution as the early probe-hole atom probe instruments. More recently, a 3DAP has been designed which uses a reflectron as an energy-compensating element [18]. The reflectron has the advantage that it acts as nothing more than a mirror in ion-optical terms, and so is free from image aberrations. However, care must be taken in the design of such a system, as the energy spreads are converted into position spreads, and therefore the initial energy spreads of ions must be kept to a minimum. The result is a 3DAP instrument with a mass resolution of  $m/\Delta m = 600$ .

### **2.4.3 Three-Dimensional Reconstruction of Atomic Chemistry**

The combination of data provided by a 3DAP can obviously be used to construct both a mass analysis of the volume of material and also a map of the distribution



**Figure 20.** Element distributions from the PoSAP analysis of the permanent magnet alloy Alnico 2 (Fe-20.8 at.%Al-14.8%Ni-11.3 Co-5.6%Cu) from an area approximately 15 nm in diameter and 2–3 nm in depth. The three-phase nature of the material, as indicated in Fig. 17, is more clearly seen in these images. Note again how the copper precipitate is totally contained within the  $\alpha_2$  phase.

of an individual element on the specimen surface. It is also possible to map out any number of elements. An example of a PoSAP analysis of a complex alloy is shown in Fig. 20. The multielement mapping provided by the PoSAP gives a detailed insight into the atomic-scale phase chemistry of the material.

However, it is the ability to reconstruct the three-dimensional microstructure present in the specimen which is the most significant advantage of the 3DAP, and makes the technique almost unique. The TAP analysis from a nickel based superalloy is shown in Fig. 21 [19]. The  $\gamma$  and  $\gamma'$  phases are seen clearly in the atomic reconstructions, as is the segregation of Mo, B, and C to both interphase and grain boundaries. Note the individual

(100) superlattice planes in the Al distributions, showing the ordering in the  $\gamma'$ -phase.

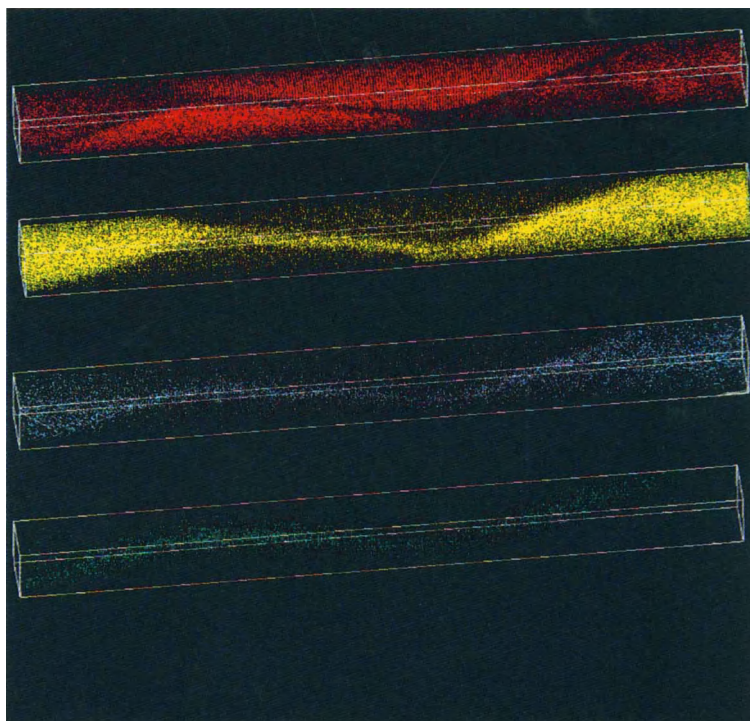
Miller et al. [20] and Hyde et al. [21, 22] have used the PoSAP to study phase separation in iron–chromium alloys, and compared the experimental results with computer simulations. With the PoSAP data, these authors have been able to compare directly atomic-scale data from experiment and the simulation, using chemical information, such as the amplitude of composition fluctuations, and also microstructural information.

## 2.5 Survey of Commercially Available Instrumentation

Field emission and field ion microscopes are relatively simple to construct, and many researchers build their own instruments from standard ultrahigh vacuum components. A commercial microscope for demonstrating field emission phenomena is available from Leybold Didactic GmbH, P.O. Box 1365, D-50330 Hürth, Germany. A compact table-top field ion microscope is manufactured by the Kitano Seiki Company Ltd, 7-17-3, Chuoh, Ohtaku, Tokyo 143, Japan.

An energy-compensated atom probe instrument, the FIM100, which incorporates a Poschenrieder analyzer, is produced by VG Scientific, The Birches Industrial Estate, Imberhorne Lane, East Grinstead, West Sussex RH19 1UB, UK. Another energy-compensated atom probe, the APFIM 220, incorporating a reflectron analyzer, is available from Applied Microscopy, The Courtyard, Whitwick Business

**Figure 21.** Tomographic atom probe analysis of the nickel-based superalloy Astroloy (Ni-8.5 at.%Al-16.0%Co-15.9%Cr-4.0%Ti-3.0%Mo-0.13%C-0.11%B-0.03%Zr) showing the atomic distributions of Al + Ti (red), Cr (yellow), Mo (blue) and B + C (green). The size of the volume analyzed is  $12\text{ nm} \times 12\text{ nm} \times 98\text{ nm}$ .  $\gamma'$ -precipitates (Al rich) are clearly visible in the  $\gamma$  matrix (Cr rich), and Mo, B, and C are seen to segregate to grain and interphase boundaries. Reproduced courtesy of L. Letellier, M. Guttman and D. Blavette.



Park, Stenson Road, Whitwick, Leicestershire LE67 3JP, UK.

Position-sensitive detection systems for 3DAP are available from Kindbrisk Ltd, 8 Tilgarsley Road, Eynsham, Oxford OX8 1PP, UK (PoSAP and OPoSAP systems, with or without reflectron-based energy compensation) and from Cameca, 103 Boulevard Saint-Denis, B.P. 6, 92403 Courbevoie Cedex, France (TAP detector). Computer software for atom probe control, data acquisition and data analysis is also available from Kindbrisk Ltd.

## 2.6 References

- [1] R. Gomer, *Field Emission and Field Ionization*, Harvard University Press, Harvard, CT **1961** (reprinted 1993, American Vacuum Society).
- [2] E. W. Müller, T. T. Tsong, *Field Ion Microscopy, Principles and Applications*, Elsevier, Amsterdam **1969**.
- [3] T. T. Tsong, *Atom Probe Field Ion Microscopy*, Cambridge University Press, Cambridge **1990**.
- [4] M. K. Miller, A. Cerezo, M. G. Hetherington, G. D. W. Smith, *Atom Probe Field Ion Microscopy*, Oxford University Press, Oxford **1996**.
- [5] T. T. Tsong, *Surf. Sci.* **1978**, *70*, 211.
- [6] W. P. Poschenrieder, *Int. J. Mass Spectrom. Ion Phys.* **1972**, *9*, 357.
- [7] V. I. Karataev, B. A. Mamyrin, D. V. Schmikk, *Sov. Phys.-Tech. Phys.* **1972**, *16*, 1177.
- [8] M. Hellsing, L. Karlsson, H.-O. André, H. Nördén, *J. Phys. E* **1985**, *18*, 920.
- [9] G. L. Kellogg, T. T. Tsong, *J. Appl. Phys.* **1980**, *51*, 1184.
- [10] A. Cerezo, T. J. Godfrey, G. D. W. Smith, *Rev. Sci. Instrum.* **1988**, *59*, 862.
- [11] H. O. Anger, *U.S. Patent No. 3 209 201*, September 28, **1965**.
- [12] D. Blavette, A. Bostel, J. M. Sarrau, B. Decohout, A. Menand, *Rev. Sci. Instrum.* **1993**, *64*, 2911.
- [13] T. F. Kelly, D. C. Mancini, J. J. McCarthy, N. A. Zreiba, *Surf. Sci.* **1991**, *246*, 396.
- [14] J. A. Panitz, *J. Vac. Sci. Technol.* **1974**, *11*, 206.

- [15] M. K. Miller, *Surf. Sci.* **1991**, 246, 428.
- [16] M. K. Miller, *Surf. Sci.* **1992**, 266, 494.
- [17] A. Cerezo, T. J. Godfrey, J. M. Hyde, S. J. Sijbrandij, G. D. W. Smith, *Appl. Surf. Sci.* **1994**, 76/77, 374.
- [18] S. J. Sijbrandij, A. Cerezo, T. J. Godfrey, G. D. W. Smith, *Appl. Surf. Sci.* **1996**, 94/95, 428.
- [19] L. Letellier, M. Guttman, D. Blavette, *Phil. Mag. Lett.* **1994**, 70, 189–194.
- [20] M. K. Miller, J. M. Hyde, M. G. Hetherington, A. Cerezo, G. D. W. Smith, C. M. Elliott, *Acta Met. Mater.* **1995**, 43, 3385.
- [21] J. M. Hyde, M. K. Miller, M. G. Hetherington, A. Cerezo, G. D. W. Smith, C. M. Elliott, *Acta Met. Mater.* **1995**, 43, 3403.
- [22] J. M. Hyde, M. K. Miller, M. G. Hetherington, A. Cerezo, G. D. W. Smith, C. M. Elliott, *Acta Met. Mater.* **1995**, 43, 3415.





## **Part VII**

---

# **Scanning Point Probe Techniques**



## General Introduction

Scanning near-field probe microscopy (SNPM) has developed from scanning tunneling microscopy (STM), which was invented in 1981 by G. Binnig, H. Rohrer, Ch. Gerber, and E. Weibel at the IBM Zürich Research Laboratory [1–3]. SNPM combines three important concepts: scanning, point probing, and near-field operation [4]. Scanning is achieved by means of piezoelectric drives which allow the positioning and raster scanning of a point probe relative to a sample surface to be investigated with subatomic accuracy. Nonlinearities due to piezoelectric hysteresis and creep usually have to be corrected electronically or by software to prevent image distortions. Point probing allows local information to be obtained about the physical and chemical properties of a sample surface, which facilitates the investigation of site-specific sample properties. The point probe is brought in close proximity to the sample at a distance  $s$  which is smaller than some characteristic wavelength  $\lambda$  of a particular type of interaction between the probe tip and the sample. [In the case of STM,  $\lambda$  would be the electron

wavelength whereas for scanning near-field optical microscopy (SNOM),  $\lambda$  would be the optical wavelength.] In this so-called near-field regime (where  $s \leq \lambda$ ), the spatial resolution that can be achieved is no longer limited by diffraction, but rather by geometrical parameters: the distance  $s$  between the point probe and the sample surface, and the effective radius of curvature  $R$  of the point probe. SNPM is therefore an exceptional type of microscopy because it works without lenses (in contrast to optical and electron microscopy), and achieves ‘super resolution’ beyond the Abbé limit. Another important feature of SNPM is that it can be operated in air and in liquids as well as in vacuum, which offers novel opportunities for high-resolution studies of the structure and processes at solid/fluid interfaces. In particular, in situ electrochemical studies and in vivo investigations of biological specimens at unprecedented real-space resolutions have become some of the more recent intense fields of application, besides surface science studies under ultra-high-vacuum conditions.



# 1 Scanning Tunneling Microscopy

## 1.1 Introduction

Scanning tunneling microscopy (STM) was the first near-field microscopy technique to be developed. It is based on vacuum tunneling of electrons between an electrically conducting point probe and an electrically conducting sample (metal or doped semiconductor). The schematic set-up for an STM experiment is shown in Fig. 1. An atomically sharp probe tip is brought within a distance of only a few angstroms ( $1 \text{ \AA} = 0.1 \text{ nm} = 10^{-10} \text{ m}$ ) from a sample surface by means of a piezoelectric drive in the  $z$ -direction (normal to the sample surface). If a bias voltage  $U$  has been applied between the tip and sample, a tunneling current  $I$  will flow due to the quantum mechanical tunneling effect even before mechanical point contact is reached. Since the tunneling current is strongly (exponentially) dependent on the tip-surface separation, it can be used very efficiently for distance control. By scanning the tip over the sample surface while keeping the tunneling current constant by means of an electronic feedback circuit, the surface contours can be traced by the tip. A quasi-three-dimensional 'topographic' image  $z(x, y)$  of the sample surface is gained by monitoring the vertical

position  $z$  of the tip as a function of the lateral position  $(x, y)$ , which is controlled by piezoelectric drives in the  $x$ - and  $y$ -directions. The position  $(x, y, z)$  of the tip can be calculated based on the known sensitivities of the piezoelectric drives in the  $x$ -,  $y$ -, and  $z$ -directions and the corresponding driving voltages  $U_x$ ,  $U_y$ , and  $U_z$ .

In addition to its use to control the tip-surface separation, the tunneling current contains valuable information about the local electronic properties of the sample surface and—to some extent—of the tip as well, which is unwanted in most cases. In the following, constant-current topographs and the various dependencies on experimental and sample-specific parameters will systematically be discussed. This will allow classification of the information which can be extracted from STM experiments.

## 1.2 Topographic Imaging in the Constant-Current Mode

Within Bardeen's transfer Hamiltonian formalism [5], the tunneling current  $I$  can be evaluated from the first-order time-dependent perturbation theory

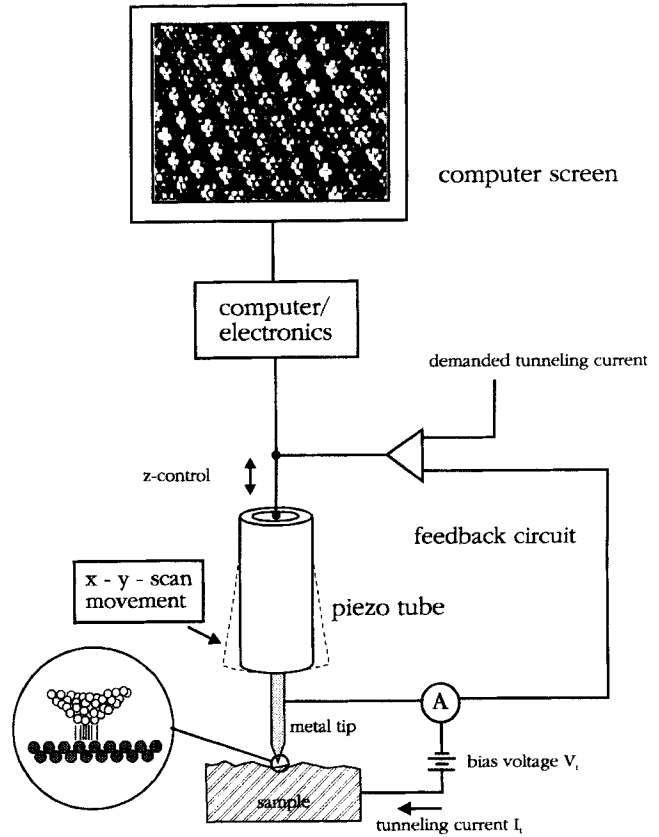


Figure 1. Schematic set-up for STM.

according to

$$I = \frac{2\pi e}{\hbar} \sum_{\mu,\nu} [f(E_\mu)(1 - f(E_\nu + eU)) - f(E_\nu + eU)(1 - f(E_\mu))] \times |M_{\mu,\nu}|^2 \delta(E_\nu - E_\mu) \quad (1)$$

where  $f(E)$  is the Fermi function,  $U$  is the applied sample bias voltage,  $M_{\mu,\nu}$  is the tunneling matrix element between the unperturbed electronic states  $\psi_\mu$  of the tip and  $\psi_\nu$  of the sample surface, and  $E_\mu(E_\nu)$  is the energy of the state  $\psi_\mu(\psi_\nu)$  in the absence of tunneling. The delta function describes the conservation of energy for the case of elastic tunneling.

(Inelastic tunneling will be considered later; see Sec. 1.6 of this Chapter.) The essential problem is the calculation of the tunneling matrix element which, according to Bardeen [5], is given by

$$M_{\mu,\nu} = \frac{\hbar^2}{2m} \int dS (\psi_\mu^* \nabla \psi_\nu - \psi_\nu \nabla \psi_\mu^*) \quad (2)$$

where the integral has to be evaluated over any surface lying entirely within the vacuum barrier region separating the two electrodes. The quantity in parentheses can be identified as a current density  $j_{\mu,\nu}$ . To derive the matrix element  $M_{\mu,\nu}$  from Eq. (2), explicit expressions for the wave functions  $\psi_\mu$  and  $\psi_\nu$  of the tip and sample

surface are required. Unfortunately, the atomic structure of the tip is generally not known. Therefore, a model tip wave function has to be assumed for calculation of the tunneling current.

Tersoff and Hamann [6, 7], who first applied the transfer Hamiltonian approach to STM, used the simplest possible model for the tip with a local spherical symmetry. In this model, the tunneling matrix element is evaluated for a s-type tip wave function, whereas contributions from tip wave functions with angular dependence (orbital quantum number  $l \neq 0$ ) have been neglected. Tersoff and Hamann considered the limits of low temperature and small applied bias voltage, for which the tunneling current becomes

$$I = \frac{2\pi e}{\hbar} U \sum_{\mu, \nu} |M_{\mu, \nu}|^2 \times \delta(E_\nu - E_F) \delta(E_\mu - E_F) \quad (3)$$

where  $E_F$  is the Fermi energy. Within the s-wave approximation for the tip, the following expression for the tunneling current is finally obtained

$$I \propto Un_t(E_F) \exp(2\kappa R) \times \sum_{\nu} |\psi_{\nu}(\bar{r}_0)|^2 \delta(E_{\nu} - E_F) \quad (4)$$

with the decay rate  $\kappa = (2m\phi)^{1/2}/\hbar$ , where  $\phi$  is the density of states at the Fermi level for the tip,  $R$  is the effective tip radius, and  $\bar{r}_0$  is the center of curvature of the tip. The quantity

$$n_s(E_F, \bar{r}_0) = \sum_{\nu} |\psi_{\nu}(\bar{r}_0)|^2 \delta(E_{\nu} - E_F) \quad (5)$$

can be identified with the surface local density of states (LDOS) at the Fermi level  $E_F$ , that is, the charge density from

electronic states at  $E_F$ , evaluated at the center of curvature  $r$  of the effective tip. The STM images obtained at low bias in the constant-current mode therefore represent contour maps of constant surface LDOS at  $E_F$  evaluated at the center of curvature of the effective tip, provided that the s-wave approximation for the tip can be justified. Since the wave functions decay exponentially in the  $z$ -direction normal to the surface toward the vacuum region,

$$\psi_{\nu}(\bar{r}) \propto \exp(-\kappa z) \quad (6)$$

it follows that

$$|\psi_{\nu}(\bar{r}_0)|^2 \propto \exp[-2\kappa(s + R)] \quad (7)$$

where  $s$  denotes the distance between the sample surface and the front end of the tip. Therefore, the tunneling current, given by Eq. (4), becomes exponentially dependent on the distance  $s$ , as mentioned in the introduction:

$$I \propto \exp(-2\kappa s) \quad (8)$$

The strong exponential distance dependence typically leads to an order-of-magnitude increase in the tunneling current for each angstrom decrease in the tip-surface separation.

Unfortunately, the simple interpretation of constant current STM images as given by Tersoff and Hamann is not valid for high bias or for tip wave functions with angular dependence.

### 1.2.1 Effects of Finite Bias

The applied bias voltage enters through the summation of states which can contribute to the tunneling current. Additionally, a finite bias can lead to a distortion of the tip and sample surface wave functions



$\psi_\mu$  and  $\psi_\nu$  as well as to a modification of the energy eigenvalues  $E_\mu$  and  $E_\nu$  [8]. The derivation of these distorted tip and sample surface wave functions and energy eigenvalues under the presence of an applied bias is, however, a difficult problem. Therefore, as a first approximation, the undistorted zero-voltage wave functions and energy eigenvalues are usually taken. Consequently, the effect of a finite bias  $U$  only enters through a shift in energy of the undistorted surface wave functions or density of states relative to the tip by an amount  $eU$ . Under this approximation, the following expression for the tunneling current as a generalization of the result of Tersoff and Hamann may be used:

$$I \propto \int_0^{eU} n_t(\pm eU \pm \mathcal{E}) n_s(\mathcal{E}, \bar{r}_0) d\mathcal{E} \quad (9)$$

where  $n_t(\mathcal{E})$  is the density of states for the tip and  $n_s(\mathcal{E}, \bar{r}_0)$  is the density of states for the sample surface evaluated at the center of curvature  $\bar{r}_0$  of the effective tip. All energies  $\mathcal{E}$  are measured with respect to the Fermi level. One can now make the following approximation motivated by a generalization of Eq. (5) together with Eq. (7):

$$n_s(\mathcal{E}, \bar{r}_0) \propto n_s(\mathcal{E}) \exp \left\{ -2(s + R) \times \left[ \frac{2m}{\hbar^2} \left( \frac{\phi_t + \phi_s}{2} + \frac{eU}{2} - \mathcal{E} \right) \right]^{1/2} \right\} \quad (10)$$

where a WKB-type expression for the decay rate  $\kappa$  in the exponential term has been used.  $\phi_t(\phi_s)$  denotes the tip (sample surface) work function. Finally, one obtains

$$I \propto \int_0^{eU} n_t(\pm eU \pm \mathcal{E}) n_s(\mathcal{E}) T(\mathcal{E}, eU) d\mathcal{E} \quad (11)$$

with an energy- and bias-dependent transmission coefficient  $T(\mathcal{E}, eU)$  given by

$$T(\mathcal{E}, eU) = \exp \left\{ -2(s + R) \times \left[ \frac{2m}{\hbar^2} \left( \frac{\phi_t + \phi_s}{2} + \frac{eU}{2} + \mathcal{E} \right) \right]^{1/2} \right\} \quad (12)$$

In Eqs. (11) and (12), matrix element effects in tunneling are expressed in terms of a modified decay rate  $\kappa$  including a dependence on energy  $E$  and applied bias voltage  $U$ . The expression (12) for the transmission coefficient neglects image potential effects as well as the dependence of the transmission probability on parallel momentum. This can be included by an increasingly more accurate approximation for the decay rate  $\kappa$ .

### 1.2.2 Effects of Tip Wave Functions with Angular Dependence

STM tips are usually made from tungsten or platinum-iridium alloy wire. For these materials, the density of states at the Fermi level is dominated by d-states rather than by s-states. Indeed, first-principle calculations of the electronic states of several types of tungsten clusters used to model the STM tip revealed the existence of dangling-bond states near the Fermi level at the apex atom which can be ascribed to  $d_{z^2}$  states [9]. Evaluation of the tunneling current according to Eqs. (1) and (2) now requires calculation of the tunneling matrix element for tip wave functions with angular dependence ( $l \neq 0$ ). Chen [10] has shown that generally the tunneling matrix element can simply be obtained from a 'derivative rule'. The angle dependence of the tip wave function in terms of

$x$ ,  $y$ , and  $z$  has to be replaced according to

$$\begin{aligned} x &\rightarrow \frac{\partial}{\partial x} \\ y &\rightarrow \frac{\partial}{\partial y} \\ z &\rightarrow \frac{\partial}{\partial z} \end{aligned} \quad (13)$$

where the derivatives have to act on the sample surface wave function at the center of the apex atom. For instance, the tunneling matrix element for a  $p_z$  tip state is proportional to the  $z$ -derivative of the sample surface wave function at the center of the apex atom at  $r_0$ .

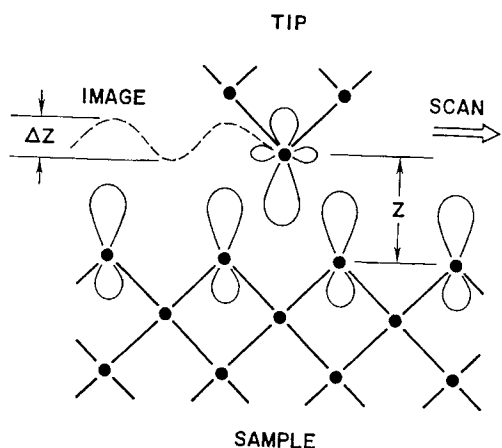
In terms of a microscopic view of the STM imaging mechanism [11] illustrated in Fig. 2, a dangling-bond state at the tip apex atom is scanned over a two-dimensional array of atomic-like states at the sample surface. Overlap of the tip state with the atomic-like states on the sample surface generates a tunneling conductance which depends on the relative position of the tip state and the sample state. The atomic corrugation  $\Delta z$  depends on the

spatial distribution as well as on the type of tip and sample surface states. Generally, for non- $s$ -wave tip states, the tip apex atom follows a contour, determined by the derivatives of the sample surface wave functions, which exhibit much stronger atomic corrugation than the contour of constant surface LDOS at  $E_F$ .

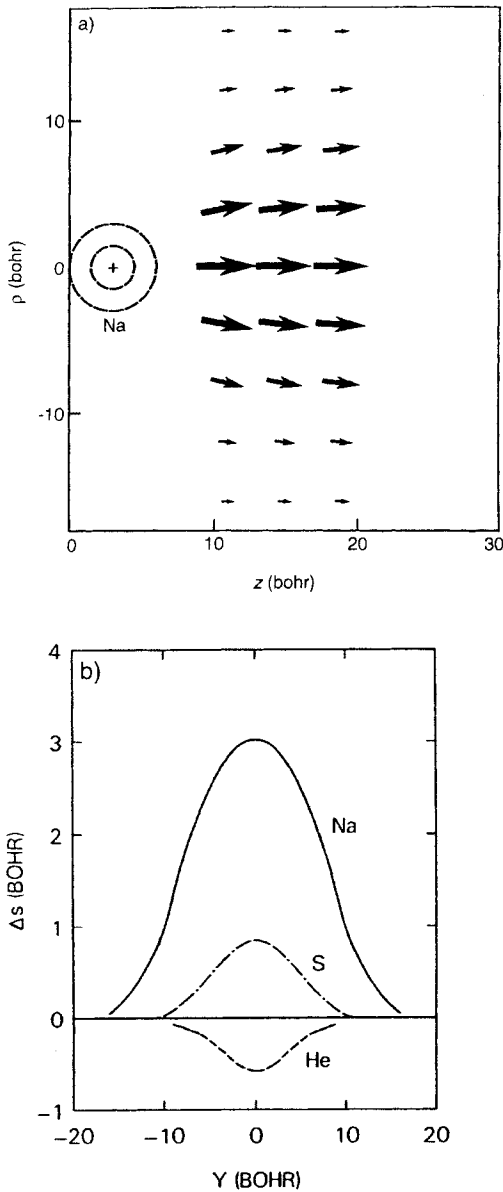
### 1.2.3 Imaging of Adsorbates

The transfer Hamiltonian approach as used by Tersoff and Hamann has further limitations. First, it is a perturbative treatment of tunneling, being appropriate only for weakly overlapping electronic states of the two electrodes. Secondly, this approach suffers from the fact that assumptions for the tip and sample surface wave functions have to be made in order to derive the tunneling current.

As an alternative, Lang [12,13] has calculated the tunneling current between two planar metal electrodes with adsorbed atoms where the wave functions for the electrodes have been obtained self-consistently within density functional theory. In Fig. 3a the calculated current density distribution from a single sodium atom adsorbed at its equilibrium distance on one of the two metal electrodes is shown. The plot illustrates how spatially localized the tunneling current is. By scanning one adsorbed atom (taken as the tip) past another adsorbed atom (taken as the sample), the vertical tip displacement versus the lateral position can be evaluated under the constant-current condition [13]. In Fig. 3b, constant-current scans at low bias of a sodium tip atom past three different sample adatoms (sodium, sulfur,



**Figure 2.** Microscopic view of the STM imaging mechanism. (Image courtesy of Chen [11].)



**Figure 3.** (a) Current density for the case in which a sodium atom is adsorbed on the left electrode. The length and thickness of the arrows are proportional to  $\ln(ej/j_0)$  evaluated at the spatial position corresponding to the center of the arrow (1 bohr = 0.529 Å). (Image courtesy of Lang [14].) (b) Change in tip distance versus lateral separation for constant tunnel current. The tip atom is sodium; sample adatoms are sodium, sulfur, and helium. (Image courtesy of Lang [13].)

and helium) are shown. Most striking is the negative tip displacement for adsorbed helium. The closed valence shell of helium is very much lower in energy with respect to the Fermi level, and its only effect is to polarize metal states away from  $E_F$ , thereby producing a decrease in the Fermi level state density. This results in a reduced tunneling current flow, that is, a negative tip displacement in a constant-current scan. This example illustrates nicely that 'bumps' or 'holes' in 'topographic' STM images may not correspond to the presence or absence of surface atoms, respectively—sometimes even the reverse is true.

### 1.2.4 Spatial Resolution in Constant-Current Topographs

According to Tersoff and Hamann [7], an STM corrugation amplitude, or corrugation in brief,  $\Delta$ , may be defined by

$$\Delta := z_+ - z_- \quad (14)$$

where  $z_+$  and  $z_-$  denote the extremal values of the  $z$ -displacement of the tip in a constant-current scan. This corrugation  $\Delta$  decreases exponentially with distance  $z$  from the surface,

$$\Delta \propto \exp(-\gamma z) \quad (15)$$

where the decay rate  $\gamma$  is very sensitive to the surface lattice constant because it depends quadratically on the corresponding Fourier component  $G$  in the expansion of the surface charge density,

$$\gamma \propto \frac{1}{4}\kappa^{-1}G^2 \quad (16)$$

with  $\kappa^2 = 2m\phi/\hbar^2$ . Consequently, only the lowest non-zero Fourier component determines the corrugation at sufficiently large

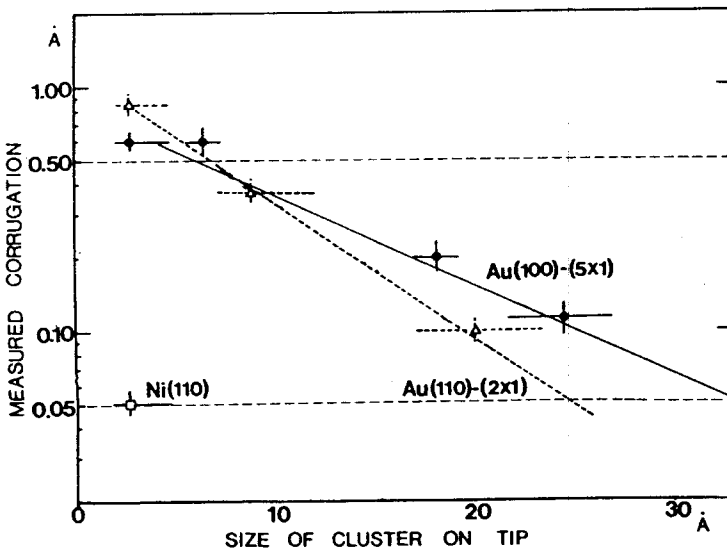
distances. Tersoff and Hamann [6, 7] argued that suppression of higher Fourier components in their expression for the tunneling current between a spherical tip of radius  $R$  and a sample surface at a distance  $s$  from the front end of the tip is equivalent to a spatial resolution determined by

$$\left( \frac{R+s}{\kappa} \right)^{1/2} \quad (17)$$

According to this expression, the lateral resolution in STM is determined by the geometrical parameters  $R$  and  $s$ , rather than by the wavelength of the tunneling electrons. This is characteristic for near-field microscopes which are operated at distances between the probe tip and the sample surface that are small compared with the wavelength, as mentioned in the introduction. For STM, typical tip-surface separations are 3–10 Å, whereas the wavelength of tunneling electrons typically varies in the range 12–120 Å for an applied bias voltage of 0.01–1 V.

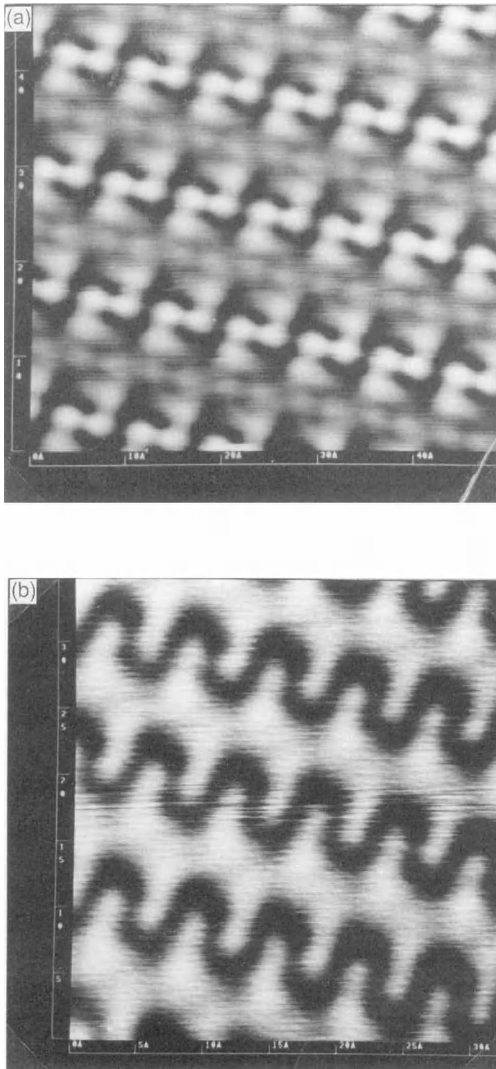
The expression, Eq. (17), for the lateral resolution in constant-current STM images implies that high spatial resolution is obtained with a small radius of curvature of the effective tip and at a small tip-surface separation, that is, at low tunneling gap resistance. Both dependencies have been verified experimentally. The dependence of the measured corrugation on the radius of curvature of the effective tip was studied by combined STM-FIM (field ion microscopy) experiments [15] where the obtained STM results could be correlated directly with the size of the effective tip as revealed by FIM (Fig. 4). As a direct consequence, measured absolute values for the corrugation  $\Delta$  are meaningless if the microscopic structure of the tip is not known.

The dependence of the measured corrugation on the tip-surface separation has experimentally been studied by systematic variation of the tunneling gap resistance [16, 17]. In particular, the suppression of higher Fourier components in the



**Figure 4.** Dependence of the measured corrugation on the size of the cluster on the tip for Au(110) ( $2 \times 1$ ) and Au(100) ( $5 \times 1$ ) reconstructed surfaces. (Image courtesy of Kuk [15].)

expansion of the surface charge density with increasing tip–surface separation, as theoretically predicted by Tersoff and Hamann [6, 7], has experimentally been verified [17]. Figure 5 shows the influence



**Figure 5.** (a) STM topograph of the W(110)/C-R(15 × 3) reconstructed surface obtained with a tunneling gap resistance  $R = 2.11 \times 10^6 \Omega$ . (b) Corresponding STM topograph with a tunneling gap resistance  $R = 1.72 \times 10^{10} \Omega$  [17].

of the chosen tip–surface separation on the spatial resolution achieved on a W(110)/C-(15 × 3) reconstructed surface. It is immediately apparent that the STM results can critically depend on the tip–surface separation, that is, on the tunneling gap resistance, particularly for surface structures with complex unit cell structure.

For close-packed metal surfaces, such as Au(111) [18] or Al(111) [16], atomic resolution could not be explained within the spherical tip model employed by Tersoff and Hamann. Baratoff [19] early on pointed out that the spatial resolution might be considerably improved compared with expression (17) if tunneling occurs via localized surface states or dangling bonds. Later, Chen [11, 20] systematically investigated the influence of different tip orbitals on the spatial resolution within a microscopic view of STM. The calculated enhancement of the tunneling matrix element by tip states with  $l \neq 0$  was shown to lead to increased sensitivity to atomic-sized features with large wavevectors. For instance, a  $p_z$  tip state acts as a quadratic high-pass filter, whereas a  $d_{z^2}$  tip state acts as a quartic high-pass filter. Consequently, the resolution of STM can be considerably higher than predicted within the s-wave tip model. The spontaneous switching of the resolution often observed in or between atomic-resolution STM images can be explained by the fact that a very subtle change of the tip involving a change of the effective orbital can induce a tremendous difference in STM resolution. In conclusion, it is the orbital at the front end of the tip which mainly determines the spatial resolution in STM. A  $p_z$  orbital typical for elemental semiconductors or a  $d_{z^2}$  orbital from d-band metals are

most favorable. Therefore, 'tip-sharpening procedures' have to aim at bringing such favorable orbitals to the front of the tip [11].

### 1.3 Local Tunneling Barrier Height

According to Eq. (8), the tunneling current  $I$  was found to depend exponentially on the tip-surface separation  $s$ :

$$I \propto \exp(-2\kappa s)$$

with a decay rate  $\kappa$  given by

$$\kappa = \frac{(2m\phi)^{1/2}}{\hbar}$$

where  $\phi$  is an effective local potential barrier height. So far,  $\phi$  was assumed to be laterally uniform. In reality,  $\phi$  exhibits spatial variations which can yield additional information about the sample surface under investigation.

Motivated by Eq. (8), an apparent local barrier height is usually defined by

$$\phi_A = \frac{\hbar^2}{8m} \left( \frac{d \ln I}{ds} \right)^2 \quad (18)$$

For large tip-surface separations outside the effective range of image forces, it is clear that  $\phi$  has to approach the surface local work function  $\phi_s$ , which is defined as the work needed to remove an electron from the Fermi level of the solid to a position somewhat outside of the surface where image force effects can be neglected. However, for small tip-surface separations (5–10 Å), image potential effects certainly have to be considered. By

assuming a model potential [21]

$$\phi(d) = \phi_0 - \frac{\alpha}{d} \quad (19)$$

where  $\phi_0$  is the average work function of the sample surface and the probe tip [ $\phi_0 = (\phi_s + \phi_t)/2$ ], and  $d$  is the distance between the two image planes ( $d \approx s - 1.5 \text{ Å}$ ), the distance dependence of the tunneling current can be calculated:

$$\begin{aligned} \frac{d \ln I}{ds} = & -\frac{2(2m)^{1/2}}{\hbar} \\ & \times \phi_0^{1/2} \left[ 1 + \frac{\alpha^2}{8\phi_0^2 d^2} + \mathcal{O}\left(\frac{1}{d^3}\right) \right] \end{aligned} \quad (20)$$

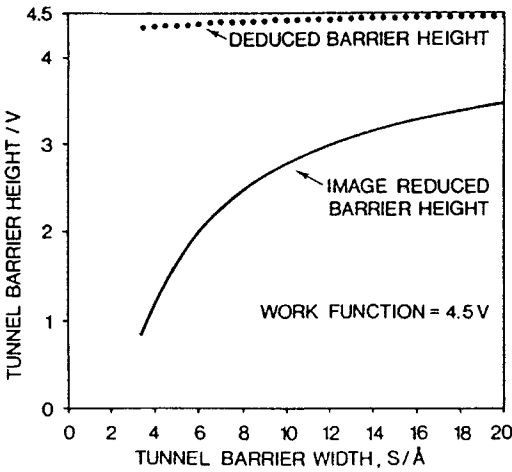
As can be seen from Eq. (20), the first-order term in  $1/d$ , although present in the potential  $\phi(d)$ , cancels exactly in the expression for  $d \ln I/ds$ . The second-order term in  $1/d$  usually contributes only a few percent of the zero-order term and can therefore be neglected to a first approximation. As a consequence, one finds

$$\frac{d \ln I}{ds} \approx \text{const.} \quad (21)$$

and

$$\phi_A \approx \phi_0 = \text{const.}$$

This means that the presence of the image potential does not show up in the distance dependence of the tunneling current although the absolute values of the current are drastically affected by the presence of the image potential. The distance independence of the apparent barrier height deduced from the  $\ln I$ - $s$  relation (Fig. 6) has been verified experimentally as well as by more detailed theoretical analysis [22].



**Figure 6.** The image-reduced mean barrier height (full line) and the apparent barrier height deduced from the  $\ln I$ - $s$  relation for this barrier (dotted line). The work function used in the calculation is 4.5 eV. It can be seen that the apparent barrier height is always within 0.2 eV of the work function despite the collapse of the image-reduced barrier. (Image courtesy of Coombs [22].)

### 1.3.1 Local Tunneling Barrier Height Measurements at Fixed Surface Locations

According to Eq. (18), the apparent barrier height  $\phi_A$  can be determined locally by measuring the slope of  $\ln I$ - $s$  characteristics at a fixed sample bias voltage  $U$  and at a fixed sample surface location. To demonstrate vacuum tunneling it is necessary to obtain reasonably high values for  $\phi_A$  of several electronvolts in addition to verifying the exponential dependence of the current on the tip-surface separation.

Alternatively, the apparent barrier height can be deduced from the slope of local  $\ln U$ - $s$  characteristics in a low applied bias voltage range and at a fixed tunneling current. In the low-bias

regime, the tunnel junction exhibits Ohmic behavior, as found earlier (4):

$$I \propto U \exp(-2\kappa s) \quad (22)$$

Therefore, one obtains

$$\phi_A = \frac{\hbar^2}{8m} \left( \frac{d \ln U}{ds} \right)^2 \quad (23)$$

at constant current.

### 1.3.2 Spatially Resolved Local Tunneling Barrier Height Measurements

The experimental determination of the spatially resolved local tunneling barrier height  $\phi_A(x, y)$  can be performed by modulating the tip-surface separation  $s$  by  $\Delta s$  while scanning at a constant average current  $I$ , with a modulation frequency  $\nu_0$  higher than the cut-off frequency of the feedback loop [23]. The modulation of  $\ln I$  at  $\nu_0$  can be measured by a lock-in amplifier simultaneously with the corresponding constant-current topograph, and directly yields a signal proportional to the square root of the apparent barrier height via the relation

$$\frac{\Delta \ln I}{\Delta s} = -\frac{2\sqrt{2m}}{\hbar} \sqrt{\phi_A} \quad (24)$$

The apparent barrier height obtained in this way is not measured at a constant tip-surface separation  $s$ . Scanning at a constant average current (and at a constant applied bias voltage) implies that the product  $\sqrt{\phi_A} s$  is kept constant, rather than  $s$ . However, since the spatial variation of  $\phi_A$  is usually small (about 10% or less of the absolute value of

$\phi_A$ ), and  $\phi_A$  enters only under the square root, the spatial variation of  $\phi_A(x, y)$  is usually measured almost at a constant tip–surface separation  $s$ .

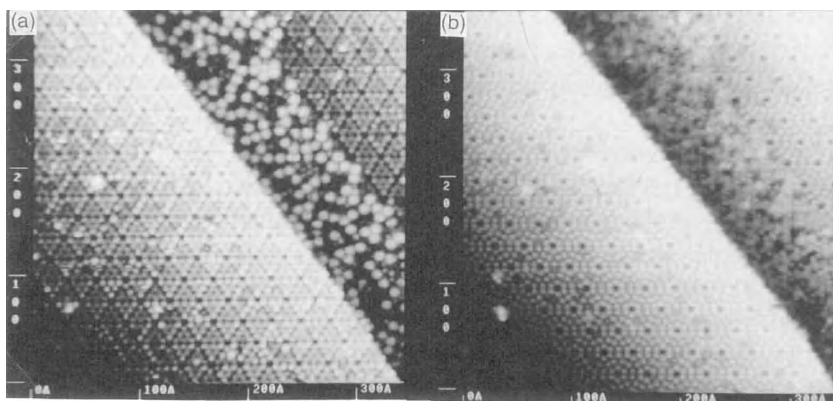
Spatially resolved measurements of the apparent potential barrier height can yield information about spatial inhomogeneities of the local sample work function  $\phi_s$ , which can be split into two contributions. The chemical component of  $\phi_s$  is determined by the chemical nature and structure of the solid only, whereas the electrical component of  $\phi_s$  depends on the chemical nature of the solid as well as on the surface crystallographic orientation. Therefore, spatially resolved measurements of  $\phi_A$  can be used, for instance, to map chemical inhomogeneities at surfaces as well as different crystallographic facets of small crystallites.

On the atomic scale, it is more appropriate to relate the measured apparent barrier height with the decay rates of the wave functions describing the sample surface and the tip. Lateral variations of  $\phi_A(x, y)$  then have to be interpreted as lateral variations in the decay rate of the surface wave function. As we know from Eq. (15), the measured surface atomic corrugation  $\Delta$  in constant-current STM images is smoothed out exponentially with an increasing tip–surface separation  $s$ . This can only occur if the decay rate  $\kappa_p$  above a local protrusion in the topography is larger than the decay rate  $\kappa_d$  above a local depression. Consequently, the apparent barrier height above a local topographic protrusion has to be larger than the barrier height above a local depression. Therefore, atomically resolved apparent barrier height images closely reflect corresponding topographic constant-current images.

## 1.4 Tunneling Spectroscopy

Besides the distance and apparent barrier height dependence of the tunneling current there also exists a bias dependence which can be studied by various tunneling spectroscopic methods. For tunneling between metal electrodes in the low-bias limit, the tunneling current is found to be linearly proportional to the applied bias voltage [Eq. (4)]. For higher bias and particularly for semiconductor samples, the bias dependence of the tunneling current generally does not exhibit Ohmic behavior, and the constant-current STM images can depend critically on the applied bias (Fig. 7). Studying this bias dependence in detail allows extraction of various spectroscopic information at high spatial resolution, ultimately down to the atomic level. The spectroscopic capability of STM combined with its high spatial resolution is perhaps the most important feature of STM, and has been applied widely, particularly for investigation of semiconductor surfaces [24–26]. Figure 8 shows a simplified one-dimensional potential energy diagram at zero temperature for the system consisting of the tip (left electrode) and the sample (right electrode), which are separated by a small vacuum gap. For zero applied bias (Fig. 8b) the Fermi levels of tip and sample are equal at equilibrium. When a bias voltage  $U$  is applied to the sample, the main consequence is a rigid shift of the energy levels downward or upward in energy by an amount  $|eU|$ , depending on whether the polarity is positive (Fig. 8c) or negative (Fig. 8d). (As discussed in Sec. 1.2.1 of this Chapter, we neglect the distortions of the wave functions and the energy eigenvalues due to the





**Figure 7.** (a) STM topograph of the Si(111) $7 \times 7$  reconstructed surface with a step along which molecules were found to be preferentially adsorbed. The image was taken with negative sample bias voltage polarity. (b) Corresponding STM image obtained with positive sample bias voltage polarity. The adsorbed molecules have become almost invisible. The Si(111) $7 \times 7$  surface also appears different depending on the bias voltage polarity.

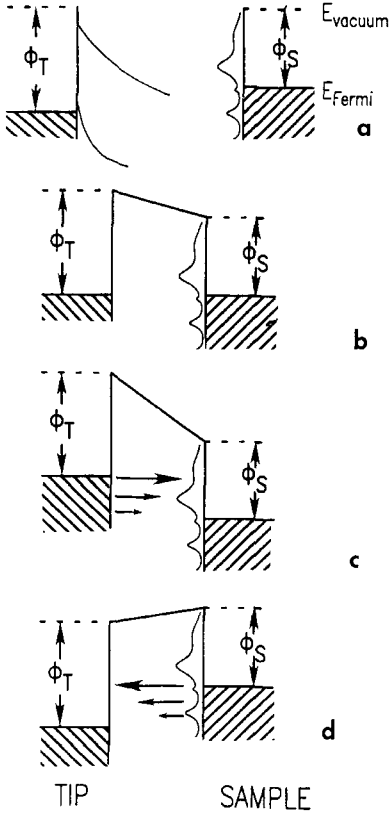
finite bias to a first approximation.) For positive sample bias, the net tunneling current arises from electrons that tunnel from the occupied states of the tip into unoccupied states of the sample (Fig. 8c), whereas at negative sample bias, electrons tunnel from occupied states of the sample into unoccupied states of the tip. Consequently, the bias polarity determines whether unoccupied or occupied sample electronic states are probed. It also becomes clear that the electronic structure of the tip enters as well, as is also obvious from Eq. (11) for the tunneling current.

By varying the amount of the applied bias voltage, one can select the electronic states that contribute to the tunneling current and, in principle, measure the local electronic density of states. For instance, the current increases strongly if the applied bias voltage allows the onset of tunneling into a maximum of the unoccupied sample electronic density of states. Therefore, the first derivative  $dI/dU(U)$  reflects the electronic density of states to a

first approximation. However, the energy and bias dependence of the transmission coefficient  $T(\mathcal{E}, eU)$  has also to be considered. Since electrons in states with the highest energy 'see' the smallest effective barrier height, most of the tunneling current arises from electrons near the Fermi level of the negatively biased electrode. This has been indicated in Fig. 8 by arrows of differing size. The maximum in the transmission coefficient  $T(\mathcal{E}, eU)$  given in Eq. (12) can be written as [26]

$$T_{\max}(U) = \exp \left\{ -2(s + R) \times \left[ \frac{2m}{\hbar^2} \left( \frac{\phi_t + \phi_s}{2} + \frac{|eU|}{2} \right) \right]^{1/2} \right\} \quad (25)$$

The bias dependence of the transmission coefficient typically leads to an order-of-magnitude increase in the tunneling current for each volt increase in magnitude of the applied bias voltage. Since the transmission coefficient increases monotonically with the applied bias voltage, it contributes only a smoothly varying



**Figure 8.** Energy level diagrams for the sample and the tip. (a) Independent sample and tip. (b) Sample and tip at equilibrium, separated by a small vacuum gap. (c) Positive sample bias: electrons tunnel from the tip to the sample. (d) Negative sample bias: electrons tunnel from the sample into the tip. (Image courtesy of Hamers [24].)

background on which the density-of-states information is superimposed.

As an important consequence of the dominant contribution of tunneling from states near the Fermi level of the negatively biased electrode, tunneling from the tip to the sample (Fig. 8c) mainly probes the empty states of the sample with negligible influence of the occupied states of the tip. On the other hand, tunneling from the sample to the tip is much more sensitive to the electronic structure of the empty

states of the tip, which often prevents detailed spectroscopic STM studies of the occupied states of the sample [27].

### 1.4.1 Scanning Tunneling Spectroscopy at Constant Current

To perform scanning tunneling spectroscopy measurements, a high-frequency sinusoidal modulation voltage can be superimposed on the constant bias voltage applied between the tip and the sample. The modulation frequency is chosen higher than the cut-off frequency of the feedback loop, which keeps the average tunneling current constant. By recording the tunneling current modulation, which is in-phase with the applied bias voltage modulation, by means of a lock-in amplifier, a spatially resolved spectroscopy signal  $dI/dU$  can be obtained simultaneously with the constant current image [28, 29]. Based on expression (11) for the tunneling current and by assuming  $dn_t/dU \approx 0$ , one obtains [24]

$$\begin{aligned} \frac{dI}{dU}(U) &\propto en_t(0)n_s(eU)T(eU, eU) \\ &+ \int_0^{eU} n_t(\pm eU \mp \mathcal{E})n_s(\mathcal{E}) \\ &\times \frac{dT(\mathcal{E}, eU)}{dU} d\mathcal{E} \end{aligned} \quad (26)$$

At a fixed location, the increase of the transmission coefficient with applied bias voltage is smooth and monotonic. Therefore, structure in  $dI/dU$  as a function of  $U$  can usually be attributed to structure in the state density via the first term in Eq. (26). However, interpretation of the spectroscopic data  $dI/dU$  as a function of position

$(x, y)$  is more complicated. As discussed in Sec. 1.3.2 of this Chapter, the apparent barrier height above a local topographic protrusion is larger, that is, the transmission coefficient is smaller, than above a local topographic depression. This spatial variation in the transmission coefficient shows up in spatially resolved measurements of  $dI/dU$  as a ‘background’ that is essentially an ‘inverted’ constant-current topography. Therefore, spectroscopic images corresponding to the spatial variation of  $dI/dU$  obtained in the constant current mode in fact contain a superposition of topographic and electronic structure information.

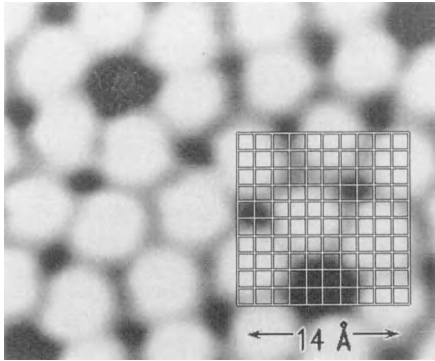
### 1.4.2 Local Spectroscopic Measurements at Constant Separation

To eliminate the influence of the  $z$ -dependence of the transmission coefficient, local  $I-U$  characteristics can be measured at a fixed tip-sample separation. This is achieved by breaking the feedback circuit for a certain time interval at selected surface locations by means of a sample-and-hold amplifier, while local  $I-U$  curves are recorded [30, 31]. The  $I-U$  characteristics are usually repeated several times at each surface location and finally signal averaged. Since the feedback loop is inactive while sweeping the applied bias voltage, the tunneling current is allowed to become extremely small. Therefore, band gap states in semiconductors, for instance, can be probed without difficulties. The first derivative  $dI/dU$  can be obtained from the measured  $I-U$  curves by numerical differentiation. The dependence of the

measured spectroscopic data on the value of the tunneling conductance  $I/U$  can be compensated by normalizing the differential conductance  $dI/dU$  to the total conductance  $I/U$ . The normalized quantity  $(dI/dU)/(I/U)$  reflects the electronic density of states reasonably well by minimizing the influence of the tip-sample separation [32]. However, the close resemblance of the  $(d \ln I / d \ln U) - U$  curve to the electronic density of states is generally limited to the position of peaks while peak intensities can differ significantly.

### 1.4.3 Current Imaging Tunneling Spectroscopy

The measurements of local  $I-U$  curves at a constant tip-sample separation can be extended to every pixel in an image, which allows performance of atomically resolved spectroscopic studies [33]. The method, denoted current imaging tunneling spectroscopy (CITS), also uses a sample-and-hold amplifier to alternately gate the feedback control system on and off. During the time of active feedback, a constant stabilization voltage  $U_0$  is applied to the sample, and the tip height is adjusted to maintain a constant tunneling current. When the feedback system is deactivated, the applied sample bias voltage is linearly ramped between two preselected values, and the  $I-U$  curve is measured at a fixed tip height. Afterwards, the applied bias voltage is set back to the chosen stabilization voltage  $U_0$  and the feedback system is reactivated. By acquiring the  $I-U$  curves rapidly while scanning the tip position at low speed, a constant-current topograph and spatially resolved  $I-U$  characteristics



**Figure 9.** STM topograph of the unoccupied states of an Si(111)7 × 7 surface (sample bias 2 V). The atoms imaged are the top-layer Si adatoms (top). The grid encompasses a 14 Å × 14 Å area of this surface for which tunneling spectra have been obtained. The 100 tunneling spectra are plotted in the  $dI/dU$  form. Such spectral maps allow one not only to obtain the energies of the occupied (negative bias) and unoccupied (positive bias) states of particular atomic sites, but also to obtain information on the spatial extent of their wave functions (bottom). (Image courtesy of Avouris [34].)

can simultaneously be obtained. To increase the possible scan speed and to decrease the amount of data to be stored, one can predefine a coarse grid of pixels in the image at which local  $I-U$  curves will be measured (Fig. 9). The ability to probe the local electronic structure down to atomic scale has great potential, for instance, for investigation of surface chemical reactivity on an atom-by-atom basis [34].

## 1.5 Spin-Polarized Scanning Tunneling Microscopy

Thus far, the dependence of the tunneling current on the tip-sample separation  $s$ , the local barrier height, and the applied sample bias voltage  $U$  has been considered:

$$I = I(s, \phi, U) \quad (27)$$

Accordingly, the corresponding modes of STM operation have been discussed: ‘topographic’ imaging, local barrier height imaging, and tunneling spectroscopy. However, the spin of the tunneling electrons and the additional spin dependence of the tunneling current, if magnetic electrodes are involved, have not yet been considered:

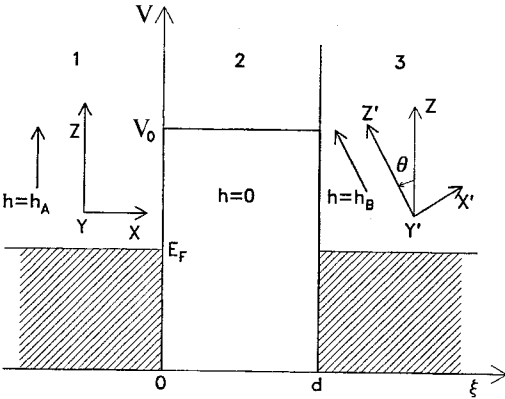
$$I = I(s, \phi, U, \uparrow) \quad (28)$$

By using this spin dependence of the tunneling current in spin-polarized STM (SPSTM) experiments, magnetic information about solid surfaces can be obtained.

Spin-dependent tunneling had been observed in the 1970s using planar tunnel junctions [35–37]. To explain the experimental results, Slonczewski [38] considered a tunnel junction with two ferromagnetic electrodes where the directions of the internal magnetic fields differ by an angle  $\theta$  (Fig. 10). Within a free-electron model and in the limit of a small applied bias voltage, the following expression for the conductance  $\sigma$  of the ferromagnet/insulator/ferromagnet tunnel junction for the case of two identical ferromagnetic electrodes was found:

$$\sigma = \sigma_{\text{fbf}}(1 + P_{\text{fb}}^2 \cos \theta), |P_{\text{fb}}| \leq 1 \quad (29)$$

Here,  $P_{\text{fb}}$  denotes the effective spin polarization of the ferromagnet/barrier interface



**Figure 10.** Schematic potential diagram for two metallic ferromagnets separated by an insulating barrier. The molecular fields  $h_A$  and  $h_B$  within the magnets form an angle  $\theta$ . (Image courtesy of Slonczewski [38].)

and  $\sigma_{fbf}$  is a mean conductance which is proportional to  $\exp(-2\kappa s)$ . If the ferromagnetic electrodes are different, the conductance becomes

$$\sigma = \sigma_{fbf}(1 + P_{fb}P_{fb'} \cos \theta) \quad (30)$$

For the two special cases of parallel and antiparallel alignment of the internal magnetic field directions, one finds

$$\begin{aligned} \sigma_{\uparrow\uparrow} &= \sigma_{fbf}(1 + P_{fb}P_{fb'}) \\ \sigma_{\uparrow\downarrow} &= \sigma_{fbf}(1 - P_{fb}P_{fb'}) \end{aligned} \quad (31)$$

Consequently, one obtains

$$\frac{\sigma_{\uparrow\uparrow} - \sigma_{\uparrow\downarrow}}{\sigma_{\uparrow\uparrow} + \sigma_{\uparrow\downarrow}} = P_{fb}P_{fb'} = P_{fbf} \quad (32)$$

where  $P_{fbf}$  is the effective polarization for the whole tunnel junction. An experimental determination of the quantity on the left-hand side of Eq. (32) by means of SPSTM offers a way to derive the effective polarization  $P_{fbf}$  locally with a spatial resolution comparable to that of topographic STM images and therefore ultimately on the atomic scale.

The spin dependence of the tunneling current in SPSTM experiments with two magnetic electrodes in a zero external magnetic field was demonstrated by using a ferromagnetic  $\text{CrO}_2$  probe tip and a  $\text{Cr}(001)$  surface [39]. The topological antiferromagnetic order of the  $\text{Cr}(001)$  surface [40] with alternately magnetized terraces separated by monoatomic steps was confirmed. In addition, a local effective polarization of the  $\text{CrO}_2/\text{vacuum}/\text{Cr}(001)$  tunnel junction was derived. Later, atomic resolution in SPSTM experiments has been demonstrated on a magnetite ( $\text{Fe}_3\text{O}_4$ ) (001) surface, where the two different magnetic ions  $\text{Fe}^{2+}$  and  $\text{Fe}^{3+}$  on the Fe B-sites in the Fe-O (001) planes could be distinguished by using an atomically sharp Fe probe tip prepared in situ [41].

With an additional external magnetic field applied, the magnetization of the sample (or of the tip) can be modulated periodically, for instance, from parallel to antiparallel alignment relative to the tip (or the sample) magnetization direction. Consequently, a portion of the tunneling current is predicted to oscillate at the same frequency, with an amplitude linearly proportional to the average tunneling current [42]. The advantage of this experimental procedure lies in the fact that lock-in detection techniques can be used, resulting in an improvement of the signal-to-noise ratio. In principle, the magnetic field can be modulated at a frequency  $\nu_0$  well above the cut-off frequency of the feedback loop, and the corresponding amplitude of the current oscillation at the frequency  $\nu_0$  can be recorded with a lock-in amplifier simultaneously with the constant-current topograph. The spatially resolved lock-in signal then provides a map of the effective spin polarization.

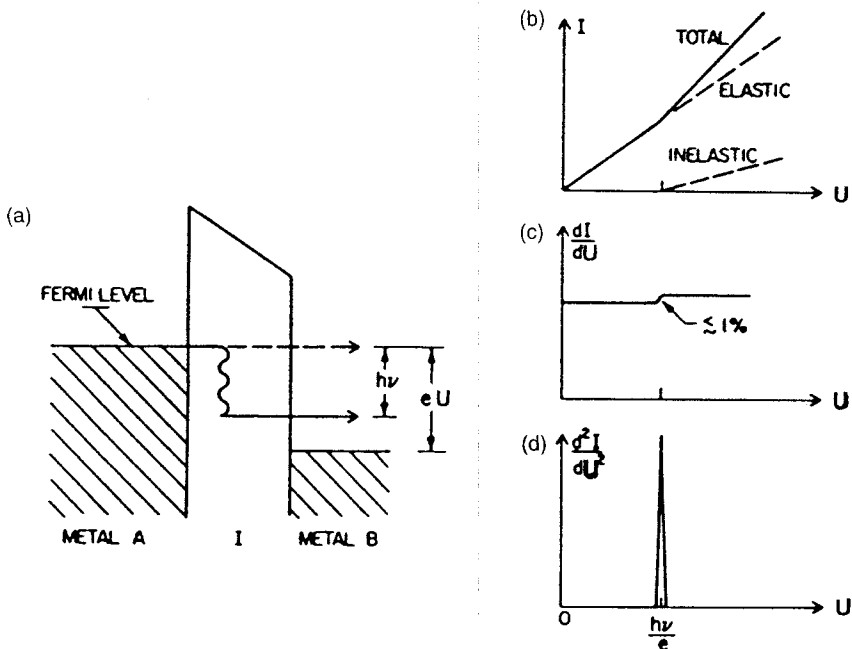
A third approach to SPSTM is to use GaAs either as samples [43, 44] or as tips. It is well known that GaAs optically pumped by circularly polarized light provides an efficient source for spin-polarized electrons. On the other hand, one can measure the circular polarization of the recombination luminescence light induced by electrons tunneling from a ferromagnetic counterelectrode.

## 1.6 Inelastic Tunneling Spectroscopy

Besides elastic tunneling processes, in which the energy of the tunneling electrons

is conserved, inelastic tunneling can occur where the electron energy is changed due to interaction of the tunneling electrons with elementary excitations. In Fig. 11 a potential energy diagram is shown, illustrating elastic and inelastic tunneling processes. In the case of inelastic tunneling, the electron loses a quantum of energy  $h\nu_0$  to some elementary excitation mode. According to the Pauli exclusion principle, the final state after the inelastic tunneling event must be initially unoccupied as depicted in Fig. 11a. Consequently, the bias dependence of the tunneling current (Fig. 11b) shows the behavior described below.

Starting from a zero applied bias voltage  $U$ , the elastic tunneling current increases linearly proportional to  $U$



**Figure 11.** (a) Tunneling electrons can excite a molecular vibration of energy  $h\nu$  only if  $eU > h\nu$ . For smaller voltages there are no vacant final states for the electrons to tunnel into. (b) Thus the inelastic current has a threshold at  $U = h\nu/e$ . (c) The increase in conductance at this threshold is typically below 1%. (d) A standard tunneling spectrum,  $d^2I/dU^2$  versus  $U$ , accentuates this small increase; the step in  $dI/dU$  becomes a peak in  $d^2I/dU^2$ . (Image courtesy of Hansma [45].)

(Eq.(4)). As long as the applied bias voltage is sufficiently small ( $U < h\nu_0/e$ , where  $\nu_0$  is the lowest energy excitation mode), inelastic tunneling processes cannot occur due to the Pauli exclusion principle. At the threshold bias  $U_0 = h\nu_0/e$ , the inelastic channel opens up, and the number of electrons which can use the inelastic channel will increase linearly with  $U$  (Fig. 11b). Therefore, the total current, including both elastic and inelastic contributions, has a kink at  $U_0 = h\nu_0/e$ . In the conductance ( $dI/dU$ ) versus voltage curve, the kink becomes a step at  $U_0$ . Since the fraction of electrons which tunnel inelastically is tiny (typically 0.1–1%), the conductance increase at  $U_0$  due to the onset of the inelastic tunnel channel is too small to be conveniently observed. Therefore, the second derivative ( $d^2I/dU^2$ ) is usually measured, which exhibits a peak at  $U_0$ . In general, there are many modes which can be excited in the tunneling process. Each excitation mode  $\nu_i$  contributes a peak in the second derivative  $d^2I/dU^2(U)$  at the corresponding bias voltage  $U_i = h\nu_i/e$  so that  $d^2I/dU^2(U)$  represents the spectrum of possible excitations. Inelastic electron tunneling can therefore be regarded as a special kind of electron energy loss spectroscopy.

To be able to detect the small changes in tunneling conductance  $\Delta G/G$  as a result of the opening of additional inelastic tunneling channels, the relative stability of the tunneling current has to be better than 1%. In addition, low temperatures are required to keep thermal linewidth broadening in the spectra, which is of the order of  $k_B T$ , small compared with the energy  $h\nu$  of the excitation modes,  $\nu$  being typically a few millielectronvolts.

### 1.6.1 Phonon Spectroscopy

Electron tunneling between the probe tip and the sample can create phonons at the interface between the conductor and the tunneling barrier. The emission of phonons is believed to take place within a few atomic layers of the interface. Low-temperature STM experiments with a tungsten probe tip and a graphite sample indeed revealed a spectrum of peaks in  $d^2I/dU^2(U)$  characteristics where the positions of the peaks corresponded closely to the energies of the phonons of the graphite sample and the tungsten tip [46]. The measured increase in conductance at the phonon energies was of the order of 5%. By analogy with elastic scanning tunneling spectroscopy (see Sec. 1.4.1 of this Chapter), spectroscopic imaging can be performed by recording  $d^2I/dU^2$  at a particular phonon energy while scanning the tip over the sample surface. This method allows one to map spatial variations of the phonon spectra, caused by spatial variations in the coupling between the tunneling electrons and the phonons, on the atomic scale.

### 1.6.2 Molecular Vibrational Spectroscopy

Inelastic tunneling spectroscopy can also yield information about vibrational modes of molecules adsorbed on a surface. By using low-temperature STM, a vibrational spectrum of an individual adsorbed molecule can be obtained by positioning the probe tip over the preselected adsorbate. It is even possible to form a map showing the sites within a molecule where particular resonances occur.

For sorbic acid adsorbed on graphite, a spectrum of strong peaks was observed in the first derivative  $dI/dU$  instead of the expected second derivative  $d^2I/dU^2$  [47]. The energies of the peaks corresponded approximately to the vibrational modes of the molecule. The measured increase in conductivity at the molecular resonances was as much as a factor of 10, which is at least two orders of magnitude larger than expected.

Future inelastic tunneling experiments have to focus on the assignment of characteristic features in the tunneling spectra to particular molecular functional groups. This will probably help to identify chemical species by STM, a problem which is not solvable by elastic tunneling spectroscopy.

## Acknowledgments

The author would like to acknowledge all colleagues who provided illustrations for this Chapter.

## 1.7 References

- [1] G. Binnig, H. Rohrer, Ch. Gerber, E. Weibel, *Phys. Rev. Lett.* **1982**, 49, 57.
- [2] G. Binnig and H. Rohrer, *Helv. Phys. Acta* **1982**, 55, 726.
- [3] G. Binnig and H. Rohrer, *Rev. Mod. Phys.* **1987**, 59, 615.
- [4] R. Wiesendanger, *Scanning Probe Microscopy and Spectroscopy: Methods and Applications*, Cambridge University Press, Cambridge **1994**.
- [5] J. Bardeen, *Phys. Rev. Lett.* **1961**, 6, 57.
- [6] J. Tersoff, D. R. Hamann, *Phys. Rev. Lett.* **1983**, 50, 1988.
- [7] J. Tersoff, D. R. Hamann, *Phys. Rev. B* **1985**, 31, 805.
- [8] C. J. Chen, *J. Vac. Sci. Technol.* **1988**, A6, 319.
- [9] S. Ohnishi, M. Tsukada, *J. Vac. Sci. Technol.* **1990**, A8, 174.
- [10] C. J. Chen, *Phys. Rev. B* **1990**, 42, 8841.
- [11] C. J. Chen, *J. Vac. Sci. Technol.* **1991**, A9, 44.
- [12] N. D. Lang, *Phys. Rev. Lett.* **1985**, 55, 230.
- [13] N. D. Lang, *Phys. Rev. Lett.* **1986**, 56, 1164.
- [14] N. D. Lang, *IBM J. Res. Develop.* **1986**, 30, 374.
- [15] Y. Kuk, P. J. Silverman, H. Q. Nguyen, *J. Vac. Sci. Technol.* **1988**, A6, 524.
- [16] J. Wintterlin, J. Wiechers, H. Brune, T. Gritsch, H. Höfer, R. J. Behm, *Phys. Rev. Lett.* **1989**, 62, 59.
- [17] M. Bode, R. Pascal, R. Wiesendanger, *Z. Phys. B*, in press.
- [18] V. M. Hallmark, S. Chiang, J. F. Rabolt, J. D. Swalen, R. J. Wilson, *Phys. Rev. Lett.* **1987**, 59, 2879.
- [19] A. Baratoff, *Physica B* **1984**, 127, 143.
- [20] C. J. Chen, *Phys. Rev. Lett.* **1990**, 65, 448.
- [21] G. Binnig, N. Garcia, H. Rohrer, J. M. Soler, F. Flores, *Phys. Rev. B* **1984**, 30, 4816.
- [22] J. H. Coombs, M. E. Welland, J. B. Pethica, *Surf. Sci.* **1988**, 198, L353.
- [23] G. Binnig, H. Rohrer, *Surf. Sci.* **1983**, 126, 236.
- [24] R. J. Hamers, *Annu. Rev. Phys. Chem.* **1989**, 40, 531.
- [25] R. M. Tromp, *J. Phys. C: Condens. Matter* **1989**, 1, 10211.
- [26] R. M. Feenstra, *NATO ASI Ser. E: Appl. Sci.* **1990**, 184, 211.
- [27] T. Klitsner, R. S. Becker, J. S. Vickers, *Phys. Rev. B* **1990**, 41, 3837.
- [28] G. Binnig, K. H. Frank, H. Fuchs, N. Garcia, B. Reihl, H. Rohrer, F. Salvan, A. R. Williams, *Phys. Rev. Lett.* **1985**, 55, 991.
- [29] R. S. Becker, J. A. Golovchenko, D. R. Hamann, B. S. Swartzentruber, *Phys. Rev. Lett.* **1985**, 55, 2032.
- [30] R. M. Feenstra, W. A. Thompson, A. P. Fein, *Phys. Rev. Lett.* **1986**, 56, 608.
- [31] J. S. Strosio, R. M. Feenstra, A. P. Fein, *Phys. Rev. Lett.* **1986**, 57, 2579.
- [32] R. M. Feenstra, J. A. Strosio, *Phys. Scripta* **1987**, T19, 55.
- [33] R. J. Hamers, R. M. Tromp, J. E. Demuth, *Phys. Rev. Lett.* **1986**, 56, 1972.
- [34] P. Avouris, I.-W. Lyo in *Chemistry and Physics of Solid Surfaces VIII. Springer Series in Surface Sciences*, Vol. 22 (Eds.: R. Vanselow, R. Howe), Springer, Berlin, **1990**, p. 371.
- [35] P. M. Tedrow, R. Meservey, *Phys. Rev. Lett.* **1971**, 26, 192.
- [36] M. Julliere, *Phys. Lett. A* **1975**, 54, 225.
- [37] S. Maekawa, U. Gäfvert, *IEEE Trans. Magn.* **1982**, 18, 707.
- [38] J. C. Slonczewski, *Phys. Rev. B* **1989**, 39, 6995.
- [39] R. Wiesendanger, H.-J. Güntherodt, G. Güntherodt, R. J. Gambino, R. Ruf, *Phys. Rev. Lett.* **1990**, 65, 247.



- [40] S. Blügel, D. Pescia, P. H. Dederichs, *Phys. Rev. B* **1989**, 39, 1392.
- [41] R. Wiesendanger, I. V. Shvets, D. Bürgler, G. Tarrach, H.-J. Güntherodt, J. M. D. Coey, *Europhys. Lett.* **1992**, 19, 141.
- [42] M. Johnson, J. Clark, *J. Appl. Phys.* **1990**, 67, 6141.
- [43] S. F. Alvarado, P. Renaud, *Phys. Rev. Lett.* **1992**, 68, 1387.
- [44] K. Sueoka, K. Mukasa, K. Hayakawa, *Jpn. J. Appl. Phys.* **1993**, 32, 2989.
- [45] P. K. Hansma, *Phys. Rep.* **1977**, 30, 145.
- [46] D. P. E. Smith, G. Binnig, C. F. Quate, *Appl. Phys. Lett.* **1986**, 49, 1641.
- [47] D. P. E. Smith, M. D. Kirk, C. F. Quate, *J. Chem. Phys.* **1987**, 86, 6034.

## 2 Scanning Force Microscopy

### 2.1 Introduction

The invention of scanning tunneling microscopy (STM) in 1982 [1] triggered the development of several further techniques which use scanned point probes to sense local properties of surfaces [2]. Among these techniques, scanning force microscopy (SFM), which was originally denoted atomic force microscopy (AFM) [3], has developed into the most widespread and commercially successful tool, and is used not only in physical, chemical, biological, and medical research laboratories, but also by many companies for tasks such as product development and routine quality control.

The history of SFM started in the fall of 1985, when Gerd Binnig and Christopher Gerber, on leave from IBM's research laboratory in Zurich, Switzerland, built the prototype of a force microscope together with Calvin Quate at his group at Stanford University, California, U.S.A. [3]. The success of SFM in the following years was due to the high resolution and the versatility of this new tool, which can map not only the pure topography of sample surfaces from the micrometer to the atomic scale, but also (often additionally to the simultaneously recorded topography) many other physical properties

which are related to forces. Moreover, due to its universality, SFM can be applied to a large variety of samples. It requires no special sample preparation and can be adapted to many different environments such as air, vacuum, and liquids. These issues are exemplified in detail below.

The principle of SFM is rather simple, and is analogous to that of a record player. A force microscope (see Fig. 1) detects forces acting between a sample surface and a sharp tip which is mounted on a soft leaf spring, the so-called cantilever. A feedback system which controls the vertical  $z$ -position of the tip on the sample surface keeps the deflection of the cantilever (and thus the force between the tip and sample) constant. Moving the tip relative to the sample in the  $(x, y)$  plane of the surface by means of piezoelectric drives, the actual  $z$ -position of the tip can be recorded as a function of the lateral  $(x, y)$  position. The obtained three-dimensional data represent a map of equal forces. The data can be analyzed and visualized through computer processing.

With the concept described above, it is possible to detect any kind of force as long as the tip is sensitive enough to the interaction, that is, the interaction causes a measurable deflection of the cantilever on which the tip is mounted. Therefore, not only interatomic forces

but also long-range forces such as magnetic or electrostatic forces can be detected and mapped. In order to manifest the general concept, the original name 'atomic force microscopy' [3] was replaced by the term 'scanning force microscopy', which is used in this chapter. In a typical force microscope, cantilever deflections in the range from 0.1 Å to a few micrometers are measured [4–6], corresponding to forces from  $10^{-13}$  to  $10^{-5}$  N. For comparison, typical forces in conventional profilometers are  $10^{-4}$  N or greater [7].

Force microscopes reach an impressive magnification. Images of atomic lattices have been demonstrated with a lateral resolution well below 1 Å. The maximum scan range is usually limited by the choice of the piezoelectric scanner. Most manufacturers supply scanners up to at least the  $100 \times 100 \mu\text{m}$  range. Even if large areas are imaged, the resolution in the *z*-direction is still sufficient to image monoatomic steps (cf. Fig. 4b).

There are only few limitations in the type of sample which can be investigated by SFM. A sample is suitable to be studied by SFM as long as it is solid and clean. For example, photographic material, which is an insulator and highly sensitive to light or electron beams, can be easily studied without any surface damage [8]. In a liquid environment, even soft samples such as the biological membrane of a virus or a cell have been successfully imaged [9, 10]. However, if powder particles [11], grains from colloidal solution [12], single molecules [13], or similar samples are to be investigated, they should adhere to a substrate by forces stronger than the lateral forces of the scanning tip. This is often realized by physisorption [14] or chemisorption [15].

The possible *sample size* is, in principle, *unlimited*, since stand-alone microscopes can be put on the sample itself [16]. Special large-sample microscopes are sold for routine quality control of whole wafers in the electronic industry at different steps of fabrication. After the spot check, the wafer can be reintroduced into the production process even if sensitive devices are being produced, since the sample is not affected by the measurement.

Probably the most important limiting factor for the successful operation of SFM on a hard solid sample is the cleanliness of the surface which is to be imaged. Images on the micrometer scale can be obtained on nearly every sample unless they show extraordinary adhesion due to, for example, surfactants or thick adsorbent films. For high-resolution SFM micrographs, however, the absence of adsorbents which disturb the correct profiling of the surface is essential. Crystals which are to be imaged on the atomic scale should preferably be cleaved before the measurements.

Samples can be probed without any special treatment and in many different environmental conditions, such as in ambient air, a large variety of liquids [17, 18], gaseous atmospheres such as dry nitrogen or argon [19–21], and ultrahigh vacuum (UHV) [4, 22–24]. Measurements can be performed at different temperatures starting from temperatures close to absolute zero [25] up to temperatures well above room temperature. Additionally, SFM can easily be combined in situ with optical microscopy.

Section 2.2 of this Chapter discusses experimental aspects such as the properties of force sensors, different experimental set-ups, and possible imaging modes. Some

theoretical aspects are considered in Sec. 2.3 of this Chapter. Further general information about SFM can be found in reviews [26] and textbooks [27].

## 2.2 Experimental Aspects

Figure 1 shows a typical set-up of a force microscope. Essential components of every force microscope are (1) a local force probe (a sharp tip) which is mounted on a soft spring (the cantilever), (2) a sensor for the accurate measurement of the cantilever deflection, (3) a device for the relative motion of the tip and sample

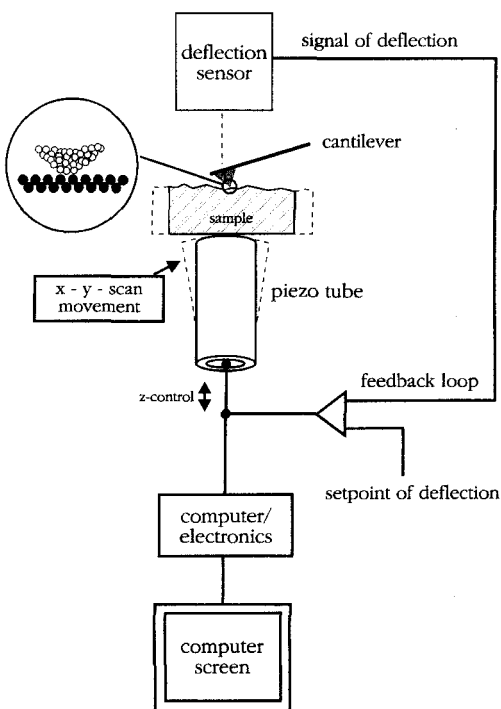
(usually a piezoelectric scanner), (4) a feedback system to control the cantilever deflection, and (5) computer-based electronics for the visualization and analysis of the measured data. Components (1)–(3) are assembled in a rigid mechanical set-up. Components (3)–(5) are identical in principle to the corresponding devices used in STM, and will not be described here. General design criteria for scanning probe microscopes are described in detail in the literature [28].

However, since the principle of SFM is based on the measurement of the forces acting between a sharp tip and the sample, the preparation of suitable tips and cantilevers is of pre-eminent importance, and will be discussed in Sec. 2.2.1 of this Chapter. Even small deflections of the cantilever have to be detected (see Sec. 2.2.2). Finally, force microscopes can be operated in different modes, which are analyzed in Secs. 2.2.3–2.2.6 of this Chapter.

### 2.2.1 The Force Sensor

In force microscopy, both the spring and the force probe (the sharp tip), which is mounted on the spring, have to satisfy certain requirements. There are basically two requirements for the spring:

(1) The transmission of an external vibration to the system is described by  $a_{\text{trans}} = a_0(f_0/f_{\text{res}})^2$ , where  $f_0$  is a frequency of the vibration amplitude  $a_0$  and  $f_{\text{res}}$  is the resonance frequency of the spring. It is usual to choose cantilevers with a high value of  $f_{\text{res}}$  to omit problems due to acoustic or mechanical vibrations of low frequency. This means, since  $f_{\text{res}}$  is



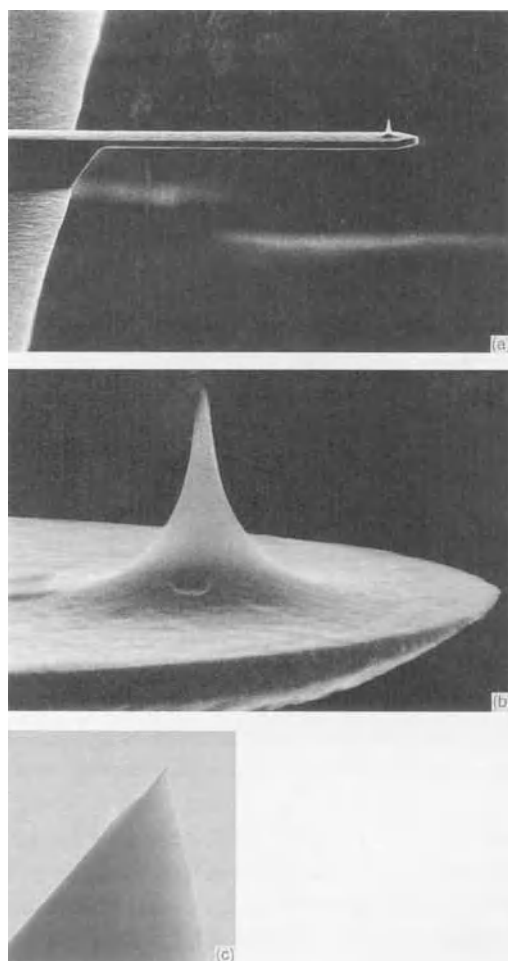
**Figure 1.** Schematic of the set-up of a typical force microscope operating in the constant force imaging mode.

proportional to  $\sqrt{c/m}$  ( $c$  = spring constant,  $m$  = mass of the spring), that the ratio  $c/m$  has to be large.

(2) The spring constant  $c$  should be small to keep the forces acting on the sample surface as small as possible. Ideally, it should be well below interatomic spring constants, which are of the order of  $10 \text{ N m}^{-1}$ .

Conditions (1) and (2) can only be fulfilled simultaneously if the mass  $m$  of the spring is small. Today, springs of a small mass are produced by standard microfabrication techniques from silicon or silicon nitride in the form of rectangular or 'V'-shaped cantilevers (Fig. 2a) [29]. The typical dimensions of such cantilevers are as follows: lengths of some hundreds of micrometers, widths of some tens of micrometers, and thicknesses of  $0.3\text{--}5 \mu\text{m}$ . Since the dimensions are within a quite narrow range, the spring constants can be calculated accurately. Spring constants and resonance frequencies of cantilevers used in contact SFM measurements are about  $0.01\text{--}1 \text{ N m}^{-1}$  and  $5\text{--}100 \text{ kHz}$ , respectively.

A tip which acts as a local probe is mounted at the end of the cantilever (Fig. 2b,c). The first tips were simply small pieces of diamond glued to the end of cantilevers which were cut from metal foil [3]. Later, microfabrication techniques were used to produce the cantilevers, to which the diamonds were glued [29]. Sometimes, thin metal wires were bent and etched at their ends. This kind of tip was frequently used in magnetic force microscopy. During recent years, however, new techniques have been developed to produce microfabricated cantilevers with integrated tips of high quality [30].



**Figure 2.** (a) Scanning electron micrograph of a rectangular silicon cantilever of  $105 \mu\text{m}$  length and  $14.5 \mu\text{m}$  width. (b) Scanning electron micrograph of the tip at the end of the cantilever shown in part (a). The tip height is  $6.5 \mu\text{m}$ . (c) Transmission electron micrograph of the tip end of a silicon cantilever. The image size is  $2 \times 2 \mu\text{m}$ , and the tip radius is estimated to be below  $10 \text{ nm}$ .

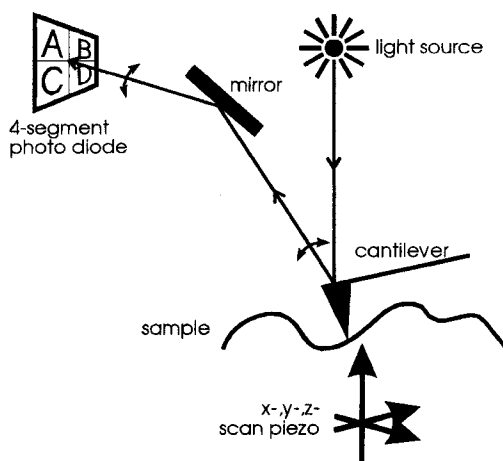
The shape of an SFM is often approximated by a cone which has a small half-sphere at its end. The tip should be as sharp as possible to measure very local properties, that is, it should possess a small opening angle for the cone and a

small radius for the sphere. On the other hand, tips with a small opening angle are often unstable and tend to flex or break. Silicon nitride tips are more robust than silicon tips due to the greater hardness of silicon nitride. However, silicon nitride tips are currently not available with such small opening angles and tip radii as silicon tips. Therefore, it is important to choose a suitable tip for a specific application [31].

## 2.2.2 Deflection Sensors

Several techniques have been developed to detect the deflection of the cantilever, which has to be measured with sub-angstrom resolution. The most frequently used techniques are reviewed in this section.

The beam deflection technique was developed independently by Meyer and Amer [4] and Alexander et al. [5] in 1988, and is currently the most widespread technique used to measure cantilever deflections in force microscopy. The idea behind this technique is presented in Fig. 3. A light beam from a laser diode or a high-power light-emitting diode is reflected from the rear side of the cantilever and focused onto a four-segment photodiode. If  $I_A$ ,  $I_B$ ,  $I_C$ , and  $I_D$  are the currents which are induced by the light in segments A–D of the photodiode, then the current  $(I_A + I_B) - (I_C + I_D)$  represents a measure of the deflection of the cantilever. Additionally, the torsion of the cantilever can be measured through the analysis of the  $(I_A + I_C) - (I_B + I_D)$  current, which is proportional to the lateral force acting on the tip (cf. Sec. 2.2.3.4 of this Chapter) [32]. This technique is used in most



**Figure 3.** Schematic drawing of the beam deflection technique for the sensing of cantilever deflections in force microscopy. The light from a light source is reflected at the back face of the cantilever and focused onto a four-segment photodiode. By analyzing the photocurrents induced in the different segments A–D, the deflection as well as the torsion of the cantilever can be detected simultaneously.

commercially available force microscopes, mainly due to a high z-resolution (typically of 0.1 Å) and the easy to measure torsion of the cantilever.

Interferometric detection schemes were introduced even before the beam deflection technique was presented [33, 34]. The idea underlying this detection scheme is explained below. A laser beam, reflected from the rear of the cantilever, interferes with a reference beam, the changes in intensity of the signal are proportional to the deflection of the cantilever. Today, most of the interferometers used to sense cantilever deflection in SFM are based on glass fiber technology [6]. This allows the positioning of the light source and the photodetector at a distance from the force microscope, enabling a compact microscope design [35]. Furthermore, miniaturized devices with dimensions of

only a few millimeters are under development. The resolution which can be obtained with this technique is similar to the beam deflection technique [36].

An alternative optical method is laser diode interferometry [37]. The light from a laser diode is reflected from the rear of the cantilever back into the laser cavity. Laser diodes are extremely sensitive to this type of feedback, which causes measurable changes in the laser intensity. In spite of the simplicity of the set-up, the analysis of the signal is rather complex [38].

A comparatively new approach is the use of piezoresistive cantilevers, which serve as deflection sensors [39]. A piezoresistive film at the rear of the cantilever changes its resistance if the cantilever is bent. The change of resistance is proportional to the deflection, and is detected by a Wheatstone bridge. The main advantage of this technique is that the size of the force microscope can be reduced since an optical part is not necessary. However, the big disadvantage of this technique is the still insufficient sensitivity when combined with soft cantilevers for contact measurements. Therefore, such microscopes are mostly operated in the non-contact dynamic mode (see Sec. 2.2.3.3 of this Chapter) [24, 40].

Historically, the first techniques were the electron tunneling technique and the capacitance technique. The electron tunneling technique was applied in the first force microscope of Binnig, Gerber, and Quate [3] and in several other early force microscopes [41]. In these the deflection of the cantilever was sensed by a tunneling junction between the back of the cantilever and an additional tip. It showed an excellent  $z$ -resolution due to the exponential dependence of the tunneling current on the separation of the electrodes (see Sec. 1 of

this Chapter). Many experimental difficulties such as the instability of the tunneling junction and a sensitivity to the surface roughness of the rear of the cantilever led to the abandonment of this technique.

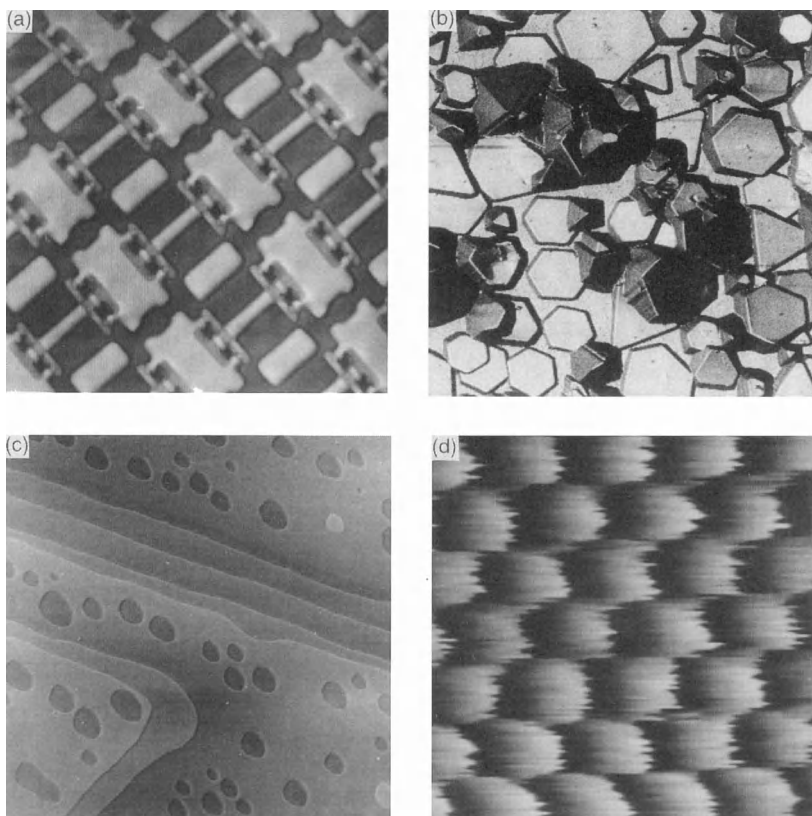
Force microscopes equipped with capacitance sensors do not yet show a performance comparable to other methods [23, 42]. In these microscopes, the change in the capacitance between the rear side of the cantilever and an external electrode is measured. In the future, microfabricated capacitances which are integrated into the cantilever design might significantly improve the reliability and sensitivity of this method.

### 2.2.3 Imaging Modes

Different imaging scenarios are realized depending on which physical parameter is taken to the feedback system. The speed of the feedback system and the number of parameters which are monitored and recorded determine the imaging operation mode. A survey of the most common modes of SFM operation is given below.

#### 2.2.3.1 Constant Force Mode

This is the standard method of SFM operation. The cantilever deflection in the  $z$ -direction (and hence the force acting between the tip and the sample) is kept constant using a feedback loop (see Fig. 1). Thus, if the sample is scanned in the  $(x, y)$  plane, the output of the feedback gives a map of equal force which usually corresponds to the topography of the sample. Some SFM micrographs acquired in this mode are presented in Fig. 4.



**Figure 4.** Examples of SFM micrographs acquired in the constant force mode, demonstrating possible applications of SFM in science and technology. (a) A  $70 \times 70 \mu\text{m}$  area of an integrated circuit. Individual components are visible. Light bright regions represent high surface areas and dark regions low surface areas. (b) Tabular silver bromide microcrystals (so-called 'T-grains') as used in modern photographic material, deposited on a glass substrate. The scanned area was  $30 \times 30 \mu\text{m}$ ; individual microcrystals are  $100\text{--}400 \text{ nm}$  high. The image is displayed as a stereoscopic projection (bird's-eye view). Despite the high sensitivity of the material to visible light or electronic beams, the microcrystals can be imaged without surface damage [8]. (c) A  $5 \times 5 \mu\text{m}$  area of a (010) cleavage face of triglycine sulfate. Light regions represent high surface areas and dark regions low surface areas. The step height between individual terraces is  $6 \text{ \AA}$ , which corresponds to half of the unit cell of the material. (d) High-resolution SFM micrograph of a mica surface, demonstrating molecular resolution. The regular protrusions have the periodicity of the lattice ( $5.2 \text{ \AA}$  distance).

### 2.2.3.2 Variable Deflection Mode

If the sample is scanned with respect to the cantilever without any feedback, an image of the surface is obtained by the direct recording of the output of the deflection sensor. Higher scan rates (to nearly video

frequencies) can be achieved in this mode. This is an advantage when imaging on the atomic scale. Large scans, however, might lead to a deformation or even a destruction of the tip and/or sample due to changing force between the tip and the sample (see Sec. 2.2.6 of this Chapter).



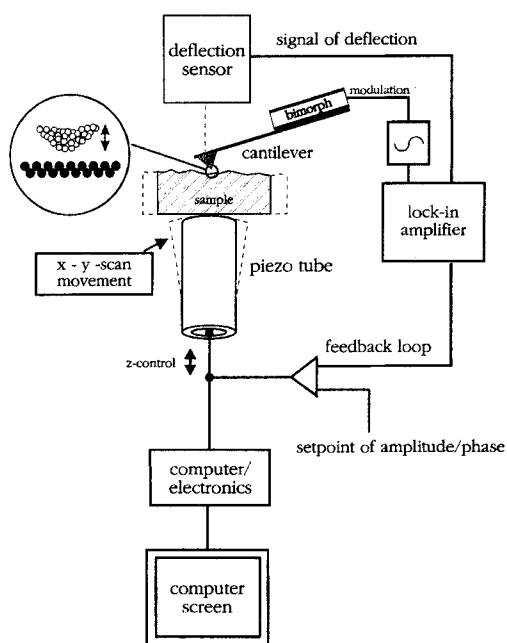
### 2.2.3.3 Noncontact Dynamic Modes

The resonance frequency  $f_{\text{res}}$  of the cantilever is proportional to  $\sqrt{c - \partial F/\partial z}$ , where  $c$  is the spring constant of the cantilever and  $\partial F/\partial z$  is the gradient of the tip-sample interaction force normal to the sample surface. A change in the force gradient results in a change of the resonance frequency which is measured [33, 34]. An example of an experimental set-up is sketched in Fig. 5. The cantilever is vibrated close to its resonance frequency by a piezo element. Then, the signal obtained by the deflection sensor is analyzed by the lock-in technique. A subsequent feedback circuit regulates on a constant phase shift between the signal of the deflection sensor and the original driving signal or on a constant amplitude

of the modulated deflection sensor signal (cf. Fig. 5 in Sec. 3 of this Chapter). Both methods keep the resonance frequency  $f_{\text{res}}$  constant.

The noncontact dynamic mode shows important differences in comparison to the imaging modes described above. First, since the resonance frequency  $f_{\text{res}}$  and not the deflection of the cantilever is kept constant, lines of constant force gradient (and not of constant force) are recorded. Second, the tip does not touch the sample during the measurement. Therefore, surface deformations and lateral forces are minimized [10, 43]. Finally, since a surface can be traced not only for a distance of a few nanometers, but also for tens of nanometers or even more, long-range forces such as electrostatic [43, 44] or magnetic (see Sec. 3.2 of this Chapter) forces can be imaged separated from the surface topography.

In order to obtain a maximum signal-to-noise ratio, the cantilever should possess a high resonance frequency  $f_{\text{res}}$  and a low force constant  $c$  [33]. However, the thermally induced motion of the cantilever, which is proportional to  $c^{-1/2}$ , reduces the minimal detectable value of a force gradient [33]. Furthermore, the lateral resolution increases with decreasing distance between the tip and the sample. In order to obtain the highest resolution, the tip must profile the sample surface as close as possible. If the spring constant is smaller than the local force gradient normal to the sample surface, the cantilever snaps to the surface (see Sec. 2.2.4 of this Chapter). This happens with soft springs which are suitable for contact force measurement (force constants between 0.01 and 1 N m<sup>-1</sup>). Therefore, hard cantilevers with spring constants between 5 N m<sup>-1</sup>



**Figure 5.** Schematic of a typical set-up of a force microscope operated in the noncontact dynamic imaging mode.

and more than  $100 \text{ N m}^{-1}$  (and corresponding resonance frequencies of some 100 kHz) are usually chosen to scan close to the surface. Then, oscillation amplitudes have to be restricted to a few angstroms.

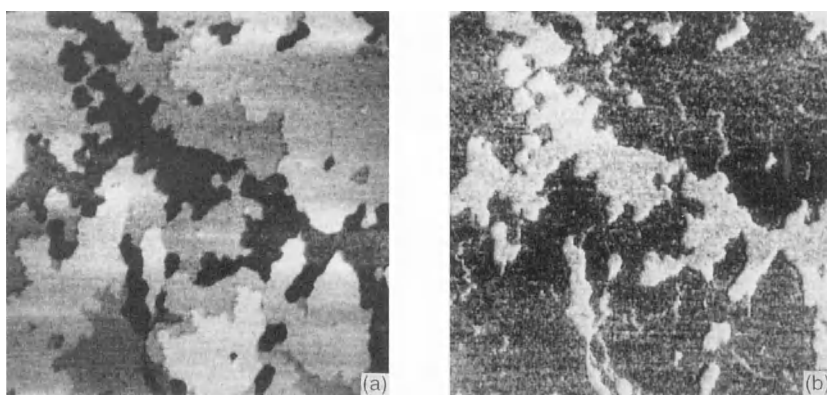
If the damping of the cantilever oscillation is very low (i.e., the quality  $Q$  of the cantilever is very high), it is difficult to regulate on a constant amplitude or constant phase shift. Therefore, Albrecht et al. [45] suggested a method for the direct determination of the cantilever resonance frequency which is especially useful for vacuum measurements. A detailed theoretical description of noncontact force microscopy is given by Hartmann [46].

#### 2.2.3.4 Imaging Friction, Elasticity, and Viscosity

In the so-called lateral force microscopy (LFM), which is often also termed friction force microscopy (FFM), the torsion of the cantilever is measured additionally to

the simultaneously recorded topography [47]. Experimentally, this is usually realized by using the beam deflection technique described in Sec. 2.2.2 of this Chapter. A map of the lateral forces acting on the tip is generated. In the absence of topographical effects (topographical steps or slopes cause a torque of the tip [47, 48]), the torsion is proportional to the frictional force between the tip and the sample. Variations of the frictional force on the atomic level have been observed [49]. Figure 6a shows the topography and Fig. 6b the simultaneously recorded friction force map of a thin film of  $\text{C}_{60}$  molecules on a GeS substrate. The  $\text{C}_{60}$  islands (dark areas in Fig. 6b) exhibit a lower friction than the GeS substrate.

The viscoelastic properties of samples can be investigated by the modulation of the vertical sample position [50]. The modulation of the sample position (or, alternatively, the position of the cantilever support) leads to a modulation of the force acting between the tip and the sample. Using the lock-in technique, the in-phase



**Figure 6.** (a) Topography and (b) simultaneously recorded friction force map of the thin film of  $\text{C}_{60}$  molecules epitaxially grown on a GeS substrate. The scanned area is  $2 \times 2 \mu\text{m}$ . The  $\text{C}_{60}$  islands are one or two monolayers high; each monolayer corresponds to about 1 nm in height. The  $\text{C}_{60}$  islands (dark regions in part (b)) exhibit a lower friction than the GeS substrate (light regions).

amplitude and the phase shift of the response of the sample can be recorded, which correspond to the elasticity and the viscosity of the surface.

### 2.2.3.5 Other Imaging Modes

Several further imaging modes or combinations of imaging modes are possible such as the so-called 'lift mode' (see Sec. 3.1 of this Chapter), which combines topography measurement and measurement of a long-range force. Another example is dynamic high-amplitude measurement ('tapping mode'), where the cantilever is vibrated as described for the noncontact modes, but purposely touches the surface at each cycle [51].

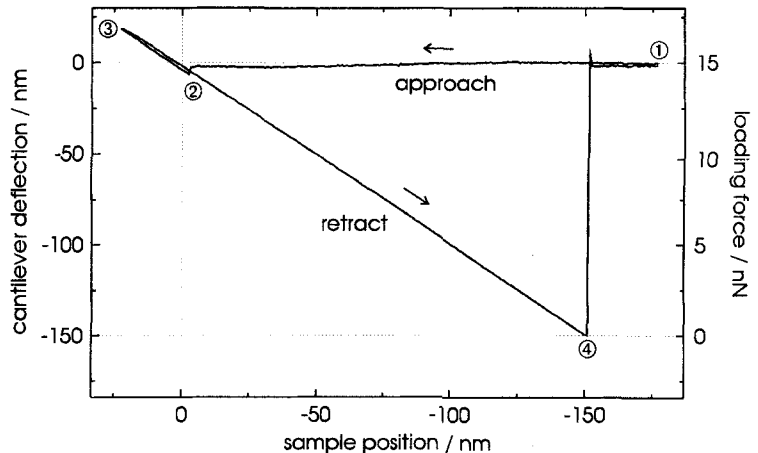
### 2.2.4 Force–Distance Curves

Since the force microscope probes the force between a tip and a sample, SFM can be used to study the tip–sample interaction as a function of their separation.

The force–distance (fd) curve provides useful information about both long- and short-range forces as well as surface hardness, etc.

Figure 7 shows a typical fd curve. The horizontal axis is calibrated so that the sample position is zero if the nondeflected cantilever touches the surface. The left vertical axis shows the deflection of the cantilever. Negative values indicate a bending of the cantilever toward the sample. The loading force  $F$  (right vertical axis) can be calculated from Hooke's law  $F = -cz$ , where  $c$  is the spring constant ( $0.1 \text{ N m}^{-1}$  for the cantilever used in Fig. 7) and  $z$  is the deflection of the cantilever.

At position 1, the tip is distant from the surface, and no interaction occurs (cantilever not deflected). When the cantilever approaches the surface, a slight deflection of the cantilever is observed starting at a distance of about 75 nm from the surface due to the long-range van der Waals forces. At position 2, close to the sample surface, the force gradient of the attractive interaction force is larger than the spring constant of the cantilever. Thus, the attractive forces acting on the tip cannot



**Figure 7.** Cantilever deflection and loading force of the tip on the sample as a function of the tip–sample separation. At position 2, the free cantilever jumps to the sample surface and sticks there until position 4 is reached.

be compensated for by the spring force of the cantilever, and the cantilever snaps to the surface (cantilever bends toward the surface by about 6 nm in the present example). The sample is retracted starting from position 3. The tip sticks to the surface until position 4 is reached due to attractive forces such as adhesion and capillary forces. At this position, just before the tip loses the contact with the sample surface, the cantilever is strongly bent toward the sample by 150 nm (see Fig. 7), corresponding to a repulsive spring force of 15 nN. If the cantilever is bent less toward the sample, the outermost tip atom is still in a repulsive force regime, causing probable local surface deformation or even surface damage, in spite of a net attractive force on the tip [52, 53]. Only at position 4 are the attractive surface forces balanced by the spring force of the cantilever, and the repulsive force of the sample surface vanishes. Therefore, this point, which represents the force experienced by the outermost atom of the tip, is set to zero force during the standard SFM force calibration.

In summary, *fd* curves are essential in order to calibrate the loading force of the tip on the sample. If *fd* curves are performed at many different surface spots, a map of local surface adhesion can be generated [54]. The nature of long-range forces can be investigated [55]. If the end of the tip is sensitized by specific molecules, the binding strength between molecules can be measured [56]. Spatially resolved measurement of adhesion with such sensitized tips leads to functional group imaging ('chemical force microscopy', see Frisbie et al. [57]). Principally, *fd* curves can also be used to determine surface energies and therefore to obtain

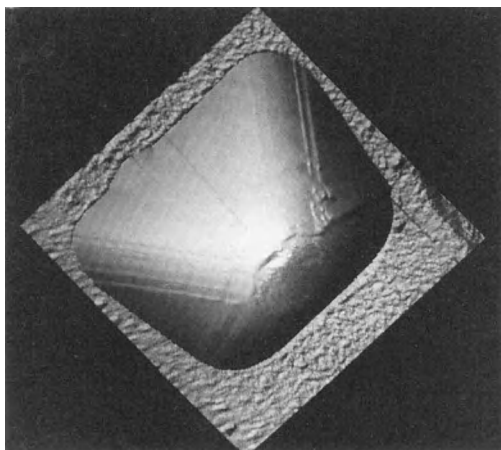
information about the local chemical composition. For this purpose, tips with well-defined shapes and advanced cantilever designs (e.g., double leaf springs [58]) are necessary. However, such tips and cantilevers are not yet commercially available.

### 2.2.5 Tip Artefacts

The information which can be obtained by SFM strongly depends on the kind of tip. In order to obtain a high lateral resolution, the interaction between the tip and the sample has to be limited to a very small surface region. Since real tips are not ideal point probes, tip effects have to be taken into account when interpreting the acquired data. These tip effects result from the fact that different sites of the probing tip interact with the sample during the scan, leading to a convolution of sample features with the tip shape.

An extreme example is presented in Fig. 8. It shows a latex ball with a diameter of 1  $\mu\text{m}$  as used for calibration purpose in electron microscopy, scanned with a pyramidal tip. Obviously, the latex ball images the pyramidal tip more than the tip images the latex ball due to the large curvature of the ball. The original spherical shape of the ball is completely hidden.

Three types of effects mainly limit the information from SFM measurements on a scale larger than the atomic scale: (1) the surface roughness might not be reflected correctly due to the finite size of the tip end; (2) the nonzero opening angle of the tip can cause artefacts at high surface steps; and (3) double tip effects might



**Figure 8.** Spherical latex ball of  $1\ \mu\text{m}$  diameter on a silicon substrate, mapped with a pyramidal tip. The scan area was  $2 \times 2\ \mu\text{m}$ . The obtained image represents a convolution of the pyramidal tip shape with the spherical ball shape.

occur on rough samples. These issues are discussed in detail by Schwarz et al. [31].

### 2.2.6 Scanning Force Microscopy as a Tool for Nanomodifications

If the force between tip and sample exceeds a certain limit, the sample and/or the tip is modified. This often happens accidentally when loading forces that are too large are used or on very soft materials. However, controlled increasing of the tip-sample interaction offers the possibility of performing hardness and wear tests as well as to machine surfaces or to create structures on the nanometer scale [59, 60]. The sample surface can be imaged before and after the modification with the same instrument. Additionally, suitable samples can also be structured by applying voltages and by combination of SFM with etching techniques [60, 61].

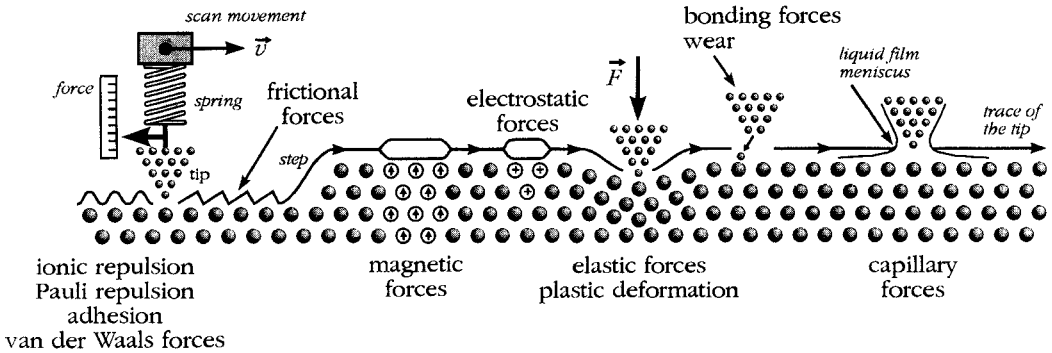
## 2.3 Theoretical Aspects

### 2.3.1 Forces in Force Microscopy

In force microscopy, a knowledge of the interactions between the tip and the sample is necessary to measure and interpret the data correctly. On the scale of atoms and molecules, the electromagnetic interaction dominates over other types of interactions. The strong and the weak interaction as well as gravitation are many magnitudes smaller. However, the electromagnetic interaction gives rise to a rich variety of different forces which not only complicate the SFM image interpretation but also have the potential of measuring many different physical properties. The potential of the force microscope lies far beyond that of a simple topography profiler. A summary of the forces relevant to force microscopy is given in Fig. 9; additional information in intermolecular and surface forces is given by Israelachvili [62].

#### 2.3.1.1 Pauli Repulsion and Ionic Repulsion

The most important forces in conventional imaging force microscopy are Pauli repulsion and ionic repulsion. The Pauli exclusion principle forbids the charge clouds of two electrons with the same quantum numbers having any significant overlap; first, the energy of one of the electrons has to be increased. This yields a repulsive force. Additionally, overlap of the charge clouds of electrons can cause an insufficient screening of the nuclear charge, leading to ionic repulsion of coulombic nature. Pauli repulsion and ionic repulsion



**Figure 9.** Summary of the forces relevant to SFM.

are very hard, that is, they vary over many orders of magnitude of angstroms [63], thereby preventing the SFM tip from penetrating into the sample. Therefore, only the foremost atoms of the tip interact with the sample surface. This very local interaction enables the imaging of features with atomic or molecular periodicities.

### 2.3.1.2 Van der Waals Forces

Van der Waals forces are forces between electric dipoles. On the atomic and molecular level, three different types of van der Waals forces are distinguished: (1) the force between two permanent dipoles; (2) the force between a permanent dipole and an induced dipole; and (3) the force between two induced dipoles. The last force arises from fluctuations in the charge distribution of the atoms and molecules, and is known as the dispersion force.

From classical electrodynamics, it is known that dipole-dipole forces are proportional to  $r^{-7}$  and, for large distances of more than 10 nm, proportional to  $\approx r^{-8}$  due to retardation effects. For assemblies of many dipoles, that is, solid bodies consisting of atoms and molecules which each

represent a small dipole, these simple laws change. For example, for a sphere over a flat surface (a geometry which is frequently used to approximate the tip and the sample), the van der Waals force in a vacuum is proportional to  $r^{-2}$  if additivity of the forces is assumed [62]. Therefore, van der Waals forces are quite long-range forces which are significant for distances up to some 10 nm (see Fig. 7).

### 2.3.1.3 Adhesion

All attractive forces occurring in SFM are often termed adhesive forces, including van der Waals forces, capillary forces, bonding forces, etc. Practical models of adhesion, however, do not consider the origin of the forces between surfaces, but describe the nature of the phenomenon in terms of surface energies. In this concept, adhesive forces arise from a reduction in surface energy. If two surfaces are brought together, a certain force has to be applied to separate them. Adhesive forces are mostly of van der Waals origin, except in the case of metals, which can show large adhesive forces due to short-range non-additive electron exchange interactions

giving rise to the formation of metallic bonds [62, 64]. The forces which have to be applied to separate tip and the sample can be derived with the theories of Johnson, Kendall, and Roberts [65] and Derjaguin, Muller, and Toporov [66] for the geometry of a sphere in contact with a flat surface. In air or liquids, surface contaminants and adsorbed molecules reduce the surface energy and, consequently, the occurring adhesive forces [18].

### 2.3.1.4 Capillary Forces

Under ambient conditions, the dominant attractive forces in force microscopy are capillary forces [19]. Water and other molecules condensing at the sample surface cause the occurrence of an adsorption layer [20, 21]. The SFM tip penetrates through this layer approaching the sample surface. At the tip-sample microcontact, a meniscus is formed which causes an additional attractive force to act on the SFM tip [20]. This force, which depends on the meniscus and hence on the shape of the tip, can be more than  $10^{-7}$  N, reducing the minimal possible loading forces for force microscopy in air to  $10^{-9}$ – $10^{-8}$  N. For soft samples such as biological membranes such forces will have often deformed the sample surface already. However, the loading forces can be reduced to below  $10^{-10}$  N if measured directly in liquids [18]. Alternatively, capillary forces can be avoided by performing the experiments in a glove box with dry gases [19, 20] or in vacuum. In vacuum, however, strong adhesive forces can occur due to the clean surfaces. Attractive forces due to capillarity can be calculated according to the theories of Fogden and White [67] and

Maugis and Gauthier-Manuel [68] for the geometry of a sphere over a flat surface.

### 2.3.1.5 Interatomic and Intermolecular Bonds

Covalent, ionic, or hydrogen bonds can be formed between the tip and the sample, giving rise to an additional attractive force. In extreme cases, for example under ultra-high vacuum conditions on samples showing dangling bonds, this can lead to the destruction of both the tip and the sample [69].

### 2.3.1.6 Frictional and Elastic Forces

If the tip and the sample are in contact and moved with respect to each other, a frictional force occurs. This force is represented by the component of the force acting on the tip which is parallel to the sample surface, hence causing a torsion of the cantilever which can be measured [32]. Frictional forces can be very large, and in special cases even larger than the normal component of the force. They depend on the surface potential and vary on the atomic scale [49]. Since different materials exhibit different surface potentials, variations of the frictional force can be associated with a different chemical or structural composition of the surface (see Fig. 6) [70].

Not only frictional forces but also elastic forces can provide information about the surface structure. Variations in the local elasticity of the sample have their origin in structural or chemical changes of the investigated sample surface [50]. Experimental set-ups for the measurement

of frictional and elastic forces are described in Sec. 2.2.3.4 of this Chapter.

### 2.3.1.7 Magnetic and Electrostatic Forces

The magnetic and electrostatic forces are, in comparison with most of the other forces described above, of long-range character and therefore they are most easily measured in noncontact modes. The imaging of magnetic materials with ferromagnetic tips, so-called magnetic force microscopy (MFM), has developed into an important field of force microscopy, and will be described in Sec. 3 of this Chapter. Using similar techniques, trapped charges on insulator surfaces [71] or the domain structure of a ferroelectric material [43, 44] can be visualized.

### 2.3.2 Contrast Mechanism and Computer Simulations

Although SFM image interpretation seems to be straightforward, many questions concerning the contrast mechanism of SFM and the behavior of atoms close to or at the tip-sample contact area are not yet fully understood. It is not the aim of this chapter to discuss these problems in detail; however, some critical points should be mentioned. A good survey on the theoretical concepts in force microscopy is given by Wiesendanger and Güntherodt [72].

Based on *ab initio* calculations of the electronic structure, total energy, and forces, Ciraci et al. analyzed the tip-sample interactions of SFM during contact imaging with atomic resolution [52]. They

found that at relatively large tip-sample separations, the tip probes the total charge density of the sample surface. However, at small tip-sample separations corresponding to a strong repulsive regime (where most SFM studies are performed), the ion-ion repulsion determines the image contrast. Therefore, the observed maxima in atomically resolved SFM images can be directly attributed to the atomic sites, contrary to STM, where such an interpretation is not generally applicable.

Tekman and Ciraci [74] showed that even in SFM measurements performed with blunt tips, features with the periodicity of the atomic lattice can be resolved although several atoms are involved in the process of contrast formation, thus preventing the observation of point-like defects. Energetic considerations suggest that a single atom at the tip end is not stable at loading forces which are practicable under ambient conditions or in an ultra-high vacuum ( $\geq 10^{-9}$  N) [74]. This might explain why scanning force micrographs of atomically resolved surfaces usually show only defect-free atomic lattice structures, in contrast to scanning tunneling micrographs, on which defects and kinks can frequently be observed. Ohnesorge and Binnig [75], measuring in liquids, demonstrated by resolving atomic scale kinks that attractive forces of about  $10^{-11}$  N can be used to obtain true atomic resolution. Nevertheless, since other authors have reported the observation of defects in standard SFM contact imaging [76], this problem has not yet been satisfactorily resolved.

Apart from the theoretical analysis, computer simulations can provide insights into the physics of the tip-sample system on the atomic level. The evolution of such



a system can be simulated with high temporal and spatial resolution via a direct numerical solution of the model equations of motion, employing a realistic interatomic interaction potential. Examples of such studies are given in the literature [77].

## Acknowledgments

It is a pleasure to thank W. Allers, H. Bluhm, M. Seider, and O. Zwörner for supplying SFM images for this chapter, and H. Eggers and P. Köster for electron micrographs. Moreover, I would like to acknowledge the people at Hamburg University who provided assistance and technical support.

## 2.4 References

- [1] G. Binnig, H. Rohrer, C. Gerber, E. Weibel, *Phys. Rev. Lett.* **1982**, 49, 57; G. Binnig, H. Rohrer, *Helv. Phys. Acta* **1982**, 55, 726.
- [2] H. K. Wickramasinghe, *J. Vac. Sci. Technol. A* **1989**, 8, 363.
- [3] G. Binnig, C. F. Quate, C. Gerber, *Phys. Rev. Lett.* **1986**, 56, 930.
- [4] G. Meyer, N. M. Amer, *Appl. Phys. Lett.* **1988**, 53, 1045.
- [5] S. Alexander, L. Hellemans, O. Marti, J. Schneir, V. Elings, P. K. Hansma, M. Longmire, J. Gurley, *J. Appl. Phys.* **1989**, 64, 164.
- [6] D. Rugar, H. J. Mamin, P. Guethner, *Appl. Phys. Lett.* **1989**, 55, 2588.
- [7] E. C. Teague, F. E. Scire, S. M. Baker, S. W. Jensen, *Wear* **1982**, 83, 1.
- [8] U. D. Schwarz, H. Haefke, T. Jung, E. Meyer, H.-J. Güntherodt, R. Steiger, J. Bohonek, *Ultramicroscopy* **1992**, 41, 435; U. D. Schwarz, H. Haefke, H.-J. Güntherodt, J. Bohonek, R. Steiger, *J. Imaging Sci. Technol.* **1993**, 37, 344.
- [9] E. Henderson, P. G. Haydon, D. S. Sakaguchi, *Science* **1992**, 257, 1944; J. H. Hoh, P. K. Hansma, *Trends Cell Biol.* **1992**, 2, 208; J. H. Hoh, G. E. Sosinsky, J.-P. Revel, P. K. Hansma, *Biophys. J.* **1993**, 65, 149.
- [10] D. Anselmetti, R. Lüthi, E. Meyer, T. Richmond, M. Dreier, J. E. Frommer, H.-J. Güntherodt, *Nanotechnology* **1994**, 5, 87.
- [11] G. Friedbacher, P. K. Hansma, E. Ramli, G. D. Stucky, *Science* **1991**, 253, 1261.
- [12] U. D. Schwarz, H. Haefke, *Appl. Phys. A* **1994**, 59, 33.
- [13] R. Wigren, H. Elwing, R. Erlandsson, S. Welin, I. Lundström, *FEBS Lett.* **1991**, 280, 225; W. A. Rees, R. W. Keller, J. P. Vesenka, G. Yang, C. Bustamante, *Science* **1993**, 260, 1646.
- [14] J. Frommer, *Angew. Chem., Int. Ed. Engl.* **1992**, 31, 1298; J. Frommer, R. Lüthi, E. Meyer, D. Anselmetti, M. Dreier, R. Overney, H.-J. Güntherodt, M. Fujihira, *Nature* **1993**, 364, 198.
- [15] S. Karrasch, M. Dolder, F. Schabert, J. Ramsden, A. Engel, *Biophys. J.* **1993**, 65, 2437.
- [16] M. Hipp, H. Bielefeld, J. Colchero, O. Marti, J. Mlynek, *Ultramicroscopy* **1992**, 42–44, 1498.
- [17] B. Drake, C. B. Prater, A. L. Weisenhorn, S. A. C. Gould, T. R. Albrecht, C. F. Quate, D. S. Cannell, H. G. Hansma, P. K. Hansma, *Science* **1989**, 243, 1586.
- [18] A. L. Weisenhorn, P. Maivald, H.-J. Butt, P. K. Hansma, *Phys. Rev. B* **1992**, 45, 11226.
- [19] T. Thundat, X.-Y. Zheng, G. Y. Chen, R. J. Warmack, *Surf. Sci. Lett.* **1993**, 294, L939.
- [20] M. Binggeli, C. M. Mate, *Appl. Phys. Lett.* **1994**, 65, 415.
- [21] J. Hu, X.-D. Xiao, D. F. Ogletree, M. Salmeron, *Science* **1995**, 268, 267.
- [22] G. Meyer, N. M. Amer, *Appl. Phys. Lett.* **1990**, 56, 2100; L. Howald, H. Haefke, R. Lüthi, E. Meyer, G. Gerth, H. Rudin, H.-J. Güntherodt, *Phys. Rev. B* **1994**, 49, 5651.
- [23] G. Neubauer, S. R. Cohen, G. M. McClelland, D. Horne, C. M. Mate, *Rev. Sci. Instrum.* **1990**, 61, 2296.
- [24] F. J. Giessibl, *Jpn. J. Appl. Phys.* **1994**, 33, 3726.
- [25] F. J. Giessibl, G. Binnig, *Ultramicroscopy* **1992**, 42–44, 281; T. R. Albrecht, P. Grütter, D. Rugar, *Ultramicroscopy* **1992**, 42–44, 1638.
- [26] H. Heinzelmann, E. Meyer, H. Rudin, H.-J. Güntherodt in *Scanning Tunneling Microscopy and Related Methods* (Eds.: R. J. Behm, N. Garcia, H. Rohrer), Kluwer, Dordrecht **1990**, p. 443; D. Rugar, P. Hansma, *Phys. Today* **1990**, Oct., 23; D. Sarid, V. Elings, *J. Vac. Sci. Technol. B* **1991**, 9, 431; J. Frommer, E. Meyer, *J. Phys. C: Condens. Matter* **1991**, 3, S1; C. F. Quate, *Surf. Sci.* **1994**, 299/300, 980.
- [27] D. Sarid, *Scanning Force Microscopy*, 1st Ed., Oxford University Press, New York **1991**; R. Wiesendanger, H.-J. Güntherodt (Eds.), *Scanning Tunneling Microscopy II, Springer Series in Surface Sciences*, Vol. 28, 1st Ed., Springer-Verlag, Berlin **1992**; R. Wiesendanger, H.-J. Güntherodt (Eds.), *Scanning Tunneling Microscopy III, Springer Series in Surface Sciences*,

- Vol. 29, 1st Ed., Springer-Verlag, Berlin **1993**;  
R. Wiesendanger, *Scanning Probe Microscopy and Spectroscopy*, Cambridge University Press, Cambridge **1994**.
- [28] D. W. Pohl, *IBM J. Res. Develop.* **1986**, *30*, 417; S. Park, C. F. Quate, *Rev. Sci. Instrum.* **1987**, *58*, 2004; Y. Kuk, P. J. Silverman, *Rev. Sci. Instrum.* **1989**, *60*, 165.
- [29] T. R. Albrecht, C. F. Quate, *J. Vac. Sci. Technol. A* **1988**, *6*, 271.
- [30] T. R. Albrecht, S. Akamine, T. E. Carver, C. F. Quate, *J. Vac. Sci. Technol. A* **1990**, *8*, 3386; O. Wolter, T. Bayer, J. Greschner, *J. Vac. Sci. Technol. B* **1991**, *9*, 1353.
- [31] U. D. Schwarz, H. Haefke, P. Reimann, H.-J. Güntherodt, *J. Microsc.* **1994**, *173*, 183.
- [32] O. Marti, J. Colchero, J. Mlynek, *Nanotechnology* **1990**, *1*, 141; G. Meyer, N. M. Amer, *Appl. Phys. Lett.* **1990**, *57*, 2089.
- [33] Y. Martin, C. C. Williams, H. K. Wickramasinghe, *J. Appl. Phys.* **1987**, *61*, 4723.
- [34] G. M. McClelland, R. Erlandsson, S. Chiang in *Review of Progress in Quantitative Non-Destructive Evaluation*, Vol. 6 (Eds.: D. O. Thompson, D. E. Chimenti), Plenum Press, New York **1987**, p. 1307; A. J. den Boef, *Appl. Phys. Lett.* **1989**, *55*, 439; C. Schönenberger, S. F. Alvarado, *Rev. Sci. Instrum.* **1989**, *60*, 3131.
- [35] P. J. Mulhern, T. Hubbard, C. S. Arnold, B. L. Blackford, M. H. Jericho, *Rev. Sci. Instrum.* **1991**, *62*, 1280; A. Moser, H. J. Hug, T. Jung, U. D. Schwarz, H.-J. Güntherodt, *Measure. Sci. Technol.* **1993**, *4*, 769.
- [36] C. A. J. Putnam, B. G. de Grooth, N. F. van Hulst, J. Greve, *Ultramicroscopy* **1992**, *42–44*, 1509.
- [37] D. Sarid, D. Iams, V. Weissenberger, L. S. Bell, *Opt. Lett.* **1988**, *13*, 1057.
- [38] D. Sarid, D. A. Iams, J. T. Ingle, V. Weissenberger, J. Ploetz, *J. Vac. Sci. Technol. A* **1990**, *8*, 378.
- [39] M. Tortonese, R. C. Barrett, C. F. Quate, *Appl. Phys. Lett.* **1993**, *62*, 834.
- [40] F. J. Giessibl, *Science* **1995**, *267*, 68.
- [41] T. R. Albrecht, C. F. Quate, *J. Appl. Phys.* **1987**, *62*, 2599; H. Heinzelmann, P. Grütter, E. Meyer, H.-R. Hidber, L. Rosenthaler, M. Ringger, H.-J. Güntherodt, *Surf. Sci.* **1987**, *189/190*, 29; O. Probst, S. Grafström, J. Kowalski, R. Neumann, M. Wörtge, *J. Vac. Sci. Technol. B* **1991**, *9*, 626.
- [42] T. Göddenhenrich, H. Lemke, U. Hartmann, C. Heiden, *J. Vac. Sci. Technol. A* **1990**, *8*, 383.
- [43] R. Lüthi, E. Meyer, L. Howald, H. Haefke, D. Anselmetti, M. Dreier, M. Rütschi, T. Bonner, R. M. Overney, J. Frommer, H.-J. Güntherodt, *J. Vac. Sci. Technol. B* **1994**, *12*, 1673.
- [44] F. Saurenbach, B. D. Terris, *Appl. Phys. Lett.* **1990**, *56*, 1703.
- [45] T. R. Albrecht, P. Grütter, D. Horne, R. Rugar, *J. Appl. Phys.* **1991**, *69*, 668.
- [46] U. Hartmann in *Scanning Tunneling Microscopy III* (Eds.: R. Wiesendanger, H.-J. Güntherodt), *Springer Series in Surface Sciences*, Vol. 29, 1st Edn., Springer-Verlag, Berlin **1993**, p. 293.
- [47] R. Overney, E. Meyer, *MRS Bull.* **1993**, *19*, 26.
- [48] S. Grafström, M. Neitzert, T. Hagen, J. Ackermann, R. Neumann, O. Probst, M. Wörtge, *Nanotechnology* **1993**, *4*, 143.
- [49] C. M. Mate, G. M. McClelland, R. Erlandsson, S. Chiang, *Phys. Rev. Lett.* **1987**, *59*, 1942; R. Erlandsson, G. Hadzioannou, C. M. Mate, G. M. McClelland, S. Chiang, *J. Chem. Phys.* **1988**, *89*, 5190; R. M. Overney, H. Takano, M. Fujihira, *Phys. Rev. Lett.* **1994**, *72*, 3546; S. Fujisawa, E. Kishi, Y. Sugawara, S. Morita, *Phys. Rev. B* **1995**, *51*, 7849.
- [50] M. Rademacher, R. W. Tillmann, M. Fritz, H. E. Gaub, *Science* **1992**, *257*, 1900.
- [51] Q. Zhong, D. Inniss, K. Kjoller, V. B. Elings, *Surf. Sci. Lett.* **1993**, *290*, L688.
- [52] S. Ciraci, A. Baratoff, I. P. Batra, *Phys. Rev. B* **1990**, *41*, 2763.
- [53] F. O. Goodman, N. Garcia, *Phys. Rev. B* **1991**, *43*, 4728.
- [54] A. Torii, M. Sasaki, K. Hane, S. Okuma, *Sensors Actuators A* **1994**, *44*, 153.
- [55] H. J. Hug, A. Moser, T. Jung, O. Fritz, A. Wadas, I. Parashnikov, H.-J. Güntherodt, *Rev. Sci. Instrum.* **1993**, *64*, 2920; H. J. Hug, A. Moser, O. Fritz, I. Parashnikov, H.-J. Güntherodt, T. Wolf, *Physica B* **1994**, *194–196*, 377.
- [56] U. Dammer, O. Popescu, P. Wagner, D. Anselmetti, H.-J. Güntherodt, G. N. Misevic, *Science* **1995**, *267*, 1173.
- [57] C. D. Frisbie, L. F. Rozsnyai, A. Noy, M. S. Wrighton, C. M. Lieber, *Science* **1994**, *265*, 2071.
- [58] H. K. Christenson, *J. Coll. Interface Sci.* **1988**, *121*, 170.
- [59] T. A. Jung, A. Moser, H. J. Hug, D. Brodbeck, R. Hofer, H. R. Hidber, U. D. Schwarz, *Ultramicroscopy* **1992**, *42–44*, 1446; Y. Kim, C. M. Lieber, *Science* **1992**, *257*, 375; G. Persch, C. Born, B. Utesch, *Appl. Phys. A* **1994**, *59*, 29; B. Bhushan, J. N. Israelachvili, U. Landmann, *Nature* **1995**, *374*, 607.
- [60] T. A. Jung, A. Moser, M. T. Gale, H. J. Hug, U. D. Schwarz in *Technology of Proximal Probe Lithography*, Vol. IS 10 (Ed.: C. R. K. Marrian), SPIE Institutes for Advanced Optical Technologies, Bellingham **1993**, p. 234.

- [61] A. Majumdar, P. I. Oden, J. P. Carrejo, L. A. Nagahara, J. J. Graham, J. Alexander, *Appl. Phys. Lett.* **1992**, 61, 2293; E. S. Snow, P. M. Campell, *Appl. Phys. Lett.* **1994**, 64, 1932; L. Tsau, D. Wang, K. L. Wang, *Appl. Phys. Lett.* **1994**, 64, 2133; M. Wendel, S. Kühn, H. Lorenz, J. P. Kotthaus, M. Holland, *Appl. Phys. Lett.* **1994**, 65, 1775.
- [62] J. N. Israelachvili, *Intermolecular and Surface Forces*, 2nd Edn., Academic Press, London **1991**.
- [63] R. G. Gordon, Y. S. Kim, *J. Chem. Phys.* **1972**, 56, 3122.
- [64] U. Dürig and A. Stalder, in *Physics of Sliding Friction* (Eds.: B. N. J. Persson and E. Tosatti), NATO ASI Series, Serie E: Applied Sciences, Vol. 311, Kluwer Academic Publishers, Dordrecht, **1996**, p. 266.
- [65] K. L. Johnson, K. Kendall, A. D. Roberts, *Proc. R. Soc. Lond. A* **1971**, 324, 301.
- [66] B. V. Derjaguin, V. M. Muller, Y. P. Toporov, *J. Colloid Interface Sci.* **1975**, 53, 314; B. D. Hughes, L. R. White, *Q. J. Mech. Appl. Math.* **1979**, 33, 445; V. M. Muller, V. S. Yushenko, B. V. Derjaguin, *J. Colloid Interface Sci.* **1980**, 77, 91; **1983**, 92, 92; D. Maugis, *J. Colloid Interface Sci.* **1992**, 150, 243.
- [67] A. Fogden, L. R. White, *J. Colloid Interface Sci.* **1990**, 138, 414.
- [68] D. Maugis, B. Gauthier-Manuel, *J. Adhesion Sci. Technol.* **1994**, 8, 1311.
- [69] L. Howald, R. Lüthi, E. Meyer, P. Güthner, H.-J. Güntherodt, *Z. Phys. B* **1994**, 93, 267.
- [70] R. M. Overney, E. Meyer, J. Frommer, D. Brodbeck, R. Lüthi, L. Howald, H.-J. Güntherodt, M. Fujihira, H. Takano, Y. Gotoh, *Nature* **1992**, 359, 133.
- [71] C. Schönenberger, S. F. Alvarado, *Phys. Rev. Lett.* **1990**, 65, 3162.
- [72] R. Wiesendanger, H.-J. Güntherodt (Eds.), *Scanning Tunneling Microscopy III, Springer Series in Surface Sciences*, Vol. 29, 1st Edn., Springer-Verlag, Berlin **1993**.
- [73] E. Tekman, S. Ciraci, *J. Phys. C: Condens. Matter* **1991**, 3, 2613.
- [74] C. F. Quate, *Surf. Sci.* **1994**, 299/300, 980.
- [75] F. Ohnesorge, G. Binnig, *Science* **1993**, 260, 1451.
- [76] S. S. Sheiko, M. Möller, E. M. C. M. Reuvekamp, H. W. Zandbergen, *Phys. Rev. B* **1993**, 48, 5675.
- [77] U. Landman, W. D. Luedtke, M. W. Ribarsky, *J. Vac. Sci. Technol. A* **1989**, 7, 2829; U. Landman, W. D. Luedtke, A. Nitzan, *Surf. Sci.* **1989**, 210, L177; U. Landman, W. D. Luedtke, N. A. Burnham, R. J. Colton, *Science* **1990**, 248, 454; J. Belak, I. F. Stowers in *Fundamental of Friction: Macroscopic and Microscopic Processes* (Eds.: I. L. Singer, H. M. Pollock), Kluwer, Dordrecht **1992**, p. 511; J. A. Harrison, C. T. White, R. J. Colton, D. W. Brenner, *Surf. Sci.* **1992**, 271, 57; *Phys. Rev. B* **1992**, 46, 9700; *J. Phys. Chem.* **1993**, 97, 6573; U. Landman, W. D. Luedtke, *MRS Bull.* **1993**, 18, 993.

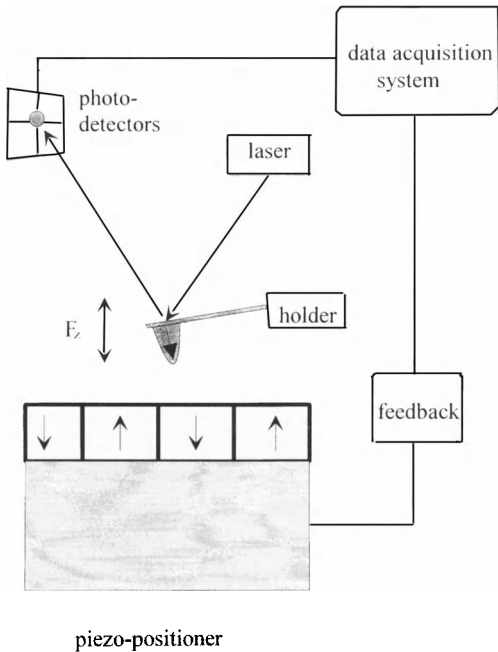
## 3 Magnetic Force Microscopy

### 3.1 Introduction

Magnetic force microscopy (MFM) is the third generation of scanning probe techniques after scanning tunneling microscopy (STM) and scanning force microscopy (SFM) [1]. MFM has been designed to study magnetic materials on the nanometer scale. The first results obtained by this method were reported by Martin and Wickramasinghe in 1987 [2]. The technique relies on a magnetostatic interaction between a magnetic sample and a probing sensor. Figure 1 is an illustration of MFM operation. The magnetic tip, which acts as the sensor, is mounted on a flexible cantilever. It is raster scanned over the sample surface, typically in the range from 20 to hundreds of nanometers. A magnetic sample with a domain structure produces a complicated stray field over the surface. The aim of MFM is to map the stray field as close to the surface as possible. The interaction which occurs when a sample is scanned by an MFM tip is measured via a deflection of the cantilever. There are many modes of MFM operation and various techniques to monitor the bending of the cantilever. There are, however, only two physical quantities directly measured by MFM: the force or force gradient acting on a magnetic tip.

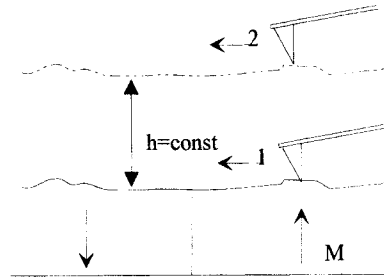
### 3.2 Force Measurement

Figure 1 is a sketch of the set-up of one particular method used to measure magnetic forces. The method of detecting cantilever bending shown in Fig. 1 is called a deflection detection scheme, and was invented by Meyer and Amer in 1988 [3]. A collimated laser beam is focused onto the back face of the cantilever and is reflected toward a set of photodetectors. In the simplest case there are two photodetectors which monitor the force causing an upward or downward cantilever deflection. One photodetector collects more light than the other one due to the deflection of the cantilever. Photocurrents produced by the photodetectors are fed to a differential amplifier. Finally, a signal is obtained that is proportional to the difference in light detected by the photodetectors, and which is a measure of the cantilever deflection. Usually, instead of two photodetectors there is a set of four forming a quadrant as in Fig. 1. This allows additional measurement of lateral forces due to the torque applied to the cantilever. However, we will concentrate on normal forces, which have been the main interest of MFM groups. Since we know now how to detect forces, the next step is to perform MFM. This means that a magnetic tip has to move



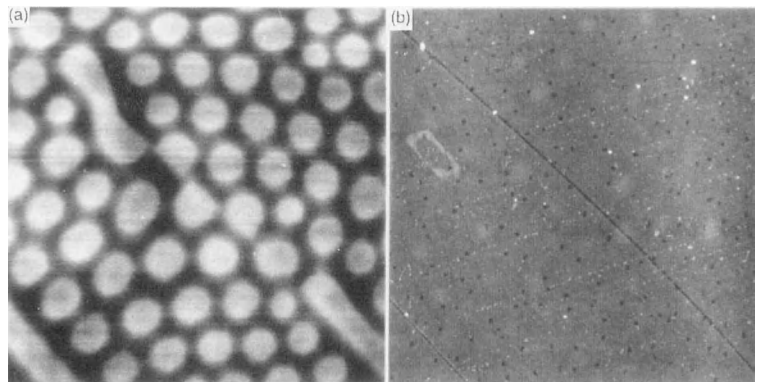
**Figure 1.** A sketch of the experimental set-up used to perform MFM.

over a sample surface at a well-defined height despite any topographic obstacles. This has been usually called the constant height mode. A very reliable way to do this has been introduced by Digital Instruments (Santa Barbara, California, U.S.A.) in one of their commercial microscopes. Figure 2 outlines the principle of their



**Figure 2.** An illustration of the constant height mode as introduced by Digital Instruments.

method, called the 'lift mode'. First, the tip is brought into contact with the surface and performs one scan line. After the topography of this scan line is stored, the tip is scanned along a path parallel to the previously stored scan line at a predefined height from the sample surface. During this second scan the tip-sample separation is nominally constant. Deflections of the cantilever due to magnetic forces are then plotted as a function of position. Such a procedure provides both the topography and the magnetic force images of the same area. This type of operation has been used to study either soft magnetic materials such as garnet films and permalloy or hard magnetic materials such as barium ferrite or hard disk [4, 5]. Figure 3 is an example of an



**Figure 3.** The MFM image of (a) bubble domains in a garnet and (b) the corresponding topographic image.

MFM image with a corresponding topographic scan obtained by the lift mode technique on an epitaxially grown  $(\text{YSmLuCa})_3(\text{FeGe})_5\text{O}_{12}$  garnet film with a bubble structure [4]. Bright circles in the MFM image represent cylindrical domains with a magnetic polarization directed perpendicularly to the surface and opposite to a surrounding area. The bright areas indicate that the MFM tip is being repelled by the sample.

The general formula describing the interaction between a tip and a magnetic sample can be written as

$$E = - \iiint_{\text{tip}} H(r) M \, dV \quad (1)$$

where  $M$  is the tip magnetization and  $H(r)$  is the magnetic field above a sample. A force acting on the tip along the  $z$ -axis (perpendicular to the surface) is related to the interaction energy by

$$F_z = - \frac{\partial E}{\partial z} \quad (2)$$

The force  $F_z$  is related directly to the bending of the cantilever detected during MFM operation by Hooke's law:

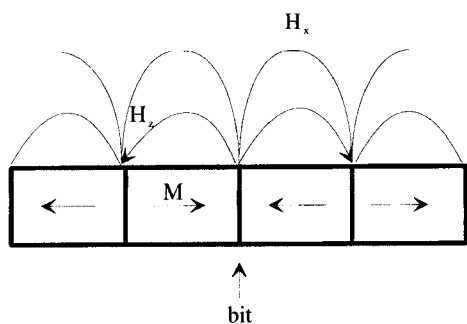
$$F_z = -kz \quad (3)$$

where  $k$  denotes the spring constant of a cantilever.

A detailed calculation of forces which subsequently can be compared with experimental data has been done with a well-defined tip geometry and distribution of the stray field  $H(r)$ . Different shapes for the top were considered: conical, pyramidal, and cylindrical. These theoretical shapes have their experimental counterparts. Examples of the calculations of forces or force gradients giving an analytical formula or a numerical approach for

a well-defined domain structure can be found in some references [6–11]. We have touched here on one of the fundamental problems of MFM, that is, to describe the tip shape, which is in practice determined by the tip preparation. An MFM tip has to fulfill certain conditions. It should be as sensitive as possible, noninvasive, and easily available. Since in magnetism we are dealing with a variety of materials with different properties it is very difficult, if not impossible, to prepare such a universal tip which will give the best MFM image. However, a search was undertaken, completed with some success to find such a tip [5]. It should be mentioned that each particularly soft magnetic sample needs a series of experiments to find the most appropriate MFM tip.

Historically, the first tips were made from magnetic wires of, for example, iron or nickel [2, 11, 12]. The end of the wire was chemically etched to form a sharp needle. Some tens or hundreds of micrometers at the very end of the wire was bent toward a sample. Such a piece of wire with an 'L' shape then acted as a cantilever with the tip at one end. Even if these tips were not magnetized they produced quite a strong signal. Due to their shape anisotropy (a long cylinder with a cone at the end), a large part of the volume probably formed a single magnetic domain responsible for the effective interaction with a magnetic sample. The word 'probably' is used here to highlight an important fact. It was not possible to determine exactly the status of the magnetic domain structure within an MFM tip. The best and most practical way to describe a newly prepared tip is to use it on a well-defined domain structure. A piece of a standard hard disk taken from a computer is a good test



**Figure 4.** The magnetic structure of a hard disk, used frequently in MFM as a test sample.

sample. Its structure is shown in Fig. 4. The stray field emerging from a hard disk has a strong in-plane component  $H_x$  and a strong component  $H_z$  just above a bit (transition area). The MFM image obtained of a hard disk gives us information about the effective  $x$ - and  $z$ -components of the tip magnetization since the energy of interaction described in Eq. (1) directly relates a measured force to the components of tip magnetization.

Magnetic wires used as cantilevers with tips have two important disadvantages. The first one is that the spring constant of the cantilever is not well defined. It always differs from preparation to preparation. In addition, etching the end might form a different domain structure inside the wire. This, however, can be partially overcome by magnetizing the tip, which then leads to the second important disadvantage of wire tips, that is, a strong stray field emerging from such tips [13]. In order to overcome this difficulty, it is necessary to scan over a sample at a sufficiently large distance, which, however, leads to a reduction in lateral resolution. A very logical and natural way to solve these problems is to completely abandon wire

tips. The other choice is to use a magnetic thin film deposited on a nonmagnetic cantilever. Experimental and theoretical results for thin film tips came almost at the same time [14, 15].

Thin film tips were first made by sputtering iron or CoZrNb on tungsten wires, giving very good results on PtCo multilayers [14]. Magneto-optically written domains in these multilayers have been imaged. Next, thin film tips were prepared on microfabricated silicon cantilevers with integrated tips [16]. This procedure assured a well-defined spring constant and a well-defined geometry for the nonmagnetic tip. Up to now there have been different magnetic materials used for the coatings, such as iron, cobalt, permalloy, nickel, and CoPtCr alloys, deposited either by sputtering or evaporation.

The newest state-of-the-art tips are also magnetic thin film tips, but deposited on so-called contamination tips or EBD (electron beam-deposited) tips [17]. It was found that an electron beam of a scanning electron microscope focused on a substrate grows pillars containing mainly carbon. These pillars have the shape of cones with a radius of curvature at the apex between 20 and 40 nm and a cone half-angle of 3–50° [17]. MFM tips prepared by the evaporation of iron on such small cones give a smaller MFM signal than other tips but always have a very good lateral resolution of around 50 nm. They have been used to study garnet films never overwriting a domain structure. The usual thickness of the iron layer has been in the range from 10 to 20 nm. The best MFM results have been obtained with the double-layer thin film tips described earlier [5].

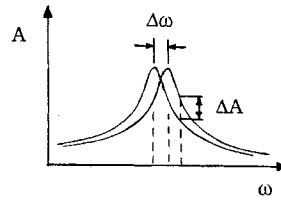
### 3.3 Force Gradient Measurement

As we mentioned above, the other physical quantity measured in MFM is the gradient of a force. This is realized in the so-called alternating current (AC) mode. If we consider the set-up in Fig. 1, some important changes have been made to perform the AC mode. First, there is a bimorph mounted instead of the part labeled 'holder' in Fig. 1. The bimorph is used to oscillate the cantilever near its resonance frequency. As the vibrating cantilever approaches the sample surface, the tip is influenced by an overall force gradient. This means that parallel to the gradient of magnetic forces the tip encounters a gradient of attractive van der Waals forces. Since a force applied to a cantilever is proportional to its bending [Eq. (3)] the force gradient is going to change the effective spring constant,  $k_{\text{eff}}$ , of the cantilever. If we consider the harmonic oscillation of a cantilever, then the resonance frequency  $\omega_0$  is given by the formula

$$\omega_0 = \sqrt{\frac{k_{\text{eff}}}{m}} \quad (4)$$

where  $m$  is the effective mass of the cantilever. The effective spring constant is  $k_{\text{eff}} = k - F'$ , where  $F'$  is the overall force gradient acting on the cantilever and  $k$  is the nominal spring constant of the cantilever.

The change in the force gradient, expected when the tip is approaching the sample surface, shifts the resonance frequency, which subsequently alters the amplitude of the spring vibration (Fig. 5). This can be detected using a laser heterodyne probe [2, 12]. If the tip is then



**Figure 5.** The shift in the resonance frequency,  $\Delta\omega$ , caused by an external force gradient. A change in frequency results in a change of the amplitude,  $\Delta A$ , at a given frequency.

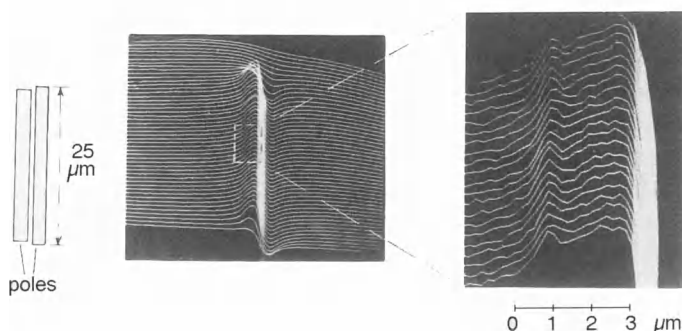
raster scanned over the surface at a constant height, as described earlier, we detect a force gradient in the constant height mode. This method of operation was introduced by Martin and Wickramasinghe in the first MFM report [2]. The sample they investigated was a thin film magnetic recording head (Fig. 6). The current of the head was modulated at a certain frequency, producing a modulated stray field of the same frequency. The vibration of the cantilever due to the modulated head field was detected using a heterodyne interferometer which was capable of detecting a vibration amplitude down to  $5 \times 10^{-5} \text{ \AA}$  in a bandwidth of 1 Hz. The detected amplitude signal from the laser heterodyne probe is then proportional to the magnetic force gradient over the recording head.

The other method of MFM operation is the constant force gradient mode, in which the signal from the laser probe is compared with an electronic reference. A feedback circuit tries to adjust the tip-sample separation by adjusting the  $z$ -signal of a piezo element upon which a sample is mounted (Fig. 1) to maintain a constant force gradient. Figure 7 shows an MFM image obtained in this way by van Kesteren et al. [12]. The image reveals an array of thermomagnetically written cylindrical



FIELD PATTERN OF AN IBM 3380 HEAD BY  
MAGNETIC MICROPROBE

**Figure 6.** Magnetic force gradient contours of a thin film recording head excited by AC as measured by MFM. (Courtesy of Martin and Wickramasinghe [2].)



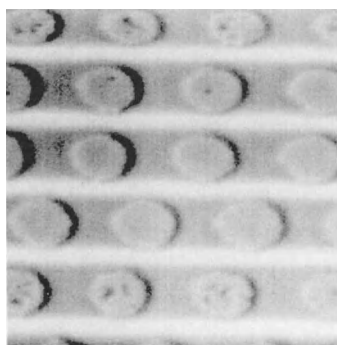
domains in Co/Pt multilayer films. Both of the examples in Figs. 6 and 7 show the direct use of MFM in magnetic recording technology.

The same technique based on the detection of a change in the oscillation amplitude,  $\Delta A$ , has been used to study permalloy films [18]. Again, the constant force gradient was traced in order to get an MFM image. In this case, however, the authors used a fiber-optic interferometer to sense the deflection of the cantilever. One of the advantages of this system over

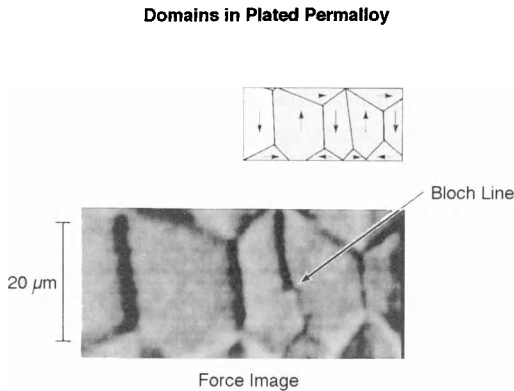
a heterodyne interferometer is its lower thermal drift, which usually affects the optical path length of a laser beam.

The fiber which carries the light is placed some tens of micrometers above the cantilever. A small part of the light is reflected by the end of the fiber. This light serves as a reference signal. The rest of the light passes the end of the fiber and spreads toward the cantilever. Again, only a part of the light reflected from the back of the cantilever enters the fiber. This fraction of the light is partly determined by the fiber-cantilever spacing. The interference of both reflected beams, by the end of the fiber and by the cantilever, determines the interferometer response [19].

Figure 8 shows the experimental data obtained by Mamin et al. on 2.3 μm thick permalloy film [18]. The sample was placed on an electrode with an applied voltage of 2–10 V, providing an offset of the electrostatic force gradient. This ensures the presence of a net attractive interaction, giving stable experimental conditions. Magnetic domains (bright areas) are surrounded by dark or bright lines representing domain walls. The arrow indicates a change in the domain wall contrast. The



**Figure 7.** MFM images of a laser-written array of domains in a Co/Pt multilayer. (Courtesy of van Kesteren et al. [12].)



**Figure 8.** MFM image of domains in a permalloy film. The arrow indicates the location of a Bloch line. (Courtesy of Mamin et al. [18].)

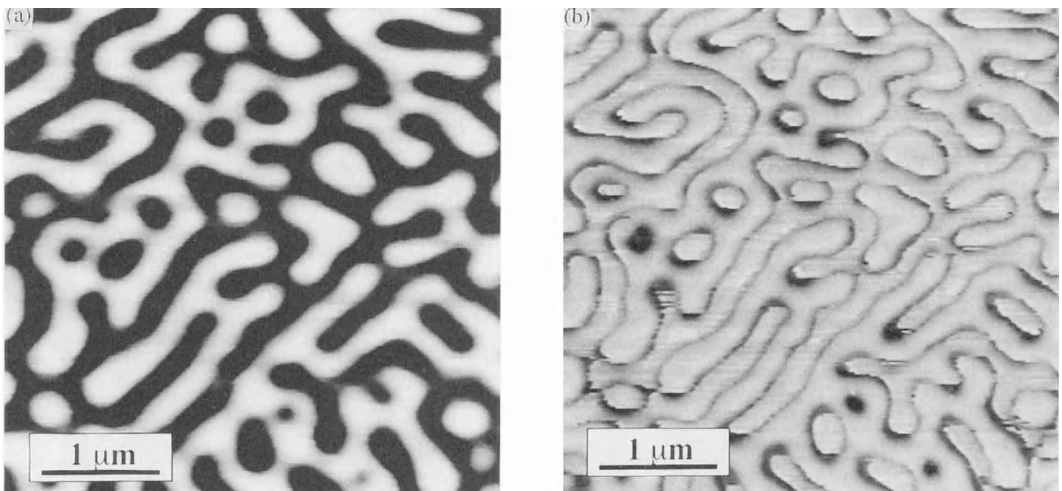
authors correlate this change with the position of a Bloch line.

Experiments on both TbFe and permalloy have been performed using an electrochemically etched iron wire cantilever. The experiment on permalloy showed that MFM is able to recognize the position of a domain wall and, additionally, is able to detect the sense of rotation of the

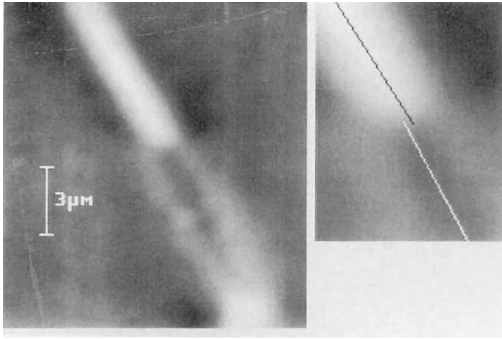
magnetization across the domain wall. This demonstrates that MFM is an important tool to study the internal structure of domain walls.

Another example using the AC technique is shown in Fig. 9 [20]. The authors measured a change in the oscillation amplitude of a vibrating cantilever to study a Co–Pd multilayer. They mapped constant force gradient contours. Figure 9 shows static images of a dynamic process which occurs when the magnetic tip is scanned over the sample at a close tip–sample distance. The image Fig. 9a, taken at a higher distance than in Fig. 9b, shows domain contrast, whereas the image in Fig. 9b, observed closer to the surface, presents a domain wall contrast of the same area. Domain walls are visible as dark lines.

All of the presented examples show various means of studying different aspects of magnetism using MFM based on the optical detection of a cantilever deflection. However, there are also other ways to



**Figure 9.** MFM images of the domain structure of a Co–Pd multilayer measured (a) at 90 nm and (b) at 40 nm tip–sample distances, respectively. (Courtesy of Barnes et al. [20].)



**Figure 10.** MFM contrast image of a  $180^\circ$  Bloch wall in an iron whisker. (Courtesy of Göddenhenrich et al. [21].)

monitor cantilever movement. Göddenhenrich et al. developed a microscope with a capacitively controlled cantilever displacement [21]. A small plate of thin aluminum foil attached to the back of a lever served as a capacitor plate. The opposite plate was mounted on a tubular control piezo element. The deflection of the cantilever due to the forces acting on the magnetic tip caused a change in the capacitance. The sensitivity of such a system is  $10^{-6}$  pF, which corresponds to a change of  $0.01 \text{ \AA}$  in the distance between the two plates. The force detection limit is of the order of  $1 \times 10^{-10} \text{ N}$  [21]. Figure 10 shows an MFM image obtained over a domain wall in a single-crystal iron whisker using the capacitive detection scheme. The change in the domain wall contrast was observed (dark to bright) due to the different polarizations of the magnetization within the domain wall.

The other nonoptical way to detect cantilever deflection is well known under the name of tunneling stabilized magnetic force microscopy (TSMFM). TSMFM is performed by using a scanning tunneling microscope with a flexible, magnetic

tunneling tip instead of the usual rigid tip. The tip position is stabilized near the surface of the magnetic sample using the STM feedback system as tunneling occurs between the tip and sample surface. If the stray field from the sample attracts the tip, then the feedback system maintaining a constant tunneling gap pulls the tip off the surface. Numerous examples using this method can be found in the references [22–24].

All of the images presented in this chapter provide evidence of the potential of MFM either to study fundamental problems in magnetism or for direct use in technology.

## Acknowledgments

I would like to thank J. R. Barnes, U. Hartmann, H. W. van Kesteren, H. J. Mamin, and Y. Martin, who provided illustrations from their work for this article.

## 3.4 References

- [1] G. Binnig, H. Rohrer, *Helv. Phys. Acta* **1982**, 55, 726.
- [2] Y. Martin, H. K. Wickramasinghe, *Appl. Phys. Lett.* **1987**, 50, 1455.
- [3] G. Meyer, N. M. Amer, *Appl. Phys. Lett.* **1988**, 53, 2400; N. M. Amer, G. Meyer, *Bull. Am. Phys. Soc.* **1988**, 33, 319.
- [4] A. Wadas, J. Moreland, P. Rice, R. R. Katti, *Appl. Phys. Lett.* **1994**, 64, 1156.
- [5] A. Wadas, P. Rice, J. Moreland, *Appl. Phys. A* **1994**, 59, 63.
- [6] A. Wadas, P. Grütter, *Phys. Rev. B* **1989**, 39, 1013.
- [7] U. Hartmann, *Phys. Lett.* **1989**, 137, 475.
- [8] U. Hartmann, *Phys. Rev. B* **1989**, 40, 7421.
- [9] U. Hartmann, *J. Vac. Sci. Technol. A* **1990**, 8, 411.
- [10] A. Wadas, P. Grütter, H.-J. Güntherodt, *J. Appl. Phys.* **1990**, 67, 3462.

- [11] D. Rugar, H. J. Mamin, P. Güthner, S. E. Lambert, J. E. Stern, I. McFadyen, T. Yogi, *J. Appl. Phys.* **1990**, 68, 1169.
- [12] H. W. van Kesteren, A. J. den Boef, W. B. Zeper, J. H. H. Spruit, B. A. Jacobs, P. F. García, *J. Appl. Phys.* **1991**, 70, 2413.
- [13] A. Wadas, H. J. Hug, *J. Appl. Phys.* **1992**, 72, 203.
- [14] A. J. den Boef, *Appl. Phys. Lett.* **1990**, 56, 2045.
- [15] A. Wadas, H.-J. Güntherodt, *Phys. Lett. A* **1990**, 146, 277.
- [16] P. Grütter, D. Rugar, H. J. Mamin, G. Castillo, S. E. Lambert, C.-J. Lin, R. M. Valletta, *Appl. Phys. Lett.* **1990**, 57, 1820.
- [17] K. I. Schiffmann, *Nanotechnology* **1993**, 4, 163.
- [18] H. J. Mamin, D. Rugar, J. E. Stern, R. E. Fontana, Jr., P. Kasiraj, *Appl. Phys. Lett.* **1989**, 55, 318.
- [19] D. Rugar, H. J. Mamin, R. Erlandsson, J. E. Stern, B. D. Terris, *Rev. Sci. Instrum.* **1988**, 59, 2337.
- [20] J. R. Barnes, S. J. O'Shea, M. E. Welland, *J. Appl. Phys.* **1994**, 76, 418.
- [21] T. Göddenhenrich, H. Lemke, U. Hartmann, C. Heiden, *Appl. Phys. Lett.* **1990**, 56, 2578.
- [22] J. Moreland, P. Rice, *J. Appl. Phys.* **1991**, 70, 520.
- [23] A. Wadas, H. J. Hug, H.-J. Güntherodt, *Appl. Phys. Lett.* **1992**, 61, 357.
- [24] I. D. Mayergoyz, A. A. Adly, R. D. Gomez, E. R. Burke, *J. Appl. Phys.* **1993**, 73, 5799.



## 4 Ballistic Electron Emission Microscopy

### 4.1 Introduction

Over the last 30 years or so, the field of surface science has addressed the structural, electronic, and vibrational properties of solid surfaces, adsorbates, and thin films [1]. More recently, the revolutionary invention of the scanning tunneling microscope [2] has impacted surface science in an unprecedented way by providing the scientific community with a tool for probing the physical and chemical properties of surfaces on an atomic scale. The capabilities of atomic-scale imaging with scanning tunneling microscopy (STM) have been coupled with local electron spectroscopy by exploiting the bias dependence of electron tunneling in the vicinity of the scanning tunneling microscope tip [3, 4]. In this regard, STM and scanning tunneling spectroscopy (STS) are extremely valuable techniques for probing both surface geometric and electronic structure as compared to more classical, spatially averaged techniques. Extensions of the principles of STM have led to the development of a host of scanning proximal probe instruments for studying surfaces and interfaces [5, 6], most notably scanning force microscopies (SFM). These scanning probe microscopies (SPMs) have significantly

impacted fields such as biology and electrochemistry which are far removed from the established ultrahigh vacuum (UHV) “classical” surface science environment.

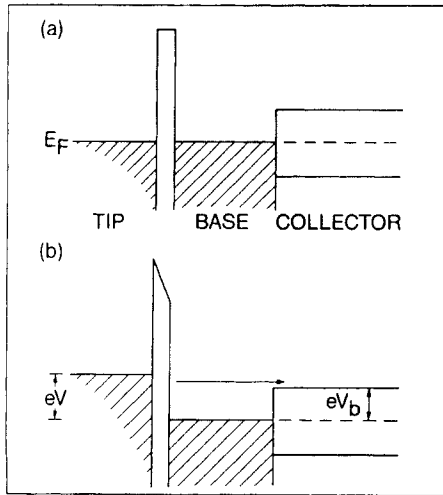
A major issue in surface science continues to be the growth of thin films and the behavior of the interfaces between films and substrates. This field of study is clearly related to electronic device technology. Metal/semiconductor interfaces can exhibit diode characteristics due to the formation of interfacial Schottky barriers [7]. Although the classical model for Schottky barrier formation predicts that the resultant barrier height can be formulated by knowing the semiconductor electron affinity and the metal work function, it has been known for several decades that other factors such as interface states, the doping density of the semiconductor, and interfacial reactions can result in large deviations from the classical model. For instance, it has been demonstrated that the Schottky barrier height can vary for the same two materials grown (epitaxially) with two different orientations at the interface [8]. This provides strong evidence that the roles of local atomic structure and interfacial band structure are very important in determining the Schottky barrier heights at metal/semiconductor interfaces.

In addition, many interfacial systems might not be atomically abrupt over macroscopic dimensions. Therefore, structural inhomogeneities coupled with relatively long screening lengths in the semiconductor play a major role in determining the Schottky barrier height of the resulting device.

A fundamental understanding of hot-electron processes at metal/semiconductor interfaces is important for the optimization of the charge transport properties of many electronic devices such as metal-base transistors and infrared Schottky photodiodes. Experimentally, probing buried interfaces, especially over lateral dimensions where inhomogeneities are likely, is difficult. Standard charged-particle probes do not provide much subsurface information. Nonlinear optical spectroscopy shows some promise in probing buried interfaces, addressing their electronic structure. However, transport measurements across device-size interfaces do not account for local inhomogeneities. STS has contributed to the body of knowledge in thin film technology but only probes the surface electronic structure. STM has provided a great deal of insight into epitaxial and nonepitaxial thin film growth processes by resolving the structure of thin film surfaces. Both of these techniques are limited for interfacial investigations because they are not capable of probing subsurface electronic and geometric structure. The combination of the proximal probe capabilities of STM and carrier transport through an interface describes the unique capability of ballistic electron emission microscopy (BEEM). With this technique, electrons are collected after they traverse a tunnel junction, a metal-base layer, and a semi-

conductor substrate. Besides information on the local Schottky barrier height, BEEM probes many aspects of hot-electron transport across a heterojunction with the nanometer resolution that is amenable to the study of variations in thin film growth and research on new device technologies that include nanometer structures.

BEEM was developed at the Jet Propulsion Laboratory in 1988 by Kaiser and Bell [9] as an STM-based technique capable of probing the electronic properties of buried metal/semiconductor interfaces with high spatial resolution via the analysis of the transport of hot electrons. BEEM is a three-probe STM-based technique where the tip is placed above a metal/semiconductor heterojunction, which is typically unbiased, and acts as a highly localized variable-energy electron injector. An experimental configuration for BEEM is depicted in Fig. 1. The metal film (thicknesses typically  $< 50$  nm), or base, is set at zero potential, as is the substrate, or collector. The tip, or emitter, bias defines the energy distribution of the injected electrons. A comparison of the tip and collector currents (of the order of 1 nA and 10 pA, respectively) as a function of tip bias provides the means to measure the fraction of the electrons that have traversed the metal film and surmounted the Schottky barrier. This describes BEEM, which is also referred to as ballistic electron emission spectroscopy, or BEES. By using a scanning tunneling microscope tip as the injector, the charge injection site is highly localized and can be scanned. Therefore, it is possible to spatially map the potential barrier as a current image. The entire process can, in principle, be modeled theoretically, provided that the electron



**Figure 1.** Illustration of the three-terminal BEEM experimental configuration for the local microscopy and spectroscopy of buried metal/semiconductor contacts. The metal overlayer is grounded, the tip bias ( $V_T$ ) defines the electron injection energy ( $eV_T$ ), and the unbiased substrate is the collector (a). A BEEM spectrum represents the collector current ( $I_c$ ) as a function of electron injection energy at constant injection current ( $I_t$ ) (b). A BEEM image is a collector current map for a specific electron injection energy. STM structural images and BEEM current images of heterojunctions are typically acquired simultaneously to distinguish morphology and interfacial electronic structure. From Ref. [9].

trajectories and scattering and energy loss mechanisms are understood. In the standard BEEM configuration electrons are injected into the metal base and couple to the conduction band states of the semiconductor interface; this requires the use of an n-type substrate that repels the negative charge carriers away from the interface into the semiconductor. Other configurations of BEEM that will be discussed briefly below can probe ballistic hole transport and electron-hole scattering by performing spectroscopy at positive and negative biases and using p-type substrates.

## 4.2 Experimental Considerations

BEEM instrumentation closely resembles that of STM, but several refinements in sample handling, electronics, and vibration isolation are required. As a three-terminal experiment (see Fig. 1), it is necessary to independently contact the front and back of the (unbiased) sample. It is extremely important that the tunneling current (injected electron flux) remains constant, which requires excellent vibration isolation since the tunneling gap is typically controlled by the STM z-axis feedback control (constant current mode). An ohmic contact to the rear of the substrate (the collector), which is often formed by using indium or highly doped silicon, is necessary to allow the collection of the BEEM electrons. Contact to the metal base can be made by a thin wire provided the base contact is continuous. Interfaces with small lateral dimensions (area  $\leq 1 \text{ mm}^2$ ) require a base contact very close to the tunneling point, which necessitates micropositioning of this contact [10]. With collector currents typically in the picoampere range, high-gain ( $\approx 10^{11}$ ), low-noise, low-input-impedance preamplifiers are used. One of the most important design criteria is the zero-bias resistance of the diode itself since the voltage noise of the preamplifier will result in noise currents across the junction. With conventional state-of-the-art solid state electronics, typical zero-bias resistances of the base-collector junction should be of the order of  $\approx 100 \text{ k}\Omega$  or higher to maintain picoampere resolution. Since the zero-bias resistance is inversely proportional to the diode area and is exponentially dependent on the inverse of the temperature, high



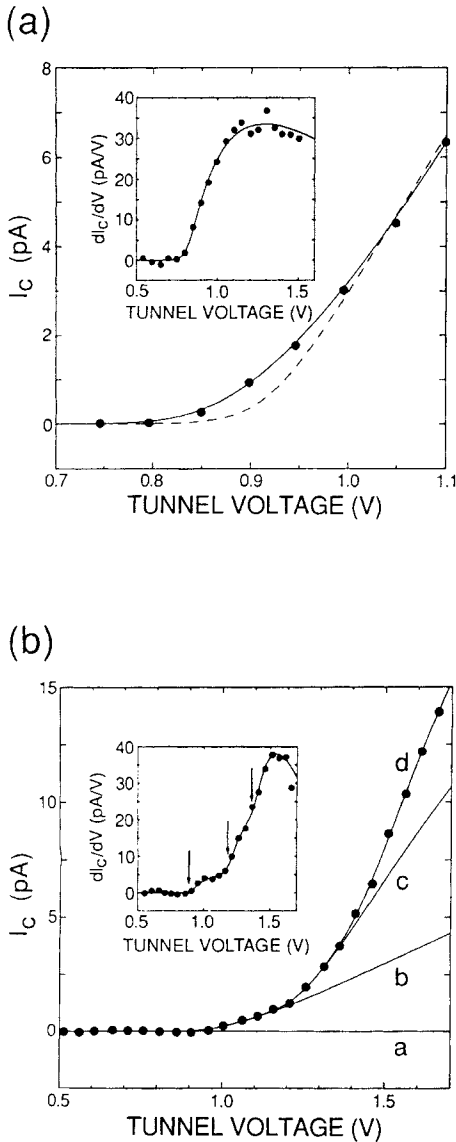
zero-bias resistances can be achieved by either decreasing the diode area or the temperature. For the low barrier heights encountered in many systems ( $\leq 0.7$  eV), low-temperature operation is necessary since the diode areas required to perform BEEM at room temperature would be impractically small. The latest generation of instruments are proceeding in this direction.

Depending on the information that is desired, there are several ways to configure a BEEM experiment. Thin film interfaces are typically grown using surface science techniques in a UHV chamber. However, just as the STM can be operated in air or a gas as well as in a vacuum, many BEEM experiments can be performed in a controlled medium at atmospheric pressure as long as the metal base resists oxidation. This simplifies low-temperature measurement, which can then be made by immersion in a cryogenic liquid or in a flow of cooled gas. In many cases, the metal base cannot be fully passivated; therefore, a totally in situ experiment is required. Recently, a BEEM system which combined UHV operation with low-temperature capabilities has been designed and demonstrated [11].

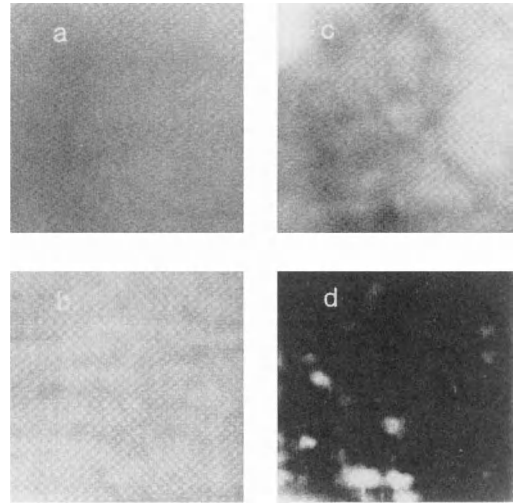
### 4.3 First Demonstrations of Ballistic Electron Emission Microscopy

The first BEEM experiments were performed on Au/Si(100) and Au/GaAs(100) heterojunctions. Those results demonstrated both the subsurface imaging

and local spectroscopic capabilities of BEEM [9]. It was previously recognized from macroscopic junction diode measurements that Au/Si exhibits fairly consistent Schottky barrier heights, while Au/GaAs exhibits a greater range of barrier heights, which was interpreted as resulting from interdiffusion and alloy formation in the GaAs system. Figure 2 shows BEEM spectra for these systems [14] where the measurement is localized in the vicinity of the scanning tunneling microscope tip. Here, the degree of localization depends on the electron trajectory through the metal film and interface. These spectra are plots of  $I_c$  versus  $V_t$  for a constant tunneling current,  $I_t$ , and represent, in the simplest interpretation, diode forward-bias characteristics. Based on modeling the threshold spectra, Schottky barrier heights of approximately 0.85 and 1.2 eV for Au/Si(100) and Au/GaAs(100), respectively, were deduced. Figure 3 shows typical STM images and BEEM images for these interfaces [9]. For the BEEM images, the electron energy was set above the threshold and a current image was recorded. Comparing STM topographic images with BEEM current images permits the distinction of morphological and interfacial effects. The Au/Si(100) STM image exhibits a smooth topograph, and the BEEM current image shows a similar uniformity. This result is consistent with the small variations in values of barrier height for Au/Si diodes in general. The surface morphology of gold on GaAs, on the other hand, is considerably rougher, and the BEEM current exhibits considerable inhomogeneity. Furthermore, features in the STM and BEEM images are not correlated, supporting the hypothesis that defect formation and interdiffusion



**Figure 2.** (a)  $I_C$  versus  $V_t$  and  $dI_C/dV_t$  versus  $V_t$  spectra near the Schottky barrier threshold region for an Au/Si(100) heterojunction. A simple one-dimensional theoretical fit (dashed line) was improved (solid line) with the inclusion of parallel momentum conservation at the interface and thermal broadening. (b)  $I_C$  versus  $V_t$  and  $dI_C/dV_t$  versus  $V_t$  spectra near the Schottky barrier threshold region for an Au-GaAs(100) heterojunction showing the direct minimum and other band structure-dependent thresholds. From Ref. [14].



**Figure 3.** Simultaneously acquired (a) STM and (b) BEEM images ( $25.0 \times 25.0 \text{ nm}^2$ ) of an Au/Si(111) heterojunction. The STM (2.2 nm height range) and BEEM (1.5 pA collector current range at  $V_t = 1 \text{ V}$  and  $I_t = 1 \text{ nA}$ ) images indicate that both the metal film structure and the Schottky barrier are very uniform over the collector. Simultaneously acquired (c) STM and (d) BEEM images ( $25.0 \times 25.0 \text{ nm}^2$ ) of an Au/GaAs(110) heterojunction. The STM (7.2 nm range) and BEEM (1.5 pA collector current range at  $V_t = 1.5 \text{ V}$  and  $I_t = 1 \text{ nA}$ ) images both show considerable nonuniformity. The fact that topographic and current features are not correlated suggests that multiphase formation creates a nonuniform Schottky barrier. [9].

exist at the interface. Note that topographic features at the 1 nm level could also be seen in the BEEM images. A closer look at the BEEM spectra for Au/GaAs(100) indicates the existence of multiple thresholds. These thresholds occurring at 0.89, 1.18, and 1.36 eV become more dramatic in the first derivative spectrum. They arise from interface band structure assuming that the parallel component of the electron momentum,  $k_{||}$ , is conserved as the electrons travel across the metal/semiconductor interface.

## 4.4 Theoretical Considerations

Electron transport in solids can be probed with an unprecedented level of spatial resolution using BEEM. As a direct result of the development of BEEM, there has been a renewed interest in understanding the details of hot-electron transport phenomena in thin-film systems. Analyses of the spectral shape of the BEEM current in the near-threshold region, where much of the transport equation can be simplified, has resulted in very accurate determinations of Schottky barrier heights. However, high-quality fits beyond the threshold require a substantially increased level of complexity. Benchmark experiments have provided insights into various components of the analysis, and at least a qualitative understanding of many novel effects has been achieved. Since the transport problem cannot be solved in its entirety by direct comparison of experimental data to first-principles calculations, many interpretations remain controversial.

The theoretical analysis of a BEEM experiment must account for the physical interactions that govern electron transport from the scanning tunneling microscope tip to the back of the semiconductor substrate. A list of several considerations which must be taken into account to properly characterize the transport of electrons across a Schottky barrier is presented in Table 1. The four major components of BEEM electron transport are: (1) vacuum tunneling of the electrons into the metal base; (2) transport through the thin film; (3) transport across the metal/semiconductor interface; and (4) transport through the semiconductor depletion region. A general expression for the BEEM current at a

**Table 1.** Theoretical considerations for BEEM spectral analysis.

---

• <i>Tunneling</i>
Tip
Bias (injection energy)
Shape/size/angular spread of electrons
Tunneling theory
Metal thin film surface
Surface electronic structure
Surface topography/gradients
• <i>Transport in the film</i>
Metal film (ballistic transport)
Growth morphology
Thickness
Elastic/inelastic scattering
Electron–electron
Electron–phonon
Defects
Quantum interference effects
Energy dependence
Band structure
• <i>Transport across the interface</i>
Schottky barrier
Epitaxy/non-epitaxy/ $k_{  }$
conservation/non-conservation
Band-structure match-up
Dislocations, defects, etc.
• <i>Transport in the semiconductor</i>
Image potential effects
Dopant density
Band structure
Elastic/inelastic scattering
Electron–electron
Electron–phonon
Impact ionization (electron–hole pair generation)
Depletion region
• <i>The effects of interlayers</i>

---

temperature  $T$  can be written as

$$I_b \propto \int D(T, E) f(E, T) E dE$$

where  $f(E, T)$  is the energy distribution of electrons reaching the interface and  $D(E, T)$  is the total transmission probability for electron transport from the point of injection to the back of the semiconductor. In their pioneering BEEM studies of the Au/Si(100) interface, Bell and Kaiser [14] derived an expression for

the bias dependence of the BEEM current by assuming that the energy dependence of both the transmission probability and the energy distribution of the injected electrons can be neglected near the threshold region, which results in a square law dependence for  $I_b$  versus  $V_t$  given by

$$I_b \propto e^2(V_t - \phi_{SB})^2$$

where  $\phi_{SB}$  is the Schottky barrier height. However, in a BEEM experiment the energy distribution of the injected electrons decays exponentially from a maximum value of  $eV_t$ . In addition, the effect of quantum mechanical reflection at the metal/semiconductor interface is strongest near the threshold region. These effects are expected to result in deviations from the square law dependence of the BEEM current within a few hundred millielectronvolts of threshold. The inclusion of the energy dependence of the transmission coefficient due to quantum mechanical reflection results in a 5/2-power law given as

$$I_b \propto e^2(V_t - \phi_{SB})^{5/2}$$

for the expansion of  $I_b$  near threshold [12]. On the other hand, it is expected that phonon scattering within the depletion region [13] will partially cancel the effect of quantum mechanical reflection at finite temperature yielding a result which again approaches the square law dependence derived by Bell and Kaiser [14].

To properly interpret BEEM data, a fundamental understanding of the complete electron transport process from injection into the metal overlayer to collection in the semiconductor is necessary. The term *hot electron* is often used to describe electrons in BEEM transport processes. Hot electrons are simply free carriers with kinetic energies greater than

approximately  $kT$  (40 meV at room temperature). Specifically, these electrons must undergo scattering processes in order to reach thermal equilibrium [15].

The energy and momentum distribution of electrons injected into the metal overlayer has a strong influence on the functional dependence of the spectroscopic lineshape and the interfacial resolution of BEEM. The energy distribution of the injected electrons will decay exponentially, with the actual strength of the decay constant depending both on the height and shape of the vacuum tunneling barrier and the tip apex and surface geometries [16]. Calculations of the energy and momentum distributions of electrons injected into a planar metallic jellium slab from a single-atom scanning tunneling microscope tip apex have indicated that the energy distribution of electrons should be 200 meV or less at energies above 5 eV and that the angular distribution of the electrons should be within about  $\pm 10^\circ$  of the surface normal [17]. This forward focusing of the BEEM electrons predicts nanometer-scale spatial resolution at the metal/semiconductor interface for metal overlayer thicknesses of 10.0 nm or less, and such spatial resolution has been observed for many systems. However, for compound metals such as  $\text{CoSi}_2$  and  $\text{NiSi}_2$ , the influence of the spatial dependence of the local tunneling probability is expected to significantly modify the energy and momentum distribution of the injected electrons. In addition, the atomic potentials of the metals atoms are also expected to affect the actual trajectories of the electrons through the metal base. For ordered metal overlayers, low-energy electron diffraction (LEED) calculations are probably more appropriate for determining the

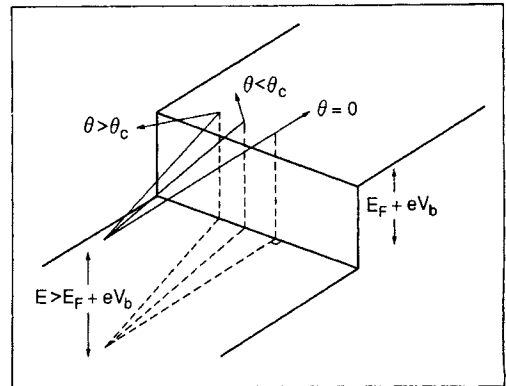
actual momentum distribution of the electrons as they approach the metal/semiconductor interface.

A determination of the relative length scales of elastic and inelastic scattering processes in the metal overlayer is of great importance since these scattering events have a direct effect on the resolution of BEEM at the interface and limit the range of metal overlayer thicknesses which can be probed. Electrons can scatter elastically off defect sites or quasielastically by emission or absorption of acoustic phonons. For metals with more than one atom per primitive basis, the emission and absorption of optical phonons is also possible. Although the energy quanta of acoustic phonons are typically of the order of a few millielectronvolts, the energies of optical phonons for compound metals, such as metal silicides, can exceed 50 meV. Therefore, the determination of whether electron scattering with optical phonons is a quasielastic process depends on the particular modes available for that system. Hot electrons may also undergo inelastic electron-electron collisions with electrons near the Fermi level, resulting in an average energy loss of one-half of their kinetic energy. When the inelastic mean free path for scattering,  $\lambda_i$ , is of the same order as the elastic mean free path,  $\lambda_e$ , the transport process is essentially *ballistic* since electrons which have undergone multiple scattering events will, on average, not have enough kinetic energy to cross the Schottky barrier. In this limit,  $I_b$  should show a clear exponential dependence on metal film thickness with the attenuation length,  $\lambda_b$ , measured by BEEM, given by

$$\frac{1}{\lambda_b} = \frac{1}{\lambda_e} + \frac{1}{\lambda_i}$$

In the limit where  $\lambda_e \ll \lambda_i$ , the transport is expected to be diffusive in nature since the probability for undergoing multiple elastic scattering events before inelastic scattering is high. This can result in an electron momentum distribution at the metal/semiconductor interface that is completely unrelated to the momentum distribution at the point of injection.

At the metal/semiconductor interface, a fraction of the incident electrons will backscatter into the metal overlayer due to quantum mechanical reflection. For non-epitaxial systems and nonabrupt interfaces, there will be a break in symmetry parallel to the interface, resulting in additional interfacial scattering. For epitaxial interfaces,  $k_{||}$  is conserved as the electron travels from the metal to the semiconductor. This condition results in a critical angle for transport into the semiconductor substrate as depicted in Fig. 4. Assuming free electron dispersion relations for the tip and base and a parabolic conduction band minimum for the semiconductor, Bell and Kaiser [14] derived an expression for the



**Figure 4.** Diagram showing a particle of energy  $E$  incident on a potential step of height  $E_F + eV_b$ . There is a critical angle,  $\theta_c$ , beyond which the particle is reflected. For angles of incidence,  $\theta < \theta_c$ , the particle is transmitted with refraction. [37].

critical angle of transmission,  $\theta_c$ , for the Au/Si(100) system, given as

$$\sin^2 \theta_c = \frac{m_s}{m} \frac{eV_t - e\phi_{sb}}{E_F + eV_t}$$

where  $E_F$  is the Fermi energy of the metal,  $m$  is the free electron mass, and  $m_s$  is the effective mass of the electron in the semiconductor. As a direct consequence of  $k_{\parallel}$  conservation, the interfacial spatial resolution of BEEM within a few electronvolts above threshold is expected to exceed the forward-focusing resolution [17]. In addition, a delayed threshold for  $I_b$  is expected for electron transport into semiconductor surfaces where the conduction band minimum is not projected onto the surface Brillouin zone center.

Of the key issues which have resulted in conflicting interpretations of BEEM data, the question of whether  $k_{\parallel}$  is conserved at the metal/semiconductor interface has remained the most controversial. Although most models for interpreting BEEM electron transport, as well as electron transport from internal photoemission data [18], assume that  $k_{\parallel}$  is conserved as the electron crosses the interface, in retrospect this appears to be a somewhat questionable assumption since most metal/semiconductor systems are nonepitaxial in nature. A more relevant issue would be the degree of scattering at nonepitaxial interfaces. In the limit of a jellium metal, the lattice positions of the metal atoms have no influence on the electron trajectory through the metal film, which implies that  $k_{\parallel}$  should be conserved. For d-band metals where the band structure deviates rather significantly from that of a simple metal, scattering at the interface might be expected to be almost isotropic. For all real interfaces, there will

be crystallographic imperfections and interface reactions which will also contribute to the scattering process. Presently, a systematic approach for accurately evaluating the degree of scattering at metal/semiconductor interfaces for any particular nonepitaxial system does not exist.

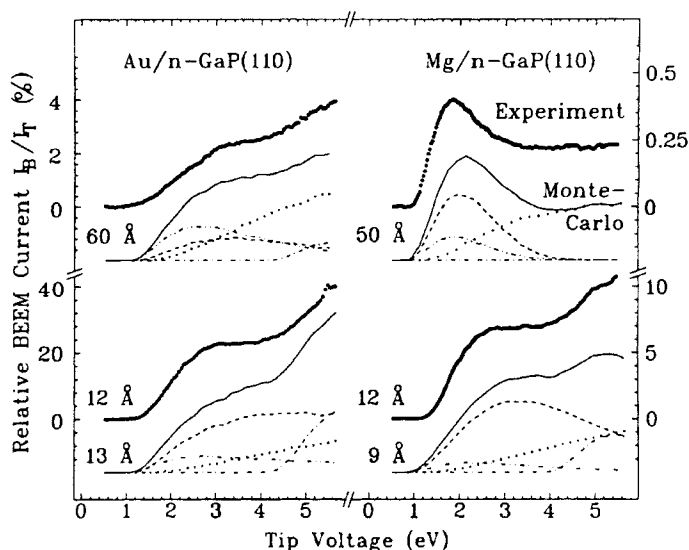
Although scattering processes within the semiconductor substrate are not expected to influence the spatial resolution of BEEM at the interface, they will affect the magnitude of  $I_b$ . Electrons can scatter by emission or absorption of either acoustic or optical phonons. Due to the image potential [19], the position of the Schottky barrier maximum is not at the metallurgical metal/semiconductor interface but is shifted into the semiconductor by a few nanometers. Electrons with energies just over the threshold for transmission that excite phonons in the region before the Schottky barrier maximum can be expected to have a high probability of re-entering the metal. Beyond the Schottky barrier maximum, the internal electric field in the depletion region accelerates the electrons toward the interior of n-type semiconductors. Therefore, the effect of phonon scattering beyond the Schottky barrier maximum on the magnitude of  $I_b$  depends on the doping density of the semiconductor, since this defines the length of the depletion region, and thus the acceleration rate. Once the kinetic energy of the electrons exceeds the band gap of the semiconductor, electron-hole pair generation, or impact ionization, becomes possible. Since the internal electric field in the semiconductor will sweep the electrons toward the interior of the semiconductor and the hole toward the metal, an electron multiplication process occurs.

## 4.5 Ballistic Electron Emission Microscopy Analysis of Schottky Barrier Interfaces

The major emphasis in BEEM analysis is on electron transport across the Schottky barrier interface. The question often comes down to whether or not  $k_{\parallel}$  is conserved in transport across the interface. To first order, this can be answered by determining whether the spectral threshold is consistent with the band structures of the metal base and the semiconductor substrate, assuming that  $k_{\parallel}$  is conserved. However, scattering in the base is a mechanism that can provide parallel momentum to a fraction of the electrons which reduces the threshold to its minimum value, relaxing band structure considerations. Beyond threshold, other scattering and band structure effects in spectra provide insights into hot-electron transport throughout the junction and electronic states at the interface.

A successful demonstration of the quantitative modeling of measured BEEM spectra in which various components of electron transport could be separated was reported recently [20]. BEEM spectra for gold and magnesium films grown on n-type GaP(110) were fitted over a range of 6 eV as shown in Fig. 5. Spectral decomposition using best-fit parameters distinguishes contributions due to zero scattering (dashed lines), elastic scattering (dashed-double-dotted lines), inelastic scattering plus secondaries (dotted lines), and impact ionization (dashed-dotted lines). One notable difference between these systems is that the overall current magnitudes for the Mg-GaP system are considerably less than that for gold ( $\approx 40\%$  of injected current for 1.2 nm-thick gold films), and the spectral shapes are different, especially near threshold. The difference in the total transmission probability between the two films was found to originate from vastly different mean free paths ( $\lambda_{in}^0 = 51.0 \text{ nm eV}^2$  for

**Figure 5.** A comparison of measured and simulated BEEM spectra for several film thicknesses of magnesium and gold on n-type GaP(110). Zero scattering (dashed lines), elastic scattering (dashed-double-dotted lines), inelastic scattering plus secondary electrons (dotted lines), and impact ionization (dashed-dotted lines) contributions to the measured spectra are isolated in the theoretical calculations. [20].



gold and  $\lambda_{\text{in}}^0 = 8.5 \text{ nm eV}^2$  for magnesium). At energies below 4 eV the strong inelastic scattering causes an overshoot in the BEEM current for the 5.0 nm-thick magnesium films and obscures impact ionization contributions.

### 4.5.1 Epitaxial Interfaces

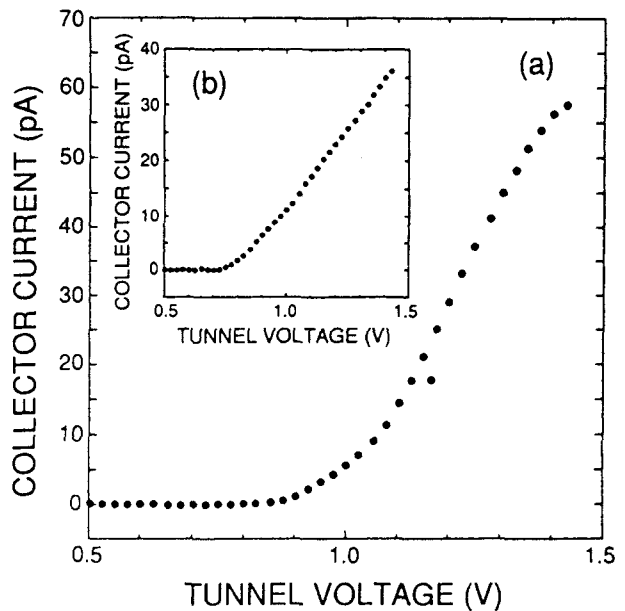
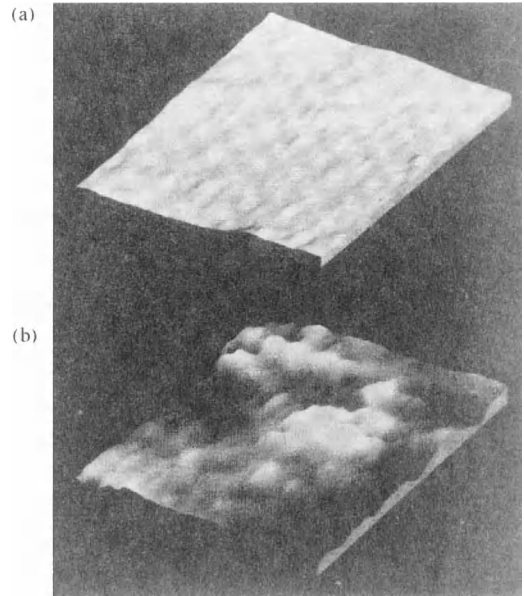
The  $\text{CoSi}_2/\text{Si}$  interface is a system under study for applications such as metal base and permeable base transistors. The Schottky barrier properties of both epitaxial  $\text{CoSi}_2/\text{Si}(111)$  and  $\text{NiSi}_2/\text{Si}(111)$  diodes, extending over macroscopic dimensions, have been studied intensively in recent years since it was determined that the barrier height could be controlled as a function of growth conditions [8]. These cubic fluorite silicides can be grown with two relative lattice orientations on  $\text{Si}(111)$ , forming high-quality atomically abrupt interfaces which account for the barrier height differences. Since both systems are epitaxial, it is expected that a  $k_{\parallel}$  model for electron transport across the interface is appropriate since there is no break in symmetry parallel to the interface.

For electrons with  $k_{\parallel}$  precisely equal to zero, the energy band alignment results in a gap approximately 1.4 eV above  $E_F$  for the  $\text{CoSi}_2/\text{Si}(111)$  interface. Calculations for the  $\text{CoSi}_2/\text{Si}(111)$  interface predict that transmission can occur about 0.2 eV above the conduction band minimum [21]. Measurements on  $\text{CoSi}_2$  films grown on n-type  $\text{Si}(111)$  and imaged in liquid nitrogen indicated typical threshold values at 0.85 eV [22] at many locations and on different samples. However, several locations gave threshold values of about

0.65 eV, which is the approximate value for the Schottky barrier height, with some difference in spectral shapes. Examples of the measured spectra are shown in Fig. 6. In addition, the BEEM current images exhibit considerable inhomogeneity at the interface which correlates with the range of threshold values. Such observations of BEEM thresholds above the Schottky barrier were the first such results to be reported. These results, including their unique spectral shapes, indicate that the electron transport measurements contain details of the band structure of the epitaxial interface when modeled with the inclusion of  $k_{\parallel}$  conservation. Although the previous experimental results seem to provide evidence for  $k_{\parallel}$  conservation for electron transport across an epitaxial interface, recent BEEM results for  $\text{CoSi}_2/\text{Si}(111)$  measured in situ at 77 K give evidence to the contrary [11]. BEEM current thresholds of 0.66 eV were measured for all regions of their samples with no additional current onset at 0.85 eV. The discrepancies between these two studies is most likely a result of the different sample preparation and measurement conditions. The lack of a delayed threshold for the in situ study indicates that an additional source of scattering at the metal/semiconductor interface was present.

BEEM studies on the  $\text{NiSi}_2/\text{Si}(111)$  system, where  $k_{\parallel}$  conservation is also expected, are more difficult to interpret since the band structure does not provide an energy gap but rather an overlap between  $\text{NiSi}_2$  and Si states at all energies above the Schottky barrier [23].  $\text{NiSi}_2$  films of various thicknesses were grown on an  $\text{Si}(111)-7 \times 7$  interface and analyzed in situ [24]. Atomically resolved STM images of the silicide surface were compared with the





**Figure 6.** *Upper:* simultaneously acquired STM (a) 2.1 nm height range and BEEM (b) 40 pA collector current range at  $V_t = 0.5$  V and  $I_t = 1$  nA images ( $25 \times 25$  nm<sup>2</sup>) of a 10 nm-thick CoSi<sub>2</sub> film on n-type Si(111). *Lower:* a representative BEEM spectrum for the CoSi<sub>2</sub> film most commonly observed at various sample positions. The inset shows representative spectra that are less commonly observed. [22].

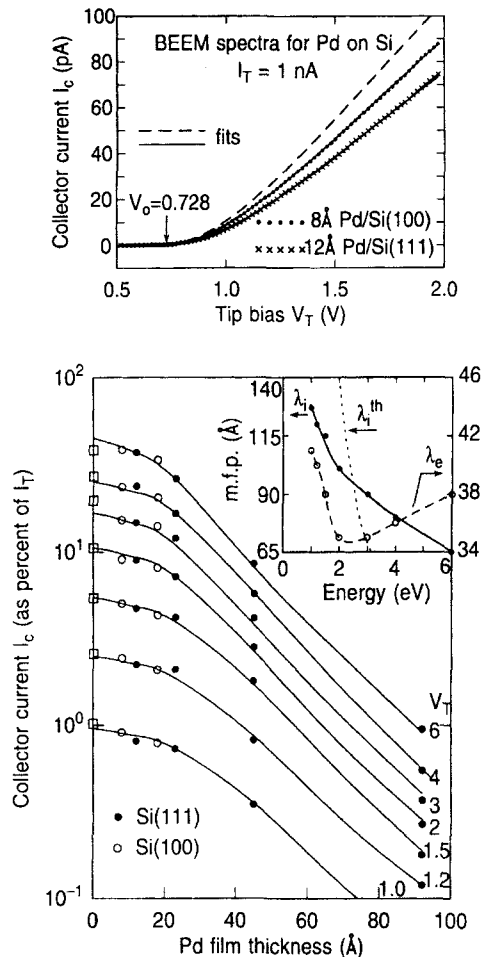
$7 \times 7$  unit cell of the bare Si(111) surface to determine the orientation of the silicon films. The thinner NiSi<sub>2</sub> films ( $< 2.5$  nm) were B-type (rotated  $180^\circ$  with respect to the substrate) whereas the thicker films were A-type and B-type mixed interfaces.

BEEM spectra were acquired on NiSi<sub>2</sub> terraces as well as in pinholes where the  $7 \times 7$  surface is viewed as a zero-thickness metal overlayer. On terraces, transmission of forward-focused electrons cannot occur without a minimum  $k_{||}$ . Thus, the

signature of  $k_{\parallel}$  conservation was taken to be the observed soft threshold, which is also predicted by theory [23]. This observation was contrasted with a sharp threshold in the pinhole regions which is attributed to disorder. In addition, a 25% increase in the BEEM current was observed in the region immediately surrounding the NiSi<sub>2</sub> terraces. This provides additional evidence for  $k_{\parallel}$  conservation because the angular distribution of the injected electrons at the terrace edges is expected to be broadened, thus resulting in increased coupling to silicon conduction band states located away from the surface Brillouin zone center.

#### 4.5.2 Nonepitaxial Interfaces

A study of electron transport in thin films where  $k_{\parallel}$  conservation was expected to break down was performed on palladium films of various thicknesses (0.8–9 nm) grown on Si(111) and Si(100) substrates [10]. The surface topography as measured by STM provided an assessment of local surface gradients which were small enough such that off-normal injection would not lead to direct transmission into conduction band minima at  $k_{\parallel} \neq 0$ . Figure 7a shows a series of BEEM spectra as a function of bias. The threshold is at approximately 0.73 eV, independent of the thickness, and the spectra are linear to approximately 2 eV, at which point there is a fall-off due to the density of states effects. A second threshold, most evident for the thinnest films, appears at approximately 2.9 eV and arises from electron–hole pair creation. For thicker films, the second threshold was less intense due to scattering in the



**Figure 7.** (a) Comparison of BEEM spectra for 0.8 nm Pd/Si(100) and 1.2 nm Pd/Si(111) illustrating the differences due to scattering at the interface. The dashed curve was generated for zero metal thickness and a transmission probability of 0.7. (b) Attenuation of  $I_c$  with palladium thickness as a function of tip bias,  $V_T$ . The inset shows the energy dependence of the elastic and inelastic electron mean free paths. [24].

metal base layer. A plot of collector current versus palladium film thickness is shown in Fig. 7b; both Si(111) and Si(100) substrate data lie along the same respective curves as expected for nonconservation of  $k_{\parallel}$ . The attenuation of the BEEM current exhibits two slopes due to the fact that the

scattering in thin metals is less significant than in thicker metal films. A model based on an isotropic momentum distribution at the interface was developed to analyze the series of plots of attenuation behavior. This result (inset to Fig. 7b) gave the energy-dependent elastic and inelastic mean free paths ( $\lambda_e$  and  $\lambda_i$ , respectively) which were fitted from the data using the phenomenological equation

$$p = \exp\frac{-w}{\lambda_b} + \left(1 - \exp\frac{-w}{\lambda_e}\right) \exp\frac{-4w^2}{\lambda_e\lambda_i}$$

where  $p$  is the probability of an injected electron reaching the interface and  $w$  is the film thickness. The first term is the probability of surviving ballistic passage and the second term is the product of the probability of elastic scattering and the probability of surviving  $\bar{n} \approx 2w/\lambda_e$  scattering events without inelastic scattering. According to this formulation, both the elastic and inelastic mean free paths can be treated independently.

### 4.5.3 Au/Si Interfaces

The Au/Si system was the first used to demonstrate the BEEM technique, and is probably one of the most comprehensively studied systems. However, there remains a curious set of inconsistencies in the data regarding these interfaces, some of which can be directed toward variations in the interfaces themselves and some in sample preparation. Several investigators have observed BEEM current for these interfaces while others have measured no current. Some investigators have concluded the validity of  $k_{\parallel}$  while others have discounted this assumption.

It has already been noted that the pioneering experiments of Kaiser and Bell demonstrated very good agreement between their data and the model assuming  $k_{\parallel}$  conservation at the Au/Si(100) interface [14]. However, similar lineshapes for Au/Si(100) and Au/Si(111) have been observed [25], which is inconsistent with  $k_{\parallel}$  conservation because of band structure differences between the two substrate faces. In essence, transport into Si(111) requires a finite  $k_{\parallel}$  to couple to the states at the conduction band minimum while transport into Si(100) does not. Their Monte Carlo calculations suggested that strong elastic (and quasielastic) scattering in the gold creates an isotropic momentum distribution which reduces the spatial resolution and relaxes the requirement of  $k_{\parallel}$  conservation at the interface. This conclusion was further supported by BEEM measurements which observed no correlation between surface gradients and BEEM current [26]. In another study of Au/Si, an SiO<sub>2</sub> interlayer grid was grown at the interface to provide information on the limits of subsurface spatial resolution [27]. It was found that BEEM currents varied over distances of about 1.0 nm, leading to a conclusion that ballistic transport must dominate through the gold layer. It was also noted that the observed spatial resolution was significantly higher than would be expected from simple planar tunneling theory. High interfacial resolution of BEEM for Au/Si(100) interfaces was also confirmed in a study in which Schottky barrier height fluctuations were directly imaged with nanometer-scale resolution for highly doped substrates [28]. However, another study found that no BEEM current was transmitted through an Au/Si(111) contact that was grown

under UHV conditions [29]. This result was attributed to a reaction at the interface forming a disordered gold–silicon alloy, which acts as a source of strong scattering. Growing gold on an H-terminated surface resulting in uniform BEEM currents. In addition, it was noted that higher biases caused an irreversible modification in BEEM transmission for these interfaces, which was attributed to an electron-induced enhancement of gold–silicon interdiffusion [29]. More recently, however, Cuberes et al. [30] found that BEEM currents could be measured up to 8 eV tip biases without resulting in interfacial modification for UHV-grown samples. Trying to reconcile all these results is difficult because there are several experimental differences, especially in sample preparation and growth and in methods of detection.

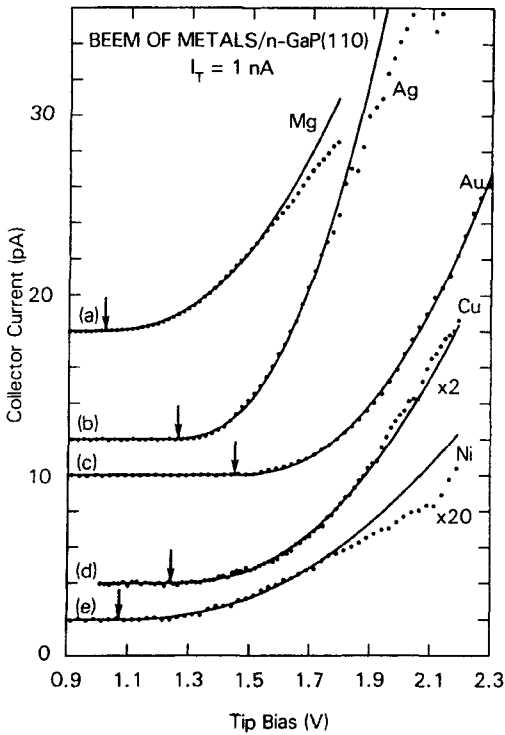
It appears that electron scattering mechanisms and the scattering strength in the metal layer are important components for electron transport analysis. Recently, a temperature-dependent BEEM study of the Au/Si(100) interface was performed for various layer thicknesses ( $\approx 6.5$ – $34.0$  nm) over a narrow range of energies [31]. The attenuation length in the gold overlayer was determined to be energy-independent with only a small temperature dependence leading to the conclusion that defect scattering is the dominant scattering mechanism in the metal film. The temperature-dependent electron transmission through the semiconductor depletion region was also probed and found to be dominated by acoustic phonon scattering. These measurements indicated that, at room temperature, almost half of the electrons which pass into the semiconductor are backscattered into the metal base.

#### 4.5.4 Metal-Film Dependence

Classical Schottky barrier theory predicts that the barrier height is dependent on the macroscopic properties of the metal and semiconductor. This theory breaks down when the actual atomic structures of interfaces of highly epitaxial systems are taken into account [8]. It has also been shown that structural inhomogeneities account for major variations in Schottky barrier heights [32]. Even without a detailed analysis of individual interfacial structures, it is still expected that different metals on the same substrate should produce different Schottky barrier heights. A systematic BEEM investigation of magnesium, silver, copper, nickel, and bismuth on GaP(110) was performed in a UHV BEEM system where the films were grown and analyzed in situ [33]. A different threshold was observed for each metal film as shown in Fig. 8, and differences in the average electron transmission coefficients were noted for the different metals. Variations in the BEEM current at different positions on the sample were attributed to the surface morphology of the respective film (see below). Despite such current variations, the threshold positions were maintained, demonstrating the uniformity of the interfacial potential barrier.

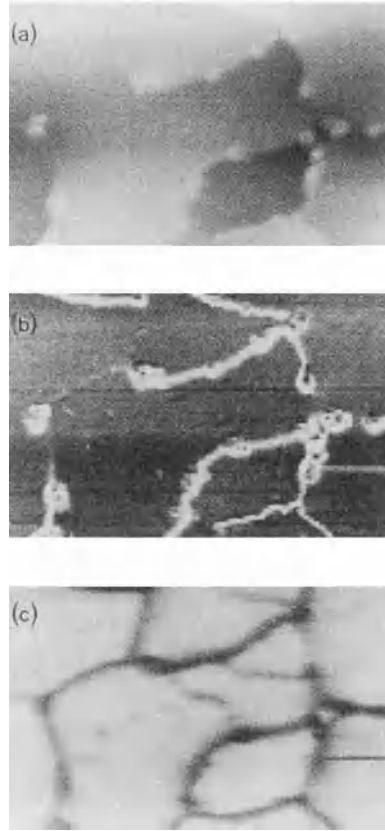
#### 4.5.5 Surface Gradients

In the systematic study of various metals grown on GaP(110) described above, it was found that the surface morphology is the major determinant for current variations in a BEEM image—not interfacial structure, which might also be present and provide contrast in a BEEM image [33]. Figure 9



**Figure 8.** BEEM spectra of various metal/GaP(110) interfaces: 8.0 nm Mg, 20.0 nm Ag, 6.0 nm Au, 15.0 nm Cu, and 5.0 nm Ni. The energies of Schottky barrier thresholds are found to depend on specific metal overlayers, while spectra taken at various positions for a given sample exhibited fixed thresholds. Ref. [33].

compares the surface morphology of a 20.0 nm magnesium film with its respective BEEM image. In the BEEM current image, the lowest currents were recorded beneath positions on the surface with high topographic gradients. A closer look at the surface morphology, where inclinations of below  $10^\circ$  existed, led to the conclusion that the transmitted current is reduced because the electrons are injected into the metal film off-normal (with respect to the interface). Neglecting scattering in the film, the forward momentum component is no longer normal to the interface, which results in reduced coupling to the substrate



**Figure 9.** (a) Topographic image (1.2 nm height range) of a  $60.0 \times 40.0 \text{ nm}^2$  area of a 20.0 nm thick magnesium film on GaP(110); (b) gradient image of (a) (0–30% gradient range); (c) BEEM image (8–28 pA collector current range) of (a) collected simultaneously. Ref. [33].

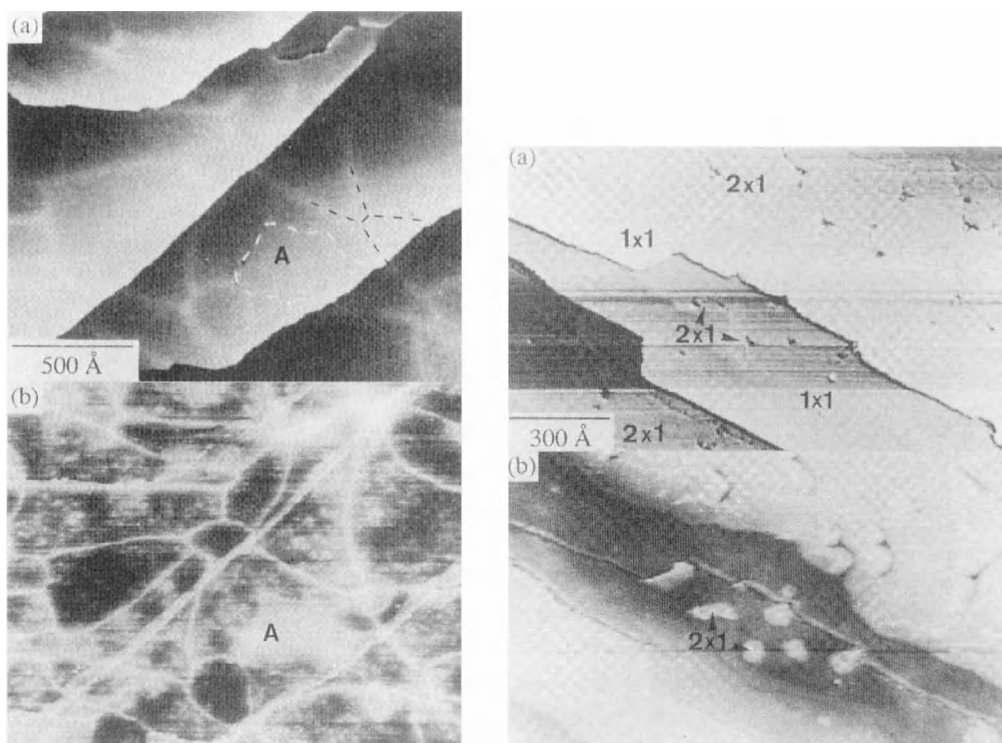
band structure and can result in total reflection away from the barrier, causing a marked reduction in collector current. The effect of surface gradients on the measured BEEM current is commonly referred to as the “searchlight effect”.

#### 4.5.6 Interfacial Nanostructures

Perfectly epitaxial, atomically abrupt interfaces are idealizations of real

heterojunctions. While considerable efforts are made to prevent interfacial inhomogeneities such as disorder, defects, dislocations, and other nanoscale structures at interfaces, the specific material properties often conspire against ideal structures. As noted above, interfacial imperfections represent scattering centers that destroy the periodicity and  $k_{\parallel}$  conservation at the interface. Besides, such features can result in variations in the local Schottky barrier height. It is reasonable that interfacial defects might create sharp contrast in BEEM images so that the underlying interfacial defect structure can be mapped.

Scattering at interfacial dislocations has been observed for partially strain-relaxed  $\text{CoSi}_2$  epitaxial films (2–7 nm thicknesses) grown on  $\text{Si}(111)$  [11, 34]. Film growth was performed in a UHV BEEM system, and imaging was performed in situ at 77 K. The STM topograph of Fig. 10 reveals steps from misorientation of the wafer as well as a hexagonal array of dislocations which result from relaxation of the strain at the interface and produce strain fields that distort the surface to such an extent that they are resolved with the scanning tunneling microscope. The corresponding BEEM images exhibit a current enhancement of approximately 20% at the dislocations



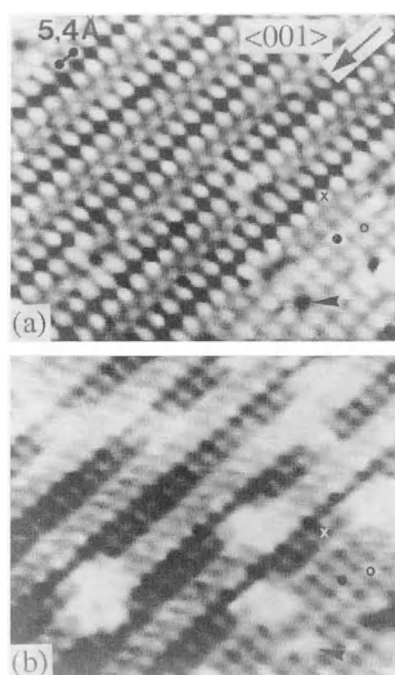
**Figure 10.** *Left:* (a) STM and (b) BEEM images in the vicinity of dislocations on a 2.5 nm-thick  $\text{CoSi}_2/\text{Si}(111)$  heterojunction indicate the variability in BEEM current due to dislocations. Region A exhibits a  $2 \times 1$  reconstruction, and the arrow is directed at a surface point defect. *Right:* (a) STM and (b) BEEM images on a 1.9 nm-thick  $\text{CoSi}_2/\text{Si}(111)$  film. The BEEM current in the  $2 \times 1$  region is approximately 40% higher than for the  $1 \times 1$  region. [34].

that decorate the BEEM image. The spatial dimension of the current enhancements at these interfacial imperfections is extremely sharp ( $\approx 0.8$  nm wide). This implies that the majority of scattering occurs at the interface, which broadens the momentum distribution and enhances the current since access to states at the conduction band minimum requires momentum transfer parallel to the interface. The enhancement was most pronounced approximately 1 eV above the Schottky barrier threshold, whereupon it begins to decrease due to band structure effects. In addition, local defects unrelated to the dislocations were observed in the current images since they also serve as scattering centers which provide parallel momentum transfer. The unstressed interface is not reconstructed. However, region A exhibits a  $2 \times 1$  reconstructed region that is created by local stress; this region shows an increased current due to a change of surface electronic structure alignment of the conduction band minimum. Small differences in contrast from step to step occur because of quantum interference effects arising from atomic variations in the film thicknesses.

#### 4.5.7 Local Electron Tunneling Effects

Much of the effort in the analysis of a BEEM experiment is devoted to understanding the details of electron transport *after* the electrons have been injected into the metal film. It is clear, however, that the surface electronic structure controls tunneling between the tip and surface atoms. It is recognized that atomic resolution can

be achieved in topographic and current STM images and that the local density of states has an effect on the bias dependence of STM images. Recent experiments conducted in a UHV BEEM apparatus at 77 K on  $\text{CoSi}_2$  grown on both  $\text{Si}(100)$  ( $\sqrt{2} \times \sqrt{2}\text{R}45^\circ$  and  $3\sqrt{2} \times \sqrt{2}\text{R}45^\circ$  structures) and  $\text{Si}(111)$  ( $1 \times 1$  and  $2 \times 1$  structures) surfaces have demonstrated that atomically resolved features can be observed in BEEM images [35]. On n-type substrates, this effect is manifest as a variation of the BEEM current on the atomic scale beyond the Schottky barrier threshold. Figure 11 compares STM and BEEM images for the ( $\sqrt{2} \times \sqrt{2}\text{R}45^\circ$ ) structure on  $\text{Si}(100)$ ; the BEEM current



**Figure 11.** (a) STM topograph and (b) simultaneously acquired BEEM image for the silicon-rich  $\text{CoSi}_2/\text{n-Si}(100)$  surface (3.8 nm film thickness). The central area exhibits a  $3\sqrt{2} \times \sqrt{2}\text{R}45^\circ$  reconstruction and the lower right shows a  $\sqrt{2} \times \sqrt{2}\text{R}45^\circ$  reconstruction. [35].

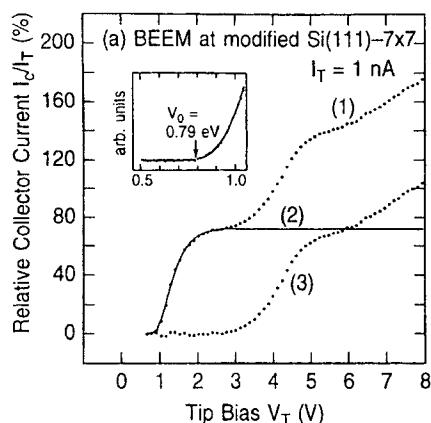
is reduced by approximately 20% at points of topographic maxima over a range of  $V_t$  between  $-1.0$  and  $-2.4$  V. This out-of-phase image contrast is similar to that observed for  $(3\sqrt{2} \times \sqrt{2}\text{R}45^\circ)$ ,  $(\sqrt{2} \times \sqrt{2}\text{R}45^\circ)$  and  $(3\sqrt{2} \times \sqrt{2}\text{R}45^\circ)$  regions can be distinguished even after atomic resolution is lost up to approximately 6 eV because of differences in the average local tunneling barrier heights. The fact that the atomic corrugation directly corresponds to the simultaneously acquired STM image is indicative that some detail of the tunneling interaction at the surface controls the process. Assuming  $k_{\parallel}$  conservation at the epitaxial interface, atomic resolution in BEEM could originate from variations in the energy distribution (details of the local tunneling barrier) or the momentum distribution (spread of the angular cone) of the carriers. The difference between injection into Si(100) and Si(111) surfaces is that the projected conduction band minimum is in the center of the Brillouin zone for Si(100). Thus, for the Si(100) substrate, a broadening in the momentum distribution would cause a reduction in the BEEM current especially near the Schottky barrier threshold. The fact that the contrast persists for higher energies, where the momentum distribution is less critical, makes variations in the energy distribution more likely. Furthermore, out-of-phase contrast between STM and BEEM are also seen for Si(111), where a broader momentum distribution should increase the current. At least for energies near threshold, changes in the energy distribution of injected carriers induced by local variations of the tunneling barrier are thus regarded as the origin of atomically resolved BEEM. However, model

calculations of transport including the energy and momentum effects still significantly underestimate the observed contrast, so that only a qualitative explanation of atomically resolved BEEM presently exists.

#### 4.5.8 Impact Ionization

An energy loss mechanism for electrons in the semiconductor substrate with a kinetic energy greater than the band gap is the generation of electron-hole pairs. In a BEEM spectrum, this is manifest in the observation of excess current, i.e. electron multiplication effects in the semiconductor, since the internal field in the semiconductor will sweep the holes toward the metal base and the electrons to the back of the depletion region. This phenomenon was first observed for disordered 5.0 nm chromium films that were grown on n-type GaP(110) [36]. More recently, impact ionization effects in BEEM spectra on very thin NiSi<sub>2</sub> films grown on Si(111) were reported [24]. BEEM spectra acquired over a high-quality epitaxial region (B) as well as at a pinhole (A), where a modified  $7 \times 7$  structure exists are shown in Fig. 12. The pinhole spectrum allows a direct analysis of scattering events in the semiconductor since the  $7 \times 7$  reconstruction results in an ultrathin metallic overlayer. This is regarded as the zero-metal-thickness limit for electron transport. Direct comparison with experiment indicated that primary electrons pass to the semiconductor with a 72% efficiency. Schottky barrier thresholds at 0.79 eV were observed for both the  $7 \times 7$  pinhole regions and the B-type NiSi<sub>2</sub>. For injection





**Figure 12.** Curve (1) is a representative BEEM spectrum acquired in a pinhole of a 2.2 nm B-type NiSi<sub>2</sub> film. Curve (2) is the primary electron component, and curve (3) is the secondary electron component which occurs due to impact ionization in the semiconductor. The inset shows the threshold region of the spectrum. [24].

energies greater than 3 eV, a second current threshold was observed, which is attributed to the onset of impact ionization in the silicon. For kinetic energies of approximately 5 eV, the quantum yield for impact ionization,  $\gamma$ , is of the order of unity, i.e. one primary electron creates a secondary electron.

## 4.6 Probing Beneath the Schottky Barrier

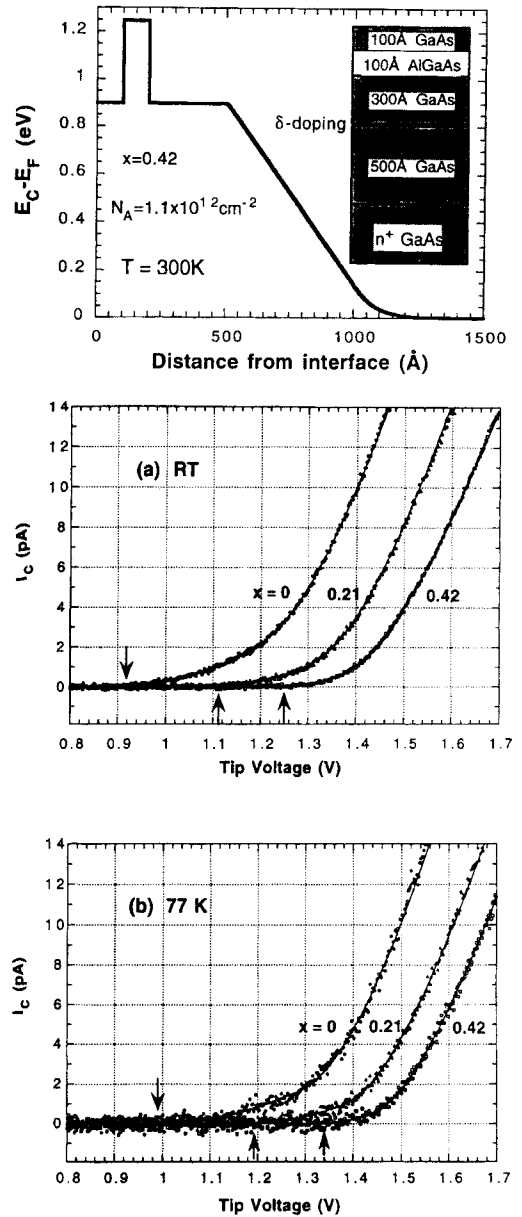
Thin film systems are not restricted to situations in which an individual layer is deposited on a substrate. Ultrathin multiple-layer structures are particularly relevant for new device applications. As an aid in the evaluation of novel devices, BEEM can be applied as a localized probe of interfacial and thin film electronic properties of semiconducting or insulating

interlayers by analyzing electrons that have traversed several layers. To perform the BEEM measurements, electrons are injected into a metal layer that has been deposited on the desired multiple-layer structure. Epitaxy and disorder, band structure and alignment, and electron transport across each interface and through each material can substantially complicate a quantitative BEEM analysis. However, some examples of recent work illustrate the extension of BEEM capabilities to such structures of increased complexity.

Typically, analysis of electron transport in the semiconductor using BEEM is complicated by scattering events in the metal overlayer and at the metal/semiconductor interface which result in modifications of the initial momentum and energy distributions of the injected electrons. However, a heterostructure configuration—a gold thin film grown on an epitaxial p-type layer on an n-type Si(100) substrate—has been used recently to address electron scattering effects in the semiconductor [48]. From these results temperature-dependent attenuation lengths within the semiconductor were determined. With a p–n configuration beneath the Schottky barrier, the actual barrier maximum can be placed tens of nanometers below the Schottky barrier. The barrier maximum is expected to be dependent on the thickness. When the p layer is thick enough so that it is not fully depleted, the barrier maximum depends only on the doping density. In addition, the barrier maximum is expected to be temperature-dependent due both to temperature effects in electron scattering and to the band structure. Experimental measurements were made for p-type layers with thicknesses of 0–100 nm, and

theoretical modeling, incorporating  $k_{||}$  conservation at the interface, produced excellent fits to the data. The barrier height saturated at an approximate thickness of 40 nm because the layer was not fully depleted, and the temperature shift in the barrier height was in agreement with previous data. Interestingly, the data analysis produced an unexpected result regarding the temperature dependence of the attenuation length,  $\lambda(T)$ . In particular, the value of approximately 45.0 nm obtained for  $\lambda(300\text{ K})$  is significantly longer than that the valued obtained for  $\lambda(77\text{ K})$  of 15.0 nm. This was tentatively attributed to the need for increased scattering at higher temperatures to provide electron momentum to couple the injected electrons to the states at the silicon conduction band minimum that exist  $k_{||} \neq 0$ .

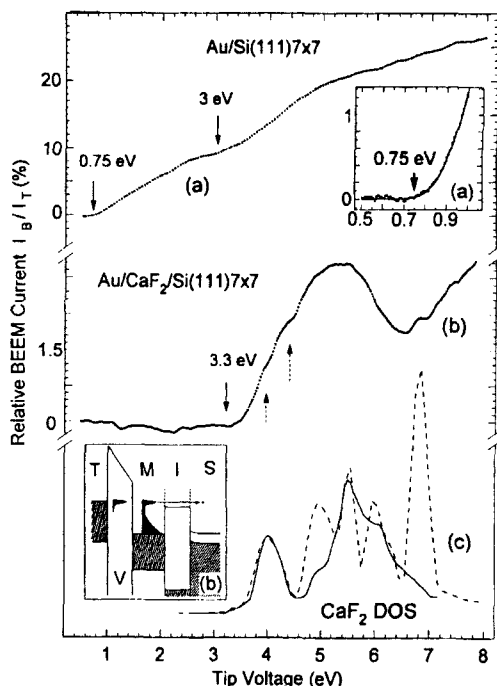
The measurement of conduction band discontinuities and transport across flat-band semiconductor heterostructures beneath the Schottky barrier has also been demonstrated. Conduction band offsets for  $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$  heterostructures as a function  $x$  (the aluminum mole fraction) were investigated [38]. The thin film configuration shown in the inset of Fig. 13a was adopted for these measurements. A p-type  $\delta$ -doped layer (beryllium) of appropriate density was inserted during growth to compensate for the band bending due to the Schottky barrier (at 300 K). The measured spectra, obtained both at 300 and 77 K for  $x = 0, 0.21$ , and  $0.42$ , are shown in Fig. 13b. The  $x = 0$  threshold (0.92 eV) is consistent with previously determined Schottky barrier values for Au/GaAs. The observed trend, an upshift in threshold energy with increasing  $x$ , can be attributed to the energy barrier of the AlGaAs interlayer where this barrier



**Figure 13.** Upper: calculated conduction band energy profile for the  $\Gamma$  valley of  $\text{Al}_{0.42}\text{Ga}_{0.58}\text{As}$  on a GaAs structure assuming a Schottky barrier height of 0.9 eV at room temperature. A p-type  $\delta$ -doping sheet is indicated. Lower: comparison of BEEM spectra for  $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$  single-barrier structures (at RT (a) and 77 K (b)). Changes in the threshold voltages correspond to band offsets of the GaAs and AlGaAs  $\Gamma$  conduction bands. [36].

height is the difference between thresholds. Two-threshold fits in accordance with the respective band structures ( $\Gamma$  and L for  $x = 0$  and  $\Gamma$  and X for  $x = 0.21, 0.42$ ) were found to best model the spectra for the determination of a precise threshold value. Although these offsets could be measured by other techniques, the fact that the BEEM measurement is localized laterally might aid in the optimization of ultrathin structures with confined (e.g. nanometer scale) lateral dimensions.

Metal/insulator/semiconductor (MIS) structures are used in many current electronic devices such as MOSFETs. Interest has recently focused on the application of  $\text{CaF}_2$  as insulating thin films because of a small lattice mismatch with silicon offering the potential for epitaxial growth. The geometric and electronic structure of  $\text{CaF}_2$  on Si(111) has been studied by a host of techniques including STM [39]. A BEEM study of the  $\text{Au}/\text{CaF}_2/\text{Si}(111)$  MIS structure focused on determining the effect of a wide band gap insulating layer on electron transport [30, 40]. The most significant observation was that BEEM spectra conform to the general width of the  $\text{CaF}_2$  density of states, that is the conduction band states “filter” the ballistic electrons that traverse the film. In addition, the  $\text{CaF}_2$  interlayer was found to inhibit the propagation of surface silicides which exist for gold growth on bare silicon at room temperature. Figure 14 compares the spectra of  $\text{Au}/\text{Si}(111)$  and  $\text{Au}/\text{CaF}_2/\text{Si}(111)$ . For  $\text{Au}/\text{Si}(111)$ , the BEEM threshold appears at 0.75 eV above  $E_F$ , and a feature at 3 eV was attributed to the onset of impact ionization. When a  $\text{CaF}_2$  interlayer of four monolayers thickness ( $\approx 1.4$  nm) was incorporated, the current threshold shifted by approximately



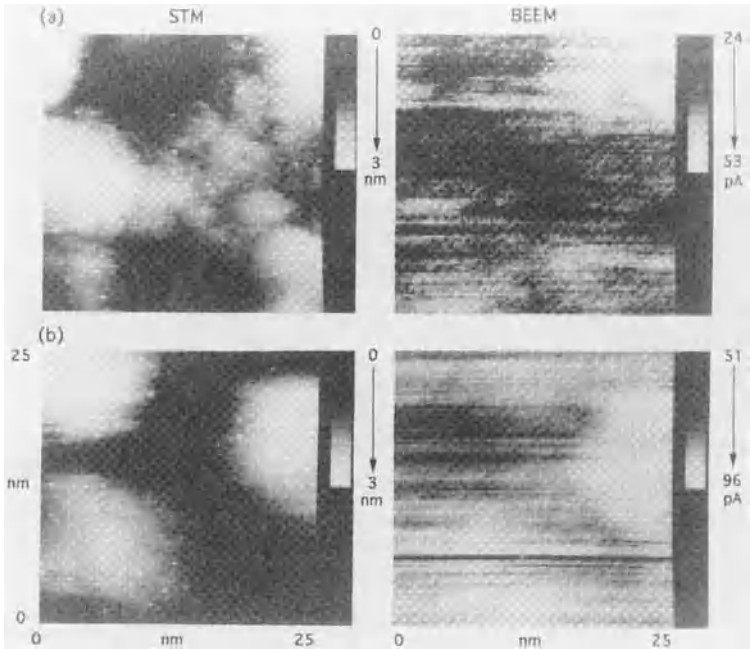
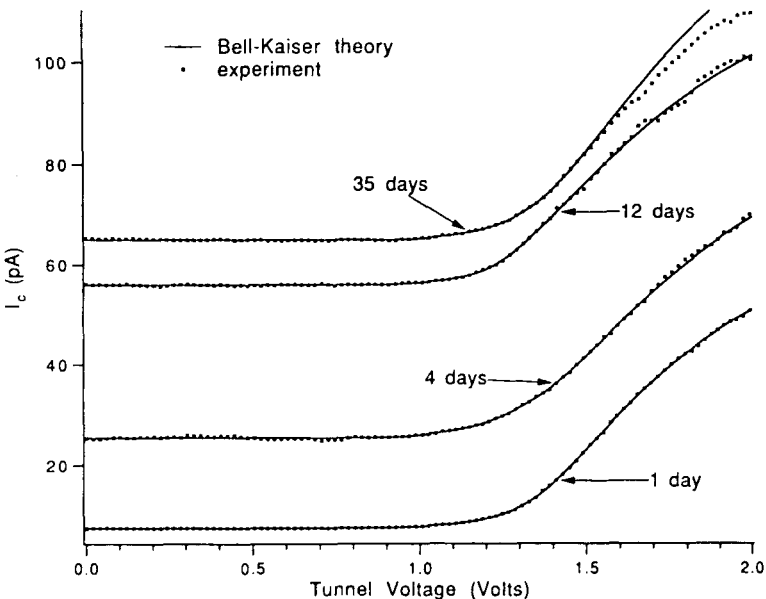
**Figure 14.** Representative BEEM spectra for (a) a 30 monolayer gold film on Si(111)- $7 \times 7$  and (b) a 30 monolayer gold film plus four monolayer  $\text{CaF}_2$  film on Si(111)- $7 \times 7$ . The inset shows the energetics of the experiment for an MIS structure. (c) The calculated bulk  $\text{CaF}_2$  density of states. [40].

2.5 eV to about 3.3 eV above  $E_F$ . In addition, the current diminished approximately 3 eV above this threshold, forming an energy window which is approximately the width of the  $\text{CaF}_2$  density of conduction states, which is also shown in Fig. 14. Small features marked by arrows relate to possible impact ionization, but other spectra features can also be related to peaks in the density of states. It should be noted that the density of states in Fig. 14 must be projected out in the direction of electron transport for any direct comparison of features. In addition, the rising background which is observed for higher energies is expected and is due to inelastic processes in electron transport.

Although  $\text{SiO}_2$  is used as a gate oxide in virtually all production MOSFETs, little experimental data exist on the hot-electron properties across these thin films on an atomic scale. This is primarily due to the amorphous nature of the  $\text{SiO}_2$  in these devices which limits the analysis of their structural and transport properties to experimental techniques which do not rely on crystalline order. Recently, an in situ BEEM study of the Pt/ $\text{SiO}_2$ /Si(100) MOS structure was performed to determine directly the phonon scattering rates of electrons through the conduction band of  $\text{SiO}_2$  with nanometer spatial resolution [41]. The samples were prepared by evaporating platinum onto device quality oxide layers approximately 6.2 nm thick which were thermally grown on n-type Si(100) wafers. In addition to the conventional BEEM set-up, a variable bias was applied to the semiconductor substrate with respect to the grounded metal base. This set-up allowed field-dependent transport measurements through the  $\text{SiO}_2$  conduction band and also provided the means to tune the energy distribution of electrons that pass into the silicon substrate. A rapid rise in  $I_c$  was observed at a threshold of approximately 4 eV for  $V_b = 0$  V and +2 V followed by an overshoot in  $I_c$  at approximately 6 eV and a more gradual rise above 8 eV, which can be attributed to the onset of electron-hole pair generation in the silicon substrate. This attenuation has been attributed to the strong electron-phonon coupling in  $\text{SiO}_2$  since the scattered electrons are increasingly accelerated toward the metal base for decreasing values of  $V_b$ . The measured attenuation rates are in good agreement with previous Monte Carlo calculations of electron-phonon interactions in  $\text{SiO}_2$  [42], and

reinforce the interpretation that hot-electron scattering with the acoustic phonon modes dominate over the optical modes in  $\text{SiO}_2$  in this energy range.

A key consideration in the development of real electronic devices is the stability of the interface in its operating environment. Au/GaAs(100) interfaces degrade at room temperature due to gold diffusion into the surface. It was demonstrated that spatial Schottky barrier inhomogeneities exist at oxide-free Au/GaAs interfaces directly after growth [9] (see Fig. 3). BEEM experiments have also shown that an AlAs interlayer inserted as a diffusion barrier between the gold and GaAs greatly improves both the homogeneity of the Schottky barrier and the stability of the device [43]. Surface science experiments typically strive to obtain atomically clean surfaces prior to deposition. However, in the fabrication of GaAs-based devices, an oxide is typically left behind, which results in the enhancement of device stability. BEEM has been used to investigate the effect of a native oxide interlayer over time on the electron transport properties of Au/GaAs(100) interfaces [44]. An approximately 2 nm-thick native oxide ( $\text{Ga}_2\text{O}_3 + \text{As}_2\text{O}_3$ ) on GaAs was left prior to gold deposition. Figure 15 shows that BEEM spectra are still obtainable after more than a month at room temperature in air. However, it was found that many regions did not support a BEEM current after this long a period, suggesting that the oxide slows but does not arrest chemical reactions at the interface. Figure 15 also shows that with increasing time the zero-bias resistance of the interface decreases, resulting in an increase of the current below the threshold bias. This was attributed to diffusion of gold atoms into



**Figure 15.** *Upper:* average BEEM spectra acquired with  $I_t = 2$  nA obtained at different locations on an Au/n-GaAs(100) sample at various times after deposition (the 35 day spectrum is displaced upward by 5 pA). *Lower:* STM and BEEM images acquired (a) 1 day after deposition and (b) 12 days after deposition. [44].

the GaAs, displacing arsenic and acting as dopants or recombination centers. The native oxide exhibits a zero or nearly zero band gap and does not cause significant

electron scattering at the interface. This is in contrast to metals grown on thin oxide layers on silicon. These layers possess a significantly large band gap that impedes

the ballistic electron current across the interface [45].

## 4.7 Ballistic Hole Transport and Ballistic Carrier Spectroscopy

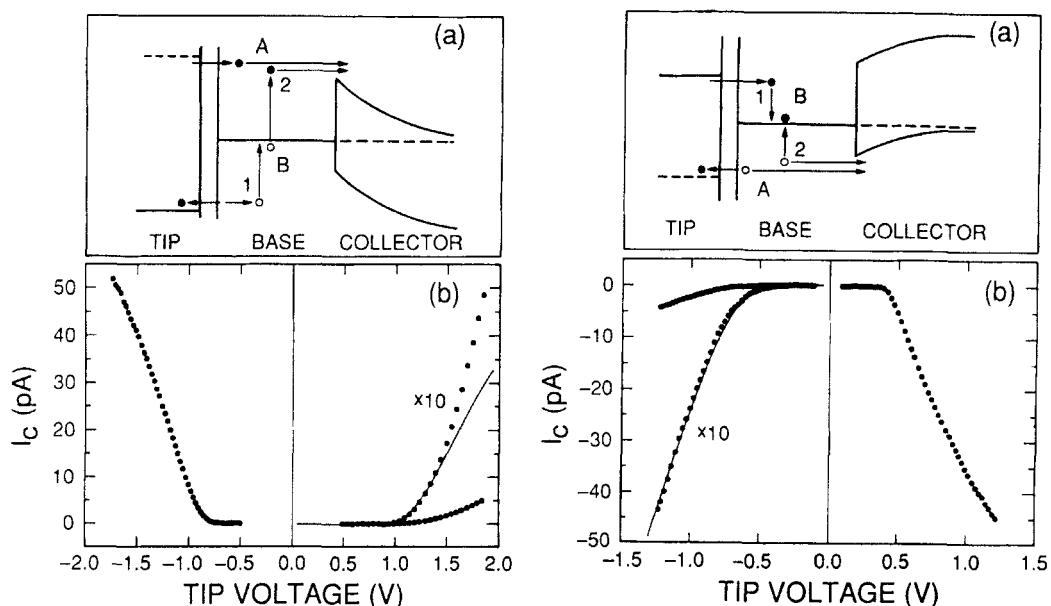
The standard BEEM configuration probes semiconductor conduction band states at the interface. This results from the injection of electrons into the metal base and their subsequent passage into the collector, which is dependent on the energetic position of the conduction band states of the semiconductor. An n-type semiconductor substrate is used to prevent majority carriers from entering the base.

Shortly after the introduction of the BEEM technique, the Jet Propulsion Laboratory group demonstrated a methodology with which to probe the semiconductor valence band states using ballistically transported holes [46, 47]. For this, the metal base was deposited on a p-type substrate, and hole injection was accomplished by reversing the electron tunneling direction with respect to BEEM on an n-type substrate, that is, from base to tip. An additional observation was made that electron and hole scattering spectroscopy in the base region can be achieved by performing  $I-V$  spectroscopy over both polarities for either metal/n-type or metal/p-type junctions.

Figure 16 shows a schematic that depicts ballistic hole transport and electron-hole scattering and measured spectra for Au/Si(100) systems acquired at low temperature. A low temperature was

required because of the magnitude of the Schottky barrier heights. We first consider the ballistic electron and ballistic hole spectra. For BEEM on n-type substrates, electrons are transported across the Schottky barrier (tip negative), and the conduction band levels of the metal/semiconductor interface are probed. For BEEM on p-type substrates, holes are transported across the Schottky barrier (tip positive), and the valence levels of the metal/semiconductor interface are probed. There is an asymmetry in the carrier distributions that are probed. With n-type substrates, the highest electronic levels of the tip are probed; with p-type substrates, the lower levels of the hole distribution take part in conduction. For these respective cases, the number of electrons created per unit energy increases with increasing bias, or the number of holes created per unit energy decreases with increasing bias. Therefore the current threshold for electron transport is at the position of the conduction band minimum, and the threshold for hole transport is at the valence band maximum. The region between these thresholds gives a direct measure of the semiconductor energy gap.

Reversing the bias for either of these cases probes electron and hole scattering in the base. The respective spectra correspond to those carriers that have undergone scattering events and can be modeled as an internal Auger process. For example, in the BEEM configuration with n-type semiconductor substrates, electrons cannot be extracted from the substrate because the Schottky barrier potential opposes the motion of majority carriers into the base. However, a bias configuration where electrons are extracted from the



**Figure 16.** *Left:* (a) The energetics of electron transport for a metal base on an n-type collector (A) using a negative tip voltage and for scattering (B) where the tip is biased positive. Process B creates a ballistic hole distribution that scatters with equilibrium electrons to produce hot electrons. (b) Spectra for Au/Si exhibit an onset at 0.82 eV, the Schottky barrier height for this n-type sample. The scattering spectrum is seen to be less intense by a factor of approximately 10. *Right:* (a) The energetics of electron transport for a metal base on a p-type collector (A) using a positive tip voltage and for scattering (B) where the tip is biased negative. For situation B, electrons injected by tunneling produce a distribution of hot holes after scattering with equilibrium electrons in the base. (b) Spectra for Au/Si exhibit an onset at 0.35 eV, the Schottky barrier height for this p-type sample. Again, the scattering spectrum is seen to be less intense by a factor of approximately 10. Combining the results for these two samples, a gap equal to the substrate band gap is obtained. [46].

base (tip positive) creates energetic holes in the metal base. Thus, ballistic holes scatter from equilibrium electrons and can excite hot electrons above the Schottky barrier. These electrons are collected following transmission across the barrier. Note that the sense of the current is the same for biasing in both polarities. For p-type semiconductor substrates, ballistic electron scattering and hot-hole creation are the analogous processes that occur. In scattering spectroscopy, the currents are attenuated by a factor of approximately 10 from the respective direct transport measurement, indicating an efficient

electron-scattering process. The theoretical fits shown in Fig. 16 are valid only around the threshold.

Ballistic carrier transport and scattering can be modeled using the BEEM formalism where momentum and energy conservation are primary considerations. Energy loss mechanisms showing fully inelastic losses for the promotion of electrons and holes and isotropic scattering angular distributions are assumed in the model that was used to fit the data in Fig. 16. It is interesting to note that comparison of ballistic hole scattering for Au/Si(100) and Au/GaAs(100) reveals that the threshold

shape for the GaAs substrate must take into account the markedly larger spin-orbit splitting of hole bands near the zone boundary [47].

## 4.8 Summary

Methods for the measurement and analysis of local Schottky barrier heights and electron transport in thin film interfaces has developed rapidly in the years since BEEM was introduced. Although the use of the BEEM technique is still restricted to relatively few groups, instrumentation is becoming more sophisticated, and interest in the technique is expanding toward applications on real devices. While first-principles BEEM calculations are on the horizon, comprehensive theoretical calculations continue to give way to more phenomenological analyses which provide new insights into transport phenomena in very thin layers and at metal/semiconductor junctions. The usefulness of BEEM is exemplified by its ability to probe the degree of homogeneity and local Schottky barrier heights by a fairly straightforward measurement—one that could not be done prior to the invention of BEEM—and connected with macroscopic measurements of metal/semiconductor junctions.

## Acknowledgments

NJD acknowledges support for this work from the National Science Foundation under grant DMR 93-13047 and via the Materials Research Laboratory Program under grant DMR 91-20668.

## 4.9 References

- [1] C. B. Duke (Ed.), *Surface Science: The First Thirty Years*, **1994**, 300.
- [2] G. Binnig, H. Rohrer, *Surf. Sci.* **1993**, 126, 236.
- [3] R. J. Hamers, R. M. Tromp, J. E. Demuth, *Phys. Rev. Lett.* **1986**, 56, 1972.
- [4] J. A. Stroscio, R. M. Feenstra, A. P. Fein, *Phys. Rev. Lett.* **1986**, 57, 2579.
- [5] N. J. DiNardo, *Nanoscale Characterization of Surfaces and Interfaces*, VCH, Weinheim **1994**.
- [6] R. Wiesendanger, *Scanning Probe Microscopy and Spectroscopy: Methods and Applications*, Cambridge University Press, Cambridge **1994**.
- [7] W. Schottky, *Naturwissenschaften* **1938**, 26, 843.
- [8] R. T. Tung, *Phys. Rev. Lett.* **1984**, 52, 461.
- [9] W. J. Kaiser, L. D. Bell, *Phys. Rev. Lett.* **1988**, 60, 1406.
- [10] A. R. Ludeke, A. Bauer, *Phys. Rev. Lett.* **1993**, 71, 1760; b. R. Ludeke, A. Bauer, *J. Vac. Sci. Technol. A* **1994**, 12, 1910.
- [11] H. Sirringhaus, E. Y. Lee, H. von Känel, *J. Vac. Sci. Technol. B* **1994**, 12, 2629.
- [12] M. Prietsch, R. Ludeke, *Surf. Sci.* **1991**, 251/252, 413.
- [13] E. Y. Lee, L. J. Schowalter, *J. Appl. Phys.* **1991**, 70, 2156.
- [14] L. D. Bell, W. J. Kaiser, *Phys. Rev. Lett.* **1988**, 61, 2368.
- [15] C. R. Crowell, S. M. Sze, *Solid-State Electron.* **1965**, 8, 979.
- [16] J. Tersoff, D. R. Hamann, *Phys. Rev. Lett.* **1983**, 50, 1998.
- [17] N. D. Lang, A. Yacoby, Y. Imry, *Phys. Rev. Lett.* **1989**, 63, 1499.
- [18] C. R. Crowell, S. M. Sze in *Physics of Thin Films* (Eds.: G. Hass, R. F. Thun), Academic Press, New York **1967**, pp. 325–391.
- [19] S. M. Sze, *Physics of Semiconductor Devices*, Wiley, New York **1984**.
- [20] A. Bauer, M. T. Cuberes, M. Prietsch, G. Kaindl, *Phys. Rev. Lett.* **1993**, 71, 149.
- [21] M. D. Stiles, D. R. Hamann, *J. Vac. Sci. Technol. B* **1991**, 9, 2394.
- [22] W. J. Kaiser, M. H. Hecht, R. W. Fathauer, L. D. Bell, E. Y. Lee, L. C. Davis, *Phys. Rev. B* **1991**, 44, 6546.
- [23] M. D. Stiles, D. R. Hamann, *Phys. Rev. Lett.* **1991**, 66, 3179.
- [24] a. A. Bauer, R. Ludeke, *Phys. Rev. Lett.* **1994**, 72, 298; b. A. Bauer, R. Ludeke, *J. Vac. Sci. Technol. B* **1994**, 12, 2667.
- [25] L. J. Schowalter, E. Y. Lee, *Phys. Rev. B* **1991**, 43, 9308.



- [26] E. Y. Lee, B. R. Turner, L. J. Schowalter, J. R. Jimenez, *J. Vac. Sci. Technol. B* **1993**, *11*, 1579.
- [27] A. M. Milliken, S. J. Manion, W. J. Kaiser, L. D. Bell, M. H. Hecht, *Phys. Rev. B* **1992**, *46*, 12826.
- [28] H. Palm, M. Arbes, M. Schulz, *Phys. Rev. Lett.* **1993**, *71*, 2224.
- [29] A. Fernandez, H. D. Hallen, T. Huang, R. A. Buhrman, J. Silcox, *Appl. Phys. Lett.* **1990**, *57*, 2826.
- [30] M. T. Cuberes, A. Bauer, H. J. Wen, D. Vandré, M. Prietsch, G. Kaindl, *J. Vac. Sci. Technol. B* **1994**, *12*, 2422.
- [31] C. A., Ventrice Jr., V. P. LaBella, G. Ramaswamy, H. P. Yu, L. J. Schowalter, *Phys. Rev. B* **1996**, *53*.
- [32] J. P. Sullivan, *J. Vac. Sci. Technol. B* **1993**, *11*, 1564.
- [33] M. Prietsch, R. Ludeke, *Phys. Rev. Lett.* **1991**, *66*, 2511.
- [34] H. Sirringhaus, E. Y. Lee, H. von Känel, *Phys. Rev. Lett.* **1994**, *73*, 577.
- [35] H. Sirringhaus, E. Y. Lee, H. von Känel, *Phys. Rev. Lett.* **1995**, *74*, 3999.
- [36] R. Ludeke, *Phys. Rev. Lett.* **1993**, *70*, 214.
- [37] L. D. Bell, W. J. Kaiser, M. H. Hecht, L. C. Davis, in *Scanning Tunneling Microscopy* (Eds.: J. A. Stroscio, W. J. Kaiser), Academic Press, San Diego **1993**, pp. 307–348.
- [38] a. J. J. O'Shea, T. Sajoto, S. Bhargava, D. Leonard, M. A. Chin, V. Narayanamurti, *J. Vac. Sci. Technol. B* **1994**, *12*, 2625; b. T. Sajoto, J. J. O'Shea, S. Bhargava, D. Leonard, M. A. Chin, V. Narayanamurti, *Phys. Rev. Lett.* **1995**, *74*, 3427.
- [39] P. Avouris, R. Wolkow, *Appl. Phys. Lett.* **1989**, *55*, 1074.
- [40] M. T. Cuberes, A. Bauer, H. J. Wen, M. Prietsch, G. Kaindl, *Appl. Phys. Lett.* **1994**, *64*, 2300.
- [41] R. Ludeke, A. Bauer, E. Cartier, *Appl. Phys. Lett.* **1995**, *66*, 730.
- [42] M. V. Fischetti, D. J. DiMaria, S. D. Bronson, T. N. Theis, J. R. Kirtley, *Phys. Rev. B* **1985**, *31*, 8124.
- [43] M. H. Hecht, L. D. Bell, W. J. Kaiser, F. J. Grunthaner, *Appl. Phys. Lett.* **1989**, *55*, 780.
- [44] A. A. Talin, D. A. A. Ohlberg, R. S. Williams, P. Sullivan, I. Koutselas, B. Williams, K. L. Kavanagh, *Appl. Phys. Lett.* **1993**, *62*, 2965.
- [45] H. D. Hallen, A. Fernandez, T. Huang, R. A. Buhrman, J. Silcox, *J. Vac. Sci. Technol. B* **1991**, *9*, 585.
- [46] L. D. Bell, M. H. Hecht, W. J. Kaiser, L. C. Davis, *Phys. Rev. Lett.* **1990**, *64*, 2679.
- [47] M. H. Hecht, L. D. Bell, W. J. Kaiser, L. C. Davis, *Phys. Rev. B* **1990**, *42*, 7663.
- [48] L. D. Bell, S. J. Manion, M. H. Hecht, W. J. Kaiser, R. W. Fathauer, A. M. Milliken, *Phys. Rev. B* **1993**, *48*, 5712.

## Part VIII

---

# **Image Recording Handling and Processing**



# 1 Image Recording in Microscopy

## 1.1 Introduction

A microscope as a highly quantitative measuring device for objects of small dimensions requires efficient means of recording the output information not only for visual observation but increasingly even for evaluation by computers. The two-dimensional distribution representing the image must be converted into signals to be stored and processed taking into account different points of view. Off-line methods such as photographic recording requiring a long processing time are increasingly being replaced by on-line converters using advanced optoelectronic technology which is much faster, frequently more accurate, and ready for immediate digitizing. The efficiency and accuracy of such devices are governed by quantum processes, and their design needs a careful analysis in order to optimize the performance with respect to the task.

In this chapter the position-sensitive detectors used in fixed-beam microscopes are treated, excluding the scanning methods. The general physical fundamentals of image characterization before and after recording are reviewed, as well as the different quantum conversion effects utilized in various optoelectronic components, and the way to achieve optimum

performance which photon and electron image recorders have in common. Parameters such as the detective quantum efficiency, resolution, dynamic range, linearity, dependence on the signal rate, and processing speed, which may have different levels of importance in the various methods of microscopy, are discussed. Subsequently, the general relations will be applied to the different types of radiation and recording systems.

## 1.2 Fundamentals

### 1.2.1 The Primary Image

In the image plane of any microscopic system a two-dimensional wave function ('primary image') is formed, the amplitude of which is registered by the recording device. Due to the quantum nature of the radiation the image suffers fluctuations ('noise') governed by Poisson statistics, according to which the probability distribution of finding  $N$  quanta in an arbitrary local and temporal measuring interval is

$$P(N, \bar{N}) = \frac{\bar{N}^N}{N!} \exp(-\bar{N}) \quad (1)$$

resulting in the well-known relation

$\text{var } N = \bar{N}$ . As a consequence, in order to detect a small image detail of area  $d^2$  and contrast  $C = \Delta n/n$  against a background of the local quantum density  $n$ , the signal-to-noise ratio  $\text{SNR} = \Delta n d^2 / \sqrt{n d^2}$  must be sufficiently large. If a detectability condition  $\text{SNR} \geq 5$  is assumed, then the quantum density should satisfy the condition [1]

$$n \geq \left( \frac{5}{Cd} \right)^2 \quad (2)$$

Thus, the size  $d$  and contrast  $C$  of detectable small details with weak contrast are limited by the storage limit of the recorder, even if an ideal detector such as a two-dimensional quantum counter without any further restriction of resolution is assumed. Approaching the storage limit, the SNR decreases. Consequently, any discussion of image recording devices must include the storage medium [2].

### 1.2.2 The General Recorder

A real recorder may be modeled by a two-dimensional device converting the primary image to a type of information to be accumulated in an integrated or subsequent frame store until the required SNR is obtained. The goal of retrieving the full information input may be obstructed by a number of restrictions and disturbances:

- (a) the saturation limit of the storage, as mentioned above;
- (b) the limited spatial field together with the local sampling intervals, frequently called 'pixels' (picture elements);
- (c) local blurring of the image by the point spread function (PSF) of the converter;

- (d) additional sources of noise, reducing the SNR to values below that in the primary image.

A useful parameter for discussing such influences is the detective quantum efficiency [3]

$$\text{DQE} := \gamma^2 \frac{\text{var } S_{\text{in}}}{\text{var } S_{\text{out}}} \quad (3)$$

where  $S_{\text{out}}$  and  $S_{\text{in}}$  denote the signals at the output and input, respectively, accumulated over a partial field and some measuring time, and  $\gamma := d\bar{S}_{\text{out}}/d\bar{S}_{\text{in}}$  the slope of the characteristic curve  $\bar{S}_{\text{out}}(\bar{S}_{\text{in}})$ . In general, the DQE is not only a function of the exposure level (e.g., if the storage limit is approached), it depends also on the size of the chosen partial field if this is small enough to become comparable to the extension of the PSF of the converter. For the time being, we ignore these complications by assuming a sufficiently large field and proportionality between  $\bar{S}_{\text{out}}$  and  $\bar{S}_{\text{in}}$ , and obtain

$$\begin{aligned} \text{DQE} &= \left( \frac{\bar{S}_{\text{out}}}{\bar{S}_{\text{in}}} \right)^2 \frac{\text{var } S_{\text{in}}}{\text{var } S_{\text{out}}} \\ &= \left( \frac{\text{SNR}_{\text{out}}}{\text{SNR}_{\text{in}}} \right)^2 = \frac{\text{SNR}_{\text{out}}^2}{\bar{N}} \end{aligned} \quad (4)$$

which shows the sense of the DQE definition. While an ideal image recorder has  $\text{DQE} = 1$ , real recorders are characterized by  $\text{DQE} < 1$ , requiring an exposure  $\bar{N}$  increasing with  $\text{DQE}^{-1}$  in order to display the same SNR. Thus, condition (2) must be replaced by

$$n \text{DQE} \geq \left( \frac{5}{Cd} \right)^2 \quad (5)$$

This condition is of importance if the number of radiation quanta must be

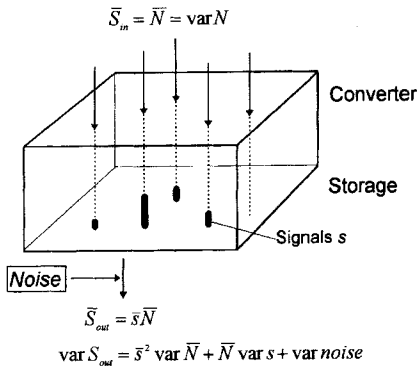


Figure 1. Signal response of an image recorder.

limited, as in imaging radiation-sensitive specimens or dynamic processes using a restricted quantum flux. In these applications particularly, the DQE should approach the ideal value of 1 as closely as possible. Unfortunately, this goal is not always achievable due to the interaction mechanisms which require compromises with the resolution, as discussed below.

An understanding of the degrading statistical processes in an image converter and the design of optimized systems may be facilitated by discussing the signal response to single primary quanta (Fig. 1). Without referring to special carriers of information, this response may be presented by an amount  $s$  of secondary quanta (e.g., electrons, photons, or grains) spread over an interaction range within the PSF of the system and fluctuating according to a probability distribution  $p(s)$ . Under this assumption and using the variance theorem [4], we have  $\bar{S}_{\text{out}} = \bar{N}\bar{s}$  and  $\text{var } S_{\text{out}} = \bar{s}^2 \text{var } N + \bar{N} \text{var } s + \text{var noise}$ ; thus, Eq. (4) may be rewritten as

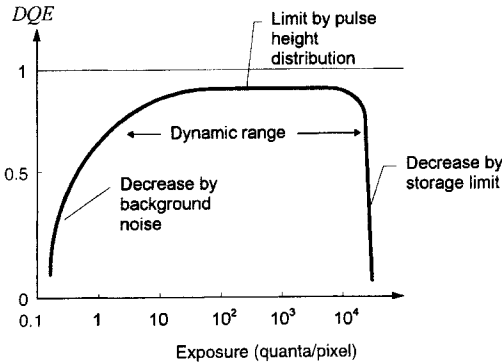
$$\text{DQE} = \left(1 + \frac{\text{var } s}{\bar{s}^2}\right)^{-1} = \frac{\bar{s}^2}{s^2} \quad (6)$$

Thus the DQE is independent of the exposure level  $\bar{N}$ . Obviously, in order to approach the ideal detector the designer should attempt to realize a distribution  $p(s)$  as narrow as possible. Unfortunately, this is not always easy, particularly if the quantum energy of the primary radiation is of the order of the formation energy of the secondary quanta.

While the noise phenomena discussed above were related to the signal to be detected, an additive noise contribution by system components may occur which is independent of the signal. Referring it to the output plane and integrating it over the same area and time as the primary quanta, its variance must be added to  $\text{var } S_{\text{out}}$ . Assuming that the noise signal makes no contribution to  $\bar{S}_{\text{out}}$  one obtains

$$\text{DQE} = \left(1 + \frac{\text{var } s}{\bar{s}^2} + \frac{\text{var noise}}{\bar{N}\bar{s}^2}\right)^{-1} \quad (7)$$

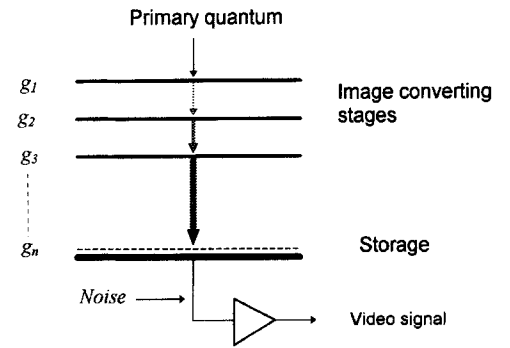
Obviously, this noise contribution makes the DQE dependent on  $\bar{N}$  at low exposure. The designer should reduce such system immanent noise as much as possible. The remaining influence may be reduced by choosing  $\bar{s}$  large enough to make  $\bar{s}^2 \gg (\text{var noise})$ . In this case, however, caution is recommended, since then a given limit of an analog storage medium may reduce the DQE at high exposure according to Eq. (3). Thus, the DQE of any recording device fades at both too low and too high exposure levels  $\bar{N}$  (Fig. 2); an intermediate working region, the 'dynamic range', with a maximum DQE determined by  $\text{var } s/\bar{s}^2$ , can be found. Its width is determined by the storage limit and the noise, and its position on the exposure scale by the signal height  $\bar{s}$ .



**Figure 2.** DQE versus exposure of a general recorder shown schematically.

A sufficiently high single-quantum response  $\bar{s}$  with low variance may be achieved using one or more quantum conversion stages between electrons and photons as components of imaging optoelectronic devices, as discussed below (Fig. 3). Every conversion process at the stage  $m$  ( $m = 1, 2, \dots, n$ ) is characterized by a probability distribution  $p_m(g_m)$  from which a quantum yield  $\bar{g}_m$  and  $\text{var } g_m$  can be deduced. By repeated application of the variance theorem [4] it can be shown [5] that the resulting signal  $s$  has an average value

$$\bar{s} = \bar{g}_1 \bar{g}_2 \bar{g}_3 \cdots \bar{g}_n \quad (8)$$



**Figure 3.** Principle of the multistage image converter.

and a statistical fluctuation

$$\frac{\text{var } s}{\bar{s}^2} = \frac{\text{var } g_1}{\bar{g}_1^2} + \frac{1}{\bar{g}_1} \frac{\text{var } g_2}{\bar{g}_2^2} + \cdots + \frac{1}{\bar{g}_1 \bar{g}_2 \cdots \bar{g}_{n-1}} \frac{\text{var } g_n}{\bar{g}_n^2} \quad (9)$$

From this expression an extremely important design rule can be concluded: the statistical contributions of the subsequent stages with  $m > 1$  may be neglected if the products

$$\prod_{i=1}^m \bar{g}_i \gg 1 \quad (10)$$

for all stages  $m = 1, 2, \dots, n-1$  (even if some of the quantum yields  $\bar{g}_m$  are smaller than unity). A logarithmic 'quantum level diagram' helps to check whether condition (10) is satisfied [6]. If it is, then the first conversion stage dominates the fluctuation  $\text{var } s$ ; consequently, its statistics requires particular attention. It is strongly determined by the interaction process of the primaries with a solid, and frequently co-determined by a compromise between high gain  $\bar{g}_1$  and resolution. If the first stage yields  $g_1 = 1$  with a probability  $p_1(1) = \eta < 1$ , and a significant fraction  $1 - \eta$  of impinging quanta yields  $g_1 = 0$ , that is,  $p_1(0) = 1 - \eta$ , then we have a binary distribution with  $\bar{g}_1 = \eta$ ,  $\text{var } g_1 = \eta(1 - \eta)$ , and

$$\text{DQE} = \eta \left( 1 + \frac{\text{var } s}{\bar{s}^2} + \frac{\text{var noise}}{\eta \bar{N} \bar{s}^2} \right)^{-1} \quad (11)$$

where  $\bar{s}$  and  $\text{var } s$  refer to the combined yields of all subsequent stages  $m > 1$ . Thus,  $\eta$  appears as an upper limit of the DQE which cannot be exceeded but only reduced by the variances of the yield  $s$  and the noise.

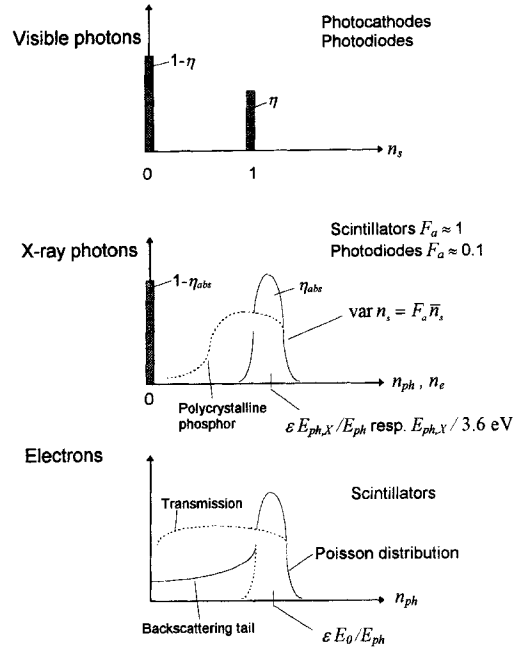
A converting chain is much more meaningfully characterized by the DQE than by a sensitivity factor defined as the ratio

$\bar{S}_{\text{out}}/\bar{S}_{\text{in}}$  of the means of the output and input, as frequently used, since a degradation of the SNR due to any misdesigned stage violating condition (10) cannot be repaired by postamplification.

Occasionally, a distinct improvement of the DQE can be achieved by a normalization procedure as used in the conventional single-channel radiation-measuring technique. Using a discriminator, the weak and strong noise signals are suppressed and the passing signal pulses are replaced by counting pulses to be stored spatially related into a digital memory. For this method Eq. (11) simplifies to  $\text{DQE} = \eta$ , where  $\eta$  denotes the probability of producing a countable pulse by the primary. This counting technique, introduced by astrophysicists [7], and applied to the three types of primary quanta under discussion, requires some electronic effort in order to determine the 'center of gravity' of the signal spots within the sampling structure and to avoid spatial and temporal multiple counting. Since the signal pulses must be processed individually before final storage, a high processing speed is required. Even then, the application is restricted to very low pulse rates, particularly at higher pixel numbers. Indeed, the counting technique cannot remove a reduction of the DQE by signal-independent statistically occurring disturbing pulses of comparable order of magnitude (e.g., ion spots of intensifiers) [6].

### 1.2.3 Quantum Efficiency of Conversion Processes

The optimum design of image converter chains in accordance with the aforementioned rules requires some knowledge



**Figure 4.** Types of pulse height distributions for the first converting stage using different forms of primary radiation shown schematically.  $F_a$ , Fano factor;  $n_s$ , number of secondaries.

about the quantum efficiency of the conversion processes employed in the optoelectronic components. In particular, the interaction of the primaries with the first converting medium must be discussed.

The ultimate limit of the DQE is set by the pulse height distribution (PHD)  $p_1(g_1)$  of the first stage, depending on the interaction of the different types of radiation quanta based on conversion processes to be discussed subsequently. Figure 4 provides a survey of the most important cases. Photons in the visible spectral range release only single photoelectrons with an efficiency  $\bar{g}_1 := \eta$ , resulting in a purely binary distribution, a model valid with photocathodes (see Sec. 1.2.3.2 of this Chapter) as well as with solid state sensors and photographic recorders (see



Sec. 1.2.3.1 of this Chapter). The PHD  $p_1(g_1)$  of signals released by weak X-ray photons contains several contributions:

- (a) A binary distribution of the photon energy deposits described by the absorption coefficient  $\alpha$  and the thickness  $d_0$  of the converting medium (see Sec. 1.5).
- (b) A distribution of the number  $n_s$  of secondaries with  $\text{var } n_s = F_a \bar{n}_s$ , where the Fano factor  $F_a \leq 1$  [8] describes a reduction of  $\text{var } n_s$  as compared to the Poisson distribution; for solid state detectors  $F_a \approx 0.1$  while in scintillators  $F_a \approx 1$ .
- (c) A reduction of the collected signals depending on the depth of the absorption events, which is typical of attenuating scintillators such as polycrystalline phosphors (see Sec. 1.2.3.3 of this Chapter).

The detection of electrons differs from that of photons with respect to the high proportion of partial energy deposits of both back-scattered and transmitted electrons which extends the PHD by a tail to lower signals, in addition to the contributions (b) and (c) which also occur in the detection of X-rays.

Different conditions exist also with respect to the spatial resolution. In a properly designed converter chain, the shape of the PSF may be dominated by the interaction range of the primaries as well as by the dissipation of the secondaries and the procedure of their collection. In photon detectors (see Secs. 1.3 and 1.5 of this Chapter), the PSF is formed mainly by the delocalized processing of the released photoelectrons (in the solid or by an electron-optical imaging system), while in converters for fast electrons using a scintillator

(see Sec. 1.4 of this Chapter) the essential contributions are caused by the electron spread in the solid and by both scattering and propagation of photons in the light-optical collection system.

The main characteristics of the most important effects will be briefly presented, with reference to the literature for greater detail.

### 1.2.3.1 Photographic Recording

Silver halide emulsions are still used as recorders with a high pixel number and a fairly good resolution determined by grain size and emulsion thickness. The fundamentals have been analyzed by, for example, Dainty and Shaw [2] and Zweig [9]. Basically, a nonlinear response occurs due to the limited number of grains each of which can detect only once. Thus, with progressing exposure the detection probability decreases. While visible photons are detected with an efficiency of the order of a few percent, energetic electrons and X-ray photons are able to hit several grains, which results in a satisfactory DQE but also in a limited dynamic range. Nonlinearity and 'fog', both dependent on the developing conditions, make the photoplate using a digitizing densitometer less suited for quantitative evaluation purposes.

### 1.2.3.2 Photoeffect

The release of electrons from photocathodes is widely used in photomultipliers and low-light-level image pick-up tubes [10, 11]. Numerous types of semitransparent transmission photocathodes are

available, optimized for certain wavelength ranges of the visible or near-infrared/ultraviolet spectrum and standardized using the abbreviations S1, ..., S25, the relevant data of which can be found in any brochure on multipliers, image intensifiers, or camera tubes. The advantage of photocathodes is that photoelectrons emitted into the vacuum can be accelerated in order to give significant signals of high SNR by every individual electron after subsequent conversion processes.

Using emissive material compositions, a strong photon absorption and long escape depth are realized as well as a low work function in order to extend the long-wavelength limit into the red spectrum as far as possible. At the other end of the spectrum the ultraviolet transparency of the window is an important secondary condition. The most popular photocathode, S20, is a multialkali layer (Sb-K-Na-Cs), the spectral response of which is closely matched to the sensitivity curve of the eye. At the emission wavelengths  $\lambda$  of some scintillators around 550 nm, corresponding to a photon energy  $E_{ph} = 2.25$  eV, these photocathodes have a sensitivity of  $S_{pc} \approx 40 \text{ mA W}^{-1}$ , resulting in a quantum efficiency of

$$\eta_{pc} = \frac{S_{pc} E_{ph}}{e} = \frac{S_{pc} (hc/e)}{\lambda} \approx 0.1 \quad (12)$$

The peak sensitivity at 400 nm allows  $\eta_{pc} \approx 0.25$ . Special photocathodes combined with suitable window materials are available, emphasizing the ultraviolet or infrared range. The most advanced negative electron affinity cathodes using GaAs(Cs) are sensitive of up to  $\lambda \approx 930$  nm.

Since in the visible region of the electromagnetic spectrum only single electrons are released with a probability  $\eta_{pc} < 1$ ,

Eq. (11) yields a  $DQE \leq \eta_{pc}$ . Thus, a quantitative photon-counting device cannot be realized. The assumption of single electron emission is no more justified if the photon energy is increased up to the weak X-ray range when only that small part  $\eta$  of the impinging photons is utilized which releases photoelectrons within a small escape depth at the exit surface. Some of these may have sufficient energy to release several secondary electrons simultaneously. Hence the resulting pulse height distribution of the emitted electrons is not restricted to  $g_1 = 0$  and  $g_1 = 1$  but it is characterized by several peaks due to multiple events causing a further increase of the variance and thus a decrease of the DQE (see Sec. 1.5.3 of this Chapter).

The intrinsic photoeffect in semiconductors is applied in numerous television (TV) pick-up tubes of the vidicon type as well as in the modern charge-coupled devices (CCDs) (see Secs. 1.2.4.1 and 1.2.4.5 of this Chapter). Due to their low band gap, a remarkably higher yield in the formation of electron-hole pairs even in the infrared range may be obtained, which, however, can only be utilized if the noise term in Eq. (11) is suppressed by low read-out noise and/or a high storage level. The scanning electron beam read-out is not very favorable in this respect and, in addition, the storage capability of the converting layers is rather limited by recombination, but it can be very much increased by cooling. Some preferred materials are  $\text{Sb}_2\text{S}_3$ , Se, PbO, CdS, and CdSe. The most advanced technique may be the silicon photodiode employed as a mosaic target in vidicons (see Sec. 1.2.4.4) and in the sensor elements of CCDs (see Sec. 1.2.4.5 of this Chapter).

Even weak X-ray photons can be efficiently detected by collecting their photoelectrons. An impinging photon with energy  $E_{\text{ph,X}}$  releases  $\bar{n}_e = E_{\text{ph,X}}/E_f$  electron-hole pairs with  $\text{var } \bar{n}_e = F_a \bar{n}_e$ , where  $E_f$  is the mean formation energy and  $F_a$  is the Fano factor. Thus, for silicon, the most frequently applied semiconductor material, with  $E_f = 3.6 \text{ eV}$  and  $F_a \approx 0.12$  at  $E_{\text{ph,X}} = 1 \text{ keV}$  a yield of 270 with a relative standard deviation of 2% can be obtained. Obviously, that fact, utilized long since in energy dispersive X-ray spectroscopy (EDX), is also very valuable in achieving a high DQE in image detectors (see Sec. 1.5.7 of this Chapter).

### 1.2.3.3 Scintillators

Luminescence in solids [12] has proved to be an efficient conversion process of energetic radiation (electrons and X-rays) into photons to be transferred to a light-sensitive sensor either by direct close contact or by an imaging optical element. In fact, even with X-radiation the luminescence effect in phosphors is always based on cathodo-luminescence due to the generated photoelectrons and their secondary electrons. The yield of produced photons

$$n_{\text{ph}} = \epsilon \frac{\Delta E}{E_{\text{ph}}} \quad (13)$$

is to a rough approximation determined by the energy deposit  $\Delta E$  of the primary quantum, the mean energy  $E_{\text{ph}}$  of emitted photons, and an energy efficiency coefficient  $\epsilon$ ; further, only a fraction of the released photons can be collected due to self-absorption and the limited solid angle of acceptance of the optical element (see Sec. 1.2.3.4 of this Chapter). The energy

deposit can have a broad variance for reasons which are partially somewhat different for electrons and X-rays (Fig. 4). Further, the trade-off between a high signal output on the one hand and resolution on the other must frequently be taken into account.

The retardation of primary electrons in the solid is governed by elastic and inelastic multiple scattering. A not negligible fraction of the electrons, increasing with atomic number, leaves the surface by back scattering with a wide energy distribution, after having lost a part of their initial energy by conversion processes in the solid. In addition, if a high resolution must be ensured by using scintillators that are thin compared to the penetration range of the primary electrons, the same is true for the transmitted electrons. By Monte Carlo simulation [13–15] the relevant quantities  $\text{var } \Delta E$ ,  $\overline{\Delta E}$ , and the mean spatial density of the energy deposit within the interaction volume can readily be determined; thus, the trade-off between DQE and PSF can be balanced, as is inevitable, particularly at high electron energies.

Soft X-ray photons are converted by photoabsorption according to the absorption law with negligible scattering (see Sec. 1.5 of this Chapter). Hence the photon loses its energy  $E_{\text{ph,X}}$  completely by an absorption event. Consequently, the distribution of  $\Delta E \equiv E_{\text{ph,X}}$  may be much narrower, a favorable prerequisite for a high DQE. However, a more detailed consideration leads to other reasons for fluctuations of the photon yield: escape of electrons through the surfaces if the thickness becomes comparable to the range of secondary electrons, and radiation-less transitions at impurities and defects; even

the reflectivity of the surface may need to be taken into account [16].

Scintillators are commercially available as powder phosphors with grain sizes of the order of 1–10  $\mu\text{m}$ , to be deposited using a binding agent on a transparent substrate and covered by a conducting aluminum mirror layer.

There are several procedures for covering a substrate: settling from suspension, cataphoretic deposition, and vapor deposition allowing a column-like orientation of CsJ crystallites with a light-guiding capability (for references, see Gruner et al. [17]). The different phosphor substances are usually characterized by a P-number [18]. If a counting technique (see Sec. 1.2.2 of this Chapter) is used, fast phosphors (e.g., P16 or P37) may be mandatory. Efficiency measurements for various phosphors with soft X-radiation are presented by Husk and Schnatterly [16] and Chappell and Murray [19].

Single crystals may be used if a restricted interaction range allows high resolution, if any fixed pattern due to the grain structure must be avoided and a narrow pulse height distribution has priority. Due to the generally relatively high refractive index  $n_0$  of the scintillator material, however, the acceptance angle is strongly reduced by total reflection at the exit face (see Sec. 1.2.3.4 of this Chapter), while an essential part of the light propagates sideways by multiple reflections. Polycrystalline phosphors, on the other hand, offer stronger signals, however, with a higher relative variance due to the inhomogeneous intrinsic structure adding its own statistics by fluctuating interaction paths. While the light output from monocrystals (or structureless scintillator plates) has a defined total reflection limit at the

emission angle  $\beta_{\max} = \sin^{-1}(1/n_0)$  which is independent of the spatial distribution of the energy deposit, the output of polycrystalline phosphors is attenuated by multiple scattering and absorption, which leads to a dependence of the response on both the thickness of the screen and the localization of the primary interaction. The resulting broadening of the signal PHD impairs the DQE according to Eq. (6). Since energy deposit and photon penetration depend in an opposite sense on the thickness, the total photon output can be maximized using an optimum thickness depending on the primary energy [16].

#### 1.2.3.4 Light Optical Elements

Only a fraction  $\eta_L$  of the whole photon output  $\bar{n}_{\text{ph}}$  in the scintillator over the solid angle  $4\pi$  as expressed by Eq. (13) can be utilized due to the limited solid angle  $\Omega$  of acceptance, determined by the numerical aperture NA corresponding to a collecting angle  $\beta = \sin^{-1}(\text{NA}/n_0)$ , and the transparency  $T$  of the chosen light optics. Thus, the light transfer element must be seen as a quantum converter stage contributing to the signal statistics by a binary distribution with a gain  $\bar{g}_L := \eta_L < 1$  and  $\text{var } g_L = \eta_L(1 - \eta_L)$  to be introduced into Eq. (9). Generally, assuming a transparent scintillator of refractive index  $n_0$  covered by a mirror layer of reflectivity  $R$ , the relation

$$\begin{aligned} n_L &= (1 + R)T \frac{\Omega}{4\pi} = (1 + R)T \sin^2 \frac{\beta}{2} \\ &= \frac{1 + R}{2} T \left[ 1 - \sqrt{1 - \left( \frac{\text{NA}}{n_0} \right)^2} \right] \\ &\approx \frac{1 + R}{4} T \left( \frac{\text{NA}}{n_0} \right)^2 \end{aligned} \quad (14)$$

shows the unfavorable consequences of a high refractive index  $n_0$  as mentioned above. The available NA of the optics is mostly co-determined by the scale factor  $M$  required for a proper matching of the resolutions.

Fiber-optic plates are composed of well-ordered  $6\mu\text{m}$  fibers consisting of a core and a coat with refractive indices  $n_1$  and  $n_2$ , respectively, which transfer the photon output of a scintillating cover layer by total reflection onto a subsequent sensor layer contacted to the plate. Large plate diameters above  $50\text{ mm}$  and numerical apertures  $\text{NA} = \sqrt{n_1^2 - n_2^2}$  of 0.66 or even 1.0 are available, resulting with  $n_0 = 1.83$  (yttrium aluminum garnet, YAG) [20] in efficiencies  $\eta_L$  of 0.047 or 0.11, respectively. The disadvantage is the limitation of the resolution and the fixed scale factor of 1:1 (although tapers are also available) [21]. The transparency of the single fiber plate is limited to  $T \approx 0.7$  due to the spatial filling factor; a sandwich of two plates, as occasionally used in coupling image intensifiers, reduces  $\eta_L$  by  $T^2$  since an exact alignment of the fibers cannot be achieved. In this case, as well as with other periodic structures (e.g., CCD sensors), some Moiré effects may be observed, which can be removed together with other contributions to this 'fixed pattern' (scintillator inhomogeneity, 'chicken wire') by image processing. Most intensifier and camera tubes as well as CCDs with a larger pixel size are available with integrated fiber plates.

Lens optics allow a wide range of matching requirements by the proper choice of distances and focal lengths. The acceptance angle on the object side is in practice not completely independent of the scaling factor and the field diameter to be

transferred. Some special cases may be distinguished, discussing the optics as a pair of two single lens systems both corrected for infinity but with different focal lengths  $f_1$  and  $f_2$  in order to realize a desired scaling factor  $M = f_2/f_1$ :

(1) A standard light microscope objective allows a high  $\text{NA} \approx 1$  with a resolution at the Abbe limit. Due to a relatively small focal length  $f_1$  ( $\approx 5\text{ mm}$ ) and the restriction of the image angle to about  $\beta_1 \approx 4.5^\circ$ , the object field is limited to  $2r_0 = 2f_1\beta_1 \approx 0.8\text{ mm}$ . In order to transfer a pixel number typical for advanced image sensors, this object field limitation can only be tolerated if the high resolution of the objective can be utilized entirely. The choice of  $f_2$  ( $\gg f_1$ ) is then determined by the pixel size of the subsequent converter stage. The design is similar or even identical to that of a microscope with a tube length  $f_2$ .

(2) Extending the field while maintaining the NA requires a larger focal length  $f_1$  with simultaneously increased pupil diameter, but this may not be feasible with standard light microscopic components. For this purpose a 'tandem' pair of two lenses (Fig. 5) with large apertures developed for transferring the output of X-ray intensifiers to TV pick-up tubes may be preferred [22]. The acceptance angle of such objectives is mostly characterized by their  $F$ -number. Introducing the relation  $\text{NA} = (4F^2 + 1)^{-1/2}$  into Eq. (14) yields

$$\eta_L \approx (1 + R)T[4n_0^2(4F_1^2 + 1)]^{-1} \quad (15)$$

for a single crystal with a reflecting layer.

Using, as an example, a front lens with  $f_1 = 50\text{ mm}$ ,  $F_1 = 0.75$ , and  $T = 0.7$  an  $\eta_L = 0.032$  can be achieved as long as the whole system is magnifying ( $f_2 > f_1$ ). If,

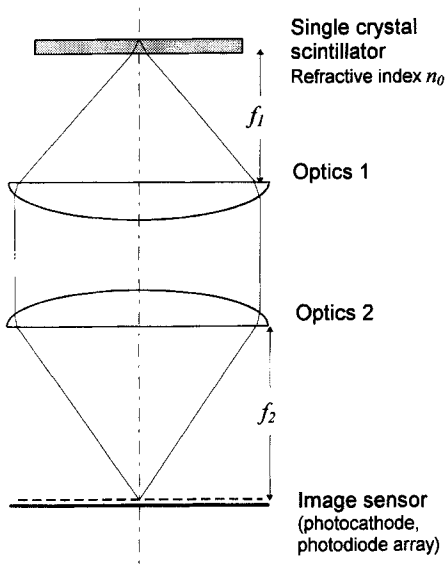


Figure 5. Schematic diagram of tandem lenses.

however, a demagnification  $M = f_2/f_1 = F_2/F_1 < 1$  is required, then  $F_2$  is limited to the same value, and we have

$$\eta_L \approx (1 + R)T \frac{M^2}{4n_0^2(4F_2^2 + M^2)} \quad (16)$$

Thus,  $\eta_L$  decreases with decreasing  $M$  since, due to the limitation of  $f_2$  and  $F_2$ , a demagnification can only be achieved by increasing  $f_1$  and thus reducing the acceptance angle. The loss of photons occurs particularly if the whole optics are replaced by a standard photo objective.

So far the scintillator has been treated as a homogeneous refractive medium which allows an easy formulation of the optical relations. Powder phosphors for which such a treatment is not applicable may be characterized by a refractive index between 1 and 1.5 [6] and a photon absorption factor depending on the thickness of the layer. In practice, the factor  $(1 + R)/n_0^2$

occurring in  $\eta_L$  may be determined empirically for each scintillator. Experimental experience with P20 has shown that below a mass thickness of about  $5 \text{ mg cm}^{-2}$ , absorption of photons within the scintillator can be neglected.

### 1.2.3.5 Secondary Emission

The emission of secondary electrons with energies below 50 eV is preferable for low-energy primary electrons. The quantitative response is well known from multipliers and scanning electron microscope instruments, and invaluable reviews are available [23, 24]. The energy dependence of the yield  $\delta_s$  typically shows a limited range with  $\delta_s > 1$  at a few kiloelectron volts, which is utilized for charge multiplication in photomultipliers and microchannel plates (MCPs) (see Sec. 1.2.4.3 of this Chapter). For fast primary electrons, secondary emission as the conversion process in the first stage is hardly practicable because of the adverse effect of its low yield on the DQE according to Eq. (11).

In almost all camera tubes, secondary emission is used to stabilize the potential of the target surface by the scanning electron beam releasing the video signal at the common electrode by recharging the areas discharged by the image (see Sec. 1.2.4.4 of this Chapter).

For secondary emission conduction (SEC) in transmission which was successfully utilized using thin insulating targets (KCl) with a porous structure of large internal surface area (see Sec. 1.2.4.4 of this Chapter), a yield above 10 at 7 keV was reported [25]. The advantage of such targets is their extremely long integration capability.

### 1.2.3.6 Electron Beam-Induced Conduction

Due to its high yield, the EBIC effect, that is, the formation of electron-hole pairs by electrons, is particularly favorable. An impinging electron with energy  $E_0 = 10 \text{ keV}$  releases  $\bar{n}_e \approx E_0/E_f = 2.7 \times 10^3$  electron-hole pairs in silicon ( $E_f = 3.6 \text{ eV}$ ), the most frequently applied semiconductor material.

### 1.2.3.7 Imaging Plate

The imaging plate (Fig. 6) was originally developed for diagnostic X-radiography [26,27], but it is applicable for electron radiation as well [28–31]. A storage phosphor, typically  $\text{BaFBr:Eu}^{2+}$  coated on a plastic sheet, traps the electron-hole pairs,

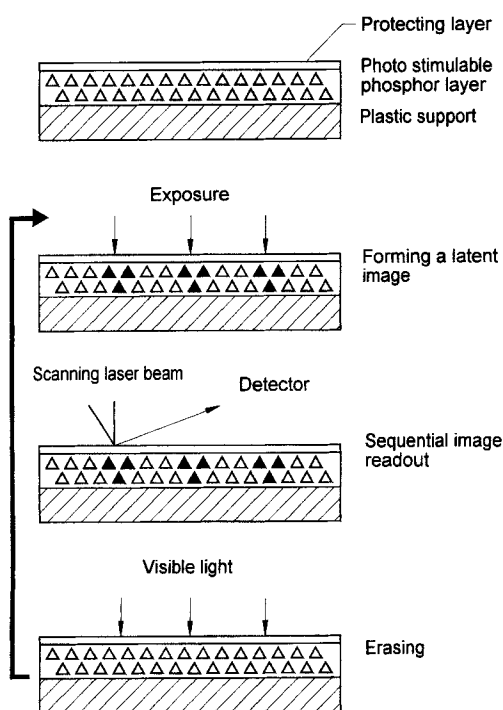


Figure 6. Working principles of the image plate.

released with high gain by the primaries, in F-centers with a high resolution determined by the range of the secondaries. This latent information is sequentially read out by a focused laser beam (630 nm) stimulating the recombination under the emission of photons (390 nm), which are detected using a photomultiplier. The merits of this device are the excellent linearity over a dynamic range of five decades and low intrinsic noise, offering the prerequisites for a high DQE. The resolution is mainly determined by scattering of the read-out beam in the layer, and depends on its thickness. The application to electrons and soft X-radiation has stimulated a reduction in the layer thickness, which was originally adapted to the absorption of hard X-rays. At present, a pixel size of  $25 \mu\text{m}$  is feasible. A critical problem is the construction of highly efficient collection optics which reduce the adverse effect of scattered laser intensity [31].

## 1.2.4 Composed Systems and Optoelectronic Components

Numerous devices employ one or more of the above-mentioned physical effects in order to convert one radiation into another, to increase the quantum level, and to offer a meaningful combination for obtaining output signals suited for storage as well as for analog or digital processing. The most important components applied to on-line image recording in the various microscopes are discussed briefly in the following sections. Further details may be found in the references (e.g. [31a]).

### 1.2.4.1 Scintillator–Photosensor Combination

Between the emission spectrum  $s(\lambda) := \Delta \bar{E}^{-1} dE/d\lambda$  of the scintillator and the spectral response  $\eta_{pc}(\lambda)$  of the following conversion stage (e.g., a photocathode or CCD) a good overlap is required by maximizing the coupling factor

$$G := (hc)^{-1} \int s(\lambda) \eta_{pc}(\lambda) \lambda d\lambda \quad (17)$$

using the data sheets of the producers (results for some combinations are given by Eberhardt [32]). Then the photoelectron number  $n_e = \Delta E \eta_1 G$  released by an energy deposit  $\Delta E$  can readily be calculated taking into account the results of Sec. 1.2.3.4 of this Chapter [nearly monochromatic emission allows the approximation  $G \approx \epsilon \eta_{pc}(E_{ph})/E_{ph}$ ]. As an example, for the most commonly applied standard combination P20/S20,  $G = 5.22 \text{ keV}^{-1}$ . For the light-optical coupling the following three options are available:

- (a) the direct contact to a photocathode, as used in X-ray intensifier tubes or sometimes with CCDs (see Sec. 1.5.4 of this Chapter), offers almost ideal collection efficiency;
- (b) the fiber plate, as well as micro-objectives of high NA, reduces to about 1 photoelectron  $\text{keV}^{-1}$ ;

- (c) tandem optics with  $F_1 = 1.6$  allow only 0.1 photoelectron  $\text{keV}^{-1}$  [6].

The refractive index of the scintillator limits the photoelectron output. Thus the performance of the scintillator–optics–sensor combination may be characterized by the expression  $\bar{n}_e/\Delta \bar{E} \approx (G/2n_0^2) \text{NA}^2$ . Table 1 gives some typical figures for a selection of scintillators coupled to both a photocathode and a CCD. Obviously, due to its higher quantum efficiency the CCD is superior, particularly with scintillators at the red end of the spectrum.

These figures allow, by Eq. (11), an estimate of the DQE depending on the energy deposit which can be achieved with different radiations.

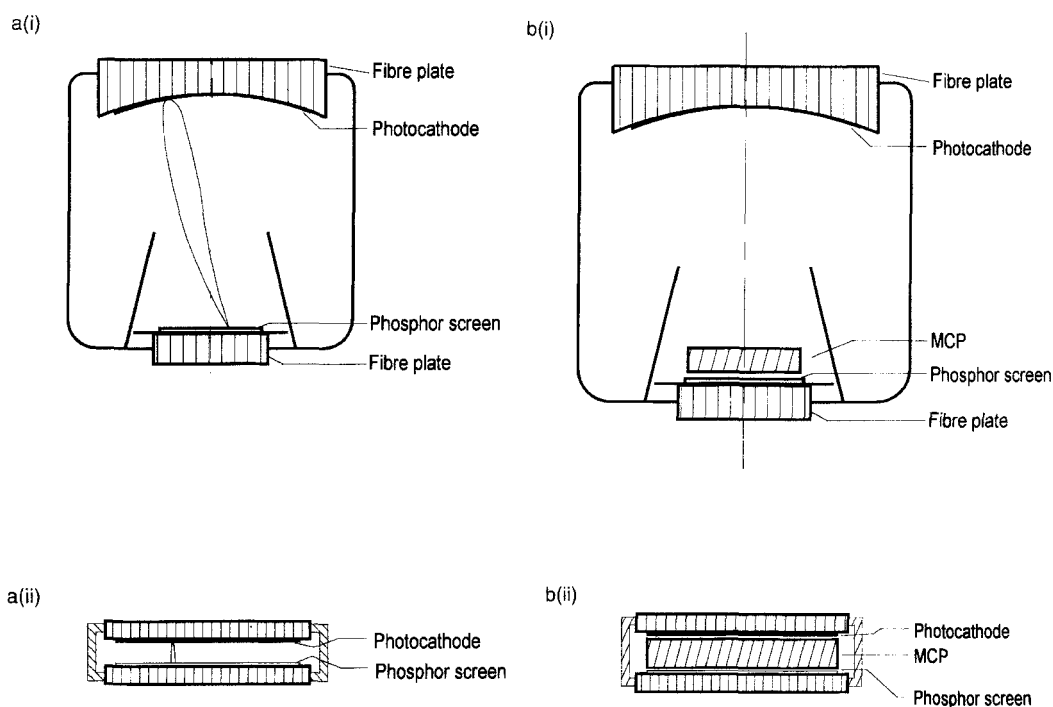
### 1.2.4.2 Image Intensifiers

First-generation image intensifier tubes are equipped with a transmission photocathode converting the input photon image into an electron distribution, an electrostatic electrode system accelerating the electrons by 10–20 kV onto a metalized output phosphor screen as the anode, where an amplified photon image is produced. The input and output faces are generally equipped with fiber plates, allowing the stacking of various tubes. Three designs are commercially available, which

**Table 1.** Data of scintillator–sensor combinations. Photoelectron yield for  $\text{NA} = 0.6$ ,  $T = 0.7$ , and  $R = 1$ .

Scintillator	YAG (Ce)	CsJ (Tl)	NaJ (Tl)	CaF (Eu)	NE 102A	P20 powder
Energy efficiency, $\epsilon$ (%)	5	11.9	11.3	6.7	3.0	20
Mean photon energy, $E_{ph}$ (eV)	2.21	2.29	2.97	2.83	2.92	2.2
Refractive index, $n_0$	1.83	1.80	1.85	1.44	1.58	$\approx 1.2$
Photoelectron yield, $\bar{n}_e/\Delta E$ ( $\text{keV}^{-1}$ )						
with the S20 photocathode	0.074	0.02	0.26	0.26	0.094	0.75
with the TEK1024 CCD	0.35	0.73	0.29	0.30	0.11	3.41





**Figure 7.** Examples of image intensifiers: (a) first generation and (b) second generation. (i) Electrostatically focused. (ii) Proximatively focused.

differ with respect to the focusing method (Fig. 7):

- (a) magnetic focusing by embedding the tube in a longitudinal magnetic field produced by a solenoid coil;
- (b) electrostatic focusing using a curved cathode and an electrical immersion lens within the tube (Fig. 7a(i));
- (c) 'proximity' focusing by keeping the distance between the plane cathode and the anode as short as possible (Fig. 7a(ii)).

While (a) offers good resolution but with the drawback of large size, (b) has the advantages of more compactness, of variable gain controlled by the voltage, and of the availability of reducing systems, but some image distortion cannot be avoided.

The Proxifier (c) is free from distortion, extremely compact, and can be built with a relatively large field diameter.

The gain of first-generation intensifiers can be estimated as the product of the yields of the cathode and the phosphor as

$$g \approx \eta_{pc}(\lambda) \epsilon e U / E_{ph} \quad (18)$$

which gives  $g \approx 50$  at  $\lambda = 500$  nm if the standard S20/P20 combination and an acceleration voltage  $U = 20$  kV are used. Practically every photoelectron releases a significant pulse of  $\epsilon e U / E_{ph}$  photons at the output. Provided proper processing of these signals, a photon DQE  $\approx \eta_{pc}$  can be assumed. Intensifiers are offered with different photocathodes, allowing an adaptation of the spectral sensitivity  $\eta_{pc}(\lambda)$  to the

incoming spectrum as well as with different phosphors, for example with a low decay time if photon counting is required.

Second-generation intensifiers (Fig. 7b) are equipped with an MCP (see Sec. 1.2.4.3 of this Chapter) in front of the phosphor screen, the high gain of which guarantees a mean brightness level even for a very weak input. Due to their wide PHD, however, these intensifiers are only recommended for use after some preamplification or if the output pulses are electronically normalized as in single-pulse counting devices.

The resolution of the intensifiers is rather limited by the electron spread in the screen, the electron optics, the fiber plates, and, in second-generation devices, the MCP. Since the image field is limited, the transferable pixel number, which may be good enough for a standard TV technique is not sufficient for high pixel recording using advanced scientific grade CCDs (see Sec. 1.2.4.5 of this Chapter), except for Proxifiers, which are available with relatively large diameters.

#### 1.2.4.3 Microchannel Plates

The MCP (for a review, see Lampton [33]) is a regular array of micro-tubes of internal diameter  $12\text{ }\mu\text{m}$  and a length below 1 mm, the inner walls of which are covered with a semiconducting material of high secondary emission yield (Fig. 8). By a longitudinal field created by a voltage of about 1 kV across the plate, an electron released at the entrance is accelerated and multiplied by  $10^4$  due to repeated secondary emission. A further plate may multiply the yield once more. The high electron output can be accelerated and proximity focused at 5 keV onto a transmission phosphor

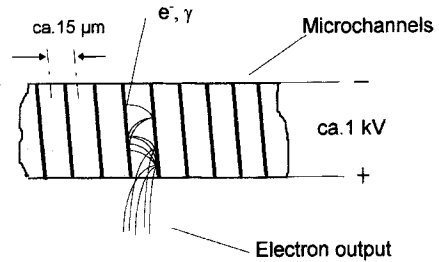


Figure 8. Principle of the MCP.

screen. The advantage of an extremely high gain is, unfortunately, accompanied by a high signal variance due to an open area ratio of 0.55, a low electron yield at the entrance for highly energetic quanta, and the fluctuation of the internal gain, which impairs the DQE according to Eq. (11) (see Sec. 1.2.2 of this Chapter) [34]. By coating the entrance faces with secondary emissive material, some improvement is possible. Moreover, the output response of high gain MCPs suffers from saturation. In pulse counting applications, this saturation effect is utilized to improve the PHD [35]; it prohibits, however, the conversion of flash images.

#### 1.2.4.4 Television Camera Tubes

The TV technique is widely applied to convert a two-dimensional photon input into a sequential video signal in order to transfer, process, observe, and record the microscopic images on-line. Due to the TV standard, the number of pixels is limited to the order of  $512^2$ , although the advanced high definition TV (HDTV) technique may offer some progress. Numerous types of camera tubes are available, designed for high photon sensitivity in different spectral ranges and low noise, the details of which may be found in the references [31a, 36, 37]

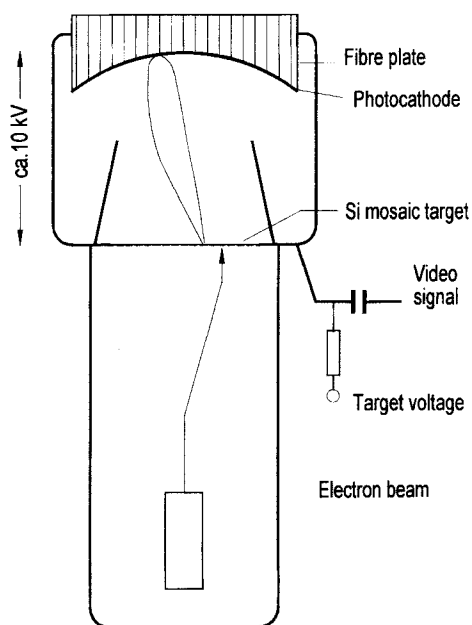
and, particularly, in the manufacturers' handbooks.

The basic element in a camera tube is the 'target' on which the image intensity is temporarily stored as a charge density distribution to be 'read out' periodically and linewise by a scanning electron beam of low energy which, by secondary emission, recharges the backface to cathode potential and thus induces the video signal in a common electrode for capacitive coupling out. Low noise of the subsequent video amplifier is one important prerequisite for camera performance.

The Vidicon types of tube are equipped with a semiconductor target at the entrance window, utilizing the intrinsic photoeffect for the discharge of the equilibrium potential. The selection of the material as well as its processing have been improving for many years, resulting in various special types. The variables of note for the targets are the quantum yield and the spectral response, the lateral conductivity and the storage capability depending on the dark current, the lag time, the maximum signal current  $I_s$ , and, further, the degree of avoidance of defects. The Newvicon and Chalnicon tubes, using ZnSe/ZnCdTe and CdSe targets, respectively, show a superior quantum yield over a wide spectral range, while the Saticon (SeAsTe) and Plumbicon (PbO) tubes have lower lag times at a narrower spectral response [36]. A mosaic target of silicon photodiodes constructed by advanced microlithography is an efficient way of avoiding lateral discharging. Although the quantum efficiency of the stored charge image is almost ideal, the read-out noise prevents the sensitivity from being sufficient for single-photon detection. The read-out noise from the amplifier, electron

beam, and, in some cases, the target is of the order of nanoamps. Thus, for good images a relatively high exposure rate  $n/\tau \approx I_s/eA$  of  $>10^{10}$  photons  $\text{cm}^{-2} \text{s}^{-1}$  is required ( $A$  is the target area of a 1 inch (2.5 cm) vidicon and  $\tau$  is the frame time). Under such conditions, an excellent image quality over a dynamic ratio of some 100:1 can easily be obtained during the TV frame time.

The goal of increasing the sensitivity up to the photon noise limit requires some preamplification of about  $10^5$ . In principle, this can be achieved using one or more intensifiers (see Sec. 1.2.4.2 of this Chapter), preferably under fiber plate coupling. More favorable, however, may be the use of a 'low-light-level' (LLL) camera tube, which are commercially available under several trade names (Fig. 9). Such tubes integrate the intensifier principle into a



**Figure 9.** Schematic diagram of the 'silicon intensifier target' camera tube.

silicon target camera tube. The photoelectrons are accelerated to 10 keV and electrostatically focused onto a thinned silicon wafer having an array of diodes on the back which collects the electrons released by the EBIC effect with a yield of  $2.7 \times 10^3$  (see Sec. 1.2.3.7 of this Chapter). The scanning electron beam has the same function as in vidicon tubes. These 'silicon intensifier target' (SIT) tubes [38] allow the pick-up of moonlit scenes; detecting single photoelectrons, however, requires an additional intensifier, generally coupled by a fiber plate, a compound tube frequently called an intensified SIT (ISIT).

SEC tubes [25] use a KCl target (see Sec. 1.2.3.6 of this Chapter), but are otherwise constructed like SIT tubes. Their sensitivity and dynamic range are lower; additionally, their rigidity against over-illumination is unsatisfactory. Their main merit, the storage capability, is also offered by the slow-scan CCDs, which have a far superior dynamic range.

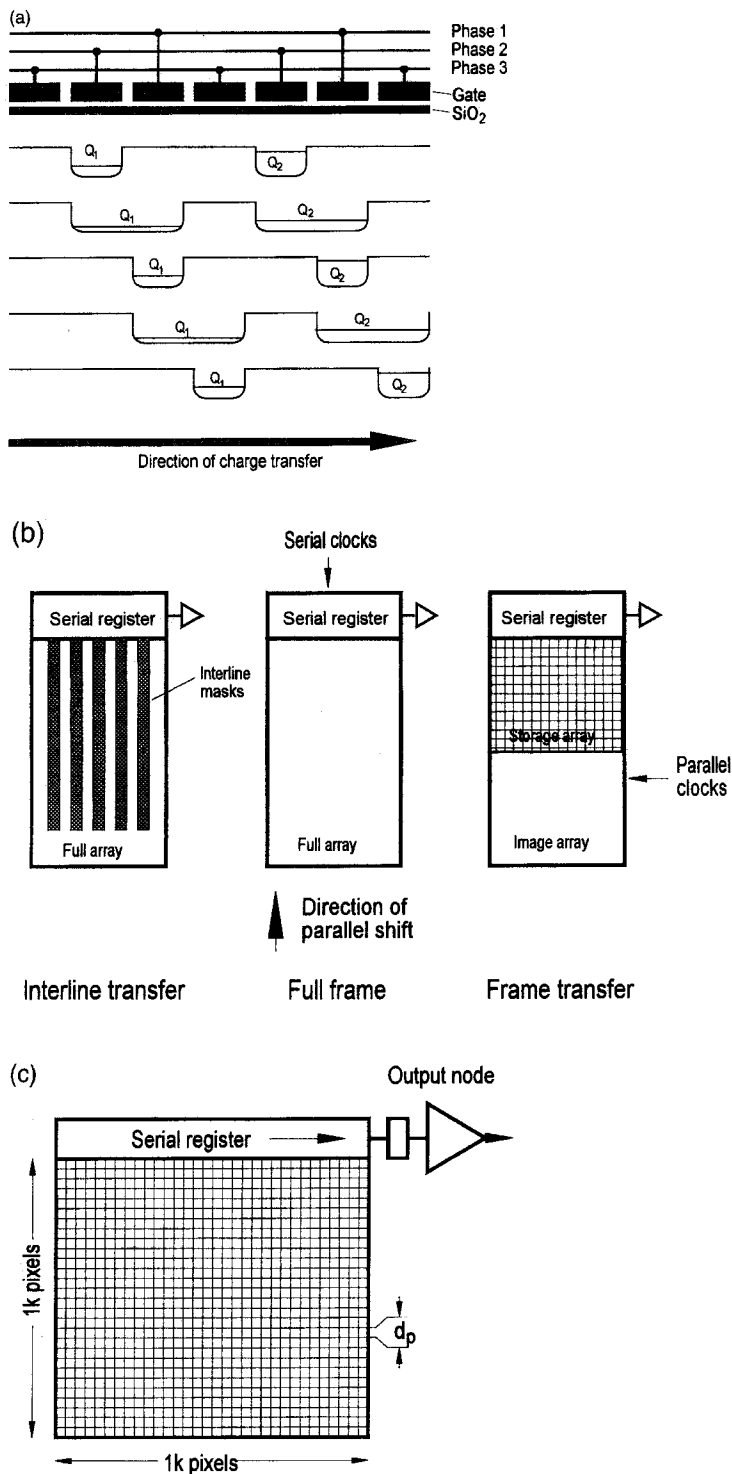
#### 1.2.4.5 Charge-Coupled Devices

Solid state sensors offer the most promising way for a reliable and accurate image conversion. Their enormous recent progress and continuous further improvement gives them promise as the final solution of many actual and future on-line processing tasks in microscopy. They are suitable for the recording of series as well as single images without some of the restrictions of conventional recorders. The incoming photons are accumulated on an array of silicon photodiodes to be read out once or periodically using microelectronic circuitry integrated on the same chip. In earlier designs the diodes were sequentially con-

nected by transistor switches to a common video line. A number of essential advantages, however, are offered by the CCD, in which, by MOS control, potential wells filled with electrons can be moved at high speed and with only negligible loss of charge [39–42]. By such an analogous shift register (Fig. 10a) the image information is transported in a suitable sequence to one or more common low-noise pre-amplifiers. Basically two read-out procedures are used (Fig. 10b):

- (a) The interline transfer CCD, in which to every photodiode line a masked transport register line is attached which accepts the charge in parallel for a sequential read-out at a standard video rate.
- (b) The full-frame CCD (Fig. 10c), in which the photodiodes form the transport register and are, after an exposure interval, sequentially read out with a speed adapted to the bandwidth of the noise-optimized video amplifier. The illumination must be blanked during the read-out phase unless a second interstorage area of equal format is provided (frame transfer CCD).

Procedure (a) is suited for motion picture recording at TV frequency, not only for common video cameras, but, after rapid development to higher pixel numbers, even for advanced HDTV. Motion recording in microscopy may profit from this technique. Procedure (b) is used in scientific grade CCDs, in which, by cooling, a drastic decrease of the dark current, resulting in extreme integration times, and by a slow-scan read-out, a concomitant reduction in amplifier noise is realized. These devices are highly promising in



**Figure 10.** Charge-coupled device. (a) Principle of charge transfer in a three-phase CCD. (b) Read-out architectures. (c) Typical format of a full-frame slow-scan CCD (e.g., Table 2).

**Table 2.** Performance figures of the CCD Tektronix TEK 1024.

Pixel size, $d_{\text{CCD}} \times d_{\text{CCD}}$	$24 \times 24 \mu\text{m}^2$
Pixel number	$1k \times 1k$ pixels
Full-well capacity, $n_{\text{max}}$	$1.5 \times 10^5 \text{ e}^- \text{ pixel}^{-1}$
Read-out noise, $n_r$	
2 MHz	$< 22 \text{ e}^- \text{ pixel}^{-1}$
1.5 kHz	$< 10 \text{ e}^- \text{ pixel}^{-1}$
Dark current ( $-30^\circ\text{C}$ ), $D$	$< 6 \text{ e}^- \text{ s}^{-1} \text{ pixel}^{-1}$
Quantum efficiency (560 nm), $\eta_{\text{CCD}}$	40%
Read-out speed	
Fast	$2 \times 10^6 \text{ pixels s}^{-1}$
Slow	$1.5 \times 10^3 \text{ pixels s}^{-1}$

approaching the ideal recorder for single images.

Slow-scan CCDs (SSCCDs) have numerous merits (Table 2) [40]: a full-well capacity  $n_{\text{max}}$  above  $10^5$  electrons per pixel and a read-out noise  $n_r$  below 10 electrons per pixel allow an extreme dynamic range not available in other analog recorders. Negligible distortion, photometric linearity, an ideal filling factor, and independence of the signal rate up to ultra-short flash exposures are further merits. The dark current  $D$  allows satisfactory operation of video cameras at room temperature using standard frame times; in slow-scan devices, however, due to the strong temperature dependence according to  $D \propto \exp[-(E_a/kT)]$  with  $E_a \approx 1.05 \text{ eV}$  [41], cooling is advisable. For integration on the chip over some minutes, Peltier cooling to about  $-50^\circ\text{C}$  is sufficient, and by liquid nitrogen cooling a prolongation of up to hours (as required, for example, in astrophysics) is feasible. The quantum efficiency  $\eta_{\text{CCD}}(\lambda)$  for electron-hole pair production by photons emphasizes the red spectral range ( $>0.6$  at 700 nm), but it can be extended to ultraviolet wavelengths by a phosphor coating. Recently, however, back-illuminated thinned CCDs have become available where by passivation

and coating techniques  $\eta_{\text{CCD}} \approx 0.8$  over a wide spectral range has been achieved [41]. These will even withstand illumination with low-energy X-rays or electrons where front illumination is inapplicable due to the undesired absorption by protection and electrode layers, as well as the radiation sensitivity.

The pixel sizes vary between 8 and  $27 \mu\text{m}$  with array sizes of  $10^5$  to above  $10^7$ . Large pixels offer a high full-well capacity and filling factor, and better matching to some optoelectronic components, avoiding, for example, a Moiré formation between fiber plates and sensor elements. The previously mentioned low noise figures can only be achieved using a bandwidth limitation for the amplifier and analog-digital converter (ADC), resulting in a slow-scan read-out using pixel frequencies of the order of some 100 kHz (although strong progress is being made by improving the amplifiers and by parallel read out of subareas). Such sensors can be operated as quantitative image digitizers with an accuracy justifying a 14-bit ADC. Since the uniformity of the pixel response as well as of other converting elements does not have this accuracy, a numerical correction of the fixed pattern by the subsequent processing

system is advisable, using as a reference both a dark current pattern for subtraction and a uniformly illuminated image ('flat field pattern') for division.

The DQE of a converter chain using the CCD as the final element can be deduced from Eq. (7). The CCD noise consists of two statistically independent contributions expressed in the data sheets as electron numbers referred to one pixel: (a) the read-out noise as a Gaussian distribution with the standard deviation  $n_r$ , and (b) the shot noise of the dark current, which may be Poisson distributed. By adding their variances over the whole measuring area one obtains

$$DQE = \left( 1 + \frac{\text{var } n_e}{\bar{n}_e^2} + \frac{n_r^2 + D\tau}{\bar{N}_p \bar{n}_e^2} \right)^{-1} \quad (19)$$

where the signal  $s$  in Eq. (7) has been replaced by the number of electrons  $n_e$ ;  $\tau$  is the exposure time and  $\bar{N}_p$  is the number of primaries on one pixel. Considering the different primary radiations and the converting mechanisms used in front of the CCD, this fundamental relation allows the estimation and optimization of the performance, as shown below. With  $(n_r^2 + D\tau)/\bar{n}_e^2 \ll 1$  a condition for single-quantum counting at low exposure is fulfilled. Already  $\bar{n}_e \approx 10$  may be sufficient, which results in a saturation limit  $N_{p,\max} = n_{\max}/\bar{n}_e \approx 10^4$  with a full-well capacity  $n_{\max} = 10^5$ . Such an extended dynamic range is hardly achievable using other recording devices.

The development of CCD technology is making good progress, as exemplified in the *Philips Journal of Research* [42] and by Blouke [43], where a comprehensive survey on the present designs and their manufacturers is available [44]. Future developments are following several

directions:

- an increase in the pixel numbers, exceeding the presently commercially offered  $4k \times 4k$  CCDs;
- an increase in the read-out speed by further improving the noise figures and new organization schemes using a high degree of parallel read out;
- a higher degree of on-chip integration of control circuitry, of active control of the pixels, and of random access capability;
- the investigation of high pixel interline transfer CCDs with improved collection efficiency by means of on-chip micro-lenses.

### 1.2.5 Resolution and Sampling

The DQE as defined in Eq. (1) depends on the measuring area  $A$ , unless  $A$  is assumed to be large compared to the dimensions of the PSF. This rather trivial fact may also be expressed in Fourier space by a spatial frequency  $u$ -dependent  $DQE(u) = DQE_0 \cdot \text{MTF}^2(u)$ , where MTF is the modulation transfer function [2]. A comprehensive theoretical treatment of this dependence taking into account internal scattering mechanisms is given in [45, 46]. The overall MTF of the converter chain is obtained by multiplying the MTFs of the individual stages, taking into account the scaling factors between them. By proper choice of these scaling factors the stage resolutions should be matched to ensure that the response is dominated only by the first interaction processes or a subsequent sampling structure, for example, a CCD. The size  $d_{\text{CCD}}$  of the CCD pixels relative to the width of the PSF, however, should be neither too small nor too large. If it is too

small, then the fine sampling of a wide PSF leads to a coverage of only a relatively small image area by the limited number of available CCD pixels. If it is too large, then the sensor structure is fully utilized but, due to the loss of shift invariance, the common optical transfer theory becomes inapplicable [47], and the occurrence of spatial frequencies above the Nyquist frequency  $(2d_{\text{CCD}})^{-1}$  may lead to artifacts in Fourier processing known as ‘aliasing’. Thus, in practice, a certain bandwidth restriction is advisable by adapting the width of the PSF to one, or a few, sensor pixels.

## 1.3 Light Microscopy

Microphotography uses standard microscopic equipment. However, there are some special demands suggesting electronic image acquisition and processing [48]:

- visualization and recording of dynamic processes using tape or optical disks;
- contrast enhancement in real time of extremely weak image details;
- recording fluorescence images at low light levels, including radiation beyond the visible spectrum requiring some integration time and high DQE converters;
- storing and analyzing the images using computer technology.

### 1.3.1 Video Recording

Video cameras equipped with various types of Vidicon tubes (see Sec. 1.2.4.4 of

this Chapter) or, recently, with CCDs (see Sec. 1.2.4.5 of this Chapter) are employed for real-time observation and image series registration [49]. Without any preamplifiers they work satisfactorily at higher light levels where enough electrons are accumulated on the target during the frame period to overcome the read-out noise discussed above [Eq. (7)]. Dedicated video microscopes are commercially available.

### 1.3.2 Low-Light-Level Detection

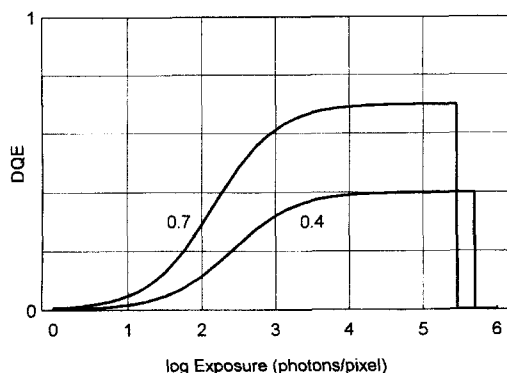
A considerable increase in sensitivity is obtained by SIT and ISIT tubes, such as for fluorescence microscopy [48, 50]; using MCP intensifiers even single photoelectrons at extremely low light levels may occasionally be counted. CCDs fiber-plate coupled to intensifiers are also used. Some solutions have been reviewed by Gursky and Fritz [51].

The merits of cooled SSCCDs (see Sec. 1.2.4.5 of this Chapter) are utilized for quantitative on-chip accumulation of single images with subsequent digital output [40, 52]. The DQE can be derived from Eqs. (11) and (19), separating the binary distribution and considering  $\bar{n}_e = \eta_{\text{CCD}}$ , with the result

$$\text{DQE} = \eta_{\text{CCD}} \left( 1 + \frac{n_r^2 + D\tau}{\bar{N}_p \eta_{\text{CCD}}} \right)^{-1} \quad (20)$$

Obviously, with a pixel exposure  $\bar{N}_p \gg (n_r^2 + D\tau)/\eta_{\text{CCD}}$  the DQE is determined by the efficiency  $\eta_{\text{CCD}}(\lambda)$ , which considerably exceeds that of photocathodes in the red spectral range and, using a fluorescent coating and/or back-illumination, in the blue and ultraviolet regions as well [41].





**Figure 11.** DQE versus exposure of a CCD for visible photons according to Eq. (20) ( $\eta_{\text{CCD}} = 0.4$  and  $0.7$ ,  $\eta_{\text{max}} = 1.5 \times 10^5 \text{ e}^-/\text{pixel}$ ,  $n_r = 10 \text{ e}^-/\text{pixel} \gg D_r$  assumed).

Consequently, the directly illuminated SSCCD allows a closer approach to the ideal detector than any photocathode device, except at low exposures (Fig. 11). Provided with digital storage and processing facilities, it may become the superior recorder in fluorescence microscopy. It has even proved successful as a sensor element in 'Nipkow disk confocal scanning microscopy' [53]. The only drawback, that is, the longer read-out time, may in time be overcome.

## 1.4 Electron Microscopy

The electron energy  $E_0$  in microscopy is sufficiently high to release significant signals within an interaction volume determined by multiple scattering. The distribution  $p(\Delta E)$  of energy deposits  $\Delta E$  consists mainly of a peak close to  $E_0$  with a steep slope on the high-energy side but a slow decay on the side of smaller pulse heights due to back-scattered electrons losing only part of their energy, and increases with the

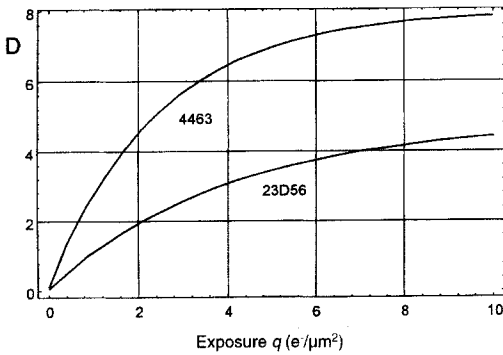
atomic number of the target material. With increasing energies  $E_0$  beyond 100 keV, however, the penetration range of the electrons becomes so long that for reasons of resolution the thickness of the converting layer must be restricted. Due to the decrease in stopping power with increasing energy the mean deposit of transmitted electrons shifts to  $\overline{\Delta E} < E_0$ , and an additional broadening of  $p(\Delta E)$  impairs the DQE. Thus, the optimization of the first converting layer with respect to resolution and the DQE is a complicated problem leading to very different results at different energies. Nevertheless, since in electron microscopes the magnification can generally be chosen within certain limits, an adaptation to a restricted resolution of the recorder is possible. Thus the image field, that is, the total number of pixels, is usually the more important parameter.

### 1.4.1 Photographic Recording

Direct photographic recording is the oldest and simplest read-out method (for a recent review, see Zeitler [54]). Commercial plates and films specialized for electron recording are characterized by the absence of a sensitizer and a larger layer thickness as compared to standard photoplates.

The secondary electrons released by one primary electron hit several silver halide grains. At increasing exposures the limited reserve of grains leads to an increasing nonlinearity and, finally, to saturation. The dependence of the optical density on the charge density exposure  $q = en$  can be described by

$$D = D_0 + D_s[1 - \exp(-c_s q)] \quad (21)$$



**Figure 12.** Density-exposure curve of Eq. (21) for two photographic emulsions (Kodak 4463, Agfa-Gevaert 23D56). (Data from Hahn [56].)

where  $D_0$  is the 'fog',  $D_s$  is the saturation density, and  $c_s$  is a speed constant. Only with a sufficiently low exposure  $q \ll c_s^{-1}$  is a linear response obtained (Fig. 12).

The  $DQE(q)$  is determined by non-linearity, by fog, and by the statistics of the sizes and position of the grains as well as by the probability distribution for the number of grains formed by the primary. At medium electron energy it is characterized by a plateau with  $DQE \approx 0.6$  over about  $10^2$  and a decrease with lower exposures due to fog and for higher exposure due to nonlinearity according to Eqs. (7) and (3), respectively. A more detailed analysis is given by Hamilton and Marchant [55] and by Hahn [56]. In the high-energy region the DQE decreases strongly [57].

The resolution should ultimately be limited by the electron interaction volume. In practice, however, the point resolution is much worse since the saturation limit prohibits the accumulation of a number of electrons sufficient for the required SNR. As a consequence, the object detail to be resolved must be expanded to more pixels using higher magnification. Fortunately,

the photoplate has a large pixel reserve. Quantitative measurements of resolution or MTF have to be processed carefully in order to correct for the electron shot noise [58]. Defining the PSF taken from such measurements as the 'true' pixel, the range of useful expositions becomes evident by plotting the DQE over the number of electrons per pixel [59]. The true pixel size is approached in low-dose microscopy of radiation-sensitive specimens where extremely noisy records of numerous identical molecular objects are superposed by the computer.

Photographic recording has the merit of simplicity, high pixel number and satisfactory DQE. Nevertheless, it does not fulfill several demands of advanced quantitative microscopy such as accuracy, reproducibility, dynamic range, and real-time response. The argument of cheapness does not hold if a costly scanning photometer for subsequent digital processing is required. Also, the decrease in cost of digital storage capacity makes electronic recording increasingly attractive.

### 1.4.2 Imaging Plate

The benefits of this recorder may be a  $DQE > 0.9$  [31] and strict linearity over a dynamic ratio greater than  $10^4:1$ , erasability of the storage medium, easy replacement of the photoplate, and the digital output of the laser beam read-out device at 14 bits. Thus, a preferred application field is quantitative high-resolution microscopy and, in particular, diffraction, where the dynamic range of the photoplate is insufficient [28, 29]. The resolution is rather limited by the scanning laser beam

to about 25  $\mu\text{m}$ , but some improvement seems to be possible [31]. The relatively high atomic number of the phosphor favors its application in high-energy microscopy [30].

### 1.4.3 Electronic Recording

In electron microscopy, immediate availability of the image information is wanted for three reasons:

- cinematographic recording of dynamic processes;
- alignment and focusing of the instrument at a high level of precision as a condition for subsequent photographic recording of highly resolved images;
- surpassing the photoplate in DQE, dynamic range, and processing speed for sophisticated digital image processing and low-dose techniques.

The standard fluorescent screen of the microscope allows the coarse adjustment and selection of specimen areas of interest; however, due to the short integration time of the eye, even if a binocular is used, the images, in general, cannot be observed in a way which yields information equivalent to that in a micrograph recorded on photoplate. This drawback gave the impulse to using TV chains, which after years of development, considering the points of view mentioned in Sec. 1.2.2 of this Chapter, have reached a high level of maturity, particularly by including digital frame stores and Fourier processors in order to display the power spectra. But presently the availability of high-pixel CCDs is going to remove the restriction of pixel numbers and surpass the photoplate with

respect to DQE, linearity, and, in particular, to dynamic range.

#### 1.4.3.1 Television Chains

Historically, the brightness distribution of the phosphor screen was first transferred using optics of medium speed onto a highly sensitive multistage intensifier equipped with a camera tube. This arrangement, however, seriously violated the dimensioning rules (see Sec. 1.2.2 of this Chapter). Although a gain in brightness was achieved, the observed output noise was mainly determined by the low efficiency of the optics. Then the Vidicon principle was applied in an 'open' tube attached to the microscope bottom, using the EBIC effect (see Sec. 1.2.3.6 of this Chapter) in an amorphous selenium target directly illuminated by the electrons. Unfortunately, however, the progressive recrystallization of the selenium target by the electron bombardment prohibited durable operation. Similar problems arose in the illumination of silicon mosaic targets with fast electrons. References to these and other early experimental devices are given by Herrmann and Krahl [6].

Thus, a combination is preferable which consists of a sequence P20 phosphor screen, a fiber plate, and an S20 photocathode placed within an SIT tube or an intensifier (see Sec. 1.2.4.1 of this Chapter) [6]. According to Eq. (17) and (Sec. 1.2.3.6 of this Chapter) it avoids a break of the quantum level, and one 100 kV electron produces in the target a signal above  $10^5$  (SIT) or even some  $10^6$  (ISIT). Thus, single-electron counting is possible with a DQE dominated by  $\text{var } \Delta E$ . This arrangement became the standard TV converter

chain, allowing observation of images which would be invisible using the final fluorescent screen. Alternative devices are the use of a second-generation intensifier combined with a Plumbicon or Newvicon tube or, recently and now most usually, CCD video cameras.

As discussed in Sec. 1.2.1 of this Chapter, the accumulation of a sufficient SNR is required for detecting small details with low contrast. Thus, a matching of the image current density  $j$  to the storage time and, additionally, of the gain  $\bar{n}_e$  to the 'dynamic window' of the target is necessary according to the balance equation  $j\bar{n}_e = I_s/A$ , where  $I_s$  is the signal current in the camera tube and  $A$  is the scanned target area. In TV chains this flexibility is realized only within certain limits. A current density yielding a sufficient SNR within one frame period (20 ms) is practicable at low and medium electron optical magnifications as used for video recording of dynamic processes. In this case the gain must be kept small enough to avoid target saturation, and occasionally even vidicon tubes without an intensifier may be the appropriate choice. At high magnification and with high coherence requirements, however, the current density decreases, and at the same time the gain must be increased, and the quantum noise becomes visible on the monitor. As the retention time of the eye may then become insufficient, persistent monitors are employed. The more advanced technique of accumulating sequences of TV frames, however, is the digital frame store now used which is available even on personal computers. A format of  $512 \times 512 \times 8$  bit may be well matched to the capability of the TV standard. Several processing schemes are possible to realize a noise-reduced observation

with a selectable persistence time, and the display of power spectra using fast Fourier processors is available. Unfortunately, every TV frame also contributes read-out noise to the image.

As a measure approaching the ideal DQE, single-electron counting (see Sec. 1.2.4.4 of this Chapter) using ISIT has been successfully applied [6, 60], which can even utilize an essential part of the back-scattering tail of the PHD, i.e. above a threshold defined by the noise of the subsequent components. The normalization procedure discriminates from low background noise as well as from fixed patterns, and the obtainable SNR is only limited by the capacity of the frame store. However, since a superposition of the spots during the lag time must be avoided this method is limited to very low current densities  $j \ll 0.1 \text{ pA cm}^{-2}$ .

#### 1.4.3.2 Slow-Scan Charge-Coupled Device Converters with a Scintillator

The demands of modern quantitative microscopy are widely met by the slow-scan CCD technique (see Sec. 1.2.4.5 of this Chapter), allowing analog image accumulation on the chip with selectable storage time and 14-bit digitization for long-term storage and numerical processing [14, 61–63]; a very comprehensive survey on the design and applications of such devices is given by de Ruijter [64]. Recently, several sophisticated schemes of information retrieval in the angstrom range, such as holography, focal series, or tilting series combination, are profiting from the merits of CCD recording, which may even be the only means for their realization. Due to the high dynamic

range, an adaptation of the gain can be dispensed with. Nevertheless, in order to utilize the advantages in an optimum manner, a quantitative analysis of the signal response over a wide range of electron energies is advisable [6, 14]. It was found that the arrangement consisting of a scintillator, light optics, and CCD offers sufficient scope to fulfill the design rules given in Sec. 1.2.2 of this Chapter. According to Eq. (13), the mean single-electron signal is  $\bar{n}_e = \bar{n}_{ph} \eta_L \eta_{CCD} = \epsilon(\Delta E/E_{ph}) \eta_L \eta_{CCD}$ . Considering the variance of the energy deposits  $\Delta E$  and the superposed Poisson distribution of the photon numbers  $n_{ph}$  emitted at a certain  $\Delta E$ , one obtains

$$\frac{\text{var } n_{ph}}{\bar{n}_{ph}^2} = \frac{\text{var } \Delta E}{\Delta E^2} + \frac{1}{\bar{n}_{ph}} \quad (22)$$

and, due to the binary distributed yields of the light optics and CCD, this relation holds if  $n_{ph}$  is replaced by the single-electron signal  $n_e$ . Thus, it follows from Eq. (19) that

$$\text{DQE} = \left( 1 + \frac{\text{var } \Delta E}{\Delta E^2} + \frac{1}{\bar{n}_e} + \frac{1}{\bar{N}_p} \frac{n_r^2 + D\tau}{\bar{n}_e^2} \right)^{-1} \quad (23)$$

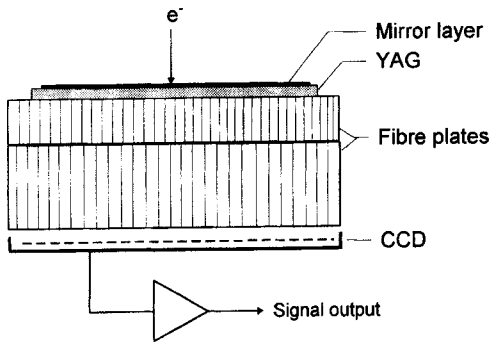
This relation can be straightforwardly evaluated using the data of the components if the probability distribution  $p(\Delta E)$  is known. The choice of the material and thickness of the scintillator as well as that of the light optical transfer element may depend on the energy. This will be discussed briefly for single-crystalline scintillators such as YAG [20], which has proven to be a good choice with respect to the maximization of  $\bar{n}_e/\Delta E$  (see Sec. 1.2.4.1 of this Chapter and

Table 1) as well as mechanical and chemical stability.

The distribution of the energy deposit  $p(\Delta E)$  as well as the interaction volume, determining the resolution  $d_r$  and the required light optical scaling factor  $d_{CCD}/d_r$ , are readily accessible by Monte Carlo simulation in a layered medium [13, 14]. For an optimization, three energy ranges may be distinguished:

(a) *Low energy, 5–100 keV.* The interaction range lies below  $1 \mu\text{m}$ . A light microscopic objective with  $\text{NA} > 0.6$  allows both a satisfactory efficiency and a magnification between  $10\times$  and  $40\times$  to be chosen according to the CCD pixel size. With decreasing energy the low single-electron signal determines the DQE, and a detection of individual electrons is no longer possible. The prerequisites, design and DQE of such a device are discussed in detail by Herrmann and Sikeler [65]. Since  $p(\Delta E)$  is exclusively due to backscattering, the relative variance of  $\Delta E$  in Eq. (23) can also be determined using analytical models of the energy distribution of back-scattered electrons [66, 67], which is complementary to  $p(\Delta E)$ .

(b) *Medium energy, 50–300 keV.* The increasing penetration requires a scaling factor of the order of 1:1 for a CCD pixel size of  $24 \mu\text{m}$ . Fiber plates have proven to be well suited (Fig. 13). The single-electron signal should exceed considerably the required  $\bar{n}_e \approx 10$ , if not, a scintillator thinner than the interaction range is used to improve the resolution. This standard converter design is commercially available and discussed in several publications [14, 62–64, 68]. By measuring the PHD using a photomultiplier with a multichannel analyzer and by Monte Carlo simulation,



**Figure 13.** Schematic diagram of the fiber plate-coupled CCD detector.

a DQE of 0.9 can be verified at 100 keV, with increasing energy; however, due to increasing transmission the full absorption peak disappears, and the DQE drops to about 0.6 at 300 keV [14]. The PSF as simulated and measured by evaluating edge transient curves remains nearly independent of the energy if the YAG thickness does not exceed 30  $\mu\text{m}$  [14]. Better matching of the thickness to the energy would intolerably impair the resolution as long as the 1:1 coupling by a fiber plate is used. More elaborate investigations of the shape of the PSF [69, 70] show a narrow peak placed on a relatively wide tail due to photon channeling. This tail can be removed in the processing system by

deconvolution. Replacing the YAG crystal by a thinner P20 phosphor avoids this effect and allows, due to the higher photon output (see Table 1), a further improvement of resolution at the expense of DQE [68].

(c) *High energy, 400–1000 keV.* In high-energy instruments, in order to conserve resolution, considerable transmission of the electrons through the phosphor layer must be tolerated; the mean energy deposit may be estimated from the approximation

$$\overline{\Delta E} = S\rho d \quad (24)$$

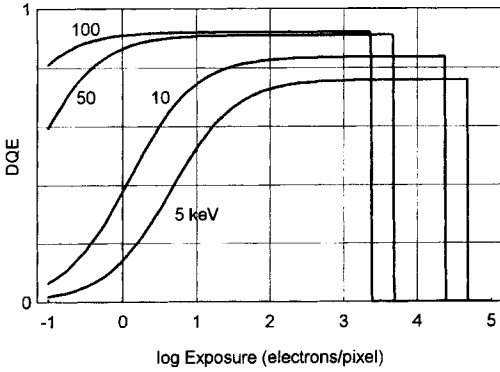
where the stopping power  $S := dE/d(\rho d)$  of the scintillator decreases with increasing energy. In order to reduce back scattering from the fiber plate as well as color center formation by the high transmitted radiation dose, a self-supporting scintillator is preferred which is slightly demagnified onto the CCD by high-speed macro-optics [71]. By Monte Carlo simulation a satisfactory DQE of 0.6 can be estimated at 1 MeV if a demagnifying tandem lens couple with  $F_1/F_2 = 1.5/0.75$  is used [72]. A phosphor powder screen sedimented onto a very thin mirrored plastic foil may be preferred as the scintillator.

Table 3 summarizes some important variables for the application of these

**Table 3.** Data for various electron image converters using a TEK 1024 CCD ( $D\tau \ll n_t^2$ ).

	Electron energy (keV)		
	10	100	1000
YAG scintillator thickness, $d_0$ ( $\mu\text{m}$ )	140	25	50
Optics	MO	FP	TO
Numerical aperture, NA	0.6	0.6	0.32
Single electron signal, $\bar{n}_c$ ( $e^-$ )	5	40	3
Storage capacity, $N_{\text{max}}$ (PE pixel $^{-1}$ )	$3 \times 10^4$	$4 \times 10^3$	$5 \times 10^4$
Maximum DQE, $\text{DQE}_{\text{max}}$	0.75	0.9	0.73
Lower detection limit at 0.9 $\text{DQE}_{\text{max}}$ , $N_{\text{min}}$	27	0.5	73

MO, micro-optics; FP, fiber plate; TO, tandem optics.



**Figure 14.** DQE of an electron image converter versus exposure at four different electron energies using Eq. (23). Assumptions: YAG single crystal,  $d_0 >$  penetration range,  $NA = 1.0$ ,  $T = 0.7$ ,  $R = 0.8$ ,  $n_{\max} = 1.5 \times 10^5 \text{ e}^- \text{ pixel}^{-1}$ ,  $n_r = 10 \text{ e}^- \text{ pixel}^{-1} \gg D\tau$ , and  $\eta_{\text{CCD}} = 0.4$ . The mean value of the energy deposits and its variance were obtained from an analytical back-scattering model [66, 67].

three devices. Figure 14 shows the evaluation of Eq. (23) using realistic parameters valid for devices (a) and (b), assuming that transmission can be neglected. The missing real-time response due to the read-out time of some seconds restricts the application of SSCCDs for immediate microscope adjustment; thus, an additional video rate camera (see Sec. 1.4.3.1 of this Chapter) is frequently used. Recently, however, a fast read-out mode ( $5 \text{ frames s}^{-1}$ ) with only slightly increased noise (e.g., see Table 2) allows the combination of both functions within the same equipment [68]. Also, a rapid readout of subarrays is possible.

#### 1.4.3.3 Directly Back-Illuminated Charge-Coupled Devices

At low energies below 5 keV the above-discussed converter chains fail due to very weak single-electron signals. While up to now the MCP (see Sec. 1.2.4.3 of this

Chapter) has been the most common choice within this range, thinned CCDs may offer a promising alternative [73]. From  $\bar{n}_e = \overline{\Delta E}/E_f$  with  $E_f = 3.6 \text{ eV}$  a signal results which is, according to Eq. (19), sufficient by far for a high DQE even with a single-electron response. At increasing energy, however, this solution fails due to the reduction of the saturation limit  $N_{p,\max} = n_{\max}/\bar{n}_e$ .

## 1.5 X-Ray Microscopy

In the energy range between 100 eV and 2 keV, preferred in X-ray microscopy, the photons interact with matter mainly by photoabsorption, which exceeds scattering by a factor of  $10^4$ . Fundamentally, the DQE in Eq. (11) is limited by the reduction factor

$$\eta_{\text{abs}} = \exp(-\alpha_d d_d)[1 - \exp(-\alpha d_0)] \quad (25)$$

where  $\alpha$ ,  $\alpha_d$  and  $d_0$ ,  $d_d$ , respectively, are the photoabsorption coefficients and the thicknesses of the converting layer and a dead layer in front of it. Thus, a precondition for a high DQE is the complete conversion of the impinging photons into detectable photoelectrons using both  $\alpha_d d_d \ll 1$  and  $\alpha d_0 \gg 1$ .

From the dependence of the mass absorption coefficient  $\alpha/\rho$  on the wavelength  $\lambda$  and atomic number  $Z$  according to

$$\frac{\alpha}{\rho} \approx CZ^4 \lambda^3 \quad (26)$$

(where  $C$  jumps at the characteristic edges) and using data from Henke et al. [74], certain basic tendencies can be concluded. In the low-energy region, particularly within the 'water window' below the  $0_K$

**Table 4.** X-ray absorption lengths  $1/\alpha$  ( $\mu\text{m}$ )

$E_{\text{ph,X}}$ (eV):	100	200	500	1000	2000	5000
$\lambda$ (nm):	12.39	6.19	2.48	1.24	0.62	0.25
$1/\alpha_{\text{PMMA}}$	0.4	2	1	2.5	20	
$1/\alpha_{\text{Si}}$	0.037	0.055	0.44	2.7	1.63	17.6
$1/\alpha_{\text{Be}}$	0.85	0.1	1.3	9.5	82.5	1469
$1/\alpha_{\text{Au}}$	0.03	0.046	0.043	0.104	0.44	0.70
$1/\alpha_{\text{CsJ}}$	0.03	0.31	0.55	0.25	1.17	2.55

After Henke et al. [74].

excitation energy preferred for hydrated biological samples, dead layers of even a few nanometers in thickness must be avoided. At increasing energy, however, a high conversion efficiency (e.g., due to  $d_0 \geq 2.3/\alpha$  for  $\eta_{\text{abs}} \geq 0.9$ ) offers increasing difficulties. Even if higher atomic numbers are used, the converter thickness  $d_0$  must be restricted to maintain the resolution at the expense of the DQE. Frequently,  $d_0$  must be chosen small enough to reduce Eq. (25) to  $\eta_{\text{abs}} \approx \alpha d_0 \ll 1$ . Table 4 gives some variables dependent on the wavelength for several typical materials [74]. This demonstrates the considerably different responses of detectors over the range of wavelengths used in microscopy.

The resolution is not limited primarily by the propagation of X-rays but rather by the range of Auger electrons ( $<5\text{ nm}$ ) [75]. Obviously, a resolution considerably below the light microscopic limit should, in principle, be possible. In practice, however, this limit is attained only in a few devices, while others are dominated by the spread functions of subsequent signal-processing steps.

The intrinsic electron signals are generally detected by one of the following effects:

(a) chemical effects (photographic emulsion, resist);

(b) luminescence in phosphors;  
(c) formation of electron-hole pairs;  
(d) electron emission.

Some corresponding devices will be discussed below, considering both off- and on-line operation, as applied to contact radiography where high resolution may be required, as well as to magnifying systems where the X-ray optics may be adapted to converters with medium resolution. Integration of extremely high signal rates may be required if flash exposures using subnanosecond laser sources are applied.

### 1.5.1 Photographic Film and Imaging Plate

Double-sided films must be dispensed with. Special gelatine-less fine-grain photo-emulsions allow a resolution of a few micrometers and a DQE of 0.1–0.3 [76].

The imaging plate (see Sec. 1.2.3.7 of this Chapter), originally developed for X-ray applications [25, 77], could be used in a thinned version as for electrons, promising a satisfactory DQE and dynamic range; however, due to the laser beam read out, the resolution may not approach the potential of soft X-radiation closely



enough. Thus, image plates appear to hold promise only for magnifying microscopes using harder radiation and for crystallography.

### 1.5.2 Resist

The solubility of high molecular chain molecules in organic developers, depending on the molecular weight, is increased or reduced, respectively, by X-radiation exposure due to bond scission (positive) or polymerization (negative resist) according to the absorbed energy. After developing the resist material, a relief is formed representing the spatial exposure distribution, which can be read out with sublight microscopic resolution using either transmission microscopy directly or a replica technique with shadowing, alternatively scanning electron or atomic force microscopy [78], as well as with confocal light optical microscopy [79]. Table 4 lists some absorption data, dependent on the photon energy, of polymethylmethacrylate (PMMA), the preferred resist (others are described by Valiev [80] and Seeger [81]). Obviously, in a layer 1  $\mu\text{m}$  in thickness, photons are almost completely absorbed only if their energy does not exceed 500 eV. With increasing energy a considerable loss in DQE is expected; this loss can be reduced by doping the resist with heavy metals [82].

Due to the several processing steps a quantitative evaluation is hardly possible. The dissolution rate  $S$  of a polymer in a developer fluid depends on the molecular mass  $M$  according to  $S \propto M^{-a}$ , with  $a \approx 1.5$  for PMMA developed in methyl isobutyl ketone, and thus the strongest

effect is expected for high  $M$ , which is chosen to be of the order of  $10^6$ . Since consecutive multiple scissions of one chain are possible, a strongly nonlinear characteristic curve  $S(E_{\text{abs}})$  (thinning rate versus energy deposit  $E_{\text{abs}}$ , that is, the volume density of absorbed radiation energy) arises depending on the development conditions but not on the wavelength [83]. The absorption law yields

$$E_{\text{abs}}(z) = E_{\text{in}}\alpha \exp(-\alpha z) \quad (27)$$

where  $E_{\text{in}} = nE_{\text{ph}}$  is the impinging local energy density. As long as  $\alpha d_0 \ll 1$ , the energy deposit  $E_{\text{abs}} = E_{\text{in}}\alpha$  is uniformly distributed over the thickness. Thus the thickness  $d$  of the resist layer decreases linearly with the developing time  $t_d$  starting with an initial thickness  $d_0$ , that is,

$$d = d_0 - S(E_{\text{abs}})t_d = d_0 - S(E_{\text{in}}\alpha)t_d \quad (28)$$

Obviously, the achievement of a large thickness modulation  $\Delta d/\Delta E_{\text{in}}$ , as desired for the weak contrast images of biological samples, for example, depends on the slope  $dS/dE_{\text{abs}}$  of the characteristic curve and on the development time  $t_d$ . If the initial thickness  $d_0$  is chosen large enough, then the weak intensity modulations can be converted to relatively large resist thickness variations compared to the remaining mean thickness. By proper matching of  $d_0$ ,  $t_d$ , and  $E_{\text{in}}$  under the consideration of  $S(E_{\text{abs}})$ , a minimum final thickness  $d_{\text{min}} > 0$  must be established. There is, however, an upper limit for the initial thickness  $d_0$  because the resolution requirements cannot be fulfilled if  $d_0$  is too thick, due to lateral under-etching.

If, on the other hand,  $\alpha d_0$  is large, as required for high DQE, then the dissolution rate  $S$  decreases with progressing

mass removal, which results in a saturation of the modulation  $\Delta d/\Delta E_{\text{in}}$ .

X-ray resist approaches a resolution below 10 nm limited by both the molecular dimensions and the Auger electron range [75,81]; the dynamic range, however, is poor [84], and the DQE is restricted to values around 0.1 due to the wide variance of molecule size. An estimate of the number of photons absorbed in a volume element of resolution dimensions shows that the effect of quantum noise plays a significant role [83]. Experimental evidence is found in the distinct granulation of uniformly exposed resist layers [84].

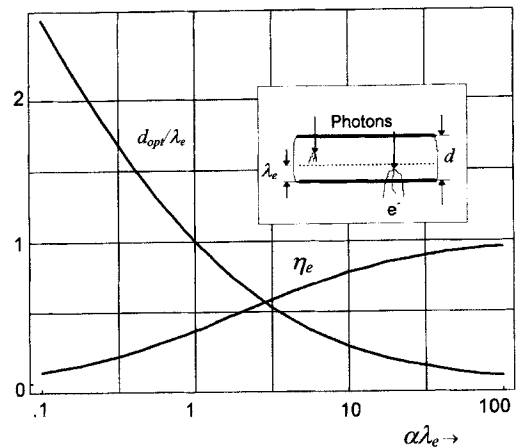
### 1.5.3 Transmission Photocathodes

On-line read out with a sublight microscopic resolution has been tried using a thin transmission photocathode, converting the photons into a spatial distribution of photoelectrons that are accelerated and magnified by means of a cathode lens and which impinge on a position-sensitive electron detector of medium resolution [85–88]. If the energy of the accelerated electrons exceeds about 10 keV, several types of highly efficient electron detectors are available (see Sec. 1.4.3 of this Chapter). A high DQE requires both the complete absorption of the incident photons and the efficient utilization of the photoelectron signals released in the solid. Following a simple model assumption [89, 90], the electron yield  $\eta_e$  can be estimated by integrating over the  $z$ -coordinate the product of the photon absorption probability, given by Eq. (27), and the escape probability, again an exponential containing the mean free path  $\lambda_e$  of the electrons,

with the result

$$\begin{aligned}\eta_e &= \int_0^{d_0} \alpha \exp(-\alpha z) \exp\left(-\frac{d_0 - z}{\lambda_e}\right) dz \\ &= \frac{\lambda_e \alpha}{1 - \lambda_e \alpha} \left[ \exp(-d_0 \alpha) - \exp\left(-\frac{d_0}{\lambda_e}\right) \right]\end{aligned}\quad (29)$$

As a consequence, there exists an optimum thickness  $d_{\text{opt}} = \lambda_e(\lambda_e \alpha - 1)^{-1} \ln(\lambda_e \alpha)$  of the photocathode for which  $\eta_e$  attains a maximum value  $\eta_{\text{max}} = (\lambda_e \alpha)^{1/(1 - \lambda_e \alpha)}$  (Fig. 15). Thus, materials are required offering a large value  $\lambda_e \alpha$ . Isolators, for example CsJ, may be superior to metals [89, 91]. Unfortunately, the DQE cannot be identified with the electron yield  $\eta_e$  since the electron emission does not follow a binary distribution as for visible radiation [90]. The energy of the X-ray photons is sufficient to release multiple electrons in each absorption event, and thus the



**Figure 15.** X-ray transmission photocathodes. Optimum relative thickness  $d_{\text{opt}}/\lambda_e$  and corresponding electron yield  $\eta_e$  versus the normalized absorption coefficient  $\lambda_e \alpha$  ( $\lambda_e$  is the mean free path length of secondary electrons in the solid).

emitted signals show an additional PHD, which, according to Eq. (11), leads to a further impairment of the DQE against  $\eta$  which must now be defined as the probability of releasing a count independent of its weight. Estimates and measurements of the DQE using gold foils [92] indicated values of the order of a few percent with  $\text{Ti K}_\alpha$  radiation. An increase of the DQE with increasing wavelength is expected.

The resolution of the transmission photocathode, on the other hand, is limited to the order of the mean free path length  $\lambda_e$ , which lies in the region of 4 nm (Au) and 25 nm (CsJ). The subsequent electron-optical processing system must be well designed in order to maintain this resolution. A cathode lens, as customary in various types of electron emission microscopes, accelerates the electrons to about 20 keV, and an additional lens inserted into the anode section produces a focused image at a magnification of about 20. The resolution of cathode lenses is dominated by the chromatic aberration, which depends to a rough approximation [93] on the energy width  $\Delta E_{pe}$  of the emitted electrons and the electric field strength  $F_e$  at the cathode surface according to

$$\delta \approx \Delta E_{pe}/F_e \quad (30)$$

For a more detailed analysis considering the other aberrations, see Polack and Lowenthal [86]. While for technical reasons the field strength is limited to about  $100 \text{ kV cm}^{-1}$ , the energy width may be considerably higher than in typical photoemission microscopes. There are two solutions: a diaphragm placed in the diffraction plane of the cathode lens and/or an imaging energy filter allowing a selectable restriction of the energy spectrum

contributing to the image. The Castaing-Henry system using an electrostatic mirror combined with a magnetic  $90^\circ$  deflector [94] as well as the purely magnetic  $\Omega$ -filter [95] are high-performance solutions to this problem, but unfortunately not achievable without some effort. In addition, the filtering may remove some signal pulses and thus further reduce the DQE.

### 1.5.4 Microchannel Plates

Due to its compactness the MCP is frequently used in X-ray astronomy combined with resistive or structurized anodes as position sensitive photon-counting detectors (review by Fraser [96]). The quantum efficiency for soft X-rays is between 1 and 10%; however, by depositing a material with a high secondary electron yield (e.g.,  $\text{MgF}_2$  or CsJ) an improvement of up to 50% can be achieved [97]. The high electron gain allows image pick-up from a subsequent phosphor screen by standard video cameras. The wide PHD, however, reduces the DQE and makes a spectroscopic resolution impossible. The saturation prohibits an application to flash exposures.

### 1.5.5 Television Chains

The design and application of TV pick-up devices for image display in real time has been reviewed [98, 99]. The rules of their design have to consider the general relations given in Sec. 1.2 of this Chapter requiring, in particular, a highly efficient photon conversion in the entrance stage by both a high photon absorption and access

to the secondary signals which is as complete as possible. While a few attempts with X-ray-sensitive pick-up tubes have successfully been made at higher photon energies, most devices for soft radiation utilize the window-less scintillators and standard optoelectronic components (see Sec. 1.2.4 of this Chapter) of LLL camera technology.

### 1.5.5.1 X-ray-Sensitive Vidicons

Several designs of X-ray-sensitive Vidicons have been reported (reviewed by Hartmann [99]) in which a beryllium window and a photoconductive target optimized for X-rays is generally used. These special tubes were developed for live topography using energies above a few kiloelectronvolts, and may, due to the beryllium window (see Table 4), be completely unable to detect very soft X-radiation with a high efficiency. In order to increase the X-ray absorption, a PbO target 15  $\mu\text{m}$  in thickness in a Vidicon tube has been used; with Cu  $K_\alpha$  radiation (8 keV) a DQE of 78% and a resolution below 25  $\mu\text{m}$  were reported [100].

### 1.5.5.2 Conversion to Visible Radiation

The basic physical principle of the classical X-ray intensifier is the conversion of the photons to electrons using a phosphor layer in close contact with a photocathode, the emitted electrons from which are accelerated and demagnified onto a highly resolving screen. Although this device has sometimes been used for topography, it is not applicable to microscopy, for the reasons mentioned above. For microscopy a

window-less conversion using an open phosphor is the best choice. The application of phosphors with subsequent LLL processing does not differ much from the devices mentioned above (see Sec. 1.2.4.4 of this Chapter). Generally the visible photons are fed using a fiber plate to the photocathode of a SIT tube or an intensifier (single or multiple stage) coupled to a Vidicon or a CCD. The resolution is limited by the light spread in the phosphor layer, depending on the thickness, and by the fiber plate structure. Phosphors with a high stopping power are preferred in order to optimize the resolution and photon output. Unfortunately, the photon attenuation in the phosphor prohibits full absorption with hard X-rays. This can be addressed using Eq. (29) by identifying  $\lambda_c$  with an empiric photon attenuation length  $\lambda_{ph}$  [16], with the conclusion that an optimum thickness of the phosphor screen exists offering an efficiency considerably below unity if  $\alpha\lambda_{ph} \ll 1$ .

Again the number of photoelectrons  $\bar{n}_e$  released by one absorbed X-ray photon of energy  $E_{ph,X}$  ( $\equiv \overline{\Delta E}$ ) can be obtained from Sec. 1.2.4.1 of this Chapter and Table 1. Obviously, only photons above 1 keV can release photoelectrons with an acceptable yield. Their multiplication in order to produce a charge spot on the target strong enough to overcome the read-out noise of the camera follows the general rules already treated in Sec. 1.2.4.4 of this Chapter.

### 1.5.6 Slow-Scan Charge-Coupled Device Chains

Cooled slow-scan CCDs may allow the construction of the most advanced X-ray

**Table 5.** Data for an X-ray image converter using a TEK 1024 CDD ( $D\tau \ll n_r^2$ ). Assumptions: CsJ(Tl) crystal,  $NA = 1.0$ ,  $T = 0.7$ ,  $R = 1$  ( $= 0$  for  $E_{ph,X} \leq 500$  eV),  $N_{p,min}$  for  $0.9 DQE_{max}$ ,  $d_{0,min}$  for  $\eta_{abs} = 0.99$ .

$E_{ph,X}$ (eV):	100	200	500	1000	2000	5000
$\lambda$ (nm):	12.39	6.19	2.48	1.24	0.62	0.25
$\bar{n}_e$ (e <sup>-</sup> )	0.12	0.24	0.6	2.03	4.06	10.1
$DQE_{max}$	0.11	0.2	0.39	0.67	0.8	0.91
$N_{p,max}$ (photon pixel <sup>-1</sup> )	$12 \times 10^5$	$6 \times 10^5$	$2.4 \times 10^5$	$7.5 \times 10^4$	$3.8 \times 10^4$	$1.5 \times 10^4$
$N_{p,min}$ (photon pixel <sup>-1</sup> )	6660	2990	921	146	43	8
$d_{0,min}$ (μm)	0.15	1.4	2.5	1.15	5.4	11.7

image read-out device, characterized by quantum-limited recording, extreme storage capability, independency of the signal rate, excellent linearity, pixel numbers comparable to photographic recording, and, due to the 14-bit digital output, on-line transfer to high-performance image-processing systems. Using an open scintillator avoids any photon losses by a window, and a versatile choice of material and thickness of the scintillator is possible. Combining Eqs. (11), (19), and (25), the DQE can be predicted to be

$$DQE = \eta_{abs} \left( 1 + \frac{1}{\bar{n}_e} + \frac{n_r^2 + D\tau}{\eta_{abs} \bar{N}_p \bar{n}_e^2} \right)^{-1} \quad (31)$$

The variance contribution of the energy deposit was neglected in this case because a partial absorption of the photon energy may not be expected.

Table 5 shows the evaluation of Eq. (31) at various photon energies assuming, as an example, a mirrored CsJ scintillator (see Table 1), optics with  $NA = 1.0$ , and a TEK 1024 CCD (see Table 2). Obviously, a satisfactory DQE is feasible above 500 eV only by using some signal integration. In order to keep the reduction factor  $\eta_{abs}$  as close as possible to unity a scintillator thickness  $d_0 \geq 2.3/\alpha$  may be chosen which sets different resolution limits over the energy range and consequently requires different scale factors for the

optics. At higher energy the DQE decreases since full absorption cannot be maintained. At low energy the decreasing single-electron signal dominates the response and, additionally, the gain by the aluminum mirror layer is overcompensated by the X-ray absorption according to Eq. (25). Thus, at very low photon energies, a mirror-less scintillator is superior. Using single-crystal scintillators, even the light microscopic resolution limit may be attained (see Sec. 1.2.3.4 of this Chapter).

### 1.5.7 Directly Illuminated Charge-Coupled Device Sensors

In principle, a soft X-ray photon produces a significant number  $\bar{n}_e = E_{ph}/3.6$  eV of electron-hole pairs with  $\text{var } n_e = F_a \bar{n}_e$  ( $F_a \approx 0.12$  Fano factor) in silicon (see Sec. 1.2.3.2). If recombination is avoided by the collecting field in the diodes this signal considerably exceeds the read-out noise contributions and allows excellent spectroscopic resolution as well as a single-quantum response over the whole range of photon energies  $E_{ph}$  relevant for X-ray microscopy. Even spatially resolved energy-dispersive spectroscopy is possible [101] as long as an overlap of signals on one pixel is avoided by low exposure and

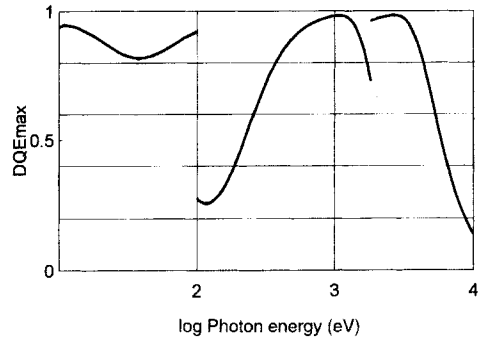
the read-out signal is fed via a multichannel analyzer to a subsequent multiple frame storage system. On the other hand, the

$$\begin{aligned} \text{DQE}_{\text{max}} &= \eta_{\text{abs}} (1 + F_a 3.6 \text{ eV} / E_{\text{ph}})^{-1} \\ &\approx \eta_{\text{abs}} \end{aligned} \quad (32)$$

is rather dominated by the probability of forming a signal within the wells, as given by their geometrical dimensions and the absorption coefficient of silicon according to Eq. (25). Unfortunately, with increasing energy the saturation limit  $n_{\text{max}}$  of the CCD limits drastically the  $\text{SNR} \leq \sqrt{\eta_{\text{abs}} N_{\text{p,max}}} = n_{\text{max}} 3.6 \text{ eV} / E_{\text{ph}}$ .

In practice some physical or technical impediments must be considered. Commercial CCDs are covered by a protective  $\text{Si}_3\text{N}_4$  layer about  $2 \mu\text{m}$  in thickness, which is transparent to visible light but strongly absorbs soft X-rays. Together with the metallic control electrodes, a dead layer is formed for soft radiation. Harder X-rays, on the other hand, are less absorbed in the dead layer, but also less so in the depletion zone. They may additionally impair the function of the MOS read-out elements by radiation damage [102]. Two possible remedies are: (a) covering the protected input face by a phosphor layer that converts each absorbed X-ray photon into a number of visible photons [103, 104], and (b) using thinned and back-illuminated CCDs [41, 105, 106] (see Sec. 1.2.4.5).

By using highly resistive material the depletion zone can be far extended, but its close approach to the back face of the thinned CCD offers a hard technological problem. A field-free charge-trapping dead layer remains. Using realistic values of  $d_d/d_0 = 50 \text{ nm}/20 \mu\text{m}$  [106] and the absorption coefficient of silicon (see



**Figure 16.**  $\text{DQE}_{\text{max}}$  of a directly X-ray back-illuminated CCD versus photon energy using Eqs. (25) and (32) and silicon absorption data from Henke et al. [74]. Assumptions: depletion depth  $d_0 = 20 \mu\text{m}$ , dead layer thickness  $d_d = 50 \text{ nm}$ , and Fano factor  $F_a = 0.12$ .

Table 4), from Eq. (25), results in  $\text{DQE} \approx 0.3/0.98/0.68$  at photon energies  $E_{\text{X,ph}} = 0.1/1/5 \text{ keV}$  (Fig. 16). Thus, using remedy (b) the dead layer problem at low energies is not entirely removed. Method (a) avoids a dead layer and allows the use of higher absorbing phosphors as well as components of consumer video cameras, while remedy (b) requires a more costly special design of a scientific-grade CCD and electronics. In any case, the resolution is limited by the sensor element size to about  $10 \mu\text{m}$  at best.

## Acknowledgment

I thank my colleague F. Lenz for critically reading the manuscript and for many valuable suggestions.

## 1.6 References

- [1] A. Rose, *J. Soc. Mot. Picture Eng.* **1946**, 47, 273.
- [2] J. C. Dainty, R. Show, *Image Science*, Academic Press, London **1974**.

- [3] C. Jones, *J. Opt. Soc. Am.* **1959**, 49, 645.
- [4] L. Mandel, *Br. J. Appl. Phys.* **1959**, 10, 233.
- [5] E. Breitenberger, *Progr. Nucl. Phys.* **1955**, 4, 56.
- [6] K.-H. Herrmann, D. Krah, *Adv. Opt. El. Micr.* **1984**, 9, 1.
- [7] A. Boksenberg, C. I. Coleman, *Adv. Electron. Electron Phys.* **1979**, 52, 355.
- [8] U. Fano, *Phys. Rev.* **1947**, 72, 26.
- [9] H. J. Zweig, *J. Opt. Soc. Am.* **1961**, 51, 310.
- [10] P. N. J. Dennis, *Photodetectors: An Introduction to Current Technology*, Plenum Press, New York **1985**.
- [11] M. Jedlicka, P. Kulhánek, *Vacuum* **1986**, 36, 515.
- [12] H. P. Kallmann, G. M. Spruch, *Luminescence of Organic and Inorganic Materials*, Wiley, New York **1962**.
- [13] D. C. Joy, *Inst. Phys. Conf. Ser., EUREM 88, York* **1988**, 93/I, 23.
- [14] I. Daberkow, K.-H. Herrmann, L. Libin, W.-D. Rau, *Ultramicroscopy* **1991**, 38, 215.
- [15] M. Kotera, Y. Kamiya, *Ultramicroscopy* **1994**, 54, 293.
- [16] D. E. Husk, S. E. Schnatterly, *J. Opt. Soc. Am. B.* **1992**, 9, 660.
- [17] S. M. Gruner, J. R. Milch, G. T. Reynolds, *Nucl. Instrum. Methods* **1982**, 195, 287.
- [18] JEDEC, *Optical Characteristics of Cathode Ray Tube Screens*, Electrical Industry Association, Washington DC. **1975**.
- [19] J. H. Chappell, S. S. Murray, *Nucl. Instrum. Methods* **1984**, 221, 159.
- [20] R. Autrata, P. Schauer, Jiri Kvapil, Josef Kvapil, *Scanning* **1983**, 5, 91.
- [21] Galileo Electro-Optics Corp. *Fibre Optics: Theory and Application, Technical Memorandum 100, Fibre Optic Faceplates, Data Sheet 1000*. Galileo Electro-Optics Corp., Sturbridge, MA.
- [22] Optische Werke G. Rodenstock, *X-Ray Manual*, Optische Werke G. Rodenstock, München **1992**.
- [23] L. Reimer, *Scanning Electron Microscopy, Springer Series in Optical Science*, Vol. 45, Springer, Berlin **1985**.
- [24] L. Reimer, *Image Formation in Low-Voltage Scanning Microscopy, Tutorial Texts in Optical Engineering*, No. 12, SPIE Optical Engineering Press, Bellingham **1993**.
- [25] G. W. Goetze, *Adv. Electron. Electron Phys.* **1966**, 22A, 219.
- [26] M. Sonoda, M. Takano, J. Miyahara, H. Kato, *Radiology* **1983**, 148, 833.
- [27] Y. Amemiya, K. Wakabayashi, H. Tanaka, Y. Ueno, J. Miyahara, *Science* **1987**, 237, 164.
- [28] N. Mori, T. Oikawa, T. Katoh, J. Miyahara, Y. Harada, *Ultramicroscopy* **1988**, 25, 195.
- [29] D. Shindo, K. Hiraga, T. Oikawa, N. Mori, *J. Electron Microsc.* **1990**, 39, 449.
- [30] S. Isoda, K. Saitoh, T. Ogawa, S. Moriguchi, T. Kobayashi, *Ultramicroscopy* **1992**, 41, 99.
- [31] C. Burmester, H. G. Braun, R. R. Schröder, *Ultramicroscopy* **1994**, 55, 55.
- [31a] L. M. Biberman, S. Nudelman (Eds.), *Photo-electronic Imaging Devices, Vol. I and II*, Plenum Press, New York **1971**.
- [32] E. H. Eberhardt, *Appl. Optics* **1968**, 7, 2037.
- [33] M. Lampton, *Sci. Am.* **1981**, 245(5), 46.
- [34] K.-H. Herrmann, M. Korn, *J. Phys. E: Sci. Instrum.* **1987**, 20, 177.
- [35] Galileo Electro-Optics Corp. *Microchannel Plates, Data Sheet 9000, Detector Assemblies, Data Sheet 7000*. Galileo Electro-Optics Corp., Sturbridge, MA.
- [36] R. G. Neuhauser, *J. SMPTE* **1987**, 96, 473.
- [37] B. L. Morgan (Ed.), *Adv. Electron. Electron Phys.* **1988**, 74.
- [38] G. A. Robinson, *J. SMPTE* **1980**, 89, 249.
- [39] D. F. Barbe (Ed.), *Charge Coupled Devices, Springer Topics in Applied Physics*, Vol. 38, Springer, Berlin **1980**.
- [40] Photometrics Ltd., *Charge-Coupled Devices for Quantitative Electronic Imaging*, Photometrics Ltd., Tucson **1989**.
- [41] R. P. Khosla, *Microelectr. Eng.* **1992**, 19, 615.
- [42] Philips, *J. Res.* [Special issue] **1994**, 48, 3.
- [43] M. L. Blouke (Ed.), *SPIE Proc.* **1994**, 2172.
- [44] E. R. Fossum, *SPIE Proc.* **1994**, 2172, 38.
- [45] M. Rabbani, R. van Metter, *J. Opt. Soc. Am. A* **1989**, 6, 1156.
- [46] H. Mulder, *J. Opt. Soc. Am. A* **1993**, 10, 2038.
- [47] J. C. Feltz, M. A. Karim, *Appl. Opt.* **1990**, 29, 717.
- [48] D. Shotton (Ed.), *Techniques in Modern Bio-medical Microscopy*, Wiley-Liss, New York **1993**.
- [49] S. Inoué (Ed.), *Video Microscopy*, Plenum Press, New York **1986**.
- [50] T. T. Tsay, R. Inman, B. Wray, B. Herman, K. Jacobsen, *J. Microsc.* **1990**, 160, 141.
- [51] H. Gursky, G. Fritz, *Scanning* **1991**, 13, 41.
- [52] R. K. Aikens, D. A. Agard, J. W. Sedat, *Methods Cell Biol.* **1989**, 29, 291.
- [53] B. R. Masters, G. S. Kino, *Electronic Light Microscopy*, Wiley-Liss, New York **1993**, p. 315.
- [54] E. Zeitler, *Ultramicroscopy* **1992**, 46, 405.
- [55] J. F. Hamilton, J. C. Marchant, *J. Opt. Soc. Am.* **1967**, 57, 232.
- [56] M. Hahn, *Electron Microscopy at Molecular Dimensions* (Eds.: W. Baumeister, W. Vogell) Springer, Berlin **1980**, p. 200.
- [57] D. G. Ast, *J. Appl. Phys.* **1974**, 45, 4638.
- [58] D. A. Grano, K. H. Downing, *38th Annu. EMSA Meeting* **1980**, 228.

- [59] K.-H. Herrmann, *Electron Microsc.* **1982**, *1*, 131.
- [60] K.-H. Herrmann, D. Krah, H.-P. Rust, O. Ulrichs, *Ultramicroscopy* **1978**, *3*, 227.
- [61] J. C. H. Spence, J. M. Zuo, *Rev. Sci. Instrum.* **1988**, *59*, 2102.
- [62] S. Kujawa, D. Krah, *Ultramicroscopy* **1992**, *46*, 395.
- [63] O. L. Krivanek, P. E. Mooney, *Ultramicroscopy* **1993**, *49*, 95.
- [64] W. J. de Ruijter, *Micron* **1995**, *26*, 247.
- [65] K.-H. Herrmann, R. Sikeler, *Optik* **1995**, *98*, 119.
- [66] H. Niedrig, *Electron Beam Interactions with Solids for Microscopy. Microanalysis & Microlithography* (Eds. D. F. Kyser, H. Niedrig, D. E. Newbury, R. Shimizu), Scanning Electron Microscopy, Inc., AMF O'Hare **1982**, p. 51.
- [67] U. Werner, H. Bethge, J. Heydenreich, *Ultramicroscopy* **1982**, *8*, 417.
- [68] I. Daberkow, K.-H. Herrmann, L. Liu, W. D. Rau, H. Tietz, *Ultramicroscopy*, submitted.
- [69] A. L. Weickenmeier, W. Nüchter, J. Mayer, *Optik* **1995**, *99*, 147.
- [70] J. M. Zuo, *Electron Microsc.* **1994**, *1*, 215.
- [71] G. Y. Fan, M. H. Ellisman, *Ultramicroscopy* **1993**, *52*, 21.
- [72] K.-H. Herrmann, L. Liu, *Optik* **1992**, *92*, 48.
- [73] D. G. Stearns, J. D. Wiedwald, *Rev. Sci. Instrum.* **1989**, *60*, 1095.
- [74] B. L. Henke, P. Lee, T. J. Tanaka, R. L. Shimabokuro, B. K. Fujikawa, *Atomic Data Nucl. Data Tables* **1982**, *27*, 1.
- [75] K. Murata, M. Tanaka, K. Kawata, *Optik* **1990**, *84*, 163.
- [76] L. Jochum, *SPIE Proc.* **1986**, *733*, 492.
- [77] J. Miyahara, K. Takahashi, Y. Amemiya, N. Kaniya, Y. Satow, *Nucl. Instrum. Methods A* **1986**, *246*, 572.
- [78] T. Tomie, H. Shimizu, T. Majima, M. Yamoda, T. Kanayama, H. Kondo, M. Yano, M. Ono, *Science* **1991**, *252*, 691.
- [79] P. C. Cheng, D. M. Shinozaki, *X-Ray Microscopy III* (Eds.: A. G. Michette, G. R. Morrison, C. J. Buckley), Springer, Berlin **1992**, p. 335.
- [80] K. Valiev, *The Physics of Submicron Lithography*, Plenum Press, New York **1992**.
- [81] D. Seeger, *IBM J. Res. Dev.* **1993**, *37*, 435.
- [82] I. Haller, R. Feder, M. Hatzakis, E. Spiller, *J. Electrochem. Soc.* **1997**, *126*, 154.
- [83] E. Spiller, R. Feder, *Topics Appl. Phys.* **1977**, *22*, 35.
- [84] D. M. Shinozaki, *X-Ray Microscopy II* (Eds.: D. Sayre, M. Howells, J. Kirz, H. Rarback), *Springer Series in Optical Sciences*, Vol. 56, Springer, Berlin **1987**, p. 118.
- [85] L. Y. Huang, *Z. Phys.* **1957**, *149*, 225.
- [86] F. Polack, S. Lowenthal, *X-Ray Microscopy* (Eds.: G. Schmahl, D. Rudolph), *Springer Series in Optical Sciences*, Vol. 43, Springer, Berlin **1984**, p. 251.
- [87] F. Polack, S. Lowenthal, D. Phalippou, P. Fournet, *X-Ray Microscopy II* (Eds.: D. Sayre, M. Howells, J. Kirz, H. Rarback), *Springer Series in Optical Sciences*, Vol. 56, Springer, Berlin **1987**, p. 220.
- [88] K. Kinoshita, T. Matsumura, Y. Inagaki, N. Hirai, M. Sugiyama, H. Kihara, N. Watanabe, Y. Shimanuki, A. Yagashita, *X-Ray Microscopy III* (Eds.: A. G. Michette, G. R. Morrison, C. J. Buckley), *Springer Series in Optical Sciences*, Vol. 67, Springer, Berlin **1992**, p. 335.
- [89] B. L. Henke, J. P. Knauer, K. Premaratne, *J. Appl. Phys.* **1981**, *52*, 1509.
- [90] G. W. Fraser, *Nucl. Instrum. Methods A* **1985**, *228*, 532.
- [91] A. Akkerman, A. Gibrekhterman, A. Breskin, R. Chechik, *J. Appl. Phys.* **1992**, *72*, 5429.
- [92] W. Mullion, Diploma Thesis, Tübingen University **1993**.
- [93] A. Recknagel, *Z. Phys.* **1941**, *117*, 689.
- [94] R. Castaing, L. Henry, *C. R. Acad. Sci. Paris* **1963**, *B255*, 76.
- [95] S. Lanio, H. Rose, D. Krah, *Optik* **1986**, *73*, 56.
- [96] G. W. Fraser, *Nucl. Instrum. Methods* **1984**, *221*, 115.
- [97] G. W. Fraser, *Nucl. Instrum. Methods* **1983**, *206*, 251, 265.
- [98] S. M. Gruner, *Rev. Sci. Instrum.* **1989**, *60*, 1545.
- [99] W. Hartmann, *Topics Appl. Phys.* **1977**, *22*, 191.
- [100] R. Nishida, S. Okamoto, *Rep. Res. Inst. Electron., Shizuoka University* **1966**, *1*, 21, 185.
- [101] D. H. Lump, G. R. Hopkinson, A. Wells, *Nucl. Instrum. Methods A* **1984**, *221*, 150.
- [102] N. M. Allinson, D. W. E. Allsop, J. A. Quayle, B. G. Magorrian, *Nucl. Instrum. Methods A* **1991**, *310*, 267.
- [103] P. Burstein, J. M. Davis, *X-Ray Microscopy* (Eds.: G. Schmahl, D. Rudolph), *Springer Series in Optical Sciences*, Vol. 43, Springer, Berlin **1984**, p. 184.
- [104] R. Germer, W. Meyer-Ilse, *Rev. Sci. Instrum.* **1986**, *57*, 426.
- [105] C. Castelli, A. Wells, K. McCarthy, A. Holland, *Nucl. Instrum. Methods A* **1991**, *310*, 240.
- [106] J. L. Culhane, *Nucl. Instrum. Methods A* **1991**, *310*, 1.





## 2 Image Processing

### 2.1 Introduction

Image processing is a generic term which covers several activities related to digital images, that is, images which have been converted into arrays of pixels, as described in Sec. 1 of this Chapter. These topics concern image preprocessing, image restoration, image segmentation, image analysis, three-dimensional reconstruction, etc.

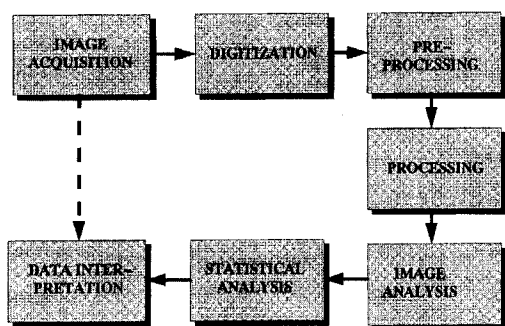
The microscopes described in the previous chapters are able to produce different kinds of image. Although the simple observation of these images is often sufficient for interpreting the specimen, it is increasingly necessary to manipulate the images, in order either to help the qualitative interpretation or to extract some quantitative information.

Unfortunately, it is not so easy to go from image acquisition to data exploitation through image analysis. Usually, a number of intermediate steps must be devised (Fig. 1). Image preprocessing consists of preparing images for subsequent operations. Image processing (in the restricted sense of the term) consists of transforming an image into another one, which is intended to be more manageable, either in terms of interpretation or in terms of subsequent analysis. Image analysis

consists of transforming an image into a data set, that is into descriptors of the image or of the 'objects' which have been depicted in the image. Such features can be used to quantify the image content or to perform a kind of pattern recognition. It should be emphasized that, at this level, the last three steps (preprocessing, processing, and analysis) are not always clearly separated. Preprocessing was previously often qualified as 'cosmetic' processing, that is a kind of processing where the aim is mainly to improve the visual appearance of images. However, this is not the unique purpose of these operations. Similarly, image processing and analysis are often mixed. This is the case with mathematical morphology, for instance, where the same operations can be used for both purposes.

Finally, after image analysis methods have been devised, a statistical approach has also to be defined for estimating the significance of the results. For instance, the number of images which have to be processed before any conclusion can be drawn must be estimated carefully.

The scheme described above concerns the processing/analysis of individual images. It must be stressed that modern microscopes can also provide image series in addition to single images. Different types of image sequence are now produced



**Figure 1.** Although we would like to go directly from the image acquisition to data interpretation, it is generally necessary to follow a much longer path, which includes different activities described by the term image processing.

more or less routinely: time series (video-microscopy), spectral series (electron energy loss, X-ray and Auger imaging), depth series (serial sections, depth profiling in SIMS), tilt series (electron or X-ray microscopies), focus series, etc. The combination of one or several of these parameters allows the recording of data in a large dimensional space. Therefore, these complex data sets cannot, in general, be interpreted and handled without the help of a computer. Specific methods have now been developed for the processing and analysis of such image sequences. These methods include, among others, the three-dimensional reconstruction (from serial sections, stereoscopy, or tilt series), the parametric imaging method (modelization approach), and the multivariate statistical analysis.

## 2.2 Image Preprocessing

Image preprocessing can be considered as an attempt to overcome some of the limitations of the imaging instrument,

which have not been accounted for at the acquisition level. (Of course, it must be stressed that everything that can be done at the specimen preparation level or at the experimental imaging level must be done – the less the computer is solicited, the better.)

Two main concepts are involved at the preprocessing stage: contrast and signal-to-noise ratio (SNR). The contrast of an image is a rather intuitive concept and need not be defined. The SNR is the ratio of the deterministic (or repetitive) contribution to an image over the stochastic (or nonrepetitive) contribution (noise). The ratio of the signal variance to the noise variance, for instance, can be used as a definition of the (quadratic) SNR (QSNR). Noise originates either from electronic fluctuations (detector) or from the quantum uncertainties associated with low-count signals. Although the SNR can often be improved at the experimental level by increasing the acquisition time, it is sometimes necessary to improve it a posteriori, in order to avoid prohibitive acquisition times.

Image preprocessing methods were developed at the early stages of digital image processing. At the beginning, these methods were what we now call *global methods*. This means that they perform the same function at any place in the image, whatever the local image content. Although this approach can be useful, it is often insufficient, as it does not modify fundamentally the image content. For this reason, these methods were criticized for doing mainly a cosmetic job (i.e. they render the images more pleasant to look at, but they do not help very much in the extraction of useful information). Nowadays, new methods (called *local*

*methods*) have been developed which do a different job according to the local content (contrast and SNR) of images.

## 2.2.1 Global Methods for Image Preprocessing

An image can be thought of as a collection of pixels along a square (or rectangular, or hexagonal) arrangement  $(x, y)$ , or a collection of gray levels (GLs) or colors. In the former space, we speak of an 'image space'. It can be described either as a real space (coordinates  $x, y$ ) or as a reciprocal space (the frequency space: coordinates  $x^{-1}, y^{-1}$ ). In the later case, we speak of a 'parameter space'. Global image processing can be performed either in the image space or in the parameter space.

### 2.2.1.1 Examples of Global Image Preprocessing in Image Space or Image Frequency Space

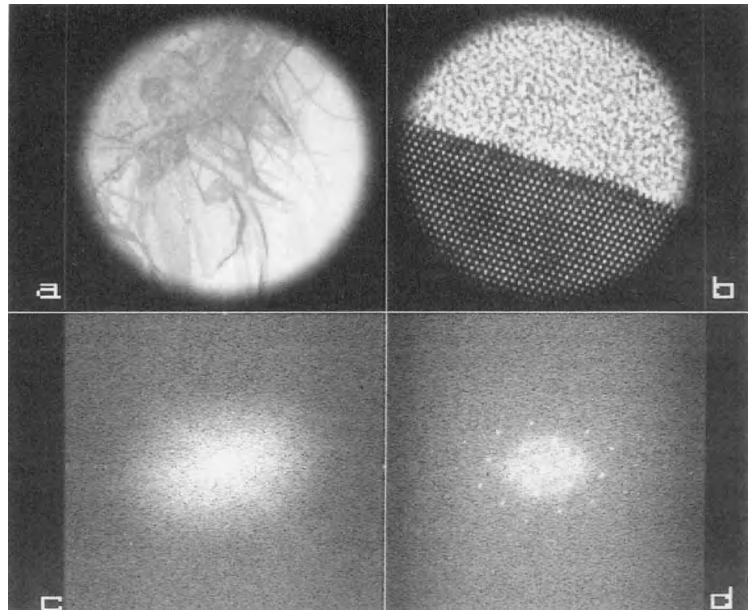
Most of the processing methods in this group are related to the concept of *filtering*. We must be aware of the fact that any image is in fact the superimposition of several components: the low frequencies, which correspond to the slowly varying components of the image (general shape of objects and regions); the intermediate frequencies, which correspond to details of the objects or regions; and the high frequencies, which correspond to rapidly varying parts of the images (i.e., to edges, very small details and also to noise), which in general vary from pixel to pixel. These different components can be

visualized in the image spectrum, which can be obtained either by diffractometry on an optical bench or by computing the power spectrum density (squared modulus of the image Fourier transform). Figure 2 shows two different images (one with a mostly low frequency content and one with a high frequency content) and their respective power spectrum density.

Since the different frequency components of an image can be identified, they can also be selectively modified. This is called *frequency filtering*. Applying a high-pass filter (which consists of lowering the amount of low frequencies in the image spectrum and then returning to the image space by an inverse Fourier transformation) results in contrast improvement, as the weight of details and edges is increased. However, this is at the expense of SNR, because the high-frequency noise components are also enhanced. Applying a low-pass filter (which consists of lowering the relative amount of high frequencies in the image spectrum before computing the inverse Fourier transform) results in an improvement of SNR, as the importance of high-frequency noise components is decreased. This approach, however, is at the expense of contrast and apparent resolution, because the high frequencies of edges and details are also weakened. These two approaches are exemplified in Fig. 3.

Applying a band-pass filter (which consists of selecting specific regions of the image frequency space) may result in contrast enhancement and SNR improvement, as selected frequencies can be reinforced at the expense of others. It must be noted, however, that this procedure, which is often used with images of

**Figure 2.** Illustration of the frequency content of an image. (a) Scanning electron microscope image of a whisker; this image contains mainly low frequency components and no periodicities. (b) Transmission electron microscope image of an interface between an amorphous structure and a crystalline structure; this image contains more high frequency components and periodicities. (c) Power spectrum density (squared modulus of the Fourier transform) of (a) displayed on a logarithmic scale. (d) Power spectrum density of (b) displayed on a logarithmic scale; note that it is necessary to apply a circular mask to the image in order to avoid artifacts in the Fourier spectrum.



periodic specimens, may also be dangerous, even if the problems associated to sharp cut-off filters are carefully avoided (in the extreme case, the filtered image of a nonperiodic specimen may appear as periodic, with a period corresponding to the central frequency of the band-pass filter).

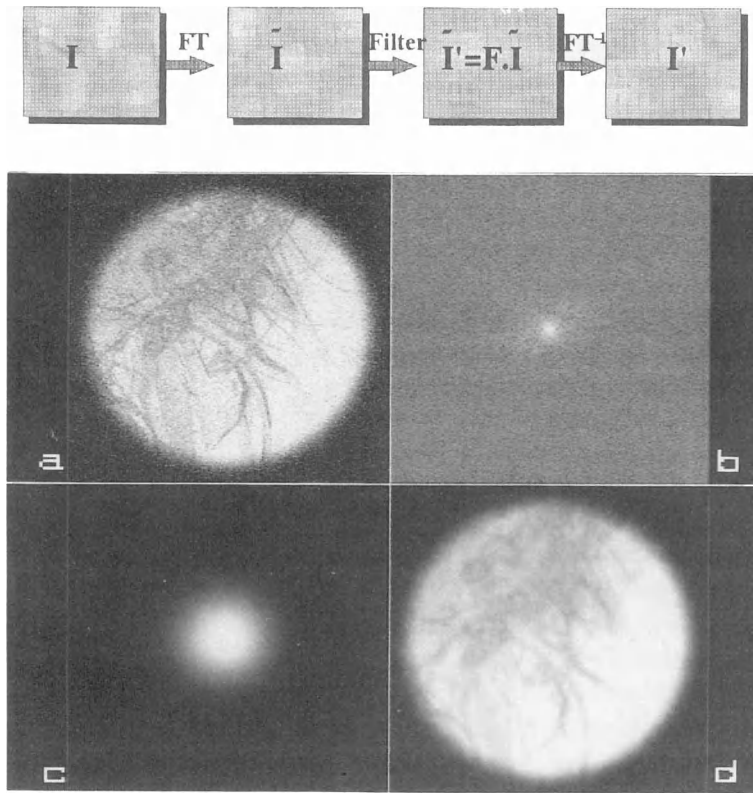
Frequential filtering is not the only way to perform these operations, because multiplying by a filter function in the image frequency space is equivalent to performing a convolution in the image space. This process is illustrated in Fig. 4. In practice, the size and the content of the convolution kernel have to be defined. We then have to scan the image to be processed with a moving window of the same size as the kernel. The result (at the center of the moving window) is thus obtained by weighting the image content

by the coefficients of the kernel. Each kind of filtering (low-pass, band-pass, high-pass filtering, first order and second order derivatives, Laplacean, etc.) is characterized by a specific convolution kernel.

Nonlinear filters can also be used. For instance, the median filter is often used to improve SNR: it consists of replacing the gray level at the center of the moving window by the median of the values in the window.

### 2.2.1.2 Examples of Global Image Preprocessing in Parameter Space

When an image is characterized by only the gray level (or color) values and not by the pixel coordinates, at best we have to compute the gray level histogram,



**Figure 3.** Schematic diagram of the procedure for filtering in the frequency space. (a) Original noisy image. (b) Power spectrum density of (a); logarithmic scale. (c) Low-pass filter applied to the image spectrum. (d) Filtered image; the noise is largely suppressed, but at the same time blurring occurs.

which represents the first-order statistics. (Note that higher order statistics can also be computed, see Sec. 2.4.3 in this Chapter.)

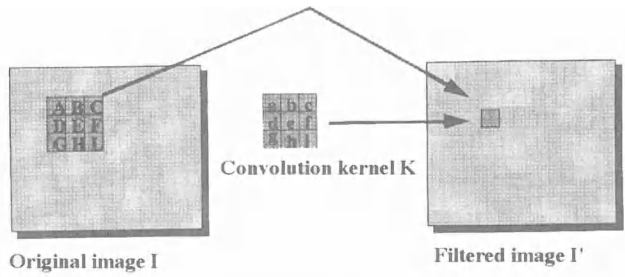
When the gray level histogram is built, it is possible to manipulate it. Replacing the original gray levels by the new ones results also in a new (processed) image. The different possibilities include: video inversion, histogram equalization, histogram hyperbolization, low gray level enhancement (logarithmic transfer function), high gray level enhancement (exponential transfer function), and false color rendering.

This classical approach for gray level manipulation, which is detailed in many

textbooks for image processing [1–4], is illustrated in Fig. 5.

## 2.2.2 Local Methods for Image Preprocessing

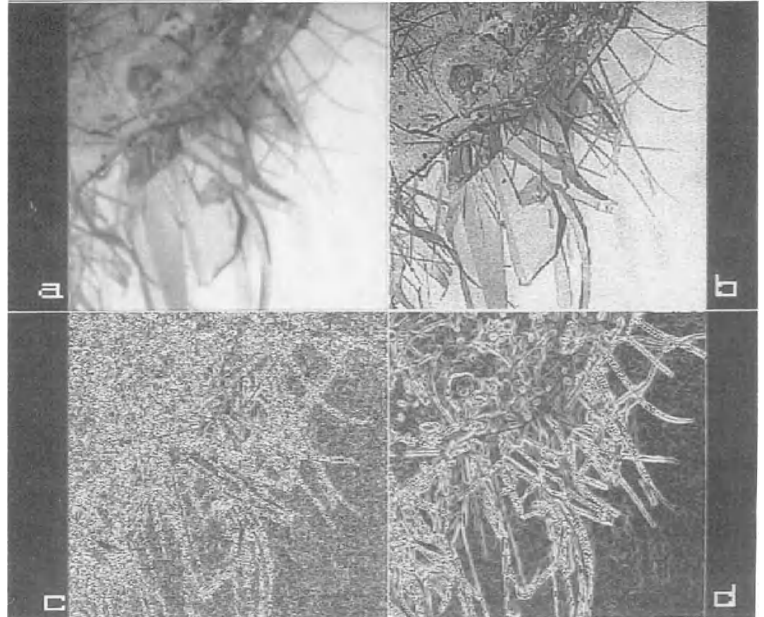
The global methods for SNR improvement or contrast enhancement described in the previous paragraph can be found in any commercial software package for image processing. However, these packages suffer from several drawbacks. During the past 10 years, new algorithms have been developed in order to overcome the drawbacks but, unfortunately, they



**Figure 4.** Illustration of the filtering principle according to a convolution process.

The result obtained for a position of the moving window is obtained by weighting the local content of the original image  $I$  ( $A, B, C$ , etc.) by the content of the convolution kernel  $k$  ( $a, b, c$ , etc.):  
 $I' = I * k = A \cdot a + B \cdot b + C \cdot c + \dots + I \cdot i$ .

(a) Smoothing with a  $7 \times 7$  Gaussian kernel (the original image is that shown in Fig. 5b).  
 (b) Contrast enhancement with a  $5 \times 5$  high-pass kernel. (c) Edge detection (Kirsch kernel) applied to the original image; the result of this high-pass filtering is much too noisy to be useful. (d) Same edge detection as applied to the smoothed image in (a).



are not included in commercial software. We describe here two examples of such algorithms.

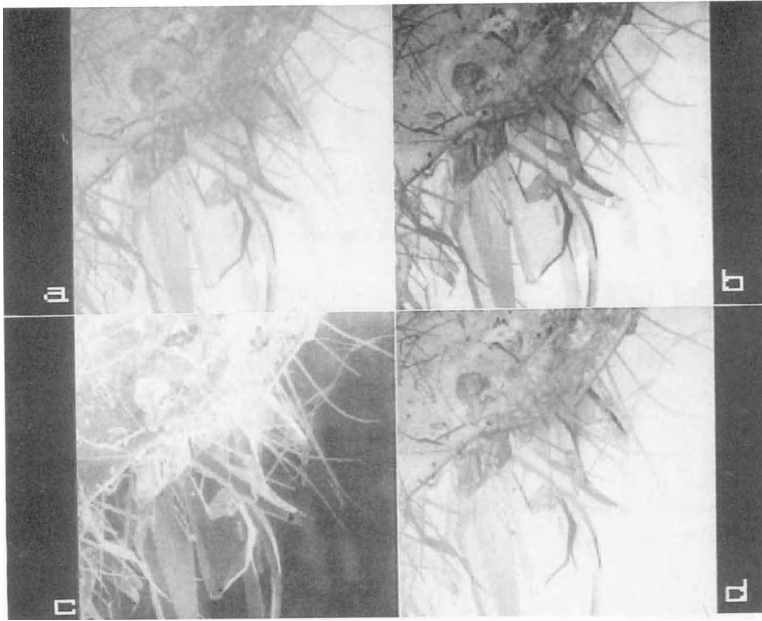
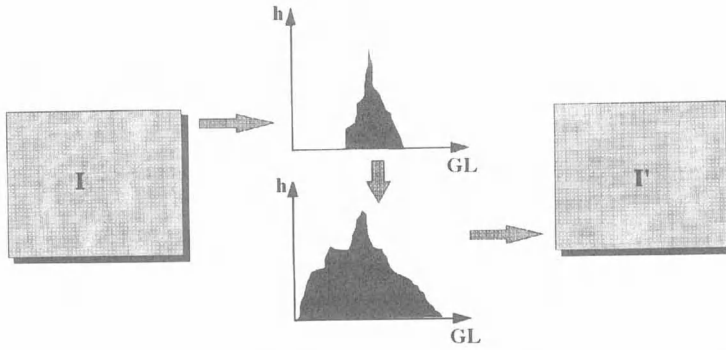
#### 2.2.2.1 Example of Algorithm for Local Contrast Enhancement

The purpose of contrast enhancement is often to increase the detectability of objects against a background (see Sec. 2.3 in this Chapter). Global contrast enhancement does not help very much in this purpose. This can only really be done by

local contrast enhancement, as was anticipated by Gabor [5] as early as 1965. An example of such a method was suggested by Beghdadi and Le Negrat [6]. The local contrast is defined for each pixel as

$$C(x, y) = \frac{|I(x, y) - E(x, y)|}{I(x, y) + E(x, y)} \quad 0 \leq C \leq 1 \quad (1)$$

where  $I$  is the pixel gray level and  $E$  is the mean gray level value of edges in a local window centered on  $(x, y)$  (Fig. 6). Once this local contrast is defined, it can be increased replacing the contrast by



**Figure 5.** Illustration of global image preprocessing by the histogram-manipulation approach (here, histogram stretching is used as an example). (a) Initial image of whiskers with low contrast. (b) Image resulting from histogram stretching. (c) Contrast inversion of (b). (d) Anamorphosis (logarithmic transfer function) of (a).

$C' = C^n$  ( $n < 1$ ). From this new contrast, the new gray level value can be computed as

$$I'(x, y) = E(x, y) \frac{1 - C'(x, y)}{1 + C'(x, y)}$$

$$I(x, y) \leq E(x, y) \quad (2a)$$

$$I'(x, y) = E(x, y) \frac{1 + C'(x, y)}{1 - C'(x, y)}$$

$$I(x, y) > E(x, y) \quad (2b)$$

This procedure has the attractive property

of reinforcing the histogram bimodality; that is it reinforces the appearance of two (or several) distinct peaks (or modes), revealing distinct features.

#### 2.2.2.2 Example of Algorithm for Improving the Signal-to-Noise-Ratio

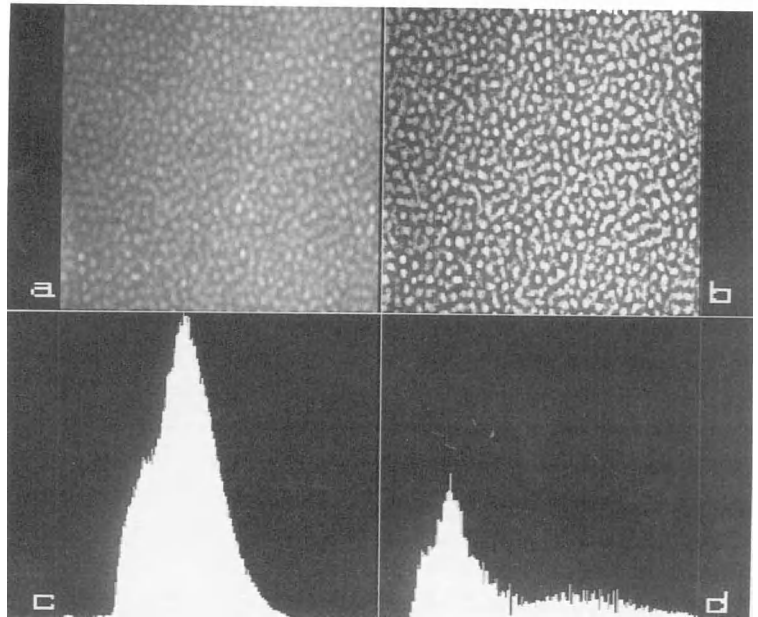
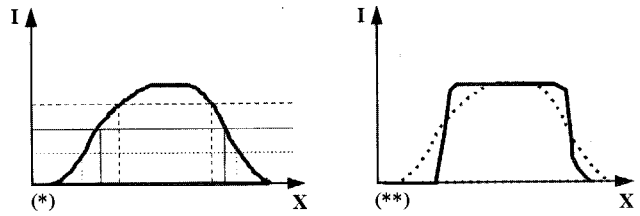
As described above, the SNR can be improved by smoothing, but this is at the expense of contrast and resolution. SNR



**Figure 6.** Illustration of a local contrast enhancement procedure. (\*) Schematic diagram of a profile across an object: the contrast is weak and thus, for different values of the threshold, different values of the object area (or perimeter) are obtained. (\*\*) The local contrast enhancement procedure helps to define steeper edges from which the binarization procedure is much less sensitive to the threshold value. (\*\*\*) Illustration of the local window procedure: within this window, the average value of the edge gray levels is computed as:

$$E(x, y) = \frac{\sum_{x', y' \in W} I(x', y') \cdot G(x', y')}{\sum_{x', y' \in W} G(x', y')}$$

where  $G$  represents the gray level gradient magnitude. If the pixel at the center of the window has a gray level value lower than  $E$  the gray level is decreased; otherwise, it is increased. This results in contrast enhancement. (a) Image of defects observed in silicon after ion implantation. (b) The result obtained after local contrast enhancement (window size 31 pixels). (c) Histogram of (a). (d) Histogram of (b). The local contrast enhancement procedure has the attractive property of reinforcing the histogram bimodality. (Unpublished. Courtesy of J. Fauré, University of Reims, and S. Simov, Institute of Solid State Physics, Sofia.)



improvement without the loss of too much contrast and resolution can be obtained if smoothing is performed on approximately constant gray-level areas while smoothing

is avoided on regions containing details and edges. Anisotropic diffusion [7] is one of the procedures which has been developed for this purpose: it is an iterative

procedure which consists of computing the image gradient, performing weighted smoothing (with weights inversely proportional to the gradient magnitude), and refining the gradient computation, etc.

Other approaches to this problem are: Unser's filter [8], Maeda and Murata's filter [9], and kriging [10, 11].

## 2.3 Processing of Single Images

Processing an image means transforming an image into another one which is easier to interpret or easier to analyze quantitatively. To illustrate these two purposes, we have chosen to describe two different groups of image processing: image restoration and image segmentation.

### 2.3.1 Image Restoration

Image restoration consists of taking into account the instrument characteristics and the experimental conditions which sometimes do not make the image a faithful representation of the object function, but instead a degraded representation. When these microscope and imaging characteristics are known, or can be determined, a restoration of the image can be attempted. It must be emphasized that, in some cases (e.g., an electron microscope working at the ultimate resolution), the restoration of the object function at the exit plane of the specimen is the main goal of image processing, as it is a prerequisite to any image interpretation.

The different kinds of degradation of the object function during image formation include: geometric distortion, uneven illumination, detector nonlinearities, aberrations, and additive and multiplicative noise. Most of these degradations can be eliminated (at least partly) a posteriori, provided they can be characterized [12]. Below we illustrate some, but not all, possibilities.

#### 2.3.1.1 Restoration of Linear Degradations

A linear system can be described by an impulse response  $h$  in the real space or a transfer function  $H$  in the frequency space

$$i(x, y) = o(x, y) * h(x, y) + n(x, y)$$

or

$$I(x^{-1}, y^{-1}) = [O(x^{-1}, y^{-1}) \cdot H(x^{-1}, y^{-1})] + N(x^{-1}, y^{-1}) \quad (3)$$

where  $o$  stands for the object function,  $i$  for the image,  $n$  for additive noise,  $O$ ,  $I$  and  $N$  for their Fourier transforms and  $*$  for the convolution product. The impulse response and the transfer function contain a contribution of the instrument characteristics (aberrations, objective diaphragm aperture, etc.) and of the imaging conditions (defocus value for instance).

Restoration by a simple inverse filtering procedure

$$\hat{O}(x^{-1}, y^{-1}) = I(x^{-1}, y^{-1}) \cdot H^{-1}(x^{-1}, y^{-1}) \rightarrow \hat{o}(x, y) \quad (4)$$

where  $\hat{\phantom{x}}$  indicates an estimated function, does not provide a correct solution, because the frequencies at which  $H = 0$  are amplified infinitely, and this results in a degraded SNR.

A Wiener filter gives a much better solution, as the second term in the denominator (a noise-to-signal ratio) prevents the filter from becoming infinite [13]:

$$\hat{O}(x^{-1}, y^{-1}) = \frac{I(x^{-1}, y^{-1}) \cdot H^*(x^{-1}, y^{-1})}{\left[ H^2(x^{-1}, y^{-1}) + \frac{N^2(x^{-1}, y^{-1})}{O^2(x^{-1}, y^{-1})} \right]} \quad (5)$$

The difficulty arises in this noise-to-signal ratio, which has to be estimated independently.

An alternative solution is given by iterative methods which consist of: (a) estimating the object function  $\hat{o}_k(x, y)$ , the first estimate  $\hat{o}_0$  being the image itself; (b) computing the image which such an object would give in the experimental conditions described by  $h(x, y)$ ; and (c) comparing the computed and the experimental image and updating the object function  $\hat{o}_{k+1}(x, y)$  according to the difference and to some additional constraints. Note that this procedure converges for strictly positive transfer functions, while in the other cases a reblurring procedure must be devised [14].

### 2.3.1.2 Restoration of Partly Linear Degradations: Very High Resolution Electron Microscopy

In the previous section we assumed that image formation can be characterized by a linear transfer function. Unfortunately, this is not always the case. In transmission electron microscopy, for instance, the image intensity (and not the image amplitude) is recorded (for books specifically devoted to image processing in electron microscopy, see [15, 16]).

This intensity can be written as

$$i = (1 + o * h)^2 = 1 + 2o * h + (o * h)^2 \quad (6)$$

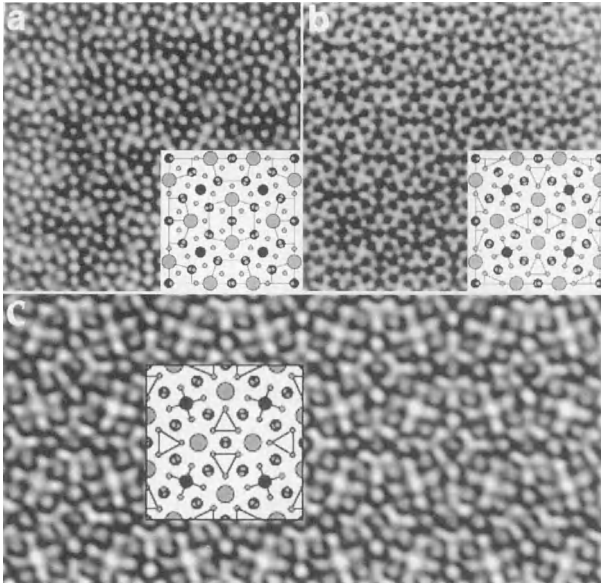
Thus, even for weak phase objects, the last term is nonlinear and prevents any restoration based on the linear system theory. Restoring the object wavefunction  $o$  therefore necessitates more sophisticated schemes, including: focus series, tilt series, and holography (see this volume, Chap. IV, Sec. 1.8). Owing to the limited space, we restrict ourself to a brief description of one of these approaches: the focus series approach.

This approach to object wave restoration was suggested as early as 1973 [17], but within the framework of the linear imaging theory. Later, image reconstruction from a focus series was considered as a nonlinear process and tackled by a recursive least-squares formalism. Recent developments [18, 19] can be described as: (a) firstly, preliminary reconstruction with a fully linear procedure; (b) secondly, a few iterations where the nonlinear terms are treated as a perturbation; and (c) thirdly, a fully nonlinear procedure, according to the paraboloid method (PAM). The updated electron wave is obtained from the previous estimation by

$$\hat{o}^{(k+1)} = \hat{o}^{(k)} + \gamma^{(k)} s^{(k)} [d(\hat{o}^{(k)})] \quad (7)$$

where  $d(\hat{o}^{(k)})$  is the difference wave computed from least-squares minimization,  $s^{(k)}$  is the best search direction determined by a conjugate gradient approach, and  $\gamma^{(k)}$  is a feedback parameter (gain).

Applications of this kind of approach concerns studies at the atomic resolution level. The amplitude and phase of the object function are reconstructed separately and can be used to determine the



**Figure 7.** Example of restoration of the object wavefunction in high resolution electron microscopy (Courtesy of D. Van Dyck and M. Op de Beeck, RUCA, Antwerp): (a) Amplitude of the reconstructed exit wavefunction starting from 20 images of  $\text{Ba}_2\text{NaNb}_5\text{O}_{15}$ . Mainly the heavy atoms are visualized. (Thickness 12 mm, microscope CM20FEG–SuperTwin). (b) Phase of the same reconstructed exit wavefunction mainly showing the light atoms. (c) Structure reconstruction towards the 1S channelling eigenstates of the different columns, showing all atomic columns in one image. (Structure models inserted).

‘true’ atomic positions. An example of reconstruction is given in Fig. 7. It concerns the structure of  $\text{Ba}_2\text{NaNb}_5\text{O}_{15}$  where the light atoms are revealed in the phase and the heavy atoms in the amplitude.

Numerous variants of this approach (tilt series reconstruction [20], stochastic algorithms [21], maximum entropy reconstruction [22], and a crystallographic approach [23, 24]) have been suggested, and have provided encouraging results. However, they require further study before they can be used routinely.

### 2.3.1.3 Example of a Completely Nonlinear Restoration: Near-Field Microscopies

In scanning tunneling microscopy and atomic force microscopy, image formation does not follow the scheme of imaging instruments with lenses (see this Volume, Chap. VII, Sec. 1). For corrugated surfaces, the image contrast is governed by both the shape of the specimen surface  $o$

and that of the tip  $s$ . It can be shown [25] that the relationship is of the form:

$$i(x, y) = \max_R [o(x', y') - s(x' - x, y' - y)] \quad (8)$$

This relation is clearly nonlinear and it has been shown [26] that the image function can be considered as the dilatation of the specimen surface function by a three-dimensional structuring element, the tip (see Sec. 2.3.2.5 in this Chapter for a definition of the dilatation operation and of the structuring element).

As image formation is not a linear operation, we cannot expect a linear operation to help in restoring the object function (the local specimen height). In fact, we can show that the best restoration can be obtained through the image erosion by the tip shape as a three-dimensional structuring element

$$\hat{o}(x, y) = \min_R [i(x', y') + s(x - x', y - y')] \quad (9)$$

The tip shape itself can be estimated when an image of a well-known object is made:

$$s(x, y) = -\min_R [i(x', y') - o(x' - x, y' - y)] \quad (10)$$

Owing to the specific properties of the dilation/erosion operations, we can also show that some estimate of the tip parameters can be obtained on the basis of the image alone, that is without any reference object (when the tip shape is known) [26].

### 2.3.2 Image Segmentation

Image segmentation is that part of image processing which consists of partitioning the whole image into different constituents. Depending on the type of image, this can mean one (or several) object(s) over a background or different regions.

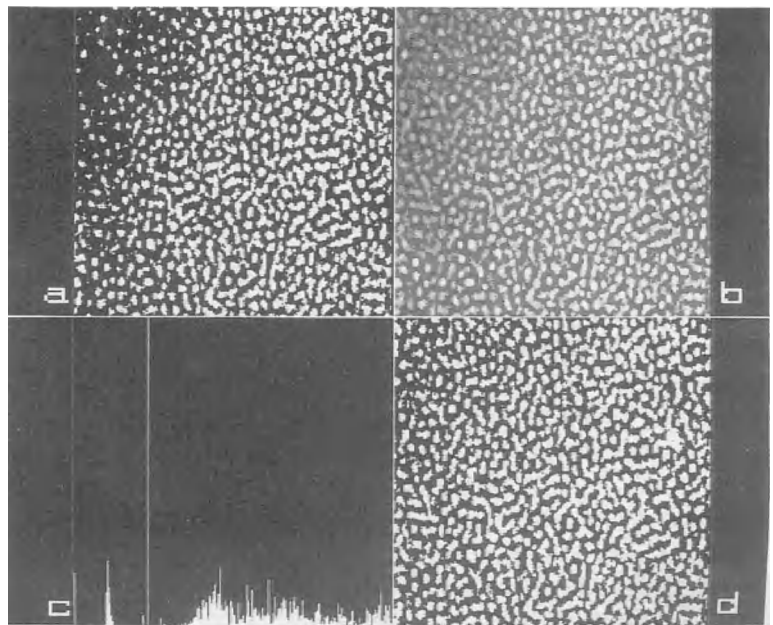
While this task is easy to perform for the human visual system (because it is associated to the observer's brain), it is a very difficult task for a computer, which does not 'know' the purpose of the segmentation (Fig. 8).

Many attributes of the image can be used in trying to perform an automatic segmentation, including: the gray levels; the gray-level gradients; the gray levels and the concept of connectivity; and the gray levels, gradients and connectivity together. This explains why hundreds of methods have been tried for performing this task (for a review see [27]).

#### 2.3.2.1 Segmentation on the Basis of Gray Levels Only

As stated above, when only gray levels (or colors) of an image are considered, we

**Figure 8.** (a) Illustration of the difficulty in segmentation using the histogram alone. The image displayed in Fig. 6b was binarized according to the histogram valley position. The segmentation is far from perfect, despite the histogram bimodality. (b) Result of applying the region-growing algorithm to the image in Fig. 6b (the threshold for growing was set at 30 gray levels). (c) Histogram corresponding to (b). Three main peaks on the left correspond to the background, while the others correspond to different granular structures. (d) The result obtained when binarizing (b) is much better.



have to work with the histogram. A necessary condition to perform segmentation on the basis of the histogram alone is that it is bimodal (or multimodal). If this is so, it can be assumed that each group of objects (or each group of regions) is characterized by a peak in the histogram. Thus, it is necessary to define the thresholds between these peaks and to attribute the pixels to the different regions according to their gray-level values (compared to the segmentation thresholds). A number of methods have been suggested in order to find, automatically, the thresholds of multimodal histograms (for reviews see [28, 29]). Since the different methods rely on different assumptions concerning the statistical characteristics of an image, none of them are expected to give the best solution in any situation. Furthermore, the histogram multimodality, though a necessary condition, is not expected to be a sufficient condition: it is easy to build examples where the peaks in a histogram are not representative of distinct regions. Thus, the segmentation on the basis of the histogram is often rather questionable.

### **2.3.2.2 Segmentation on the Basis of Gray-Level Gradients**

When we try to go beyond the simple concept of gray levels, it is essential to define a strategy for segmentation. Two strategies have previously been defined. The first consists of defining objects (or regions) as sets of connected pixels which possess some kind of attribute in common. The attribute can be a gray level, a color, or a texture. This concept of attribute homogeneity leads to the algorithm of

‘region growing’ (see below). The other strategy consists of defining contours of objects (or boundaries of regions) as pixels for which there is a strong discontinuity of an attribute. These two approaches are in fact complementary since, when regions are defined, their boundaries are also known, and when closed boundaries are defined regions inside these boundaries are also known. But the intermediate steps differ greatly for these two strategies.

The discontinuity-based strategy consists first of computing either the gray-level gradient or the gray-level Laplacean (an edge can be identified either by the local maximum of the gradient amplitude or by the zero-crossings of the second-order derivative).

Despite the fact that derivative operations are ill-conditioned (i.e., are very sensitive to a small amount of noise), very efficient digital filters are now available for performing such computations [30, 31]. However, this first step is not sufficient, because ‘edges’ obtained in this manner are noisy, unclosed, etc. Therefore, a number of postprocessing operations, such as edge closing, edge following, and hysteresis [32], are needed in order to obtain the real boundaries of regions.

### **2.3.2.3 Segmentation on the Basis of Gray-Level Homogeneity and the Concept of Connectivity**

As described above, an object or a region can be defined as a connected set of pixels for which some attribute is homogeneous. From this definition, an algorithm can be defined: from the seed of a region, consider connected pixels and incorporate them in the same region if the difference between

the pixel gray value and the average gray value of the region is smaller than a predefined threshold. Otherwise, the pixel serves as a seed for a new region. This region-growing algorithm is easy to understand and easy to implement. However, it also suffers from some drawbacks: (a) the threshold has to be defined, which is not an easy task, so several thresholding values have to be tried in order to choose the best segmentation, according to some criterion of what is a 'good' segmentation; and (b) the result is very sensitive to the choice of the seeds (i.e., to the order in which the pixels are processed).

This region-growing method is sometimes used in conjunction with a region-splitting method (split-and-merge algorithm [33, 34]).

#### 2.3.2.4 Segmentation on the Basis of Gray Levels, Gradients, and Connectivity: Functional Minimization

Owing to the difficulties associated with the methods just described, and also to the lack of criterion for good segmentation, a more recent approach consists of associating a cost function with any segmented image. The idea is to associate two terms: the first guarantees that the different regions remain sufficiently close to the original image (fidelity principle) and the other states that the boundaries of the regions must be as short as possible (simplicity principle). The cost function (also called the energy function) can then be written as:

$$E = \sum_R \sum_{x,y \in R} |i(x,y) - E_R[i(x,y)]|^2 + \alpha L \quad (11)$$

where  $E_R$  is the average gray level of region  $R$ ,  $L$  is the total length of boundaries and  $\alpha$  is a weighting coefficient proportional to the size of the smallest desired region [35].

The strategy of segmentation consists of minimizing this energy function. From this general strategy, several algorithms can be defined: either approximate contours are drawn interactively and then (automatically) deformed iteratively until a minimum energy is attained ('snakes' approach), or small homogeneous regions are merged until a minimum of the energy function is also obtained (conditional region-growing approach).

#### 2.3.2.5 Mathematical Morphology

In addition to all the approaches described above, an important class of procedures deals with the concept of neighbor relationship. These procedures form the basis of mathematical morphology [36–38], which is concerned with operations that respond to a feature shape (segment, disc, etc.) called a structuring element. The applications of mathematical morphology are not limited to image segmentation, but concern also image preprocessing (cleaning) and image analysis [39]. The basic tools are:

- (a) *Erosion*. At every pixel position, the result is the minimum gray-level value over the part of the image covered by the structuring element, placed at this position.
- (b) *Dilatation*. The maximum gray level value is selected instead of the minimum value.

From these two basic operations, many others can be defined, such as:

- (c) *Opening*. Erosion followed by dilatation.
  - (d) *Closing*. Dilatation followed by erosion.
  - (e) *Morphological gradient*. Eroded image subtracted from the dilated image.
  - (f) *Top-hat transformation*. Open image subtracted from the original image; this allows the detection of small bright particles (or details, or spots) over an uneven background.
- What could be the most powerful tool for image segmentation is the watershed transformation [40,41], which can be considered as a region-growing approach starting from markers of the regions: the regional minima. From these minima, a flooding process is realized, which results in different catchment basins separated by the watershed lines.
- (b) *Shape features*. Shape can be characterized by a number of methods: (i) global shape factors such as form factor ( $P^2/4\pi A$ ), eccentricity (or ellipticity), which characterize the elongation of the object; (ii) geometrical moments [42]; and (iii) coefficients of the Fourier series (development of  $p(s)$ , where  $p$  is the vector radius computed from the center of mass, and  $s$  is the normalized arc length) [43].
  - (c) *Densitometric features*. Provided a proper calibration can be made, the integrated gray levels within object can be related to one of the physical characteristics of that object (density, absorption coefficient, molecular mass, etc.).
  - (d) *Texture and fractal parameters*. (See Sec. 2.4.3 of this Chapter).
  - (e) *Spatial distribution features*. See a discussion in the literature [44].

## 2.4 Analysis of Single Images

Once an image is split (automatically or interactively) into several objects or several regions, image analysis can take place. This consists of computing features of the different constituents.

### 2.4.1 Object Features

A nonexhaustive list of features includes:

- (a) *Geometrical features*. Area  $A$ , perimeter  $P$ , Feret diameters (sizes of the smallest rectangle containing the object), Euler number (number of objects minus number of holes), intercept numbers, etc.

### 2.4.2 Pattern Recognition

Pattern recognition is also a part of image analysis. Several aspects can be encountered.

Pattern recognition may consist of detecting objects of known shape within images, without performing image segmentation. One of the most used approaches in this context is the generalized Hough transform. The Hough transform was originally designed for detecting straight lines. It consists of mapping each pixel of the original image into a parameter space and characterizing the object to be detected; for example, for a circle the parameter space is composed of the center coordinates and the radius. When such an



object effectively exists in the image space, an accumulation point appears in the parameter space. In electron microscopy, the Hough transform was used to detect automatically lines in Kikuchi patterns [45]. The Hough technique has been extended to the detection of circles, ellipsoids, polygons or objects of any known shape [46].

Pattern recognition may consist of the detection of any object, characterized by its gray-level pattern. In this case, a *matched-filter* technique can be used, which is generally implemented as a cross-correlation technique [47]. The object to be detected is placed at the center of an empty image of the same size as the real image where it is to be searched for. The two images are then cross-correlated:

$$C(\Delta x, \Delta y) = \frac{E[i_1(x, y) \cdot i_2(x + \Delta x, y + \Delta y)] - E[i_1(x, y)] \cdot E[i_2(x + \Delta x, y + \Delta y)]}{\sigma(i_1) \cdot \sigma(i_2)} \quad (12)$$

where  $E$  is the expectation value and  $\sigma$  the standard deviation. It can be shown that peaks of the  $C(\Delta x, \Delta y)$  function appear in places where the objects are located.

Pattern recognition may consist of recognizing or classifying objects on the basis of the features which have been computed from them. Recognizing objects assumes that a bank of objects (with their associated features) already exists. The problem is then one of *supervised* pattern recognition, for which numerous methods exist, ranging from the computation of the smallest distance to neural network approaches. Classifying a set of objects without reference to previously known classes is a problem of *unsupervised* pattern recognition (or automatic clustering) [48].

## 2.4.3 Image Analysis without Image Segmentation

The usual procedure of image processing (segmentation) followed by image analysis (computation of object features) is sometimes inoperative because the whole image is either composed of only one region, which has to be characterized globally, or is composed of a very large number of very small or overlapping objects that cannot be isolated and characterized individually. In these two extreme cases specific tools have to be used for performing image analysis. Some examples of these tools are described below.

### 2.4.3.1 Texture Analysis

The problem of texture recognition is one of the problems for which the human visual system is much more efficient than any computer system. However, many tools are now available. Texture may be recognized through different attributes, including graininess, directionality, uniformity, and elongation. These features can be characterized and quantified with concepts generally related to second-order statistics, that is statistical parameters relating the gray level of one pixel to the gray level of another pixel (a distance  $d$  apart). Such parameters include co-occurrence matrices and their descriptors [49], output amplitudes of real space filters [50, 51], Gabor frequency space filters [52] and wavelet decomposition.

Pixels (or groups of pixels) of textured images can thus be described by vectors of attributes and thus subjected to pattern recognition or pattern classification (texture segmentation).

### 2.4.3.2 Fractal Analysis

The pioneering work of Mandelbrot [53] has shown that natural scenes often exhibit self-similarity, that is they have similar behavior on different scales. This property can be used to characterize images which otherwise appear very chaotic. Here, again, many different methods have been suggested for estimating fractal attributes, the most used of which is the fractal dimension  $D$ :

- (a) *Area-perimeter method*. Following the suggestion of Mandelbrot, the fractal dimension can be estimated by performing image binarization at different thresholds (pixel values below the gray level threshold are set to zero, while pixel values above the threshold are set to one). For each threshold, the total area  $A$  and total perimeter  $P$  of the thresholded objects are computed. The linear regression of the  $\log(P)$  versus  $\log(A)$  data provides an estimation of  $D$ . An application to STM images is given in [54].
- (b) *Power spectrum method* [55]. For a purely fractal object, the power-spectrum density is proportional to the spatial frequency at the power of a parameter  $H$  connected to  $D$ :  $P(|f|) = |f|^{-2(H+1)}$ . Thus, a  $\log(P)$  versus  $\log(|f|)$  regression allows us to estimate  $D = 2 - H$ .
- (c) *Hurst coefficient method* [56]. For each pixel in an image, the first, second, third, etc., neighbors are considered and the variance of the gray levels for each neighborhood plotted as a function of the distance (on a log-log scale). Again, the slope of the line

provides an estimation of the local fractal dimension  $D$ .

- (d) *Blanket method* [57]. The gray levels are considered as the third coordinate of a  $z(x,y)$  plot. The surface thus created is iteratively expanded towards the higher and lower gray-level values by computing the maximum and minimum values over a given neighborhood. The surface variation is then plotted (on a log-log scale) as a function of neighborhood size. Again, the slope of the curve gives an estimate of the fractal dimension  $D$ . The advantage of this method (which is closely related to the mathematical-morphology approach) is that both the upper and lower sides of the surface can be characterized individually.

Comparative studies are presently being performed in order to evaluate the advantages and inconveniences of the different approaches in specific situations.

### 2.4.3.3 Stereology

Another way to obtain estimates of certain geometrical quantities without segmenting an image into its different constituents is provided by the stereological approach. More specifically, stereology allows us to estimate quantities in an  $N$ -dimensional space from measurements (or countings) performed in a space of dimension  $(N - 1)$  [58, 59]. For instance, the ratio between the volume of two compartments ( $V_A/V_B$ ) can be estimated from their average area ratio ( $S_A/S_B$ ), measured in random two-dimensional sections through the three-dimensional volume. The area ratio itself can be

estimated from the average ratio of the number of counts ( $N_A/N_B$ ) of points placed in a systematic (or random) way on to the images of the sections ( $N_A$  is the number of points which fall inside structure A, and  $N_B$  the number of points within the image of structure B). Several other, but not all, figures of interest can be estimated using similar stereological methods. Increasingly efficient stereological procedures have recently been developed for obtaining less biased estimations, even for anisotropic structures and objects [60, 61].

## 2.4.4 Mathematical-Morphology Approach to Image Analysis

As for image processing, mathematical morphology provides a large set of specific tools for image analysis. Some details of a limited number of these tools are given below.

### 2.4.4.1 Granulometry

The case of images composed of a large number of small (possibly overlapping) objects has been mentioned previously. Suppose that we want to compute the size-distribution histogram of these objects. Apart from the possibility of characterizing each object individually, a more fruitful approach consists of working with the whole image. This can be done by opening granulometry: when performing opening with a structural element of size  $\lambda$ , objects with a smaller size than  $\lambda$  are suppressed. Therefore, performing opening with structuring elements of increasing

size acts like a sieve. It can be shown [37, 63] that the size-distribution histogram can be obtained as

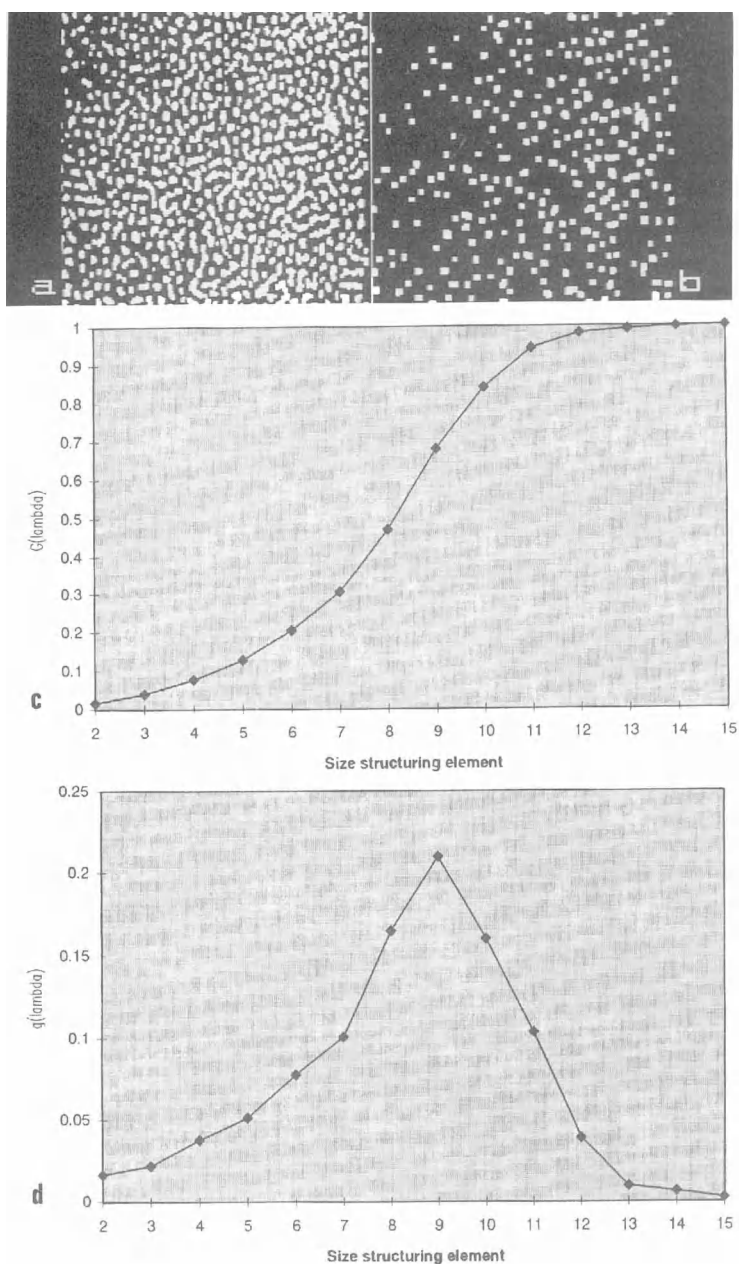
$$g(\lambda) = \frac{dG(\lambda)}{d\lambda} \quad (13)$$

where  $G(\lambda) = [i - O_\lambda(i)]/i$  and  $O_\lambda(i)$  is the open image with a structuring element  $\lambda$  (see Fig. 9). It should be noted that opening granulometry can be performed on binary images as well as on gray-level images [62, 63].

This approach of characterizing complex images by a statistical parameter can be largely extended to other parameters. We then reach the field of random image models [64, 65]: from the random sets theory [66], numerous random function (RF) models (such as the Boolean RF model, the dead-leaves RF, the alternate sequential RF, the reaction-diffusion RF or the dilution RF) can be built. The main purposes of these scalar or multivariate models are: to sum up the microstructure in a limited number of parameters; to predict morphological properties; and to provide a means of simulation. They are applied, for example, in studies of powders, fibers, aggregates, microsegregations, multiphase microstructures and multi-spectral mapping.

### 2.4.4.2 Distance Function

Image analysis may concern the spatial repartition of objects within regions. One useful tool in this context is the distance function. Assuming that a region is defined, the aim of the distance function is to define the distance of any interior pixel to the closest boundary. This can be done by eroding the binary region



**Figure 9.** Illustration of the granulometry procedure. Examples of intermediate results obtained after opening the image in Fig. 8d by a structuring element of size (a) 5 and (b) 9 pixels, respectively (images contain  $512 \times 512$  pixels). (c) Curve representing  $G(\lambda)$ . Curve representing  $g(\lambda)$ .

iteratively. The distance function is built by assigning to each pixel the number of erosions needed to attain it. Once the distance function has been computed,

concentric zones can be defined easily. Computing the watershed of the distance function may also help in separating touching particles.

**2.4.4.3 Skeleton**

The skeleton of an object (also called the medial axis transform) is the set of pixels which correspond to the midline of that object. It is generally composed of a trunk and branches. It can be described by several attributes (number of branches, end-points, triple points, loops, etc.). It can be shown that the skeleton concentrates the information on the object into this limited set of features, and can thus be used for pattern recognition. It can be computed according to several algorithms (iterative conditional erosion, regional maxima of the distance function).

The skeleton by influence zones (SKIZ) is the skeleton of the surroundings of one or several objects, which can be obtained by negating the binary image. It can be used for separating touching features, or for defining the zones of influence of objects. It is also called the Voronoï

diagram (see Fig. 10). A recent extension of this approach is the calypter [67].

**2.5 Processing/Analysis of Image Series**

As stated in the Introduction, image sequences are increasingly becoming recorded in addition to individual images. We will consider two groups of applications in this very important extension to imaging: the three-dimensional reconstruction of objects and the analysis of a spectral/temporal/spatial image series.

**2.5.1 Three-Dimensional Reconstruction**

For a long time imaging methods have suffered from the fact that a two-dimensional

**Figure 10.** Illustration of some tools of binary mathematical morphology. (a) Binary image containing several objects. (b) Distance function, with the object contours overlaid. (c) Skeletons of the objects. (d) SKIZ of the objects, representing their zones of influence.

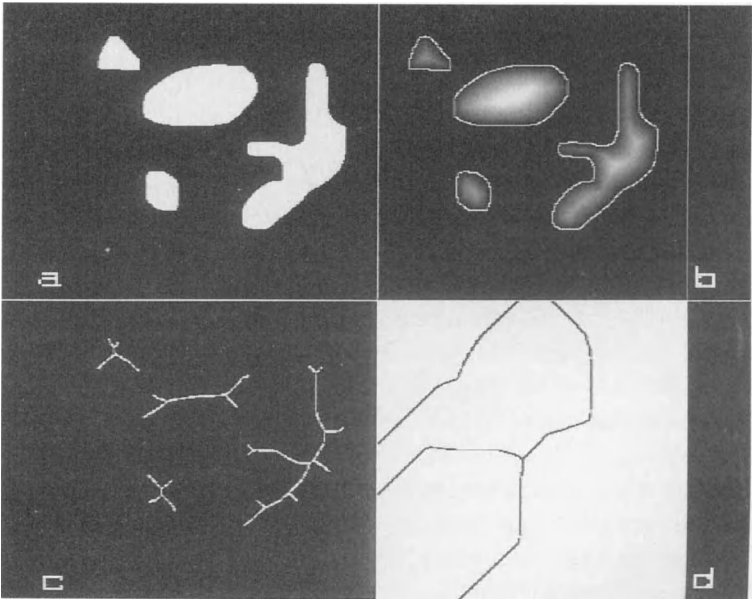


image can only give a limited representation of the three-dimensional world. Nowadays, things are changing rapidly, since for most transmission imaging methods tools have been developed for performing the three-dimensional reconstruction. Of course, several images must be recorded for this purpose. Three main groups of methods are available: serial sectioning, stereoscopy, and microtomography.

### 2.5.1.1 Serial Sections

With this approach, many images of longitudinal sections of the specimen are recorded. The way in which these images are obtained depends on the depth of field of the imaging instrument. In optical or confocal microscopy, it is possible to focus at different depths in the specimen and to record 'optical' sections. In transmission electron microscopy, this is not possible and the only way to obtain serial sections is to cut the specimen ('mechanical' sections). In SIMS, the principle consists of eroding the specimen and recording images of the successive eroded slices.

The principle of three-dimensional reconstruction from serial sections is very simple: it consists of stacking the different digital images into a digital volume, taking into account the height increment ( $\Delta z$ ) between images and verifying that no lateral shift occurred during image acquisition.

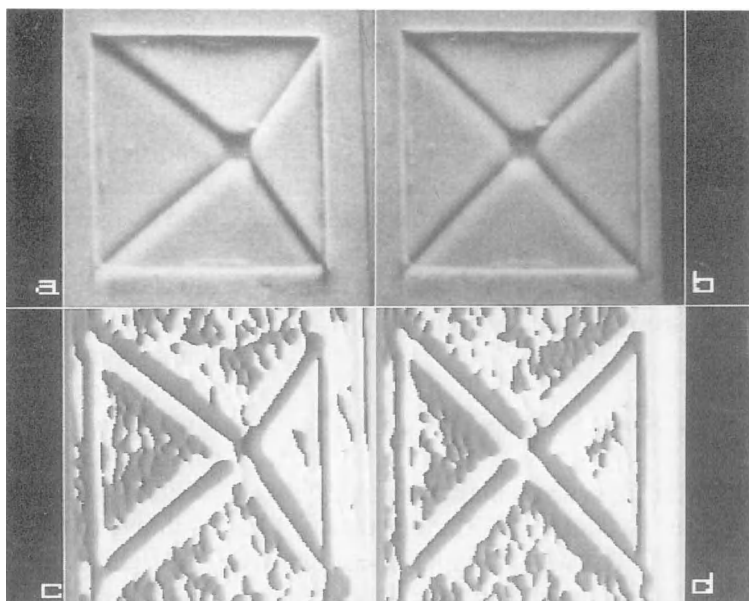
### 2.5.1.2 Stereoscopy

Instead of looking 'through' the object, another possibility of recording information relevant to the third dimension is to look at the object from different

viewpoints'. This is done in stereoscopy and microtomography. In stereoscopy (mainly used in transmission electron microscopy), two images of the object are recorded for two directions of illumination relative to the specimen [68]. Within these two images (which constitute a stereopair), the distance between two details situated at different depths is different (the difference is called the *parallax*). This coded three-dimensional information must be decoded a posteriori [69]. This can be done either analogically or by digital means. Analog modes consist of making use of the human visual system: one image of the stereopair is shown to the left eye and the other one to the right eye. The human brain reconstitutes the third dimension and the object can be seen in relief. This approach can be carried out using either a stereoscope or computer procedures and a display screen (anaglyph procedure with red/green spectacles, polarized light procedure, etc.).

Digital three-dimensional reconstitution consists of computing the height ( $z$ ) of any object detail from its coordinates ( $x_1, y; x_2, y$ ) in the stereopair ( $y$  is the tilt axis and must be common to both images: the epipolar constraint). When images are simple (i.e., contain only a few well-separated object details), correlation procedures allow us to put these homologous details into correspondence and thus to compute the parallax automatically. For more complicated images, most of the reconstructions that have been made to date in microscopy have been performed interactively: the user has to define (with a graphic cursor) homologous details in the two images, a procedure which is rather tedious. However, according to recent progress described in the image processing

**Figure 11.** Illustration of three-dimensional reconstruction by stereoscopy. (a, b) Stereopair of an etchpits, recorded in MEB (courtesy of S. Simov, Institute of Solid State Physics, Sofia; for more details see [69]). From the coordinates of specific features (corners), specified manually or detected automatically in the two images, the third coordinates (height) can be estimated, as can the angles between segments and planes. From these angles, the crystallographic indices of the faces can be deduced. (c, d) Illustration of the phase (in the horizontal direction) of images (a) and (b). Comparing the phases of the left and right images allows us to deduce the local parallax, and to build the elevation map (unpublished results).



literature (see for example the phase method [70, 71] for computing disparity), we can hope that in the near future three-dimensional maps will be built automatically from a stereopair (see Fig. 11).

### 2.5.1.3 Microtomography

In transmission imaging instruments, which can be equipped with a goniometer stage (transmission electron microscopy (TEM), X-ray microtomographs), it is possible to perform at the microscopic level the equivalent of the medical X-ray scanner. Experimentally, a tilt series (i.e., a large number of images corresponding to as many different orientations of the specimen as possible) has to be recorded.

In a sense, microtomography can be understood as a generalization of stereoscopy (as  $N$  views of the object are recorded, instead of two), but the philosophy is also different, because the gray-level distribution within the object is restored, instead of only the geometrical position of the details.

Numerous algorithms for performing this reconstruction have been suggested [72–74]; here we give just a few indications about two of them. We consider a single tilt axis geometry (images are recorded when tilting the specimen around a horizontal  $y$  axis) and a parallel-beam illumination system (for a description of a cone beam geometry see this Volume, Chap. II Sec. 3). In this configuration, vertical planes ( $x, z$ ) of the object are reconstructed

from homologous lines in the images. The general reconstruction principle is that of back-projection: since images are projections (and integration) of the object function along the illumination direction, the object function must be a combination of the back-projections of the images along their respective illumination directions. The problem is to know how to combine these different back-projections.

One approach (originating from medical imaging) is the filtered back-projection: the projection/back-projection process can be characterized by a (low-pass) linear transfer function. Thus the 'true' reconstruction consists of a filtered back-projection: during the course of back-projection, a high-pass filter performing the inverse of the (low-pass) projection filtering is applied. A variant of this approach can be used for crystalline materials for which very high resolution three-dimensional reconstructions can be obtained (see e.g., [75]).

Another approach consists of iteratively refining the reconstructed object by comparing experimental images with projections computed from the object reconstructed in the previous step (note the analogy with iterative methods for image restoration; see Sec. 2.1.1 in this Chapter). An example of the application of this approach to X-ray microtomography is given in this Volume, Chap. II Sec. 3. Numerous variants of this approach have been (and still are) suggested, the aim being to reduce as much as possible the imperfections of the reconstruction (the most important of which is its anisotropy, which comes from the fact that the goniometer stages generally allow only partial coverage of the solid angle).

#### 2.5.1.4 Three-Dimensional Display

Whatever the method used for the three-dimensional reconstruction, the next problem is to display the three-dimensional digital volume, a problem which is much less trivial than displaying images. Fortunately, to do this we can take advantage of the advances made in the field of image synthesis. Many methods are thus now available for efficient rendering of three-dimensional images: surface rendering, volume rendering, ray-tracing [76], and other new methods. However, whatever the performance of a three-dimensional static rendering approach, the best way to render the whole information contained in three-dimensional digital data sets, is to perform dynamic rendering (i.e., to produce movies). These movies can now be played easily on computer screens.

### 2.5.2 Processing and Analysis of Spectral, Temporal and Spatial Image Series

Besides image sequences for image reconstruction, other types of sequences are also produced for other purposes. Some examples are:

- (a) Filtered images recorded at different energy losses of the incident electrons, in order to deduce the chemical map of the specimen (see this Volume, Chap. IV Sec. 1.3), constitute an example of spectral imaging.
- (b) Time series of images recorded in order to study the dynamical behavior of a specimen.
- (c) Multielemental maps of the same specimen. These can be recorded



using X-ray diffraction (XDS), electron energy loss (EELS), SIMS or Auger electron (AES) spectroscopies, for instance.

- (d) Very high resolution electron microscopy of crystalline specimens, recorded with the aim of deducing chemical concentration variations across an interface (subimages of crystalline units play the role of individual images).
- (e) Multimodality imaging, where the same object area is analyzed using different imaging techniques.

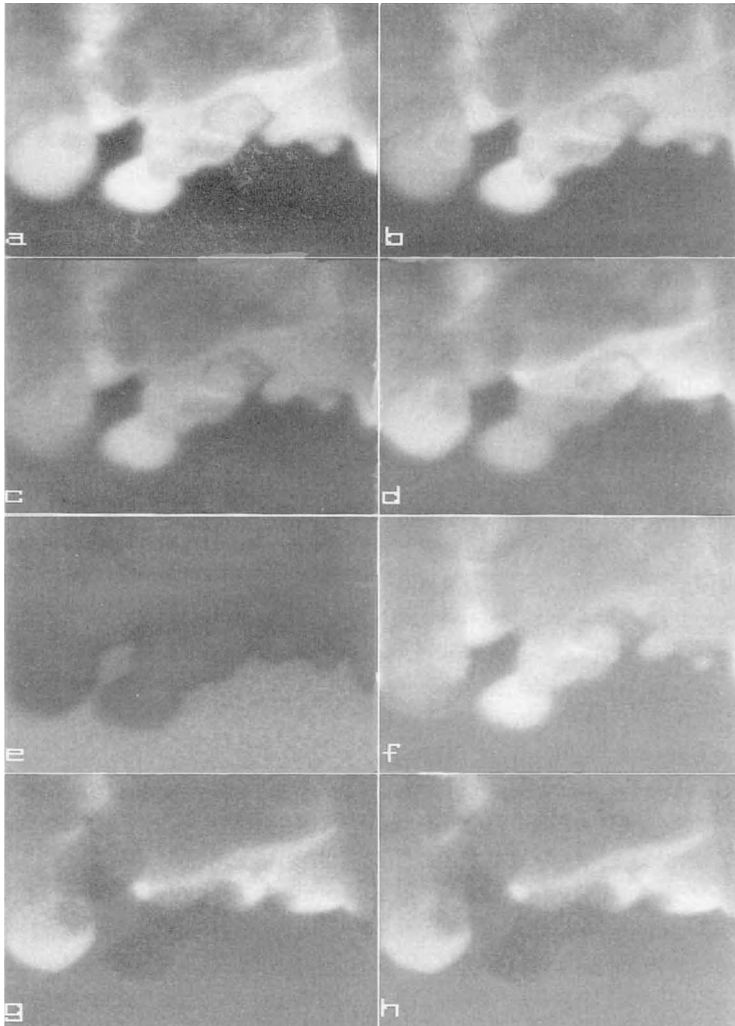
In disregarding practical experimental problems, all these series have in common the fact that every pixel is now characterized by several values (one in each of the  $N$  images of the series). Each pixel can be described by a vector with  $N$  components, which represent either different attributes or the evolution of an attribute (as a function of time, energy, space coordinate, etc.). How do we analyze (or process) such data sets? Several groups of methods are available.

One method follows the *modelization approach*: a model is assumed (or tried) for describing the evolution of any pixel. The data for each pixel are then fitted to this model and the parameter(s) of the model deduced. Since the parameter value(s) is obtained for any pixel, new (synthesized) image(s) can be built. Parametric images are thus obtained, which concentrate the information contained within the whole experimental series, into one (or a few) image(s), that are more easily interpretable. One example of such an approach is the 'standard' method used for processing electron-energy-loss image series, in order to obtain the chemical map

of an element [77, 78]. Time series can be modeled in a similar way.

Another group of methods is multivariate statistical analysis (MSA). Its aim is to extract the significant information from large data sets, while discarding redundant contributions. This is done by decomposing the total 'useful' information into a linear combination of 'basic' information components ('basic' essentially means simple and uncorrelated). For this purpose, the variance-covariance matrix ( $\mathbf{Y}$ ) of the multidimensional data set is first built:  $\mathbf{Y} = \mathbf{X}^t \cdot \mathbf{X}$ , where  $\mathbf{X}$  is the (centered) data matrix and  $\mathbf{X}^t$  the transposed matrix. The variance-covariance matrix contains the variances of the different images (along its diagonal) and the covariances of the different image pairs. The eigenvalues and eigenvectors of this matrix are first computed. The eigenvectors associated with the decreasing eigenvalues correspond to new axes of representation. The new space of representation is orthogonal (uncorrelated components) and the number of useful components (i.e. those containing interpretable information) is generally smaller than the number of initial images. Several tools are available as aids to interpreting the factorial decomposition: the scores of images on the different factorial axes, and the scores of pixels on these axes (which can be visualized as factorial images, or eigenimages). From this first step (decomposition and interpretation), other approaches can follow:

- (a) *Filtering and reconstitution*. Once the meaning of the factorial axes has been identified, an image series can be reconstituted by retaining only the 'useful' axes. This is an efficient method for performing multivariate



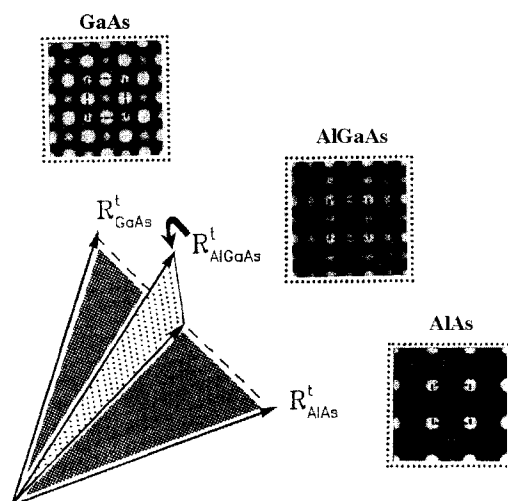
**Figure 12.** Illustration of different approaches for mapping chemical concentration from a sequence of electron energy filtered images (see this Volume Chap. 4 Sec. 1.3 for more details). Sequence of experimental images, recorded with a scanning transmission electron microscope and an electron-energy spectrometer at energy losses of (a) 82, (b) 99, (c) 114 and (d) 129 eV (courtesy of C. Colliex). The specimen is a Co/CeO<sub>2</sub> catalyst. The aim of the experiment was to map the Ce (and also Co) distribution. (g) Characteristic map of cerium. This map was obtained by subtracting the estimated background at 129 eV (f) from the experimental image at 129 eV (d). The background was estimated by extrapolating the scores of the pre-edge background images (a–c) and combining the extrapolated score with the first factorial image (e). (h) Characteristic map of cerium obtained using the standard procedure (modelization of the background according to the  $A \cdot E^{-R}$  model). The similarity between the two results (g) and (h) means that the power-law model is valid in this case. But the MSA result does not assume any model. This could be an advantage in other situations. (Reproduced from [79] by permission of *Ultramicroscopy*.)

noise filtering for instance, since in the decomposition, noise is often rejected in the last factorial components.

- (b) *Interpolation and extrapolation.* Instead of reconstituting the images with the initial image scores, it is also possible to reconstitute them with new coefficients. When these new values are deduced from the initial ones by interpolation or extrapolation, interpolated or extrapolated images are easily obtained [79, 80].

MSA is now increasingly applied to EDX [81–84], EELS [85–88] (see Fig. 12), Auger microanalysis [89], chemical mapping by HREM [90], and SIMS [91].

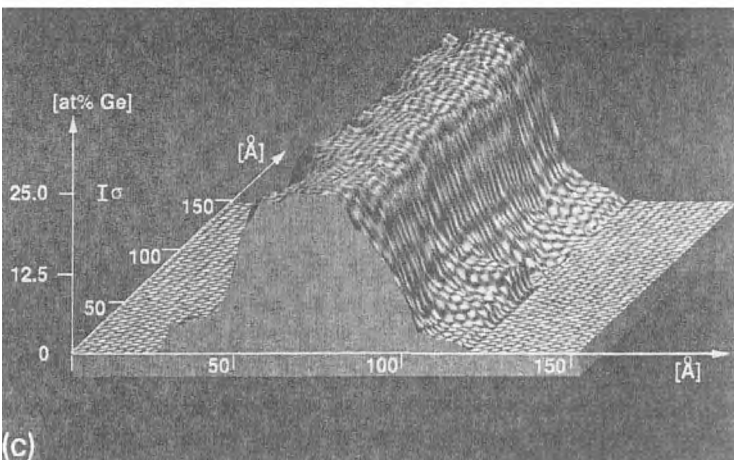
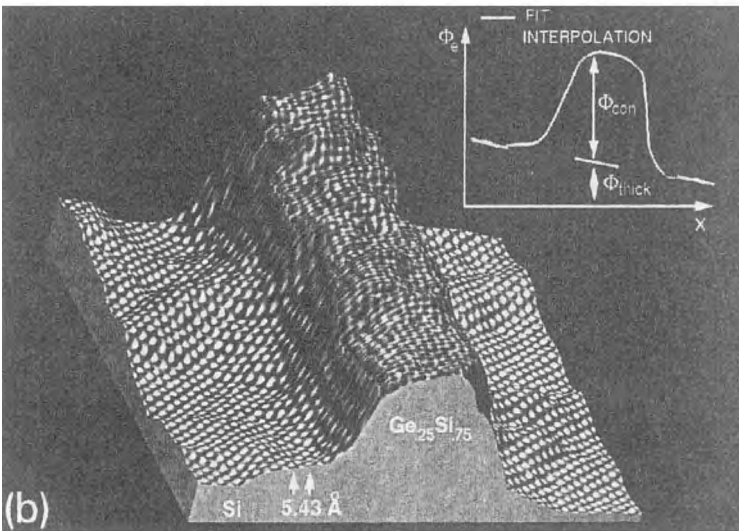
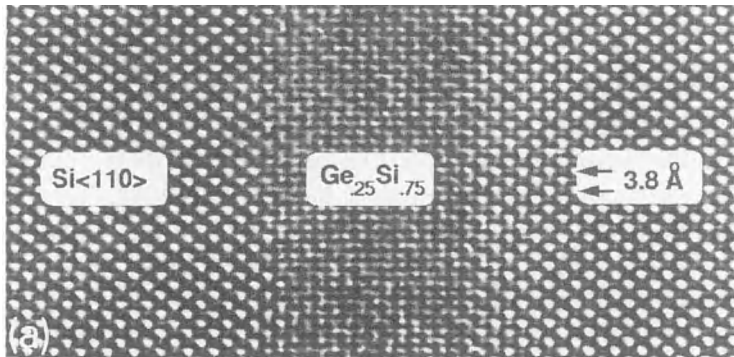
Another group of methods for the processing of image series is connected with pattern recognition. As every pixel is described by a vector of attributes, we can apply to this set of vectors the methods of pattern recognition, which consist of computing distances between vectors, angles between vectors, automatic clustering, etc. An example of this approach is the method used for quantifying the chemical-content change across an interface for HREM (lattice imaging) of crystals [92]: the image of a unit cell is represented by a multidimensional vector, the components of which are the gray levels of the different pixels that compose the unit cell (say,  $30 \times 30 = 900$  components). If we assume that two vectors are available as reference vectors (those vectors which correspond to average unit cells far from the interface, for instance, where the chemical composition is known), indicators can be built which relate an unknown vector to these reference vectors. Thus, provided experimental conditions (specimen thickness and defocus) are carefully chosen so that these



**Figure 13.** Principle of quantitative lattice imaging: lattice images are considered as multidimensional vectors. In general, vectors corresponding to columns with unknown composition (here  $\text{Al}_x\text{Ga}_{1-x}\text{As}$ ) do not lie in the plane defined by reference vectors (here GaAs and AlAs), but the projection on this plane allows the composition parameter  $x$  to be deduced. (Reproduced from [92] by permission of *Ultramicroscopy*.)

indicators can be related linearly to the concentration variation, the concentration at any point of the interface can be determined. This procedure is illustrated in Fig. 13. Recently, it has been shown that this type of method can be extended to measure the projected atomic potential, with no knowledge of the imaging conditions [93]. This approach, named QUANTITEM (quantitative analysis of the information from transmission electron micrographs) is illustrated in Fig. 14.

Another example is in the preliminary attempt to perform automatic correlation partitioning (ACP) from multielemental maps, that is to define automatically regions with homogeneous composition within a specimen for which several composition maps have been recorded. When only two maps have been recorded, the



**Figure 14.** Illustration of the QUANTITEM procedure. (a) Experimental  $\langle 110 \rangle$  lattice image of a Si/GeSi/Si structure. (b) Map of phase angle over the sample. (Inset) Schematic representation of the way in which QUANTITEM interpolates the sample thickness over regions of unknown composition, and separates the phase angle into parts according to thickness change and composition change. (c) Composition map of the Si/GeSi/Si structure (height represents composition). (Reproduced from [93] by permission of *Ultramicroscopy*.)

usual way to proceed consists of computing and displaying the scatter diagram (or scatterplot), where the content of one image is plotted as a function of the content of the other image [94–96]. The different clusters can then be selected interactively and parts of the original images (with selected properties) can be restored (interactive correlation partitioning (ICP) [97]). The purpose of ACP is to generalize the procedure to more than two maps, and towards automation. In a preliminary attempt [98], we used methods of automatic clustering (the  $K$ -means approach) and of image segmentation (multivariate region-growing approach) for this purpose. Here again, further developments can benefit from recent advances in the general domain of pattern recognition (fuzzy logic, etc.).

## 2.6 Conclusion

Image processing now plays a big part, at different levels, in the interpretation of images recorded with microscopes. The human visual system is able to perform qualitative image interpretation in most situations, but an important exception to this is high resolution microscopy, where the transfer of the information from the object wavefunction to the image intensity is very sensitive to the imaging conditions. However, when it is necessary to go from a qualitative description to a quantitative interpretation, a number of intermediate steps have to be followed, ranging from image preprocessing to image segmentation and image analysis. Moreover, one single image is often insufficient to characterize an object, and image series (focus

series, tilt series, depth series, time series, energy series, etc.) are thus increasingly being produced. In this case, the need for computer algorithms to extract the useful information is even greater.

Image processing is still an evolving discipline. During the last 20 years tremendous progress has been made in different directions. New tools are becoming available every day. In addition, the theory of image processing is also being progressively improved; for example, the theory of image algebra [99–101], which is able to establish links between tasks as different as image processing by convolution, mathematical morphology, three-dimensional reconstruction and multidimensional filtering. Of course, such progress is also facilitated by the new possibilities offered by the improved capabilities of personal computers and workstations.

Perhaps the need to moderate this optimistic view about the development of image processing comes from that fact that it often takes a long time for the new possibilities, discovered by researchers in the domain of ‘pure’ image processing, to reach the larger community of microscopists.

## 2.7 References

- [1] W. K. Pratt, *Digital Image Processing*, Wiley, New York, 1978.
- [2] R. C. Gonzalez, P. Wintz, *Digital Image Processing*, Addison-Wesley, New York, 1987.
- [3] A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice Hall, Englewood Cliffs, NJ, 1989.
- [4] J. C. Russ, *Computer-Assisted Microscopy*, Plenum Press, New York, 1990.
- [5] D. Gabor, *Lab. Invest.* **1965**, 14, 63.

- [6] A. Beghdadi, A. Le Negrat, *Computer Vision, Graphics and Image Processing* **1989**, 46, 62.
- [7] P. Perona, M. Malik, *IEEE Trans. Pattern Analysis Machine Intelligence* **1990**, 12, 629.
- [8] M. Unser, *Signal Proc.* **1990**, 20, 3.
- [9] J. Maeda, K. Murata, *Optics Commun.* **1986**, 59, 11.
- [10] C. Daly, C. Lajaunie, D. Jeulin, *Geostatistics* **1988**, 1, 749.
- [11] C. Daly, D. Jeulin, D. Benoit, *Scanning Microsc.* **1992**, Suppl. 6, 137.
- [12] R. H. Bates, M. J. McDonnell, *Image Restoration and Reconstruction*, Clarendon Press, Oxford **1986**.
- [13] N. Wiener, *The Extrapolation, Interpolation and Smoothing of Stationary Time Series*, Wiley, New York **1949**.
- [14] S. Kawata, Y. Ichioka, *J. Opt. Soc. Am.* **1980**, 70, 768.
- [15] W. O. Saxton, *Computer Techniques for Image Processing in Electron Microscopy*, Academic Press, New York **1978**.
- [16] P. W. Hawkes (Ed.), *Computer Processing of Electron Micrographs*, Springer, Berlin **1980**.
- [17] P. Schiske, in *Image Processing and Computer-aided Design in Electron Optics* (Ed.: P. W. Hawkes), Academic Press, London **1973**, 82–90.
- [18] W. Coene, A. Janssen, M. Op de Beeck, D. Van Dyck, *Phys. Rev. Lett.* **1991**, 69, 37.
- [19] D. Van Dyck, M. Op de Beeck, W. Coene, *Optik* **1993**, 93, 103.
- [20] A. I. Kirkland, W. O. Saxton, K. L. Chau, K. Tsuno, M. Kawasaki, *Ultramicroscopy* **1995**, 57, 355.
- [21] A. Thust, M. Lentzen, K. Urban, *Ultramicroscopy* **1994**, 53, 101.
- [22] S. J. Pennycook, D. E. Jesson, M. F. Chisholm, A. G. Ferridge, M. J. Seddon, *Scanning Microsc.* **1992**, Suppl. 6, 233.
- [23] . Downing, *Scanning Microsc.* **1992**, Suppl. 6, 43.
- [24] J. M. Rodenburg, B. C. McCallum, *Scanning Microsc.* **1992**, Suppl. 6, 223.
- [25] D. J. Keller, F. S. Franke, *Surf. Sci.* **1993**, 294, 409.
- [26] N. Bonnet, S. Dongmo, P. Vautrot, M. Troyon, *Microsc. Microanal. Microstruct.* **1994**, 5, 1.
- [27] R. Haralick, L. Shapiro, *Scanning Microsc.* **1988**, Suppl. 2, 39.
- [28] P. K. Sahoo, S. Soltani, A. K. C. Wong, Y. C. Chen, *Computer Vision, Graphics and Image Processing* **1988**, 41, 233.
- [29] C. A. Glasbey, *Computer Vision, Graphics and Image Processing* **1993**, 55, 532.
- [30] R. Deriche, *IEEE Trans. Pattern Analysis Machine Intelligence* **1990**, 12, 78.
- [31] J. Shen, S. Castan, *Computer Vision, Graphics and Image Processing: Graphical Models Image Processing* **1992**, 54, 112.
- [32] M. Petrou, *Adv. Electron. Electron Phys.* **1994**, 88, 297.
- [33] S. L. Horowitz, T. Pavlidis, *J. Assoc. Comput. Machine* **1976**, 23, 368.
- [34] S. Chen, W. Lin, C. Chen, *Computer Vision, Graphics and Image Processing: Graphical Models Image Processing* **1991**, 53, 457.
- [35] M. Kass, A. Witkin, D. Terzopoulos, *Int. J. Comput. Vis.* **1988**, 1, 321.
- [36] J. Serra, *Image Analysis and Mathematical Morphology*, Academic Press, London **1982**.
- [37] M. Coster, J. L. Chermant, *Précis d'Analyse d'Images*, CNRS, Paris **1989**.
- [38] E. R. Dougherty, *Mathematical Morphology in Image Processing*, Dekker, New York **1992**.
- [39] D. Jeulin, *Scanning Microsc.* **1988**, Suppl. 2, 165.
- [40] S. Beucher, *Scanning Microsc.* **1992**, Suppl. 6, 299.
- [41] S. Beucher, F. Meyer, in *Mathematical Morphology in Image Processing* (Ed.: E. R. Dougherty) Dekker, New York, **1992** 433.
- [42] S. O. Belkasim, M. Shridhar, M. Ahmadi, *Patt. Rec.* **1991**, 24, 1117.
- [43] L. M. Luerkens, J. K. Beddow, A. F. Vetter, *Powder Technol.* **1982**, 31, 209.
- [44] L. M. Karlsson, A. Liljeborg, *J. Microsc.* **1994**, 175, 186.
- [45] J. C. Russ, D. S. Bright, J. C. Russ, T. M. Hare, *J. Comput. Assist. Microsc.* **1989**, 1, 3.
- [46] E. R. Davies, *Machine Vision*, Academic Press, London **1990**.
- [47] J. Frank, in *Computer Processing of Electron Micrographs* (Ed.: P. W. Hawkes), Springer, Berlin **1980**, 187.
- [48] R. O. Duda, P. E. Hart, *Pattern Recognition and Scene Analysis*, Wiley, New York **1973**.
- [49] R. M. Haralick, *Proc. IEEE* **1979**, 67, 786.
- [50] K. Laws, *Textured Image Segmentation*, Report 940, USC Image Processing Institute, Los Angeles, CA **1980**.
- [51] M. Unser, M. Eden, *IEEE Trans. Syst. Man Cybern.* **1990**, 20, 804.
- [52] A. K. Jain, F. Farrokhnia, *Patt. Rec.* **1991**, 24, 1167.
- [53] B. Mandelbrot, *The Fractal Geometry of Nature*, Freeman, San Francisco **1982**.
- [54] M. Aguilar, E. Anguiano, F. Vasquez, M. Pancorbo, *J. Microsc.* **1992**, 167, 197.
- [55] A. P. Pentland, *IEEE Trans. PAMI* **1984**, 6, 661.
- [56] J. C. Russ, *J. Comput. Assist. Microsc.* **1990**, 2, 249.

- [57] S. Peleg, J. Naor, R. Hartley, D. Avnir, *IEEE Trans. Pattern Analysis Machine Intelligence* **1984**, 6, 518.
- [58] E. E. Underwood, *Quantitative Stereology*, Addison-Wesley, Reading, **1970**.
- [59] J. C. Russ, *Practical Stereology*, Plenum Press, New York **1986**.
- [60] L. M. Cruz-Orive, *Acta Stereol.* **1987**, 6, 3.
- [61] H. J. Gundersen, *Acta Stereol.* **1987**, 6, 173.
- [62] M. Prod'homme, L. Chermant, M. Coster, *J. Microsc.* **1992**, 168, 15.
- [63] M. Prod'homme, M. Coster, L. Chermant, J. L. Chermant, *Scanning Microsc.* **1992**, Suppl. 6, 255.
- [64] D. Jeulin, *Signal Proc.* **1989**, 16, 403.
- [65] D. Jeulin, *Scanning Microsc.* **1992**, Suppl. 6, 121.
- [66] G. Matheron, *Random Sets and Integral Geometry*, Wiley, New York **1975**.
- [67] E. Pirard, *J. Microsc.* **1994**, 175, 214.
- [68] S. K. Gosh, *Methods Cell Biol.* **1981**, 22, 155.
- [69] I. Stoev, S. Simov, E. Simova, N. Bonnet, G. Balossier, *Sensors Actuators* **1987**, 12, 1.
- [70] D. J. Fleet, A. D. Jepson, R. M. Jenkin, *Computer Vision, Graphics and Image Processing: Image Understanding* **1991**, 53, 198.
- [71] J. Weng, *Int. J. Comput. Vision* **1993**, 11, 211.
- [72] E. Zeitler, *Optik* **1974**, 39, 396.
- [73] G. T. Herman, *Image Reconstruction from Projections: the Fundamentals of Computerized Tomography*, Academic Press, New York **1980**.
- [74] J. Frank (Ed.), *Three-Dimensional Imaging with the Transmission Electron Microscope*, Plenum Press, New York **1992**.
- [75] K. H. Downing, H. Meisheng, H.-R. Wenk, M. O'Keefe, *Nature* **1990**, 348, 525.
- [76] L. Lucas, N. Gilbert, D. Ploton, N. Bonnet, *J. Microsc.* **1996**, 181, 238.
- [77] C. Jeanguillaume, C. Colliex, P. Trebbia, *Ultramicroscopy* **1978**, 3, 137.
- [78] N. Bonnet, C. Colliex, C. Mory, M. Tence, *Scanning Microsc.* **1988**, Suppl. 2, 351.
- [79] N. Bonnet, E. Simova, S. Lebonvallet, H. Kaplan, *Ultramicroscopy* **1992**, 40, 1.
- [80] N. Bonnet, P. Trebbia, *Scanning Microsc.* **1992**, Suppl. 6, 163.
- [81] R. Browning, P. King, J. M. Paque, P. Pianetta, *Microbeam Anal.* **1990**, 199.
- [82] C. Quintana, N. Bonnet, *Scanning Microsc.* **1994**, 8, 563.
- [83] C. Quintana, N. Bonnet, *Scanning Microsc.* **1994**, Suppl. 8, 83.
- [84] P. Trebbia, J. M. Wulveryck, N. Bonnet, *Microbeam Anal.* **1995**, 4, 85.
- [85] P. Hannequin, N. Bonnet, *Optik* **1988**, 81, 6.
- [86] P. Trebbia, N. Bonnet, *Ultramicroscopy* **1990**, 34, 165.
- [87] P. Trebbia, C. Mory, *Ultramicroscopy* **1990**, 34, 237.
- [88] E. Gelsema, A. Beckers, W. De Bruijn, *J. Microsc.* **1994**, 174, 161.
- [89] M. Prutton, M. El Gomati, P. Kenny, *J. Electron Spectrosc. Relat. Phenom.* **1990**, 52, 197.
- [90] J. L. Rouvière, N. Bonnet, *Inst. Phys. Conf. Ser.* **1993**, 134, 11.
- [91] P. Van Espen, G. Janssens, W. Vanhoolst, P. Geladi, *Analysis* **1992**, 20, 81.
- [92] A. Ourmazd, F. H. Baumann, M. Bode, Y. Kim, *Ultramicroscopy* **1990**, 34, 237.
- [93] C. Kisielowski, P. Schwander, F. H. Baumann, M. Seibt, Y. Kim, A. Ourmazd, *Ultramicroscopy* **1995**, 58, 131.
- [94] C. Jeanguillaume, *J. Microsc. Spectrosc. Electron.* **1985**, 10, 409.
- [95] D. Bright, D. Newbury, *Anal. Chem.* **1991**, 63, 243.
- [96] P. Kenny, I. Barkshire, M. Prutton, *Ultramicroscopy* **1994**, 56, 289.
- [97] J. M. Paque, R. Browning, P. L. King, P. Pianetta, *Microbeam Anal.* **1990**, 195.
- [98] N. Bonnet, *Ultramicroscopy* **1995**, 57, 17.
- [99] G. Y. Ritter, *Adv. Electron. Electron Phys.* **1991**, 80, 243.
- [100] P. W. Hawkes, *Optik* **1993**, 93, 149.
- [101] P. W. Hawkes, *Scanning Microsc.* **1994**, Suppl. 8, 289.

# Special Topics





# 1 Coincidence Microscopy

## 1.1 Introduction

The impact of a single fast electron on a specimen in the electron microscope can cause several events: impact ionization accompanied by emission of an electron, subsequent decay of the excited state accompanied by the emission of an X-ray or an Auger-electron and, of course, energy loss of the primary electron. All these events may occur in combination with elastic scattering, backscattering, other secondary processes, or even absorption. The time scale on which these events follow each other is usually in the  $10^{-14}$  s range and hardly ever exceeds  $10^{-9}$  s. Since it is technically feasible to detect these events with a time resolution of the order of  $10^{-9}$  s, there is a possibility of obtaining more than one signal from a single scattering event. Such a measurement, involving two signals that coincide in time, is called a coincidence measurement. It contains more information than the two signals involved can give separately. The best way to think about this is to consider one signal as a filter for the spectrum of the other signal: this spectrum can be split up in the different parts that constitute it in the first place. For example, in the energy loss signal it is possible to select a specific excitation and then, in

the coincidence–Auger spectrum, measure the decay products of only that excitation. Or, the one signal can be secondary electrons and then, in the coincidence–energy loss spectrum, it is possible to find only the events that were responsible for the emission of a secondary electron.

Historically, the coincidence technique was developed in the field of elementary particle physics, but the power was soon recognized in atomic and molecular physics. The most common application is to determine the electron momentum distribution, for which a relatively low energy electron beam scatters on a gas and both the energy and angular direction of the two outgoing electrons is selected [1]. More recently, these (e,2e) experiments have also been performed on solids, at higher energies, but not yet with a focussed primary beam [2–6]. VanderWiel and Brion [7] have reported on near-dipole (e,2e) spectroscopy, which uses the similarity between photon and electron-induced excitations. By acquiring the Auger electrons in coincidence with certain well-defined energy losses in a gas, Ungier and Thomas [8] could decompose Auger spectra to study specific decay processes. A different form of coincidence spectroscopy measures the photoelectron and the Auger electron after absorption of an X-ray [9, 10]. Coincidence measurements in the

electron microscope have been reported [11–17], but hardly using the high spatial resolution of the microscope and not yet applied to obtain information on the specimen. After these proofs of principle, the method awaits improvements in instrumentation. One potential improvement is the development of parallel detectors with the required timing resolution; another improvement would be a more generally available high vacuum, since the most promising form of coincidence measurement involves spectroscopy of emitted electrons.

## 1.2 Instrumentation

For coincidence microscopy, an STEM is the most appropriate instrument, although it is possible to use a TEM with a small probe or, for some forms of coincidence measurement, an SEM. The two signals that are to be measured in coincidence must be detected with a high detection efficiency, since coincidence rates are usually low. More specific for the coincidence measurements, however, is that single events must be detected with a good time resolution. For the signal processing, some dedicated electronics is necessary.

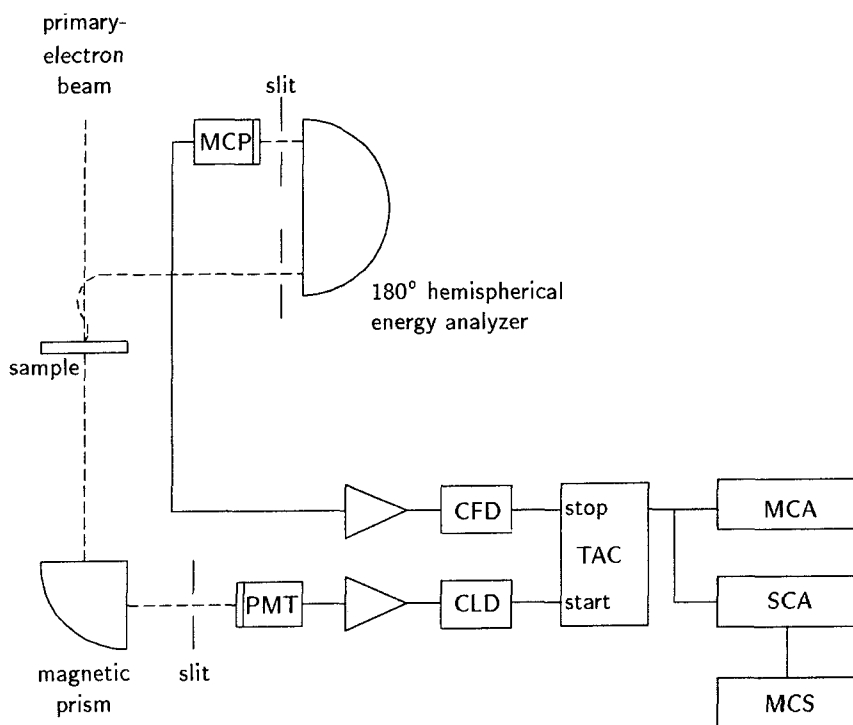
Intrinsic to electron energy loss spectroscopy (EELS) in the STEM is the very high collection efficiency, because the fast primary electrons are scattered within a small solid angle. In fact, this is the main reason for an electron microscope being the ideal instrument for coincidence measurements. If the system has a serial acquisition mode, there usually is a slit behind the spectrometer and a scintillator–photomultiplier tube (PMT) for detection. If the scintillator has a short decay time, as with plastic

scintillators although not the case with some other materials, it is possible to measure single event pulses at the exit of the PMT. These can be used for the coincidence measurement. A parallel detector behind the spectrometer is usually a photodiode array or a CCD camera. In such devices all timing information is lost. Parallel detectors for high energy electrons with sufficient timing resolution and the ability to detect up to  $10^6$  or  $10^7$  electrons per second are expected to become available in the near future.

Energy dispersive X-ray detectors are single event sensitive and are parallel in the sense that the energy of each X-ray photon is determined. However, the time resolution is very limited due to the small number of electron hole pairs that are created in the pn-junction. The time necessary to determine the X-ray energy is several hundreds of nanoseconds although, using a second signal path, the timing can be improved to several tens of nanoseconds [13]. For any further improvement it would be necessary to use a wavelength dispersive spectrometer with a position sensitive detector. The disadvantage of WDS detectors, however, lies in their lower collection acceptance angle.

Secondary electron detectors using fast scintillator PMT combinations and cathodoluminescence detectors using PMTs are well suited for coincidence measurements.

Energy spectrometers for secondary and Auger electrons in a microAuger instrument usually have single electron detectability, although the collection efficiency is relatively poor and there is no EELS signal, unless the experiment is performed in reflection [18]. The Auger spectrometers especially designed for



**Figure 1.** Experimental setup for EELS-emitted electron coincidence measurements. Key: (MCP) micro-channel plate; (PMT) photomultiplier tube; (CFD) constant fraction discriminator; (CLD) constant level discriminator; (TAC) time-to-amplitude converter; (MCA) multichannel analyzer; (SCA) single channel analyzer; (MCS) multichannel scaler.

STEM instruments [19] combine a high collection efficiency and a high spatial resolution. The secondary and Auger electrons spiral up through the magnetic field of the objective lens pole piece and are subsequently deflected to enter a spectrometer. Behind the spectrometer, the electrons are detected with a channeltron or multichannelplate, giving excellent time resolution. The multichannelplate can be combined with a position-sensitive detector for parallel detection, either based on the resistive anode technique or with multiple anodes. The latter can accept a higher count rate.

The electronics necessary for serial-serial coincidence measurements are

schematically shown in Fig. 1. Pulses from the detectors are amplified and fed into discriminators to obtain TTL or ECL level pulses. The best timing resolution is obtained when using constant fraction discriminators, which trigger the output pulse not at a predefined level, but rather at a level which is defined as a fraction of the amplitude of the input pulse. A time-to-amplitude converter, or a time-to-digital converter, creates a signal proportional to the time interval between the arrival of pulses from the two detectors. A multichannel scaler can form a time spectrum in which it is possible to find the coincidence peak on top of a false coincidence background. Setting a

window on the time axis, the events that follow within the window can be counted as a function of the energies selected in the spectrometers. This yields the coincidence spectrum. The use of parallel detectors involves a more complicated setup of the electronics and data processing [20].

### 1.3 Coincidence Count Rates

Given a primary beam current  $I$  and the cross section for a specific event  $\alpha$  to occur  $\sigma_\alpha$ , the count rate connected with this event is

$$R_\alpha = \frac{I}{e} \sigma_\alpha n t f_\alpha T_\alpha \eta_\alpha \quad (1)$$

where  $e$  is the electron charge,  $n$  is the atomic density,  $t$  the thickness of the specimen,  $f_\alpha$  the fraction of the thickness from which information can escape and contribute to the signal,  $T_\alpha$  the transmission efficiency to the detector and  $\eta_\alpha$  the efficiency of the detector. Alternatively, the countrate can be expressed in terms of escape depth when this is smaller than the thickness of the specimen [10]. If the same primary electron can also cause an event  $\beta$ , a similar equation holds for the count rate  $R_\beta$ . For example,  $\alpha$  can be the K shell excitation of carbon, with an energy loss between 300 and 305 eV,  $\beta$  can be the Auger decay emitting an electron of energy between 260 and 265 eV. If both events are measured in coincidence, the coincidence count rate is

$$R_T = \frac{I}{e} \sigma_{\text{true}} n t f_\alpha f_\beta T_\alpha T_\beta \eta_\alpha \eta_\beta \quad (2)$$

where  $\sigma_{\text{true}}$  is the cross section for the combined event  $\alpha$  and  $\beta$ . There is a possibility that two unrelated events in the

detectors occur at the same time, or more precisely, within the time window  $\tau$  which is set to define which events are coincident. The associated count rate, called false coincidence rate, is given by

$$R_F = R_\alpha R_\beta \tau \quad (3)$$

From the ratio of Eq. (2) to Eq. (3), and substituting for  $R_\alpha$  and  $R_\beta$  using Eq. (1), we find the true to false ratio

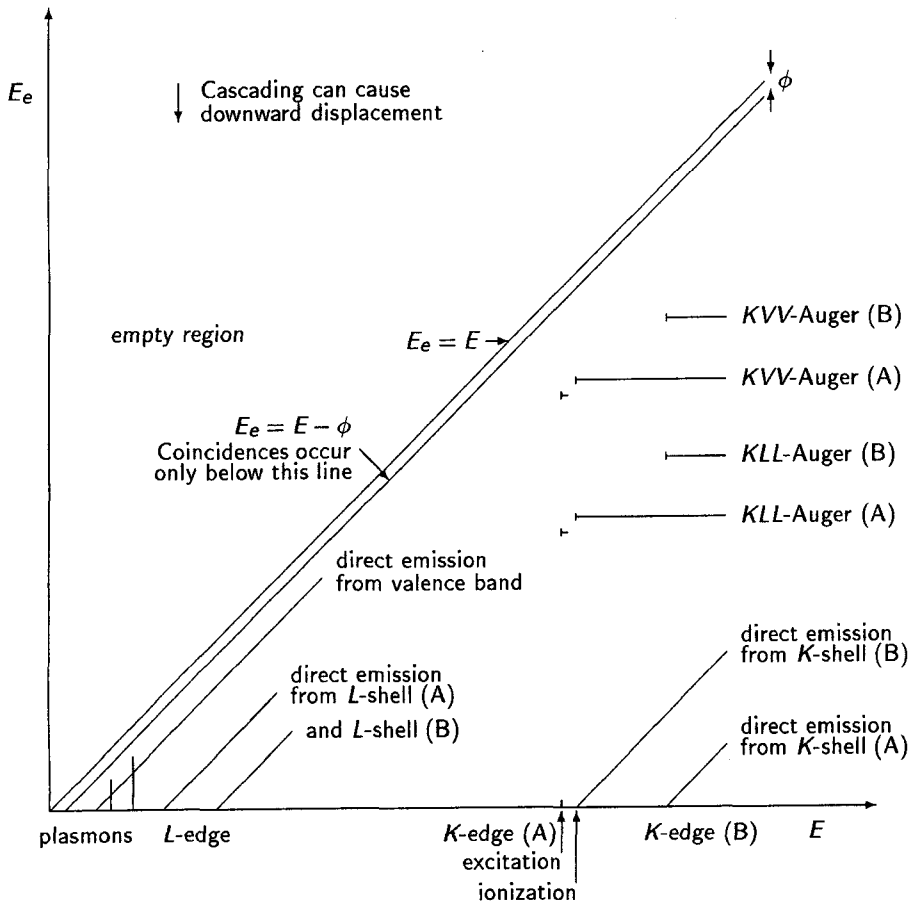
$$\frac{R_T}{R_F} = \frac{1}{n t} \times \frac{\sigma_{\text{true}}}{\sigma_\alpha \sigma_\beta} \times \frac{e}{I} \times \frac{1}{\tau} \quad (4)$$

This shows how important it is to obtain a good time resolution so that  $\tau$  can be made small. Although it seems advantageous to also make the beamcurrent small, this is somewhat misleading: the false coincidence background can easily be subtracted, so it becomes a question of statistical noise in the background subtracted spectra. From that, an optimized current follows for each experiment [21]. The numerical value of the count rate depends very much on the process under observation and on the instrumental setup. For EELS—secondary electron measurements at the plasmon excitation, with energy windows of 1 eV in both spectra, coincidence rates of the order of 1000 counts per second were obtained for a primary beam current less than 1 nA. The false coincidence background was an order of magnitude smaller.

## 1.4 Signal Combinations

### 1.4.1 EELS—Emitted Electron

Since the energy distributions of both the EELS and the emitted electrons contain



**Figure 2.** Schematic 2D EELS-emitted electron coincidence spectrum, for a specimen with two elements, A and B. Along the horizontal axis is the energy loss of the transmitted electron, along the vertical axis is the kinetic energy of the emitted electron. (From Pijper [21], reproduced by permission of the author.)

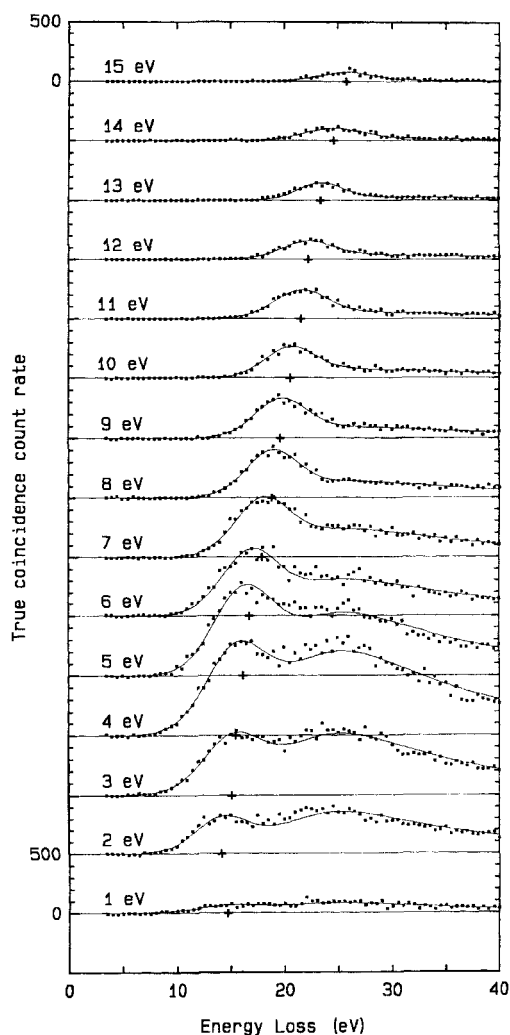
information, a coincidence measurement between these two signals is a two-dimensional spectrum. Figure 2 gives a schematic representation of such a spectrum, which can be used to discuss some of the applications of this signal combination.

The low energy part of the EELS spectrum gives information on plasmon excitation and single electron excitation from the valence band. If counts in the EELS spectrum are only accepted if in coincidence with a secondary electron, one effectively

obtains the energy loss events which are responsible for emitting a secondary electron. Voreades [11] first performed this experiment although without energy analysis of the secondary electrons. His conclusion was that most secondary electrons are produced by energy loss events of about 20 eV, but it was impossible to decide whether plasmon decay processes or interband transitions were predominantly responsible for the secondary electrons. Pijper and Kruit [14] worked with a

better EELS resolution and were able to discriminate for different secondary electron energies, thus, they could identify the surface plasmon excitation and decay as an important contributor to secondary electrons in the case of amorphous carbon films. For all other energy loss processes, including plasmon excitation, they established the validity of the stopping power rule: the probability of emitting a secondary electron is directly proportional to the amount of energy lost by the primary electron. In addition, it was clear from the spectra, of which an example is shown in Fig. 3, that direct excitation from the valence band can be seen as a feature in the 2D-spectrum that runs diagonally from the bottom-left to the top-right: the energy of the emitted electron is directly proportional to the energy lost by the primary electron. Similar experiments with EELS-secondary electron coincidence were performed by Mullejans et al. [15], and Scheinfein et al. [16]. The latter showed that in Si, it is not the plasmon which is responsible for most of the secondary electrons. Drucker et al. [22] extended the technique by also discriminating energy loss events on the basis of scattering angle and concluded that secondary electrons are more efficiently produced by large momentum transfer scattering, thus explaining the very high spatial resolution obtainable in secondary electron imaging.

Perhaps the most promising application of coincidence microscopy is virtual photoelectron spectroscopy. Because of the similarity between forward electron scattering and photoabsorption, a vertical section through the 2D coincidence spectrum resembles a photoelectron spectrum. Selecting a different energy loss at which



4

4

**Figure 3.** Energy loss coincidence spectra (shown as dots) for 15 consecutive values of the secondary electron energy, 1 eV apart. The spectra have been vertically displaced over an equivalent of 500 counts. False coincidence counts have been subtracted. Solid curves are fitted to the data with a model including a plasmon function and a Gaussian. The position of the Gaussian peak is indicated with a +.

the section is taken is equivalent with tuning to a different photon energy in a synchrotron. The positions of the photoelectron peaks shift with the selected energy loss value, indicated by the diagonal lines in Fig. 2. The advantage over photoelectron spectroscopy lies in the obtainable spatial resolution. The energy resolution is limited by the EELS resolution. Whether spatially resolved virtual-photoelectron spectroscopy will indeed become a useful technique will depend on the possibility of obtaining a parallel detection coincident spectra in an ultra-high vacuum electron microscope. It would then be relatively easy to extend the technique to select specific scattering angles of the primary electron. However, to determine also the angular distribution of the emitted electrons, in order to have a full (e,2e) experiment, seems to be much more complicated.

A normal Auger spectrum is the vertical spectrum that is obtained after integration over all possible energy losses in Fig. 2. A coincident Auger spectrum is the spectrum obtained after integration only over a selected range of energy losses. If only those energy loss events are selected which have a relatively high probability of producing an Auger electron, an improvement of peak-to-background ratio in the Auger spectrum can be expected, comparable to the improvement obtained in X-ray excited Auger spectra. This might be the most promising approach to single atom identification, since it combines a high signal-to-background spectrum with an extremely small analyzed volume. Another application of the coincidence Auger spectroscopy might be to study the decay of weakly bound states near the absorption edge in the

EELS spectrum, equivalent to the experiments of Ungier and Thomas [8] on gases. For solids, this would yield information on local chemical bonds.

Quite similar to the coincident Auger spectrum, a coincident EELS spectrum is obtained after integration over a selected range of Auger energies in Fig. 3. Again, the effect is background reduction, but now accompanied with an appreciable loss of signal, because Auger electrons are only detected if they come from the surface of the specimen.

### 1.4.2 EELS–X-ray

The first effort using coincidence techniques in the electron microscope that was directed at improving the detection limits of microanalysis used the coincidence between X-rays and energy loss events [13]. An energy loss spectrum around the Ca K-edge collected in coincidence with CaK X-rays, showed an almost complete removal of the background under the edge. A background reduction could also be obtained in the X-ray spectrum. However, due to the relatively slow EDX detector, the primary current had to be low and thus there was no improvement of the signal-to-noise ratio. Attempts by Nicholls et al. [23] were no more successful.

### 1.4.3 EELS–Cathodoluminescence

By measuring the time delay between an electron energy loss event and the corresponding cathodoluminescence event, lifetimes of excited states may be determined



[12,24]. As in the case of secondary electron production, the CL production efficiency is linearly proportional to the energy loss, except at low energies. The technique has been applied to scintillators and semiconductors. The originators of the method mention the possibility of producing submicron resolution maps of lifetimes in the vicinity of inhomogeneities in semiconductor crystals. Selecting only short lived excitations might improve the resolution in cathodoluminescence images.

#### 1.4.4 Backscattered Electron–Secondary Electron

For bulk samples, the energy loss signal must be obtained in reflection mode, or be accompanied by a backscatter event. Preliminary experiments have been performed by Kirschner et al. [17] and Cazaux et al. [18]. Further experiments are proposed by Gergely [25].

#### 1.4.5 Other Combinations

Other signal combinations are possible. For example, accepting X-rays only when they are in coincidence with secondary electrons would give surface sensitivity to the X-ray microanalysis technique. Many energy loss events give rise to more than one emitted electron, or to one electron plus a photon or X-ray. Technically it is possible to perform triple coincidence measurements, which should lead to results comparable with Auger-photo-electron coincidence experiments [9, 10].

## 1.5 References

- [1] I. E. McCarthy, E. Weigold, *Rep. Prog. Phys.* **1991**, *54*, 789.
- [2] U. Amaldi Jr., A. Egidi, R. Marconero, G. Pizzella, *Rev. Sci. Instrum.* **1969**, *40*, 1001.
- [3] R. Camilloni, A. Giardini Guidoni, R. Tiribelli, G. Stefani, *Phys. Rev. Lett.* **1972**, *29*, 618.
- [4] N. A. Krasil'nikova, N. M. Persiantseva, *Phys. Lett.* **1978**, *69A*, 287.
- [5] A. L. Ritter, J. R. Dennison, R. Jones, *Phys. Rev. Lett.* **1984**, *53*, 2054.
- [6] Chao Gao, A. L. Ritter, J. R. Dennison, N. A. W. Holzwarth, *Phys. Rev. B* **1988**, *54*, 3914.
- [7] M. J. van der Wiel, Th. M. El-Sherbini, C. E. Brion, *Chem. Phys. Lett.* **1970**, *7*, 161.
- [8] L. Ungier, T. D. Thomas, *J. Chem. Phys.* **1985**, *82*, 3146.
- [9] H. W. Haak, G. A. Sawatsky, T. D. Thomas, *Phys. Rev. Lett.* **1978**, *41*, 1825.
- [10] E. Jensen, R. A. Bartynski, S. L. Hulbert, E. D. Johnson, *Rev. Sci. Instrum.* **1992**, *63*, 3013.
- [11] D. Voreades, *Surf. Sci.* **1976**, *60*, 325.
- [12] R. J. Graham, J. C. H. Spence, H. Alexander, *Mat. Res. Soc. Symp. Proc.* **1987**, *82*, 235.
- [13] P. Kruit, H. Shuman, A. P. Somlyo, *Ultramicroscopy* **1984**, *13*, 205.
- [14] F. J. Pijper, P. Kruit, *Phys. Rev. B* **1991**, *44*, 9192.
- [15] H. Mulleijans, A. L. Bleloch, *Phys. Rev. B* **1992**, *46*, 8597.
- [16] M. R. Scheinfein, J. Drucker, J. K. Weiss, *Phys. Rev. B* **1993**, *47*, 4068.
- [17] J. Kirschner, O. M. Artomonov, A. N. Trekhov, *Phys. Rev. Lett.* **1992**, *69*, 1711.
- [18] J. Cazaux, O. Jbara, K. H. Kim, *Surf. Sci.* **1991**, *247*, 360.
- [19] P. Kruit, J. A. Venable, *Ultramicroscopy* **1988**, *25*, 183.
- [20] R. W. Hollander, V. R. Bom, C. W. E. van Eijk, J. S. Faber, H. Hoevers, P. Kruit, *Nucl. Instrum. Methods A* **1994**, *348*, 664.
- [21] F. J. Pijper, PhD Thesis, Delft University of Technology **1993**, p. 97.
- [22] J. Drucker, M. R. Scheinfein, *Phys. Rev. B* **1993**, *47*, 15973.
- [23] A. W. Nicholls, I. P. Jones, M. H. Loretto, *IOP Conf. Proc. (EMAG'85)*, **1985**, *78*, 205.
- [24] C. C. Ahn, O. L. Krivanek, *Proc. 43rd EMSA*, San Francisco Press **1985**, p. 406.
- [25] G. Gergely, *Acta Physica Hungarica* **1991**, *70*, 47.

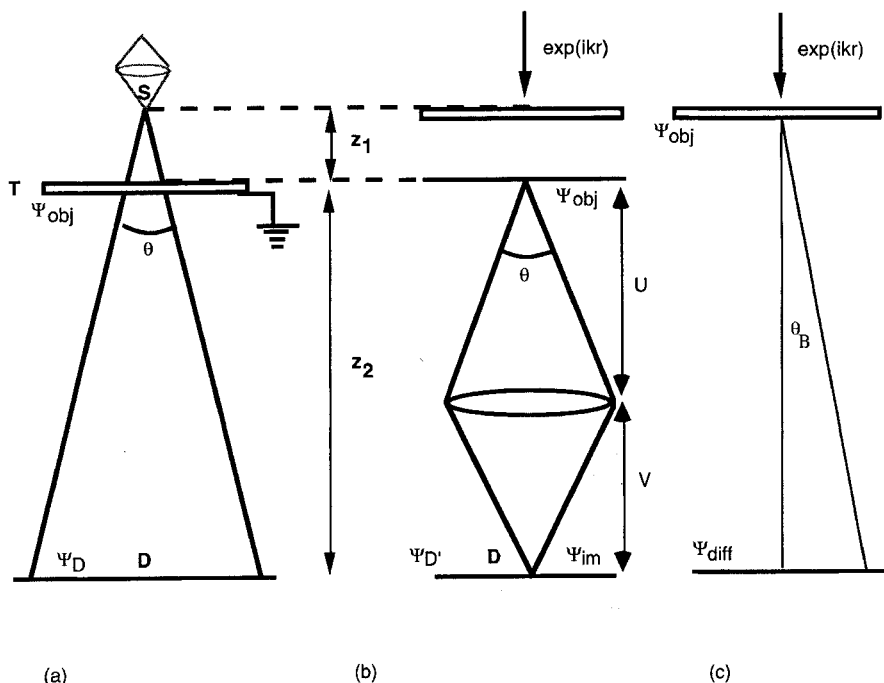
## 2 Low Energy Electron Holography and Point-Projection Microscopy

### 2.1 Introduction and History

Electron holography has its origins in the first observations of Fresnel fringes using electrons in 1940, and with shadow imaging by projection from a small source.

Shadow images cast on a wall by an opaque object placed in front of a lamp have their exact parallel in electron microscopy – we see them in the central disc of every convergent-beam microdiffraction pattern if the probe is focused slightly above the sample at  $z_1$ , as shown in Fig. 1a. If the source is small, Fresnel fringes will be seen around object edges (indicating that the image is out of focus by  $z_1$ ), and the magnification is approximately  $M = z_2/z_1$ , as shown. It is therefore not surprising that this mode was one of the first to be used to obtain electron microscope images in 1939 [1]. At that time, because of the large sources used, the troublesome Fresnel edge fringes were not seen. A year later, Fresnel fringes were observed for the first time using electrons by Boersch [2] (using the point-projection geometry), and by Hillier [3] (using the more familiar geometry shown in Fig. 1b). Boersch's remarkable pre-war paper gives the relationship between source size

and fringe spacing. Hillier interpreted his result in terms of Fresnel fringes and spherical aberration. In the recent light of De Broglie's matter-wave hypothesis and Davison, Germer and Thomson's work the significance of these results was seen as a confirmation of the wave nature of the electron. Six years after that, an entirely new interpretation was given by Gabor and Boersch's Fresnel edge fringes [4] when he named them in-line electron holograms and pointed out that these fringes might be removed by a focus-restoration procedure involving optical reconstruction. In crudest terms, it was planned to sharpen the edge of a shadow image in order to improve resolution. The shadow image of the hologram (if the source is small), and reconstruction removes the Fresnel edge fringes. For these mask-like opaque objects, the unobstructed wave passing around the edge provides a limited reference wave. Gabor extended the analysis to include partially transparent objects, and the aim of his work was to use in-line holography to eliminate the aberration of the electron lens (shown in gray in Fig. 1a) used to focus the probe to a 'point' source. Three-dimensional reconstructions and off-axis optical holography came much later. The first experiments



**Figure 1.** (a) Point-projection geometry in convergent-beam microdiffraction. The probe is focus distance  $z_1$  slightly above the sample. (b) Transmission geometry with focus error  $z_1$  ( $1/f = 1/U + 1/V$ ). (c) Diffraction pattern.

test of Gabor's ideas came in 1952 when Haine and Mulvey [5] compared the 'projection' (Fig. 1a) and 'transmission' (Fig. 1b) geometries, the equivalence of which was not fully appreciated at that time [6]. The success of this approach (usually using the 'transmission' geometry) has had to wait more than 40 years for the commercial availability of brighter electron sources and better lenses, and is described in detail in Sec. 2.2 of Chap. 4, this volume. Section 1.9 of Chap. 4, this Volume describes high-energy electron holography experiments in the projection geometry, using scanning transmission electron microscopy (STEM).

The appeal of the point-projection geometry lies in the fact that, if a sufficiently small source can be found, almost

unlimited magnification can be obtained apparently without aberrations. The aberrations of a probe-forming magnet can be avoided by using a physical electron emitter as the source, and the first experimental shadow images formed at low energy by this method appear to be those of Morton and Ramberg [7] in 1939. In 1952, researchers used a glass ultra-high vacuum system and electron field-emitter source close to a transmission electron microscope (TEM) grid, with observation at a distant screen to obtain a shadow image. This approach was later developed by Melmed and co-workers for both electron and ion point-projection microscopy. Since any electric field distribution with cylindrical symmetry forms a lens, a field-emission tip itself forms a lens.

emitting area on the surface of the tip is imaged onto a virtual source inside the tip. The aberration coefficients of such a nano-tip lens are finite, and will be discussed below. The point-projection geometry was proposed again for the atomic-resolution electron imaging of thin crystals by Cowley and Moodie in 1957 [9] with their development of the theory of Fourier imaging. Finally, the development of field-emission STEM instruments [10] has allowed the continuous development of in-line electron holography in the point-projection geometry shown in Fig. 1a, using one stationary probe position for the collection of each hologram [11–13].

Claims for atomic resolution at low voltage using the lensless, field-emission arrangement of Morton and Ramberg have recently revived interest in this technique [14]. Using specially prepared ‘single atom’ field emission tips [15], Fink and co-workers have obtained point-projection images of greatly improved quality [16], showing many Fresnel edge fringes. Since the reconstructed-image resolution is approximately equal to the source size, an atomic-sized emitter should produce atomic-resolution images; however, unless there is clear evidence of penetration of the beam through the sample, the reconstructed images will show only the outline of the object (mask reconstruction). By using Fraunhofer reconstruction methods we can test for transmission of the beam through the sample. Several similar point-projection microscopes (PPM) have since been constructed by other groups [17–19]. This low-voltage approach solves the problems of lack of source brightness and lens aberrations which limited Haine and Mulvey’s work; however, the severely limited penetration of low voltage

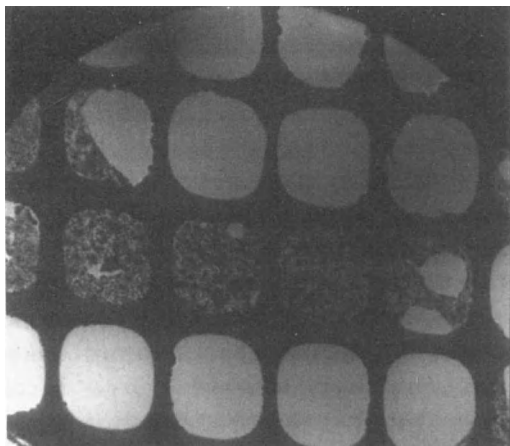
electrons through matter is a severe constraint. Solutions to this problem using either the reflection geometry or very low or higher voltage are discussed later in this section.

In a typical point-projection low voltage microscope [16], a tungsten field emitter at a potential of about  $-100$  V is placed a distance  $z_1$  (a few thousand Angstrom) from a grounded sample, which acts as the anode. A shadow image is formed on a channel plate a distance of  $z_2 = 14$  cm away, with a magnification  $M = z_2/z_1$ . Images of the fiber-optic face-plate may be recorded using a liquid-nitrogen-cooled CCD camera. Typical exposure conditions are about 0.1 s with less than 1 nA beam current at 100 V. The image (see Fig. 1) is superimposed on a point-projection field electron image of the tip (the tip emission pattern, projected from the virtual source inside the tip). This causes a slowly varying background modulation. Note that only movement of the tip relative to the sample is magnified; as in field ion microscopy (FIM), movement of the tip does not cause magnified movement of the field-emission pattern from the tip, since no movement of the virtual source relative to the surface of the tip occurs.

If the tip (and detector) are modeled as coaxial parabola, the on-axis electric field  $E(z)$  is given by

$$E(z) = \frac{U_0}{z \ln(2z_1/r)} \approx \frac{U_0}{6z}$$

if  $z_1 = 200$  nm and  $r = 1$  nm, where  $r$  is the tip radius,  $z$  is a coordinate along the axis from the focus and  $U_0$  is the beam potential. The field varies slowly with  $z_1$  but rapidly with  $r$ , and falls off very rapidly away from the tip. The virtual source size  $d \approx r/2$  [20]. Field emission from tungsten



**Figure 2.** Low magnification image of a porous carbon film obtained at ca 400 V.

requires about  $0.4 \text{ V } \text{\AA}^{-1}$  at the surface, so fixing  $z = r/2$  fixes  $U_0 \approx 12 \text{ V}$ . (A more realistic value would be  $U_0 = 100 \text{ V}$ ; accurate numerical computations for tip fields and electron trajectories can be found in Scheinfein et al. [20].) In practice then, reducing the tip-to-sample distance  $z_1$  increases the magnification (as  $1/z_1$ ), reduces the defocus (see below), and reduces the voltage needed for the same tip field and emission current.

Figure 2 shows a low magnification image of a hole-containing carbon film obtained at about 400 V, using an instrument in which the tip motion along the axis is controlled by an inchworm.

## 2.2 Electron Ranges in Matter: Image Formation

Before discussing the image (hologram) formation process it is important to determine what fraction of elastically scattered

100-V electrons traverse a thin film of, say, carbon. This, together with the source size, will determine the nature of the scattering theory required, which might be based in a transport equation (if multiple inelastic processes dominate) or perhaps on transmission low energy electron diffraction (LEED) theory if the film thickness is much less than the inelastic mean free path (IMP). Since the beam energy is less than that of the inner shell atomic ionization energies, surface plasmons, phonon excitation, valence electron and plasmon excitation are the main energy-loss mechanisms. Extensive calculations of the inelastic mean free path  $\lambda_E$  have been made in the field of photoemission spectroscopy, and in support of LEED calculations [21]. This is defined such that the intensity of the elastic portion of the electron wavefield decays on entering a solid as

$$I(z) = I_0 \exp(-z/\lambda_E) = I_0 \exp(-2\sigma V'_0 z) \\ = I_0 \exp(-zN\sigma_t) \quad (1)$$

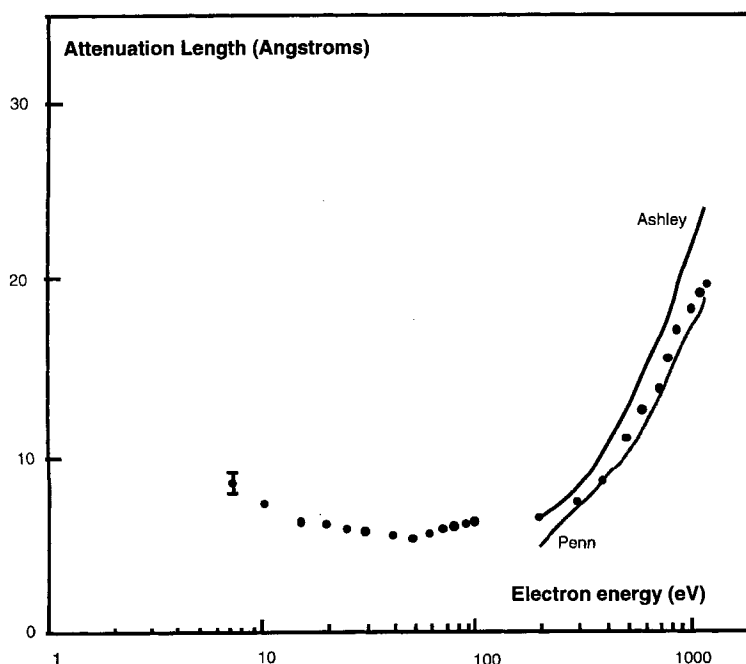
where  $\sigma = \pi/\lambda U_0$  ( $\lambda$  is the electron wavelength),  $V'_0$  is the imaginary part of the mean inner potential (in volts),  $N$  is the density of atoms, and  $\sigma_t$  is the total inelastic cross-section. Measurements and calculations covering the range below 100 V (in which the theory becomes most difficult) can be found in the literature [22]. The results generally show a minimum of a few Angstrom for  $\lambda_E$  at about 80 eV for carbon, below which  $\lambda_E$  rises steeply with the onset of ballistic transport. Thus  $\lambda_E$  may be 100 nm or more near the Fermi energy for a pure metal at room temperature. An empirical formula has been proposed [23] of the form

$$\lambda_E = AE^{-2} + BE^{\frac{1}{2}} \quad (2)$$

where  $E$  is the electron energy in eV,  $\lambda_E$  is given in Angstrom, and  $A$  and  $B$  are fitting parameters. The values  $A = 1430$  and  $B = 0.54$  for elements,  $A = 6410$  and  $B = 0.96$  for inorganic compounds, and  $A = 310$  and  $B = 0.87$  for organic compounds are given. The traditional theoretical approach has been based on a modified Bethe theory of inelastic electron scattering. Another recent approach derives values of  $\lambda_E$  from measured optical dielectric functions  $\epsilon(\omega)$  [24] using a model (such as the Lindhard dielectric function) for the dependence of scattering vector  $q$  on energy loss  $\hbar\omega$ . At low energies, exchange and correlation effects become important, and multiple inelastic scattering must be considered.

Figure 3 shows experimental measurements of the attenuation length  $L$  for low-energy electrons in the 7–1000 eV range [25]. These measurements were obtained

using free-standing thin carbon films about 4 nm thick in transmission experiments, using an electron spectrometer to isolate the transmitted elastic scattering.  $L$  may be identified with  $\lambda_E$  under the approximation that the detector includes all scattering angles. Thus they are uniquely relevant to PPM imaging of organic films. Any undetected pinhole lying under the beam (of width about 60  $\mu\text{m}$ ) would increase the contribution of elastic scattering and so make this estimate of  $\lambda_E$  (about 0.6 nm at 100 V) too large. We see that thin films of this material only become reasonably transparent to electrons at energies below about 10 V. At energies around 100 V (where much point-projection work is done), the IMP is close to its minimum value. Since experience from TEM shows that it is extremely difficult to prepare extended areas of thin films less than 0.6 nm thick, we shall



**Figure 3.** Measurements of the attenuation length  $L$  for low energy electrons in the 7–1000 eV range [23]. Dashed and continuous lines show theoretical estimates.

assume as an approximation that negligible transmission occurs, and give here the theory of point-projection image formation for masks, a mask being defined as a two-dimensional 'transmission' object whose transmission function is equal to zero (in opaque regions) or unity (in transparent regions). Such a theory may be useful for the important problem of determining the shape of molecules, if methods can be found for placing them, unsupported, across holes in carbon films. Clear experimental evidence for significant transmission of elastically scattered electrons at around 100 V is presently lacking. However, should such evidence be obtained, the theory of transmission LEED [26, 27] may be useful for edges of thin crystals, in which case Fourier imaging effects might also be observed. The theory of Fourier imaging in transmission at low voltage with multiple scattering is given by Spence and Qiau [28]. Experiments using an energy loss spectrometer attached to a PPM will clarify this point.

Figure 3 is consistent with the recent work of Shedd [29], who found that good agreement between simulated and experimental PPM images of carbon films with holes could be obtained by treating them as opaque masks containing holes, ignoring completely any partial transmission at edges. The two-dimensional patterns of fringes seen previously [14] were thereby found to be Young's fringes due to the interference between waves passing through three different pinholes, rather than atomic-resolution transmission lattice images. The two effects can easily be distinguished if the magnification (and hence the defocus) is accurately known. A more refined treatment would consider

the contribution to the image of the large fraction of electrons which lose energy in traversing the sample, and the declining efficiency of channel plate detectors with electron energy below 100 V. An energy-filtering mesh has been fitted to a PPM instrument [14].

As discussed further in Sec. 2.7 of this Chapter, our experimental results suggest that unstained, uncoated purple membrane (thickness 5 nm) is opaque at 100 V, but reasonably transparent at 1 kV, and that significant transmission occurs through samples of the lipid monolayer C16 (thickness 2.6 nm) at 100 V. We are not able to determine the fractions of elastic and inelastic scattering in the transmitted beam.

We now consider the image-formation process in PPM. First, we demonstrate that a shadow image of a mask (or a weak phase object) is equivalent to a conventional out-of-focus image, and so discuss focus restoration schemes for masks. Consider a mask illuminated by a spherical wave originating from a source at distance  $z_1$  from the mask. The transmission function for the sample is  $q(x)$ , but, for masks, we cannot assume the weak phase object approximation. Choose Cartesian coordinates with  $z$  along the beam path. Let the spherical wave incident on the sample be represented in the parabolic approximation by  $t_{z_1}(x) = \exp(-i\pi x^2/z_1\lambda)$  (with Fourier transform  $T_{z_1}(u) = C \exp(i\pi z_1 \lambda u^2)$ , where  $C$  is a complex constant) and let the electron wavefunction across the downstream side of the slab be  $\Psi_i(x)$ . Then, with  $u = \Theta/\lambda$ , where  $\Theta$  is the scattering angle

$$\psi_i(x) = q(x)t_{z_1}(x) \quad (3)$$

the detected wavefunction in the far field

will be the Fourier transform

$$\psi_h(u) = \text{FT}(\psi_i(x)) = Q(u) * T_{z_1}(u) \quad (4)$$

where the asterisk denotes convolution. Evaluating the convolution in Eq. (4), and ignoring unimportant phase factors gives

$$\begin{aligned} \psi_h(u) = & \int [Q(U) \exp(i\pi z_1 \lambda U^2)] \\ & \times \exp(-2\pi i z_1 \lambda u U) dU \end{aligned} \quad (5)$$

Now, consider

$$\begin{aligned} z_1 \lambda u U &= z_1 \lambda (\Theta/\lambda) U \\ &= z_1 \lambda (X/z_2 \lambda) U = (X/M) U, \end{aligned}$$

where  $X$  is the spatial coordinate on the detector and  $M = z_2/z_1$  is the magnification of the shadow image. Hence

$$\psi_h(X = z_2 \lambda u) = q(X/M) * t_{z_1}(X/M) \quad (6)$$

This important result establishes that, for masks (or for any thin sample for which a transmission function can be defined), the shadow image consists of an 'ideal' image which is out of focus by the tip-to-sample distance  $z_1$ . Equation (6) is identical to the expression for an out-of-focus high-resolution transmission electron microscope (HREM) image in the absence of lens aberrations. (Note, however, that for HREM plane-wave illumination is used.) However, the point-projection image has been magnified by  $M = z_2/z_1$  without the use of lenses. We have compared simulated images based on Eqs. (5) and (6) and found them to be identical. As may readily be confirmed using an optical laser and a slide transparency, increasing  $z_2$  increases the overall magnification of a shadow image, but not the focus defect, which is fixed by  $z_1$ . No assumption of periodicity has been made

for the sample. An in-focus image can only be obtained using Eq. (6) if  $z_1 = 0$ , in which case  $M$  is infinite and, if a physical emitter is used, field emission then becomes impossible. (In addition, if  $z_1 = 0$ , zero contrast is predicted for the image of an in-focus transmission phase object.) The magnification and the focus setting of the PPM are not independent – for given  $z_2$  both are fixed by  $z_1$ . For a crystalline sample for which the transmission diffraction orders overlap, the above analysis does not apply in the presence of multiple scattering, since then a transmission function cannot be used, as discussed elsewhere [30].

It follows from Eq. (6) that existing algorithms for focus correction (under plane-wave illumination) can be used to reconstruct low-voltage in-line holograms (formed with spherical-wave illumination). For masks, Fresnel edge fringes and Young's fringes between nearby holes are the main contrast features, and the aim of reconstruction is to remove these fringes, thus revealing the shape of the holes. The problem of focus correction (or hologram reconstruction) in the in-line geometry has been studied extensively, both for optical and electron holography. The first solution to the unavoidable twin-image problem which arises in this geometry may be that of Thompson and co-workers [31], who developed the optical method of in-line Fraunhofer holography for small particles. Solutions to the twin-image problem based on recording images at different defoci or lateral tip positions are described by Lin and Cowley [12]. Shadow images of masks have recently become important in the field of semiconductor lithography, where the inverse problem arises. For finer line-widths and sharper edges one requires the



edge transmission function whose defocused shadow image is most abrupt.

## 2.3 Holographic Reconstruction Algorithms

Given the PPM image (necessarily out of focus by  $z_1$ ) of a mask-like object, the question arises as to whether the true shape of the mask can be recovered. For a mask we define the transmission function

$$\begin{aligned} q(x) &= 1 - p(x) \\ &= 1 && \text{Within a hole} \\ &= 0 && \text{Within opaque regions} \end{aligned}$$

Then, assuming the equivalence between point-projection shadow images and conventional images (Eq. 6), the wavefunction on a plane distance  $z_1$  downstream from the sample is, for a magnification of unity [32],

$$\psi_h(x) = q(x)^* t_{z_1}(x) \quad (7)$$

The recorded hologram intensity is

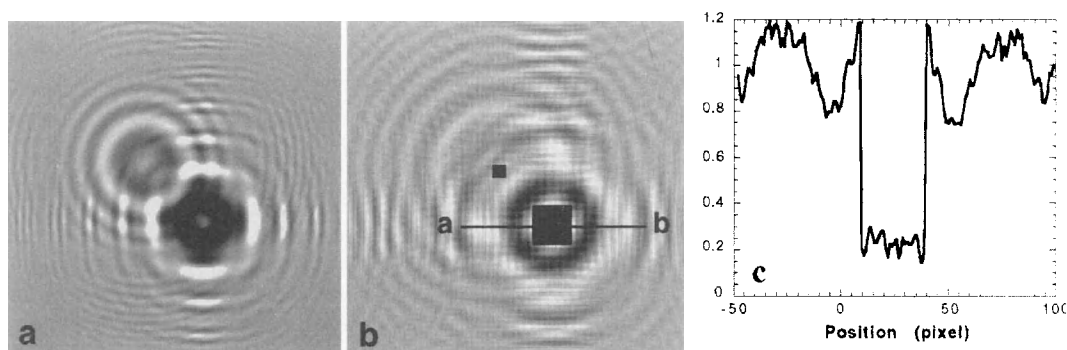
$$\begin{aligned} I(x) &= \psi_h(x) \psi_h^*(x) \\ &= 1 - p(x)^* t_{z_1}(x) - p^*(x)^* t_{z_1}^*(x) \\ &\quad + |p(x)^* t_{z_1}(x)|^2 \end{aligned} \quad (8)$$

To reconstruct the object from the hologram, we first Fourier transform, then multiply by  $T_{z_1}^*(u)$ , and finally perform an inverse transform:

$$\begin{aligned} \text{FT}^{-1} \{ T_{z_1}^* \cdot \text{FT}[I(x)] \} \\ &= 1 - p(x) - p^*(x)^* t_{2z_1}^*(x) \\ &\quad + |p(x)^* t_{z_1}(x)|^2 * t_{z_1}^*(x) \\ &= 1 - p'(x) \end{aligned} \quad (9)$$

The first two terms give a perfect reconstruction of the original transmission function  $q(x)$ , the real image. The third term gives the complement of the object, which is out of focus by twice the tip-to-sample distance  $z_1$ . If the object (of width  $d$ ) is small, and  $z_1$  large enough, this virtual twin image will produce a broad and slowly varying background which allows the real image to be isolated. [The requirement for this is  $z_1 \gg d^2/\lambda$  (Fraunhofer holography).] The fourth term is a second-order term (actually the reconstruction of the hologram of the complementary object), which is small for weak phase objects, but not necessarily small for masks. Physically, this term might be expected to be troublesome in regions well inside the shadow edge, where the 'reference wave' is weak. The ratio of the wave scattered from the edge to the direct wave controls this term and, for a finite object, this ratio depends on the Fresnel number  $N(d) = \pi d^2/(\lambda z_1)$ . Thus reconstruction for masks introduces greater errors than for weakly scattering objects; however, for the purpose of identifying simple discontinuities at edges, we shall see that the Fraunhofer condition can be relaxed. We have investigated these issues using computational trials on simulated and experimental data [33].

In Fig. 4 we illustrate the use of the above algorithm for simulated data. The simulations require periodic boundary conditions (with supercell period  $L$ ). Interference from neighboring cells can be avoided if the Fresnel number  $N(L) \gg 1$ . The reconstruction of two opaque square objects of width 1 and 3 nm (Fresnel numbers 0.43 and 3.86) is shown using 100-V electrons. The hologram was simulated using Eq. (8) and reconstructed using



**Figure 4.** Simulation of data using Eq. (9): supercell period  $L$ , Fresnel number of  $N(L) \gg 1$ , 100 V electrons. Reconstruction [using Eq. (10)] of two opaque square objects of (a) width 1 nm; Fresnel numbers 0.43. (b) 3 nm; Fresnel number 3.86. Note how smaller square with smaller  $N$  is clearer. (c) The edge of the objects can be readily located from the profile.

Eq. (9). The edge of the objects can be readily located from the profile, despite the disturbance from the background.

These results suggest that any sharp discontinuity (such as an edge) in a finite object will produce a corresponding sharp discontinuity in the reconstructed image at Gaussian focus if  $N$  is not too large. This would allow the shape of objects to be determined with high resolution. We now apply this method to experimental data.

Before attempting reconstruction, the magnification  $M = z_2/z_1$  of the images must be known. Five methods have been used:

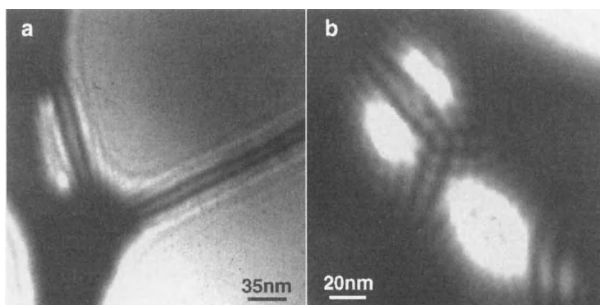
- (1) Use of a through-magnification series and an object (such as a grid bar) of known size. This is very inaccurate

since errors between successive images compound.

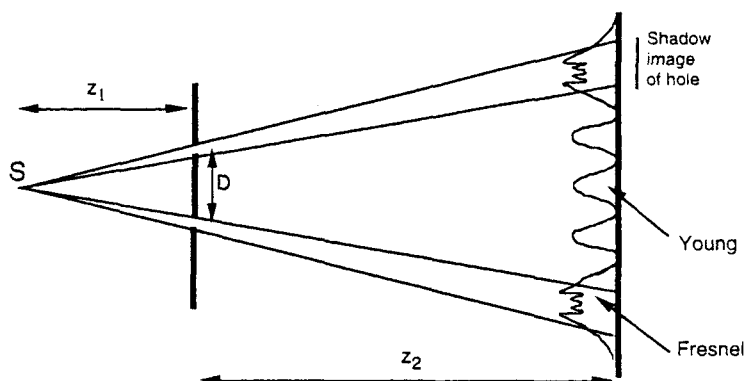
- (2) Use of parallax-shift of images with lateral tip movement [15].
- (3) Analysis of Fresnel edge fringes [34].
- (4) Use of Young's fringes observed between two pinholes of spacing  $D$ , using  $d = z_2\lambda/D$ , where  $d$  is the fringe spacing and  $z_2$  is the sample-to-screen distance.
- (5) Trial-and-error values of  $z_1$  may be used in the reconstruction scheme until a sharp image is obtained, as in Fig. 4b.

We have used methods (3)–(5).

Figure 5 shows two experimental point-projection holograms of a hole-containing carbon film obtained at 90 V. Figure 5b



**Figure 5.** Experimental point-projection holograms of a hole-containing carbon film obtained at 90 V. (a) In-focus image. (b) Crossed sets of Young's interference fringes between three pinholes.



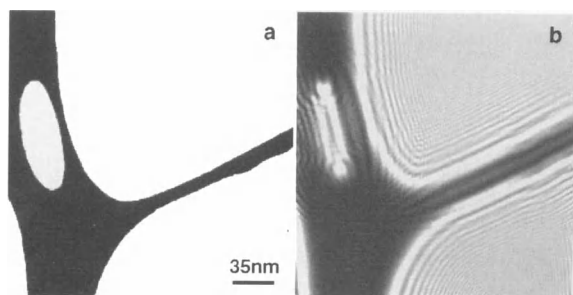
**Figure 6.** In the near-field of a mask containing holes, both a shadow image of the holes and interference fringes are formed.

shows the crossed sets of Young's interference fringes between three pinholes which might be confused with an atomic-resolution lattice image. The fringes may be simply interpreted using Eq. (6), expressed in two dimensions. Then we see that the fringes are equivalent to near-field Young's fringes, as would be obtained using plane-wave illumination. In the near-field of a mask containing holes, both a shadow image of the holes and interference fringes between different holes are observed [29], as indicated in Fig. 6.

The holograms shown in Fig. 5 were also reconstructed using Eq. (9) and a program which allows for the change of object pixel size with variation of defocus

$z_1$ , which changes the magnification. An approximate value of  $z_1$  was assumed and the corresponding magnification  $M = 14 \text{ cm}/z_1$  used. Since the computations are fairly rapid, the effect of varying the trial value of  $z_1$  simulates changing experimental focus near Gaussian focus, and by comparing reconstructed amplitude and phase images, the in-focus image may readily be identified by eye, despite the background from the twin image and non-linear term. The focus correction needed to obtain the in-focus image is the experimental tip-to-sample distance.

The shape of the fibers retrieved from the in-focus image reconstructed from Fig. 5a is shown in Fig. 7a, while Fig. 7b shows



**Figure 7.** (a) Shape of the fibers retrieved from the in-focus image reconstructed from Fig. 5a; (b) forward simulation of the hologram based on (a).

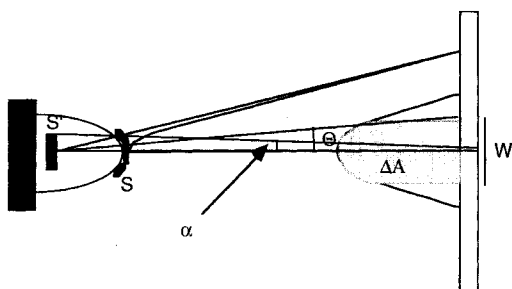
a forward simulation of the hologram based on Fig. 7a, in excellent agreement with the experimental hologram (Fig. 5a) if a defocus of 1850 nm is used. An integration over the finite source size would improve the fit by washing out the higher order Fresnel fringes. Reconstructions obtained from Fig. 5b eliminate entirely the Young's fringes at certain focus settings [33].

The sharpness of the edges in the reconstruction depends on the Fresnel number  $N = \pi d^2 / \lambda z_1$ , which describes the effect of defocus  $z_1$ , wavelength  $\lambda$ , and object diameter  $d$ . For the edge of an 'infinite' half-plane ( $N$  infinite) the reconstruction is noticeably poorer than for a small object. We find that for a simple mask object the Fraunhofer condition  $N \ll 1$  for elimination of the twin image is an excessively stringent requirement. In addition, any phase shift on scattering at edges, or partial transmission, will also affect the reconstruction.

In summary, we find that any sharp discontinuity (such as an edge) in a finite object will produce a corresponding sharp discontinuity in the reconstructed image at Gaussian focus if  $N$  is not too large. As Fig. 4 shows, this focus setting (the tip-to-sample distance) can easily be identified from a focal series of reconstructed images. This discontinuity, identifiable against the background of the twin image and the non-linear terms, easily allows the outline shape of small opaque objects to be identified. Resolution in the reconstructed image will still be governed by the electron source size (which determines the highest order Fresnel fringe, and hence the size of the hologram, since the Fringes get progressively finer according to Eq. (13)).

## 2.4 Nanotips, Tip Aberrations, Coherence, Brightness, Resolution Limits, and Stray Fields

The observation of Fresnel fringes (i.e., the ability to form a hologram) in the point-projection mode requires either demagnification of a conventional electron source (as in the original experiments of Boersch and in STEM work), or a physical source of subnanometer dimensions. This may be obtained either by using naturally occurring asperities (as in STM), or by preparing an atomic structure of desired shape. An extensive literature exists in the field of FIM on tip preparation procedures, based on oxygen etching, field evaporation, heating or sputtering [35]. More recently, using FIM imaging to assist in characterization, single-atom and few-atom 'nanotip' field emission tips have been prepared [15], for which claims of very high brightness and other unique properties have been made, including departures from the Fowler–Nordheim law, 'focusing' effects and unusual emission energy spectra [36]. Theoretical treatments for diffraction of a wavepacket at a laterally confined tunneling barrier have also been given (for a review of all this work see Vu Thien Binh et al. [37]). We now discuss the nanotip properties most relevant to PPM, their aberrations and brightness. Figure 8 shows an idealized tip and the curved trajectory of electrons accelerated from the tip. Since most of the potential is dropped within a few hundred nanometers of the tip, the electrons reach their final kinetic energy rapidly. The rearward



**Figure 8.** Idealized tip used in PPM and curved trajectory of electrons accelerated from the tip.

asymptotic extension of these rays defines the virtual source inside the tip, whose size  $d$  depends on the failure of rays leaving at different angles to meet at a point in the Gaussian image plane inside the tip (defined by paraxial rays). Thus the tip acts as a lens of nanometer dimensions, whose aberration coefficients can therefore be expected to be also of nanometer dimensions. Aberration coefficients for such a nanotip have been computed by modern ray-tracing techniques, together with numerical solutions of Laplace's equation [20]. For a tip with 1 nm radius sitting on a boss with 100 nm radius, we find  $C_s = 0.177$  nm and  $C_c = 0.142$  nm at 100 V, with a tip field strength at the surface of the tip of  $5 \text{ V nm}^{-1}$ . We shall see that the effects of these aberrations on resolution may not be negligible, because the virtual source size depends also on emission angle and wavelength. The emission angle  $\theta$  for this tip was found to be  $7.4^\circ$  (halfwidth at  $e^{-1}$ ).

From Fig. 8 the coherence width at the detector is seen to be approximately  $X_c = \lambda/\alpha$ . It is also useful to define a coherence angle  $\beta = \lambda/d$ , which is equal to the angular range over which emission from the tip is coherent – this is the angle subtended by  $X_c$  at the tip. The virtual

source diameter  $d$  for the 1 nm radius nanotip discussed above is found to be 0.43 nm (i.e., demagnified). For such a source the coherence width on a sample at  $z_1 = 100$  nm is 54 nm, while that on a detector at  $z_2 = 10$  cm is almost 6 cm ( $\lambda = 0.126$  nm, 100 V). The coherence angle is greater than the emission angle. The coherence width is approximately equal to the width of all the Fresnel edge fringes which can be seen, and hence to the size of the hologram. Because of their high brightness, nanotips may therefore be useful for electron interferometry experiments in which the beam must pass around macroscopic objects smaller than  $X_c$ . Measurements of source size have recently been made for a nanotip using a biprism [38]. The longitudinal coherence length  $L_c = \lambda(E/2\Delta E) = 63$  nm for a nanotip with  $\Delta E = 0.1$  eV at 100 V is considerably less than for high-voltage field emission guns.

The brightness  $B = j/\Delta\Omega$  (see Fig. 8) of a nanotip has been measured by matching ray-tracing computations and Fowler–Nordheim curves to experimental nanotip intensity distributions [39]. By comparing field ion microscopy and field emission electron images with the ray-tracing results, the effective source size could be obtained. Tungsten (111) single-crystal tips were fabricated using Ne sputtering and field evaporation. The average brightness of single-atom-terminated nanotips was found to be  $3.3 \times 10^8 \text{ Å cm}^{-2} \text{ sr}^{-1}$  at 470 V, or  $7.7 \times 10^{10} \text{ Å cm}^{-2} \text{ sr}^{-1}$  when extrapolated to 100 kV. This produces a beam with greater particle flux per unit energy range than obtainable using current synchrotrons. Although this value is almost two orders of magnitude brighter than the values quoted for STEM

cold-field emission guns, it should be borne in mind that the STEM values are based on measurements of the focused probe intensity, which is affected by objective lens aberrations. When corrections are made for this effect, the nanotip brightness is found to be greater than cold-field emission by a factor of about 10 at 100 kV. The most valuable aspect of the nanotip for STEM instruments may be the increased emission stability obtained at low extraction voltages, the reduced emission energy spread  $\Delta E$ , and the increased tolerance to poorer vacuum conditions. It has been speculated [40] that this is due to the reduced energy of back-streaming ions, which at higher extraction voltages may modify the workfunction of the tip.

Source brightness can be measured more directly using the concept of degeneracy  $\delta = B/B_0$ , where  $B_0$  is the upper quantum limit on brightness corresponding to two Fermions per cell in phase space (if  $\delta = 1$ ). The degeneracy can also be expressed as

$$\delta = (j/e)A_c T_c \quad (10)$$

This is the mean number of particles per coherence time  $T_c = h/\Delta E$  traversing a coherence area  $A_c = X_c^2$  (normal to the beam). In practice, this means that  $\delta$  (and hence the brightness) can be measured if the beam energy spread and the current density within a patch of Fresnel fringes are known, since the width of these fringes is given approximately by  $X_c$ .  $\delta$  is unity if  $j = e/(A_c T_c)$ , so that one electron crosses the coherence area per coherence time. Using the expression for the constant  $B_0$ , it follows [34] that the experimental brightness can be determined from the wavelength, the current density  $j$  and the

coherence area at the image plane using

$$B = 2jA_c/\lambda^2 \quad (11)$$

The effect of source size on resolution may be understood by considering the effect of displacing an ideal emission point transversely by  $b$ . This translates the entire image by  $Mb$ , where  $M = z_2/z_1$ . This imposes a resolution limit on the images, since the complete image intensity must be integrated over the effective source size  $d$ , which cannot be smaller than the electron wavelength. This integration takes into account all partial spatial coherence effects. For masks, the resolution of the PPM is most conveniently defined operationally as the width of the finest Fresnel fringe (in the object space), since this is readily shown to be equal to the size of the electron source (including the effects of all instabilities), as we now show. It was first shown by Sommerfeld that a good approximation to the Fresnel integral may be obtained by summing over just two optical paths (rather than infinitely many), the direct ray from source to detector, and that which passes via the edge of the mask. This gives the correct position of maxima and minima, but not the intensity distribution, if path differences  $\Delta_0 = 3\lambda/8$  for the edge wave scattered outside the shadow edge and  $\lambda/8$  for the edge wave scattered inside are used. Then the lateral position  $X_n$  of the  $n$ th Fresnel fringe on the screen is

$$X_n = M\sqrt{2z_1(n\lambda + \Delta_0)} \quad (12)$$

The width of the  $n$ th fringe on the screen is

$$\delta X_n = \frac{M^2 z_1 \lambda}{X_n} \quad (13)$$

Setting  $X_c = X_n/M$  at the sample we find the width of the highest order fringe at the

sample is approximately equal to the source size  $d = 2z_1\lambda/X_c$ . In practice, however, the contrast of the fringes decreases with order, making measurements of source size and coherence by this method approximate. (For a full analysis of Fresnel fringes in PPM see Spence et al. [34]. The image magnification, effective source size, transverse coherence width, instrumental resolution, and source brightness are all obtained from an analysis of the fringe spacings and intensity.)

The resolution limit in PPM may be thought of in other ways. By limiting the size of the hologram, the coherence width  $X_c$  imposes a diffraction limit on the reconstruction, since the hologram acts rather like a lens (or zone-plate) during reconstruction. Thus the coherence angle  $\beta$  plays a similar role to the angular limit set by the aperture of the objective lens in HREM (as suggested by  $\theta$  in Fig. 1) – a larger angle is needed for higher resolution  $d_r$ . The diameter of the zone plate is approximately equal to the transverse coherence width  $MX_c$  at the detector screen. Thus  $d_r = 1.6\lambda z_2/MX_c = 0.8\lambda/\beta$ , in agreement with the idea that the coherence angle from the source acts as a diffraction-limiting aperture. Higher resolution is obtained by increasing the emission coherence angle  $\beta = \lambda/d$ , and this can be achieved by reducing the source size  $d$ . A useful result is obtained by noting that, since  $X_c = 2z_1\lambda/d = \lambda/2\alpha$  (where  $\alpha$  is the angle subtended by the source at the sample) and, since the resolution is  $d_r = 1.6\lambda z_1/X_c$ , then  $d_r = 3.2z_1\alpha$ . Thus resolution in the reconstruction is improved by using small tip-to-sample distances and the smallest sources. For a thin crystal, we may again think of  $\alpha$  as limiting the ‘numerical aperture’ of the

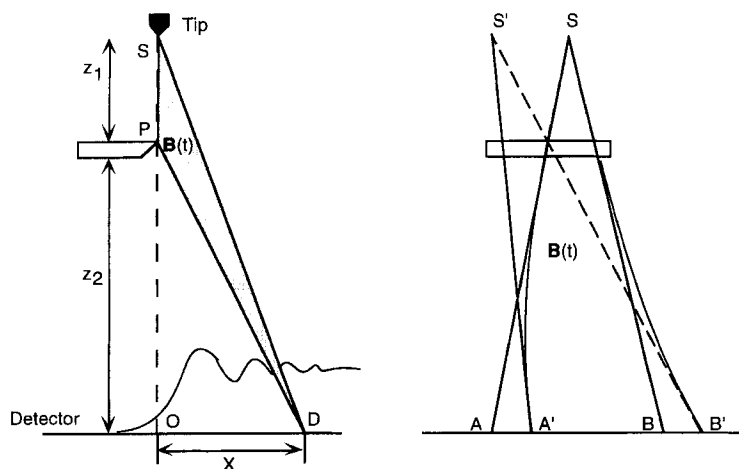
nanotip lens. Then the finest fringes contributing to a Fourier image will be those with Bragg angle  $2\Theta_B = a$  for plane spacing  $d_{hkl}$ . Using  $l/d_{hkl} = 2\Theta_B/\lambda$ , we again see that the resolution is approximately equal to the source size  $d = d_{hkl}$ .

The effects on resolution of the spherical aberration of the tip must be considered. Taking the resolution to be very crudely equal to the virtual source diameter  $\Delta_{C_s}$ , we have, from the definition of spherical aberration,

$$\begin{aligned}\Delta_{C_s} &= 2r_i \\ &= 2M(C_{s3}\sin^3\theta + C_{s5}\sin^5\theta)\end{aligned}\quad (14)$$

where  $\theta$  is measured in the object space (the emitting surface) and, from calculations,  $M = 0.531$  at 100 V. For emission at  $\theta = 45^\circ$ ,  $\Delta_{C_s} = 0.84 \text{ \AA}$  for the 1 nm radius nanotip. It is instructive to compare this with the performance of a magnetic lens for high-energy electrons. Here, at 100 kV ( $\lambda = 0.037$ ) one might typically have  $C_s = 0.5 \text{ mm}$ , but an angular aperture set at about  $\theta = 10 \text{ mRad}$ . Then  $\Delta_{C_s} = 2.5 \text{ \AA}$  because of the much smaller angles but larger aberrations. The idea that the aberrated virtual source size is equal to the resolution, however, is not strictly correct, since no simple expression for the resolution-limiting effects of spherical aberration on shadow images has been given, even for weakly scattering objects [41].

The effects on resolution of mechanical vibration of the tip and an enlarged emission area may be included by summing the image intensity distribution given by Eq. (5) over a range of source points using a suitable source intensity distribution function. This incorporates all partial spatial coherence effects. It is instructive to compare the effects of tip motion in several



**Figure 9.** A transverse time dependent magnetic field will introduce a time dependent phase shift between the optical paths SPD and SD (left); construction for estimating the degrading effect of stray magnetic fields on the resolution of a general PPM image (right).

point projection microscopies. For both electron field emission microscopy and FIM, the image is a projection of the surface of the tip from a virtual source inside the tip which cannot move relative to the surface of the tip. The effects of vibration in these instruments are therefore not severe, since they consist of whole-body translations of the source and 'sample' together, which are not magnified. A 1-nm movement of the tip appears as a 1-nm displacement on the screen. For the PPM and STEM in-line holography, any independent vibration at the source appears magnified at the detector. It is perhaps this relative immunity of the field ion microscope to vibration effects which accounts for the fact that atoms were first seen with these instruments. For conventional microscopy (TEM) it is the relative motion of the sample and the detector (rather than the source) which is crucial, since plane-wave illumination is used.

The effects of a spread of energies  $\Delta E$  in the beam has been discussed previously [34]. For Fresnel fringes, it was found that energy broadening of the source imposes a chromatic resolution limit such

that the highest order fringe which may be observed has order  $n_{\max} = 2E/\Delta E$ . The width of this fringe (referred to the object space) is

$$\Delta C_r = \sqrt{z_1 \lambda / (2n_{\max})} \quad (15)$$

which, by substituting for  $n_{\max}$ , gives the resolution limit due solely to the energy spread of the beam. For  $z_1 = 1000$  nm and  $\Delta E = 0.1$  eV we obtain  $\Delta C_r = 0.17$  nm at  $E = 100$  eV.

As shown in Fig. 9a, a transverse time-dependent magnetic field will introduce a time-dependent phase shift between the optical paths SPD and SD. This phase shift is

$$\theta(t) = \frac{e}{h} \int \mathbf{B}(t) \cdot d\mathbf{r} \quad (16)$$

where the integral is carried out over the shaded area in the figure. The use of just two optical paths rather than the Fresnel integral can be shown to give the correct positions of the maxima and minima in the Fresnel fringes. This phase shift may be included in the Fresnel fringe analysis, and a time average of the intensity taken in order to compute the effects of stray fields.



The result will depend sensitively on the order of the fringe. A rough estimate may be made by setting the maximum allowable phase shift equal to  $\pi/4$  at the position of the  $n$ th Fresnel fringes and solving for  $B$ . This fringe occurs at  $X_n = M[2z_1(n\lambda + \Delta_0)]^{\frac{1}{2}}$  on the detector, with  $\Delta_0 = \frac{3}{8}\lambda$  the approximate phase shift on scattering at the edge. This procedure gives

$$\langle B \rangle_{\max} = \frac{\pi}{4} \frac{h}{e} \frac{2}{z_1 X_n} \quad (17)$$

The fringe width  $\delta X_n$  at the detector is related to  $X_n$  by  $\delta X_n = M^2 z_1 \lambda / X_n$ . The smallest (highest order) fringe is equal to the incoherent virtual source size. If we take this to be, say 1 nm (referred to the specimen plane), then about six fringes will be seen. The maximum tolerable time-dependent field which allows observation of the sixth fringe is then  $\langle B \rangle_{\max} = 6 \times 10^{-7}$  T (a few milligauss) at 100 V with  $z_1 = 100$  nm and  $z_2 = 14$  cm. Only the component of the field normal to the shaded area is important. The constant Earth's field (about 0.6 T) can be ignored.

Figure 9 also suggests a construction for estimating the degrading effect of stray magnetic fields on the resolution of a general PPM image. If a homogeneous field  $\mathbf{B}(t)$  fills the space between sample and detector in the direction shown (worst case), the electron's trajectory will be circular, and the rearward asymptotic extension of these deflected rays at the detector will define a displaced source point  $S'$ . The distance  $S-S'$  gives a measure of the source enlargement, and hence resolution loss, due to a time dependent field. We note that  $S-S'$  also contains a  $z$ -component, and so there is a defocusing effect.

For steady fields, the angular deflection of a non-relativistic electron of energy  $eU_0$  acted on over path length  $L$  by a transverse electric field  $F$  is

$$\theta = \frac{FL}{2U_0} \quad (18)$$

while for a magnetic field with component  $B_t$  transverse to the beam the deflection is

$$\theta = \left( \frac{e}{2m} \right)^{\frac{1}{2}} \left( \frac{B_t L}{U_0^{\frac{1}{2}}} \right) \quad (19)$$

In summary, the width of the finest Fresnel fringe from a carbon foil edge is the most useful practical measure of resolution for the PPM. Energy-broadening and source-size effects may limit resolution to approximately the electron wavelength. They do not impose an absolute limit, however, and, as in near-field and other forms of microscopy, resolution will ultimately be limited by the signal/noise ratio. The non-linear terms in holographic reconstruction methods also limit resolution. In the presence of multiple scattering, resolution cannot easily be defined. Resolution increases with emission coherence angle, which acts as a diffraction-limiting aperture, as for a lens. This angle  $\beta = \lambda/d$  (where  $d$ , the effective source size, is a measure of the resolution) can be measured using the methods outlined above [34] or, better, by using an electron biprism [38].

## 2.5 Instrumentation

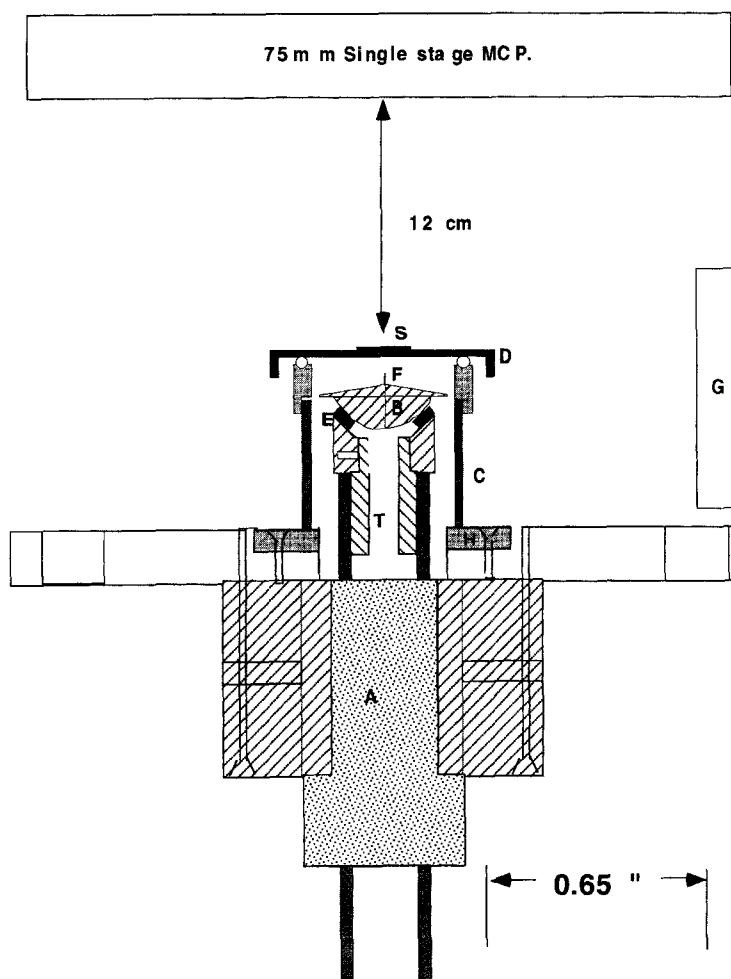
Several designs for low-voltage point-projection instruments have appeared in the literature. All consist of a tip motion, a sample holder and a detector, for which all

recent instruments have used a single-stage channel-plate followed by a fluorescent screen. These items are supported within an ultra-high vacuum chamber which must be bakeable (at least initially) and supported on a vibration isolation system. A leak valve for sputtering and vacuum gauges may be required. Pumps must pump the noble gas used for sputtering. Fine and coarse motion of the tip relative to the sample along  $x$ ,  $y$ , and  $z$  directions is essential to bring a region of interest onto the optic axis, defined by the beam, and to control the magnification and focus. Tip motions have been based on the original piezo tripod scanner used in STM [14] or on a piezo tube scanner for lateral motion and an inchworm for coarse  $z$  motion [18]. Tip cooling to 4.2 K has been provided [19]. Tip heating (for cleaning) may be provided by resistance heating (in which case the tip support follows the design of the tip assembly in field-emission STEM instruments), or by indirect electron-beam heating, which achieves lower temperatures but allows a higher mechanical resonant frequency for the tip, and does not require thick high-current leads to the tip assembly. If field-ion imaging of the tip is planned, electrical insulation for voltages of at least 10 kV must be provided to the tip. Since resolution is degraded by tip vibration, resistance heaters must be short and rigid, and all the principles of good STM design apply [42]. For a machine supported on a spring system with natural resonance frequency  $f_{sp}$ , the attenuation  $K$  of vibration amplitude reaching a tip with natural resonance frequency  $f_{tip}$  is approximately

$$K = \left( \frac{f_{sp}}{f_{tip}} \right)^2 = \frac{a_{tip}}{a_{ext}} \quad (20)$$

if  $f_{sp} \ll f_{tip}$ . Here  $a_{tip}$  is the vibration amplitude of the tip, while  $a_{ext}$  is the amplitude supplied to the frame of the support. The frequencies are proportional to the square root of the relevant spring constant divided by a mass. One makes  $K$  as small as possible by making  $f_{sp}$  as small as possible (e.g., by using an air-table or elastic suspension system with large mass and small spring constant), and  $f_{tip}$  as large as possible (by making the tip as short and rigid as possible). Typical values of  $f_{tip} = 2$  kHz and  $f_{sp} = 2$  Hz give  $K = 10^{-6}$  or 120 dB attenuation. Tips with  $f_{tip}$  above 10 kHz are highly desirable. A short mechanical path between tip and sample is required to minimize relative displacements of the tip and sample – a translation of both together is not magnified.

As for STM, the most difficult design problem is the coarse lateral motion of the sample relative to the tip. A design has been described [18] which provides one-dimensional transverse motion of a carriage riding on three sapphire balls, based on an inertial stick-slip motor. This, however, requires careful leveling to avoid consistent downhill motion. The use of a molybdenum plate riding on sapphire balls has been reported to solve this problem. (Unlike STM stick-slip stages, magnets cannot easily be used to stabilize the motion against gravity.) Transverse inchworm motion has been used; however, jumps in the motion and the loss of a short, rigid mechanical path between tip and sample may limit resolution. The crucial requirement is to be able to control *both*  $z_1$  and the coarse lateral stage motion interactively, in order to center a feature of interest on the optic axis. This is simple once one is within range of the fine lateral motion provided by a tune scanner



**Figure 10.** A recent vertical axis PPM which allows convenient tip and sample exchange, together with a stick-slip goniometer [43] for alignment of the beam.

supporting the tip. High-stability, low-noise electrical supplies (e.g., batteries) may be required for the scanner, since lateral sensitivities of  $50 \text{ nm V}^{-1}$  are common.

Figure 10 shows a recent vertical axis design [43] which allows convenient tip and sample exchange, together with a stick-slip goniometer [44] for alignment of the beam. This allows one to take advantage of natural asperities on a single-crystal tungsten tip, which, in our

experience (and that of STM workers) can be extremely sharp in the sense of producing very fine Fresnel edge fringes. The goniometer is also used for experiments in the reflection mode. Since natural asperities are unlikely to be aligned with the  $[111]$  axis, the goniometer allows them to be rotated onto the axis. This use of chemically etched tips without heating, sputtering or field evaporation then allows organic samples, which cannot be baked, to be studied at  $10^{-8} \text{ T}$ . Lower vacuum

produces excessive ion background in the MCP. In Fig. 10, A is an inchworm and E is three shear-mode piezo stacks which rotate the hemisphere B about two orthogonal axes by the stick-slip mechanism, thus aligning the tip F eucentrically. The stage D (with samples S on a TEM grid) sits on three balls on top of a piezo outer tube C, whose four electrodes are driven by a ramp, giving lateral fine or (stick-slip) coarse motion of the sample. A pumped load-lock at G allows either the sample to be removed using a pick-up fork, or, subsequently, after extending the inchworm upward, the hemisphere containing the tip may be exchanged by a smaller fork. The previously baked UHV environment within is thus preserved. An internal  $\mu$ -metal cylinder surrounds the instrument, which sits on a 60 l/s ion pump.

A simple technique for initial chemical etching of tungsten tips has been described which allows optical examination during etching [45]. Such a tip will field-emit at about 600 V or less initially, and this voltage may fall to 50 V or so as the tip is brought to within a few tens of nanometers of the conducting sample. One of the most important findings from nanotip research is the unexpected stability of field-emitters at low voltage, and their tolerance to rather poor vacuum conditions. It has been suggested [40] that this occurs because electrons striking the 'anode' (the sample) have insufficient energy to sputter ions efficiently, which are normally focused back onto the tip, causing local modifications in the workfunction. By thus avoiding the need for baking after each sample exchange, the study of organic films is made possible using a nanotip field emitter.

## 2.6 Relationship to Other Techniques

We have seen that the point-projection method has its origins in the very beginnings of electron microscopy, where much higher voltages were used. Many of the effects seen at higher voltage can also be expected in the PPM if a transmitted (or reflected) beam can be obtained, using any of the methods listed in Sec. 2.7 of this Chapter. The central disk of an out-of-focus coherent convergent-beam pattern may be interpreted in an in-line electron hologram if the sample is sufficiently thin (for a review see Spence and Zuo [46]). The theory of STEM holography and shadow imaging using 100-kV electrons is given elsewhere [11,40], and this is closely related to the theory of lattice imaging in STEM, which depends on interference between coherent overlapping diffraction orders [47,48]. The resulting fringes may be interpreted as a shadow image of the crystal lattice [46], and such images have been obtained at atomic resolution [49]. For crystalline samples, a family of real or virtual sources (probe images) is formed by Bragg diffraction, lying on the reciprocal lattice and forming a point diffraction pattern [46] (exactly as in the Tanaka wide-angle CBED method, which isolates one of these). If the illumination aperture is much larger than the Bragg angle, the patterns are known as 'ronchigrams' and provide a useful aberration figure of the probe-forming lens, as described in Sec. 2.2, Chapter 4 of this Volume. Point projection has also been used as an imaging method using ions [50] and X-rays [51]. Possibilities using matter waves in general (e.g., atom interferometry or neutrons) are

discussed by Clauser and Reinsch [52]. Since spin effects between different beam electrons are negligible, an exact analogy exists between the theory of in-line optical holography and electron PPM imaging of masks, so that, as we have seen, the methods of Fraunhofer optical holography [31] may be used to solve the twin-image problem for small objects. Other approaches to solving the twin-image problem have been proposed [12].

## **2.7 Future Prospects, Radiation Damage, and Point Reflection Microscopy**

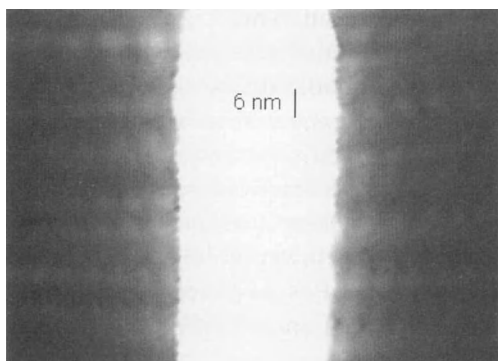
The present generation of instruments are capable of providing subnanometer-resolution images which reveal the external shape of mask-like objects. They may thus be useful for determining the shape of molecules drawn across holes in thin carbon films. In order to obtain images of internal atomic structures, one must use either higher or lower voltage (see Fig. 3).

Higher voltages become possible if the tip field is shielded by a cup-shaped surround, so that a higher voltage is needed for the same tip field. This shield (which is actually a lens) must not obstruct the sample. Lithographic techniques may allow such structures to be formed. Withdrawing the tip and using lenses after the sample to compensate for lost magnification allows higher voltage to be used (up to about 1 kV), but the image then has a very large focus defect. The aberrations of a second stage lens would be unimportant because of the small angles involved. Our experience has been that purple membrane

(thickness about 5 nm) becomes reasonably transparent to electrons above about 1 kV. Using a blunt tip, lower-resolution transmission images have been obtained at this voltage.

Research in surface science indicates that greater penetration should be obtainable at much lower voltage (see Fig. 3). Preliminary PPM experiments down to 7 eV have been encouraging [17]. Wedges of silicon may be useful test samples, and the resonances seen in LEED and elsewhere should be observable. Image interpretation becomes extremely complex at very low voltage (since exchange and correlation effects become important) and the increasing electron wavelength may limit resolution, unless near-field conditions apply. Full band-structure computations may be needed to interpret the images.

An important question concerns the effects of radiation damage at low voltage. The Bethe stopping power law predicts a cross-section for inelastic scattering increasing as  $1/v^2$  as beam energy decreases, but not all excitations result in damage, and this law fails below the ionization energy threshold. Rather little is known regarding damage in transmission samples at low energy; however, there are indications that it may decline in some systems [53]. Figure 11 shows a PPM image of a cut in a film of unstained, uncoated purple membrane (bacteriorhodopsin) showing no damage at 2-nm resolution after long exposure [54]. However, damage may have saturated in a surface layer, and samples of this thickness (5 nm) then still act as masks at 100 V. By contrast, for C16 lipid monolayers (thickness 2.6 nm) significant transmission was obtained at 100 V (perhaps mostly inelastic



**Figure 11.** PPM image of a cut in a film of unstained, uncoated Purple Membrane (bacteriorhodopsin) showing no visible damage at the 2 nm resolution level after long exposure [53].

scattering) and damage also observed. The interpretation of the purple membrane image in Fig. 11 has been confirmed by image simulations based on Eq. (5). The important point is that the fringes running normal to the edges of the vertical cut in the film can be explained by roughness at the edges alone, and do not indicate significant transmission of electrons. Because this roughness reflects the periodicity of the crystal, the fringes also show the period of the crystal [55].

It has been suggested that inner-shell ionization may be the primary cause of damage at higher voltages in some materials [56]. Although less probable than valence excitation, these processes involve larger, more localized energy transfers. Then damage would be greatly reduced at beam energies below the carbon K ionization energy (284 eV). (In fact, competition with other loss processes and multiple scattering may raise this threshold appreciably [53].) Secondly, a recent study of the energy dependence of damage found a distinct threshold for aromatic (but not aliphatic) compounds, such that

damage fell abruptly below about 1 kV [53]. Finally, it is commonly found [57] that radiolysis following valence excitation is suppressed in aromatic macromolecules with low  $G$  factors rich in  $\pi$  bonds (such as the nucleic acids), due to the delocalized nature of these resonant states, or by rapid screening of the hole on a time-scale faster than any nuclear response. A study of energy-loss spectra obtained from PPM images at various beam energies would contribute greatly to our understanding in this poorly understood and complex field. The elastic and inelastic cross-sections are comparable for electron beams at low energy, but inelastic scattering dominates for X-ray scattering. We therefore expect much less radiation damage per unit information using point projection electron methods than using comparable X-ray holography methods (see [58] for a detailed comparison).

More useful images might be obtained in the reflection geometry from bulk crystals with atomically smooth surfaces. The instrument shown in Fig. 10 was designed to investigate this possibility, using the tip goniometer to control the beam direction. Two methods have been proposed – one based on the Lloyd's mirror geometry [59] and the other using the specular reflected cone of radiation from a 'point' source [60] (i.e., RHEED with a spherical incident wave). The second method may be demonstrated using a diverging laser beam reflected from a mirror, on which is placed a small object. Both reflection high-energy electron diffraction (RHEED) computations (at a few hundred volts) and ray-tracing calculations [61] suggest that useful shadow images of surface steps can be obtained in this way if the experimental conditions can be controlled accurately.

In these low energy multiple scattering calculations, the results for reflection from an atomically smooth surface are modified to include the effect of a surface step using a simple phase shift. The experimental method relies on the fact that the electric field falls off rapidly near the tip, so that electron trajectories are relatively unaffected by the orientation of a reflecting surface a few micrometers away. Reflected rays appear to come from a virtual source below the mirror-like surface, whose distance from the surface defines the defocus. A diffusely scattering object on the surface scatters electrons weakly which interfere with the specular RHEED beam, forming a shadow image and hologram at infinity, as shown elsewhere [60, 61]. Multiple scattering computations show that the image is modulated by the convergent-beam RHEED rocking curve for the specular beam. Preliminary results are encouraging, and we note that similar shadow images can be seen in defocused reflection convergent beam patterns. Details are given elsewhere [61]. The first experimental low-voltage point reflection images were obtained recently, and are described elsewhere [62]. This instrument holds promise as a small, lensless, inexpensive alternative to the more complex low-energy electron microscope (LEEM).

In summary, the most promising areas of research in this field appear to be as follows.

(1) The development of imaging energy filters to exclude inelastic scattering. A filter of the imaging retarding Wien type may have advantages [63]. We note that the declining detective quantum efficiency of channel plates below about 100 V has a filtering effect.

(2) Development of the reflection mode [62]. If successful, this could be applied to problems of crystal growth in surface science and magnetic and ferroelectric domain imaging, as in the fields of (REM) and STM.

(3) Development of methods for preparing self-supporting molecule assemblies, perhaps drawn across holes in carbon films. The important problem of determining molecular shape for molecules which cannot be crystallized might be tackled in this way with existing instruments. In order to avoid the restriction to small objects ( $N \ll 1$ ), we are exploring the use of recursive calculation between two defocii, and the imposition of the a priori constraint that  $q = 1$  or  $0$  in image processing. The use of lower voltage (where IMF increases) may increase transmission, but reconstruction is needed to test this, since defocus produces delocalized images, and Young's fringes between different pinholes may otherwise give a false impression of transmission.

(4) Development of lithographically formed tip assemblies (mini-lenses) to enable operation at higher voltages, allowing greater penetration, but still below the inner shell ionization energies. For the study of lipid monolayers and Langmuir-Blodgett films, a compromise voltage might be found; however, it is to be expected that the effects of radiation damage will be highly materials dependent.

(5) For research in electron interferometry, electron anti-bunching, electron holography and the Aharonov-Bohm effect, the PPM provides an inexpensive experimental arrangement with excellent counting statistics due to the high source brightness. Chromatic coherence, however, is limited. Nanotips may also be

useful for electron spectroscopy in view of reports of very low energy spreads in the emission distribution [36].

## Acknowledgements

I am most grateful to Drs W. Qian, J. M. Zuo, X. Zhang and K. Taylor for their assistance with this project which was supported by NSF award DMR95-26100.

## 2.8 References

- [1] H. Boersch, *Z. Techn. Physik.* **1939**, 12, 346.
- [2] H. Boersch, *Naturwissenschaften* **1940**, 28, 709.
- [3] J. Hillier, *Phys. Rev.* **1940**, 58, 842.
- [4] D. Gabor, *Proc. R. Soc., Ser. A* **1949**, 197, 454.
- [5] M. E. Haine, T. Mulvey, *J. Opt. Soc. Am.* **1952**, 42, 763.
- [6] T. Mulvey, Personal communication **1994**.
- [7] G. A. Morton, E. G. Ramberg, *Phys. Rev.* **1939**, 56, 705.
- [8] A. J. Melmed, J. Smit, *J. Phys. E* **1979**, 12, 355.
- [9] J. M. Cowley, A. F. Moodie, *Proc. Phys. Soc. B* **1957**, 70, 486.
- [10] A. Crewe, D. N. Eggenberger, J. Wall, L. M. Welter, *Rev. Sci. Instrum.* **1968**, 39, 576.
- [11] J. Munch, *Optik* **1975**, 43, 79.
- [12] J. A. Lin, J. M. Cowley, *Ultramicroscopy* **1986**, 19, 31.
- [13] J. Rodenberg, R. Bates, *Phil. Trans. R. Soc., Ser. A* **1992**, 339, 521.
- [14] H. W. Fink, H. Schmid, H. J. Kreuzer, A. Wierzbicki, *Phys. Rev. Lett.* **1991**, 67, 1543.
- [15] H. Fink, W. Stocker, H. Schmid, *Phys. Rev. Lett.* **1990**, 65, 1204.
- [16] H. Fink, W. Stocker, H. Schmid, *J. Vac. Sci. Technol. B* **1990**, 8, 1323.
- [17] R. Morin, A. Gargani, *Phys. Rev. B* **1993**, 48, 6643.
- [18] J. C. H. Spence, W. Qian, A. J. Melmed, *Ultramicroscopy* **1993**, 52, 473.
- [19] V. T. Binh, V. Semet, N. Garcia, *Ultramicroscopy* **1995**, 58, 307.
- [20] M. Scheinfein, W. Qian, J. Spence, *J. Appl. Phys.* **1993**, 73, 2057.
- [21] D. K. Saldin, J. C. H. Spence, *Ultramicroscopy* **1994**, 55, 397.
- [22] D. R. Penn, *Phys. Rev. B* **1987**, 35, 482.
- [23] M. Seah, W. Dench, *Surf. Interface Anal.* **1979**, 1, 2.
- [24] S. Tanuma, C. Powell, D. Penn, *Surf. Interface Anal.* **1991**, 17, 911.
- [25] C. Martin, E. Arakawa, T. Callcott, R. Warmack, *J. Electron Spectrosc. Rel. Phenom.* **1987**, 42, 171.
- [26] W. Qian, J. Spence, J. Zuo, *Acta Crystallogr., Part A* **1993**, 49, 436.
- [27] J. Kreuzer, K. Nakamura, A. Wierzbicki, H. Fink, H. Schmid, *Ultramicroscopy* **1992**, 45, 381.
- [28] J. C. H. Spence, W. Qian, *Phys. Rev. B* **1992**, 45, 10 271.
- [29] G. M. Shedd, *J. Vac. Sci. Technol. A* **1994**, 12, 2595.
- [30] J. C. H. Spence, *Optik* **1992**, 92, 57.
- [31] J. DeVelis, G. Parrent, B. Thompson, *J. Opt. Soc. Am.* **1966**, 56, 423, and references therein.
- [32] J. Spence, *Experimental High Resolution Electron Microscopy*, Oxford University Press, Oxford, 1988.
- [33] J. C. H. Spence, X. Zhang, W. Qian, *Proc. Int. Workshop on Electron Holography 1995* (Eds.: A. Tonomura, L. Allard, D. Joy, Y. Ono), Elsevier, Amsterdam, 267.
- [34] J. C. H. Spence, W. Qian, M. P. Silverman, *J. Vac. Sci. Technol. A* **1994**, 12, 542.
- [35] A. P. Janssen, J. P. Jones, *J. Phys. D: Appl. Phys.* **1971**, 4, 118. See also: D. Bettler, S. Charbonnier, *Phys. Rev.* **1960**, 119, 85.
- [36] Vu Thien Binh, S. T. Purcell, N. Garcia, J. P. Doglioni, *Phys. Rev. Lett.* **1992**, 69, 2527.
- [37] Vu Thien Binh, N. Garcia, K. Dransfeld (Eds.), *Manipulation of Atoms under High Fields and Temperatures: Applications*, NATO-ASI-E Series, Vol. E235, Plenum Press, New York **1993**.
- [38] A. Degiovanni, R. Morin, *Proc. 13th Int. Congress on Electron Microscopy*, Les Editions de Physique (Les Ulis), **1994**, Vol. 1, p. 331.
- [39] W. Qian, M. Scheinfein, J. Spence, *J. Appl. Phys.* **1993**, 73, 7041.
- [40] P. Bovet, Personal communication.
- [41] J. M. Cowley, *Ultramicroscopy* **1979**, 4, 435.
- [42] P. Chen, *Introduction to Scanning Tunnelling Microscopy*, Oxford University Press, Oxford **1993**.
- [43] J. C. H. Spence, *Proc. Int. Congress on Electron Microscopy* Les Editions de Physique (Les Ulis) **1994**, 329.
- [44] L. Howald, H. Rudin, H. Guntherodt, *Rev. Sci. Instrum.* **1992**, 63, 3909.
- [45] A. J. Melmed, *J. Vac. Sci. Technol. B* **1991**, 9, 601.
- [46] J. C. H. Spence, J. M. Zuo, *Electron Microdiffraction*, Plenum Press, New York **1992**, Fig. 8.11.



- [47] J. C. H. Spence, J. M. Cowley, *Optik* **1978**, 50, 129.
- [48] J. Rodenberg, B. McCallum, P. Nellist, *Ultramicroscopy* **1993**, 48, 304.
- [49] J. M. Cowley, *J. Electron. Microscop. Technol.* **1986**, 3, 25.
- [50] H. Schmid, H. W. Fink, *Appl. Surf. Sci.* **1993**, 67, 436 and references therein.
- [51] V. E. Cosslett, W. C. Nixon, *J. Appl. Phys.* **1953**, 24, 616.
- [52] J. Clauser, M. Reinsch, *Appl. Phys. B* **1992**, 54, 380.
- [53] A. Howie, M. Muhid, F. Rocca, U. Valdre, *Inst. Phys. Conf. Ser. 90 (EMAG 87)* **1987**, 155.
- [54] J. Spence, W. Qian, X. Zhang, *Ultramicroscopy* **1994**, 55, 19.
- [55] X. Zhang, J. Spence, *Proc. 53rd Annual Meeting of the Electron Microscopy Society of America* (Ed.: G. Bailey), San Francisco Press, San Francisco **1995**, p. 846.
- [56] M. Isaacson, D. Johnson, A. V. Crewe, *Rad. Res.* **1973**, 55, 205.
- [57] A. J. Swallow, *Radiation Chemistry*, Longmans-Halsted-Wiley, New York **1973**.
- [58] R. Henderson, *Quarterly Reviews of Biophysics* **1995**, 28, 171.
- [59] J. Spence, W. Qian, *Proc. 50th Annual Meeting of the Electron Microscopy Society of America* (Ed.: G. Bailey), San Francisco Press, San Francisco **1992**, p. 938.
- [60] J. Spence, W. Qian, W. Lo, S. Mo, U. Knipping, X. Zhang, in *Determining Nanoscale Physical Properties of Materials by Microscopy and Spectroscopy* (Ed.: M. Sarikaya) MRS, Pittsburgh **1994**, Vol. 332, 405.
- [61] J. C. H. Spence, X. Zhang, J. M. Zuo, U. Weierstall, E. Munro, J. Rouse, in *Proc. MSA* (Ed.: G. Bailey), San Francisco Press, San Francisco **1995**, 610.
- [62] J. C. H. Spence, X. Zhang, J. Zuo, U. Weierstall, *Surf. Rev. Lett.* **1996**, in press.
- [63] H. Rose, *Optik* **1987**, 77, 26.

# General Reading

## Chapter IV.2.1

- J. N. Chapman and A. J. Craven (Ed.), *Quantitative Electron Microscopy*, Scottish University Summer Schools in Physics, **1984**.
- J. I. Goldstein, D. E. Newbury, P. E. 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**.
- D. B. Holt and D. C. Joy, *SEM Microcharacterization of Semiconductors*, Academic Press, London, **1989**.
- C. E. Lyman, D. E. Newbury, J. I. Goldstein, D. B. Williams, A. D. Romig, J. T. Armstrong, P. Echlin, D. C. Joy, E. Lifshin and K. R. Peters, *SEM, X-Ray Microanalysis and AEM – A Laboratory Notebook*, Plenum Press, NY, **1990**.
- D. E. Newbury, D. C. Joy, P. Echlin, C. E. Fiori and J. I. Goldstein, *Advanced Scanning Electron Microscopy and X-Ray Microanalysis*, Plenum Press, NY, **1986**.
- C. W. Oatley, *The early history of the SEM*, J. Appl. Phys., **53**, R1, **1982**.
- L. Reimer, *Image Formation in Low Voltage SEM*, SPIE Tutorial Texts in Optical Engineering, (SPIE: Bellingham, WA), TT12, **1993**.
- L. Reimer, *Scanning Electron Microscopy Physics of Image Formation and Microanalysis*, Springer Series in Opt. 5a. 45, Springer, Berlin, **1985**.
- O. C. Wells, *Scanning Electron Microscopy*, McGraw-Hill, NY, **1974**.
- B. G. Yacobi and D. B. Holt, *Cathodoluminescence of Inorganic Solids*, Plenum Press, NY, **1990**.
- H. Rose, *Non-standard imaging methods in electron microscopy*, *Ultramicroscopy*, **1977**, **2**, 251.
- M. Isaacson, D. Kopf, M. Utlaut, N. W. Parker, A. V. Crewe, *Direct observations of atomic diffusion by scanning transmission electron microscopy*, Proc. Natl. Acad. Sci. USA, **1977**, **74**, 1802.
- J. C. H. Spence and J. M. Zuo, *Electron Microdiffraction*, Plenum Press, New York, **1992**.
- J. M. Cowley, *Electron nanodiffraction; progress and prospects*, *J. Electron Microscopy*, **1996**, (in press).
- J. M. Cowley, J. C. H. Spence, *Principles and theory of electron holography*, Chapter 1 of *Introduction to Electron Holography*, L. F. Allard et al. (Eds.), Plenum Press, New York, **1996** (in press).
- S. J. Pennycook, D. E. Jesson, *High-resolution Z-contrast imaging of crystals*, *Ultramicroscopy*, **1991**, **37**, 14.
- J. M. Cowley, *Scanning microscopy at Arizona State University*. In: EMSA Bulletin, **1991**, **21**, 57.
- J. Liu, J. M. Cowley, *Scanning reflection electron microscopy and associated techniques for surface studies*, *Ultramicroscopy*, **1993**, **48**, 381.

## Chapter IV.2.3

- D. M. Bird, *Theory of Zone Axis Electron Diffraction*, J. Elect. Micr. Tech. **13**, **1989**, 77.
- P. R. Buseck, J. M. Cowley and L. Eyring (Eds.), *High Resolution Transmission Electron Microscopy*, Oxford University Press, New York, **1988**.
- J. N. Chapman and A. J. Craven (Eds.), *Quantitative Electron Microscopy*, Scottish University Summer School in Physics, Edinburgh, **1984**.
- M. M. Disko, C. C. Ahn and B. Fultz (Eds.), *Transmission Electron Energy Loss Spectrometry in Materials Science*, The Minerals, Metals & Materials Society, Warrendale PA, **1992**.
- R. F. Egerton, *Electron Energy-Loss Spectroscopy in the Electron Microscope*, Plenum Press New York, **1986**.

## Chapter IV.2.2

- J. M. Cowley, *Special Electron Microscopy Techniques*, Chapt. 7, in *Physical Methods of Chemistry*, Vol. IV, *Microscopy*, (Eds. B. W. Rossiter, J. F. Hamilton), John Wiley & Sons, New York, **1991**, p. 239.
- A. V. Crewe, *The physics of the high-resolution scanning microscope*, *Rep. Prog. Phys.*, **1980**, **43**, 621.

- C. J. Humphreys, *The Scattering of Fast Electrons by Crystals*, Rep. Prog. Phys., 42, **1979**, 1825.
- P. G. Merli and M. Vittori (Eds.), *Electron Microscopy in Materials Science*, Antisari, World Scientific, Singapore, **1992**.
- S. J. Pennycook, *High Resolution Electron Microscopy and Microanalysis*, Contemporary Physics 23, **1982**, 371.
- S. J. Pennycook, *Z-Contrast Imaging in the Scanning Transmission Electron Microscope*, J. Microsc. Soc. Am. 1, **1995**, 231.
- L. Reimer, *Transmission Electron Microscopy*, Springer-Verlag, Berlin, **1993**.
- J. Wall, J. Langmore, M. Isaacson and A. V. Crewe, *Scanning Transmission Electron Microscopy at High Resolution*, Proc. Nat. Acad. Sci. USA, 71, **1974**, 1.

## Chapter IV.2.4

- G. C. Smith, *Surface Analysis by Electron Spectroscopy*, Plenum Press, New York and London, **1994**.
- D. Briggs and M. P. Seah, *Practical Surface Analysis*, Second Edition, Vol. 1, John Wiley and Sons, **1990**.
- L. C. Feldman and J. W. Mayer, *Fundamentals of Surface and thin film Analysis*, North Holland Elsevier Science Publishing Co. Inc. **1986**.
- W. F. Egelhoff, Jr., *Surface Science Reports* 6, **1987**, 253-415.
- A. W. Czanderna, *Methods of Surface Analysis*, Elsevier Scientific Publishing Company, Amsterdam, **1975**.
- W. Dekeyser, L. Fiermans, G. Vanderkelen and J. Vennik, *Electron Emission Spectroscopy*, D. Reidel Publ. Co., Dordrecht, **1973**.
- L. Fiermans, J. Vennik and W. Dekeyser, *Electron and Ion Spectroscopy of Solids*, Plenum Press, New York, **1978**.
- T. A. Carlson, *Photoelectron and Auger spectroscopy*, Plenum Press, New York, **1975**.

## Chapter IV.2.5

- M. A. Baker and J. E. Castle, in: *Material Science and Technology: a Comprehensive Treatment*, (Ed. R. W. Cahn, P. Haasen and E. I. Kramer), VCH, Weinheim, **1994**, Vol. 2B, Chapter 13.
- N. D. Browning and S. J. Pennycook, *Microbeam Analysis*, **1993**, 1, 19.

- J. I. Goldstein, D. E. Newbury, P. Echlin, D. C. Joy, C. E. Fiori and E. Lifshin, *Scanning Electron Microscopy and X-Ray Microanalysis*, Plenum Press, New York, **1981**.
- J. I. Goldstein and D. B. Williams, *Microbeam Analysis*, **1992**, 1, 29.
- J. J. Hren, J. I. Goldstein and D. C. Joy (Ed.), *Introduction to Analytical Electron Microscopy*, Plenum Press, New York, **1979**.
- D. C. Joy, *Microbeam Analysis*, **1992**, 1, 19.
- D. C. Joy, A. D. Romig and J. I. Goldstein (Ed.), *Principle of Analytical Electron Microscopy*, Plenum Press, New York, **1986**.
- L. Reimer, *Transmission Electron Microscopy. Physics of Image Formation and Microanalysis*, 2nd ed., Springer-Verlag, Berlin, **1989**.
- K. F. J. Heinrich and D. E. Newbury (Ed.), *Electron Probe Quantification*. Plenum Press, New York, **1991**.
- V. D. Scott, G. Love and S. J. B. Reed, *Quantitative Electron-Probe Microanalysis*, 2nd ed., Ellis Horwood, New York, **1995**.

## Chapter IV.2.6.

- A. Benninghoven, F. G. Rüdenauer, H. W. Werner (Eds.), *Secondary Ion Mass Spectrometry: Basis Concepts, Instrumental Aspects, Applications and Trends*, Wiley, New York, **1987**.
- A. Benninghoven, *Angew. Chem. Int. Ed. Engl.*, **1994**, 33, 1023.
- 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.
- A. Lodding, in: *Inorganic Mass Spectrometry* (Eds.: F. Adams, R. Gijbels, R. Van Grieken) Wiley, New York, **1988**, p. 125.
- R. W. Odom, in: *Microscopic and Spectroscopic Imaging of the Chemical State* (Ed.: M. D. Morris), Marcel Dekker, New York, **1993**, p. 345.
- J. C. Vickerman, A. Brown, N. M. Reed (Eds.), *Secondary Ion Mass Spectrometry: Principles and Applications*, Charendon Press, Oxford, **1989**.
- R. G. Wilson, F. A. Stevie, C. W. Magee, *Secondary Ion Mass Spectrometry*, Wiley, New York, **1989**.

## Chapter V.1

- B. Blümich, W. Kuhn, *Magnetic Resonance Microscopy*, VCH, Weinheim, **1992**.  
P. Blümler, B. Blümlich, *NMR-Basic Principles and Progress*, Springer-Verlag, Berlin, **1993**.  
P. T. Callaghan, *Principles of Nuclear Magnetic Resonance Microscopy*, Oxford University Press, Oxford, **1991**.  
D. G. Cory, *Annu. Rep. NMR Spectrosc.*, **1992**, 24, 88.  
P. Jezzard, J. J. Attard, T. A. Carpenter, L. D. Hall, *Prog. NMR Spectrosc.*, **1992**, 23.  
P. Mansfield, E. L. Hahn, *NMR Imaging*, The Royal Society, London, **1990**.

## Chapter V.2

- R. J. Celotta, D. T. Pierce and J. Unguris, *MRS Bulletin*, **1995**, 30.  
E. D. Dahlberg and J. G. Zhu, *Physics Today*, **1995**, 34.  
J. P. Jacobovics, *Electron Microscopy in Materials Science Part IV*, Commission of European Communities, Brussels, **1973**, p. 1305.  
J. Kessler, *Polarized Electrons*, 2nd ed., Springer-Verlag, Berlin, **1985**.  
J. Kirschner, *Polarized Electrons at Surfaces*, Springer-Verlag, Berlin, **1985**.  
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**.  
D. T. Pierce, J. Unguris, R. J. Celotta, in *Ultrathin Magnetic Structure II*, Springer-Verlag, Berlin, **1994**, p. 117.  
J. Unguris, M. R. Scheinfein, R. J. Celotta, D. T. Pierce, in *Chemistry and Physics of Solid Surfaces VII*, Springer-Verlag, Germany, **1990** p. 239.

## Chapter V.3

- E. Bauer, *Rep. Prog. Phys.*, **1994**, 57, 895.

## Chapter VI.1

- E. Bauer, W. Telieps, in: *Surface and Interface Characterization by Optical Methods* (Eds.: A. Howie and U. Valdré), Plenum, New York, **1988**, p. 195.

- E. F. Burton, W. H. Kohl, *The Electron Microscope*, 2nd ed., Reinhold, New York, **1948**.  
H. Bethge, J. Heydenreich, *Electron Microscopy in Solid State Physics*, Elsevier, Amsterdam, **1987**.  
V. E. Cosslett, *Introduction to Electron Optics*, 2nd ed., Oxford University Press, Oxford, **1950**.  
O. H. Griffith, W. Engel, *Ultramicroscopy*, **1991**, 36, 1.  
O. H. Griffith, G. F. Rempfer, *Advances in Optical and Electron Microscopy*, Vol. 10, **1987**, p. 269.  
C. E. Hall, *Introduction to Electron Microscopy*, McGraw-Hill, New York, **1953**.  
O. Klemperer, *Electron Optics*, 2nd ed., Cambridge University Press, Cambridge, **1953**.  
C. Ramsauer, *Electronenmikroskopie (Report of the Work of the AEG Research Institute 1930-1942)*, Springer, Berlin, **1943**.  
R. A. Schwarzer, *Microsc. Acta*, **1981**, 84, 51.

## Chapter VI.2

- A. Cerezo, G. D. W. Smith, *Materials Science and Technology*, Vol. 2B: Characterisation of Materials Part II, (Ed. E. Lifshin), VCH, Weinheim, pp. 513-559.

## Chapter VII.1

- R. Wiesendanger, *Scanning Probe Microscopy and Spectroscopy: Methods and Applications*, Cambridge University Press, Cambridge, **1994**.  
C. J. Chen, *Introduction to Scanning Tunneling Microscopy*, Oxford University Press, New York and Oxford, **1993**.  
H.-J. Güntherodt and R. Wiesendanger (Eds.), *Scanning Tunneling Microscopy I*, Springer, Berlin, Heidelberg, New York, 2nd edition, **1994**.  
R. Wiesendanger and H.-J. Güntherodt (Eds.), *Scanning Tunneling Microscopy II*, Springer, Berlin, Heidelberg, New York, 2nd edition, **1995**.  
R. Wiesendanger and H.-J. Güntherodt (Eds.), *Scanning Tunneling Microscopy III*, Springer, Berlin, Heidelberg, New York, 2nd edition, **1996**.  
J. Stroscio and W. Kaiser (Eds.), *Scanning Tunneling Microscopy*, Academic Press, **1993**.

- D. A. Bonnel (Ed.), *Scanning Tunneling Microscopy and Spectroscopy: Theory, Techniques and Applications*, VCH-Verlag, **1993**.
- S. Gauthier and C. Joachim (Eds.), *Scanning Probe Microscopy: Beyond the Images*, Les éditions de physique, **1992**.
- P. Avouris (Ed.), *Atomic and Nanometer-Scale Modification of Materials: Fundamentals and Applications*, Kluwer, **1993**.
- C. R. K. Marrian (Ed.), *Technology of Proximal Probe Lithography*, SPIE Press, **1993**.
- P. N. J. Dennis, *Photodetectors: An introduction to current technology*, Plenum Press, New York, **1985**.
- L. M. Biberman, S. Nudelman (Eds.), *Photoelectronic Imaging Devices, Vol. I and II*, Plenum Press, New York, **1971**.
- M. M. Blouke (Ed.), *Charge-coupled Devices and Solid State Optical Sensors IV*, SPIE Proc., **1994**, 2172.
- Photometrics Ltd., *Charge-coupled Devices for Quantitative Electronic Imaging*, Photometrics Ltd., Tucson, **1989**.
- D. Shotton (Ed.), *Electronic light microscopy, Techniques in Modern Biomedical Microscopy*, Wiley-Liss, New York, **1993**.
- K.-H. Herrmann and D. Krahel, *Electronic image recording in conventional electron microscopy*, Adv. Opt. El. Micr. (Ed. R. Barer, V. E. Cosslett), **1984**, 9, 1.
- W. J. de Ruijter, *Imaging properties and applications of slow-scan charge-coupled device cameras suitable for electron microscopy*, Micron, **26**, **1995**, 247.
- G. Schmahl and D. Rudolph (Eds.), *X-Ray Microscopy*, Springer Series in Optical Sciences, **1984**, 43.
- A. G. Michette, G. R. Morrison and C. J. Buckley (Eds.), *X-Ray Microscopy III*, Springer Series in Optical Sciences, **1992**, 67.

## Chapter VII.2

- H.-J. Güntherodt, D. Anselmetti and E. Meyer (Eds.): *Forces in Scanning Probe Methods*, NATO-ASI Series, Series E: Applied Sciences, Vol. 286, Kluwer Academic Publishers, Dordrecht, **1995**.

## Chapter VII.3

- Dror Sarid, *Scanning Force Microscopy*, Oxford University Press, Oxford, New York, **1991**.
- H.-J. Güntherodt, D. Anselmetti and E. Meyer (Eds.), *Forces in Scanning Probe Methods*, Kluwer Academic Publishers, Dordrecht, Boston, London, **1995**.
- Sushin Chikazumi, *Physics of Magnetism*, John Wiley & Sons, Inc., **1964**.
- Richard M. Bozorth, *Ferromagnetism*, D. van Nostrand Comp., Inc., New York, **1951**.
- W. F. Brown, Jr. Robert, E. Krieger, *Micromagnetics*, Publish. Comp., Huntington, New York, **1978**.
- D. J. Craik and R. S. Tebble, *Ferromagnetism and Ferromagnetic Domains*, North Holland Publishing Comp., Amsterdam, **1965**.
- A. P. Malozemott and J. C. Slonczewski, *Magnetic Domain Walls in Bubble Materials*, Academic Press, INC., New York, **1979**.

## Chapter VIII.1

- J. C. Dainty and R. Show, *Image Science*, Academic Press, London, **1974**.

## Chapter IX.2

- J. W. Goodman, *Introduction to Fourier Optics*, McGraw-Hill, Current Edition, New York.
- J. C. H. Spence and J. M. Zuo, *Electron Microdiffraction*, Plenum Press, New York, **1992**, (ISBN 0-306-44262-0).
- J. C. H. Spenc, *Experimental High Resolution electron Microscopy*, Oxford University Press, New York, **1988**, 2nd. Edn. (ISBN 0-19-505405-9).
- R. Gomer, *Field Emission and field Ionization*, American Institute of Physics, New York, **1993**.
- R. Collier, C. Burkhardt, L. Lin, *Optical Holography*, Academic Press, New York, **1971**.

# List of Symbols and Abbreviations

Symbol	Designation	Chapter
$A$	amplitude	VII.4
$A$	area	IV.2.6, VIII.2, VIII.1
$A$	atomic weight of element $a$ ; constant	IV.2.5
$A$	fitting parameter	IX.2
$A$	scattering asymmetry	V.2
$a$	constant	V.1
$a$	periodicity	IV.2.2
$A(u)$	aperture function	IV.2.2
$A, B$	linear dimensions	IV.2.1
$a, b$	constant	VI.2
$A_C$	coherence area	IX.2
$A_{ex}$	exchange asymmetry	V.3
$a_{ext}$	amplitude of vibration of support	IX.2
$a_{tip}$	amplitude of vibration of tip	IX.2
$B$	fitting parameter; brightness	IX.2
$B$	magnetic field	V.1, IX.2
$B$	magnetic flux density	V.3
$b$	Kramers' constant ( $2 \times 10^{-9}$ photons $s^{-1}$ eV $^{-1}$ electron $^{-1}$ )	IV.2.5
$B_0$	upper quantum brightness	IX.2
$b^j(K, R)$	Bloch state amplitude	IV.2.3
$b^{RMS}$	Root-mean-square impact parameter for inelastic scattering	IV.2.3
$C$	complex constant	IX.2
$C$	constant	IV.2.5
$C$	image contrast	VIII.1
$c$	spring constant of the cantilever	VII.3
$c$	velocity of light	VIII.1
$C(x, y)$	contrast	VIII.2
$C, C'$	local contrast	VIII.2
$c_0$	estimated concentration	VI.2
$C_1, C_2, C_3$	constants	VI.2
$C_c$	Chromatic aberration coefficient	IX.2
$C_i, C_j$	concentration of element $i, j$	IV.2.5
$C_h$	mass concentration of element $a$ in all elements $j$	IV.2.5
$C_M$	concentration of species $M$	IV.2.6
$C_s$	Spherical aberration coefficient	IX.2, IV.2.3, IV.2.1, IV.2.2
$c_s$	speed constant of a photoemulsion	VIII.1
$C_X$	atomic concentration of element X	IV.2.4
$D$	fractal dimension	VIII.2

Symbol	Designation	Chapter
$D$	optical density; charge-coupled device dark current ( $\text{e}^- \text{ pixel}^{-1} \text{ s}^{-1}$ )	VIII.1
$D$	pinhole spacing	IX.2
$D$	specimen–screen distance	VI.2
$d$	crystal lattice spacing	IV.2.4
$d$	diameter of electron beam probe	IV.2.1
$d$	distance between image planes	VII.2
$d$	distance between pixels	VIII.2
$d$	flight length	VI.2
$d$	fringe spacing; tip size; virtual source size; width of object	IX.2
$d$	lattice spacing	IV.2.5
$d$	object size to be resolved	VIII.1
$D(E, T)$	total transmission probability for electron transport to the back of the semiconductor	VII.4
$D(\mathbf{K})$	Reciprocal space annular detector function	IV.2.3
$d(\mathbf{R})$	Real space annular detector function	IV.2.3
$D(u)$	transmission function of detector	IV.2.2
$d_{(hkl)}$	interplanar spacing for $(hkl)$ crystallographic pole	VI.2
$d(\delta^{(k)})$	difference wave	VIII.2
$D_0$	fog density of photoplates	VIII.1
$d_0$	thickness of an absorbing layer	VIII.1
$D_A$	detector efficiency	IV.2.4
$d_{\text{CCD}}$	pixel size of a charge-coupled device	VIII.1
$d_d$	thickness of a dead layer	VIII.1
$D_s$	saturation density of photoplates	VIII.1
$D_x$	relative scale factor	IV.2.4
$E$	electron energy	VII.4, IX.2
$E$	energy	IV.2.5, V.1, VII.2, VIII.1
$\mathcal{E}$	energy with respect to the Fermi level	VII.2
$E$	incident electron energy	IV.2.3
$E$	mean grey level of edges; expectation value	VIII.2
$E$	tip-sample interaction energy	VII.4
$e$	electron charge	IV.2.2, IV.2.3, VI.2, VII.2, VII.4, VIII.1, IX.1, IX.2
$E(z)$	on-axis electric field	IX.2
$E_0$	energy of a primary electron	VIII.1, IV.2.5
$E_0$	incident electron beam energy	IV.2.1
$E_A$	Auger kinetic energy	IV.2.4
$E_{\text{abs}}$	volume density of absorbed energy	VIII.1
$E_b$	ground state binding energies	IV.2.4
$E_c$	energy of critical excitation band	IV.2.5
$E_d$	material „dead voltage“	IV.2.1
$E_{\text{eh}}$	energy required to form a one-hole electron pair	IV.2.1
$E_F$	Fermi energy	IV.2.5, VII.2, VII.4
$E_f$	mean formation energy of an electron-hole pair	VIII.1
$E_{\text{gap}}$	material band-gap energy electron	IV.2.1

Symbol	Designation	Chapter
$E_{\text{in}}$	spatial density of energy	VIII.1
$E_K, E_L, \dots$	energy of the Kth and Lth shells	IV.2.4
$E_{\text{kin}}$	kinetic energy	IV.2.4
$E_\mu$	Energy of the state $\Psi_m$	VII.2
$E_\nu$	Energy of the state $\Psi_n$	VII.2
$E_{\text{ph}}$	photon energy	VIII.1
$E_{\text{ph}, x}$	energy of an X-ray photon	VIII.1
$eu$	incident electron energy (eV)	IV.2.5
$F$	electric field strength	VI.2
$F$	F-number of a lense	VIII.1
$F$	figure of merit	V.2
$F$	force	VII.3
$F$	fraction of total scattering passing through the hole in the annular detector	IV.2.3
$F$	transverse electric field	IX.2
$F'$	overall force gradient	VII.4
$f$	focal length of a lense	VIII.1
$f$	function	IV.2.5
$f(E, T)$	energy distribution of electrons reaching the interface	VII.4
$f(K), f(s)$	atomic form factor for elastic scattering	IV.2.3
$f(u)$	atomic scattering factor	IV.2.2
$f(X)$	fraction of carrier which diffuse to junction	IV.2.1
$f(\chi)$	absorption factor	IV.2.5
$f(E)$	Fermi function	VII.2
$F_a$	Fano factor	VIII.1
$F_e$	electrical field strength	VIII.1
$F_i$	fluorescence correction for element $i$	IV.2.5
$F_z$	force acting along the z-axis	VII.4
$f_{\text{res}}$	resonance frequency of the cantilever	VII.3
$f_{\text{sp}}$	resonance frequency of spring	IX.2
$f_{\text{tip}}$	resonance frequency of tip	IX.2
$G$	Fourier component	VII.2
$G$	spectral matching factor for a phosphor sensor	VIII.1
$g$	normalization factor	IV.2.5
$G(u, \rho)$	a four-dimensional function	IV.2.2
$G(x, y)$	gray level gradient magnitude	VIII.2
$g_m$	quantum yield of a converting stage	VIII.1
$H$	transfer function; parameter	VIII.2
$h$	constant	VII.4
$h$	impulse response	VIII.2
$h$	Planck's constant	IV.2.1, IV.2.4, VII.2, VIII.1, IV.2.2
$\hbar$	Dirac's constant	V.1, VII.2
$\hbar$	Planck's constant	IV.2.3
$H(r)$	magnetic field above a sample	VII.4
$H(x^{-1}, y^{-1})$	transfer function of an imaging instrument	VIII.2
$H_x$	stray magnetic force along the x-axis	VII.4
$H_z$	stray magnetic force along the z-axis	VII.4
$h(x, y)$	impulse response of an imaging instrument	VIII.2



Symbol	Designation	Chapter
$I$	current	VII.4, VI.2
$I$	current of primary electron beam	IX.1
$I$	Fourier transform of $i$	VIII.2
$I$	nuclear spin	V.1
$I$	reflected intensity	V.3
$I$	spin angular momentum	V.1
$I$	tunneling current	VII.2
$i$	image function	VIII.2
$I(\Theta)$	angular distribution of backscattered current in Eq. 3	V.2
$I(\mathbf{R}_0)$	Image intensity for probe in position $\mathbf{R}_0$	IV.2.3
$I(\mathbf{R}_0, t)$	Thickness dependent image intensity incorporating dynamical effects	IV.2.3
$I(x)$	intensity of hologram	IX.2
$i(x, y)$	image function	VIII.2
$I(z)$	Intensity of electron beam as a function of depth within sample	IX.2
$I, I'$	pixel grey level	VIII.2
$I_M^\pm$	secondary in current of species $M$	IV.2.6
$I_0$	current of primary electron beam	IV.2.4
$I_0$	Intensity of electron beam incident on sample	IX.2
$I_0$	polarization-independent intensity	V.3
$I_{00}$	incident intensity	V.3
$I_A$	Auger yield	IV.2.4
$I_A/I_B$	Auger current ratio for elements A and B	IV.2.4
$I_{Ag}$	peak-to-peak amplitude of silver standard	IV.2.4
$I_B$	current required in the beam	IV.2.1
$I_b$	ballistic electron emission microscopy current	VII.4
$I_b$	current contained in electron-beam probe	IV.2.1
$I_c$	collector current	VII.4
$I_{cc}$	charge collected signal	IV.2.1
$I_{CL}$	brightness dependence of cathodoluminescence signal	IV.2.1
$I_{ex}$	exchange intensity	V.3
$I^H(\mathbf{R}_0)$	Image intensity using a high angle annular detector	IV.2.3
$I_i$	current induced in segment $i$ of a photodiode	VII.3
$I_{sp}^i$	intensity of line $i$ in specimen	IV.2.5
$I_{st}^i$	intensity of line $i$ in standard	IV.2.5
$I_i, I_j$	X-ray intensity of line $i, j$	IV.2.5
$I_k$	$k$ th ionization energy	VI.2
$I_P$	primary in current	IV.2.6
$I_s$	signal current of a camera tube	VIII.1
$I_t$	injection current	VII.4
$I_X$	peak-to-peak amplitude of element X	IV.2.4
$I_X(u)$	intensity distribution	IV.2.2
$j$	current density	IX.2
$j$	electron current density	VIII.1
$j$	inner quantum number	IV.2.5
$J(\mathbf{r}-\mathbf{r}_i)$	exchange coupling strength	V.3
$j, k, l$	energy levels relating to multiplet splitting and electronic relaxation	IV.2.4
$j_{\mu, \nu}$	current density between the states $\Psi_m$ and $\Psi_n$	VII.2
$K$	attenuation of vibration	IX.2
$K$	instrumental constant	IV.2.4

Symbol	Designation	Chapter
$K$	transverse component of the incident wavevector $\chi$	IV.2.3
$k$	Boltzmann's constant	V.1, VIII.1, VII.4
$k$	constant relating electric field at specimed apex to the applied voltage	VI.2
$k$	electron wave vector	V.3
$k$	scattering vector	IV.2.3
$k$	spring constant	VII.4
$K, L, M, \dots$	atomic energy shells	IV.2.4
$k_{\parallel}$	electron momentum parallel to interface	VII.4
$k_B$	Boltzmann's constant	VI.2
$k_B$	Boltzmann's constant	VII.2
$K_c$	cut-off wavevector	IV.2.3
$K_f$	transverse component of the final state wavevector	IV.2.3
$k_{\text{eff}}$	effective spring constant	VII.4
$k_a^i$	$I_{\text{sp}}/I_{\text{st}}$ ratio for element $a$	IV.2.5
$k_i, k_f$	incident and scattered wave vectors, respectively	V.2
$k_{ij}$	Cliff-Lorimer sensitivity factor	IV.2.5
$k_z, k_A, k_F$	atomic number, absorption, fluorescence correction factor	IV.2.5
$L$	attenuation length; supercell period	IX.2
$L$	minority carrier diffusion length	IV.2.1
$L$	orbital angular momentum	V.2
$L$	total length of boundaries	VIII.2
$l$	azimuthal quantum number	IV.2.5
$L_c$	longitudinal coherence length	IX.2
$l_c$	coherence length	IV.2.3
$M$	atomic mass	IV.2.6
$M$	Debye-Waller factor	IV.2.3
$M$	instrumental magnification; linear magnification	IV.2.1
$M$	magnetic moment	V.1
$M$	magnetization	V.3
$M$	magnification	IX.2
$M$	scaling factor, molecular mass	VIII.1
$M$	tip magnetization	VII.4
$m$	cantilever mass	VII.4
$m$	free electron mass	VII.4
$m$	mass	IV.2.3, VII.2, IV.2.2, VI.2
$\bar{m}_s$	electron effective mass in a semiconductor	VII.4
$M_{\mu, n}$	tunneling matrix element between the states $\Psi_m$ and $\Psi_n$	VII.2
$M_x, M_y, M_z$	components of magnetization $M$	V.2
$N$	Avogadro's number	IV.2.5
$N$	density of atoms	IX.2
$N$	Fourier transform of $n$	VIII.2
$N$	number of primary quanta	VIII.1
$N$	total number of atoms; number of image rings	VI.2
$n$	additive noise function	VIII.2
$n$	atomic density	IX.1
$n$	constant	IV.2.1
$n$	diffraction order	IV.2.4

Symbol	Designation	Chapter
$n$	number of atoms in a column	IV.2.3
$n$	number of electrons removed during ionisation of an atom	VI.2
$n$	number of occupied d orbitals	V.3
$n$	principal quantum number; Bragg constant	IV.2.5
$n$	spatial density of primary quanta	VIII.1
$\bar{n}$	unit vector normal to the scattering plane	V.2
$\bar{n}$	average number of scattering events	VII.4
$N$	number of electrons detected with spins parallel and antiparallel to $x$ direction	V.2
$n$	number of spins per unit volume parallel and antiparallel to a particular orientation	V.2
$N(d)$	Fresnel number	IX.2
$N(E)$	distribution of incident particles	IV.2.4
$N(E) d(E)$	number of photons with energy in the range $E$ to $E + dE$	IV.2.5
$n_0$	refractive index of a scintillator	VIII.1
$n_1$	refractive index of fiber-optic core	VIII.1
$n_2$	refractive index of a fiber-optic coat	VIII.1
$N_A$	analyte concentration	IV.2.4
$NA$	numerical aperture	VIII.1
$N_A, N_B$	number of points inside structures A, B	VIII.2
$n_B$	number of Bohr magnetons per atom	V.2
$n_e$	electron response of a single primary quantum	VIII.1
$N_L, N_R$	number of electrons scattered to left and right, respectively	V.2
$n_{\max}$	full-well capacity of a charge-coupled device ( $e^- \text{ pixel}^{-1}$ )	VIII.1
$N_p$	number of primary quanta on a pixel	VIII.1
$n_{ph}$	photon response of a single primary quantum	VIII.1
$n_r$	read-out noise ( $e^- \text{ pixel}^{-1}$ )	VIII.1
$n_s$	number of atoms of species $s$	VI.2
$n_s$	number of secondaries per primary quantum	VIII.1
$n_s(E)$	density of states for the sample	VII.2
$n_t(E)$	density of states for the tip	VII.2
$n_v$	number of valence electrons per atom	V.2
$O$	Fourier transform of $o$	VIII.2
$o$	object function	VIII.2
$O'(\mathbf{K})$	Reciprocal space object function for inelastically scattered electrons	IV.2.3
$O'(\mathbf{R})$	Real space object function for inelastically scattered electrons	IV.2.3
$O(\mathbf{K})$	Reciprocal space object function for elastically scattered electrons	IV.2.3
$O(\mathbf{R})$	real space object function for elastically scattered electrons	IV.2.3
$O(\mathbf{R}, t)$	thickness dependent object function incorporating dynamical effects	IV.2.3
$O(x^{-1}, y^{-1})$	object spectrum = Fourier transform of the object function	VIII.2
$o(x, y)$	object function	VIII.2
$O^B(\mathbf{R})$	Born approximation object function	IV.2.3
$O^{TDS}(\mathbf{K})$	Reciprocal space object function for thermally scattered electrons	IV.2.3
$O^{TDS}(\mathbf{R})$	real space object function for thermally scattered electrons	IV.2.3
$P$	degree of polarization	V.3
$P$	perimeter	VIII.2
$P$	polarization	V.3
$\mathbf{P}$	polarization vector	V.2
$p$	Complement of transmission function	IX.2

Symbol	Designation	Chapter
$p$	probability of an injected electron reaching the interface	VII.4
$p$	transition probability	IV.2.5
$P(\mathbf{R})$	probe amplitude profile	IV.2.3
$p(s)$	probability distribution of a stochastic quantity $s$	VIII.1
$P_0$	polarization of beam of $N_0$ electrons	V.2
$P^2(\mathbf{R})$	probe intensity profile	IV.2.3
$P_{\text{eff}}^2(\mathbf{R})$	effective probe intensity profile	IV.2.3
$P_{\text{fb}}$	effective spin polarization of the ferromagnetic/barrier interface	VII.2
$P_x, P_y$	polarization components	V.2
$Q$	Fourier transform of transmission function	IX.2
$Q$	ionization cross-section of atoms $a$	IV.2.5
$Q$	quality of the cantilever	VII.3
$q$	electric charge	IV.2.4
$q$	electron charge density	VIII.1
$\mathbf{q}$	momentum transfer on inelastic scattering	IV.2.3
$\mathbf{q}$	scattering vector	IX.2
$q(x, y)$	transmission function	IV.2.2
$q(x)$	transmission function	IX.2
$Q_{0(n)}$	energy required to remove surface atom as an ion with charge $n$ in the absence of an electric field	VI.2
$Q_{\text{sp}}^a, Q_{\text{st}}^a$	ionization cross-section of element $a$ in specimen, standard	IV.2.5
$q_{\text{B}}$	Brillouin zone boundary	IV.2.3
$Q_n$	activation energy for surface atom removal as an ion with charge $n$ in the presence of an electric field	VI.2
$q_z$	momentum transfer parallel to the incident beam direction	IV.2.3
$R$	effective tip radius	VII.2
$R$	incident electron beam range	IV.2.1
$R$	radius	IV.2.5
$R$	reflectivity of the mirror layer	VIII.1
$R$	region	VIII.2
$R$	relative energy resolution; Auger backscattering factor	IV.2.4
$\mathbf{R}$	transverse positional coordinate on the specimen entrance surface	IV.2.3
$\mathbf{r}$	atomic position	IV.2.3
$r$	charge separation	VII.3
$\mathbf{r}$	incident electron position	V.3
$r$	radius of curvature; radius of specimen	VI.2
$r$	tip radius	IX.2
$\bar{r}_0$	center of curvature of the effective tip	VII.2
$\mathbf{R}_0$	probe position	IV.2.3
$r_0$	radius of the object field	VIII.1
$r_1, r_2$	radii of concentric cylinders	IV.2.4
$R_{\text{sp}}^a, R_{\text{st}}^a$	electron backscattering of element $a$ in the specimen, standard	IV.2.5
$R_a, R_\beta$	count rate for events $a, \beta$	IX.1
$\mathbf{r}_i$	target electron position	V.3
$r_{\text{M}}(E_{\text{A}a})$	backscattering coefficient at angle $a$ to surface normal	IV.2.4
$r_{\text{mn}}$	separation of $m$ th and $n$ th atoms in a column	IV.2.3
$S$	dissolution rate of the resist development	VIII.1
$S$	sputter yield	IV.2.6
$S$	surface	VII.2

Symbol	Designation	Chapter
$s$	incident electron spin	V.3
$s$	$\mathbf{K}/4\pi$	IV.2.3
$s$	normalized arc length	VIII.2
$s$	single-quantum response	VIII.1
$s$	tip-sample distance	VII.2
$S(X)$	signal function	IV.2.2
$s^{(k)}$	best search direction	VIII.2
$s(\lambda)$	emission spectrum of a scintillator	VIII.1
$S(\Theta)$	Sherman function	V.2
$s(x, y)$	three-dimensional tip shape	VIII.2
$s, p, d, \dots$	orbitals within energy levels	IV.2.4
$S_{sp}^a, S_{st}^a$	electron stopping power of element $a$ in the specimen, standard	IV.2.5
$S_A, S_B$	area of structures A, B	VIII.2
$s_i$	target electron spin	V.3
$Si(x)$	sine integral function	IV.2.3
$S_{in}$	input signal	VIII.1
$s^j(\mathbf{K})$	transverse energy of Bloch state $j$	IV.2.3
$SNR$	signal-to-noise ratio	V.2
$S_{out}$	output signal	VIII.1
$S_{pc}$	sensitivity of a photocathode	VIII.1
$T$	temperature	VI.2, V.1, VII.2, VII.4
$T$	transmission coefficient of light optics; absolute temperature	VIII.1
$T$	transmission efficiency to detector in Eq. 1	IX.1
$t$	flight time	VI.2
$t$	specimen thickness	IV.2.3, IX.1
$t$	time	IV.2.6, V.1
$T(u)$	transfer function	IV.2.2
$T(u)$	transfer function of the lens	IV.2.2
$t(x), t(x, y)$	spread function	IV.2.2
$T_{(E_A)}$	transmission efficiency of analyzer at $E_A$	IV.2.4
$T(E)$	transmission coefficient	VII.2
$t(x)$	spread function	IV.2.2
$T_1$	spin-lattice relaxation time	V.1
$T_2$	spin-spin relaxation time	V.1
$T_c$	critical temperature	IV.2.3
$t_d$	developing time	VIII.1
$T_{max}(U)$	maximum transmission coefficient	VII.2
$T_{z1}$	inverse of $t_{z1}$	IX.2
$t_{z1}$	parabolic approximation	IX.2
$U$	accelerating voltage	VIII.1
$U$	applied sample bias current	VII.2
$U$	overvoltage	IV.2.5
$U$	Reciprocal space coordinate	IX.2
$u$	$(2/\lambda) \sin(\phi)$	IV.2.2
$u$	deviation from atomic position due to thermal motion	IV.2.3
$u$	spatial frequency	VIII.1
$U_0$	beam potential	IX.2
$V$	valence band	IV.2.4

Symbol	Designation	Chapter
$V$	voltage	VII.4
$V$	volume	VII.4
$v$	electron velocity	IV.2.3
$v$	final velocity of ion	VI.2
$V_0^i$	imaginary part of mean inner potential	IX.2
$V_0$	threshold voltage	VII.4
$V(\mathbf{R})$	specimen projected potential	IV.2.3
$V(r)$	central potential	V.2
$V_0$	applied voltage	VI.2
$V_1, \dots V_n$	electrode potentials on a series of electrodes	IV.2.4
$V_A, V_B$	volume of structures A, B	VIII.2
$V_b$	ballistic electron emission microscopy voltage	VII.4
$V_{ex}$	exchange potential	V.3
$V_t$	injection voltage	VII.4
$v_i$	potential difference, $V_i - V_0$	IV.2.4
$W_{mn}$	correlation factor between mth and nth atoms in a column	IV.2.3
$\mathbf{X}$	a multiplet state resulting from coupling two holes, $k$ and $l$ , in the final state	IV.2.4
$\mathbf{X}$	data matrix	VIII.2
$X$	distance between incident beam and junction	IV.2.1
$X$	spatial coordinate on detector	IX.2
$x^{-1}, y^{-1}$	spatial frequency coordinates	VIII.2
$x, y$	local window coordinates	VIII.2
$x, y$	spatial coordinates	VIII.2
$x_0$	distance from specimen surface	VI.2
$X_c$	coherence width	IX.2
$x_c$	critical distance	VI.2
$X_n$	lateral position of $n$ th fringe	IX.2
$\mathbf{X}^t$	transpose of $\mathbf{X}$	VIII.2
$\mathbf{Y}$	variance-covariance matrix	VIII.2
$Z$	atomic number	IV.2.1, IV.2.3, IV.2.5, VIII.1
$z$	cantilever deflection	VII.3
$z$	coordinate on $z$ axis	IX.2
$z$	depth	IV.2.5, IV.2.3
$z$	height of object detail	VIII.2
$z$	width within which Auger emission occurs	IV.2.4
$Z(x, y)$	atomic number	IV.2.2
$z_1$	tip-to-sample distance	IX.2
$z_2$	Distance from sample to detector screen in point projection microscope	IX.2
$z_2$	sample-to-screen distance	IX.2
$\alpha$	absorption coefficient of X-rays	VIII.1
$\alpha$	angle subtended by source at the sample	IX.2
$\alpha$	angle to surface normal	IV.2.4
$\alpha$	phase difference between two reflections	IV.2.2
$\alpha$	weighting coefficient	VIII.2
$\alpha^\pm$	transmission efficiency	IV.2.6
$\alpha, \beta$	events $\alpha, \beta$	IX.1
$\beta$	acceptance half-angle of light optics	VIII.1
$\beta$	brightness of electron gun	IV.2.1

Symbol	Designation	Chapter
$\beta$	coherence angle	IX.2
$\beta$	image compression factor	VI.2
$\beta_I$	image half-angle of a micro-objective	VIII.1
$\chi$	$(\mu r)\text{cosec}\psi$	IV.2.5
$\chi$	incident electron wavevector	IV.2.3
$\chi(u)$	phase factor	IV.2.2
$\Delta$	measure of defocus	IV.2.2
$\delta$	continuum fluorescence contribution/primary characteristic X-ray intensity ratio	IV.2.5
$\delta$	corrugation	VII.2
$\delta$	degeneracy	IX.2
$\delta$	instrumental resolution	IV.2.2
$\delta$	number of secondary electrons per incident electron (-)	IV.2.1
$\delta$	resolution	VI.2
$\delta$	resolution of emission electron microscope	VIII.1
$\delta(\mathbf{R})$	Dirac delta function	IV.2.3
$\Delta_0$	path difference	IX.2
$\Delta E$	change in energy	IV.2.3
$\Delta E$	energy deposit of a primary quantum in a solid	VIII.1
$\Delta E$	energy loss	IV.2.5
$\Delta E$	energy spread	IV.2.4
$\Delta E$	energy range	IX.2
$\Delta E_{\text{pe}}$	energy width of photoelectrons	VIII.1
$\Delta f$	objective lens defocus	IV.2.3
$\delta_{mn}$	Kroenecker delta	IV.2.3
$\Delta R$	transverse separation of two objects	IV.2.3
$\delta_s$	secondary emission coefficient	VIII.1
$\Delta W$	Solid angle	IX.2
$\varepsilon$	efficiency of the SEMPA system in Eqs. 6,7	V.2
$\varepsilon$	energy conversion efficiency of a scintillator	VIII.1
$\varepsilon(\omega)$	optical dielectric function	IX.2
$\varepsilon_0$	permittivity of free space	VI.2
$\varepsilon'(K)$	excitation coefficient of Bloch wave j	IV.2.3
$\Phi$	work function	IV.2.4
$\phi$	density of states at the Fermi level for the tip	VII.2
$\phi$	scattering angle	IV.2.2
$\phi$	Schottky barrier height	VII.4
$\phi$	work function	VI.2
$\phi(\Delta rz)$	emission from an isolated thin film of mass thickness $\Delta rz$	IV.2.5
$\Phi_A$	apparent barrier height	VII.2
$\Phi_P$	primary current density	IV.2.6
$\Phi_s$	sample work function	VII.2
$\Phi_t$	tip work function	VII.2
$\gamma$	decay rate	VII.2
$\gamma$	fluorescence intensity/primary characteristics X-ray ratio	IV.2.5
$\gamma$	quantum yield for impact ionization	VII.4
$\gamma$	gyromagnetic ratio	V.1
$\gamma$	slope of the characteristic curve of a recorder	VIII.1
$\gamma^\pm$	ionization probability	IV.2.6

Symbol	Designation	Chapter
$\sigma$	conductance of a tunnel junction	VII.2
$\sigma$	cross section of spin-dependent scattering	V.2
$\sigma$	electron interaction constant	IV.2.3
$\sigma$	interaction constant	IV.2.2
$\sigma$	Lennard's constant	IV.2.5
$\sigma$	standard deviation	VI.2, VIII.2
$\sigma_{\uparrow\uparrow}$	conductance of a tunnel junction with parallel internal magnetic field directions	VII.2
$\sigma_{\uparrow\downarrow}$	conductance of a tunnel junction with antiparallel internal magnetic field directions	VII.2
$\sigma_A(E_p)$	electron impact ionization cross section at primary energy $E_p$	IV.2.4
$\sigma_a, \sigma_b$	cross section for events $a, b$	IX.1
$\sigma_{at}$	atomic cross-section for scattering to a high angle annular detector	IV.2.3
$\sigma_{fbf}$	mean conductance of a tunnel junction	VII.2
$\sigma_t$	total inelastic cross-section	IX.2
$\sigma_{true}$	cross section for the combined events $a, \beta$	IX.1
$\tau$	pixel dwell time in Eqs. 6,7	V.2
$\tau$	storage time	VIII.1
$\tau$	time window	IX.1
$\tau^j(\mathbf{K}, \mathbf{R})$	two-dimensional Bloch state $j$ of the fast electron	IV.2.3
$\Omega$	solid acceptance angle	IV.2.4
$\Omega$	acceptance solid angle of light optics	VIII.1
$\omega$	fluorescence yield; X-ray emission yield	IV.2.5
$\omega_0$	photon energy	V.1
$\psi$	angle between emitted X-rays and the specimen surface	IV.2.5
$\psi(\mathbf{R}, z)$	electron wavefunction within the specimen	IV.2.3
$\Psi_h$	Electron wavefunction detected in point projection microscope	IX.2
$\Psi_i$	Electron wavefunction across downstream face of sample	IX.2
$\Psi_i(x)$	electron wavefunction	IX.2
$\psi_m$	electronic state of the tip	VII.2
$\Psi_n$	electronic state of the sample	VII.2
$\psi_s$	amplitude of elastic scattering in the first Born approximation	IV.2.3

Abbreviation	Explanation
3DAP	three-dimensional atom probe
AC	alternating current
ACP	automatic correlation partitioning
ADC	analog-to-digital convertor
ADF	annular dark field
AES	Auger electron spectroscopy
AFM	antiferromagnetism
AP	atom probe
APB	antiphase boundary
BBT	butylbenzotriamide
BF	bright field
BIF	best image field



Abbreviation	Explanation
BSE	backscattered electrons
CCD	charged coupled device
CFD	constant fraction discriminator
CHA	concentric hemispherical analyzer
CID	charge injection device
CL	cathodoluminescence
CLD	constant level discriminator
CMA	cylindrical mirror analyzer
CRT	cathode ray tube
DC	direct current
DD	detector defined
DQE	detective quantum efficiency
EBD	electron beam deposited
EBIC	electron beam induced conductivity
EBIC	electron-beam-induced current
ECL	emitter-coupled logic
fd	force-distance
FEG	field emission gun
FEM	field emission microscope
FID	free induction decay
FIM	field ion microscope
FM	ferromagnetism
FWHM	full width at half-maximum
GL	gray level
HAADF	high-angle annular dark field
HDTV	high-definition television
HOLZ	High order Laue zone
HREM	high resolution electron microscopy
ICP	interactive correlation partitioning
ISIT	intensified silicon intensifier target
LDOS	local density of states
LEDS	low energy diffuse scattering
LEED	low-energy electron diffraction
LIMA	laser ionization mass analysis
LLL	low light level
MCA	multichannel analyzer
MCD	magnetic circular dichroism
MCP	microchannel plate
MCS	multichannel scaler
MIS	metal/insulator/semiconductor
MISR	matrix/ion species ratio
MLCFA	maximum likelihood common factor analysis
MMF	minimum mass fraction
MOSFET	metal oxide semiconductor field effect transistor
MSA	multivariate statistical analysis
PC	personal computer
PHD	pulse height distribution
PLAP	pulsed laser atom probe
PMMA	polymethylmethacrylate

Abbreviation	Explanation
PMT	photomultiplier tube
PoSAP	position-sensitive atom probe
PSF	point spread function
QSE	quantum size effect
QUANTITEM	quantitative analysis of the information from electron micrographs
RAE	resistive anode encoder
RF	radio-frequency
SCA	single channel analyzer
SD	source defined
SE	secondary electrons
SEC	secondary electron conduction
SED	secondary electron detectors
SIT	silicon intensifier target
SKIZ	skeleton by influence zones
SNR	signal-to-noise ratio
SSCCD	slow-scan charge-coupled device
TAC	time-to-amplitude converter
TAP	topographic atom probe
TOF	time of flight
TTL	transistor-transistor logic
TV	television
UHV	ultrahigh vacuum
VCR	video cassette recorder
YAG	yttrium aluminum garnet

Techniques	Explanation
AEEM	Auger electron emission microscopy
AEM	analytical electron microscopy
AES	atomic emission spectroscopy
AES	Auger electron spectroscopy
AFM	atomic force microscopy
APFIM	atom probe field ion microscopy
ARPES	angle-resolved photoelectron spectroscopy
ATRS	attenuated total reflectance spectroscopy
BEEM	ballistic electron emission microscopy
BEES	ballistic electron emission spectroscopy
BF CTEM	bright field conventional transmission electron microscopy
CBED	convergent beam electron diffraction
CBRHEED	continuous beam reflective high-energy electron diffraction
CCSEM	computer-controlled scanning electron microscopy
CITS	current imaging tunneling spectroscopy
CL	cathodoluminescence
CLSM	confocal laser scanning microscopy
CT	computer-aided tomography
CTEM	conventional transmission electron microscopy
CVD	chemical vapor deposition

Techniques	Explanation
DLTS	deep level transient spectroscopy
EBT	electron beam testing
EDS	electron diffraction spectrometry
EDS	energy-dispersive spectroscopy
EDX	energy dispersive X-ray spectroscopy
EELS	electron energy loss spectroscopy
EFTEM	energy filtered transmission electron microscopy
EM	electron microscopy
EPMA	electron probe microanalysis
EPXMA	electron probe X-ray microanalysis
ESCA	electron spectroscopy for chemical analysis
ESEM	environmental scanning electron microscopy
ESI	electron spectroscopic imaging
ESI	element-specific imaging
FFM	friction force microscopy
FIB	focused ion beam milling
FIM	field ion microscopy
FMT	fluorescent microthermography
FT-IR	Fourier transform infrared spectroscopy
HREM	high resolution electron microscopy
HRSEM	high resolution scanning electron microscopy
HRTEM	high resolution transmission electron microscopy
HVEM	high voltage electron microscopy
LACBED	large angle convergent beam electron diffraction
LCT	liquid crystal thermography
LEEM	low-energy electron microscopy
LFM	lateral force microscopy
LM	light microscopy
LMMS	laser microprobe mass spectrometry
LOM	light optical microscopy
LPCVD	low-pressure chemical vapor deposition
LTSLEM	low-temperature scanning laser electron microscopy
M-PIXE	micro-(proton-induced X-ray emission spectroscopy)
MBE	molecular beam epitaxy
MEM	mirror electron microscopy
MFM	magnetic force microscopy
MOVPE	metal-organic vapor phase epitaxy
MRI	magnetic resonance imaging
MULSAM	multispectral Auger microscopy
NMR	nuclear magnetic resonance
OM	optical microscopy
PCA	principal components analysis
PEELS	photoelectron energy loss spectroscopy
PEEM	photoemission electron microscopy
PFA	principal factor analysis
PIXE	proton induced X-ray emission spectroscopy
PL	photoluminescence
PPM	point-projection microscopy
RBS	Rutherford backscattering spectroscopy

Techniques	Explanation
RDE	reactive deposition epitaxy
REM	reflection energy microscopy
REM	reflection electron microscopy
RHEED	reflection high-energy electron diffraction
SAM	scanning acoustic microscopy
SAM	scanning Auger microscopy
SAXS	small-angle X-ray scattering
SCM	scanning capacitance microscopy
SDLTS	scanning deep level transient spectroscopy
SECM	scanning electrochemical microscopy
SEELS	serial electron energy-loss spectroscopy
SEEM	secondary electron emission spectroscopy
SEM	scanning electron microscopy
SEMPA	scanning electron microscopy with polarization analysis
SFM	scanning force microscopy
SIMS	secondary ion mass spectrometry
SLEEM	slow low-energy electron microscopy
SNOM	scanning near-field optical microscopy
SNPM	scanning near-field probe microscopy
SPE	solid phase epitaxy
SPLEED	spin-polarized low-energy electron diffraction
SPLEEM	spin-polarized low energy electron microscopy
SPM	scanning probe microscopy
SPSTM	spin-polarized scanning tunneling microscopy
SQUID	superconducting quantum interference device
SREM	scanning reflection electron microscopy
STEM	scanning transmission electron microscopy
STM	scanning tunneling microscopy
STOM	scanning tunneling optical microscopy
STS	scanning tunneling spectroscopy
STXM	scanning transmission X-ray microscopy
TED	transmission electron diffraction
TEEM	thermionic electron emission microscopy
TEM	transmission electron microscopy
TL	thermoluminescence
TS	tunneling spectroscopy
TSMFM	tunneling stabilized magnetic force microscopy
TXM	transmission X-ray microscopy
UFM	ultrasonic force microscopy
UMT	ultra microtomography
VPE	vapor phase epitaxy
WDS	wavelength dispersive spectroscopy
XES	X-ray emission spectroscopy
XPS	X-ray photoelectron spectroscopy
XPS	X-ray photoemission spectroscopy
XTEM	cross-sectional transmission electron microscopy



# Index

- adhesion, scanning force microscopy 839
- adsorbates, scanning tunneling microscopy 811
- advanced universal Auger electron microscope, experimental set-up 677
- Ag–O–Cs, photoemission threshold 765
- $\text{Al}_x\text{Ga}_{1-x}\text{As}$ /GaAs, ballistic electron emission microscopy 875
- aliasing 905
- Alnico 2, atom probe field ion microscopy 794
- Alnico 5, atom probe field ion microscopy 791
- alternating current mode, magnetic force microscopy 849
- amorphous carbon, coincidence microscopy 960
- analog detection, scanning electron microscopy with polarization analysis 743
- analytical electron microscopy, experimental set-up 674
- annular dark-field imaging
  - heavy atoms 564, 568
  - high angle 567
  - platinum on zeolite 568
  - uranium 564
- antiphase boundaries, field ion microscopy 785
- area analysis, energy-dispersive spectrometry 681
- area-perimeter method 939
- artefacts 837 f
  - scanning Auger microscopy 654
- astrology, atom probe field ion microscopy 800
- atom probe field ion microscopy 775 ff
  - experimental set-up 789
  - microanalysis 788 ff
  - semiconductors 795
  - three-dimensional 795
- atomic bonding, scanning force microscopy 840
- atomic number correction, quantitative X-ray scanning microanalysis 683
- attenuation length, point-projection microscopy 967
- Au/CaF<sub>2</sub>/Si(111), ballistic electron emission microscopy 876
- Au/GaAs(100)
  - heterojunction, ballistic electron emission microscopy 858, 859
  - interfaces, ballistic electron emission microscopy 877
- Au/Si
  - heterojunction, ballistic electron emission microscopy 858
  - interfaces, ballistic electron emission microscopy 868
- Auger electron microanalyzers 677 ff
- Auger electron microscopy, scanning see scanning Auger microscopy
- Auger electron spectroscopy 621 ff
  - beam effect 625
  - detection limit 627
  - information depth 628
  - maximum likelihood common factor analysis 629
  - quantitative analysis 627
  - sample preparation 625
- Auger electrons 622
- Auger spectrometers, coincidence microscopy 956
- Auger spectrum
  - line shape 624
  - notations 624
- Auger transition, kinetic energy 623, 624
- automatic correlation partitioning 948
- background
  - Auger electron spectroscopy 630
  - slope effects 656
  - X-ray photoelectron spectroscopy 630
- back-illumination, charge-coupled devices 903, 912
- back-projection, 3D image reconstruction 945
- backscattered electron imaging 548 ff
  - applications 549
  - atomic number contrast imaging 549
  - detectors 548
  - limitation 548 f
- backscattered electrons
  - range 550
  - yield 548
- backscattering coefficient 550
- bacteriorhodopsin, point-projection microscopy 982, 983
- ballistic electron emission microscopy 855 ff
  - applications
    - Au–Si interfaces 862
    - epitaxial interfaces 865
    - metal/insulator/semiconductor structures 876
    - Schottky barrier interfaces 864 ff
  - carrier spectroscopy 879
  - carrier transport 880
  - comparison with STM topographic imaging 858 f
  - experimental set-up 856 ff
  - hole transport 862, 879

- instrumentation 857
- quantitative modeling 864
- theory 860
- tunneling effects 872
- beam current fluctuation effects, scanning Auger microscopy 657
- beam deflection technique, scanning force microscopy 831
- beam effect
  - Auger electron spectroscopy 625
  - X-ray photoelectron spectroscopy 626
- Beamson approach 654
- best image field, field ion microscopy 782
- bias voltage, scanning tunneling microscopy 809
- binary mathematical morphology 942
- biological applications
  - imaging secondary ion mass spectrometry 713
  - slice selected, two-dimensional NMR imaging 732 f
- Bitter method 736
- blanket method 939
- borosilicate glass, scanning microanalysis 668
- bremsstrahlung 669
- bright field imaging mode, transmission electron microscopy 56 ff
- brightness
  - low energy electron holography 974
  - point-projection microscopy 974
- bulk magnetic structure, scanning electron microscopy with polarization analysis 738
- buried interfaces 856
- butylbenzotriamide, secondary ion mass spectrometry 711
- CaF<sub>2</sub>/silicon interface, channeling pattern, selected area 552
- Cameca IMS secondary ion mass spectrometer
  - experimental set-up 698
  - ion microscope mode 699
  - mass range 699
- Cameca TOF secondary ion mass spectrometer
  - detection system 701
  - experimental set-up 700
  - ion sources 701
  - mass resolution 701
- cantilever
  - piezoresistive 832
  - requirements for SFM 830
- cantilever deflection
  - scanning force microscopy 836
- cantilever displacement 852 f
  - capacitively controlled 852
  - optical detection 851
  - tunneling stabilized magnetic force microscopy 852
- capacitance sensors
  - magnetic force microscopy 852
- scanning force microscopy 832
- capillary forces, scanning force microscopy 840
- carbon nanotubes
  - nanodiffraction patterns 577
- scanning transmission electron microscopy 591
- carrier diffusion length 556
- carrier transport, ballistic 880
- catalyst
  - photoelectron emission microscopy 770
- supported, scanning tunneling electron microscopy 578
- Z-contrast scanning transmission electron microscopy 618
- cathodoluminescence
  - II–VI compound semiconductors 558
  - modes of operation 560
  - scanning electron microscopy 559
  - experimental set-up 559
  - signal collection 559
  - spatial resolution 560
- CdTe(001)–GaAs(001) interface
  - edge dislocation in 611
  - Z-contrast scanning transmission electron microscopy 611 f
- ceramics, Z-contrast imaging 613
- Chalnicon tubes, image recording 900
- channeling pattern
  - backscattered electron imaging 550
  - CaF<sub>2</sub>/silicon interface 552
  - information about crystal structure 551
  - limitations 552
  - selected area 552
- charge collection mode
  - electron beam induced channeling 554
  - semiconductor devices 555
- charge-coupled device 891, 901 ff, 912 ff
  - converters 910 f
  - technical specifications 911
  - detectors 585, 911
  - energy-dispersive spectroscopy 918
  - fiber plate-coupled 911
  - illumination 912, 981 f
  - image recording 894, 901 f
  - scanning transmission electron microscopy 585
  - slow scan 909, 917
  - Tektronix, performance figures 903
- chemical composition determination, quantitative
  - Auger electron spectroscopy 630
  - X-ray photoelectron spectroscopy 630
- chemical force microscopy 837
- chemical polishing 788
- chemical shifts, Auger electron spectroscopy 625
- chromatic contrast 706

- Cliff–Lorimer sensitivity factor 687
- Co–Pd multilayer
  - domain structure 851
  - magnetic force microscopy 850f
- Co on Mo{011}, photoelectron emission microscopy 768
- cobalt 752
  - quantum size effects 752
  - spin-polarized low-energy electron microscopy 753, 756f
- coherence angle, point-projection microscopy 974
- coherent nanodiffraction, scanning transmission electron microscopy 578
- coincidence microscopy 955ff
  - Auger spectrum 961
  - collection efficiency 956
  - count rates 958
  - instrumentation 956
  - signal combinations
    - backscattered electron–backscattered or secondary electron 962
    - electron energy loss spectroscopy – cathodoluminescence 961
    - electron energy loss spectroscopy – emitted electron 958
    - electron energy loss spectroscopy – X-ray 961
- columnar coherence, Z-contrast scanning transmission electron microscopy 602
- composition analysis, field ion microscopy 793
- composition–composition histogram 681
- computer simulation, scanning force microscopy 841
- II–VI compound semiconductors 558
- concentric hemispherical analyzer 637f
- conditional region-growing approach 936
- $\beta$ -conductivity 554f
- constant current mode, scanning tunneling microscopy 807, 819
- constant force gradient mode, magnetic force microscopy 849, 850
- constant force mode, scanning force microscopy 832, 833
- constant height mode, magnetic force microscopy 846
- continuous beam reflective high-energy electron diffraction 587
- contrast
  - chromatic 706
  - crystallographic, secondary ion mass spectrometry 706
  - enhancement 928, 930
    - algorithm for 928
  - field emission microscopy 777
  - matrix
  - scanning force microscopy 841
- convergent-beam microdiffraction
  - point-projection geometry 964
  - transmission geometry 964
- convergent-beam scanning transmission electron microscopy 577
- CoPt, scanning electron microscopy with polarization analysis 746
- corrugation
  - dependence on tip–surface separation 813
  - scanning tunneling microscopy 812
- CoSi<sub>2</sub>–Si(111)
  - Z-contrast imaging 606
- CoSi<sub>2</sub>/Si, ballistic electron emission microscopy 865
- Coster–Kronig transitions 623
- count rate, coincidence microscopy 958
- current imaging tunneling spectroscopy 820
- current–voltage characteristics, scanning tunneling microscopy
  - fixed tip height 820
  - fixed tip–sample separation 820
- cylindrical mirror analyzer
  - double-pass 637
  - focusing properties 638
  - imaging properties as a function of energy 649
  - single-pass 637
  - use in X-ray photoelectron spectroscopy 639
- datolite, scanning microanalysis 668
- deep level transient spectroscopy 557
- defects in a bulk wafer, scanning electron microscopy 557
- deflection sensors
  - piezoresistive cantilevers 832
  - scanning force microscopy 831
- deflection type analyzers 634, 635
  - cylindrical sector fields 634
  - energy resolution 634
  - hemispherical analyzers 634
  - parallel plate analyzers 634
- depth profiling
  - atom probe field ion microscopy 793
  - secondary ion mass spectrometry 691
- detected quantum efficiency 886ff
- detection defined imaging
  - scanning Auger microscopy 632
  - X-ray photoelectron spectroscopy 649
- detection limit 627
  - Auger electron spectroscopy 628
  - X-ray photoelectron spectroscopy 627
- detector defined imaging
  - 2D position analysis 649
  - efficiency 649
  - X-ray photoelectron spectroscopy 644, 648
- diamond
  - dislocations, cathodoluminescence image 560



- planar defects, scanning transmission electron microscopy 590
- diatoms, scanning electron microscopy 546
- differential phase contrast microscopy 736
- digital frame store 909
- distance function 940
- dynamic processes, image recording 887
- dynamic secondary ion mass spectrometry 692
- edge effects, scanning Auger microscopy 657
- elastic forces, scanning force microscopy 840
- elastic scattering, incoherent imaging 598
- elastic tunneling spectroscopy 823
- electron backscattering factor, quantitative X-ray scanning microanalysis 683
- electron beam-deposited tips 848
- electron beam-induced conduction 896, 908
- electron beam-induced conductivity 554 ff, 662
  - $\beta$ -conductivity 555
  - p–n junction 555
  - Schottky surface barrier 555
- electron beam interactions, analytical electron microscopy 662
- electron channeling pattern, InP 551
- electron energy loss spectroscopy, incoherent structure imaging 612
- electron holography see holography
- electron microanalysis 661 ff
- electron microscopy see also main entries of individual techniques
  - imaging secondary ion mass spectrometry 691 ff
  - imaging X-ray photoelectron microscopy 621 ff
  - scanning Auger 621 ff
  - scanning beam methods 537 ff
  - scanning microanalysis 661 ff
  - scanning reflection 539 ff
  - scanning transmission 563 ffff
  - stationary beam methods see Volume 1 of this Handbook
  - Z-contrast scanning transmission 595 ff
- electron momentum conservation 862
- electron optical column, SEMPA analysis 739 f
- electron probe X-ray microanalysis 661 ff
  - instrumentation 669 f
  - low voltage 673
- electron scattering effects, ballistic electron emission microscopy 874
- electron scattering mechanisms, ballistic electron emission microscopy 869
- electron–solid interactions 663
- electron–specimen interactions, electron probe X-ray microanalysis 669 ff
- electron transport, ballistic electron emission microscopy 860, 872
- electron tunneling
  - ballistic electron emission microscopy 872
  - phonon spectroscopy 824
- electron tunneling technique, scanning force microscopy 832
- electronic image recording 908, 912
- electropolishing, field ion microscopy 788
- electrostatic forces, scanning force microscopy 841
- electrostatic transfer lens 642
- element mapping
  - position sensitive atom probe 799
  - scanning Auger microscopy 645
- elemental distributions, X-ray mapping 681
- energy compensation, atom probe field ion microscope 791
- energy-dispersive spectrometry 680
- energy-dispersive spectroscopy 666
  - comparison with wavelength dispersive spectroscopy 679
  - $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$  666
- energy dispersive X-ray detector, coincidence microscopy 956
- energy loss distribution, scanning transmission electron microscopy 575
- energy loss mechanism
  - ballistic electron emission microscopy 873
  - point-projection microscopy 966
- energy resolution, deflection type analyzers 635
- epitaxial growth, photoelectron emission microscopy 768
- epitaxial interfaces, ballistic electron emission microscopy 865, 871
- evaporation fields (table) 782
- Everhart–Thornley detector 545
- extrinsic luminescence 558
- F*-number 894
- Fe/Cr/Fe(100)
  - sandwich structure 747
  - scanning electron microscopy with polarization analysis 746 f
- ferromagnetic domains, photoelectron emission microscopy 769
- ferromagnetic materials
  - magnetic fields in 582
  - off-axis holography 583
  - scanning electron microscopy with polarization analysis 735
  - spin-polarized low-energy electron microscopy 752
- Fe–SiO<sub>2</sub>, scanning electron microscopy with polarization analysis 747, 748
- fiber-optic plates, image recording 894
- field absorption 784
- field-emission gun 586
- field emission microscopy 775 ff
  - applications 777 f

- experimental set-up 775 f
- field ion microscopy 775 ff, 781, 787
  - applications 777
  - atom probe 775 f
  - basic design 778
  - image formation 781
  - limitations 777
  - specimen preparation 787
- film electronic properties, ballistic electron emission microscopy 874
- filtered back-projection, 3D image reconstruction 945
- filtering 925 ff
  - convolution process 928
  - frequency space 927
  - frequential 925 f
- filters, nonlinear 926
- fluorescence correction, quantitative X-ray scanning microanalysis 684
- focus series approach, image restoration 932
- force–distance curve, scanning force microscopy 836 f
- force gradient measurements 849, 850
- force measurement 845
- forces relevant to scanning force microscopy 839
- Fourier imaging 968
- Fourier lenses
  - energy selection 651
  - imaging 651
- Fowler–Nordheim equation 775
- fractal analysis 939
- Fraunhofer holography 969
- Fraunhofer reconstruction methods, point-projection microscopy 965
- frequential filtering 925 f
- friction force microscopy 835
- frictional force mapping 835
- frictional forces 835, 840
- full-frame charge coupled device 901

## GaAs

- cathodoluminescence imaging 560
- dislocations 560
- doping variations 557
- electron beam induced channeling image 557
- spin-polarized scanning tunneling microscopy 823
- Z-contrast scanning transmission electron microscopy 610
- Gabor holography, limitations 580
- GaP(110)/Au, ballistic electron emission microscopy analysis 864
- GaP(110)/Mg, ballistic electron emission microscopy analysis 864
- garnet, magnetic force microscopy 846
- global image preprocessing 925, 929

- gradient echo, nuclear magnetic resonance microscopy 728, 733
- grain boundaries
  - field ion microscopy 784, 785
  - structure–property relationships 616
  - superconducting properties 615
  - Z-contrast scanning transmission electron microscopy 615
- grain boundary engineering 616
- granulometry 940, 941

## heavy atoms

- annular dark-field imaging 564, 567 ff
- on zeolites 589
- scanning transmission electron microscopy 587 ff
- high-angle detector, incoherent imaging 602
- high definition image recording 899
- holography see also Gabor holography
  - comparison with SEMPA 736
  - detectors 582 f
  - electron microscopy 582
  - Fresnel imaging 581
  - image reconstruction 583 ff
  - in-line 963 f
  - low energy electron 966
  - off-axis 582
  - reconstruction algorithms 970 ff
- hot-electron transfer processes, ballistic electron emission microscopy 856, 860 f, 864
- Hough transform 937
- Hurst coefficient method 939

## image analysis

- field ion microscopy 786
- mathematical morphology 936 ff
- image contrast, scanning transmission electron microscopy 570
- image conversion, quantum efficiency 889
- image density distribution, scanning transmission electron microscopy 571
- image depth sequences, secondary ion mass spectrometry 707 f
- image formation
  - field ion microscopy 781
  - low energy electron holography 966
  - point-projection microscopy 968
- image intensifier 897 ff
- image intensity distribution, scanning transmission electron microscopy 572 f
- image preprocessing
  - contrast 924
  - global methods 925 ff
    - in image frequency space 925
    - in image space 925
    - in parameter space 926
  - local methods 927

- mathematical morphology 936
- quadratic signal-to-noise ratio 924
- image processing 923 ff
- scanning electron microscopy with polarization analysis 744
- secondary ion mass spectrometry 707
- image reconstruction
  - low energy electron holography 969, 971
  - nuclear magnetic resonance microscopy 730
  - point-projection holography 971
  - point-projection microscopy 969
  - three-dimensional 942
- image recording
  - directly back illuminated charge coupled device 909
  - noise contribution 887
  - restrictions and disturbances 886
  - rules for the design of instrumentation 886
  - secondary ion mass spectrometry 696
  - slow-scan charge coupled device 909
  - television chains 908
- image restoration 931 ff
  - focus series approach 932
  - iterative methods 932
  - nonlinear procedures 932
  - partly linear procedures 932
  - simple inverse filtering procedure 931
- image segmentation
  - conditional region-growing approach 936
  - connectivity 935
  - functional minimization 936
  - gray-level homogeneity 935
  - mathematical morphology 936
  - on the basis of gray level gradients 934 f
  - region-growing method 936
  - region-splitting method 936
  - snakes approach 936
- image sequence processing 924
- image series 945
- image space 925
- imaging fields (table) 780
- imaging modes
  - scanning force microscopy 832
  - – lift mode 836
  - – noncontact dynamic modes 834
  - – tapping mode 836
  - – variable deflection mode 833
  - scanning transmission electron microscopy 566 ff
- imaging plates
  - detected quantum efficiency 907
  - X-ray optics 896, 913
- imaging secondary ion mass spectrometry 691 ff
  - applications 713 f
  - comparison with other imaging microanalytical and surface analytical techniques (table) 693
  - data acquisition modes 692, 704
  - experimental set-up 692
  - image analysis 706 ff
  - instrumentation 695 ff
  - sample requirements
- immersion lens, scanning electron microscopy 541
- impact ionization, ballistic electron emission microscopy 873
- imperfections in crystals, scanning transmission electron microscopy 578
- incoherent imaging 595 ff
  - direct structure inversion 598
  - dynamic diffraction conditions 597
  - elastically scattered electrons 598
  - experimental set-up 598
  - high resolution 599
  - image intensity 601
  - residual correlation 602
  - resolution 599
  - resolution criteria 595
  - STEM geometry 598
  - theory 598 ff
    - – elastically scattered electrons 600
    - – inelastically scattered electrons 604
    - with thermally scattered electrons 601
  - Z-sensitivity 597
- incoherent structure imaging and EELS 611
- inelastic scattering, scanning transmission electron microscopy 574
- inelastic tunneling spectroscopy 823 f
- infinite-magnification circle 579
- information depth
  - Auger electron spectroscopy 628
  - spin-polarized low-energy electron microscopy 753
- in-line electron holography 963 f
- inner-shell energy losses, scanning transmission electron microscopy 575
- InP, electron channeling pattern 551
- input lenses, cylindrical mirror analyzer 641
- interactive correlation partitioning 950
- interface
  - epitaxial 865, 871
  - nonepitaxial 867
  - semiconductor 865 f
- interfacial imperfections, ballistic electron emission microscopy 872
- interfacial nanostructures 870 f
- interferometric detection, scanning force microscopy 831
- intrinsic luminescence 558
- ion detection 696
- ion implantation, secondary ion mass spectrometry 695
- ion microprobe, comparison with ion microscopes (table) 702

- ion microscope, comparison with ion micro-probes (table) 702
- ion microscope mode, secondary ion mass spectrometry 692
- ionic repulsion, scanning force microscopy 838
- ionization potentials (table) 780
- iron, scanning electron microscopy with polarization analysis 745
- k*-space 723 f, 730
- knife-edge test, scanning transmission electron microscopy 584
- Kramers' equation 669
- LaB<sub>6</sub> emitters, scanning Auger microscopy 645
- large-area surface analysis, secondary ion mass spectrometry 691
- Larmor frequency 721
- laser ionization mass analysis 795
- lateral force microscopy 835
- lateral resolution
  - ion microscope 702
  - scanning transmission electron microscopy 675
  - scanning tunneling microscopy 813
  - secondary ion mass spectrometry 703, 703
- lateral resolution limit, secondary ion mass spectrometry 695
- lattice defects, field ion microscopy 784
- lens optics, image recording 894
- lift mode
  - magnetic force microscopy 846
  - scanning force microscopy 836
- light optical elements, image recording 893
- lipid monolayers, point-projection microscopy 982
- local contrast enhancement 930
- local tunneling barrier height
  - measurement at fixed surface locations 816
  - scanning tunneling microscopy 815
  - spatially resolved measurements 816 f
- lock-in detection techniques
  - deflection type analyzers 636
  - spin-polarized scanning tunneling microscopy 822
- Lorentz microscopy
  - reflection mode 736
  - comparison with SEMPA 736
- low energy diffuse scattering spin polarization analyzer 741 f
- low energy electron holography 963 f
- low-light-level detection 900, 905
- macromolecules, mass determination by STM 587 f
- magnetic circular dichroism, comparison with SEMPA 737
- magnetic domain structure
  - backscattered electron imaging 550
  - spin-polarized low-energy electron microscopy 755
- magnetic domains, spin-polarized low-energy electron microscopy 755
- magnetic field
  - gradients, nuclear magnetic resonance microscopy 723 f
  - imaging XPS 652
- magnetic force microscopy 845 ff
  - comparison with SEMPA 736
  - experimental set-up 845 f
- magnetic forces, scanning force microscopy 841
- magnetic imaging, comparison of different techniques (Table) 736
- magnetic layers, exchange coupling 746
- magnetic materials, photoelectron emission microscopy 769
- magnetic media, secondary electron imaging 546, 547
- magnetic microstructure, scanning electron microscopy with polarization analysis 735
- magnetic objective lens, X-ray photoelectron imaging 654
- magnetic phase transitions, spin-polarized low-energy electron microscopy 758
- magnetic sector mass spectrometer 696
- magnetic singularities, scanning electron microscopy with polarization analysis 747
- magnetic thin film tips 848
- magnetization gratings, nuclear magnetic resonance microscopy 724
- magneto-optic Kerr effect 737
- magneto-optic recording media 746
- magnification
  - determination of for point-projection microscopy 971
  - scanning force microscopy 828
- mass resolution 704, 791 ff
  - field ion microscopy 780 f
  - secondary ion mass spectrometry 705
  - three-dimensional atom probe 798
- mathematical morphology 936
- matrix contrast, secondary ion mass spectrometry 706
- maximum likelihood common factor analysis, Auger electron spectroscopy 630
- medial axis transform see skeleton
- metal/insulator/semiconductor structures 876
  - ballistic electron emission microscopy 862
  - degree of scattering 863
  - momentum distribution of electrons 862
- metals, photoemission threshold 765
- MgO
  - in-line holography 579

- nanodiffraction patterns 578
- scanning transmission electron microscopy 565
- microanalysis, scanning transmission electron microscopy 575
- microchannel plates 895, 899, 916
- microfabrication 830
- micromagnetic imaging, scanning electron microscopy with polarization analysis 737
- microphotography, image recording 905
- microtomography, 3D image reconstruction 944
- MIDAS system 586
- modelization approach 946
- modulation transfer function 904
- molecular motions, nuclear magnetic resonance microscopy 732
- molecular vibrational spectroscopy 824
- molybdenum–sulphide fullerene
  - high-resolution image 676
  - X-ray energy-dispersive spectra 676
- momentum distribution, ballistic electron emission microscopy 873
- monolayers
  - chemical kinetic reaction-diffusion fronts 769
  - photoelectron emission microscopy 769
- Moseley's law 666
- Mott analyzer 742
- multielement characterization, secondary ion mass spectrometry 693
- multielement mapping, position sensitive atom probe 799
- multilayer structure, secondary ion mass spectrometry 708
- multiphase materials, field ion microscopy 786
- multiple detector systems, analytical electron microscopy 662
- multiple-layer structures, ultrathin, ballistic electron emission microscopy 874
- multiple-pulse experiments, nuclear magnetic resonance microscopy 727, 729, 730
- multislice calculation, scanning transmission electron microscopy 574
- multispectral Auger microscope, reduction of image artefacts 655
- multivariate data analysis
  - multivariate statistical analysis 946
  - secondary ion mass spectrometry 710
- nanodiffraction, scanning transmission electron microscopy 577 ff
- nanotip field emission tips 973
- near-field microscopy, image restoration 933
- Newvicon tubes 900
- nickel aluminum alloys, atom probe field ion microscopy 796
- Nipkow disk confocal scanning microscopy 906
- NiSi<sub>2</sub>/Si(111) diodes, ballistic electron emission microscopy 865
- noise 904
  - charge coupled device 904
  - contribution in image recording 887
  - reduction during image recording 887 ff
- noncontact dynamic mode, scanning force microscopy 834
- nonepitaxial interfaces, ballistic electron emission microscopy 867
- nonlinear filters 926
- nuclear magnetic resonance microscopy 719 ff
  - image acquisition 719
  - instrumentation 727
  - solid-state imaging 733
  - theory 720
  - two-dimensional imaging 730
- object features 937
- off-axis STEM holography 582
- optical PoSAP 798
- organic coatings, secondary ion mass spectrometry 713
- oxide interlayers, ballistic electron emission microscopy 877
- p–n junction
- P-number 893
- Packwood–Brown model 685
- parallax 943
- parallel detection in XPS and AES 640
- parameter space 925
- particles, analytical electron microscopy 687
- passivation layers, secondary ion mass spectrometry 710
- pattern recognition 937, 948
- Pauli repulsion, scanning force microscopy 838
- Pd on C support, bright-field STEM image 590
- Pd/Si(100), ballistic electron emission microscopy 867
- Pd/Si(111), ballistic electron emission microscopy 867
- permalloy, magnetic force microscopy 850 f
- phonon spectroscopy, scanning tunneling microscopy 824
- phosphors, image recording 893
- photocathodes 891
- photoeffect 890 f
- photoelectron emission microscopy 763 ff
  - applications
  - catalyst surfaces 769
  - Co on Mo{011} 768
  - epitaxial growth 768
  - ferromagnetic domains 769
  - magnetic recording materials 769 f
  - monolayers, chemical kinetic reaction-diffusion fronts 769

- reactive monolayer films 769
- instrumentation 766, 767
- resolution 767
- sample preparation 771
- photoelectron spectroscopy, virtual 961
- photographic film 913
- photographic recording 890, 906 f
- resolution 907
- photoionization 764
- photomultiplier 890
- platinum on zeolite, annular dark-field imaging 568
- Plumbicon, image recording 900
- point-projection geometry, convergent-beam microdiffraction 964
- point-projection microscopy 963 ff
  - comparison with other techniques 981
  - holographic reconstruction algorithms 970 ff
  - image formation 968
  - instrumentation 978
  - reflection geometry 980, 983
  - resolution limit 978
  - tip aberrations 973
- polarization analyzers 740 f
- polished surfaces, electron probe X-ray microanalysis 669
- polymethylmethacrylate, resist material 914
- polystyrene/polybutadiene blend, nuclear magnetic resonance microscopy 733
- position-sensitive atom probe 797
- power spectrum method, fractal analysis 939
- primary ion sources 695
- probe channeling, Z-contrast scanning transmission electron microscopy 606 ff
- probe configurations, nuclear magnetic resonance microscopy 727
- profile analysis, energy-dispersive spectrometry 681
- Pt/SiO<sub>2</sub>/Si(100), ballistic electron emission microscopy 877
- ptychology 580
- pulse experiments, nuclear magnetic resonance microscopy 728
- pulse height distribution 889
  - contributions to 890
  - image conversion 889
- pulsed laser atom probe 795
- purple membrane
  - low energy electron holography 983
  - point-projection microscopy 982
- quadrupole mass spectrometer 696
- quantitative analysis, atom probe analysis 792
- quantitative lattice imaging 948
- quantitative X-ray scanning microanalysis 682
  - applications
  - particles 687 f
  - thin films 687 f
- atomic number correction 683
- fluorescence correction 684
- X-ray absorption correction 684
- standardless analysis
- $\phi(\rho z)$  methods 685
- QUANTITEM 948, 949
- quantum efficiency,
  - effect of exposure levels on 887
- image conversion 889 ff
- image recording 886
- quantum level diagram 888
- quantum size effect, ultrathin magnetic layers 752
- radiation damage, point-projection microscopy 982
- radiation-sensitive specimens, image recording 887
- random function models 940
- reactive monolayer films, photoelectron emission microscopy 769
- reconstruction algorithms, holography 970 ff
- reduced aberration 639
- reflection geometry, point-projection microscopy 980, 983
- reflection high-energy electron diffraction patterns
- region-growing method 936
- region-splitting method 936
- resist materials 914 f
- resistive anode encoder, secondary ion mass spectrometry 697, 704
- resolution
  - field emission microscopy 776
  - image recording 904
  - improvement by holography 583
  - nuclear magnetic resonance microscopy 725, 726
  - photoelectron emission microscopy 767
  - photographic recording 907
  - point-projection microscopy 975 f
  - scanning Auger microscopy 644
  - scanning electron microscopy 542 ff
  - scanning electron microscopy with polarization analysis 740
  - scanning transmission electron microscopy 567, 573, 583
- resolution limit
  - atom probe field ion microscope 791
  - point-projection microscopy 976, 978
- retarding field analyzers 632
  - concentric hemispherical analyzer 641
  - cylindrical mirror analyzer 641
- rhodium catalyst on  $\gamma$ -Al<sub>2</sub>O<sub>3</sub>, Z-contrast images 618

- sample damage, scanning Auger microscopy 646
- sample preparation
  - Auger electron spectroscopy 625 f
  - field ion microscopy 787 f
  - scanning force microscopy 828 f
- Saticon tubes, image recording 900
- scanning Auger microscopy 576, 621 ff, 630 ff, 654
  - analyzers 633 ff
  - background slope effects 656
  - beam current fluctuation effects 657
  - detection defined imaging 632, 644
  - edge effects 657
  - emitters 645
  - image artefacts 654
  - imaging technique 630
  - parallel detection 639
  - source defined imaging 632, 644
  - substrate backscattering effects 656
  - topographic effects 656
- scanning beam methods 537 ff
- scanning electron beam analysis 669
  - instrumentation 669 ff
- scanning electron microscopy with polarization analysis
  - applications
    - Fe/Cr/Fe(100) 746
    - Fe-SiO<sub>2</sub> films 747
    - iron single crystals 745
    - magnetic layers, exchange coupling 746
    - magnetic singularities 747
    - magnetooptic recording media 746
  - comparison with other magnetic imaging techniques 736
  - instrumentation 739
  - operating principle 737
  - resolution 740
  - system performance 743
  - theory 737 ff
- scanning electron probe methods, basic characteristics (table) 670 f
- scanning force microscopy
  - experimental set-up 827 ff
  - imaging modes 832 ff
  - limitations 828, 837
  - magnification 828
  - nanodiffraction 838 f
  - theory 838
- scanning microanalysis 661 ff
  - instrumentation 669 f
  - methods 669
  - quantitative 682 ff
- scanning near-field probe microscopy 805 ff
  - operating principle 805
- scanning reflection electron microscopy 539 ff, 546
  - capabilities 539
  - comparison with scanning transmission microscopy 546 f
  - instrumentation 540, 541
  - operating principle 539
- scanning transmission electron microscopy
  - applications
    - carbon nanotubes 591
    - diamond, planar defects in 590
    - heavy atoms on light atom supports 587
    - heavy atoms on zeolites 589
    - mass determinations of macromolecules 588
    - small particles 588
    - surface superstructures 591
  - comparison with conventional transmission electron microscopy 565
  - comparison with SEMPA 736
  - holography 581 ff
  - imaging modes 566 ff
  - instrumentation 563 ff, 584
  - resolution 567, 573
  - Z-contrast 595 ff
- scanning tunneling microscopy
  - adsorbates 811
  - comparison with BEEM images 858, 859
  - experimental set-up 807, 808
  - spin polarized 820
  - phonon spectroscopy 824
  - tips 810
  - topographic images 858, 859
  - tunneling spectroscopy 817 ff
- Schottky barrier
  - ballistic electron emission microscopy 874
  - depletion depth 557
  - electron beam induced channeling image 557
  - height, ballistic electron emission microscopy 856, 860
  - interfaces, ballistic electron emission microscopy analysis 864
- Schottky field emitters, scanning Auger microscopy 645
- Schottky surface barrier, electron beam induced conductivity 555
- scintillator 892 ff
  - inhomogeneity 894
  - photoelectron output 892
- scintillator-photomultiplier detector 956
- searchlight effect 870
- secondary electron detector 644
- secondary electron imaging 544 ff
  - Everhart-Thornley detector 545
  - voltage contrast 547
- secondary electrons
  - generation 544
  - yield 544
- secondary emission, image recording 895
- secondary ion formation 694

- secondary ion mass spectrometry 691 ff
  - applications 712 f
  - dynamic mode 692
  - experimental set-up 692
  - image recording 696
  - instrumentation 695
  - ion detection 696
  - ion microprobe mode 692
  - ion microscope mode 692
  - sample requirements 712
  - sensitivity 693
  - static mode 692
- segmentation see image segmentation
- self-supporting molecule assemblies, point-projection microscopy 984
- semiconductor interfaces
  - ballistic electron emission spectroscopy 865 ff
  - Z-contrast scanning transmission electron microscopy 611
- semiconductors
  - as-grown, electron beam induced channeling image 557
  - atom probe field ion microscopy 795
  - photoemission threshold 765
  - scanning electron microscopy 553
  - Z-contrast scanning transmission electron microscopy 610
- sensitivity
  - atom probe field ion microscope 790
  - secondary ion mass spectrometry 693
- sequential energy analysis 631
  - scanning Auger microscopy 631
- serial imaging mode, scanning transmission microscopy 566
- serial sections, 3D image reconstruction 943
- shape of molecules, point-projection microscopy 968
- Sherman function 741
- Si see silicon
- Si(111), scanning tunneling microscopy 818
- Si(111)/NiSi<sub>2</sub>, ballistic electron emission microscopy 873
- signal collection, cathodoluminescence 559
- signal processing, scanning electron microscopy with polarization analysis 742 f
- signal response, image recording 887
- signal-to-noise ratio
  - algorithm for improving 929
  - electronic image recording 909
  - primary image formation 885
  - quadratic 924
  - scanning electron microscopy with polarization analysis 744
  - scanning force microscopy 834
  - spin-polarized scanning tunneling microscopy 822
- silicon
  - nanocrystals 617
  - photodiode 891
  - scanning tunneling microscopy 818
  - tilt grain boundaries in 611
  - Z-contrast images 617
- silicon intensifier target camera tube 900
- single electron counting 900, 908 ff
- single-phase alloys, field ion microscopy 785
- single-quantum response 888
- SiO<sub>2</sub> gate oxide layers, scanning Auger microscopy 645
- site occupancy, atom probe field ion microscopy 794
- skeleton 942
- slice selection, nuclear magnetic resonance microscopy 731
- slow-scan charge coupled device
  - converters with a scintillator 909 ff
  - energy ranges 909
- slow-scan charge-coupled device chains 917
- small particles, scanning transmission electron microscopy 588
- snakes approach, image segmentation 936
- snorkel lens, scanning electron microscopy 541
- soft X-ray spectra, scanning microanalysis 668
- solid state imaging, nuclear magnetic resonance microscopy 733
- solid state sensors 901
- source defined imaging
  - Auger electron spectroscopy 644
  - scanning Auger microscopy 632
  - X-ray photoelectron spectroscopy 647
- spatial displacement, scanning Auger microscopy 631
- spatial resolution
  - backscattered electron imaging 550
  - ballistic electron emission microscopy 863
  - cathodoluminescence 560
  - electron beam induced channeling image 556
  - field ion microscopy 784
  - image recording 890
  - scanning electron microscopy 542 ff
  - scanning electron microscopy with polarization analysis 738
  - scanning tunneling microscopy 812, 814
  - X-ray microanalysis 676
- specimen geometry, field ion microscopy 787
- spin dependent scattering, scanning electron microscopy with polarization analysis 738
- spin echo 729
- spin-lattice relaxation 721, 731
- spin-polarized low-energy electron microscopy 737, 751 ff
  - applications
    - in situ studies 756
    - magnetic domain structure 755



- magnetic phase transitions 758
- superlattices 755
- thin films 755
- comparison with SEMPA 737
- information depth 753
- instrumentation 753, 754
- resolution 755
- theory 751
- spin-polarized scanning tunneling microscopy 821 f
- spin-spin relaxation 721
- spin warp imaging 731
- sputter yield, secondary ion mass spectrometry 694
- SrTiO<sub>3</sub>, [001] tilt grain boundaries 613 f
- standardless analysis, quantitative X-ray scanning microanalysis 686
- static secondary ion mass spectrometry 692
- stereology 939 f
- stereoscopy 943, 944
- stimulated echo 730
- stray fields, point-projection microscopy 978
- substrate backscattering effects, scanning Auger microscopy 656
- superlattices, spin-polarized low-energy electron microscopy 755
- supported metal catalysts, scanning transmission electron microscopy 578
- surface morphology, effect on ballistic electron emission microscopy current image 869, 870
- surface structuring, scanning force microscopy 838
- surface superstructures, scanning transmission electron microscopy 591
- synchrotron radiation, photoelectron imaging 766
- tandem lenses, image recording 895
- tapping mode, scanning force microscopy 836
- television camera tubes, image recording 899
- television chains, image recording 908
- texture recognition 938
- thermal vibrations, incoherent imaging 601
- thin-annular bright-field imaging 573
- thin-film criterion 687
- thin films
  - analytical electron microscopy 687
  - magnetic fields 582
  - spin-polarized low-energy electron microscopy 755
- three-dimensional image reconstruction, microtomography 943 ff
- three-dimensional atom probe 795 ff
  - experimental set-up 797
  - mass resolution 798
- three-dimensional imaging, nuclear magnetic resonance microscopy 732
- time-of-flight analysis
  - field ion microscopy 792
  - mass spectrometry 696
- tip aberrations, point-projection microscopy 973
- tip motion, point-projection microscopy 977
- tip requirement, magnetic force microscopy 847, 848
- tip-sample interaction
  - magnetic force microscopy 847
  - scanning tunneling microscopy 819
- tip-surface separation, scanning tunneling microscopy 813
- tip wave functions 810 f
- topographic atom probe 797
- topographic effects, scanning Auger microscopy 656
- topographic imaging, constant-current mode 807
- transfer lenses 641 f
  - cylindrical mirror analyzer 641
  - retarding effect 642
  - system for X-ray photoelectron imaging 651
- transmission geometry, convergent-beam microdiffraction 964
- transmission photocathodes
  - optimum thickness 915
  - resolution 916
- transport optics, scanning electron microscopy with polarization analysis 740
- TRIFT mass spectrometer 702
- tungsten, field ion microscopy 783
- tunneling
  - density of states 818, 819
  - gap resistance 814
- tunneling current
  - bias dependence 817, 818
  - scanning tunneling microscopy 809 f
  - spin dependence 822
- tunneling spectroscopy 817 ff
  - elastic 823
  - inelastic 823
- tunneling stabilized magnetic force microscopy 852
- twin image formation, point-projection microscopy 973, 982
- two-dimensional imaging, slice selected 732
- ultrathin magnetic layers, spin-polarized low-energy electron microscopy 752
- universal escape depth, Auger electron spectroscopy 628
- University of Chicago UC-HRL scanning ion microprobe 702
- uranium, annular dark-field imaging 564

- V<sub>2</sub>O<sub>5</sub>
  - Auger spectrum 633
  - X-ray photoelectron spectroscopy 627
- Van der Waals forces, scanning force microscopy 839
- variable deflection mode, scanning force microscopy 833
- vector magnetization maps, scanning electron microscopy with polarization analysis 738, 745
- video recording 704, 905
- Vidicon tubes 900, 917
- virtual photoelectron spectroscopy 960
- viscoelasticity, scanning force microscopy 835
- voltage contrast, secondary electron imaging 547
- Voronoi diagram 942
- Warren approximation 603
- water, nuclear magnetic resonance imaging 726
- wavelength-dispersive spectrometry
  - comparison with energy-dispersive spectroscopy 679
- YBa<sub>2</sub>Cu<sub>3</sub>O<sub>7-x</sub> 666
- X-ray absorption correction 683 f
- X-ray analysis 673
- X-ray detector, scanning transmission electron microscopy 586
- X-ray distribution function, analytical electron microscopy 663, 686
- X-ray emission spectra 664 f
- X-ray emission spectrometry 679
- X-ray image converter 918
- X-ray intensifier 917
- X-ray mapping
  - ceramics 682
  - energy-dispersive spectrometry 681
  - wavelength-dispersive spectrometry 681
- X-ray microanalysis, scanning transmission electron microscopy 675 ff
- X-ray optics
  - imaging plates 896, 913
  - image recording 912, 913
  - photographic film 913
  - resolution 913
- X-ray photoelectron spectroscopy 621 ff, 625, 626, 647
  - analyzers 632 ff
  - beam effects 627
  - detector defined imaging 648
  - kinetic energy 626
  - quantitative analysis 627
  - source defined imaging 647
- X-ray scanning microanalysis, quantitative 682
- X-ray spectra, characteristic, scanning microanalysis 666
- X-rays
  - generation 646
  - monochromatizing 646
- YBa<sub>2</sub>Cu<sub>3</sub>O<sub>7-x</sub>
  - energy-dispersive spectroscopy 666
  - wavelength-dispersive spectroscopy 666
  - X-ray emission spectra 664, 665
  - Z-contrast imaging 617
- YBa<sub>2</sub>Cu<sub>3</sub>O<sub>7-x</sub> grain boundaries, Z-contrast scanning transmission electron microscopy 615
- Young's fringes 968
- Z-contrast scanning transmission electron microscopy 595 ff
  - applications
    - catalysts 618
    - CdTe(001)–GaAs(001) interface 611, 612
    - ceramics 613
    - GaAs 610
    - nanocrystalline materials 616
    - semiconductors 610 f
    - tilt boundaries in silicon 611
    - YBa<sub>2</sub>Cu<sub>3</sub>O<sub>7-x</sub> grain boundaries 615
  - incoherent imaging 598 ff
- ZAF method, quantitative X-ray scanning microanalysis 683
- Zernike–van Cittert theorem 571