

NH PL



DIFFRACTION PHYSICS

Third Revised Edition

JOHN M. COWLEY

NORTH-HOLLAND
PERSONAL LIBRARY

**NORTH-HOLLAND
PERSONAL LIBRARY**

DIFFRACTION PHYSICS

THIRD REVISED EDITION

This Page Intentionally Left Blank

DIFFRACTION PHYSICS

THIRD REVISED EDITION

JOHN M. COWLEY

*Regents' Professor emeritus
Arizona State University*



1995

ELSEVIER

Amsterdam – Lausanne – New York – Oxford – Shannon – Tokyo

ELSEVIER SCIENCE B.V.
Sara Burgerhartstraat 25
P.O. Box 211, 1000 AE Amsterdam, The Netherlands

First edition	1975
Second revised edition	
First impression (hardbound)	1981
Second impression (paperback)	1984
Third impression (paperback)	1986
Fourth impression (paperback)	1990
Third revised edition	
First impression (paperback)	1995

Library of Congress Cataloging-in-Publication Data

Cowley, J. M. (John Maxwell), 1923-
Diffraction physics / John M. Cowley. -- 3rd rev. ed.
p. cm. -- (North-Holland personal library)
Includes bibliographical references and index.
ISBN 0-444-82218-6 (alk. paper)
1. Diffraction. I. Title.

QC415.C75 1995
535'.4--dc20

95-38709
CIP

ISBN: 0-444-82218-6

© 1995 ELSEVIER SCIENCE B.V. All rights reserved.

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of the publisher, Elsevier Science B.V., Copyright & Permissions Department, P.O. Box 521, 1000 AM Amsterdam, The Netherlands.

Special regulations for readers in the U.S.A. - This publication has been registered with the Copyright Clearance Center Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923. Information can be obtained from the CCC about conditions under which photocopies of parts of this publication may be made in the U.S.A. All other copyright questions, including photocopying outside of the U.S.A., should be referred to the publisher.

No responsibility is assumed by the publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein.

This book is printed on acid-free paper.

Printed in The Netherlands

PREFACE TO THE FIRST EDITION

This book has evolved, over the last ten years, from lecture notes for classes in physical optics, diffraction physics and electron microscopy given to advanced undergraduate and graduate students. It reflects my own particular interests in electron diffraction and diffraction from disordered or imperfect crystals and employs an approach which is particularly suited to the treatment of these topics. This approach, using the Fourier transform from the beginning instead of as an extension of a Fourier series treatment, is not only more satisfying from a conceptual and theoretical basis but it provides the possibility for a unified treatment of all the different branches of diffraction physics, employing electrons, X-rays or neutrons.

Because this approach has been adopted, the generation of the familiar ideas of diffraction of radiation by solids is slow and detailed. Bragg's Law does not appear until Chapter 6 and then only as a corollary to the Ewald sphere construction. This may create difficulties for readers or students unfamiliar with the more conventional approaches. Hence this book is probably more appropriate for those who have had one elementary course in diffraction or who are already engaged in some area of diffraction work.

Much of the content relating to electron diffraction has been generated in collaboration with A.F. Moodie, of the Division of Chemical Physics, C.S.I.R.O., Australia, who has for many years been my mentor in matters of diffraction theory. My heartfelt gratitude goes to him and to all my colleagues in the C.S.I.R.O., the University of Melbourne and Arizona State University whose valuable and friendly collaboration has made this work possible and who have allowed me to use their results and their pictures to illustrate the text. I wish to thank Drs. Kato, Borrmann and Beauvillain for permission to reproduce Figs. 9.10, 14.4, 14.5 respectively.

PREFACE TO THE SECOND EDITION

In preparing the second edition of this book, I have been greatly assisted by my students and others who have pointed out a number of errors in the first edition. Some valuable comments came from Professor Z.G. Pinsker and colleagues who prepared the Russian language edition. I am very grateful to these people and also to Dr. Peter Goodman for discussions on some of the additional sections which have been included.

Most of the changes which have been made to the first edition come within the final few chapters which deal with applications of the diffraction methods since it is in these areas that progress has been most rapid. In fact, the expansion of the field has been so great that it is impossible in a book of this size to provide any coherent account of the many interesting recent research developments. However, it is hoped that the mention of some of these topics, together with the appropriate references, will lead the reader towards more complete information on topics of a more specialized interest.

PREFACE TO THE THIRD EDITION

In this third edition an attempt has been made to make this book more consistent with the present status of diffraction activity which has undergone important developments in the last fifteen years. It is not feasible to attempt an adequate description of current techniques or the most recent theoretical developments, but it is hoped that sufficient reference has been made to these things to allow them to be placed in the context of the elementary theoretical ideas which form the basis for this book. In preparing this Edition, I have benefitted greatly from interactions with, and assistance from, the faculty, research associates, staff and students associated with the Center for High Resolution Electron Microscopy at Arizona State University. In particular, I would like to express my gratitude to Kaye Frix for her painstaking work on preparing the camera-ready copy.

CONTENTS

<i>Preface to the first edition</i>	V
<i>Preface to the second edition</i>	VI
<i>Preface to the third edition</i>	VI

Section I - PHYSICAL OPTICS

<i>Chapter 1. Fresnel and Fraunhofer diffraction</i>	3
1.1. Introduction	3
1.2. Wave equations and waves	5
1.2.1. Wave functions	5
1.2.2. Electromagnetic waves	6
1.2.3. Particle waves	7
1.3. Superposition and coherence	8
1.3.1. Superposition	8
1.3.2. Independent point sources	9
1.4. Huygen's principle	11
1.4.1. Kirchhoff's formulation	11
1.4.2. Application of the Kirchhoff formula	12
1.5. Scattering theory approach	13
1.5.1. Integral form of wave equation	13
1.5.2. Born series	14
1.6. Reciprocity	16
1.7. Fresnel diffraction	16
1.7.1. Small angle approximation	16
1.7.2. Fresnel integrals	17
1.7.3. Periodic objects - "Fourier images"	19
1.8. Fraunhofer diffraction	21
Problems	24
 <i>Chapter 2. Fourier transform and convolutions</i>	 25
2.1. Preliminaries	25
2.1.1. Delta-functions and discontinuities	25
2.1.2. Convolutions	26
2.1.3. Examples of convolutions	27
2.2. Fourier transforms: general	30
2.2.1. Definitions	30

2.2.2. Properties of Fourier transforms	31
2.2.3. Multiplication and convolution	33
2.2.4. Space and time	34
2.3. Fourier transforms and diffraction: examples	35
2.3.1. Point source or point aperture	35
2.3.2. A plane wave: the inverse of 2.3.1.	35
2.3.3. Translation of an object	36
2.3.4. Slit function	36
2.3.5. Slit function-alternative	37
2.3.6. Straight edge	37
2.3.7. Rectangular aperture	38
2.3.8. Circular aperture	39
2.3.9. Two very narrow slits	39
2.3.10. Two slits of appreciable width	40
2.3.11. Finite wave train	40
2.3.12. Periodic array of narrow slits	42
2.3.13. Arbitrary periodic function	43
2.3.14. Diffraction grating: thin slits	43
2.3.15. Diffraction grating: general	44
2.3.16. Gaussian function	45
2.3.17. Row of circular holes	46
2.3.18. Complementary objects-Babinet's principle	47
2.3.19. Total intensities: Parseval's theorem	48
Problems	48

Chapter 3. Imaging and diffraction 51

3.1. Wave theory of imaging	51
3.1.1. Coherent wave optics	51
3.1.2. Incoherent wave imaging	54
3.2. Abbe theory	55
3.3. Small angle approximation	56
3.4. Phase contrast	59
3.4.1. Phase and amplitude objects	59
3.4.2. Out-of-focus contrast	60
3.4.3. Aperture limitation	62
3.4.4. Zernike phase contrast	62
3.5. Holography	63
3.6. Multi-component systems	67
3.7. Partial coherence	69
Problems	72

Section II - KINEMATICAL DIFFRACTION

Chapter 4. Radiations and their scattering by matter 77

4.1. X-rays	77
4.1.1. X-ray sources	77
4.1.2. Scattering by electrons	78

4.1.3. Scattering by atoms	80
4.1.4. Dispersion corrections	81
4.2. Electrons	81
4.2.1. Sources of electrons	81
4.2.2. Atom scattering amplitudes	82
4.2.3. Phase object approximation	84
4.2.4. Failure of first Born approximation	85
4.2.5. "Absorption" effects	86
4.3. Neutrons	88
4.3.1. Atomic scattering factors	88
4.3.2. Nuclear spin scattering	89
4.3.3. Isotopic disorder	90
4.3.4. Thermal and magnetic scattering	91
Problems	92

Chapter 5. Scattering from assemblies of atoms 93

5.1. The kinematical approximation	93
5.2. Real and reciprocal space	95
5.2.1. Reciprocal space distribution	95
5.2.2. The reciprocal lattice	96
5.2.3. Friedel's law and the phase problem	97
5.3. The generalized Patterson function	98
5.4. Examples of correlation functions	101
5.4.1. Finite volume limitations	101
5.4.2. Finite crystals	102
5.5. Correlation in space and time	104
5.5.1. Four-dimensional Patterson	104
5.5.2. Special cases	105
5.5.3. Ideal monatomic gas or liquid	106
5.5.4. Real monatomic gases and liquids	109
5.5.5. The hydrogen atom	112
5.6. Diffraction geometry and intensities	113
5.7. Practical considerations	115
5.7.1. Finite sources and detectors	115
5.7.2. Wavelength spread	117
5.7.3. Integrated intensities	118
5.8. Sections and projections	119
Problems	121

Chapter 6. Diffraction from crystals 123

6.1. Ideal crystals	123
6.2. Diffraction geometry	126
6.2.1. Laue and Bragg diffraction conditions	126
6.2.2. Shape transforms	127
6.2.3. Special cases for electron diffraction	128

6.3. Crystal structure analysis	131
6.3.1. The phase problem	131
6.3.2. Supplementary information	133
6.4. Structure analysis methods	134
6.4.1. Trial and error	134
6.4.2. Patterson function	135
6.4.3. Heavy-atom & isomorphous replacement methods	135
6.4.4. Direct methods	136
6.5. Neutron diffraction structure analysis	138
6.5.1. Nuclear scattering	138
6.5.2. Magnetic scattering	139
6.6. Electron diffraction structure analysis	141
Problems	142

Chapter 7. Diffraction from imperfect crystals 145

7.1. Formulation of the diffraction problem	145
7.1.1. Types of defects	145
7.1.2. General diffraction formulation	146
7.2. Patterson function approach	147
7.2.1. Patterson with average periodic structure	147
7.2.2. Patterson with no average structure	149
7.3. Deviations from an average lattice	150
7.3.1. Random vacancies: no relaxation	150
7.3.2. Clustered vacancies	152
7.3.3. Lattice relaxation	154
7.3.4. Thermal vibrations - Einstein model	156
7.4. Imperfect crystals with no average lattice	157
7.4.1. Uneven separation of lattice planes	157
7.4.2. Disordered orientations	160
Problems	163

Section III - DYNAMICAL SCATTERING

Chapter 8. Diffraction by perfect crystals 167

8.1. Multiple coherent scattering	167
8.2. Theoretical approaches	168
8.3. Bethe theory	170
8.3.1. The dispersion equations	170
8.3.2. Solutions of the equations	172
8.3.3. Boundary conditions	173
8.4. Two-beam approximation	175
8.4.1. Bloch waves and dispersion surfaces	175
8.4.2. Conduction electrons-energy representation	177
8.4.3. X-ray diffraction; polarization	178

8.5. The Laue (transmission) case	178
8.5.1. Electron diffraction for a thin crystal	178
8.5.2. Small angle approximation	181
8.5.3. Laue case with absorption	182
8.6. Bethe potentials	182
8.7. The Bragg case	184

Chapter 9. Dynamical diffraction effects 189

9.1. Thickness fringes, rocking curves-electron diffraction	189
9.1.1. Intensity formulas	189
9.1.2. Real space picture	189
9.1.3. Rocking curves	191
9.1.4. Extinction contours	191
9.1.5. Convergent beam diffraction	193
9.1.6. Diffraction and imaging of crystal wedges	195
9.1.7. Absorption effects for wedges	197
9.2. Dynamical effects of X-ray and neutron diffraction	199
9.2.1. Techniques for X-ray diffraction	199
9.2.2. Energy flow	202
9.2.3. Dispersion surface picture	203
9.2.4. Neutron diffraction	204
9.3. Borrmann effect	205
Problems	207

Chapter 10. Extension to many beams 209

10.1. Dynamical n -beam diffraction	209
10.2. Extension of Bethe theory - transmission	211
10.2.1. Matrix formulation	211
10.2.2. Small angle approximation	213
10.2.3. Bloch waves and boundary conditions	214
10.2.4. The scattering matrix	216
10.2.5. Derivation of the two-beam approximation	218
10.3. The Darwin-type approach	220
10.4. Special cases - beam reduction	222
10.5. Computing methods	224
10.6. Column approximation	227
Problems	229

Chapter 11. Multi-slice approaches 231

11.1. Propagation of electrons in crystals	231
11.1.1. Transmission through thin slices	231
11.1.2. Three-dimensional objects	233
11.1.3. Diffraction by a crystal	234
11.1.4. General expression; excitation errors	236

11.2. Multiple-scattering series	237
11.2.1. Zero-order scattering	237
11.2.2. Single scattering-kinematical approximation	238
11.2.3. Multiple scattering	238
11.3. General double-summation solution	240
11.3.1. General series solution	240
11.3.2. Phase grating approximation	241
11.4. Computing methods	243
11.4.1. "Slice method" calculations	243
11.4.2. Steps in a computation	245
11.4.3. Possible errors	246
11.4.4. Consistency tests	247
11.4.5. Computing times	248
11.5. Intensities from non-periodic objects	248
11.6. Real-space formulations	250
11.6.1. High-energy approximation	250
11.6.2. Useful approximations	252
11.6.3. A real-space basis for computing	253
Problem	254

Section IV - APPLICATIONS TO SELECTED TOPICS

<i>Chapter 12. Diffuse scattering and absorption effects</i>	257
12.1. Thermal diffuse scattering	257
12.1.1. Phonons and vibrational waves	257
12.1.2. Scattering for a longitudinal wave	258
12.1.3. Diffuse scattering component	259
12.1.4. Dispersion curves	261
12.1.5. Three-dimensional generalizations	262
12.2. Static atom displacements	262
12.2.1. Relaxation around point defects	262
12.2.2. Diffraction intensities for displaced atoms	263
12.2.3. The Bragg peaks	265
12.2.4. The diffuse scattering	266
12.3. Electron excitations	269
12.3.1. Inelastic X-ray scattering	269
12.3.2. Electron excitation by electrons-plasmons	270
12.3.3. Single-electron excitations	272
12.4. Dynamical effects in diffuse scattering	274
12.4.1. Scattering and re-scattering	274
12.4.2. Coherent and incoherent scattering	276
12.4.3. Analysis of diffuse scattering	278
12.5. Absorption effects	279
12.5.1. The nature of absorption parameters	279
12.5.2. Absorption of X-rays and neutrons	280
12.5.3. "Absorption" for electrons	281
12.5.4. Absorption due to thermal vibrations	282
12.5.5. Absorption from electron excitations	284
12.5.6. Values of absorption coefficients	285

<i>Chapter 13. Electron microscope imaging</i>	287
13.1. Electron microscopes	287
13.1.1. Conventional transmission e. m.	287
13.1.2. Scanning transmission electron microscopes	289
13.2. Image formation	292
13.3. Contrast for thin specimens	294
13.3.1. Phase-object approximation	294
13.3.2. Weak-phase object approximation	296
13.3.3. Failure of weak-phase object approximation	299
13.3.4. Dark-field images	300
13.4. The imaging of crystals	301
13.4.1. Imaging of thin crystals; structure images	301
13.4.2. Calculation of images of crystals: envelope	305
13.4.3. Imaging of crystals - inelastic scattering	307
13.4.4. Lattice fringe imaging	309
13.4.5. Crystal imaging without lattice resolution	312
13.5. Image contrast in STEM	313
13.5.1. STEM imaging of thin crystals	313
13.5.2. STEM imaging of thicker crystals	317
13.6. Electron holography	318
13.7. Combining high-resolution imaging with diffraction	326
Problems	328
 <i>Chapter 14. K-line patterns and channelling</i>	 329
14.1. Kossel lines	329
14.1.1. Geometry of Kossel lines	329
14.1.2. Dynamical theory of Kossel intensities	330
14.1.3. Kossel lines with limited resolution	332
14.2. Kikuchi lines	335
14.3. External sources of divergent radiation	339
14.4. Information from K-line patterns	341
14.5. Channelling	343
14.6. Secondary radiations	346
 <i>Chapter 15. Application of dynamical effects in single crystals</i>	 349
15.1. Dependence of dynamical effects on crystal parameters	349
15.2. X-ray interferometry	349
15.3. n -beam and 2-beam dynamical diffraction	351
15.4. Accurate determinations of structure amplitudes	354
15.4.1. Measurements of thickness fringes	355
15.4.2. Structure amplitudes from rocking curves	357
15.4.3. Convergent beam electron diffraction method	358
15.4.4. The use of critical voltages	359
15.4.5. Intersecting K-lines	362
15.5. The determination of crystal symmetries	363
15.6. Coherent convergent-beam electron diffraction	367

<i>Chapter 16. Mosaic crystals and polycrystalline materials</i>	369
16.1. General	369
16.2. Mosaic crystals	370
16.2.1. The mosaic crystal model	370
16.2.2. Kinematical integrated intensities	370
16.2.3. Extinction effects	372
16.2.4. Dynamical electron diffraction intensities	374
16.3. Polycrystalline material	374
16.3.1. Idealized models	374
16.3.2. Kinematical diffraction intensities	377
16.3.3. Line profile analysis	377
16.3.4. Rietveld refinements	379
16.3.5. Dynamical diffraction intensities	380
16.3.6. n -beam diffraction effects	383
 <i>Chapter 17. Ordering of atoms in crystals</i>	 385
17.1. The nature and description of disordered states	385
17.2. Order parameters	387
17.2.1. Short-range order	387
17.2.2. Long-range order	389
17.3. Patterson function	389
17.4. Size effects	390
17.5. Kinematical diffraction	392
17.5.1. Diffraction with ordering only	392
17.5.2. Diffraction with ordering and size effects	395
17.6. Relationship with ordering energies	400
17.7. Dynamical scattering from disordered crystals	401
17.7.1. Dynamical effects in diffuse scattering	401
17.7.2. Calculations of diffuse scattering	402
17.7.3. Strong scattering, multi-atom correlations	403
17.7.4. High resolution imaging disordered crystals	404
17.8. Out-of-phase domains	405
17.8.1. Ordered out-of-phase superlattices	405
17.8.2. Out-of-phase domains in disordered alloys	407
17.8.3. Modulated structures	408
Problems	410
 <i>Chapter 18. Extended defects</i>	 411
18.1. Introduction	411
18.2. Stacking faults—statistical, kinematical theory	412
18.2.1. Patterson method for a simple case	412
18.2.2. A general treatment	414
18.2.3. Faults in close-packed structures	420
18.3. Dynamical diffraction by stacking faults	422
18.4. Dislocations	424
18.4.1. Diffraction effects	424

CONTENTS

xv

18.4.2. The imaging of dislocations	426
18.4.3. Averaging over angles of incidence	427
18.4.4. n -beam diffraction effects	427
<i>Chapter 19. Diffraction from surfaces</i>	433
19.1. Introduction	433
19.2. Surface imaging and diffraction with electrons	435
19.2.1. Phase-contrast imaging	435
19.2.2. Crystal terminations and superlattices	436
19.2.3. Structure analysis of surface superlattices	438
19.2.4. Crystal profile imaging	439
19.3. Reflection from surfaces: grazing incidence	440
19.3.1. Kinematical approximation: x-rays, neutrons	440
19.3.2. Standing wave techniques	443
19.3.3. RHEED and REM	444
19.4. Reflection at normal incidence: LEED	451
19.5. Diffraction of emitted electrons	453
<i>References</i>	457
<i>Index</i>	477

This Page Intentionally Left Blank

Section I

PHYSICAL OPTICS

This Page Intentionally Left Blank

Fresnel and Fraunhofer diffraction

1.1. Introduction

In attempting a unified treatment of the field of X-ray, neutron and electron diffraction, including electron microscopy and other imaging techniques based on diffraction, we are faced with the task of bringing together a number of theoretical treatments developed for convenience in handling particular topics. Often the "popular" treatment used by experimentalists is not just a simplification of the more rigorous methods of the theorist. It may have a different basis which is more conceptually pleasing but less tractable mathematically or is easier to visualize as in the geometric optics approximation for imaging theory.

What we hope to do is to provide the experimentalist with a more coherent account of the basic theoretical ideas of this field, using arguments which may be followed with a certain minimum of mathematics. For both theorist and experimentalist we hope to provide an appreciation of how the physical observations using different radiations and different types of sample may be knit together with a common thread of theory so that ideas and understanding in one area may be related and transferred to other areas.

For most people using X-ray diffraction for crystal structure analysis, diffraction is a three-dimensional extension of simple Fraunhofer diffraction theory applied to the idealized case of infinite periodic objects with sharply defined diffracted beam directions and a lattice of weighted points in reciprocal space. The basic mathematical tool is the Fourier series. With greater sophistication, the concept of a Fourier transform is introduced to deal with cases of finite or imperfect crystals in the same single-scattering, kinematic approximation.

This is far removed from the theoretical treatment of X-ray diffraction by Ewald [1916] or von Laue [1931] or even from the simpler, intuitive approach of Darwin [1914]. Yet these are the sources which must be relied upon for discussion of the limitations of the assumption of the simple practical theory. Also, for the understanding and interpretation of important new techniques such as X-ray topography and X-ray interferometry and the older methods using Kossel lines, there is the same basic diffraction theory but the simplifications for practical experimental use are developed along different lines.

Neutron diffraction was developed first by nuclear physicists with a jargon of differential scattering cross sections rather than atomic scattering amplitudes. Then it diversified, with structure analysis people transferring the X-ray diffraction ideas and solid-state physicists describing their experiments in terms of k vectors, Brillouin zones and so on. The added complication of inelastic scattering studies for time-dependent processes, including phonons and magnons, brought mainly an elaboration of this solid-state physics approach rather than an extension of the Fourier transform methods.

The development of electron diffraction paralleled that of X-ray diffraction in that the simple kinematic, Fraunhofer diffraction, approximation was used where possible and the dynamical theory of Bethe [1928], analogous to von Laue's X-ray theory, was used where necessary. The difference from the X-ray case came from the greater importance of dynamical scattering effects and the greater complication of the dynamical effects which did occur. Consequently there was a need for relatively simple approximations for practical use, and several were developed to deal with different types of experiment.

An additional complication to the electron diffraction scene arose from the derivation of atomic scattering factors from scattering theory, the domain of nuclear and atomic physicists, with the attendant jargon. A legacy of this origin remains in that those engaged in electron diffraction from gases consider atomic scattering factors (in a unit of Å) as a function of the variable $s = 4\pi\lambda^{-1} \sin \theta$ rather than the variable $\lambda^{-1} \sin \theta$ used in X-ray diffraction or electron diffraction of the solid state, or the reciprocal lattice distance $u = 2\lambda^{-1} \sin \theta$.

The theory of image formation in the electron microscope naturally follows that of light microscopy, but this must be combined with the theory of the scattering or diffraction of electrons in the object. These two aspects of electron microscopy are brought together naturally by the unifying concept of the Fourier transform which has a basic significance for electron diffraction as for X-ray diffraction in the simple kinematical approximation and has also formed the basis for recent developments in imaging theory, introduced by Duffieux [1946].

The fact that the basic dynamical theories of X-ray diffraction and electron diffraction have been developed in terms of differential equations rather than in terms of integrals such as the Fourier transform then suggests a major cleavage of the necessary theory into two parts, formally related but very different in practice. However this gap has been bridged, at least for electron diffraction, by development of dynamical diffraction theories formulated in

terms of integrals, with the Fourier transform playing a major part. A further major component can be recognized as related to the elementary physics text book treatment of Fresnel diffraction.

With this in mind, we start our discussion of diffraction with a reminder to the reader of the elements of physical optics and introduce in this way the description of diffraction, scattering and imaging in terms of the Fourier transform integral and the important associated integral, the convolution.

It would probably be rather too indigestible for most practising electron microscopists and diffractionists if we attempted a complete treatment of diffraction along these lines. Instead we relate our treatment where appropriate, to the more familiar ground of the differential equation methods. However we would like to suggest that these may be considered as parallels drawn at places to an underlying unified line of development and introduced in order to satisfy the needs of convenience or convention.

For ease of understanding, we have avoided over-rigorous arguments or mathematical complexity to the point where the purist may well criticize the logic of the development. In many cases we have duplicated mathematical statements with statements in words which are frequently less precise but may be of assistance for the less mathematically-minded reader.

To start with, we assume the reader is familiar with the use of a complex exponential to represent a wave function, the use of Fourier series to represent a periodic function and the elements of integration and differentiation. For those not having this background we recommend any one of a large number of text books of physical optics such as Ditchburn [1976] or Nussbaum and Phillips [1976]. Excellent treatments at a slightly more advanced level are given by Lipson and Lipson [1969], Goodman [1968], Petykiewicz [1992] and Born and Wolf [1989].

1.2. Wave equations and waves

1.2.1. Wave functions

As a means for obtaining the intensities of diffracted beams of radiation, we make use of the convenient conventional idea of a wave function. For none of the electromagnetic radiations and particle beams we will be considering is it possible to observe any oscillatory wave motion. The wave function, a complex function of spatial coordinates which we write $\psi(\mathbf{r})$, is a convenient mathematical device for obtaining the observable quantity, the intensity or energy transfer in an experiment, given by $|\psi(\mathbf{r})|^2$. By analogy with

water waves or waves in a string, we can visualize a wave function and appreciate the concepts of the wavelength λ , the wave vector k which indicates the direction of propagation and is of magnitude $1/\lambda$, the frequency ν or angular frequency in radians per second, ω , the phase velocity of the wave v and the group velocity.

If we consider them as particles, the X-ray photons, electrons and neutrons have very different properties as shown by their collisions with other particles or, more practically, their generation and detection. However if we consider only their propagation through space and their scattering by matter or fields with no appreciable loss of energy all these radiations may be considered as waves, described by wave functions which are solutions of the same type of differential equation, the wave equation. Thus we can deal with a relatively simple semi-classical wave mechanics, rather than the full quantum mechanics needed for interactions of quanta involving changes of energy. The practical differences in experimental techniques and interpretation of measured intensities for the different radiations derive from the differing values of the parameters in the wave equation.

1.2.2. Electromagnetic waves

For electromagnetic waves, the wave equation is usually written for the electric or magnetic field vector E or H ; for example,

$$\frac{\partial^2 E}{\partial x^2} + \frac{\partial^2 E}{\partial y^2} + \frac{\partial^2 E}{\partial z^2} = \nabla^2 E = \epsilon\mu \frac{\partial^2 E}{\partial t^2}, \quad (1)$$

where ϵ is the permittivity or dielectric constant and μ is the magnetic permeability of the medium. The velocity of the waves in free space is $c = (\epsilon_0\mu_0)^{1/2}$.

For most purposes the vector nature of the wave amplitude will not be important. For simple scattering experiments with unpolarized incident radiation, for example, the only consequence of the vector properties is multiplication of the scattered intensity by a polarization factor, dependent only on the angle of scattering. Hence we may usually simplify our considerations by assuming that we are dealing with a scalar amplitude function $\psi(\mathbf{r})$ which is a solution to the wave equation,

$$\nabla^2 \psi = \epsilon\mu \frac{\partial^2 \psi}{\partial t^2}. \quad (2)$$

Except in special situations we may assume $\mu = 1$. The permittivity ϵ , the refractive index $n = (\epsilon/\epsilon_0)$ or the velocity $v = c/n$, is a function of position in space corresponding to the variation of electron density.

For purely elastic scattering, we need consider only a single frequency, ω . As we will see later, if more than one frequency is present, the intensity may be calculated for each wavelength separately and then these intensities are added. Hence we can assume that the wave-function is of the form

$$\psi = \psi_s \exp\{i\omega t\} = \psi_s \exp\{2\pi i\nu t\}.$$

Then, since $\partial^2\psi/\partial t^2 = -\omega^2\psi$, the wave equation (2) becomes

$$\nabla^2\psi = -\epsilon\mu\omega^2\psi \quad \text{or} \quad \nabla^2\psi + 4\pi^2k^2\psi = 0, \quad (3)$$

where $2\pi k$ is $(\epsilon\mu)^{1/2}\omega$ or ω/v , the wave number. Here we use the convention, now standard in the electron diffraction literature (Cowley 1992) that the modulus of the wave vector k has magnitude $|k| = 1/\lambda$, rather than the convention, current in much of solid-state physics, that $|k| = 2\pi/\lambda$.

For X-rays the dielectric constant and refractive index are very close to unity.

1.2.3. Particle waves

For electrons or other particles, the corresponding wave equation is the Schrodinger equation which may be written

$$-\frac{\hbar^2}{2m} \nabla^2\psi + e\phi(r)\psi = -i\hbar \frac{\partial\psi}{\partial t}, \quad (4)$$

where ψ is a wave function such that $|\psi|^2$ represents the probability of an electron being present in unit volume and $\phi(r)$ is the electrostatic potential function which can be assumed to vary sufficiently slowly for our purposes even though the variation is caused by the distribution of positive and negative charges in the atoms and ions in matter. If, as before, we consider only a single frequency ω , and put $\hbar\omega = eE$, where eE is the kinetic energy for $\phi = 0$ i.e. in field-free space, then the time independent wave equation is

$$\nabla^2\psi + \frac{2me}{\hbar^2} \{E + \phi(r)\} \psi = 0, \quad (5)$$

where the bracket multiplied by e gives the sum of kinetic and potential energies of the electron.

This is of exactly the same form as (3), with

$$k^2 = (2me/\hbar^2) \{E + \phi(r)\}.$$

Correspondingly the refractive index for electrons having kinetic (or incident) energy eE in a potential field $\phi(\mathbf{r})$ is

$$n = \frac{k/\omega}{(k/\omega)_0} = \frac{(E + \phi(\mathbf{r}))^{1/2}}{E^{1/2}} \approx 1 + \frac{\phi}{2E}. \quad (6)$$

The convention for the sign of ϕ used here is such that $\phi/2E$ is positive for an electron in a positive field which accelerates it. Hence the refractive index of matter for electrons is slightly greater than unity.

The simple, non-trivial solution of the wave equation (3) or (5), which can be verified by substitution, is

$$\psi = \psi_0 \exp\{-2\pi i \mathbf{k} \cdot \mathbf{r}\}, \quad (7)$$

representing a plane wave proceeding in the direction specified by the vector \mathbf{k} which has a magnitude $|\mathbf{k}| = k = 1/\lambda$. Adding the time dependence, this becomes

$$\psi = \psi_0 \exp\{2\pi i(\nu t - \mathbf{k} \cdot \mathbf{r})\}. \quad (8)$$

The choice of the positive sign of the exponent is arbitrary. A negative sign before the i is equally valid. The choice as to whether the ωt or the $\mathbf{k} \cdot \mathbf{r}$ is given the minus sign is a matter of convention. It is important only to be consistent in the choice. The form of (8) implies that for t constant the phase decreases for increasing positive distance in the direction of \mathbf{k} and for a fixed position the phase increases with time.

A further solution of importance, corresponding to a spherical wave radiating from a point source, taken to be the origin, is

$$\psi = \psi_0 \frac{\exp\{2\pi i(\nu t - kr)\}}{r}, \quad (9)$$

where $r = |\mathbf{r}|$. Since the direction of propagation is always radial, $\mathbf{k} \cdot \mathbf{r} = kr$. This form is consistent with the inverse square law, since the energy per unit area, proportional to $\psi\psi^*$, will be proportional to r^{-2} .

1.3. Superposition and coherence

1.3.1. Superposition

We have seen that the wave functions for both electromagnetic waves and particles are solutions of linear differential equations such as (3). A fundamental property of the solutions of such equations is that the sum of any two solutions is also a solution. In terms of waves, we can state

that if any number of waves, represented by the wave-functions ψ_n , coexist in space, the resultant disturbance will be represented by the wave-function $\psi = \sum_n \psi_n$. This property may be referred to as the Principle of Superposition. This is fundamental to all considerations of diffraction.

In general, waves may be considered to originate from sources having approximately the dimensions of atoms. X-rays arise when electrons jump from one energy level to another in an atom, or when an incident electron collides with an atom. Electrons may be considered to be emitted from the region of high electron density surrounding an atom. However the "uncertainty principle" implies that the origin of a photon or electron can not be specified with a precision of better than about half a wavelength, so that "point sources" must be of at least this size.

For all our present purposes, half a wavelength is sufficiently small to be neglected in comparison with all other dimensions considered so that we may think of the radiation as coming from one or more independently-emitting point sources. Hence, first of all, we must consider how we may treat the usual case of radiation emitted from a large number of such independent sources.

1.3.2. Independent point sources

We start by considering two sources at point Q_1 , and Q_2 . At some point of observation, P , the wave function is

$$\begin{aligned} \psi_P &= \psi_{1P} + \psi_{2P} = A_1 \exp\{2\pi i(\nu_1 t - k_1 \cdot r_1)\} \\ &+ A_2 \exp\{2\pi i[(\nu_2 t - k_2 \cdot r_2) + \alpha(t)]\}, \end{aligned}$$

where the phase factor $\alpha(t)$ has been added to suggest that the phase of ψ_2 relative to ψ_1 varies in a random manner with time as separate, independent quanta of radiation are generated at the two points. The amplitudes A_1, A_2 are assumed real. The intensity at P at any one time is

$$\begin{aligned} I(t) &= |\psi_{1P} + \psi_{2P}|^2 \\ &= A_1^2 + A_2^2 + 2A_1 A_2 \cos 2\pi\{(\nu_1 - \nu_2)t - (k_1 \cdot r_1 - k_2 \cdot r_2) - \alpha(t)\}. \end{aligned} \quad (10)$$

The observed intensity is the time average of $I(t)$ since the frequencies involved and the rate of fluctuations of the phase factor $\alpha(t)$ are both unobservably high; so that

$$\begin{aligned} I_{\text{obs}} &= \langle I(t) \rangle_t = \langle A_1^2 \rangle_t + \langle A_2^2 \rangle_t \\ &+ 2\langle A_1 A_2 \cos 2\pi\{(\nu_1 - \nu_2)t - (k_1 \cdot r_1 - k_2 \cdot r_2) - \alpha(t)\} \rangle_t. \end{aligned} \quad (11)$$

The time-average, indicated by the angular bracket $\langle \dots \rangle_t$, may be defined, more precisely, as

$$I_{\text{ob}}(r) = \langle I(r, t) \rangle_t \equiv \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T I(r, t) dt$$

For different frequencies, the last term fluctuates with a frequency $\nu_1 - \nu_2$ and so averages out to zero even if $A_1 = A_2$, $r_1 = r_2$ and $\alpha(t) = 0$, i.e. if the two waves come from one source. If the frequencies are the same for two separate sources, the phase term $\alpha(t)$ varies in a random manner so that the cosine term has positive and negative values with equal probability and so, again, averages to zero. Hence

$$I_{\text{obs}} = \langle A_1^2 \rangle_t + \langle A_2^2 \rangle_t = I_1 + I_2, \quad (12)$$

where I_1 and I_2 are the intensities at the point of observation due to the two sources separately.

From this rather rough, non-rigorous argument we may conclude that for any single source the observed intensity will be the sum of the intensities for different frequencies for two or more independent sources having the same or different frequencies, the observed intensity will be the sum of the intensities given by the individual sources separately.

Hence we are justified in following the usual procedure of considering diffraction problems for the ideal case of a monochromatic point source either at a finite distance or at infinity (parallel, coherent illumination) and then generalizing, if necessary, to take account of the finite source size or frequency range. A source of finite size may be considered as made up of a collection of very small, independently-emitting points. For each one of these points the intensity at a point of observation is calculated taking account of the relative intensity of the source at that point. Then the intensities are added for all source points. If there is a finite range of frequencies, the intensity is calculated for each frequency and then the summation is made for all frequencies.

The actual intensity measurement recorded depends on the characteristics of the detector, including the variation of its response with the frequency of the incident radiation and also with the position and direction of the radiation.

Most, if not all, detectors may be considered incoherent in that they are not sensitive to phase relationships of the incident radiation. They record only incident energies. We are justified in considering the intensity at each point of the detector and for each wavelength separately and then adding intensities for all points of the detector and all

wavelengths, taking the response functions of the detector into account, to find the actual measured intensity.

1.4. Huygens' principle

1.4.1. Kirchhoff's formulation

The original concept of Huygens pictures the propagation of waves through space as involving the generation of secondary waves at each point of a wavefront so that the envelope of the secondary waves becomes a new wave-front. This is a simple intuitive picture which permits an "understanding" or interpretation of the formula of Kirchhoff which may be derived directly from the wave equation by application of Green's theorem. This derivation is a standard one, reproduced in many physics texts and will not be repeated here.

The Kirchhoff formula may be written thus; the disturbance at a point P due to any wave-field u , representing a solution of the wave equation, is given by integrating over any closed surface containing P as

$$u_p = \frac{1}{4\pi} \oint_S \left(\frac{\exp\{-2\pi ikr\}}{r} \cdot \text{grad } u - u \cdot \text{grad} \left[\frac{\exp\{-2\pi ikr\}}{r} \right] \right) dS. \quad (13)$$

If we apply this to the wave field u due to a point source of unit strength at Q, we obtain

$$u_p = \frac{1}{4\pi} \int \frac{\exp\{-2\pi ikr\}}{r} \cdot \frac{\exp\{-2\pi ikr_q\}}{r_q} \\ \times \left\{ \left(\frac{1}{r} + 2\pi i k \right) \cos n^{\wedge} r - \left(\frac{1}{r_q} + 2\pi i k \right) \cos n^{\wedge} r_q \right\} dS,$$

where $n^{\wedge} r$ and $n^{\wedge} r_q$ are the angles between the surface normal n and the vectors r and r_q to the points P and Q, as in Fig. 1.1 the usual convention in this case being that all distances are measured from the surface.

Making the approximation that all the dimensions involved are much greater than the wavelength, r^{-1} and r_q^{-1} can be neglected in comparison with k and we obtain

$$u_p = \frac{i}{2\lambda} \oint \frac{\exp\{-2\pi ikr_q\}}{r_q} \cdot \frac{\exp\{-2\pi ikr\}}{r} \{ \cos n^{\wedge} r - \cos n^{\wedge} r_q \} dS. \quad (14)$$

This may be interpreted in terms of the Huygens concept

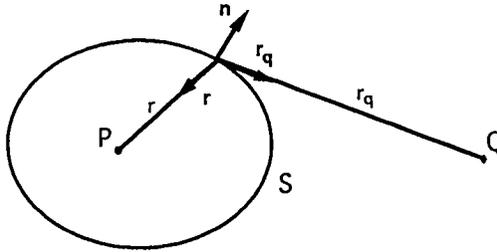


Fig. 1.1. Diagram illustrating Kirchhoff's derivation of the amplitude at P due to a source point Q.

as implying that the secondary spherical waves emitted from each surface element, $r^{-1} \exp\{-2\pi ikr\}dS$, have relative amplitude proportional to the amplitude of the incident wave from Q, $r_q^{-1} \exp\{-2\pi ikr_q\}$.

There is a change of phase of $\pi/2$, represented by multiplying by i , a scale factor of λ^{-1} , and an obliquity factor, $\frac{1}{2}[\cos n^{\wedge}r - \cos n^{\wedge}r_q]$, which ensures that the waves add up with maximum amplitude in the forward direction and are not propagated backwards. Thus the expression (14) may be considered as a mathematical way of writing a slightly modified, more accurate, version of Huygens' Principle.

1.4.2. Application of the Kirchhoff formula

The main area of usefulness of the Kirchhoff formula is not for propagation through free space, as assumed for (14), but rather for cases in which the wave function $\psi(x,y)$ represents the wave modified by the presence of an object and the surface of integration is conveniently the exit surface of the object. We neglect the many complications which may arise, particularly for wavelengths comparable with the dimensions of the structure of the object, with confidence that these complications need not concern us.

For the simple, idealized case of a planar, two dimensional object placed between P and Q, we may define a "transmission function" $q(X,Y)$ which is multiplied by the incident wave function to represent the effect of the object on the amplitude and phase of the incident wave. Then, for a point source of radiation, the wave incident on the object is $r_q^{-1} \exp\{-2\pi ikr_q\}$, the wave emerging from the object is $q(X,Y)r_q^{-1} \exp\{-2\pi ikr_q\}$ and the wave-function $\psi(x,y)$ at a point on a plane of observation is given by

$$\psi(x, y) = \frac{i}{2\lambda} \iint \frac{\exp\{-2\pi i k r_q\}}{r_q} \times \int q(X, Y) \frac{\exp\{-2\pi i k r\}}{r} \{\cos Z^{\wedge} r + \cos Z^{\wedge} r_q\} dX dY, \quad (15)$$

where X, Y and Z are coordinates referred to orthogonal axes. The surface S is considered to be the X, Y plane plus a closing surface at infinity. We follow the convention by which the Z axis is the direction of propagation.

Equations such as (15) may be applied directly to the elastic interaction of fast electrons (energies greater than about 20 keV) with matter since the wavelengths are less than 10^{-1} Å and so considerably smaller than atomic dimensions. They may also be applied to the scattering of X-rays or thermal neutrons (wavelength of the order of 1 Å) by accumulations of matter which are considerably larger than the wavelength i.e. for small-angle scattering experiments. However these equations can not be used for the consideration of the scattering of X-rays by electrons or of neutrons by nuclei. For such purposes we must rely on scattering theory, which follows a somewhat different line in its development from the wave equation, or from the fundamental postulates of quantum mechanics.

1.5. Scattering theory approach

1.5.1. Integral form of wave equation

Again, for simplicity, we consider only a scalar wave, neglecting the complications of treating vector quantities, and so deal with scattering theory as developed for the scattering of particles by a potential field. (See, for example, Wu and Ohumara [1962].)

The wave equation (5) may be written

$$[\nabla^2 + 4\pi^2 k_0^2 + \mu\phi(r)] \psi = 0, \quad (16)$$

where k_0 represents the wave number for the incident wave in free space and μ is a parameter which specifies the strength of the interaction with the potential field.

As an alternative, most useful for scattering theory, we may write the equivalent integral equation by making use of the Green's function $G(r, r')$. For scattering radiation from a potential field, $G(r, r')$ represents the amplitude at a point of observation r , due to a point of unit scattering strength at r' in the field, (see Fig. 1.2):

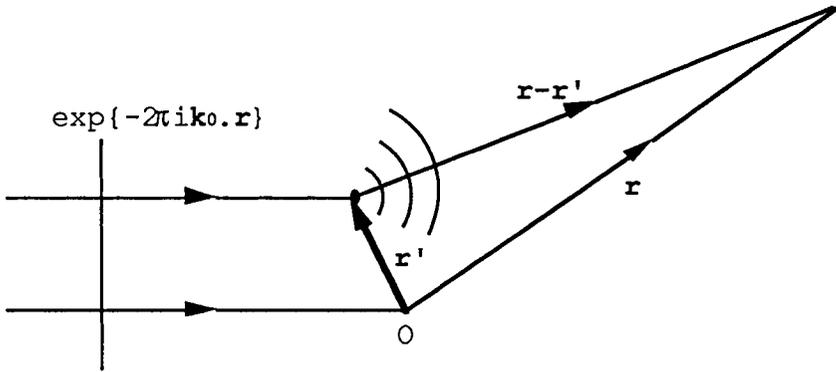


Fig. 1.2. Definition of distances for scattering problems.

$$\psi(\mathbf{r}) = \psi^{(0)}(\mathbf{r}) + \mu \int G(\mathbf{r}, \mathbf{r}') \phi(\mathbf{r}') \psi(\mathbf{r}') d\mathbf{r}' \quad (17)$$

where $\psi^{(0)}(\mathbf{r})$ represents the wave incident on the scattering field and the integral represents the scattered radiation. The appropriate form of the Green's function is

$$G(\mathbf{r}, \mathbf{r}') = \frac{\exp[2\pi i k |\mathbf{r} - \mathbf{r}'|]}{4\pi |\mathbf{r} - \mathbf{r}'|}, \quad (18)$$

which from (9) is the amplitude at \mathbf{r} of a spherical wave emitted from the point \mathbf{r}' .

Thus the expression (17) may be compared with the expression (15) derived from the Kirchhoff integral. We may interpret it as indicating that each point of the scattering field gives off a spherical wave (18), and the strength of this wave depends on the value of the scattering potential $\phi(\mathbf{r}')$ and the wave function $\psi(\mathbf{r}')$. We would have the exact three-dimensional equivalent of (15) if we could say that the amplitude of the scattered wave was proportional to the incident wave amplitude $\psi^{(0)}(\mathbf{r})$, but this is not possible in general because the scattered radiation itself forms part of the wave function $\psi(\mathbf{r})$. Hence we have an integral equation which is much more difficult to solve.

1.5.2. Born series

If the amplitude of the scattered wave is much less than the incident wave amplitude, we may assume as a first approximation that $\psi(\mathbf{r})$ in the integral may be replaced by $\psi^{(0)}(\mathbf{r})$, the incident wave amplitude. This is the "First Born Approximation". Higher order Born approximations are found by iteration. Thus the second order approximation is given by replacing $\psi(\mathbf{r})$ in the integral by $\psi^{(0)}(\mathbf{r}) + \psi^{(1)}(\mathbf{r})$ and so on.

For a plane incident wave $\psi^{(0)}(\mathbf{r}) = \exp\{-2\pi i \mathbf{k}_0 \cdot \mathbf{r}\}$, the first Born approximation gives

$$\begin{aligned} \psi^{(0)}(r) + \psi^{(1)}(r) &= \exp\{-2\pi i k_0 \cdot r\} \\ &+ \frac{\mu}{4\pi} \int \frac{\exp(-2\pi i k |r-r'|)}{|r-r'|} \varphi(r') \exp\{-2\pi i k_0 \cdot r'\} dr'. \end{aligned} \quad (19)$$

Under the usual conditions of a scattering experiment we may assume that the point of observation is at $r = R$ where R is very large compared with the dimensions of the scattering field. Then, writing $q = k - k_0$, we obtain the asymptotic form,

$$\begin{aligned} \psi^{(0)}(r) + \psi^{(1)}(r) &= \exp\{-2\pi i k_0 \cdot R\} \\ &+ \frac{\mu}{4\pi} \frac{\exp\{-2\pi i k_0 R\}}{R} \int \varphi(r') \exp\{-2\pi i q \cdot r'\} dr'. \end{aligned} \quad (20)$$

The characteristic scattering from the potential field may then be sorted out from this by defining a scattering amplitude $f(q)$ such that the right-hand side of (20) is

$$\exp\{-2\pi i k_0 \cdot R\} + R^{-1} \exp\{-2\pi i k_0 R\} f(q).$$

$$\text{Then} \quad f(q) = \frac{\mu}{4\pi} \int \varphi(r) \exp\{-2\pi i q \cdot r\} dr. \quad (21)$$

This is the first Born approximation for the scattering amplitude, given by assuming that the incident wave amplitude is equal to the total wave amplitude in the scattering field: i.e. that the scattered amplitude is negligibly small. The scattered wave is made up of contributions scattered directly from the incident wave. Hence this is a single-scattering approximation.

In general this first Born approximation is very good and useful for weakly scattering fields or objects. For stronger scattering, further terms may be calculated in the Born series

$$\psi(r) = \psi^{(0)}(r) + \psi^{(1)}(r) + \psi^{(2)}(r) + \dots$$

by use of the recurrence relationship

$$\psi^{(n)}(r) = \frac{\mu}{4\pi} \int \frac{\exp(-2\pi i k |r-r'|)}{|r-r'|} \varphi(r') \psi^{(n-1)}(r') dr'. \quad (22)$$

However the convergence of this Born series is usually poor if the first order approximation fails. The addition of the second order term improves the approximation over a rather limited range of scattering strengths and is sometimes useful in suggesting the nature of the modifications needed when the first order approximation fails. But the higher order terms become rapidly more complicated and difficult to calculate and it is not often useful to evaluate them.

1.6. Reciprocity

One point of particular interest for the practical application of these mathematical treatments is illustrated by the formulas we have given. In the expressions derived from the Kirchhoff formula relating to radiation from a point Q and a point of observation P it is seen that the expressions are symmetrical with respect to P and Q. If the source were at P and the point of observation were at Q, so that the vectors r and r_q were interchanged, the same amplitude would result.

The expression (20) for single scattering, and the more general series (17) relate an incident plane wave defined by k_0 and an out-going plane wave k , corresponding to source and observation points at infinity. Again, if the source and observation points were interchanged there would be no change in the resulting amplitude. These observations are consistent with the general Reciprocity theorem of wave optics introduced by Helmholtz [1886]. This may be restated as follows:

The amplitude of the disturbance at a point P due to radiation from a point Q which has traversed any system involving elastic scattering processes only, is the same as the amplitude of the disturbance which would be observed at Q if the point source were placed at P.

The first application of this theorem to diffraction phenomena was made by von Laue [1935] who employed it to simplify the theoretical treatment of Kossel lines due to diffraction of radiation originating from point sources within a crystal. More recently Pogany and Turner [1968] showed its general applicability and usefulness in electron diffraction and electron microscopy under n -beam dynamical diffraction conditions. They further showed that the same relationship applies for intensities (but not amplitudes) if inelastic scattering is involved, provided that the energy loss in the inelastic scattering process is negligible.

The applications of reciprocity have recently multiplied in number and importance. We will meet several in later chapters.

1.7. Fresnel diffraction

1.7.1. Small angle approximation

From the general Kirchhoff formula it is possible to derive relatively simple forms appropriate to the conditions which typify particular classes of diffraction effects. "Fresnel diffraction" refers usually to, although not

exclusively, the phenomena observed close to a two-dimensional object illuminated by plane parallel incident light. If the object plane is perpendicular to the direction of incidence we may replace the incident radiation in (15) by $\Psi_0 = 1$, which represents a plane wave of unit amplitude having zero phase at $Z = 0$. Then the amplitude on any plane of observation a distance R beyond the object is

$$\Psi(x, y) = \frac{i}{2\lambda} \iint q(X, Y) \frac{\exp(-2\pi i k r)}{r} (1 + \cos Z^{\wedge} r) dX dY \quad (23)$$

where $r^2 = (x - X)^2 + (y - Y)^2 + R^2$.

If, as was assumed earlier, the wavelength is small compared with the dimensions of the object the angles of deflection of the radiation will be small and we may make a small angle approximation, putting $\cos Z^{\wedge} r = 1$, $r = R$ in the denominator and, in the exponent,

$$\begin{aligned} r &= \{R^2 + (x - X)^2 + (y - Y)^2\}^{1/2} \\ &\approx R + \frac{(x - X)^2 + (y - Y)^2}{2R}, \end{aligned} \quad (24)$$

so that

$$\begin{aligned} \Psi(x, y) &= \frac{i \exp\{-2\pi i k \cdot R\}}{R\lambda} \\ &\times \iint q(X, Y) \exp\left\{\frac{-2\pi i [(x - X)^2 + (y - Y)^2]}{R\lambda}\right\} dX dY. \end{aligned} \quad (25)$$

For the special case that the object has a transmission function $q(X)$ which varies in one dimension only, as in the idealized cases of straight edges, slits and so on, the integral over Y can be carried out to give

$$\int_{-\infty}^{\infty} \exp\left\{\frac{-2\pi i k (y - Y)^2}{2R}\right\} dY = \left(\frac{R\lambda}{i}\right)^{1/2}$$

so that

$$\Psi(x) = \left(\frac{i}{R\lambda}\right)^{1/2} \exp\{-2\pi i k \cdot R\} \int_{-\infty}^{\infty} q(X) \exp\left\{\frac{-\pi i (x - X)^2}{R\lambda}\right\} dX. \quad (26)$$

1.7.2. Fresnel integrals

For the special cases of objects which are either completely opaque or transparent so that $q(X) = 1$ in some

parts and $q(X) = 0$ elsewhere, this integral reduces to a sum of integrals of the form

$$\int_a^b \exp\left\{\frac{-\pi i(x-X)^2}{R\lambda}\right\} dX.$$

These integrals can be expressed in terms of the Fresnel integrals

$$C(x) = \int_0^x \cos\left(\frac{1}{2}\pi u^2\right) du,$$

$$S(x) = \int_0^x \sin\left(\frac{1}{2}\pi u^2\right) du.$$

For example let us consider a single slit of width a for which the transmission function in one dimension is

$$q(x) = \begin{cases} 1 & \text{if } |x| \leq a/2 \\ 0 & \text{if } |x| > a/2 \end{cases}$$

Then the integral of (26) is

$$\int_{-a/2}^{a/2} \exp\left\{\frac{-\pi i(x-X)^2}{R\lambda}\right\} dX.$$

Putting $u^2 = 2(x-X)^2/R\lambda$, this becomes

$$\int_A^B \cos\left(\frac{1}{2}\pi u^2\right) du - i \int_A^B \sin\left(\frac{1}{2}\pi u^2\right) du$$

where

$$A^2 = 2\left(x + \frac{1}{2}a\right)^2/R\lambda \quad \text{and} \quad B^2 = 2\left(x - \frac{1}{2}a\right)^2/R\lambda;$$

and this is equal to

$$\{C(B) - C(A)\} - i\{S(B) - S(A)\}.$$

The Fresnel integrals may be evaluated numerically by reference to tabulations of these functions. Alternatively the amplitudes may be found graphically by use of the very ingenious device of the Cornu Spiral.

These methods and results for the most usual applications are described in detail in most text books of optics and so will not be reproduced here. The typical Fresnel fringes

formed by diffraction at a straight edge are well known in both visible-light optics and electron optics and are used in electron microscopy as an aid to focussing.

1.7.3. Periodic objects - "Fourier images"

There are very few functions $q(X)$ for which the integral (26) can be evaluated analytically, but the few favorable cases include the important one of periodic functions which we consider in some detail here because of its interesting implications for the electronoptical imaging of crystals.

Consider a planar object with transmission function $q(X) = \cos(2\pi X/a)$. This would be difficult to achieve in practice, but not impossible since the negative sign required could be obtained by use of a "half-wave plate" which changes the phase of the incident radiation by π . The integral (26) becomes

$$\Psi(x) = \left(\frac{i}{R\lambda}\right)^{\frac{1}{2}} \exp\{-2\pi i kR\} \int_{-\infty}^{\infty} \cos\left(\frac{2\pi X}{a}\right) \exp\left\{-\frac{\pi i(x-X)^2}{R\lambda}\right\} dX, \quad (27)$$

or, putting $X = x - W$,

$$\begin{aligned} \Psi(x) &= \left(\frac{i}{R\lambda}\right)^{\frac{1}{2}} \exp\{-2\pi i kR\} \cos\left(\frac{2\pi x}{a}\right) \\ &\quad \times \int_{-\infty}^{\infty} \cos\left(\frac{2\pi W}{a}\right) \exp\left\{-\frac{\pi i W^2}{R\lambda}\right\} dW. \end{aligned}$$

The corresponding sine term is an integral over an odd function and so is zero. Making use of the standard integral form

$$\int_{-\infty}^{\infty} \exp\{-a^2 x^2\} \cos bx \, dx = \frac{\pi^{1/2}}{a} \exp\{-b^2/4a^2\}, \quad (28)$$

we obtain

$$\Psi(x) = \exp\{-ikR\} \exp\{i\pi R\lambda/a^2\} \cos(2\pi x/a), \quad (29)$$

so that the intensity distribution on the plane of observation is

$$I(x) = \cos^2(2\pi x/a). \quad (30)$$

This is independent of R . Hence on any plane of observation at any distance from the object, the intensity distribution will be exactly the same as at the exit face of the object. If such an object, illuminated by plane parallel

light, were imaged, the image could never be made to go out of focus.

For a general periodic object, we write

$$q(X) = \sum_h F_h \exp\{2\pi i h x/a\}.$$

Then (26) becomes

$$\psi(x) = \left(\frac{i}{R\lambda}\right)^{\frac{1}{2}} \exp\{-2\pi i k \cdot R\} \sum_h F_h \int \exp\{2\pi i \frac{hx}{a}\} \exp\left\{\frac{-\pi i(x-X)}{R\lambda}\right\} dX$$

The integral over X becomes

$$\exp\left\{\frac{-\pi i x^2}{R\lambda}\right\} \int \exp\left\{\frac{-\pi i X^2}{R\lambda}\right\} \exp\left\{2\pi i \left(\frac{h}{a} + \frac{x}{R\lambda}\right) X\right\} dX$$

Making use of (28) then gives

$$\psi(x) = \exp\{-2\pi i k \cdot R\} \sum_h F_h \exp\{\pi i R \lambda h^2/a^2\} \exp\{2\pi i h x/a\}. \quad (31)$$

This is again a periodic distribution, having the same periodicity as the object but with relative phases of the Fourier coefficients which vary with the distance R .

For special values of R given by $R = 2na^2/\lambda$, the exponent in (31) is $2N\pi$ where $N = nh^2$ is an integer. Then

$$\psi(x) = \exp\{-2\pi i k R\} \sum_h F_h \exp\{2\pi i h x/a\} \quad (32)$$

and the intensity distribution is exactly the same as at the exit surface of the object.

For the values $R = na^2/\lambda$ where n is odd, the exponent will be $(2N + 1)\pi$, so that, since h is odd if nh^2 is odd,

$$\begin{aligned} \psi(x) &= \exp\{-2\pi i k R\} \sum_h F_h \exp\{-i\pi h\} \exp\{2\pi i h x/a\} \\ &= \exp\{-2\pi i k R\} \sum_h F_h \exp\{2\pi i \frac{h}{a}(x \pm a/2)\} \end{aligned} \quad (33)$$

Thus again the intensity distribution is exactly the same as for $R = 0$ except that it is translated by half the periodicity.

Thus for parallel incident illumination the intensity distribution is exactly repeated at intervals of $R = na^2/\lambda$. This phenomenon was first reported by Talbot [1836], was partially explained by Rayleigh [1881], explored further by Weisel [1910] and Wolfke [1913] and "rediscovered" and

examined in detail for possible applications to electron diffraction and electron microscopy by Cowley and Moodie [1957a,b,c; 1960] who named these self-images of a periodic object "Fourier Images". A number of different aspects of the phenomenon will be explored in various contexts in later pages. Here we mention only the case that the incident radiation comes from a point source at a finite distance, say R_1 , from the periodic object. Then (26) is replaced by

$$\psi(x) = C \int_{-\infty}^{\infty} \exp\left\{\frac{-\pi i X^2}{R_1 \lambda}\right\} q(X) \exp\left\{\frac{-\pi i (x-X)^2}{R \lambda}\right\} dX, \quad (34)$$

where the first exponential function represents a spherical wave from the point source incident on the object having transmission function $q(X)$.

It is readily shown that Fourier images will be magnified by a factor $(R + R_1)/R_1$ and will occur at positions given by

$$\frac{1}{R} + \frac{1}{R_1} = \frac{\lambda}{na^2}. \quad (35)$$

The real Fourier images occur at successive larger intervals in R up to some maximum positive value of n , beyond which λ/na^2 is less than $1/R_1$. Then there are virtual Fourier images on the source side of the object corresponding to negative values of n and to values from plus infinity down to the n value for which $R_1 = na^2/\lambda$, as suggested in Fig. 1.3.

1.8. Fraunhofer diffraction

The approximation to the general Kirchhoff formula which typifies the Fraunhofer diffraction condition is that the overall dimensions of the object should be very much smaller than the distances to source or point of observation; or, in the more usual parlance, the source and point of observation are effectively at infinity. Thus we assume an incident plane wave of unit amplitude and write

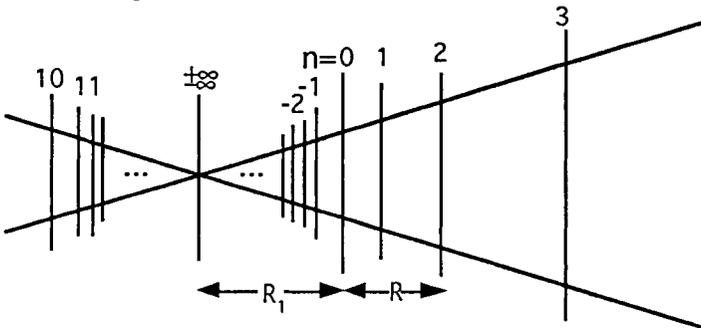


Fig. 1.3. Positions of the Fourier images of a periodic object, $n = 0$, illuminated by a point source. The Fourier image, $n = 5$, is at infinity.

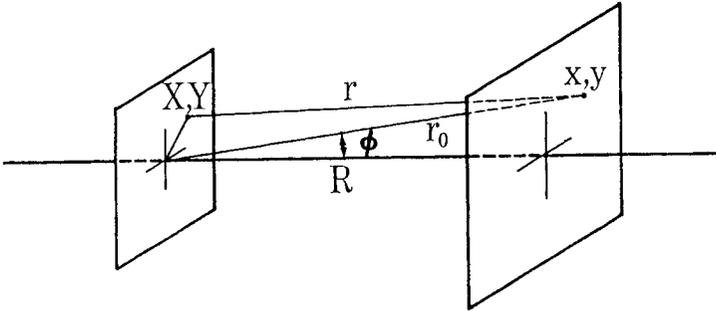


Fig. 1.4. Coordinate systems for the description of Fraunhofer diffraction.

$$\psi(x, y) = \frac{i}{\lambda} \left(\frac{1 + \cos \phi}{2} \right) \iint q(X, Y) \frac{\exp[-2\pi i k r]}{r} dX dY. \quad (36)$$

The obliquity factor is taken outside the integral since it will not vary appreciably for the small range of X and Y considered. It may be included with the other constants and terms of modulus unity which are combined into a factor C and usually ignored when only relative, and not absolute intensities, are of interest. The distance r is compared to the distance r_0 from the origin of coordinates in the object to the point of observation, as suggested in Fig. 1.4.

Then

$$\begin{aligned} r &= \{R^2 + (x-X)^2 + (y-Y)^2\}^{1/2} \\ &\approx \{ (R^2 + x^2 + y^2) - 2(xX + yY) \}^{1/2} \\ &\approx r_0 - \frac{x}{r_0}X - \frac{y}{r_0}Y. \end{aligned}$$

Putting

$$\begin{aligned} x/r_0 &= l = \sin \phi_x, \\ y/r_0 &= m = \sin \phi_y, \end{aligned}$$

where ϕ_x and ϕ_y are the components of the scattering angle, we have the amplitude as a function of the angular variables;

$$\psi(l, m) = C \int_{-\infty}^{\infty} \int q(X, Y) \exp\{2\pi i k (lX + mY)\} dX dY \quad (37)$$

and

$$C = (1 + \cos \phi) i \frac{\exp[-2\pi i k r_0]}{2r_0 \lambda}.$$

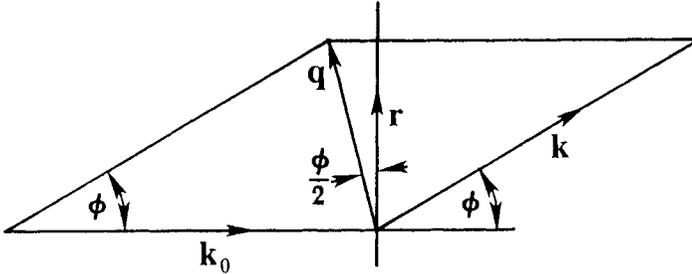


Fig. 1.5. Definition of scattering vectors.

From this formula, all the well-known results on the Fraunhofer diffraction from one and two dimensional objects, such as slits, apertures, gratings etc., may be derived. However we note at this stage that the integral in (37) has the form of a Fourier transform integral. In the next chapter we outline the properties of the Fourier transform and work out the form of the diffraction pattern for a number of simple objects as examples in the use of this type of transform which represents the basic tool for most kinematical diffraction theory and a great deal of dynamical theory.

At this stage we note the equivalence of (37) and the expression for the diffracted amplitude in the First Born approximation of scattering theory (20). If we consider the transmission function $q(X, Y)$ to be derived from a planar distribution of scattering potential $V(\mathbf{r}')$, the integral of (20) is confined to values of \mathbf{r}' lying in the X, Y plane.

As seen from Fig. 1.5, the magnitude of q is

$$|q| = 2k \sin(\phi/2),$$

so that, for the X direction, for example,

$$q \cdot \mathbf{r}' = 2kr' \sin(\phi_x/2) \cos(\phi_x/2) = kr \sin\phi_x = k l X.$$

Hence the integrals of (20) and (37) are equivalent. The integral of (20) has the advantage that it may be immediately applied to three-dimensional distributions provided that the necessary conditions of weak scattering apply. The form (37) necessarily deals with objects which can be considered two-dimensional and is essentially a single-scattering form, since no distance can exist between double scattering events, but there is no restriction on the strength of the scattering.

The presence of the obliquity factor in (37) is a consequence of the assumption of a planar distribution of the scattering function. It does not exist in (20) because no such restriction is placed on $V(r')$.

The manner in which (37) may be used to develop a treatment for the scattering from extended three-dimensional distributions will be discussed in a later chapter. In this way we provide an additional approach to the problem of the many-beam dynamical theory for the diffraction from strongly scattering crystals.

Problems

1. Starting from the expression (34) show that for a point source of radiation Fourier images of a periodic object will occur at positions given by equation (35). Find the magnification of the images. [Note: make use of the standard integral, (28).]

2. Make use of a Cornu spiral to plot, to scale, the intensity distribution in a plane of observation 5000 \AA behind an opaque straight edge illuminated by a parallel beam of electrons of wavelength 0.04 \AA (and energy 87 keV). [Note: good Cornu spirals are given in many text books including Stone, *Radiation and Optics* and Joos, *Theoretical Physics* and in Jahnke and Emde, *Tables of Functions*.]

Give qualitative arguments to suggest how this intensity distribution would be affected if

(a) the plane of this Fresnel diffraction pattern is imaged by an electron microscope having a resolution of 10 \AA ,
or,

(b) the incident electron beam comes from a finite source which subtends an angle of 10^{-3} radians at the straight edge.

3. A thin object of uniform small thickness may be approximated by a pure phase object for transmission of high energy electrons. Derive an expression for the intensity in an out-of-focus image (i.e. the intensity in the Fresnel diffraction pattern) of a straight edge of such an object (e.g. the edge of a thin crystal of uniform thickness) in terms of the Fresnel integrals. Show that, if the phase change in the object, relative to that in vacuum is π , the intensity distribution can be described in terms of the square of the distance from one point in the Cornu spiral diagram to a point on the Cornu spiral. (This exercise may be extended to treat phase objects having arbitrary phase changes and also to mixed phase- and amplitude-objects, if desired).

Fourier transforms and convolutions

2.1. Preliminaries

2.1.1. Delta-functions and discontinuities

In this chapter we provide the mathematical background for much of what follows. Most considerations of kinematical diffraction involve the use of the Fourier transform in one form or another. One of the most important properties of the Fourier transform involves the concept of a Convolution, or Convolution integral, and for both this and the Fourier transform it is convenient to use delta-functions. Hence we define and discuss this latter function first so that we may make use of it later.

In this, as in all subsequent mathematical exposition, we do not strive for mathematical rigor. We assume that all functions considered are sufficiently well-behaved in a mathematical sense to allow us to use them in a way which makes physical sense. When we make use of delta-functions and other discontinuous functions we do so as a matter of convenience. When we wish to represent real situations, in which there can be no discontinuities, we use the discontinuous mathematical functions as a short-hand notation for the continuous functions of physical reality to which they approximate.

For example, a Dirac delta function at $x = a$ is defined by

$$\delta(x - a) = \begin{cases} 0 & \text{for } x \neq a \\ \infty & \text{for } x = a \end{cases}$$

and

$$\int_{-\infty}^{\infty} \delta(x - a) dx = 1. \quad (1)$$

The delta function at $x = 0$, $\delta(x)$, can be considered as the limit of a set of real continuous functions, such as gaussians:

$$\delta(x) = \lim_{a \rightarrow \infty} [a\pi^{-1/2} \exp\{-a^2x^2\}]. \quad (2)$$

As a tends to infinity the gaussian function has a maximum value tending to infinity and a half-width ($1/a$) tending to zero but the integral over the function is always unity. Then we can use a delta function as a convenient notation for any function of integral unity in the form of a sharp peak having a width so small that it is not experimentally significant.

Similarly a weighted delta function $c.\delta(x)$ is used to indicate a sharp peak for which the integral is c . The definition of the function as the limit of a series of functions can be convenient to clarify or prove various relationships. For example,

$$\delta(bx) = \lim_{a \rightarrow \infty} [a\pi^{-1/2} \exp\{-a^2 b^2 x^2\}] = b^{-1}.\delta(x).$$

We may define a delta function in two dimensions, $\delta(x,y)$ which is zero except at $x = y = 0$ and for which

$$\iint \delta(x,y) dx dy = 1.$$

Similarly for any number of dimensions we may define $\delta(r)$ or $\delta(r-a)$ where r and a are vectors in the n -dimensional space.

We note that in two dimensions $\delta(x)$ represents a line and in three dimensions $\delta(x)$ represents a plane.

For a detailed discussion of delta functions see, for example, Lighthill [1960] or Arsac [1966] or Shmueli [1992].

We note in passing the important definition of a delta function,

$$\delta(x) = \int_{-\infty}^{\infty} \exp\{2\pi ixy\} dy, \quad (3)$$

which will occur later in connection with Fourier transforms (equations (33) to (36)).

2.1.2. *Convolution*

In one dimension, the Convolution integral (or, convolution product, or convolution, or faltung, or folding) of two functions $f(x)$ and $g(x)$ is defined as

$$C(x) = f(x) * g(x) \equiv \int_{-\infty}^{\infty} f(X)g(x - X) dX. \quad (4)$$

By simple change of variable, we find that

$$f(x) * g(x) = \int_{-\infty}^{\infty} g(X)f(x - X) dX = g(x) * f(x). \quad (5)$$

For two or more dimensions we may use the vector form,

$$f(\mathbf{r}) * g(\mathbf{r}) = \int f(\mathbf{R})g(\mathbf{r} - \mathbf{R}) d\mathbf{R}. \quad (6)$$

The identity operation is the convolution with the Dirac delta function:

$$\begin{aligned} f(x) * \delta(x) &= f(x), \\ f(x) * \delta(x - a) &= f(x - a). \end{aligned} \quad (7)$$

2.1.3. Examples of convolutions

The convolution integral (4) or (6) appears with great frequency in many areas of scientific work, being fundamental to the interpretation of most experimental measurements and an essential component of many sophisticated theoretical developments such as the Green's function methods of theoretical physics. To get a clearer understanding of what is involved we analyse the integral (4) in detail. It may be written thus; the function $f(X)$ is multiplied by the function $g(X)$ which has been shifted to an origin at $X = x$ and inverted to give $g(x - X)$. The value of the product of $f(X)$ and $g(x - X)$ is integrated over X and the result plotted as a function of x to give $C(x)$.

This is exactly the process involved, for example, in measuring the intensity of a spectral line by scanning it with a detector having a finite slit as input aperture, as suggested in Fig. 2.1. The coordinate X may represent the angle of scattering of light by a prism or diffraction grating and the intensity distribution $I(X)$ shows the spectral lines of interest.

The intensity distribution in the spectrum is measured by recording the intensity passing through a slit whose transmission function is

$$g(X) = \begin{cases} 0 & \text{for } |X| \geq a/2, \\ 1 & \text{for } |X| < a/2, \end{cases}$$

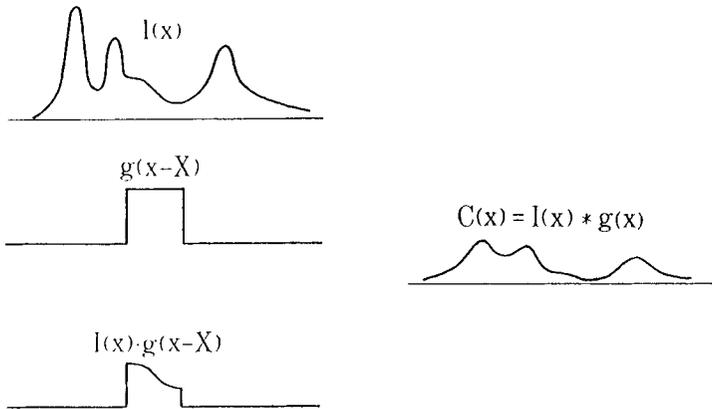


Fig. 2.1. A convolution operation. An intensity function $I(X)$ representing a spectral line, is multiplied by the transmission function $g(x-X)$ of a slit, centered at $X=x$; the product of the two functions is integrated to give the measured intensity $I_{\text{obs}}(x)$.

i.e. the slit transmits all radiation within a range of X of width a and excludes all other radiation. When the slit is placed at $X = x$, the transmitted intensity as a function of X is $I(X)g(x - X)$. What is recorded is the total transmitted intensity. If this is plotted as a function of x , the position of the slit, we get

$$I_{\text{obs}} = I(x) * g(x) = \int_{-\infty}^{\infty} I(X)g(x - X) dX.$$

A single very sharp spectral line gives an observed intensity of unity for a range of x values equal to the slit width a ;

$$I_{\text{obs}} = \delta(x) * g(x) = \int_{-\infty}^{\infty} \delta(X)g(x - X) dX = g(x).$$

For a general intensity distribution, $I(X)$, each sharp spectral line or each part of a broader spectral line will be "spread out" by the "spread-function" $g(X)$, so that the recorded intensity I_{obs} is less sharply peaked or less well resolved than the original spectrum.

Likewise, the blurring of the image due to the imperfections of a camera lens may be described in terms of convolution of the ideally perfect image intensity with some function $g(x,y)$. For a point source of light, the ideal image would be a delta function. The spreading gives

$$I_{\text{obs}} = g(x,y) * \delta(x,y) = g(x,y). \quad (8)$$

For a general object, consisting of a large number of independently emitting point sources, having an ideal image, $I_0(x, y)$,

$$I_{\text{obs}} = g(x, y) * I_0(x, y). \quad (9)$$

Thus each point of the original intensity distribution is spread into a disc of intensity and the overlapping of these discs gives a blurring and loss of resolution of the image. This illustrates the concept that the "spread function" represents the response of the system to a delta function input, in this case a point source. This is the basis for the Green's function method useful in scattering theory and many other fields of physics and also for the analysis of the properties of an electronic circuit by measuring its response to a sharp voltage or current pulse.

An excellent example of a convolution is provided by Huygens' Principle as expressed by the Kirchhoff formulas. Each point on a wave front is considered to give rise to a spherical secondary wave having an initial amplitude proportional to the amplitude of the incident wave. Then the amplitudes of the secondary waves are added to give the amplitude on a plane of observation. Thus the amplitude function, $q(x, y)$, on the original wave front is spread out by a function which represents the spherical secondary wave emitted by a single point on the wave front.

We have written this explicitly as a convolution integral in the equation (1.25) for Fresnel diffraction in the small angle approximation. This may then be rewritten

$$\psi(x, y) = q(x, y) * \left[\frac{i \exp\{-2\pi i k R\}}{R \lambda} \exp\left\{ \frac{-2\pi i k (x^2 + y^2)}{2R} \right\} \right], \quad (10)$$

and the function in square brackets may be referred to as the "propagation function", or the wave function obtained for a point source, $q(x, y) = \delta(x, y)$.

Likewise the First Born approximation result given in eq. (1.19) may be rewritten in terms of a convolution. The integral giving the singly scattered amplitude becomes,

$$V(r) \exp\{-2\pi i k_0 \cdot r\} * \frac{\exp\{-2\pi i k r\}}{r}. \quad (11)$$

The first function here is the incident wave modified by the potential field $V(r)$. This is convoluted with the amplitude due to a point source, namely the amplitude of a spherical wave from the origin. Thus the equation (1.19) or (11) simply states that the observed amplitude is the sum of the amplitudes of spherical waves from all points of the scatterer, and the amplitude of scattering from each point is proportional to the product of the incident wave amplitude and the value of the potential function, $V(r)$ at the point.

2.2. Fourier transforms: general

2.2.1. Definitions

The Fourier transform of a one-dimensional function $f(x)$ is defined as

$$\mathfrak{F}[f(x)] \equiv F(u) = \int_{-\infty}^{\infty} f(x) \exp\{2\pi i u x\} dx. \quad (12)$$

The inverse transform, \mathfrak{F}^{-1} , is defined so that

$$\begin{aligned} f(x) &= \mathfrak{F}^{-1}[\mathfrak{F}[f(x)]] \\ &= \int_{-\infty}^{\infty} F(u) \exp\{-2\pi i u x\} du. \end{aligned} \quad (13)$$

Here we follow the convention of including 2π in the exponent. This is the convention commonly used in considerations of diffraction and is convenient in that it avoids the necessity of adding a constant multiplier in either (12) or (13). In other conventions, often used in solid-state physics, the 2π is omitted from the exponent. It must then be included as a constant; the integral in either (12) or (13) is multiplied by $(2\pi)^{-1}$ or both integrals are multiplied by $(2\pi)^{-1/2}$. Also some authors choose to define the Fourier transform with a negative sign in the exponential and the inverse transform with a positive sign. It is necessary, therefore, to use considerable caution in making use of Fourier transform expressions derived from various sources.

For more than one dimension we may use the vector form of (12)

$$F(\mathbf{u}) = \int f(\mathbf{r}) \exp\{2\pi i \mathbf{u} \cdot \mathbf{r}\} d\mathbf{r}. \quad (14)$$

The vector \mathbf{u} may be considered as a vector in "Fourier transform space". For the three-dimensional case, for example, the vector \mathbf{r} may be considered to have coordinates x, y, z and \mathbf{u} may be considered to have coordinates u, v, w .

Then the scalar product is $\mathbf{u} \cdot \mathbf{r} = ux + vy + wz$ and

$$F(u, v, w) = \int \int \int_{-\infty}^{\infty} f(x, y, z) \exp\{2\pi i (ux + vy + wz)\} dx dy dz, \quad (15a)$$

and

$$f(x, y, z) = \int \int \int_{-\infty}^{\infty} F(u, v, w) \exp\{-2\pi i (ux + vy + wz)\} du dv dw \quad (15b)$$

We have seen in Chapter 1 that the amplitude of scattering from an object in the Fraunhofer diffraction approximation, whether derived from the Kirchhoff formulation or from scattering theory, is described by a Fourier transform integral. In (1.37) for example, we put $u = l/\lambda$, $v = m/\lambda$ to get the two-dimensional form of (15b). We may thus describe the diffraction amplitude in terms of a distribution in Fourier transform space, often referred to, as we shall see, as "Reciprocal space". This will be our most common means for deriving or understanding diffraction effects, so we now proceed to describe and illustrate the most important properties and behavior of the Fourier transform.

2.2.2. Properties of Fourier transforms

Instead of using the complex exponential, we may write (12) as

$$F(u) = \int_{-\infty}^{\infty} f(x) \cos(2\pi ux) dx + i \int_{-\infty}^{\infty} f(x) \sin(2\pi ux) dx. \quad (16)$$

If the function $f(x)$ is real and an even function, so that $f(-x) = f(x)$, the sine integral is zero so that

$$F(u) = \int_{-\infty}^{\infty} f(x) \cos(2\pi ux) dx = 2 \int_0^{\infty} f(x) \cos(2\pi ux) dx, \quad (17)$$

and $F(u)$ is a real function.

If $f(x)$ is a real odd function so that $f(-x) = -f(x)$ then the cosine integral is zero and

$$F(u) = i \int_{-\infty}^{\infty} f(x) \sin(2\pi ux) dx = 2i \int_0^{\infty} f(x) \sin(2\pi ux) dx, \quad (18)$$

and the function $F(u)$ is pure imaginary.

Since any real function can be written as a sum of an even and an odd function,

$$f(x) = \frac{1}{2} \{f(x) + f(-x)\} + \frac{1}{2} \{f(x) - f(-x)\} = f_e(x) + f_o(x), \quad (19)$$

we may write

$$F(u) = A(u) + iB(u)$$

where $A(u)$ and $B(u)$ are real functions given by

$$A(u) = 2 \int_0^{\infty} f_e(x) \cos(2\pi ux) dx,$$

and

$$B(u) = 2 \int_0^{\infty} f_0(x) \sin(2\pi ux) dx. \quad (20)$$

It is these cosine and sine integrals that are tabulated, for the most part, in the lists of Fourier integrals given, for example, in the compilations of Erdelyi [1954] and Bracewell [1986]. Photographic representations of Fourier transforms, obtained by optical diffraction, are given by Harburn et al. [1975].

A number of general relationships may be written for any function $f(x)$, real or complex, thus;

<i>Real space</i>	<i>Fourier transform space</i>	
$f(x)$	$F(u)$	(21)

$f(-x)$	$F(-u)$	(22)
---------	---------	------

$f^*(x)$	$F^*(-u)$	(23)
----------	-----------	------

$f(ax)$	$\frac{1}{a} F(u/a)$	(24)
---------	----------------------	------

$f(x) + g(x)$	$F(u) + G(u)$	(25)
---------------	---------------	------

$f(x - a)$	$\exp\{2\pi i a u\} F(u)$	(26)
------------	---------------------------	------

$df(x)/dx$	$(-2\pi i u) F(u)$	(27)
------------	--------------------	------

$d^n f(x)/dx^n$	$(-2\pi i u)^n F(u)$	(28)
-----------------	----------------------	------

These relationships may be readily proved by writing out the relevant integrals. For (24):

$$\begin{aligned} & \int_{-\infty}^{\infty} f(ax) \exp\{2\pi i u x\} dx \\ &= \frac{1}{a} \int_{-\infty}^{\infty} f(X) \exp\{2\pi i u X/a\} dX \\ &= \frac{1}{a} F(u/a). \end{aligned}$$

For (26):

$$\begin{aligned} & \int_{-\infty}^{\infty} f(x - a) \exp\{2\pi i u x\} dx \\ &= \int_{-\infty}^{\infty} f(x) \exp\{2\pi i (uX + ua)\} dX \\ &= F(u) \exp\{2\pi i ua\}. \end{aligned}$$

For (27):

$$\begin{aligned} & \int \frac{d}{dx} f(x) \exp\{2\pi i u x\} dx \\ &= \iint \frac{d}{dx} [F(v) \exp\{-2\pi i u x\}] \exp\{2\pi i u x\} dx \\ &= \int (-2\pi i v) F(v) \int \exp\{2\pi i (u-v)x\} dv dx \\ &= (-2\pi i u) F(u) \text{ since } \int \exp\{2\pi i (u-v)x\} dx = \delta(u-v). \end{aligned}$$

The relation (28) follows by repetition of the derivation of (27).

2.2.3. Multiplication and convolution

We add the two important relationships, the Multiplication theorem

$$\mathfrak{S} [f(x) \cdot g(x)] = F(u) * G(u), \quad (29)$$

i.e. the Fourier transform of a product of two functions is the convolution of their Fourier transforms, and the Convolution theorem,

$$\mathfrak{S} [f(x) * g(x)] = F(u) \cdot G(u), \quad (30)$$

i.e. the Fourier transform of the convolution of two functions is the product of their Fourier transforms.

Here we have followed the convention that functions in real space are represented by small letters and the Fourier transforms are represented by the corresponding capital letters.

These theorems are again easily proved, using manipulations which are in general non-rigorous but are valid for the types of functions which we will be employing. For example, to derive (30), putting $x - X = y$ in the left side of the integral form gives,

$$\begin{aligned}
& \iint f(X)g(x - X) dX \cdot \exp\{2\pi iux\} dx \\
&= \int f(X)g(y) \exp\{2\pi iu(X+y)\} dX dy \\
&= \int f(X) \exp\{2\pi iuX\} dX \int g(y) \exp\{2\pi iuy\} dy \\
&= F(u) \cdot G(u) .
\end{aligned}$$

2.2.4. Space and time

In addition to providing the relationship between spatial distributions $f(r)$ and the diffraction amplitudes, $F(u)$ the Fourier transform also relates the variation of a function in time, $f(t)$, and the corresponding frequency distribution. Thus we may write

$$F(v) = \int_{-\infty}^{\infty} f(t) \exp\{2\pi i vt\} dt, \quad (31)$$

and

$$f(t) = \int_{-\infty}^{\infty} F(v) \exp\{-2\pi i vt\} dv, \quad (32)$$

where we use the frequency v , rather than the angular frequency ω . To make the analogy complete it is necessary to introduce the artifice of negative frequencies. While this is inconsistent with common parlance, it is readily seen that a negative frequency can correspond to a negative progression of the phase of a wave in time, i.e. to a wave going "backwards".

For a function of both space and time, $f(r, t)$, a Fourier transform may be made with respect to any one or all of the coordinates or with respect to time, or with respect to both spatial coordinates and time. Hence where any ambiguity is possible it is necessary to specify the variable or variables involved in the Fourier transform, possibly by use of a subscript on \mathfrak{F} . For example,

$$\begin{aligned}
\mathfrak{F}_{x,t} [f(x, y, z, t)] &= F(u, y, z, v) \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y, z, t) \exp\{2\pi i (u x + v t)\} dx dt .
\end{aligned}$$

2.3. Fourier transforms and diffraction: examples

We now provide a series of examples to serve the two functions of familiarizing the reader with common Fourier transforms and demonstrating the application of Fourier transforms to kinematical diffraction. We refer for the most part to diffraction from simple one- or two-dimensional objects.

2.3.1. Point source or point aperture

The amplitude distribution of a very small source or the transmission through a very small aperture (or slit) in one dimension may be described as $\delta(x)$, or by $\delta(x - a)$ when it is not at the origin. The Fourier transform used to derive the Fraunhofer diffraction pattern is

$$\begin{aligned}\mathfrak{F} \delta(x) &= 1, \\ \mathfrak{F} \delta(x - a) &= \exp\{2\pi i u a\}.\end{aligned}\tag{33}$$

To show this we write the integral

$$\int_{-\infty}^{\infty} \delta(x - a) \exp\{2\pi i u x\} dx.$$

The integrand is zero except for $x = a$. Hence the integral may be written

$$\exp\{2\pi i u a\} \int_{-\infty}^{\infty} \delta(x - a) dx = \exp\{2\pi i u a\}.$$

The amplitude of a diffraction pattern is proportional to $F(u) = \mathfrak{F} \delta(x - a)$ where $u = l/\lambda$. The intensity observed is then proportional to $|F(u)|^2 = 1$.

Thus, as is well known, the Fraunhofer diffraction pattern from a point source has uniform intensity, apart from the factor $1/R^2$ and the obliquity factor (if applicable) which have been omitted.

2.3.2. A plane wave: the inverse of 2.3.1

Fourier transform of a plane wave (equation 1.8) with respect to t only gives

$$\mathfrak{F}_t \exp\{2\pi i (v_1 t - x/\lambda_1)\} = \delta(v + v_1) \exp\{-2\pi i x/\lambda_1\},\tag{34}$$

i.e. a delta function in frequency and a plane wave in real space. Fourier transform with respect to x gives

$$\mathfrak{F}_x \exp\{2\pi i(v_1 t - x/\lambda_1)\} = \exp\{2\pi i v_1 t\} \delta(u - 1/\lambda_1), \quad (35)$$

i.e. a sinusoidal variation in time and a delta function in reciprocal space. Transforming with respect to both variables gives,

$$\mathfrak{F}_{x,t} \exp\{2\pi i(v_1 t - x/\lambda_1)\} = \delta(v + v_1) \cdot \delta(u - 1/\lambda_1), \quad (36)$$

which is a delta function in both frequency and reciprocal space.

If we take the Fourier transform of all possible waves which can exist in a given medium with respect to both x and t we obtain a set of points defining the relationship between the frequency ν and λ^{-1} (or between the angular frequency ω and $2\pi k$) which is known as the dispersion relation for that type of wave and the particular medium.

2.3.3. Translation of an object

$$\mathfrak{F}f(x - a) = \mathfrak{F}[f(x) * \delta(x - a)] = F(u) \exp\{2\pi i u a\}, \quad (37)$$

where we have used (33) and the convolution theorem (30).

Thus translation of the object in real space has the effect of multiplying the amplitude in reciprocal space by a complex exponential. The intensity distribution of the Fraunhofer diffraction pattern is given by $|F(u)|^2$, which is independent of the translation.

2.3.4. Slit function

The transmission function of a slit of width a in an opaque screen is given by

$$g(X) = \begin{cases} 0 & \text{if } |x| > a/2, \\ 1 & \text{if } |x| \leq a/2. \end{cases}$$

The Fourier transform of this is

$$F(u) = \int_{-a/2}^{a/2} \exp\{2\pi i u x\} dx = \sin(\pi a u) / (\pi u). \quad (38)$$

Putting $u = l/\lambda$, we obtain the well-known form of the diffraction pattern

$$F(l) = a \sin(\pi l a / \lambda) / (\pi l a / \lambda), \quad (39)$$

with an intensity distribution

$$I(l) = a^2 \sin^2(\pi la/\lambda) / (\pi la/\lambda)^2, \tag{40}$$

which has a central maximum value a^2 for $l = 0$, and has decreasing subsidiary maxima with increasing $|l|$ with zero values for $l = n\lambda/a$ for $n \neq 0$ (see Fig. 2.2).

2.3.5. Slit function-alternative

To illustrate the application of (27), we note that for the slit function defined in 2.3.4,

$$g(x) \equiv dq(x)/dx = \delta(x + \frac{1}{2}a) - \delta(x - \frac{1}{2}a),$$

as suggested in Fig. 2.2. Then

$$G(u) = \exp\{-\pi i a u\} - \exp\{\pi i a u\} = -2i \sin(\pi a u).$$

But, since from (27)

$$G(u) = \{-2\pi i u\} F(u),$$

it follows that, as in equation (38),

$$F(u) = \sin(\pi a u) / (\pi u).$$

2.3.6. Straight edge

The transmission function is

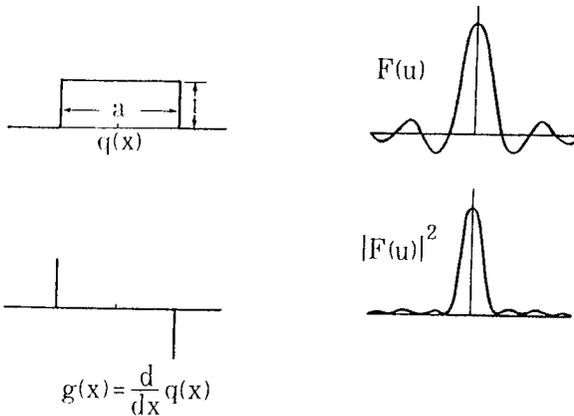


Fig. 2.2. Derivation of the Fourier transform of a slit function by consideration of its differential.

$$f(x) = \begin{cases} 0 & \text{if } x < 0, \\ 1 & \text{if } x \geq 0. \end{cases}$$

Using the same procedure as in 2.3.5 we put

$$g(x) = df(x)/dx = \delta(x)$$

Then

$$G(u) = 1$$

and

$$F(u) = (2\pi i u)^{-1}.$$

But an indeterminate constant term has been omitted in the integration. We note that

$$\int_{-\infty}^{\infty} (f(x) - \frac{1}{2}) dx = 0.$$

This suggests that the missing constant term is $1/2$. Inserting this gives the correct result,

$$F(u) = +\frac{1}{2}\delta(u) - (2\pi i u)^{-1}. \quad (41)$$

2.3.7. Rectangular aperture

In the two dimensional form of 2.3.4 we define the transmission function of a rectangular aperture as

$$f(x, y) = \begin{cases} 1 & \text{if } |x| < a/2 \text{ and } |y| < b/2, \\ 0 & \text{elsewhere.} \end{cases}$$

Then

$$\begin{aligned} F(u, v) &= \int_{-a/2}^{a/2} \exp\{2\pi i u x\} dx \int_{-b/2}^{b/2} \exp\{2\pi i v y\} dy \\ &= ab \sin(\pi a u) / (\pi a u) \sin(\pi b v) / (\pi b v), \end{aligned} \quad (42)$$

so that, for diffraction from a rectangular aperture the intensity distribution is

$$I(u, v) = a^2 b^2 \sin^2(\pi a u) / (\pi a u)^2 \sin^2(\pi b v) / (\pi b v)^2. \quad (43)$$

The maximum intensity at $u = v = 0$ is $a^2 b^2$. Zeros occur at intervals of a^{-1} along the u direction, parallel to the x axis and b^{-1} along the v direction, as suggested in Fig. 2.3. Thus

the intensity function falls off through decreasing oscillations in each direction. The dimensions of the distribution in reciprocal space are inversely proportional to the dimensions in real space. Integrating over the whole peak gives an integrated intensity equal to ab , the area of the aperture, as expected.

2.3.8. Circular aperture

The transmission function for a circular aperture is written

$$f(x,y) = \begin{cases} 1 & \text{if } (x^2+y^2)^{1/2} < a/2, \\ 0 & \text{elsewhere.} \end{cases}$$

The Fourier transform is best carried out by writing the Fourier integral in plane polar coordinates. The result is

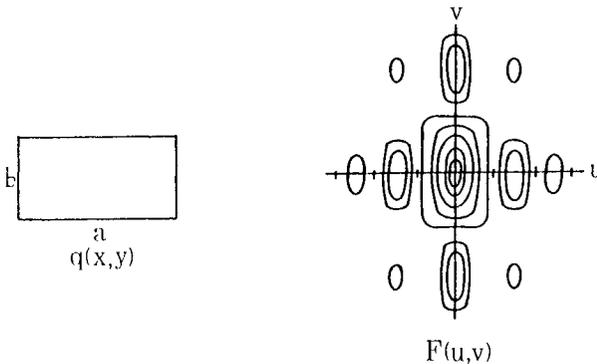


Fig. 2.3. Diagram suggesting the form of the Fourier transform of a rectangular aperture, equation (2.43).

$$F(u) = (\pi a^2/2) J_1(\pi a u) / (\pi a u), \quad (44)$$

where u is a radial coordinate and $J_1(x)$ is the first order Bessel function. The function $J_1(x)/x$ is similar in form to $(\sin x)/x$ but has a somewhat broader central maximum with the first zero at $1.22a^{-1}$ instead of a^{-1} .

2.3.9. Two very narrow slits

For two slits a distance A apart we take the origin half-way between them and write

$$f(x) = s(x+A/2) + s(x-A/2),$$

where $s(x)$ is the transmission function for one slit of width a as in 2.3.4. For very narrow slits we let a go to zero. But to keep the intensity finite we imagine the incident amplitude to be proportional to $1/a$. Then $s(x)$ becomes $\delta(x)$. We put

$$f(x) = \delta(x+A/2) + \delta(x-A/2),$$

and

$$\begin{aligned} F(u) &= \exp\{-\pi iAu\} + \exp\{\pi iAu\} \\ &= 2 \cos(\pi Au), \end{aligned} \tag{45}$$

and the diffraction pattern intensity is

$$I(l) = 4 \cos^2(\pi Al/\lambda). \tag{46}$$

Thus we get simple sinusoidal fringes of uniform amplitude. Comparison with the Fresnel diffraction pattern shows that for this particular object the diffraction pattern is independent of the approximation made to the general Kirchoff integral (see Problem 2).

2.3.10. Two slits of appreciable width

For two slits of width a , a distance A apart, we may write

$$f(x) = s(x) * [\delta(x+A/2) + \delta(x-A/2)],$$

where $s(x)$ is the transmission function for a single slit defined as in 2.3.4. Using the convolution theorem and the results (39) and (45)

$$F(u) = 2a \cos(\pi Au) \cdot \sin(\pi au) / (\pi au). \tag{47}$$

The intensity of the diffraction pattern is then as shown in Fig. 2.4; \cos^2 fringes of period $1/2A$ are modulated by a $(\sin^2 x)/x^2$ function which goes to zero for $u = a^{-1}$. This is a description of the result of the Young's fringe experiment, well known as of fundamental significance for the development of physical optics.

2.3.11. Finite wave train

The inverse of 2.3.9 gives

$$\mathfrak{I}\{2 \cos \pi Ax\} = \delta(u+A/2) + \delta(u-A/2).$$

If the wave train represented by the cosine function is cut to a finite length by multiplying it by a slit function of width B we have

$$f(x) = 2 \cos(\pi Ax) s_B(x),$$

so that, using the multiplication theorem,

$$F(u) = \{ \delta(u+A/2) + \delta(u-A/2) \} * B \sin(\pi Bu) / (\pi Bu). \tag{48}$$

Thus the amplitude distribution in reciprocal space is the sum of two functions of the form $S(u) = \Im s_B(x)$ centered on $u = \pm A/2$ (Fig. 2.5). If B is much greater than the periodicity $2/A$ of the wave train, the two peaks in $F(u)$ are much narrower than their separation and they do not overlap appreciably: Then the intensity distribution $|F(u)|^2$ is very nearly given by

$$I(u) = B^2 \sin^2(\pi Bu) / (\pi Bu)^2 * \{ \delta(u+A/2) + \delta(u-A/2) \}.$$

This is a useful approximation for many purposes, but it can not be used if the length of the wave train B is only a few times the periodicity.

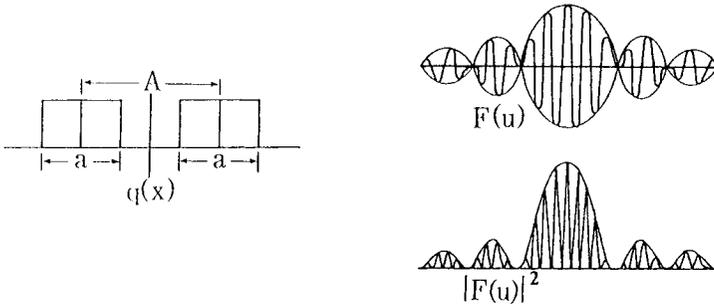


Fig. 2.4. Diagram suggesting the diffraction pattern from two parallel slits, equation (2.47).

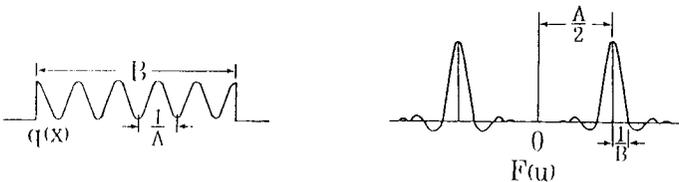


Fig. 2.5. A wave train cut off by multiplying by a slit function and its Fourier transform, equation (2.48).

2.3.12. *Periodic array of narrow slits*

We assume a periodic array of slits having zero width and repeated at regular intervals, a indefinitely. Then

$$f(x) = \sum_{n=-\infty}^{\infty} \delta(x - na).$$

The Fourier transform is, using (33),

$$F(u) = \sum_{-\infty}^{\infty} \exp\{2\pi i u n a\}.$$

This summation of a Fourier series with all coefficients unity gives a well-known result which we can derive by elementary methods as follows.

Since

$$\sum_0^{\infty} x^n = (1 - x)^{-1},$$

we may write

$$\begin{aligned} F(u) &= \sum_0^{\infty} [\exp\{2\pi i u a\}]^n + \sum_0^{\infty} [\exp\{-2\pi i u a\}]^n - 1 \\ &= [1 - \exp\{2\pi i u a\}]^{-1} + [1 - \exp\{-2\pi i u a\}]^{-1} - 1 \\ &= 0, \end{aligned}$$

except that

$$F(u) = \infty \quad \text{for} \quad \exp\{2\pi i u a\} = 1,$$

i.e. if $2\pi u a = 2h\pi$ where h is an integer, or $u = h/a$. Then

$$F(u) = a^{-1} \sum_h \delta(u - h/a), \quad (49)$$

where the factor a^{-1} gives the delta functions the correct weight.

Hence the Fourier transform is a set of equally spaced delta functions of period a^{-1} in reciprocal space.

2.3.13. *Arbitrary periodic function*

For an object with an arbitrary periodic transmission function we write

$$f(x) = \sum_{-\infty}^{\infty} F_h \exp\{-2\pi i h x/a\}. \quad (50)$$

Then

$$F(u) = \sum_{-\infty}^{\infty} F_h \int_{-\infty}^{\infty} \exp\left\{2\pi i \left(-\frac{hx}{a} + ux\right)\right\} dx.$$

From (7) this is

$$F(u) = \sum_{-\infty}^{\infty} F_h \delta(u - h/a). \quad (51)$$

Hence the diffracted amplitude is represented by a set of delta functions equally spaced with separation a^{-1} in u , each delta function having the "weight" F_h equal to the corresponding Fourier coefficient of (50). This result and those which follow form the basis for much of our consideration of diffraction of X-rays and electrons by crystals.

2.3.14. *Diffraction grating: thin slits*

A set of N parallel, equally spaced thin slits constitutes a primitive form of diffraction grating. The transmission function may be written;

$$f(x) = \sum_{-(N-1)/2}^{(N-1)/2} \delta(x - na). \quad (52)$$

Then

$$\begin{aligned} F(u) &= \sum_{-(N-1)/2}^{(N-1)/2} \exp\{2\pi i u n a\} \\ &= \exp\{-\pi i u (N - 1) a\} \sum_0^{N-1} \exp\{2\pi i u n a\} \\ &= \exp\{-\pi i u (N - 1) a\} \frac{\exp[2\pi i u N a] - 1}{\exp[2\pi i u a] - 1}, \end{aligned}$$

i.e.

$$F(u) = \sin(\pi N a u) / \sin(\pi a u). \quad (53)$$

On the other hand, we may use the result of (49) and write the transmission function as

$$f(x) = s(x) \sum_{n=-\infty}^{\infty} \delta(x - na). \quad (54)$$

where $s(x)$ is a slit function of width Na , which cuts off the transmission of all but N of the thin slits. Then, as in (48) we may write

$$F(u) = \sum_h \delta(u - h/a) * Na \sin(\pi Nau) / (\pi Nau). \quad (55)$$

The functions (53) and (55) are similar (see Fig. 2.6). For each there are sharp peaks with side ripples, of the form $(\sin x)/x$, at intervals $u = a^{-1}$. The width of the peak, the distance from the maximum to the first zero, is $1/Na$.

It is not obvious from the form of the functions (53) and (55) that they are identical, but since the functions in real space, (52) and (54) are identical, the transforms must be also.

2.3.15. Diffraction grating: general

The individual lines of a diffraction grating, used in transmission, may be slits of finite width, rulings on glass, replicas in plastic or rulings on metal and so on. The transmission function of the whole grating is given by the convolution of the transmission function of a single ruling $g(x)$ with a set of N delta functions, (52) or (54)

$$f(x) = \sum_{-(N-1)/2}^{(N-1)/2} \delta(x - na) * g(x) \quad (56)$$

so that

$$F(u) = G(u) \sin(\pi Nau) / \sin(\pi au), \quad (57)$$

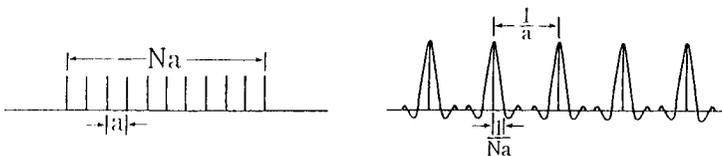


Fig. 2.6. The amplitude distribution for diffraction from a large number of parallel very fine slits, as given by equation (2.53) or (2.55).

or

$$F(u) = G(u) \left[\sum_h \delta(u - h/a) * Na \sin(\pi Nau) / (\pi Nau) \right]. \quad (58)$$

Thus the height of each of the diffraction maxima is proportional to the value of the Fourier transform of $g(x)$ at that u value.

Alternatively we could say that the diffraction grating may be represented by cutting off a periodic transmission function by multiplying it with a slit function $s(x)$ of width Na . This is not the same as (56) since in this case we may be using $s(x)$ to cut off a continuous function. Correspondingly, the diffraction amplitude, written from (51) as

$$F(u) = \sum_h F_h \delta(u - h/a) * Na \sin(\pi Nau) / (\pi Nau). \quad (59)$$

is not exactly the same as (57) or (58).

2.3.16. Gaussian function

A Gaussian function rarely occurs in practice as the transmission function of an object but is frequently used as an approximation, to "round-off" a discontinuity, or to achieve a suitable convergence of an analytic function, because of its desirable properties in relation to Fourier transform and convolution.

Putting

$$f(x) = \exp\{-a^2x^2\}, \quad (60)$$

we make use of the standard integral (1.28) to obtain

$$F(u) = (\pi^{1/2}/a) \exp\{-\pi^2u^2/a^2\}. \quad (61)$$

Hence, if $f(x)$ is a Gaussian of half-width a^{-1} then its Fourier transform is also a Gaussian having half-width a/π .

We note also that, if

$$f(x) = \exp(-a^2x^2) * \exp\{-b^2x^2\},$$

$$F(u) = (\pi/ab) \exp\{-\pi^2u^2(a^2+b^2)/(a^2b^2)\},$$

and, applying the inverse Fourier transform

$$f(x) = (\pi/(a^2+b^2))^{1/2} \exp\{-(a^2b^2/(a^2+b^2))x^2\}, \quad (62)$$

i.e. the convolution of two Gaussians is again a Gaussian.

2.3.17. Row of circular holes

The extension to two dimensions allows the above results to be applied to calculate the diffraction patterns from many simple two-dimensional distributions. However, in using delta functions and convolutions it is important to distinguish carefully the delta functions, and convolutions in one and two dimensions.

We take as example the case of a linear row of equally spaced circular holes in an opaque screen. For this the transmission function may be written

$$f(x, y) = \sum_n \delta(x - na, y) * O(x, y) \quad (63)$$

where we have used $\delta(x - na, y)$ to indicate a delta function at $x = na, y = 0$, and $O(x, y)$ is the transmission function for a circular aperture as defined in 2.3.8, above. See Fig. 2.7.

The Fourier transformation of the set of delta functions with respect to x is

$$\sum_h \delta(u - h/a)$$

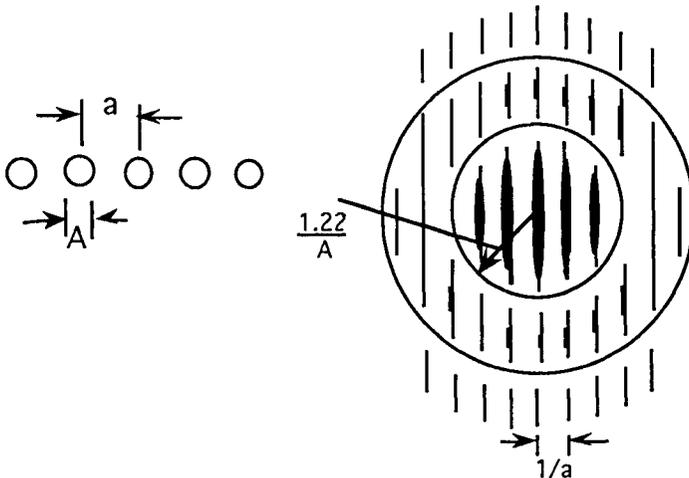


Fig. 2.7. The diffraction pattern from a row of circular holes equally spaced in an opaque screen. The widths of the lines are made to suggest their intensities.

and the Fourier transform with respect to y is unity. Hence the Fourier transform in x and y is a set of straight lines parallel to the v axis and equally spaced at intervals of a^{-1} in the u direction. This distribution will be modulated by multiplying by the Fourier transform of $O(x,y)$, namely

$$(\pi A^2/2) J_1(\pi AU) / (\pi AU),$$

where A is the diameter of the holes and $U = (u^2+v^2)^{1/2}$.

This gives the set of lines of varying intensity as suggested by Fig. 2.7 where the width of the line has been used as an indication of relative intensity.

This example has some relevance in radio astronomy. A common form of radio interferometer used for detecting the radio waves from distant radio stars, consists of an equally spaced line of paraboloidal antennas or "dishes" each of which sums the incident amplitude over the circular aperture of the dish. It can be seen, by application of the reciprocity theorem that the amplitude obtained by summing the amplitudes from all dishes coherently is exactly the same as the amplitude which would be observed at the distant radio star if the apertures of the dishes were to be illuminated from behind by a plane parallel incident wave. Thus, as a radio star moves across the sky, the amplitude measured by the interferometer corresponds to a linear trace across the diagram of the diffraction pattern in Fig. 2.7.

2.3.18. *Complementary objects - Babinet's Principle*

Consider two objects for which the transmission functions are complementary so that

$$q_1(x,y) + q_2(x,y) = 1. \tag{64}$$

For black and transparent objects, this means that the black areas of the first are the transparent parts of the second and vice versa.

Then, for the second object

$$q_2(x,y) = 1 - q_1(x,y).$$

The diffraction pattern amplitude is

$$Q_2(u,v) = \delta(u,v) - Q_1(u,v),$$

so that the intensity in the diffraction pattern is proportional to $|Q_2(u,v)|^2$ which is equal to $|Q_1(u,v)|^2$ except at the origin.

This exception, the exclusion of the origin point, was overlooked in the original formulation, known as Babinet's Principle or Babinet's Theorem.

2.3.19. Total intensities: Parseval's Theorem

In a diffraction experiment, conservation of energy implies that the total intensity in the diffraction pattern is equal to the total intensity at the exit surface of the object, i.e.

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |q(x, y)|^2 dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |Q(u, v)|^2 du dv. \quad (65)$$

This is a special case of a general relationship of Fourier transform theory

$$\int |f(r)|^2 dr = \int |F(u)|^2 du, \quad (66)$$

which in turn is a special case of the more general Parseval's Theorem relating to any two functions $f(r)$ and $g(r)$;

$$\int f(r) g^*(r) dr = \int F(u) G^*(u) du. \quad (67)$$

From this theorem we see that

$$\int f(R) f(r+R) dR = \int |F(u)|^2 \exp\{2\pi i u \cdot r\} du \quad (68)$$

which is a special case of the inverse of the multiplication theorem (29);

$$\mathfrak{I}[F(u)F^*(u)] = f(r) * f(-r). \quad (69)$$

This convolution of a function with the same function inverted is the autocorrelation function which has considerable application in many fields of science and, as we will see later, has special significance for diffraction under the title of a "generalized Patterson function".

Problems

1. Write down an expression for, and sketch the form of, the diffraction pattern given when an incident plane wave is diffracted by:

- (1) a row of N circular holes, equally spaced in an opaque screen,
- (2) a similar row of opaque circular discs,
- (3) a similar row of holes in an opaque screen which are alternately circular and square,

- (4) a row of circular holes in an opaque screen when each second one is covered by a half-wave plate (phase change π),
(5) two parallel rows of circular holes.

2. Find the Fresnel diffraction pattern given by a pair of parallel, very fine slits. Compare this with the Fraunhofer diffraction pattern.

3. A radio-interferometer consists of two perpendicular, intersecting lines of regularly spaced circular dishes (paraboloidal antennas) forming an equi-armed cross (known as a "Chris-cross"). If the signals from all dishes are added in phase, how does the response of the interferometer vary with the position of a radio source in the sky? By what manipulation of the electronic signals could the interferometer be "pointed" to receive maximum radiation from a particular direction?

4. Compare the response of the radio interferometer described in 3 with that of a complete two-dimensional square array of $N \times N$ paraboloidal antennas.

5. In a Young's fringe experiment, two fine parallel slits, separation a , are illuminated by monochromatic light from a distant, parallel slit of width d , which may be regarded as an incoherent source, so that intensities due to each point of the source are added on the plane of observation. Find the visibility (defined as $(I_{\max} - I_{\min}) / (I_{\max} + I_{\min})$) of the interference fringes produced as a function of a and d . Taking the visibility of the fringes as a measure of the degree of coherence of the illumination at the two slits, verify the Zernike-van Cittert theorem which states that the degree of coherence is given by the Fourier transform of the intensity distribution of the source.

This Page Intentionally Left Blank

Imaging and diffraction

3.1. Wave theory of imaging

3.1.1. Coherent wave optics

While, as we shall see, lenses may be used to obtain either diffraction patterns or images, the major development of the theory of lens action has been directed towards imaging. We give some brief account of this development here but prefer a somewhat different approach which allows the diffraction patterns and images to be treated with equal facility and stresses the relationship between them. This has considerable practical importance, especially in electron microscopy.

For many years, imaging was considered in terms of the approximations and from the point of view of geometric optics. The wave nature of light was taken into account only as a refinement when resolution was affected by diffraction effects due to finite aperture size. The formulation of imaging theory completely in terms of waves was introduced by Duffieux [1946] and subsequently developed by many authors including Hopkins [1953, 1950], Fellgett and Linfoot [1955] and Linfoot [1955]. An outline of the development has been given by Born and Wolf [1989]. This approach to imaging theory is based essentially on the Fourier transform. Here we describe first of all a more graphic version of the theory and then a more formal and elegant form.

We consider first the system suggested by Fig. 3.1. Incident radiation, passing through a small object or a small part of an object gives an amplitude distribution which, over a limited region in one dimension, is given by the function $f(x)$. Fraunhofer diffraction then gives rise to a distribution $F(u)$ on a spherical reference surface at the entrance to a lens system. The coordinate u measured along the spherical surface is φ/λ , where φ is the scattering angle, or s/R where s is the distance measured on the sphere.

The function of the lens is to transfer the distribution $F(u)$ into the distribution $F'(u')$ on a spherical reference surface centered on the part of the image, $\Psi'(x)$, corresponding to $f(x)$. Following a common convention, quantities

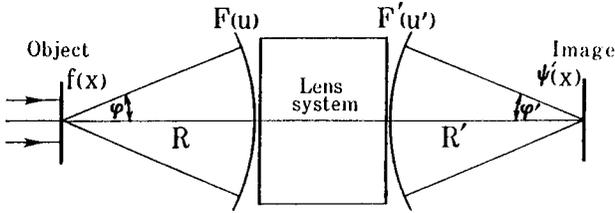


Fig. 3.1. Diagram suggesting the imaging process in an optical system. The complex amplitude distribution on a spherical reference surface in object space is converted into an amplitude distribution on a spherical reference surface in image space.

referring to the image space are primed. If the transfer of F to F' were perfect, with no perturbation of the function or aperture limitation, we would have $F'(s) = F(s)$. From the geometry of the situation,

$$s = uR = u'R'$$

so that, for this case,

$$F'(u) = F(u'R'/R). \quad (1)$$

The distribution $F(u)$ is $\mathfrak{S}f(x)$. By symmetry we put $F'(u') = \mathfrak{S}\psi'(x)$. The transfer from F' to ψ' might be thought of as an inverse transform but since the direction of propagation from ψ' to F' is opposite to that in object space, the sign of the exponent is changed and a positive sign in the exponential of the transform of F' to ψ' makes it a direct Fourier transform.

Hence

$$\begin{aligned} \psi'(x) &= \mathfrak{S} F(R'u'/R) \\ &= \int \int f(X) \exp\{2\pi i (u'R'/R)X\} dX \exp\{2\pi i u'x\} du' \\ &= \int f(X) \delta(X + Rx/R') dX = f(-Rx/R'), \end{aligned} \quad (2)$$

i.e., the image is a recreation of $f(x)$, inverted and magnified by a factor R'/R .

In practice, of course, the transfer from $F(u)$ to $F'(u')$ is not perfect. There is always a limitation due to the finite aperture size. In addition, the lens aberrations give rise to phase changes which vary as a function of u' . Classically, this phase change is described by a power series in u' or, in two dimensions, by a power series in ρ (proportional to $(u^2 + v^2)^{1/2}$) and the polar angle ϕ . In this way the traditional aberration coefficients are generated, the most important for our purposes being the Third Order Spherical Aberration coefficient which derives from the coefficient of the ρ^4 term.

In the wave-theory treatment, the changes of amplitude and phase due to the limitations of the lens are represented by a Transfer Function $T(u')$, characteristic of the lens. Then

$$F'(u') = F(R'u'/R) T(u'). \quad (3)$$

Correspondingly the modification of the image is represented by

$$\psi'(x) = f(-Rx/R') * t(x), \quad (4)$$

where $t(x)$, the Fourier transform of $T(u)$, is the spread function. For the simplest case, in two dimensions, of limitation by a circular aperture only, we have that, as in (2.44),

$$T(u', v') = \begin{cases} 1 & \text{if } (u'^2 + v'^2)^{1/2} < u_0/2 \\ 0 & \text{elsewhere.} \end{cases}$$

so that

$$t(x, y) \equiv t(r) = J_1(\pi u_0 r) / u_0 r,$$

and the intensity distribution in the image is

$$I'(x, y) = \psi\psi^* = \left| f\left(\frac{-R}{R_1}x, \frac{-R}{R_1}y\right) * J_1(\pi u_0 r) / (\pi r) \right|^2, \quad (5)$$

which represents a loss of resolution due to a blurring of the amplitude distribution.

When, in addition to the aperture limitation, the aberrations of the lens are appreciable, the transfer function includes the phase change term so that (3) may be written

$$F'(u') = F(R'u'/R) T_a(u') \exp\{i\Phi(u')\},$$

where $T_a(u')$ is the aperture function and $\Phi(u)$ represents the change of phase, which may be expanded as a power series in ρ and ϕ to introduce the aberration coefficients. Then the observed intensity is given by

$$I'(x, y) = \left| f\left(\frac{-R}{R_1}x, \frac{-R}{R_1}y\right) * J_1(\pi u_0 r) / (\pi r) * \mathfrak{I}[\exp\{i\Phi(u)\}] \right|^2.$$

The effects of the added convolution here are not immediately apparent. They may be calculated in detail for particular cases but a simple intuitive understanding is difficult. We will discuss these effects later in terms of the Abbe treatment.

3.1.2. Incoherent wave imaging

The same sort of treatment applies if we consider incoherent imaging, i.e., imaging of an object which is self-luminous or which is illuminated by incoherent incident light or scatters the incident light incoherently. Then we consider the imaging of light from each point of the object separately and add the resulting intensities.

Referring to (4), the amplitude distribution for a point source on the axis of the lens is $t(x)$ and the intensity distribution is $|t(x)|^2$. Then the image intensity is given by

$$I'(x, y) = I((-R/R_1)x, (-R/R_1)y) * |t(x)|^2. \quad (6)$$

The spread function $|t(x)|^2$ can be referred to a Contrast Transfer Function which characterizes the lens;

$$O(u') = T(u') * T^*(-u'). \quad (7)$$

For the simple case of a circular aperture, the image intensity is convoluted by the spread function

$$|t(x, y)|^2 = \frac{J_1^2(\pi u_0 r)}{(\pi r)^2}.$$

If we then apply the Rayleigh criterion for the resolution of two adjacent point sources, that the images can be resolved if the maximum of one point image falls at the first minimum of the other, i.e., at $r = 1.22/u_0$, we obtain the least resolvable distance for the lens as

$$\Delta x = 1.22\lambda/2\phi_0 = 1.22\lambda/\alpha \quad (8)$$

where α is the angular aperture subtended by the lens at the object. For convenience in considerations of microscopy we have transferred the dimensions back into object space.

The more formal description of the imaging process eliminates the need for the picture of a transfer from one reference surface to another as suggested in Fig. 3.1. The spread function, $t(x, y)$ for amplitudes for coherent imaging, or $|t(x, y)|^2$ for the incoherent case, is defined as the response to a point source in the object. The Fourier transform, $T(u, v)$ or $O(u, v) = T(u, v) * T^*(-u, -v)$, is the appropriate transfer function which is characteristic of the lens. This contrast transfer function may be derived, independently of any aberration theory, by studying the image intensity distribution for particular test objects; for example, the contrast of the images of objects with transmission function $1 + \cos(2\pi ax)$ give the amplitudes and

phases of the transfer for particular values of the "spatial frequency", a . From the contrast transfer function the spread function is derived by Fourier transform.

3.2. Abbe theory

For our purposes, the formal use of Fourier transforms in imaging theory is interesting but limited, in the form described in the previous section, since we are interested in the concrete embodiment of the Fourier transform in the form of the diffraction pattern as well as in the image. We turn therefore to the imaging theory originating from the ideas of Abbe. Although this is essentially a wave theory, it is conveniently illustrated by reference to the geometric-optics diagram of Fig. 3.2.

Parallel light is incident on an object having transmission function $f(x,y)$ which is imaged by a lens having focal length, f . Light scattered by a point A of the object is brought to a focus at A' in the image plane. From the geometry of the diagram the image is inverted and magnified by a factor R'/R so that the image is

$$\psi(x,y) = f((-R/R_1)x, (-R/R_1)y) .$$

Also it may be noted that all light scattered through an angle ϕ is brought to a focus at one point in the back-focal plane. This is equivalent to interference at a point at infinity. Hence the amplitude distribution on the back-focal plane is that of the Fraunhofer diffraction pattern given by

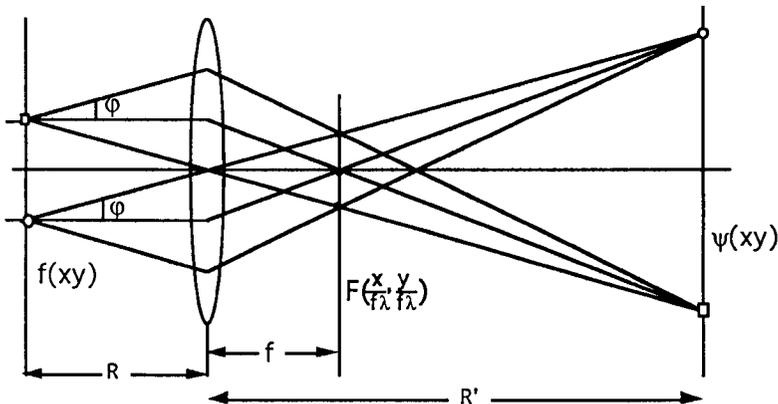


Fig. 3.2. Geometric optics ray diagram used to suggest the plausibility of the wave-optics description of the imaging process in the Abbe theory.

the Fourier transform function, $F(u, v)$. In this case $u = (\sin\phi_x)/\lambda$ and if ϕ is not too large we may write $u = x/f\lambda$, $v = y/f\lambda$. Thus the imaging process may be described in terms of two Fourier transforms: The scattered radiation from the object interferes on the back-focal plane to give the Fraunhofer diffraction pattern described by a Fourier transform; then the radiation from the back-focal plane again forms an interference pattern on the image plane which is effectively at infinity so that the amplitude distribution in the image is given by Fourier transform of that in the back-focal plane.

Two immediate consequences are of interest. The first is that lenses may be used for light or electrons to give the Fraunhofer diffraction pattern of an object at a convenient location and on a convenient scale, depending on the focal length of the lens. Secondly, the limitations of the lens system in reproducing the object transmission function on the image plane may be described in terms of modifications of the amplitude and phase of the distribution on the back-focal plane.

Apart from their direct interest for consideration of optical and electron-optical imaging and diffraction, these points are of importance for X-ray diffraction because of the very graphic use made of the optical diffractometer by Taylor and Lipson [1964] to simulate the production of X-ray diffraction patterns and the reconstruction of crystal structures from diffraction amplitudes.

3.3. Small angle approximation

Although it is strictly limited in its validity, the small angle approximation which we introduced for the discussion of Fresnel Diffraction forms a very convenient basis for describing the essential behavior of imaging systems. It provides a model which reproduces all important features of the properties of optical systems with a relative mathematical simplicity and wide versatility.

All objects are considered to be composed of planar distributions having transmission functions $q_n(x, y)$. Propagation through a medium of constant refractive index is given by convolution with a propagation function which, in the small angle approximation, is $(i/R\lambda) \exp\{-\pi i(x^2+y^2)/R\lambda\}$. We introduce the concept of an ideal thin lens, which is a planar object having a transmission function $\exp\{\pi i(x^2+y^2)/f\lambda\}$. It is readily confirmed that

$$\exp\{\pi i(x^2+y^2)/f\lambda\} * \exp\{-\pi i(x^2+y^2)/f\lambda\} = \delta(x, y) \quad (9)$$

i.e. if a plane wave, amplitude unity, passes through the ideal thin lens, propagation of the wave through the focal length f gives a delta-function or a point cross-over. Similarly a point source placed a distance f before the ideal thin lens gives a plane wave:

$$\left[\delta(x,y) * \exp\left\{\frac{-\pi i(x^2 + y^2)}{f\lambda}\right\} \right] \exp\left\{\frac{\pi i(x^2 + y^2)}{f\lambda}\right\} = 1. \tag{10}$$

In this approximation the amplitude given on a plane of observation when a plane wave passes through an object of transmission function $q(x,y)$ and then an ideal thin lens, as suggested in Fig. 3.3, is, in one dimension and omitting constant multipliers,

$$\psi(x) = \left[\left[\left[q(x) * \exp\left\{\frac{-\pi i x^2}{R\lambda}\right\} \right] \exp\left\{\frac{\pi i x^2}{f\lambda}\right\} \right] * \exp\left\{\frac{-\pi i x^2}{R'\lambda}\right\} \right], \tag{11}$$

where the operations in the successive brackets represent propagation through a distance R , transmission through the lens, and propagation through a distance R' . By writing out the convolution integrals in detail it is readily shown that if $R' = f$, $\psi(x) = Q(x/f\lambda)$, and if $(1/R) + (1/R') = 1/f$, then $\psi(x) = q(-Rx/R')$. To prove the second case, (11) is written in detail, with dummy variables X and Y ,

$$\begin{aligned} \psi(x) = & \iint q(Y) \exp\left\{\frac{-\pi i Y^2}{R\lambda}\right\} \exp\left\{\frac{-\pi i}{\lambda} \left(\frac{1}{R} + \frac{1}{R'} - \frac{1}{f}\right) X^2\right\} \\ & \times \exp\left\{\frac{-\pi i x^2}{R'\lambda}\right\} \exp\left\{\frac{\pi i}{\lambda} \left(\frac{Y}{R} + \frac{x}{R'}\right) X\right\} dx dy. \end{aligned}$$

Then if $(1/R) + (1/R') = 1/f$ the integral over X is seen to be $\delta(Y+Rx/R')$ and the integral over Y then gives $\psi(x) = Cq(-Rx/R')$,

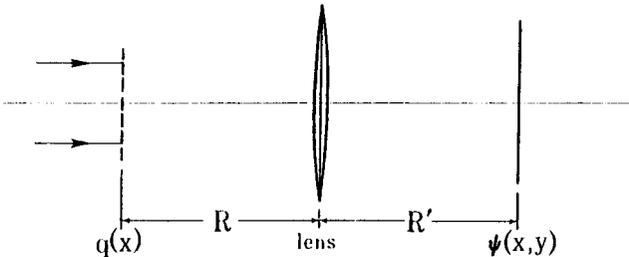


Fig. 3.3. Diagram defining the imaging system described by equation (3.11).

where C has modulus unity and may be absorbed with other similar terms omitted from the propagation functions.

The proof of the first case, for $R' = f$, is left as an exercise for the reader.

Thus the properties of producing diffraction patterns and images are reproduced. Obviously the action of any combination of sources, object and lenses may be reproduced by writing down the appropriate series of operations of convolution with a propagation function and multiplication with a transmission function. For example, if there is a point source at $x = X$ a distance R_0 before the object of Fig. 3.3, the amplitude distribution on the plane of observation is

$$\begin{aligned} \psi(x) = & \left[\left[\left[\delta(x-X) * \exp\left\{\frac{-\pi i x^2}{R_0 \lambda}\right\} \right] q_0(x) \right] * \exp\left\{\frac{-\pi i x^2}{R \lambda}\right\} \right] \\ & \times \exp\left\{\frac{\pi i x^2}{f \lambda}\right\} \right] * \exp\left\{\frac{-\pi i x^2}{R' \lambda}\right\}. \end{aligned} \quad (12)$$

By evaluation of the integrals it can be shown that the diffraction pattern is given for $1/(R_0+R) + (1/R') = 1/f$, and the image is given, as before, $(1/R) + (1/R') = 1/f$. The effect on the diffraction pattern or on the image of incomplete coherence of the incident radiation, i.e., of illumination of the object by an incoherent source of appreciable extent, is given by summing the intensities for each point of the source separately. Thus the intensity for the source point $x = X$, given by $|\psi_X(x)|^2$, from (12), is calculated and multiplied by the intensity of the source point, $I_0(X)$. Then the observed intensity distribution is derived by integrating over X .

The effect of the limitation of the aperture of the lens on the image is investigated both experimentally and theoretically by placing the aperture in the back-focal plane to multiply the diffraction pattern by the transmission function of the aperture. The result is exactly the same as expressed in equations (5) and (6).

Excellent examples of the effects on the image of the limitation of the diffraction pattern by apertures of various sizes and shapes, produced by use of the optical diffractometer, are given by Taylor and Lipson [1964] (their Plates 43 to 46).

The effect of defocussing of the lens is readily deduced. If the plane of observation is a distance Δ' away from the plane of the in-focus image, which has amplitude distribution $\psi_0(x)$, then,

$$\psi(x) = \psi_0(x) * \exp\{-\pi i x^2 / \Delta' \lambda\}. \quad (13)$$

Alternatively we may say that the amplitude on a plane a distance Δ away from the object is brought to a focus on the plane of observation so that the distribution which is imaged is

$$q(x) * \exp\{-\pi i x^2 / \Delta \lambda\} \quad (14)$$

The amplitude distribution in the back-focal plane, given by Fourier transform, is then

$$Q(u) \exp\{-\pi i \lambda \Delta u^2\}.$$

Thus the effect of defocus can be considered as equivalent to the addition of a second order phase term in the back-focal plane. Higher order terms in the exponent are introduced by aberrations of the lens. The third-order spherical aberration, for example, adds a term proportional to u^4 .

3.4. Phase contrast

3.4.1. Phase and amplitude objects

A pure phase object is an idealized concept. It is essentially a two-dimensional object which changes the phase but not the amplitude of the incident wave. We may use it for convenience to describe a thin object having a varying refractive index when the angles of refraction of the incident waves are so small that the lateral spread of the waves within the object thickness can be ignored. In the one-dimensional form which we use for convenience, the transmission function is of the form

$$q(x) = \exp\{i\phi(x)\} \quad (15)$$

where the phase change $\phi(x)$ depends on the thickness and refractive index of the material.

If a plane incident wave falls on this object, the transmitted intensity is $|q(x)|^2 = 1$. Similarly, an ideally perfect lens would give an exact re-creation of $q(x)$ so that the intensity distribution of the image would be

$$|\psi(x)|^2 = |q(-Rx/R')|^2 = 1. \quad (16)$$

Thus there is no contrast in the image corresponding to the structure of the object. For most objects some absorption and multiple scattering takes place so that the object, if thin enough, must be regarded as a mixed phase and amplitude object with a transmission function which may be written

$$q(x) = \exp\{i\phi(x) - \mu(x)\}, \quad (17)$$

$$|q(x)|^2 = \exp\{-2\mu(x)\}. \quad (18)$$

However for a large and important class of objects the pure phase object forms a reasonable approximation. Most thin biological samples are very nearly phase objects for visible light.

All electron microscope specimens of less than a certain thickness are essentially phase objects for electrons. The refractive index for electrons, given by equation (1.6), varies with the value of the electrostatic potential.

While for both light and electrons staining techniques have been developed to increase the effective absorption and so provide contrast in the in-focus image, these techniques are of limited use for some purposes and may introduce artifacts which complicate the image interpretation. We therefore consider in some detail the so-called "phase-contrast" techniques for obtaining contrast in the image of a pure phase object.

3.4.2. Out-of-focus contrast

This is the most common form of phase contrast, especially in electron microscopy. It is a well-known experimental observation that the contrast is a minimum near exact focus. Contrast appears off-focus and reverses when one goes through focus.

We consider the amplitude distribution on a plane a distance Δ from the object, written as

$$\psi(x) = q(x) * \exp\{-\pi i x^2 / \Delta \lambda\}. \quad (19)$$

Instead of the form (15), it is more convenient to work with the Fourier transform function $\Phi(u)$, defined by

$$q(x) = \exp\{i\phi(x)\} \equiv \int \Phi(u) \exp\{-2\pi i u x\} du. \quad (20)$$

The convolution in (19) is

$$\iint \Phi(u) \exp\{-\pi i X^2 / \Delta \lambda\} \exp\{-2\pi i u(x-X)\} dX du.$$

Making use of the standard integral (1.28) gives us

$$\int \Phi(u) \exp\{-2\pi i u x\} \exp\{i\pi \Delta \lambda u^2\} du. \quad (21)$$

Then if Δ is sufficiently small we may put

$$\exp\{i\pi \Delta \lambda u^2\} \approx 1 + i\pi \Delta \lambda u^2,$$

and we obtain for the image, referred back to the object plane,

$$\begin{aligned}\psi(x) &= \int \Phi(u) \exp\{-2\pi i u x\} (1+i\pi\lambda\Delta u^2) du \\ &= \exp\{i\varphi(x)\} + i\pi\Delta\lambda \int u^2\Phi(u) \exp\{-2\pi i u x\} du.\end{aligned}\quad (22)$$

The relation (2.28) suggests that the integral in this expression may correspond to a second differential and, in fact

$$\frac{d^2}{dx^2} [\exp\{i\varphi(x)\}] = \int (-4\pi^2 u^2) \Phi(u) \exp\{-2\pi i u x\} du;$$

but straight-forward differentiation gives

$$\frac{d^2}{dx^2} [\exp\{i\varphi(x)\}] = -\exp\{i\varphi\} [\{\varphi'(x)\}^2 + i\varphi''(x)],$$

where the primes represent differentiation with respect to x . Hence (22) is

$$\psi(x) = \exp\{i\varphi(x)\} [1 + (\Delta\lambda/4\pi)\varphi''(x) + (i\Delta\lambda/4\pi)(\varphi'(x))^2]. \quad (23)$$

Then the intensity of the out-of-focus object is, to first order in the small quantity, $\Delta\lambda$,

$$I(x) = 1 + (\Delta\lambda/2\pi) \varphi''(x). \quad (24)$$

Thus the contrast depends on the second differential of the phase function $\varphi(x)$, and is reversed when the sign of the defocus, Δ , is reversed.

As we noted previously, going out of focus is equivalent to changing the phase of the amplitude of the diffraction pattern in the back focal plane by a factor proportional to u^2 . The effect of lens aberrations can be represented by changing the phase of the diffraction pattern by a function in the form of a power series in even orders of u and defocus is included as a first-order aberration, coming from the second-order term in u . The fourth-order terms include third-order spherical aberration.

When the lens is exactly focussed the phase changes due to these higher order terms in u remain and provide some amplitude contrast. To an even greater extent than with the second-order defocus terms, these phase shifts are small for small u but increase rapidly for large u so that in the image contrast most contributions come from the outer part of the diffraction patterns. Hence contrast fluctuations may be expected where $\varphi(x)$ changes rapidly.

The treatment given above in equations (20) to (24) may be extended to include higher order terms in u , but in practice the result contains combinations of higher-order differentials of $\varphi(x)$ which are not so readily visualized and so are less useful.

For out-of-focus images, the phase changes due to spherical aberration are either added to or subtracted from the phase changes for defocus, depending on the direction of defocus, and either assist or hinder the phase contrast imaging. This is an important consideration for electron microscopy of thin objects and will be discussed in more detail in Chapter 13.

3.4.3. Aperture limitation

For a phase object the formation of an image of zero contrast depends on interference of waves from the back-focal plane with exactly the right amplitude and relative phase. Any modification of the back-focal plane amplitude distribution will upset this balance and produce some contrast.

If, for example, a circular aperture is placed centrally in the back-focal plane, the image has the form

$$I(x,y) = |\exp\{i\varphi(x,y)\} * J_1(\pi ar) / (\pi r)|^2. \quad (25)$$

It is not immediately obvious that this represents amplitude contrast. However from qualitative reasoning we may predict that, firstly, since the outer parts of the diffraction pattern are affected, only the higher-order Fourier components of $q(x,y)$ are modified; hence contrast appears where there are sudden changes in $\varphi(x,y)$ as at the edges of particles. In fact the image contrast shows some similarity to the function $|\varphi'(x)|$. Secondly, since the image is convoluted by a spread function the resolution is not better than the width of the spread function. Hence again, the gain in contrast is made at the expense of resolution.

Other forms of aperture limitation include the well-known case of "Schlieren" optics used for visualizing air flow in wind tunnels. A straight edge is introduced in the back-focal plane to cut off half the diffraction pattern. From the relation (2.41) it is seen that the effect may be represented in part by convolution of the image $\exp\{i\varphi(x)\}$ with a function of the form $(2\pi ix)^{-1}$, which gives contrast somewhat resembling $\varphi'(x)$.

3.4.4. Zernike phase contrast

The form of phase contrast introduced by Zernike is the most effective in that it produces maximum contrast with no

loss of resolution and the contrast is proportional to the phase function $\varphi(x)$ itself, rather than to any differential of this function. It is most easily understood for the case of a small phase change, $|\varphi(x)| \ll 1$. For this condition we can assume that all higher powers of $\varphi(x)$ are negligible and write

$$q(x) = 1 + i\varphi(x). \quad (26)$$

Then in the back-focal plane the amplitude is

$$Q(u) = \delta(u) + i\Phi(u), \quad (27)$$

where the delta function represents the directly transmitted plane-parallel incident beam and $\Phi(u)$ is the radiation scattered into the rest of the diffraction pattern. No contrast results when these two portions are re-united in the image because they are $\pi/2$ out of phase, giving

$$\psi(x) = 1 + i\varphi(x)$$

and $\psi\psi^* = 1$ to first order terms only. The central beam and the rest of the diffraction pattern can be brought back into phase if the phase of the central beam is changed by $\pi/2$ by inserting a quarter-wave plate. Then (27) becomes

$$Q'(u) = i[\delta(u) + \Phi(u)]$$

and

$$\psi(x) = i[1 + \varphi(x)],$$

so that, to the first order in $\varphi(x)$,

$$I(x) = \psi\psi^* = 1 + 2\varphi(x). \quad (28)$$

This simple description does not often represent the practical situation. If the phase change is not small, as in many important cases in both optical and electron microscopy, the theory becomes much more complicated. Also it is impossible to change the phase of a delta-function central beam only. The phase plate has finite size and so changes the phase of part of the diffraction pattern also. Further, we have assumed a plane incident wave. This is a reasonable approximation for the usual conditions of electron microscopy. It is not so in optical microscopy since for a plane wave the intensity of the image is insufficient. Instead, a hollow cone of illumination is used with an annular phase-plate near the back-focal plane.

3.5. Holography

Gabor (1948,1949) introduced the concept of holography as the basis for a scheme to overcome the limitations of resolution in electron microscopy due to the unavoidable effects of lens

aberrations (see Chapter 13). In order to correct for the phase changes due to lens aberrations, it is necessary to influence the relative phases of the waves contributing to the image, but in the recording of the image intensities, the phases are normally lost. Gabor proposed the use of a reference wave which could interfere coherently with the image wave, giving rise to intensity data which contained the phase information of the image. In his initial proposal, a strong electromagnetic lens is used to form a small cross-over close to the plane of a weakly scattering thin object, as in Fig. 3.4(a), so that a highly magnified "shadow image" is formed on a distant plane of observation. The electron source is assumed to be so small that the electron beam

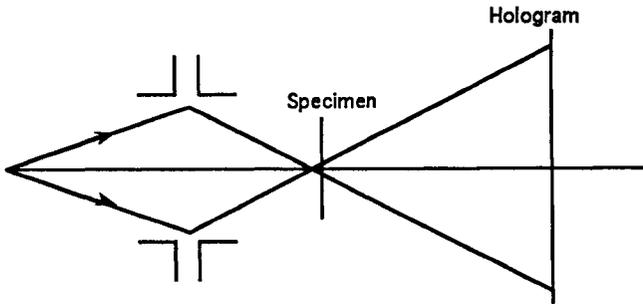


Fig. 3.4(a)

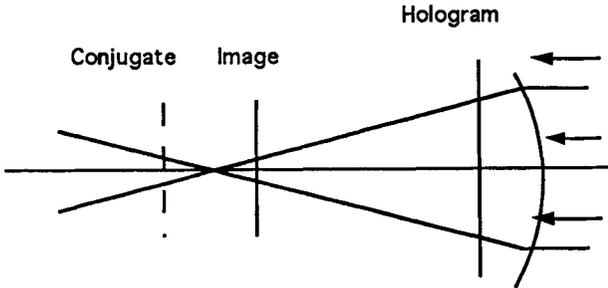


Fig. 3.4(b).

Fig. 3.4. Diagram suggesting the principles of Gabor holography. In (a) a lens is used to form a cross-over close to a thin specimen and a magnified shadow-image or hologram. In (b), the hologram is illuminated by a convergent wave to form the reconstruction of the specimen wave as a conjugate image.

incident on the specimen is completely coherent. The wave directly transmitted through the object is taken as the reference wave and the much weaker wave, scattered by the

object, is the image signal. In the shadow-image plane, or hologram, these two waves interfere, so that the intensity distribution reflects their relative phases. In a reconstruction process, the recorded hologram is illuminated by coherent visible light, as in Fig. 3.4(b), and the image wave function of the object is recreated in the equivalent plane. Unfortunately, because it is the intensity of the hologram which is recorded, the reconstructed wave function is accompanied by a defocused "conjugate" image, providing an unwanted background.

In a simplified, one-dimensional description, we may consider the electron cross-over as a point source created by a lens with no aberrations. The wave incident on the specimen is then $\exp(-\pi i x^2/R_0\lambda)$. The transmission function for the weakly scattering object may be written $q(x) = 1 + \epsilon(x)$, where $\epsilon(x)$ is small and, in general, complex. The intensity distribution on the distant hologram is then

$$\begin{aligned} |\Psi(u)|^2 &= |Q(u) * \exp(\pi i R_0 u^2)|^2 \\ &= \delta(u) + E(u) * \exp(-\pi i R_0 u^2) + E^*(u) * \exp(+\pi i R_0 u^2), \end{aligned} \quad (29)$$

where the term of second order in $E(u)$, the Fourier transform of $\epsilon(x)$, has been neglected.

Performing the inverse Fourier transform in the reconstruction process gives the wave function

$$\Psi(x) = 1 + \epsilon(x) \cdot \exp(-\pi i x^2/R_0\lambda) + \epsilon^*(x) \cdot \exp(+\pi i x^2/R_0\lambda). \quad (30)$$

If this reconstruction is imaged a distance R_0 out of focus, to give $\epsilon(x)$ directly, the image is accompanied by the conjugate image $\epsilon^*(x)$, imaged a distance $-2R_0$ out of focus. If the distance R_0 can be made large, and if the desired image consists of small, sharply defined objects, a reasonable representation of the desired object may be achieved as in the original reconstructions from holograms made with visible light (Gabor, 1949) in the absence of suitable bright electron sources at that time.

The presence of aberrations in the lens used to form the small electron crossover of Fig. 3.4(a) has the effect of adding additional phase terms to the $\exp(\pi i R_0 \lambda u^2)$ term in (29). If an equivalent aberration phase term is subtracted from the real image term in the reconstruction process of (30), to give an image free of aberration, a doubled, reversed phase term is added to the conjugate image, making it more strongly aberrated and closer to random noise.

Realization of the Gabor scheme, using a small electron crossover, and showing an improvement of resolution by correction for aberrations, was not achieved for many years (Lin and Cowley, 1986) when the use of high-brightness electron sources involving field-emission guns became

feasible. Lin and Cowley showed that reconstructions could be conveniently made by the computer manipulation of digitally recorded holograms and it is possible to remove the contributions of the unwanted conjugate image by, for example, superimposing a number of reconstructions from holograms obtained with lateral shifts of the electron beam relative to the object (see also Wang and Cowley, 1995).

The great expansion of the application of the holographic principle to the many areas of imaging with visible light and other radiations, including high energy electrons, came with the realization by Leith and Upatnieks (1962) that the reference wave used to define the image wave phases need not be one transmitted through the object but could be one passing around the object in free space. For visible light it is easy to split a coherent wave into two portions, widely separated in space, by means of prisms or partially transparent mirrors.

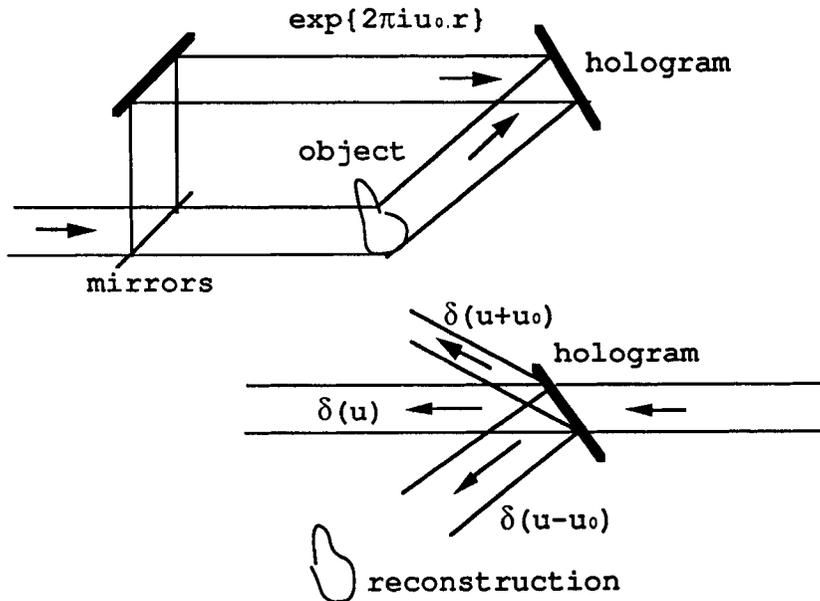


Fig. 3.5. Diagrams suggesting the formation of an off-axis hologram with light when a plane reference wave interferes with the wave scatter by an object and the reconstruction of the object wave in the direction $\delta(u-u_0)$ when the hologram is illuminated by a plane wave.

For a system represented as in Fig. 3.5, the reference plane wave $\exp(2\pi i u_0 x)$ may be added to the image signal, $\Psi_1(x)$. At a large distance, the diffraction pattern intensity is then represented by

$$|\Psi(u)|^2 = \{\delta(u) + (\Psi_1(u) * \Psi_1^*(u))\} * \delta(u) \\ + \Psi_1(u) * \delta(u - u_0) + \Psi_1^*(u) * \delta(u + u_0). \quad (31)$$

The distribution around the direction $\delta(u-u_0)$ may be readily separated from the distributions around the directions $\delta(u)$ and $\delta(u+u_0)$ by use of apertures. Then the inverse Fourier transform of $\Psi_1(u)$ gives the desired image wave, $\psi_1(x)$, directly, without confusion. For visible light, the possibility of using large angles between the object wave and the reference wave introduces the possible use of holography for three-dimensional imaging. Applications for electron waves will be discussed further in Chapter 13.

3.6. Multi-component systems

The treatment of optical systems in the small angle approximation reproduces most of the properties of real optical systems and is a very good approximation for electron optics of medium- and high-energy systems since the scattering of electrons of these energies by atoms is essentially a small-angle phenomenon. In Section 3.3 we showed how to write down the expressions for the diffraction patterns, images or the phase distributions on any plane in a simple ideal-thin-lens system, using this approximation. We now consider the extension of this treatment to multi-component systems. We consider, for brevity and convenience, only one-dimensional objects. The extension to two dimensions is obvious.

Radiation from a source with amplitude distribution $q_0(x)$ passes through a series of planar objects having transmission functions $q_n(x)$. Propagation over the distance R_n , from the n th to the $(n + 1)^{\text{th}}$ object is represented by convolution with the propagation function $p_n(x)$ (see Fig. 3.6). In the small-angle approximation,

$$p_n(x) = (i/R\lambda)^{1/2} \exp\{-2\pi i k R_n\} \exp\{-\pi i x^2/R\lambda\}, \quad (32)$$

and neglecting the $\exp\{-2\pi i k R_n\}$ term, the Fourier transform of $p_n(x)$ is

$$p_n(u) = \exp\{\pi i R_n \lambda u^2\}. \quad (33)$$

Then the amplitude on the plane of observation, which may be regarded as the plane of the $(N + 1)^{\text{th}}$ object is written

$$\Psi_{N+1}(x) = q_N(x) \left[\dots \left[q_2(x) \left[q_1(x) \left[q_0(x) * p_0(x) \right] * p_1(x) \right] * p_2(x) \right] \dots \right] * p_N(x) \quad (34)$$

where the brackets have been numbered for clarity. The contents of the N bracket represent the amplitude of the wave

incident on the N^{th} object. This amplitude is multiplied by the transmission function $q_N(x)$ and the product is then convoluted with $p_N(x)$.

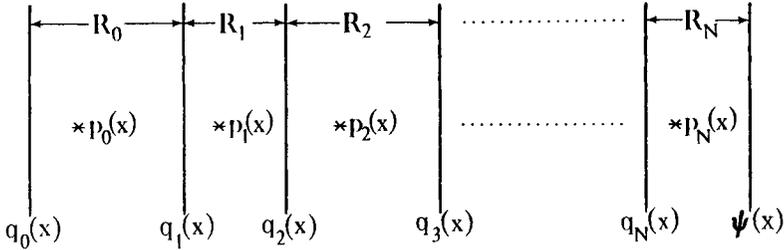


Fig. 3.6. Diagram illustrating the description of wave propagation through a multi-component system.

By making use of the convolution and multiplication theorems of Fourier transforms, the Fourier transform of (34) is written

$$\Psi_{N+1}(u) = Q_N(u) * \dots \left[Q_2(u) * \left[Q_1(u) * Q_0(u) P(u) \right] P_1(u) \right] P_2(u) \dots P_N(u), \quad (35)$$

where again, the multiplication takes place before the convolution. This represents the Fraunhofer diffraction pattern given by the radiation emerging from the N objects.

In the case of the small angle approximation using (32) and (33), the expression for either the amplitude, (34), or the diffraction pattern, (35), may be written in terms of any combination of the functions in real space, $q_n(x)$, $p_n(x)$, or the functions of Fourier transform space, $Q_n(u)$, $P_n(u)$. This follows from the special properties of the complex exponentials with respect to Fourier transform and convolution.

Thus we have

$$\begin{aligned} q(x) * \exp\{-i\alpha x^2\} &= \int q(X) \exp\{-i\alpha(x - X)^2\} dX \\ &= \exp\{-i\alpha x^2\} \int q(X) \exp\{-i\alpha X^2\} \exp\{2i\alpha x X\} dX \\ &= (\pi/i\alpha)^{1/2} \exp\{-i\alpha x^2\} [Q(\alpha x/\pi) * \exp\{i\alpha x^2\}]. \end{aligned} \quad (36)$$

Similarly

$$Q(u) * \exp\{i\alpha u^2\} = (\pi i/\alpha)^{1/2} \exp\{i\alpha u^2\} [q(\alpha u/\pi) * \exp\{-i\alpha u^2\}] \quad (37)$$

Thus any bracketed term of (34) or (35) may be converted to a form involving Q instead of q or q instead of Q . This is

of particular value, for example, for periodic objects for which it is often more convenient to deal with the Fourier transform, a set of weighted delta functions, rather than a continuous function.

Another useful relationship is

$$\begin{aligned}
 & [q(x) \exp\{-i\alpha x^2\}] * \exp\{-i\beta x^2\} \\
 & = (i\pi/(\alpha + \beta))^{1/2} \exp\{-i\beta x^2\} [Q(\beta x/\pi) * \exp\{i\beta^2 x^2/(\alpha + \beta)\}] \quad (38)
 \end{aligned}$$

Using the relationship (37) to modify the bracketed term, this becomes

$$\begin{aligned}
 & [q(x) \exp\{-i\alpha x^2\}] * \exp\{-i\beta x^2\} \\
 & = \exp\{-i\alpha\beta x^2\} [q(\beta x/(\alpha + \beta)) * \exp\{-i\beta^2 x^2/(\alpha + \beta)\}]. \quad (39)
 \end{aligned}$$

Similarly,

$$\begin{aligned}
 & [Q(u) \exp\{i\alpha u^2\}] * \exp\{i\beta u^2\} \\
 & = \exp\{i\alpha\beta u^2\} [Q(\beta u/(\alpha + \beta)) * \exp\{i\beta^2 u^2/(\alpha + \beta)\}], \quad (40)
 \end{aligned}$$

and so on. Such relationships are of use when a transmission function is multiplied by an incident wave from a point source, $\exp\{-\pi i x^2/R_0\lambda\}$, or when a transmission function $q_n(x)$ modifies the transmission function of an ideal thin lens, $\exp\{\pi i x^2/f\lambda\}$.

In the limit that the number of two-dimensional objects tends to infinity and the distance between them tends to zero, expressions of the form (34) or (36) may be used to represent rigorously the scattering from any three-dimensional object (Moodie [1972]). To represent the scattering from a three-dimensional object with any desired degree of accuracy, the number of planar objects must be made sufficiently large and the distance between them sufficiently small. An example of this approach will be given in our consideration of the scattering of electrons by crystals (Chapter 11).

3.7. Partial Coherence

In this chapter we have considered the two limiting cases of coherent imaging, when the object is illuminated by a plane wave or radiation from a monochromatic point source, and incoherent imaging which applies to the imaging of a self-luminous source consisting of independently emitting point sources. Most practical situations are intermediate between these limiting, ideal cases. As has been pointed out in

Section 1.3, it is always possible to derive the observed intensity in an experiment by considering separately the intensity given by each point of the source and by each different wavelength, and then adding all these intensities. However, it is often convenient to make use of the concept of a degree of coherence as a measure of partial coherence.

In its most complete form, the concept of a degree of coherence as a function of space and time provides a description of imaging or diffraction processes in terms of observable quantities so that there is no need to postulate an unobservable and hence fictitious wave function. (See, for example, Born and Wolf [1989]). In its simplest form, the use of the degree of coherence or the approximations of a "coherence width" and "coherence length" provide very convenient indications of the amount and type of the interference effects which might be observed in any experimental situation. We will confine our attention to these latter, less exacting aspects.

The degree of coherence is not a property of radiation at any one point. It is a measure of the correlation which exists between the phases of the radiation at two points. If there is a correlation between the phases at the two points, then light proceeding from those two points can give interference effects if it is allowed to overlap. The degree of coherence is defined in terms of the strength of these interference effects. Thus a practical interference experiment can be set up to measure the degree of coherence. If the two points are separated by a vector perpendicular to the direction of propagation we can measure the "lateral coherence". If they are separated by a vector parallel to the direction of propagation we measure the "longitudinal coherence", also called the "chromatic coherence" because the correlation of the phases then depends only on the spread of wavelengths or frequencies present.

To measure the lateral coherence between two points we can envisage an ideal Young's fringes experiment, in one dimension and with monochromatic radiation for simplicity. Narrow slits are placed at the two points and the interference pattern is observed on some plane beyond them. For perfect coherence, the interference pattern is given as squared cosine fringes [equations (1.30) and (2.46)]. For complete incoherence, the intensities from the slits are added separately and there are no fringes. The degree of coherence between the two points is given by the visibility of the fringes, defined as

$$\text{Visibility} = (I_{\max} - I_{\min}) / (I_{\max} + I_{\min}).$$

For a fringe intensity function $A + B \cos bx$, the visibility is B/A .

As an example, we consider the coherence due to a small but finite source which may be considered as ideally incoher-

ent, with an intensity distribution $I_0(x)$. If this is placed a distance R_0 from the two points, which are separated by a distance x_1 and the interference pattern is observed at a distance R , the observed intensity due to a point in the source at $x = X$ is, from (1.30).

$$I_X(x) = 4 \cos^2(\pi x_1/\lambda) (x/R+X/R_0). \quad (41)$$

The total intensity observed is then given by integrating over the source, as

$$\begin{aligned} I(x) &= \int I_0(X) 4 \cos^2[(\pi x_1/\lambda) (x/R+X/R_0)] dX \\ &= 2 \int I_0(X) dX + 2 \cos^2\{\pi x_1/R\lambda\} \cdot \int I_0(X) \cos\{2\pi x_1 X/R_0\lambda\} dX. \end{aligned} \quad (42)$$

So that the visibility of the fringes and the degree of coherence is

$$\gamma_{12} = \left[\int I_0(X) \cos(2\pi x_1 X/R_0\lambda) dx \right] / \int I_0(X) dX. \quad (43)$$

Thus the degree of coherence is given by the normalised Fourier transform of the source intensity distribution on a scale of $2x_1/R_0\lambda$. This, in simplified form, is the Zernike-van Cittert theorem.

For a circular source of uniform intensity and diameter a , for example, the degree of coherence takes the form $J_1(\pi a u) / (\pi a u)$ where $u = x_1/R\lambda$. This defines the "coherence patch". As a measure of the width of this patch we take the radius to the first zero, giving $x_1 \approx \lambda/\alpha$ where $\alpha = a/R_0$, the angle subtended by the source at the points of interest. This gives a rough indication of the distances between two points for which interference effects are observable. Conversely, if the interference fringe visibility is measured as a function of the separation of two slits, it is possible to deduce α , the angle subtended by the source. This is the basis for the use of the Michelson Stellar Interferometer to measure apparent diameters of stars.

In an entirely analogous fashion it is possible to define the longitudinal or chromatic coherence in terms of the visibility of fringes produced in a Michelson interferometer or any similar interferometer which takes radiation going in one direction, splits it into two paths of unequal length and recombines it to produce intensity fluctuations as a function of path length difference. For monochromatic radiation one gets fringes

$$I(x) = \cos(\pi x/c),$$

where ν is the frequency, x is the path length difference and c is the velocity.

For a small range of frequencies $\nu = \nu_0 + \nu^1$ around an average frequency ν_0 , the observed intensity is given by adding intensities for all frequencies as

$$I(x) = \int I(\nu^1) \cos^2(\pi\nu x/c) d\nu$$

and the longitudinal coherence is given by the visibility,

$$[(\int I(\nu) \cos(2\pi\nu x/c) d\nu)^2 + (\int I(\nu) \sin(2\pi\nu x/c) d\nu)^2]^{1/2} / \int I(\nu) d\nu$$

or

$$|\int I(\nu) \exp(2\pi i \nu x/c) d\nu| / \int I(\nu) d\nu.$$

Thus for a spectral line which is gaussian in form of width $\nu^1 = a$, the longitudinal coherence function is a gaussian of width $x = c/\pi a$, which may be called the coherence length.

In Chapter 4 we give estimates of the coherence width and coherence lengths commonly encountered in diffraction experiments with the usual radiations and sources.

Problems

1. A lens is used to image a coarse diffraction grating composed of slits in an opaque screen. An aperture is inserted in the back-focal plane of the lens. Find what image will be produced if the aperture lets through

- (a) the central spot and the first diffracted spot on each side of it,
- (b) the central spot and the first diffracted spot on one side only,
- (c) the first diffracted spot on each side but not the central spot.

2. An object having a transmission function $A + B \cos(2\pi x/a)$ is imaged by an ideal thin lens as defined in Section 3.

(a) Find the variation of image intensity as a function of the position of the plane of observation.

(b) Determine how the sequence of out-of-focus images will be affected by spherical aberration of the lens, limitation of the lens aperture, and by imaging in white light rather than monochromatic light.

3. (a) Make a simple argument to show that if a lens is used to image a perfectly incoherent source, the coherence in the image plane depends only on the aperture of the lens and not at all on the extent of the source (provided that the source is larger than some minimum size).

(b) Using the definition of a degree of coherence between two points as the visibility of interference fringes resulting when small apertures in an opaque screen are placed at the two points, derive the result (a) mathematically.

4. A high-contrast (black and transparent) photographic negative of a chess board (8 squares each way with a white surround) is imaged in an optical system. The transparent parts of the negative have a uniform phase retardation. What is the form of the diffraction pattern in the back-focal plane of the imaging lens? If a screen with a slit parallel to one edge of the chess board is used to cut out all the diffraction maxima except those in a row through the central maximum, how will the image be affected? What image will result if the slit is rotated through 45° ?

(Note. The amplitude along a line in a diffraction pattern is the Fourier transform of the projection of the object transmission function on to a parallel line.)

This Page Intentionally Left Blank

Section II

KINEMATICAL DIFFRACTION

This Page Intentionally Left Blank

Radiations and their scattering by matter

4.1. X-rays

4.1.1. X-ray sources

The X-ray radiation normally used in diffraction experiments is the relatively sharp lines of the characteristic K_{α} radiation spectrum emitted when an electron beam of sufficient energy strikes the metallic anode of an X-ray tube. The wavelengths vary from 2.28 Å for chromium to 0.71 Å for molybdenum, the most commonly used being the copper radiation of wavelength 1.54 Å (actually the $K_{\alpha 1}$ and $K_{\alpha 2}$ doublet having wavelengths approximately 1.537 and 1.541 Å). The radiation from a typical X-ray tube contains, in addition to these strong maxima, one or more K_{β} lines of shorter wavelength, some weak long-wavelength lines of the L series, some weak lines from impurities or contamination of the X-ray tube target, and a continuous background of white radiation starting abruptly at a short wavelength limit (for which the energy of the emitted X-ray, hc/λ , is equal to the energy of the electrons of the exciting electron beam, eE) passing through a maximum and then decreasing with increasing wavelength. Confusion of the diffraction patterns by unwanted radiation may be reduced by use of various monochromating devices, including absorbing filters, crystal monochromators and energy-discriminating detectors.

The isolated K_{α} lines are rather broad compared with spectral lines in the visible region, having a relative width, $\Delta\lambda/\lambda$, of the order of 10^{-4} . This limits the chromatic coherence of the radiation so that the coherence length is of the order of $10^4\lambda$ or approximately 1 μm . The more serious complication for many purposes comes from having the two adjacent $K_{\alpha 1}$ and $K_{\alpha 2}$ lines.

Recently increasing use has been made of the very intense synchrotron radiation emitted in a comparable wavelength range when electrons of sufficient energy (in the GeV range) are accelerated by the bending magnets of the synchrotron or by special "wigglers" inserted in the beam path. The synchrotron radiation is "white" radiation, i.e. it is emitted with a continuous range of wavelengths. The maximum intensity, for a circular electron beam orbit, is given for a wavelength $\lambda \approx 1.4\lambda_c$, where the critical wavelength $\lambda_c = (4\pi/3) R(E/mc^2)^{-3}$ where R is the radius of

curvature and E is the electron energy, e.g. maximum intensity is at $\lambda = 1.5 \text{ \AA}$ for $E \approx 5.5 \text{ GeV}$ (less with a wiggler). The radiation is highly polarized and is pulsed with the orbital frequency of the electron bunches in the synchrotron. For recently-constructed synchrotron radiation sources, this X-ray intensity obtained when a monochromator is used to select a small range of wavelengths is much higher than that for the characteristic X-ray line from a conventional X-ray tube having a comparable range of wavelengths (Coppens, 1992; Margaritondo, 1988). Special experimental geometries and equipment are being developed to take advantage of the specific properties of the synchrotron radiation sources (Winick, 1980; Bachrach, 1992).

4.1.2. Scattering by electrons

The scattering of X-rays by matter is usually considered in terms of the scattering by a single electron, localized at the origin of coordinates by some sort of restoring force and so having some resonant frequency, ν_0 , associated with it. For incident radiation

$$E = E_0 \exp\{2\pi i(\nu t - \mathbf{k} \cdot \mathbf{r})\}$$

having a frequency $\nu \gg \nu_0$, the simple Thomson scattering theory gives a radiated wave of amplitude

$$E_s = -E_0 (e^2/mc^2) R^{-1} \sin\psi \cdot \exp\{2\pi i(\nu t - \mathbf{k} \cdot \mathbf{r})\} \quad (1)$$

where R is the distance to the point of observation and ψ is the angle between the scattered beam and the direction of the acceleration of the electron.

If, as in Fig. 4.1 we consider the electron at the origin of orthogonal axes and the direction of propagation is the z -axis, then we consider the components of the incident radiation polarized in the $x - z$ and $y - z$ planes. The scattered radiation is in, say, the $y - z$ plane. Then for the component polarized in the $y - z$ plane, the direction of the electron acceleration is the y axis and $\psi = 90 - \phi$, where ϕ is the scattering angle. Then

$$|E| = E_0 (e^2/mc^2) R^{-1} \cos\phi.$$

For the polarization in the $x - z$ plane, $\psi = \pi/2$ and $\cos\phi = 1$ so that

$$|E| = E_0 (e^2/mc^2) / R.$$

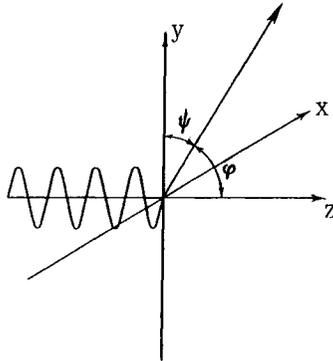


Fig. 4.1. Coordinate system for discussion of scattering of X-rays by an electron.

For unpolarized incident radiation the total scattered intensity is the sum of the intensities for the two polarizations, i.e.

$$I = I_0 (e^2/mc^2)^2 R^{-2} (1 + \cos^2\phi) / 2. \quad (2)$$

The last term here is the polarization factor. It is customary to treat the scattering as for a scalar wave function, omit the polarization factor until the final reckoning of intensities and take $(e^2/mc^2)R^{-1}$ as 1 "electron unit" of scattering.

The factor m^{-1} in (2) ensures that for scattering from atoms the scattering from the nucleus can be ignored. We consider the scattering from the cloud of electrons around the nucleus and use the nucleus only as an origin of coordinates. Then for each electron we may define a distribution function or electron density function $\rho_n(\mathbf{r})$ which gives the probability that the electron should be contained in unit volume at the position defined by \mathbf{r} .

As we will see later, in Chapter 7, the purely elastic scattering is given by such a time-averaged scattering function. Hence, from equation (1.19) and (1.21), using $\rho_n(\mathbf{r})$ instead of $-\varphi(\mathbf{r})$ for the scattering strength and expressing the result in electron units in order to eliminate the constants, we find the scattering amplitude for the n th electron is

$$f_n(\mathbf{q}) = \int \rho_n(\mathbf{r}) \exp\{-2\pi i \mathbf{q} \cdot \mathbf{r}\} d\mathbf{r}, \quad (3)$$

or, putting $\mathbf{q} = \mathbf{u}$,

$$f_n(\mathbf{u}) = \int \rho_n(\mathbf{r}) \exp\{2\pi i \mathbf{u} \cdot \mathbf{r}\} d\mathbf{r}. \quad (4)$$

The total scattering from the electron is given by (1) and is equal to one electron unit. The difference between the elastic scattering, (4), and the total scattering is then the inelastic scattering, or Compton scattering, corresponding to the case that the X-ray photon collides with the electron, with loss of energy and change of momentum calculated in the usual way. The intensity of the inelastic scattering is then

$$I'_n = 1 - |f_n|^2.$$

4.1.3. Scattering by atoms

For all the electrons associated with an atom, the elastic scattering is given by scattering from the time averaged electron density function

$$\rho(r) = \sum_n \rho_n(r),$$

so that

$$f(u) = \sum_n f_n(u)$$

and the so-called "atomic scattering factor" is

$$f(u) = \int \rho(r) \exp\{2\pi i u \cdot r\} dr \quad (5)$$

or, if spherical symmetry of the atom is assumed, as it may always be for free atoms,

$$f(\theta) = \int_0^\infty 4\pi^2 \rho(r) \sin(2\pi ur) / (2\pi ur) dr, \quad (6)$$

where θ is half the scattering angle φ .

For inelastic scattering, the scattering from the various electrons is incoherent, so we add intensities to give

$$\begin{aligned} I_{inel.} &= \sum_n \{1 - |f_n|^2\} \\ &= Z - \sum_n |f_n|^2, \end{aligned} \quad (7)$$

where Z is the atomic number. Thus the calculation of the Compton scattering is in general more complicated than for elastic scattering since the distribution functions for all electrons and their Fourier transforms must be calculated separately. The elastic atomic scattering factors, calculated from various approximations for the electron density function, are listed in the International Tables for Crystallography, Vol. C (Wilson, 1992).

4.1.4. Dispersion corrections

The assumption which forms the basis for equation (1) and all the subsequent derivation, is that the frequency of the incident radiation is much greater than the frequency corresponding to any energy for excitation of the atom. This is clearly not the case if, as frequently happens, the atom has an absorption edge not far from the incident X-ray frequency. In the presence of absorption, the refractive index and also the atomic scattering factor become complex. An imaginary part and a small real part are added to $f(u)$ as defined by (5), so we write

$$f = f_0 + f' + if'' \quad (8)$$

The variation of f' and f'' with wavelength are roughly as indicated in Fig. 4.2 with discontinuities at λ_K , the wavelength of the K absorption edge for the atom.

Since f' and f'' arise from the excitation of electrons from the inner electron shells, they may be considered as Fourier transforms of highly localized scattering distributions and so decrease with scattering angle much more slowly than $f_0(u)$ which is given by Fourier transform of the total distribution of electrons of the atom.

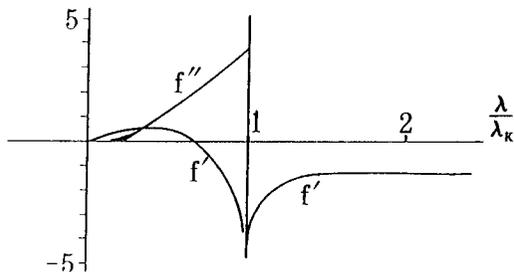


Fig. 4.2. The variation of the magnitude of the real and imaginary "anomalous scattering" components f' and f'' of the atom scattering factor for X-rays as a function of wavelength. The K absorption edge of the atoms is at λ_K .

4.2. Electrons

4.2.1. Sources of electrons

The electrons to be considered for electron diffraction and microscopy have energies mostly in the range 100 to 400 keV although with the advent of high voltage electron microscopes this range needs to be extended to 1 MeV or more. For electrons in the range 10-200 eV used in Low Energy

Electron Diffraction (LEED) the interactions with matter are sufficiently different to require a separate treatment (see Chapter 19).

The wavelength of the electron beam including the relativistic correction is

$$\begin{aligned}\lambda &= h/mv = \lambda_0(1 + eE/(2m_0c^2))^{-1/2}, \\ &= 12.2643/(E_0 + 0.97845 \times 10^{-6} E_0^2)^{1/2}\end{aligned}\quad (9)$$

where E_0 is the accelerating voltage in volts and λ is in angstroms (\AA), and λ_0 is the non-relativistic wavelength. This is 0.0548 \AA for 50 keV, 0.0370 for 100 keV and 0.00867 for 1 MeV electrons.

The electron beam is produced in an electron gun in which the effective source size has a diameter of about 10 μm for a normal "hair-pin" filament, approximately 1-2 μm for heated pointed filaments and as little as 20-50 \AA for field emission tips. In combination with electron lenses these sources can give well collimated beams. A divergence of 10^{-3} or 10^{-4} rads. is used for most purposes but a divergence of 10^{-6} or less can be achieved fairly readily.

In the 100 keV range, the voltage supplies may have a stability of better than 10^{-5} giving sufficient chromatic coherence for most experiments, with a coherence length of the order of 1 μm or better, often limited only by the thermal energy spread of the electron source.

4.2.2. Atom scattering amplitudes

To an approximation which is sufficiently accurate for all calculations of elastic scattering of electrons, the property of matter which is of interest is the potential distribution $\varphi(\mathbf{r})$. The potential distribution is related to the electron density by Poisson's equation

$$\nabla^2\varphi(\mathbf{r}) = \frac{|\mathbf{e}|}{\epsilon_0} \{ \rho_n(\mathbf{r}) - \rho_e(\mathbf{r}) \}, \quad (10)$$

where ρ_n is the charge density due to the atomic nuclei and ρ_e is that due to the electrons.

The atom scattering amplitude for electrons is then appropriately defined as the Fourier transform of the potential distribution for the atom, $\varphi(\mathbf{r})$ measured in volts;

$$f_e(\mathbf{u}) = \int \varphi(\mathbf{r}) \exp\{2\pi i \mathbf{u} \cdot \mathbf{r}\} d\mathbf{r}. \quad (11)$$

The f values are then defined as properties of the atoms, independent of any assumption as to the scattering process or the theoretical approximation used to describe it (Dawson et al. [1974]).

The relationship of this scattering factor to that for X-rays, f_X , is derived by inserting the inverse Fourier transforms in Poisson's equation (10);

$$\begin{aligned} \nabla^2 \left[\int f_e(u) \exp\{-2\pi i u \cdot r\} du \right] \\ = \frac{|e|}{\epsilon_0} \int f_X(u) \exp\{-2\pi i u \cdot r\} du - \frac{|e|}{\epsilon_0} \int Z \exp\{-2\pi i u \cdot r\} du, \end{aligned}$$

where the final integral is a delta-function of weight Z , the atomic number, due to the positive charge on the nucleus. Writing the left-hand side as

$$\int (-2\pi i |u|)^2 f_e(u) \exp\{-2\pi i u \cdot r\} du$$

we equate integrands to obtain the Mott formula

$$f_e(u) = |e| (Z - f_X(u)) / 4\pi^2 \epsilon_0 u^2. \quad (12)$$

We note that for scattering at large angles for which f_X is small, f_e is approximately proportional to Z and decreases with angle as $(\sin^2\theta)/\lambda^2$. For $(\sin\theta)/\lambda$ tending to zero, since for neutral atoms f_X tends to Z , the value of f_e becomes indeterminate. However the limiting value is obtained from (11) as

$$f_e(0) = \int \varphi(r) dr. \quad (13)$$

By historical accident, the atomic scattering factor for electrons has been defined as the quantity occurring in the first Born approximation for the theory of the scattering of electrons by atoms, equation (1.21); this gives

$$f_{FB}(u) = (2\pi m e / h^2) f_e(u) = (\sigma / \lambda) f_e(u),$$

where σ is the interaction constant to be defined below.

Since this quantity is not a property of the atom itself and since the first Born approximation has a very restricted range of validity for electron scattering, particularly from solids, the use of this basis for a definition is inappropriate. Its use has led to considerable confusion in the literature. However it is $f_{FB}(u)$ measured in \AA for which tables of values are given, for example in the International Tables for Crystallography, Vol. C (Wilson, 1992).

4.2.3. Phase object approximation

Since appreciable scattering takes place through relatively small angles only, we may use the Fraunhofer approximation of equation (1.37) instead of the Born approximation (1.21) to describe scattering by atoms. Regarding the atom as a quasi-two-dimensional object, we derive the transmission function $q_e(x, y)$.

From (1.6) the refractive index for the electron wave is $1 + \varphi(r)/2E$. Then the phase difference for the wave passing through the potential field $\varphi(r)$ in the z direction, relative to the wave in vacuum is $(\pi/\lambda E) \int \varphi(r) dz$.

Putting the projection expressed by the integral equal to $\varphi(x, y)$ and $\sigma = \pi/\lambda E$, we have

$$q_e(x, y) = \exp\{-i\sigma\varphi(x, y)\}, \quad (14)$$

and the diffraction amplitude is

$$\Psi(u, v) = \mathfrak{S}[\exp\{-i\sigma\varphi(x, y)\}]. \quad (15)$$

With the relativistic correction, we write

$$\sigma = (2\pi/E\lambda) \{1 + (1 - \beta^2)^{1/2}\}^{-1} = (2\pi m_0 e \lambda / h^2) (1 + h^2/m_0^2 c^2 \lambda^2)^{1/2} \quad (16)$$

where $\beta = v/c$. This relativistic interaction factor σ is seen to tend to a constant value as E increases, the limit corresponding to the non-relativistic value for $\lambda = 0.02426$, the Compton wavelength, or $E = 212$ keV. The relativistic correction for the wavelength, on the other hand, is seen from (9) to make the decrease of λ with E more rapid.

A form very close to (15), but somewhat more accurate at higher angles, is known in nuclear scattering theory as the "Molière high-energy approximation". It is derived from the general partial wave theory of scattering from a central force field by making a small angle approximation (see, for example, Wu and Ohmura [1962]).

It has been shown by Doyle [1969] by detailed calculation that this approximation and (15) are good for most atoms over the range of scattering angles normally used in diffraction experiments with solids.

If $\sigma\varphi(x, y)$ is small so that the exponential in (15) may be written

$$\exp\{-i\sigma\varphi(x, y)\} \approx 1 - i\sigma\varphi(x, y), \quad (17)$$

then the scattering is given by

$$\Psi(u, v) = \delta(u, v) - i\sigma\Phi(u, v). \quad (18)$$

The delta function represents the transmitted, undiffracted beam and $\Phi(u, v)$ is equal to a planar section of $f_e(u)$ defined by (11).

4.2.4. Failure of First Born Approximation

Numerical evaluation shows that, except for very light atoms, the assumption that $\sigma\phi(x, y)$ is much smaller than unity is not justified. Then the First Born Approximation is no longer adequate. The effect of this failure may be estimated by considering the second Born Approximation (Schomaker and Glauber [1952], Glauber and Schomaker [1953]) or by invoking the partial-wave scattering theory (Hoerni and Ibers [1953]). However, the essential points are seen more immediately and graphically by use of (15), which may be written

$$\begin{aligned}\psi(u, v) &= \Im[1 + \{\cos\sigma\phi(x, y) - 1\}] - i\Im\{\sin\sigma\phi(x, y)\} \\ &= \delta(u, v) + \Im[\cos\sigma\phi - 1] - i\Im[\sin\sigma\phi].\end{aligned}\quad (19)$$

or, to make the analogy with (10) closer

$$\psi(u, v) = \delta(u, v) - i\{\Im\{\sin\sigma\phi\} + i\Im\{\cos\sigma\phi - 1\}\},\quad (20)$$

and the atomic scattering factor $f_e(u, v)$ is the complex function in the square brackets, divided by σ .

If $\sigma\phi$ is not too large it may be useful to expand the sine and cosine terms and take the Fourier transforms so that we get a real part of $f_e(u, v)$ equal to

$$\Phi(u, v) - (\sigma^2/3!)\{\Phi(u, v) * \Phi(u, v) * \Phi(u, v)\} + \dots,\quad (21)$$

and an imaginary part equal to

$$(\sigma/2!)\{\Phi(u, v) * \Phi(u, v)\} - (\sigma^3/4!)\{\Phi * \Phi * \Phi * \Phi\} + \dots,\quad (22)$$

Thus the real part is equal to $\Phi(u, v)$, the planar section of $f_e(u)$ given by (11), plus higher order terms. The imaginary term is given approximately by the two-dimensional self-convolution of $f_e(u, v)$ and so increases more rapidly than f_e with atomic number and falls off more slowly than the real part with $|u|$ or $(\sin\theta)/\lambda$.

If the complex atomic scattering factor is written

$$f(u) = |f_e(u)| \exp\{i\eta(u)\},\quad (23)$$

then the phase angle η increases with λ , with atomic number and with scattering angle. For the uranium atom, for example, $\eta = 0.29$ radians for $(\sin\theta)/\lambda = 0$ and $\eta = 2.4$ rad for $(\sin\theta)/\lambda = 1.15 \text{ \AA}^{-1}$ when $E = 39.5$ keV.

The effects of the failure of the first Born approximation for electron scattering from atoms were first recognized as such in the explanation by Schomaker and Glauber [1952] of the intensities of diffraction from gaseous molecules of UF_6 . Tables giving the values of the amplitudes and phases of the complex atomic scattering factors, used in the calculation of intensities for gas diffraction have been prepared by Ross, Fink & Hilderbrandt (International Tables for Crystallography, Vol. C (Wilson, 1992)).

It was shown later by Gjønnes [1964] and Bunyan [1963] that it may not be an adequate approximation for some gas molecules to merely replace the first Born approximation atom scattering factors by the complex scattering factors (23). In terms of the phase object approximation, (14), it is seen that if $\varphi(x,y)$ is the projected potential distribution for a molecule, the values of the real and imaginary components, (21) and (22), will depend on whether or not the atoms overlap in projection. If two atoms overlap, the contributions to $\Phi(u,v)$ will be multiplied by 2, but the contributions to the second and third order terms will be multiplied by 4 and 8. Bartell [1975] has shown that this effect can be described by an extension of Glauber's theory which provides a convenient basis for calculations.

The effects of the overlapping of atoms in projection is much more important in solids. The "pseudo-kinematical" theory proposed by Hoerni [1956], in which complex atom scattering factors replace the real ones in kinematical intensity formulas, has a very limited range of validity. It is appropriate only for "solids" consisting of single layers of atoms perpendicular to the incident beam. Coherent multiple scattering, or "dynamical scattering", must be taken into account for most experiments by use of the special theoretical treatments to be described in later chapters.

4.2.5. "Absorption" effects

As in the case of X-ray scattering, the presence of absorption gives the effect of a complex refractive index and so a complex scattering potential for electrons. In the simple formulation of (14), absorption modifies the transmission function of an atom to give

$$\begin{aligned} q(x,y) &= \exp\{-i\sigma\varphi(x,y) - \mu(x,y)\} \\ &= \exp\{-i[\sigma\varphi(x,y) + i\mu(x,y)]\}. \end{aligned} \quad (24)$$

Then the atomic scattering factor for elastic scattering is convoluted by

$$\begin{aligned} \mathfrak{S}\{\exp\{-\mu(x,y)\}\} &= \delta(u,v) - \mathfrak{S}\{1 - \exp\{-\mu(x,y)\}\} \\ &= \delta(u,v) - M(u,v) + \frac{1}{2}\{M(u,v) * M(u,v)\} + \dots \end{aligned} \quad (25)$$

where $M(u, v)$ is the Fourier transform of $\mu(x, y)$.

Then, in place of the bracketted part of (20), we obtain

$$f_e(u, v) = \Im\{\sin\sigma\phi\} + i\Im\{\cos\sigma\phi - 1\} - iM(u, v) \\ - M(u, v) * \Im\{\sin\sigma\phi\} - iM(u, v) * \Im\{\cos\sigma\phi - 1\} + \dots, \quad (26)$$

It is to be noted that the effect of absorption in rendering the atomic scattering factor complex is essentially different from the effect of the failure of the single-scattering approximation since for the latter the object remains a pure phase object and no energy is lost.

The absorption function $\mu(x, y)$, which may be considered as the projection of a three-dimensional function $\mu(r)$, arises from any scattering process for which the scattering is experimentally distinguished from and not interacting coherently with the scattering which is of interest.

For isolated single atoms, the only appreciable contribution to the absorption originates from the excitation of the atomic electrons. Electrons from the incident beam which have undergone an inelastic scattering process involving this excitation have lost energy of the order of 10 eV and may be separated from the elastically scattered electrons by use of an energy analyser. Hence the inelastic scattering process may give the effect of an absorption function for elastic scattering.

For X-rays the main contribution to the absorption comes from the excitation of electrons from the inner shells of atoms and so, except in the immediate vicinity of the absorption edge, shows very little dependence on the association of atoms in molecules, liquids or solids. For electrons however, the most important inelastic scattering processes involve the outer electron shells and energy losses in the range 0 to 50 eV. The absorption coefficients are thus strongly dependent on the state of bonding or ionization of the atoms. For solids the most important contributions come from the excitation of collective electron oscillations (plasmons), the production of excited states of the crystal electrons and the excitation of lattice vibrations (phonons).

Again, for X-rays, the absorption processes subtract most of the energy from the incident photons involved so that they no longer contribute to measured intensities. This is not the case for electrons. The energy losses of the incident electrons are frequently so small that they can not be detected except by the use of special devices, so that often no distinction is made experimentally between elastic and inelastic scattering. Furthermore, if only one sort of elastic scattering is considered, e.g. only the sharp Bragg scattering, then other elastic scattering as well as the inelastic scattering may be excluded from the measurement and will then contribute to the effective absorption coefficient.

Hence the absorption coefficients for electron scattering is strongly dependent not only on the state of aggregation of the atoms but also on the type of measurement made and the equipment used in the measurement. We therefore defer detailed discussion of the absorption coefficients for electron scattering until Chapter 12, after we have considered the elastic scattering of electrons from solids in more detail. At this stage we remark only that the imaginary part of the effective complex potential, $\mu(x,y)$ in (24) is usually smaller than the real part $\sigma\phi(x,y)$ by a factor of from 5 to 50, and its effect may often be treated as a perturbation on the purely elastic scattering case.

4.3. Neutrons

4.3.1. Atomic scattering factors

The neutrons of interest for diffraction experiments are usually the thermal neutrons which are in thermal equilibrium with the atoms in a nuclear reactor and so have an average energy of about 0.025 eV for which the corresponding wavelength is about 1.5 Å. However the spread of energies or wavelengths in the beam of neutrons obtained from a reactor is quite broad and for diffraction experiments a narrow range of wavelengths is usually selected out by use of a crystal monochromator or, especially for long wavelengths, by a time-of-flight chopper device which selects a range of neutron velocities. Time-of-flight energy (and wavelength) analysis may also be used with the pulsed spallation sources in which fast neutrons are generated by bombarding heavy metal targets with high-energy proton beams from an accelerator. Small blocks of moderator material near the target serve to slow the neutrons to thermal velocities. For recent accounts of experimental arrangements for neutron diffraction see Craven [1987] and Majkrzak and Wood [1992].

Neutrons are not appreciably scattered by electrons. Their main interaction is with the nucleus. We quote here the results derived initially in the context of nuclear physics and retaining something of that flavor.

Since the nucleus of an atom is very much smaller than the wavelength of the thermal neutron, the atomic scattering factor for neutrons is isotropic, independent of scattering angle and represented by a single-valued "scattering length", b . Contributions to b include a "potential" scattering term ξ , corresponding to scattering from a hard sphere of appropriate radius and "resonance" scattering terms arising from reactions of the neutron with the nucleus. The Breit-Wigner formula for scattering from a zero-spin, isolated nucleus gives

$$b = \xi + \frac{\frac{1}{2}\Gamma_n(r)/k}{(E - E_r) + \frac{1}{2}i(\Gamma_n(r) + \Gamma_a(r))}, \quad (27)$$

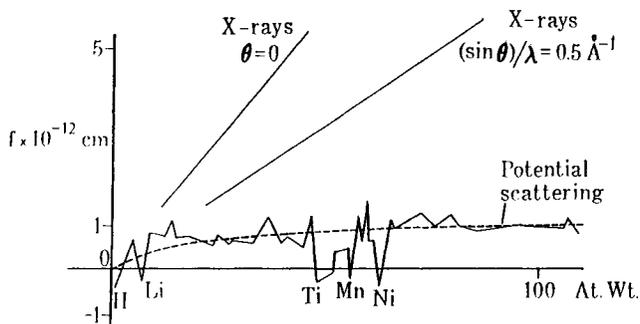


Fig. 4.3. Variation of the atomic scattering factor for neutrons, b , and the atomic scattering factor for X-rays (at particular values of $(\sin\theta)/\lambda$) with atomic weight (after Bacon [1975]).

where E_r is the energy for resonance, $\Gamma_n^{(r)}$ is the width of the resonance for re-emission of the neutron with its original energy and $\Gamma_a^{(r)}$ is the width of the resonance for absorption. More usually the scattering is described in terms of a scattering cross-section,

$$\sigma = 4\pi b^2. \quad (28)$$

It is seen from (27) that, depending on the interactions with the nucleus, b may be positive or negative, real or complex. In practice ξ is relatively small, and being proportional to the nuclear radius, increases as the one-third power of the atomic number. The imaginary part is appreciable for only a few isotopes such as those of boron and cadmium which are strong absorbers of thermal neutrons. For a number of isotopes, including the important cases of ^1H , ^7Li , ^{48}Ti , ^{51}V , ^{53}Mn , and ^{62}Ni , b is negative. Tables of the values of b and of scattering cross-sections are given, for example by Bacon [1975] and the International Tables for Crystallography, Vol. C (Wilson, 1992). The variation with atomic number is compared with that for X-rays in Fig. 4.3.

4.3.2. Nuclear spin scattering

For neutron scattering the nuclear spin is also of importance. For a nucleus with spin, I , two scattering lengths b_+ and b_- must be defined corresponding to the formation of compound nuclei of spin $I + 1/2$ and $I - 1/2$ with relative probabilities w_+ and w_- respectively. For an assembly of atoms, or for the time average of the scattering of neutrons from a single atom, the two states contribute at random, subject to the probabilities w_+ and w_- . Then, following the

concepts to be developed in Chapter 7, the scattering consists of "coherent" scattering from the average of the two states with a cross-section

$$\sigma_c = 4\pi(w_+b_+ + w_-b_-)^2, \quad (29)$$

and "incoherent" scattering, given by the difference between σ and the average cross-section, as

$$\sigma_{inc} = \sigma - \sigma_c = 4\pi(w_+b_+^2 + w_-b_-^2) - 4\pi(w_+b_+ + w_-b_-)^2, \quad (30)$$

The coherent scattering cross-section represents the scattering giving rise to diffraction effects. The incoherent scattering cross-section contributes a uniform background with no diffraction effects. This is particularly important for hydrogen for which $I = 1/2$, $b_+ = 1.04 \times 10^{-12}\text{cm}$, $b_- = -4.7 \times 10^{-12}\text{cm}$, so that $\sigma = 81 \times 10^{-24}\text{cm}^2$ and $\sigma_c = 2 \times 10^{-24}\text{cm}^2$. Then most of the scattering is "incoherent" background. In what follows, however, we will ignore this background scattering and consider only the b and σ_c values appropriate for the "coherent" scattering.

In the above, we have used the terms "coherent" and "incoherent" in the incorrect manner which, unfortunately, has become conventional. It must be emphasized that the results (29) and (30) are derived by considering the coherent scattering of radiation from a disordered array of spins. Coherence refers to the properties of the incident radiation. The order or disorder of the scattering material is properly described in terms of correlation functions. Thus (29) refers to coherent scattering from the correlated part of the scattering material distribution, and (30) refers to the coherent scattering from the non-correlated part of the scattering material distribution.

4.3.3. Isotopic disorder

For assemblies of atoms a further complication is that, for most elements the atoms present consist of a random array of the various isotopes, present either with their natural abundance or sometimes with artificial enrichment of particular isotopes. Since neutron scattering depends on the reactions of the neutrons with the nucleus, the b value may vary widely from one isotope to another. For example, b for ^{58}Ni is $1.44 \times 10^{-12}\text{cm}$, for ^{60}Ni it is $0.30 \times 10^{-12}\text{cm}$ and for ^{62}Ni it is $-0.87 \times 10^{-12}\text{cm}$, giving an average b for naturally occurring Ni of $1.03 \times 10^{-12}\text{cm}$.

Following the considerations of disordered variation of scattering factors given in Chapter 7 we see that for diffraction effects from the correlated component of the

material (e.g. for giving Bragg diffraction maxima) the effective scattering factor is given by the average value

$$b = \sum_n w_n b_n \quad (31)$$

where w_n is the relative abundance of the n th isotope. The difference between scattering from this average and the total scattering is the diffuse background scattering which has intensity proportional to

$$\begin{aligned} I_{\text{diff}} &= I_{\text{tot}} - |b|^2 \\ &= \sum_n w_n b_n^2 - \left| \sum_n w_n b_n \right|^2. \end{aligned} \quad (32)$$

Again, the portions of the scattering represented by (31) and (32) are often, inappropriately, referred to as "coherent" and "incoherent".

4.3.4. Thermal and magnetic scattering

Two further aspects of neutron scattering are of considerable importance for practical purposes of solid-state investigations and will be treated in context in later chapters.

Firstly, because the energies of the incident neutrons are approximately the same as the thermal energies of vibration of the scattering atoms, the energy changes involved with the excitation of thermal lattice vibrations, or the creation or annihilation of phonons, are important and can be measured to provide information on phonon energies as well as momentums. Alternatively one can consider that the velocity of the thermal neutrons is so low that the diffraction process in matter is sensitive to the variation of atom positions in time so that the Fraunhofer diffraction pattern gives information concerning the relative positions of atoms in both space and time.

Secondly, because of their spin, the neutrons can interact with unpaired electron spins as well as with nuclear spins. For the near-random spin arrays of paramagnetic material this gives rise to diffuse background scattering, falling off with increasing scattering angle rather more rapidly than the X-ray scattering from an atom because the scattering is done by the electrons having unpaired spins, which are usually confined to the outer electron shells. When the electron spins are ordered in parallel or anti-parallel array as in ferromagnetic and anti-ferromagnetic materials the spin scattering gives rise to appropriate diffraction maxima. Hence neutron diffraction has become a major tool for the study of the magnetic properties of materials.

Problems

1. Estimate the magnitude of the correction to X-ray atomic scattering factors due to the inclusion of multiple-scattering terms (e.g. the second Born approximation).

2. Will Compton scattering give rise to an absorption effect for X-rays? If so, estimate its magnitude.

3. By use of the Mott formula consider the effect of ionization of an atom on its electron atomic scattering factor as the scattering angle tends to zero. What do the infinities imply? For atoms in solids, the formulas derived for isolated atoms do not apply. Suggest how scattering from ions may be treated in this case. (Then see Doyle and Turner [1968].)

Scattering from assemblies of atoms

5.1. The kinematical approximation

The first Born approximation for scattering from a three-dimensional distribution otherwise known as the "kinematical" or "single scattering" approximation, is given in equation (1.20). It is not limited in its application to the scattering from single atoms but may be applied to any collection of scattering matter. Normally we think in terms of assemblies of distinguishable atoms although for X-rays when $\varphi(\mathbf{r})$ is replaced by the electron density distribution, $\rho(\mathbf{r})$, the modifications of the electron distribution due to bonding may make it difficult to give a meaningful assignment of separate components of $\rho(\mathbf{r})$ to separate atoms. For electrons when the $\varphi(\mathbf{r})$ of (1.20) becomes the electrostatic potential distribution, $\varphi(\mathbf{r})$, the assignment to atoms may be even more difficult especially when the scattering involves, in the usual theoretical approximation, excitation of a whole crystal from one state to another i.e. the transfer from one non-localized wave function of the crystal electrons to another. However these limitations are important only for special considerations and will be treated separately when the need arises.

We will proceed under the assumption that the electron density of a collection of atoms may be written

$$\rho(\mathbf{r}) = \sum_i \rho_i(\mathbf{r}) * \delta(\mathbf{r} - \mathbf{r}_i) \quad (1)$$

where $\rho_i(\mathbf{r})$ is the electron density associated with the atom centered at $\mathbf{r} - \mathbf{r}_i$ and is not necessarily assumed to be the same as for a free atom. For convenience we use the notation appropriate to X-ray diffraction with the understanding that exactly the same considerations apply for electron and neutron scattering within the limits of applicability of the single-scattering approximation.

Validity of the single-scattering approximation implies that the amplitude of the single scattered radiation is very small compared with the incident beam amplitude. Then the amplitude of the doubly- and multiply-scattered radiation is very small and negligible compared with the singly scattered.

The amplitude of a scattered beam in a particular direction depends to a very great extent on the possibility of cooperative scattering from an ordered array of atoms. If, as in the case of X-rays and neutrons, the interaction with atoms is so weak that, in a crystal, the diffracted energy can be concentrated into one or two sharply defined directions by three dimensional diffraction before the incident beam has lost much energy, then we may consider the possibility of multiple scattering of well-defined beams. A diffracted beam will be diffracted again if it passes through another crystal region which is set at the right angle for Bragg reflection. This condition is always ensured in a large perfect single crystal, but becomes less probable in the presence of crystal defects, grain boundaries and so on. For a perfect crystal and a strong crystalline reflection, multiple scattering becomes appreciable for X-rays for path lengths of the order of $1 \mu\text{m}$. For neutrons the necessary path lengths are several times greater. If the atoms are not sufficiently ordered to give well-defined diffracted beams, the diffracted intensity in any direction is much less and multiple scattering effects are correspondingly less important.

On the other hand, we have seen that for electrons the interaction with atoms is much stronger so that multiple scattering may be important in the scattering from a single heavy atom. Well-defined diffracted beams are not generated before the amplitude of scattering is appreciable. Hence multiple scattering effects become important within a distance which is of the order of one or two hundred Å for light atoms and less for heavy atoms, and are almost as strong for non-crystalline as for crystalline specimens.

It is worthwhile for any radiation to draw the distinction between "coherent" and "incoherent" multiple scattering. Here the word "coherent" is used, in accordance with a common terminology, to refer not to the incident radiation but to the correlation of the atomic positions and so to the relationships of phases of scattered radiation. It would be more appropriate to refer to coherent multiple scattering from "correlated" and "un-correlated" arrays of atoms.

At one extreme, the positions of the atoms in a perfect crystal are well correlated in that they may all be related to the sites of a periodic lattice. Then the relative phases of waves scattered by any atoms may be uniquely defined and the amplitudes of the waves may be added. The correlated multiple scattering is then referred to as "dynamical" scattering.

In the other extreme, the correlation of atomic positions does not extend over the distances required to generate appreciable scattered amplitude. The relative phases of waves scattered by strongly correlated groups of atoms varies at random from one group to another throughout the crystal. Then multiple scattering intensities are added

incoherently. This case has been referred to as "multiple elastic scattering". Obviously in practice any intermediate situation, or combination of situations, may exist so that an adequate description of the scattering may be very complicated.

Fortunately, for many specimens of interest, the scattering may be described in terms of one of the two extreme cases, pure kinematical or pure dynamical diffraction, or else in terms of small perturbations of one or the other. For other specimens the physical dimensions or state of crystallinity can often be modified to simplify the scattering theory.

5.2. Real and reciprocal space

5.2.1. Reciprocal space distribution

We have seen that kinematical scattering amplitudes may be expressed in terms of the Fourier transform of a distribution in real or direct space. In real space we consider a position vector \mathbf{r} with coordinates x, y, z . In reciprocal space we consider a position vector \mathbf{u} with coordinates u, v, w . Then, in X-ray terminology, a distribution $\rho(\mathbf{r})$ in real space is related to a distribution $F(\mathbf{u})$ in reciprocal space by the Fourier transform,

$$F(\mathbf{u}) = \int \rho(\mathbf{r}) \exp\{2\pi i \mathbf{u} \cdot \mathbf{r}\} d\mathbf{r}, \quad (2)$$

or

$$F(u, v, w) = \iiint \rho(x, y, z) \exp\{2\pi i (ux + vy + wz)\} dx dy dz. \quad (3)$$

If an electron density distribution, $\rho(\mathbf{r})$, is considered to be the sum of the distributions $\rho_i(\mathbf{r})$ attributed to individual atoms as in equation (1) and if

$$f_i(\mathbf{u}) \equiv \int \rho_i(\mathbf{r}) \exp\{2\pi i \mathbf{u} \cdot \mathbf{r}\} d\mathbf{r}, \quad (4)$$

then Fourier transform of (1) gives

$$F(\mathbf{u}) = \sum_i f_i(\mathbf{u}) \exp\{2\pi i \mathbf{u} \cdot \mathbf{r}_i\}. \quad (5)$$

If it is further assumed that the electron density $\rho_i(\mathbf{r})$ is sufficiently close to the electron density of an isolated free atom, the value of $F(\mathbf{u})$ can be found from the tables of atomic scattering factors given, for example, in the International Tables for Crystallography, Vol. C (Wilson, 1992).

The form of $F(u)$ corresponding to various forms of the function $\rho(r)$ may be seen by an extension to three-dimensions of the relationships and examples of Fourier transforms given in Chapter 2. For example, by extension of (2.38) and (2.42), if $\rho(r) = 1$ inside a rectangular box of dimensions a, b, c and zero outside, then

$$F(u) = abc \frac{\sin(\pi au)}{\pi au} \frac{\sin(\pi bv)}{\pi bv} \frac{\sin(\pi cw)}{\pi cw}. \quad (6)$$

This is the three dimensional analogue of (2.42), with a central peak of height abc and falling off through diminishing oscillations along each axis.

In this way we define a "shape function" and the "shape transform" often used to describe rectangular volumes of material.

5.2.2. The reciprocal lattice

By extension to three dimensions of our results for a diffraction grating, we see that, corresponding to (2.51) we have that, if

$$\rho(r) = \sum_n \sum_m \sum_p \delta(x - na, y - mb, z - pc),$$

then

$$F(u) = \sum_h \sum_k \sum_l \delta(u - (h/a), v - (k/b), w - (l/c)). \quad (7)$$

Thus for a periodic lattice in real space with periodicities a, b, c the corresponding distribution in reciprocal space is a lattice of points with periodicities a^{-1}, b^{-1}, c^{-1} . This is the "reciprocal lattice" for the special case of rectangular axes.

Corresponding to (2.50) and (2.51) we have the relationship, that, if $\rho(r)$ is a periodic function expressed by the Fourier series

$$\rho(r) = \sum_h \sum_k \sum_l F_{hkl} \exp\{-2\pi i((hx/a) + (ky/b) + (lz/c))\}, \quad (8)$$

then

$$F(u) = \sum_h \sum_k \sum_l F_{hkl} \delta(u - (h/a), v - (k/b), w - (l/c)), \quad (9)$$

which represents a reciprocal lattice with each reciprocal lattice point weighted by the Fourier coefficient.

For a periodic function in real space chopped off by rectangular shape-function having dimensions A, B, C , we have from (2.59), (6) and (9),

$$F(\mathbf{u}) = \sum_h \sum_k \sum_l F_{hkl} \delta(u - (h/a), v - (k/b), w - (l/c)),$$

$$* ABC \frac{\sin(\pi Au)}{\pi Au} \frac{\sin(\pi Bv)}{\pi Bv} \frac{\sin(\pi Cw)}{\pi Cw}, \quad (10)$$

which implies that each point of the weighted reciprocal lattice, (9), is spread out into a continuous distribution given by a shape transform of the form (6).

5.2.3. Friedel's law and the phase problem

There are several relations of importance which follow from the definitions of Chapter 2.

Provided that no absorption effect is important, $\rho(\mathbf{r})$ may be assumed to be a real function. Then

$$F(-\mathbf{u}) = \int \rho(\mathbf{r}) \exp\{2\pi i(-\mathbf{u}) \cdot \mathbf{r}\} d\mathbf{r}$$

$$= \int \rho(\mathbf{r}) \exp\{-2\pi i\mathbf{u} \cdot \mathbf{r}\} d\mathbf{r}$$

$$= F^*(\mathbf{u}), \quad (11)$$

so that

$$|F(-\mathbf{u})|^2 = |F(\mathbf{u})|^2.$$

Then it follows from (2.22) that

$$|\Im\rho(-\mathbf{r})|^2 = |\Im\rho(\mathbf{r})|^2. \quad (12)$$

Since the intensity of diffraction is proportional to $|F(\mathbf{u})|^2$, this is a convenient statement of Friedel's law which implies that inversion of a crystal through a center of symmetry does not change the diffraction intensities in a kinematical approximation.

The inversion of the relationship (2) gives

$$\rho(\mathbf{r}) = \int F(\mathbf{u}) \exp\{-2\pi i\mathbf{u} \cdot \mathbf{r}\} d\mathbf{u}. \quad (13)$$

If the diffraction amplitudes $\psi(l, m, n)$ could be measured so that $F(\mathbf{u})$ could be derived, then the distribution $\rho(\mathbf{r})$ could be deduced by numerical evaluation of this integral. However the measurement of wave amplitudes is not possible for the radiations we consider. Only the intensities given by $\psi\psi^*$ can be recorded. Thus information on relative phases of

the diffracted beams is lost and the function $\rho(\mathbf{r})$ cannot be deduced directly.

For most experimental situations complete knowledge of $\rho(\mathbf{r})$ is neither necessary nor desirable since it would involve knowledge of the relative positions and bonding of all the atoms of the sample and represent a quantity of data which would be difficult to handle. We therefore proceed to investigate the type of information which may be derived directly from the observable intensities.

5.3. The generalized Patterson function

The observable intensity of diffracted radiation is related in a known way to $\psi\psi^*$ and so to $|F(\mathbf{u})|^2$. This function, $|F(\mathbf{u})|^2$, sometimes written $J(\mathbf{u})$, is a function in reciprocal space. It is a real positive function and, from (11), is centro-symmetric about the reciprocal space origin. The form of the function for particular cases is readily derived from the relations for $F(\mathbf{u})$ given, for example, in equations (6) to (9).

By inverse Fourier transform we see that the corresponding function in real space is

$$\begin{aligned} \mathfrak{F}^{-1} \{ |F(\mathbf{u})|^2 \} &\equiv P(\mathbf{r}) \\ &= \rho(\mathbf{r}) * \rho(-\mathbf{r}) \\ &= \int \rho(\mathbf{R}) \rho(\mathbf{r} + \mathbf{R}) d\mathbf{R}. \end{aligned} \quad (14)$$

This function of real space, $P(\mathbf{r})$, we call the generalized Patterson function to distinguish it from the Patterson function employed for crystal structure analysis and usually referring to periodic structures only. When no chance of confusion exists, we will call it simply the "Patterson". It is similar to the Q -function of Hosemann and Baggchi [1962]. It may also be called a "self-correlation" function, as we shall see. It is the real space function directly derivable from the observable intensities.

The greatest value of $P(\mathbf{r})$ occurs at the origin, $\mathbf{r} = 0$, since

$$P(0) = \int \rho^2(\mathbf{R}) d\mathbf{R} = \int |F(\mathbf{u})|^2 d\mathbf{u}, \quad (15)$$

from (2.66). It tends to have a relatively large value when \mathbf{r} is a vector which connects points having high $\rho(\mathbf{r})$ values.

When $\rho(\mathbf{r})$ represents the electron density for a collection of atoms, we write, as before,

$$\rho(\mathbf{r}) = \sum_i \rho_i(\mathbf{r}) * \delta(\mathbf{r} - \mathbf{r}_i). \quad (16)$$

Then

$$\begin{aligned} \rho(r) * \rho(-r) &= \sum_i \sum_j \rho_i(r) * \rho_j(-r) * \delta(r - r_i) * \delta(r + r_j) \\ &= \sum_i \sum_j [\rho_i(r) * \rho_j(-r)] * \delta\{r - (r_i - r_j)\}. \end{aligned} \quad (17)$$

Since $\rho_i(r)$ is a peaked function, roughly represented by a gaussian, such that $\int \rho_i(r) dr = Z_i$, the convolution $\rho_i * \rho_j$, will represent a slightly broader gaussian-like peak of integrated weight $Z_i Z_j$. This is placed at the end of the interatomic vector $r_i - r_j$ relative to the origin of $P(r)$. There is an identical peak at $r_j - r_i$. Thus $P(r)$ is made up of peaks corresponding to all interatomic vectors present. The weight of each peak corresponds to the sum of all products $Z_i Z_j$ for the pairs of atoms having that interatomic vector. This is illustrated for simple two- and three- and four-atom objects in Fig. 5.1.

The peak at the origin of $P(r)$ corresponds to all vectors between centers of atoms of length zero and so is

$$\delta(r) * \sum_i [\rho_i(r) * \rho_i(-r)],$$

and has weight $\sum_i Z_i^2$.

If there are two vectors of the same length and direction in the object, as in Fig. 5.1(c), the corresponding Patterson peaks are twice as large. Thus the Patterson function may be regarded as a mapping of the weighted probability that two atoms are found a distance r apart, the weighting being according to the atomic numbers. Or we may say that $P(r)$ gives the probability that if any one atom is taken as origin, there will be another atom at a vector distance r from it. Thus it is, in essence, a correlation function giving the spatial correlation of electron densities.

Another way of looking at $P(r)$ is illustrated in Fig. 5.2, where we have redrawn the Patterson of the three-atom structure of Fig 5.1(b) in two ways to emphasize that it may be regarded as a superposition of "images" of $\rho(r)$ or its inverse, $\rho(-r)$. We may consider each of the atoms in turn to be placed at the origin and the resulting "images" added with a weighting factor given by the Z for the atom at the origin. Alternatively, we may consider the distribution $\rho(r)$ to be translated so that one of its atoms is placed, in turn, at each of the atom positions of $\rho(-r)$, the points marked by P in the diagram.

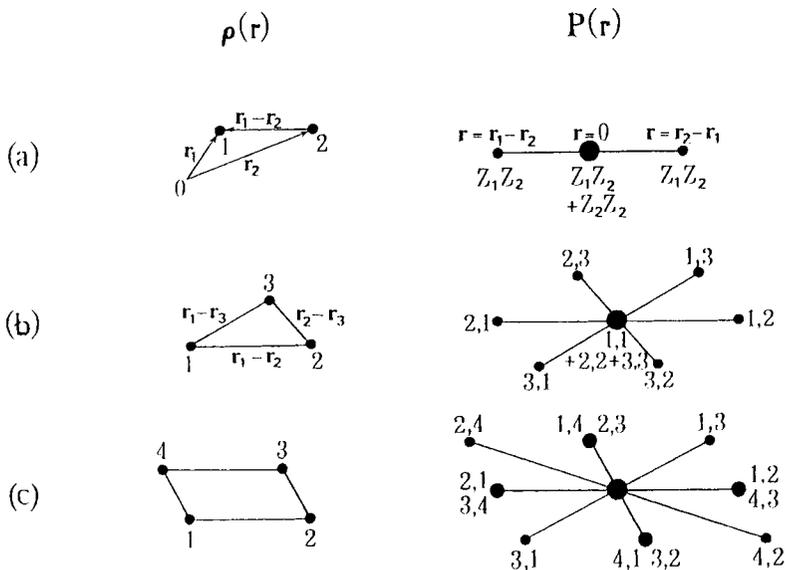


Fig. 5.1. Diagrams to suggest the form of the electron density distribution, $\rho(\mathbf{r})$, and the Patterson function, $P(\mathbf{r})$ for (a) two atoms, at positions defined by vectors \mathbf{r}_1 and \mathbf{r}_2 ; (b) three atoms, and (c) four atoms forming a parallelogram.

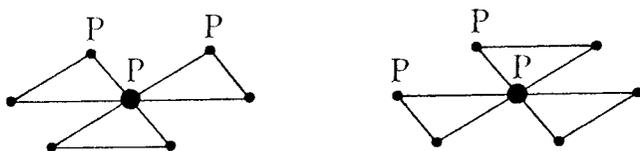


Fig. 5.2. Representation of the Patterson function for three atoms, Fig. 5.1(b), as a superposition of images of the object or its inverse.

For systems composed of a very large number, N , of atoms, the Patterson function has $N^2 - N + 1$ peaks (since N peaks are superimposed at the origin), and so is too complicated to be interpreted in a statistical sense as a probability-density or self-correlation function. We now consider a few representative cases to illustrate the application of these concepts.

5.4. Examples of correlation functions

5.4.1. Finite volume limitations

A distribution of atoms over a finite volume is often described in terms of a function representing an infinite distribution $\rho'(\mathbf{r})$ multiplied by a shape-function $s(\mathbf{r})$ i.e.

$$\rho(\mathbf{r}) = \rho'(\mathbf{r}) s(\mathbf{r}).$$

In reciprocal space,

$$F(\mathbf{u}) = F'(\mathbf{u}) * S(\mathbf{u}).$$

If $s(\mathbf{r})$ has dimensions which are very large compared to atomic dimensions so that $S(\mathbf{u})$ is a very narrow function compared with the variations of $F'(\mathbf{u})$ which are of interest, then it is a reasonable approximation to put

$$|F(\mathbf{u})|^2 = |F'(\mathbf{u})|^2 * |S(\mathbf{u})|^2. \quad (18)$$

The most obvious error in this assumption comes where $F'(\mathbf{u})$ changes sign, when $|F'(\mathbf{u}) * S(\mathbf{u})|^2$ has zero values but $|F'(\mathbf{u})|^2 * |S(\mathbf{u})|^2$ does not. Also, $|F'(\mathbf{u}) * S(\mathbf{u})|^2$ has a greater region of near-zero values, but the region affected has approximately the half-width of $S(\mathbf{u})$.

Fourier transform of (18) gives the corresponding assumption for real space,

$$P(\mathbf{r}) = P'(\mathbf{r}) [s(\mathbf{r}) * s(-\mathbf{r})], \quad (19)$$

i.e. the effect of spatial limitation of the sample is to multiply the Patterson for the infinite crystal (the Patterson "per unit volume") by the self-convolution of the shape transform. In one dimension, if $s(x)$ is the slit function given by

$$s(x) = \begin{cases} 0 & \text{if } |x| \leq A/2, \\ 1 & \text{if } |x| > A/2. \end{cases}$$

Then

$$s(x) * s(x) = \begin{cases} x + A & \text{if } -A < x < 0, \\ A - x & \text{if } 0 < x < A, \\ 0 & \text{elsewhere,} \end{cases}$$

which is the Fourier transform of $|S(\mathbf{u})|^2$; see Fig. 5.3.

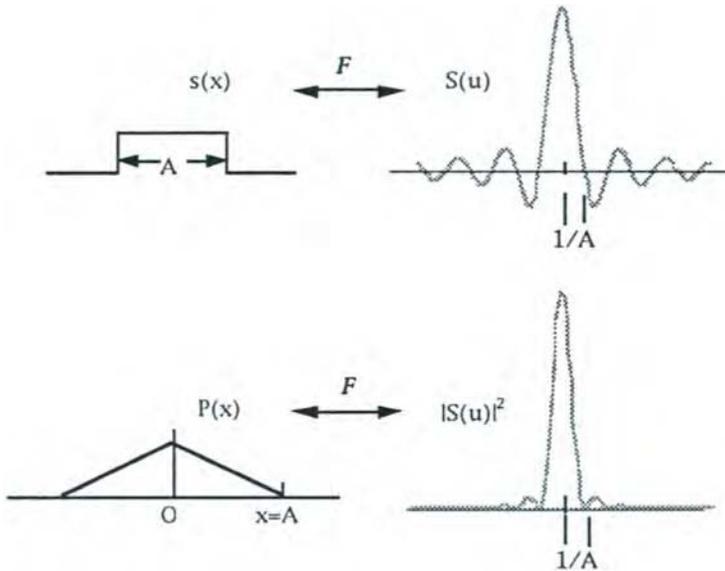


Fig. 5.3. Diagram of a shape function $s(x)$ and its Patterson function $P(x)$, and the corresponding Fourier transforms.

Similarly in three dimensions, if the sample is a rectangular block limited by planes $x = \pm A/2$, $y = \pm B/2$, $z = \pm C/2$, where the dimensions A, B, C are much greater than the dimensions of the groups of atoms of interest, the Patterson per unit volume is multiplied by a factor $ABC = V$ at the origin and decreases linearly to zero in a distance A in the x direction, B in the y direction, C in the z direction, corresponding to the fact that no interatomic vector can have x, y, z components greater than A, B, C .

Usually, on the assumption that A, B and C are very much greater than any interatomic vectors of interest, this "shape-convolution" function is ignored or omitted.

5.4.2. Finite crystals

As a simple model, we consider a crystal having one atom in a rectangular unit cell which has axes $\mathbf{a}, \mathbf{b}, \mathbf{c}$. The crystal has dimensions $A = N_1 a$, $B = N_2 b$, $C = N_3 c$ and so can be represented by

$$[\rho_0(\mathbf{r}) * \sum_n \delta(\mathbf{r} - \mathbf{R}_n)] s(\mathbf{r}), \quad (20)$$

where $s(\mathbf{r})$ is the shape function having dimensions A, B, C and $\mathbf{R}_n = n_1 \mathbf{a} + n_2 \mathbf{b} + n_3 \mathbf{c}$.

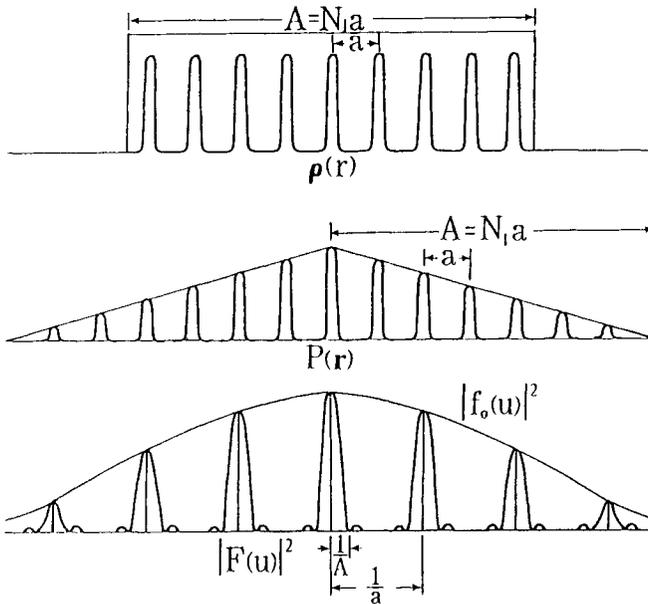


Fig. 5.4. Representation of a finite crystal as the product of a periodic object and a shape function, the corresponding Patterson function and its Fourier transform.

Then, following the considerations of 5.4.1, the Patterson function may be written,

$$\rho_0(r) * \rho_0(-r) * \left[\sum_n \sum_m \delta(r - R_n + R_m) \right] [s(r) * s(-r)], \quad (21)$$

$$= N\rho_0(r) * \rho_0(-r) * \left[\sum_n \delta(r - R_n) \right] [s(r) * s(-r)], \quad (22)$$

since the convolution of the two sets of N delta functions gives N times the set of all vectors between lattice points multiplied by the shape convolution, as suggested in the one-dimensional example of Fig. 5.4.

Then the Fourier transform of (25) gives the corresponding distribution in reciprocal space as

$$|F(u)|^2 = N|f_0(u)|^2 \sum_{h,k,l} \delta(u - (h/a), v - (k/b), w - (l/c)) * |S(u)|^2, \quad (23)$$

which is the reciprocal lattice having spacings a^{-1} , b^{-1} , c^{-1} with reciprocal lattice points weighted by the function $|f_0(u)|^2$, and with each reciprocal lattice point spread out by convolution with the function

$$|S(\mathbf{u})|^2 = A^2 B^2 C^2 \frac{\sin^2(\pi A u)}{(\pi A u)^2} \frac{\sin^2(\pi B v)}{(\pi B v)^2} \frac{\sin^2(\pi C w)}{(\pi C w)^2}, \quad (24)$$

which is the square of the function (6).

This case represents the most perfect ordering of atoms within a finite volume with the correlation function of atomic positions limited only by the shape-correlation. There are many stages of ordering intermediate between this and the case of an ideal gas representing the minimum of ordering. Several examples of intermediate degrees of order, particularly those approximating to the order of the perfect crystal, will be considered in later chapters.

5.5. Correlation in space and time

5.5.1. Four-dimensional Patterson

The concept of a Patterson function, which is a correlation function for the distribution of electron density in space, can be extended to include correlations in time. In space we ask; what is the probability that an atom will be separated from a given atom by a vector \mathbf{r} ? We may ask, similarly: if an atom is at the point \mathbf{r} at time 0, what is the probability that there will be an atom at that point (or at any other point) after a time t ?

The function which includes such correlations is the four-dimensional Patterson function

$$\begin{aligned} P(\mathbf{r}, t) &\equiv \int \rho(\mathbf{R}, T) \rho(\mathbf{r} + \mathbf{R}, t + T) d\mathbf{R}dT \\ &= \rho(\mathbf{r}, t) * \rho(-\mathbf{r}, -t). \end{aligned} \quad (25)$$

Fourier transformation gives a function in four-dimensional reciprocal space. The transform with respect to time, following equation (2.31), gives a function of frequency so we write

$$|F(\mathbf{u}, \nu)|^2 = \iint P(\mathbf{r}, t) \exp\{2\pi i(\mathbf{u} \cdot \mathbf{r} + \nu t)\} d\mathbf{r}dt. \quad (26)$$

Just as in the spatial relationship, \mathbf{u} is related to the vector \mathbf{q} of a diffraction experiment, giving the change in the wave vector \mathbf{k} , or in the momentum $\hbar\mathbf{k}$, so ν is related to the change in frequency from the incident frequency, or the change in energy $\hbar\nu$ of the incident photons (or particles).

This type of formalism, introduced by van Hove [1954], is of special value for neutron diffraction for which changes in both momentum and energy of the incident neutrons may be measured. For scattering from a crystal having thermal

vibration of the atoms or, more appropriately, a distribution of phonons, both the momentums and energies of the phonons may be deduced.

5.5.2. Special cases

A few special cases will serve to illustrate the properties and relationships of the function $P(\mathbf{r}, t)$. For the point $\mathbf{r} = 0$ we have

$$P(0, t) = \iint \rho(\mathbf{R}, T) \rho(\mathbf{R}, t + T) dT d\mathbf{R}. \quad (27)$$

The integral over T gives the probability that if there is an atom at \mathbf{R} at time 0 there will be an atom at \mathbf{R} at time t , and this is averaged over all points \mathbf{R} . This gives an indication of the rate of diffusion of atoms away from their initial positions.

For purely elastic scattering there is no change of frequency of the incident radiation so that $\nu = 0$. Then the reciprocal space function of interest is

$$\begin{aligned} |F(\mathbf{u}, 0)|^2 &= \iint P(\mathbf{r}, t) \exp\{2\pi i(\mathbf{u} \cdot \mathbf{r})\} d\mathbf{r} dt \\ &= \int [\int P(\mathbf{r}, t) dt] \exp\{2\pi i\mathbf{u} \cdot \mathbf{r}\} d\mathbf{r}. \end{aligned} \quad (28)$$

This is also equal to the Fourier transform of the three-dimensional Patterson function of the time average of the electron density.

Hence the section of the four-dimensional reciprocal space distribution, $\nu = 0$, corresponds to the Fourier transform of the time-average of the four-dimensional Patterson function.

If a measurement is made in which the scattered radiation intensities for all energies are added together, the reciprocal space function of interest is

$$\begin{aligned} \int |F(\mathbf{u}, \nu)|^2 d\nu &= \iint P(\mathbf{r}, t) \exp\{2\pi i(\mathbf{u} \cdot \mathbf{r})\} \delta(t) d\mathbf{r} dt \\ &= \int P(\mathbf{r}, 0) \exp\{2\pi i\mathbf{u} \cdot \mathbf{r}\} d\mathbf{r}. \end{aligned} \quad (29)$$

Hence the total scattered intensity, elastic plus inelastic, is related to $P(\mathbf{r}, 0)$, which gives the correlations of atomic positions for no difference in times; corresponding to the sum of all correlation functions for instantaneous "pictures" of the atomic configuration.

Expressions such as (28), however, are unsatisfactory as they stand because they lead to infinities. If $P(\mathbf{r}, t)$ refers to a system which is a constant, or repeats periodically in time, the integral $\int P(\mathbf{r}, t) dt$ is infinite. This unsatisfactory result must arise because we have not described the experimental system adequately.

Measurements of intensities are always made with instruments of finite resolution. The finite energy resolution may be taken into account by replacing $|F(u, v)|^2$ by, say,

$$|F(u, v)|^2 * (\pi M^2)^{-1/2} \exp\{-v^2/M^2\},$$

where M is a measure of the range of frequencies included. The Fourier transform of this function is then

$$P(r, t) \exp\{-2\pi^2 M^2 t^2\}.$$

Then the measurement of intensities for all v values is given in the limit $M \rightarrow \infty$, giving, by Fourier transform, $P(r, 0)$.

In the limit that M tends to zero we return to the ideal case of measuring $|F(u, v)|^2$ and obtaining $P(r, t)$ by Fourier transform.

The case where near-elastic scattering only is measured is represented by selecting only a small range of frequency around the origin:

$$\int |F(u, V)|^2 \exp\{-V^2/M^2\} dV.$$

The Fourier transform of this is

$$P(r, t) * (\pi^{1/2} M) \exp\{-2\pi^2 M^2 t^2\}.$$

In the limiting case $M \rightarrow 0$, corresponding to purely elastic scattering this integral tends to $\langle P(r, t) \rangle_t$, i.e. the time average of $P(r, t)$, since $P(r, t)$ is convoluted with a function which has a width tending to infinity but an integrated value always equal to unity.

In what follows we will not complicate the mathematics by using these more complete formulations, but will bear in mind that the difficulties arising from the over-simplified expressions may be avoided in this way.

5.5.3. *Ideal monatomic gas or liquid*

We consider the idealised case of N non-interacting atoms contained in a volume V such that the distribution of atom positions is completely random, i.e. there is no correlation of atom center positions except for the limitations of the finite volume. Each atom has an electron density distribution $\rho_0(r)$. We ignore the usual practical limitation that the electron distributions of separate atoms cannot overlap appreciably.

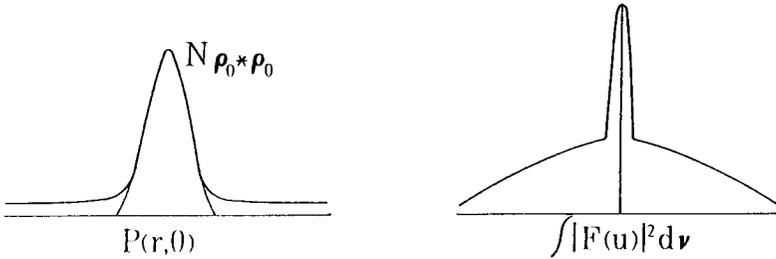


Fig. 5.5. Instantaneous correlation function, $P(\mathbf{r}, 0)$, for an ideal monatomic gas and the corresponding distribution in reciprocal space.

The form of the Patterson function $P(\mathbf{r}, t)$ can be written down immediately. The instantaneous correlation function $P(\mathbf{r}, 0)$ has a peak at the origin given by $N\rho_0(\mathbf{r}) * \rho_0(\mathbf{r})$. For any non-zero length vector the probability of finding an atom at that vector distance from a given atom has the same value which depends only on the gas or liquid density, apart from the finite volume effect discussed in 5.4.1. Hence we can write, as a reasonable approximation except near $\mathbf{r} = 0$,

$$\begin{aligned} P(\mathbf{r}, 0) &= (N/V) \int \{ \rho_0(\mathbf{r}) * \rho_0(\mathbf{r}) \} d\mathbf{r} [s(\mathbf{r}) * s(\mathbf{r})] \\ &= Z^2 (N/V) [s(\mathbf{r}) * s(-\mathbf{r})]. \end{aligned} \quad (30)$$

The form of $P(\mathbf{r}, 0)$ is then as shown in Fig. 5.5. This function could be derived from the scattering power distribution for all energies, $\int |F(\mathbf{u}, \nu)|^2 d\nu$, which consists of a central peak coming from the Fourier transform of (30) plus $N|f_0(\mathbf{u})|^2$ from the origin peak of $P(\mathbf{r}, 0)$. Thus the scattering is N times that for a single atom.

The correlation $P(\mathbf{r}, t)$ for \mathbf{r}, t not equal to zero is the same. For $\mathbf{r} = 0$, the value of $P(0, t)$ gives the probability that if an atom is at a point at time 0, there will be an atom there at time t . For large times this tends to the random value of (30). The rate at which $P(0, t)$ approaches this random value depends on the average velocity of the atom or the self-diffusion coefficient.

More formally and generally we may describe the electron density function as

$$\rho(\mathbf{r}, t) = \rho_0(\mathbf{r}) * g(\mathbf{r}, t)$$

where

$$g(\mathbf{r}, t) = \sum_n \delta\{\mathbf{r} - \mathbf{r}_n(t)\}. \quad (31)$$

Thus $g(\mathbf{r}, t)$ is a distribution function: a set of delta functions describing the positions of the atom centers in space and time. Then

$$P(\mathbf{r}, t) = \rho_0(\mathbf{r}) * \rho_0(-\mathbf{r}) * g(\mathbf{r}, t) * g(-\mathbf{r}, -t). \quad (32)$$

The convolutions are over space and time. Since $\rho_0(\mathbf{r})$ does not depend on time the first convolution is

$$\iint \rho_0(\mathbf{R}) \rho_0(\mathbf{r} + \mathbf{R}) d\mathbf{R} dT,$$

where the normalizing factor must be introduced as before so that the integral $\int dT$ does not give an infinity. The distribution function convolution

$$G(\mathbf{r}, t) = g(\mathbf{r}, t) * g(-\mathbf{r}, -t),$$

gives the probability that, given an atom centered $\mathbf{r} = 0$ at time $t = 0$, there will be an atom centered at the position \mathbf{r} at time t . Provided that we are dealing with a very large number of atoms, $G(\mathbf{r}, t)$ can be considered to be a continuous function.

Then in a diffraction experiment in which all frequencies are recorded with equal efficiency, the observation measures $\int |F(\mathbf{u}, \nu)|^2 d\nu$ from which one may deduce $P(\mathbf{r}, 0)$ which depends on

$$\begin{aligned} G(\mathbf{r}, 0) &= [g(\mathbf{r}, t) * g(-\mathbf{r}, -t)]_{t=0} \\ &= \int g(\mathbf{R}, T) g(\mathbf{r} + \mathbf{R}, T) d\mathbf{R} dT \\ &= \langle g(\mathbf{r}, t) * g(-\mathbf{r}, -t) \rangle_t \end{aligned} \quad (33)$$

and this is the time average of the instantaneous distribution Patterson.

The purely elastic intensity, $|F(\mathbf{u}, 0)|^2$, gives $\int P(\mathbf{r}, t) dt$, which depends on

$$\begin{aligned} \int G(\mathbf{r}, t) dt &= \iint g(\mathbf{R}, T) g(\mathbf{r} + \mathbf{R}, t + T) dt dT d\mathbf{R} \\ &= \int \langle g(\mathbf{R}, t) \rangle_t \langle g(\mathbf{r} + \mathbf{R}, t) \rangle_t d\mathbf{R} \\ &= \langle g(\mathbf{r}, t) \rangle_t * \langle g(-\mathbf{r}, -t) \rangle_t, \end{aligned} \quad (34)$$

and this is the self convolution of the time average of the distribution function.

5.5.4. Real monatomic gases and liquids

An important difference between (33) and (34) appears when we introduce the restriction that atoms cannot overlap appreciably, i.e. that at a particular time t the distribution Patterson $g(\mathbf{r}, t) * g(-\mathbf{r}, -t)$ and therefore in general $G(\mathbf{r}, 0)$ cannot contain any points between $|\mathbf{r}| = 0$ and $|\mathbf{r}| = 2r_0$ where r_0 is the effective radius of an atom. For $|\mathbf{r}| > 2r_0$ some fluctuations in $G(\mathbf{r}, 0)$ occur because atoms usually have an attractive interaction. There tends to be an excess of atoms at the smallest possible, "nearest neighbor" distance; then especially for liquids, an appreciable but less marked tendency for atoms to occur also at a second nearest, third nearest, etc. neighbor distance, so that $G(\mathbf{r}, 0)$ has a form as suggested, in one dimension, in Fig. 5.6(a), with a delta-function at the origin. The form of $P(\mathbf{r}, 0)$ is then as suggested in Fig. 5.6(b) and the distribution of scattering power oscillates with $|\mathbf{u}|$ as suggested in Fig. 5.6(c).

On the other hand the purely elastic scattering $|F(\mathbf{u}, 0)|^2$ derives from the self convolution of the time average of the distribution function, $\langle g(\mathbf{r}, t) \rangle_t$. But, since all atoms are

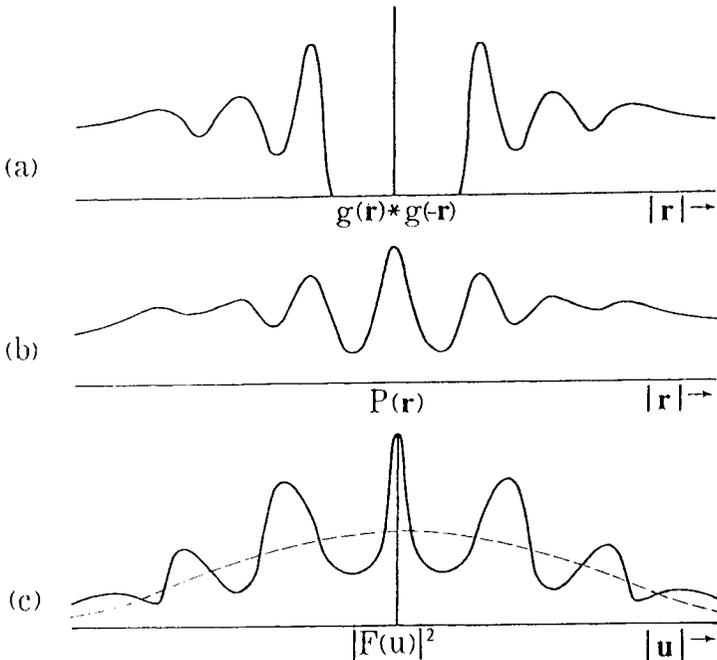


Fig. 5.6. Diagrams suggesting the form of, (a) the distribution Patterson, (b) the actual Patterson function and (c) the scattering power distribution for a real monatomic gas or liquid.

moving and all positions for an atom are equally probable, this averaged distribution function is a constant within the volume of the system. The convolution (34) and $\int P(r, t) dt$ are then also constants, multiplied by the convolution of the shape function $[s(r) * s(-r)]$. Then the purely elastic scattering consists of the single sharp peak $|S(u)|^2$.

Thus we deduce that, apart from the sharp central, zero angle peak, all diffracted intensity is inelastic. This is a somewhat academic point for X-ray or electron diffraction because the rate of decay of $P(0, t)$ with time is commensurate with the Boltzmann distribution of atom velocities for normal temperatures and average energies of the atoms of $kT (=0.02 \text{ eV})$. Energy changes of the incident radiation of this magnitude are not detectable and the normal measurements reflect the Patterson $P(r, 0)$ as given by (33). For neutron diffraction however such changes in energy, or frequency, can be detected and the situation represented by (34) may be approached.

A further stage of complication is added if we consider gases or liquids composed of molecules rather than independent atoms. Then instead of $\rho_0(r)$ for a single atom we must consider $\rho_0(r, \theta_n, \varphi_n)$ where the angles θ_n, φ_n specify the orientation of the molecule and are functions of time. The total scattered intensity then corresponds to a time average of the instantaneous Pattersons,

$$P(r, 0) = \left\langle \sum_n \sum_m \rho_0(r, \theta_n, \varphi_n) * \rho_0(-r, \theta_m, \varphi_m) * \delta(r - r_n + r_m) \right\rangle_t. \quad (35)$$

If it can be assumed as a first approximation that the relative positions of the centers of all molecules are completely random, all these terms for $n \neq m$ give only a continuous smooth background and the Patterson is the same as for a monatomic gas except that the origin peak is replaced by

$$\left\langle \sum_n \rho_0(r, \theta_n, \varphi_n) * \rho_0(-r, \theta_n, \varphi_n) \right\rangle_t.$$

For each orientation θ_n, φ_n , $P(r, 0)$ contains a set of peaks corresponding to the interatomic vectors within the molecule. The summation over n and the averaging over time reproduces this set of peaks in all possible orientations about the origin with equal probability. Then the radial distribution $P(r, 0)$ contains a set of peaks corresponding to the interatomic vector as suggested in Fig. 5.7. (a) and the intensity function correspondingly shows fluctuations about the smooth $N|f_0(u)|^2$ curve, Fig. 5.7 (b).

A more detailed and complete discussion of diffraction from liquids along lines related to the above is given by Guinier [1963]. Diffraction from gases, as a means for analysing the structures of gas molecules, has been developed most extensively using electron beams. The scattering by the low-density gases is much stronger for electrons than for X-rays or neutrons, and the shorter wavelength of electrons in the 30-40 keV range implies that intensity measurements can be made for much greater values of $(\sin\theta)/\lambda$. An account of the gas electron diffraction techniques and a survey of the structural information obtained by the method have been given by Hargittai and Hargittai (1988), (see also Hargittai, 1992).

For an array of atoms, as in an isolated molecule, separated by vectors $r_{ij} = r_i - r_j$, the diffracted intensity given by the time-averaged, spherically symmetrical distribution is given (as in (4.6)) by

$$I(s) = \langle K^2 I_0 / R^2 \rangle \sum_{i=1}^N \sum_{j=1}^N f_i(s) f_j^*(s) (\sin sr_{ij} / sr_{ij}) \quad (36)$$

where $K = 2\pi me^2 / h^2 \epsilon_0$ and, as is conventional for gas diffraction, the angular variable is $s = (4\pi \sin\theta) / \lambda = 2\pi |u|$. The distance r_{ij} varies with time because the molecule has thermal vibrations. This complicates the expression (36) somewhat (Hargittai, 1992). The terms with $i - j$ in (36) give the sum of intensities of scattering from all atoms considered separately, $\sum_i |f_i(s)|^2$, and this constitutes a smoothly falling background to the oscillations produced by the $i \neq j$ terms, the molecular scattering denoted by $I_m(s)$. By Fourier transform of this function, $I_m(s)$, obtained experimentally by subtraction of the smooth background, one obtains the radial distribution function

$$D(r) = \int_0^{\infty} s I_m(s) \sin(sr) ds \quad (37)$$

which gives, directly the distribution of probabilities for the occurrence of the interatomic distance r , weighted by the scattering strengths of the contributing atoms i.e. a radial Patterson function. From this, the interatomic distances and hence the structure of the molecule may be derived. Also, in favorable cases, the amplitudes of thermal vibration of the atoms may be deduced.

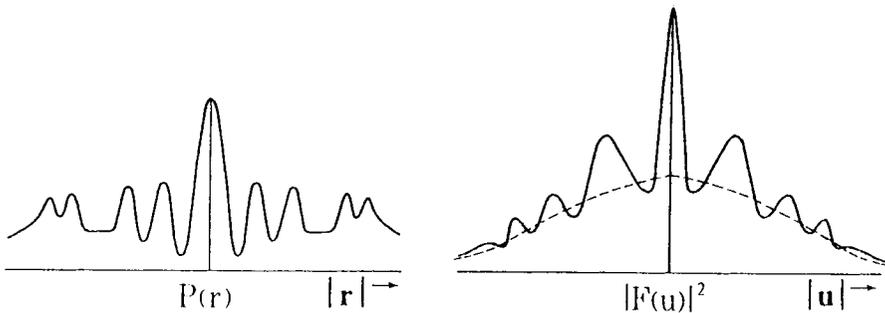


Fig. 5.7. Diagrams suggesting the Patterson function for a molecular gas, and the corresponding variation of scattering power in reciprocal space.

5.5.5. The hydrogen atom

When dealing with quantum processes of excitation of the scattering object, considerable caution is required and it cannot necessarily be assumed that the simple classical considerations we have used can be carried over (van Hove [1954]). However a direct application of the above theory gives an adequate treatment for one case; the scattering of X-rays by a hydrogen atom, i.e. by the electron of a hydrogen atom. For this electron $P(\mathbf{r}, 0)$ is a delta function at the origin, since all instantaneous pictures show a single point, one electron. Then the total scattering, in electron units squared, is given by

$$\int |F(\mathbf{u}, \nu)|^2 d\nu = 1.$$

For the elastic scattering we consider

$$\int P(\mathbf{r}, t) dt = \rho_e(\mathbf{r}) * \rho_e(-\mathbf{r}),$$

where $\rho_e(\mathbf{r})$ is the time average distribution of the electron position or the electron density function for the electron in a hydrogen atom. Then in reciprocal space

$$|F(\mathbf{u}, 0)|^2 = f_e^2(\mathbf{u}). \quad (38)$$

The inelastic (Compton) scattering is then given by the difference between the total and the elastic scattering as,

$$1 - f_e^2(\mathbf{u}), \quad (39)$$

in accordance with our result of equation (4.7).

The application of this formalism to the scattering of neutrons by phonons will be given in Chapter 12.

5.6. Diffraction geometry and intensities

We have established that kinematical, elastic diffraction amplitudes and intensities, obtained by the scattering of X-rays from electron density distributions, may be related to the distributions in reciprocal space given by Fourier transform of $\rho(\mathbf{r})$ or $P(\mathbf{r})$. The next step is to show how the amplitudes or intensities for particular experimental arrangements may be derived from the reciprocal space distributions. The argument may be carried through either for diffraction amplitudes in terms of the reciprocal space distribution $F(\mathbf{u})$ or for diffraction intensities in terms of the distribution $|F(\mathbf{u})|^2$, which may be called the distribution of the scattering power. For the moment we restrict ourselves to the latter as being more appropriate to general diffraction experiments in which intensities or energy fluxes are measured.

From equations (1.20) and (1.21), the scattered amplitude is given in the asymptotic limit of large R , as a function of $\mathbf{q} = \mathbf{k} - \mathbf{k}_0$;

$$\psi(\mathbf{q}) = -(\mu/4\pi) \int \phi(\mathbf{r}) \exp\{-2\pi i \mathbf{q} \cdot \mathbf{r}\} d\mathbf{r}. \quad (40)$$

In the appropriate units and putting $\mathbf{q} = -\mathbf{u}$, we have $\psi(\mathbf{q}) = F(\mathbf{u})$, given by equation (2) and the intensity is

$$I(\mathbf{q}) = \psi\psi^*(\mathbf{q}) = \int P(\mathbf{r}) \exp\{2\pi i \mathbf{u} \cdot \mathbf{r}\} d\mathbf{r}. \quad (41)$$

Thus for an incident monochromatic beam in a direction defined by the wave vector \mathbf{k}_0 , the intensity diffracted elastically in a particular direction defined by wave vector \mathbf{k} is equal to the value of the function $|F(\mathbf{u})|^2$ at the position in reciprocal space defined by $\mathbf{u} = (\mathbf{k}_0 - \mathbf{k})$.

This relationship is expressed by the Ewald sphere construction in reciprocal space, Fig. 5.8. A vector of length λ^{-1} ($=|\mathbf{k}_0|$) is drawn to the origin, O, of reciprocal space, in the direction of \mathbf{k}_0 from the point P. A sphere of radius λ^{-1} is drawn around P as center. Then for any point on the sphere, \mathbf{u} , the radial vector (length λ^{-1}) from P represents the direction of the diffracted beam \mathbf{k} such that $\mathbf{u} = (\mathbf{k}_0 - \mathbf{k})$. The intensity of this diffracted beam is $|F(\mathbf{u})|^2$. Thus this Ewald sphere construction gives the directions and intensities for all diffracted beams produced for a given incident beam direction.

Unless $|F(\mathbf{u})|^2$ is isotropic, its orientation in reciprocal space is defined in terms of the orientation of $P(\mathbf{r})$ in real space. Then a rotation of the sample in real space produces the corresponding rotation of the scattering power distri-

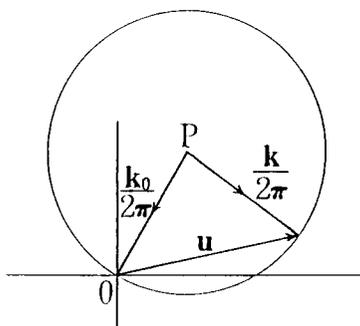


Fig. 5.8. The Ewald sphere construction.

bution $|F(\mathbf{u})|^2$ in the reciprocal space. For a constant incident beam direction the diffracted intensities then vary as regions of different scattering power are rotated through the Ewald sphere. Naturally, exactly the same sequence of intensities is produced if the distribution $|F(\mathbf{u})|^2$ is kept stationary and the Ewald sphere is rotated in the opposite direction, i.e. if the sample is kept stationary and the direction of the incident beam is rotated.

The form of the diffraction pattern recorded in practice depends on the geometry of the recording system and the wavelength of the radiation. The function $|F(\mathbf{u})|^2$ decreases, on the average, with the square of the atomic scattering factor for an average atom, $|f(\mathbf{u})|^2$. If the mean radius of an atom is taken to be about 0.5 \AA , the half-width of the distribution $|f(\mathbf{u})|^2$ is of the order of 2 \AA^{-1} and the range of $|\mathbf{u}|$ which is normally of interest is several times this: say 5 \AA^{-1} .

For X-ray and neutron diffraction, the wavelengths are of the order of 1 \AA so that the diameter of the Ewald sphere is 2 \AA^{-1} . Thus the whole of the intersection of the sphere with the function $|F(\mathbf{u})|^2$ is of interest and scattering through all angles from 0 to π is normally recorded as suggested by Fig. 5.9(a). Thus for photographic recording it is customary to use a cylindrical film with the sample on the axis. For electronic recording with a photon- or particle-counting detector, a goniometer stage is used which allows the detector to be swung through diffraction angles which are as large as is convenient.

For the diffraction of electron having a wavelength of about 0.04 \AA , the diameter of the Ewald sphere is 50 \AA^{-1} . On this sphere, only the small region of radius about 5 \AA^{-1} around the reciprocal space origin is of interest and the scattering is predominantly through small angles as suggested by Fig. 5.9(b).

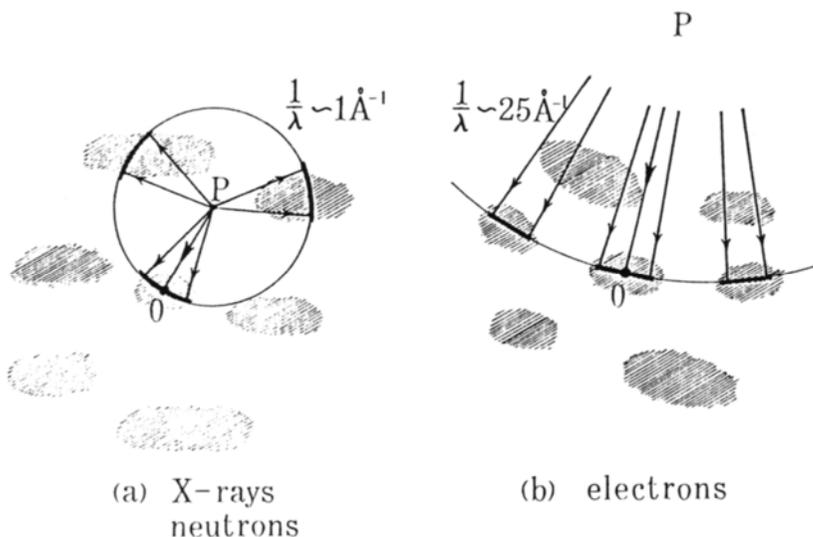


Fig. 5.9. Comparison of the scales of the Ewald spheres for X-rays, neutrons and electrons, in relation to normal scattering power distributions.

The diffraction pattern may be recorded on a flat plate or film placed perpendicular to the incident beam at some distance behind the specimen and represents an almost planar section of the distribution of scattering power in reciprocal space.

In this way it would seem that the intensities to be observed for a given radiation and for a particular geometry of the experiment may be deduced for scattering from any system for which the Patterson function can be derived or postulated. However our discussion so far has been for the idealized case of perfectly plane and monochromatic waves. These considerations must be broadened in order to make contact with experimental situations.

5.7. Practical considerations

5.7.1. Finite sources and detectors

In X-ray diffraction experiments, the finite source size results in an angle of convergence of the incident beam at any point of the sample of something like 10^{-4} to 10^{-3} radians although with synchrotron radiation much smaller convergence angles may be achieved. The intensity is not necessarily uniform over this range of angles. In neutron diffraction the angle of convergence is often made greater than this since the source intensity is low.

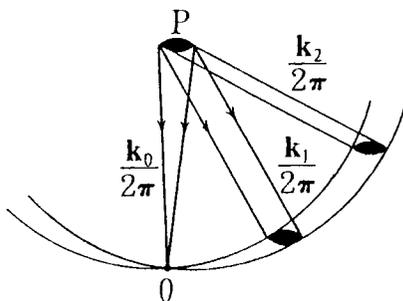


Fig. 5.10. The effect of a finite source size in spreading the Ewald sphere into a spherical shell of varying thickness.

For electron diffraction the angle of convergence may be made much less, although it may be as high as 10^{-3} radians when the specimen is immersed in the magnetic field of the objective lens in an electron microscope and the "selected-area electron diffraction" technique is used (see Chapter 13).

Such a convergence of the incident beam modifies the picture of the Ewald sphere in reciprocal space in a way suggested in Fig. 5.10. Drawing the incident beam directions as vectors k_0 to the reciprocal lattice origin, O , the origin points P are distributed over a disc having the shape of the source and a relative weighting of the points given by the intensity distribution of the source. Corresponding to each point of this disc there is a differently-oriented Ewald sphere, so that we may think of an Ewald sphere thickened into a spherical shell of thickness which varies with distance from O .

For a given diffracted beam direction, the vectors k from the source disc at P , each drawn to the corresponding Ewald sphere, define a disc-shaped section of the "Ewald shell". This we may call the "scattering disc". The total intensity scattered in the direction k is then given by integrating $|F(u)|^2$ over this scattering disc with a weighting factor corresponding to the intensity of the source disc at P . Clearly, the size and shape of the scattering disc will vary with the angle of scattering, so the effect on the intensities cannot be represented by a simple convolution of some shape function with $|F(u)|^2$ unless all scattering angles are small as in the case of electron diffraction.

When the recording of the diffraction pattern is photographic, the resolution of the photographic plate or film is normally sufficient to allow each diffracted beam direction of interest to be recorded separately. For electronic recording, however, the detector usually accepts a finite angular range of diffracted beams from each point of the sample. Then for each incident beam direction the diffracted intensity depends on the value of not just for one

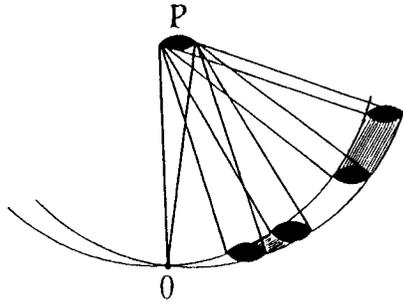


Fig. 5.11. The combined effects of finite source size and finite detector aperture size to define "scattering volumes" in reciprocal space over which the scattering power is integrated.

point on the Ewald sphere but for a finite area of the Ewald sphere. Combining this with the effect of the finite source size, as suggested by Fig. 5.11, the recorded intensity for any setting of the crystal and the detector comes from a volume of reciprocal space. Both the size and shape of this "scattering volume" vary with scattering angle.

5.7.2. Wavelength spread

A further complication comes from the finite range of wavelengths present for any real source. For X-rays the natural half-widths of the characteristic emission lines are of the order of $10^{-4} \lambda$ or more. For neutron diffraction or for synchrotron X-rays, since the radiation used is selected out of a continuous "white radiation" distribution, the range of wavelengths used may be made greater in order to increase the total intensity of incident radiation. For electrons the radiation is usually much more nearly monochromatic with a width of about $10^{-6} \lambda$.

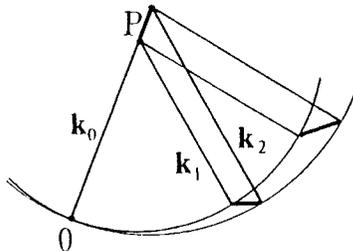


Fig. 5.12. The effect of a spread of wavelength in producing a spread in values of the Ewald sphere radius.

The spread of wavelengths produces a spread in the values of the Ewald sphere radius, as suggested in Fig. 5.12 for the case of very small incident and diffracted beam divergence. This results in a broadening of the scattering region into a line rather than a disc since the origin points P lie on a line. There is an essential difference from the case illustrated in Fig. 5.10 in that the wavelength spread gives a scattering line which varies in its orientation as well as length, being small and parallel to k_0 in the limit of small scattering angles; of medium length and roughly perpendicular to k_0 for a scattering angle like the larger one shown in the figure, and of maximum length and oppositely directed to k_0 for a scattering angle of π .

Adding this effect to the effects of the divergence of the incident and scattered beams suggested in Fig. 5.11 gives a scattering volume having a shape which is complicated and varies with scattering angle in a complicated way. Thus, in general, the relationship of the observed intensity to the function $|F(u)|^2$ is only to be derived by laborious calculation from a detailed knowledge of the parameters of the experimental arrangement.

5.7.3. Integrated intensities

The usual procedure to simplify matters for the measurement of continuous distributions of scattering power is to ensure that the scattering volume is much smaller than the region over which $|F(u)|^2$ varies significantly. In the case of diffraction from well-crystallized materials, when the object is to measure the integrated scattering power contained in sharp maxima separated by a low flat background, the usual scheme is to generate a scattering volume in the shape of a thin disc, e.g. by having a well-collimated, nearly-monochromatic beam and a relatively wide angle of acceptance of the detector so that the disc is a portion of the Ewald sphere. Then the disc is swept through the sharp maximum of scattering power, e.g. by rotating the crystal, and the observed intensity is integrated over time.

For each experimental situation involving these simplifying special cases, the variation of the shape and size of the scattering volume and the rate at which it is scanned through reciprocal space are taken into account by multiplying the observed intensities by an appropriate factor, known as the Lorentz factor, which may usually be taken as a function of the magnitude of the scattering angle only. Detailed derivations are given, for example, by Warren [1969] or in the International Tables for Crystallography, Vol. C (Wilson, 1992).

For the diffraction of electrons with medium or high energies (greater than 50 keV say) all these considerations of scattering volumes are greatly simplified by the small

scattering angles involved. The effect of the wavelength spread is negligible. The effect of the finite source size is to convolute the two-dimensional scattering power distribution, $|F(u, v)|^2$, with a source function, say $S(u, v)$ since this may be assumed independent of scattering angle. Then

$$I(u, v) = |F(u, v)|^2 * S(u, v). \quad (42)$$

Also if this intensity distribution is observed with a detector of finite aperture or photographic plate of finite resolution with sensitivity represented by a function $D(u, v)$, the measured intensity is given by

$$\begin{aligned} I_{\text{obs}} &= I(u, v) * D(u, v) \\ &= |F(u, v)|^2 * S(u, v) * D(u, v). \end{aligned} \quad (43)$$

The Lorentz factor then derives solely from the way in which intensities are integrated over time when a sample is rotated or, more commonly, integrated over crystal orientation when the sample contains a range of crystal orientations.

In neutron diffraction involving measurements of intensity as a function of energy loss as well as diffraction angles, the complications may be correspondingly greater than for X-ray diffraction. We prefer not to deal with them here.

5.8. Sections and projections

For a number of purposes it is convenient to deal with two- or one-dimensional sections or projections of the three-dimensional functions $\rho(r)$, $P(r)$, $F(u)$ and $|F(u)|^2$. It is much easier to represent and appreciate functions plotted in one or two dimensions and the amount of data required to define them is often very much less. For the case of a radially symmetric function, such as $P(r)$ for a gas or liquid, no new information is gained by going to more dimensions than one.

Experimentally, electron diffraction patterns represent, to a first approximation, planar sections of reciprocal space. In X-ray diffraction a planar section of reciprocal space may be obtained by use of the Buerger Precession Camera or similar devices, and the simplest form of operation of the X-ray diffractometer, the $\theta - 2\theta$ scan, gives a section along a radial straight line.

It is therefore appropriate to summarize now the general relationships between sections and projections in real and reciprocal space. We do so with reference to the functions $\rho(r)$ and $F(u)$ for convenience but draw our examples from our broader range of interest.

The projection of the function $\rho(r)$ in the direction of the z -axis is

$$\rho(x, y) = \int_{-\infty}^{\infty} \rho(r) dz. \quad (44)$$

In terms of $F(u)$ this is

$$\rho(x, y) = \iint F(u) \exp\{-2\pi i(ux + vy + wz)\} dz du. \quad (45)$$

The integral over z is the delta function $\delta(w)$ so that $\rho(x, y)$ is the inverse Fourier transform of $F(u, v, 0)$, the section of $F(u, v, w)$ on the plane $w = 0$.

Similarly, a planar section of $\rho(r)$ on the $z = 0$ plane is

$$\begin{aligned} \rho(x, y, 0) &= \int F(u) \exp\{-2\pi i(ux + vy)\} du dv dw \\ &= \int [\int F(u, v, w) dw] \exp\{-2\pi i(ux + vy)\} du dv. \end{aligned} \quad (46)$$

Thus we have the general relationship, that a planar section through the origin in real space corresponds to a projection of the reciprocal space distribution on a parallel plane, and vice versa.

For example, the electron diffraction pattern, to a first approximation is a planar section of reciprocal space so that Fourier transform of the intensity distribution of the diffraction pattern gives the projection of the Patterson function $P(r)$ in the beam direction. This is the approximation that the object may be treated as a two-dimensional phase- and amplitude-object.

We have seen examples in Section (5) above for the case of the four-dimensional distributions in space and time when intensity is measured as a function of scattering angles and frequency change. Thus the section of reciprocal space on the $v = 0$ plane, appropriate for purely elastic scattering, equation (28), gives the projection of the Patterson function in the time direction or the time-average of the correlation function. The projection of the four-dimensional distribution scattering power in reciprocal space in the v direction, given by the integral over v in equation (29), is the Fourier transform of the section of the Patterson function $P(r, 0)$ which is the sum of the instantaneous spatial correlations of the object.

If the section in real space is not through the origin, a corresponding phase factor is introduced in reciprocal space. Thus the Fourier transform of the section $\rho(x, y, c)$ gives the modulated projection

$$\int F(u, v, w) \exp\{-2\pi iwc\} dw. \quad (47)$$

Extending the considerations to one-dimensional sections and projections, we have

$$\mathfrak{S}\rho(x, 0, 0) = \iint F(u, v, w) dv dw, \quad (48)$$

and

$$\int \int \rho(x, y, z) dy dz = F(u, 0, 0). \quad (49)$$

Thus the value of $\rho(r)$ along the line of the x axis is given by Fourier transform of the projection of $F(u)$ along both v and w directions on to the u -axis.

An example from the four-dimensional situation is that of equation (27), where the function of interest is $P(0, t)$, the correlation in time for the position at the origin. This function would be given by projecting the reciprocal space function $|F(u, v)|^2$ in the u , v and w directions, i.e. by integrating over all scattering directions for each frequency change, ν .

Problems

1. Sketch the Patterson function in two dimensions for a stationary molecule in the form of an equilateral triangle of atoms. Is there a centro-symmetrical object or a group of centro-symmetric objects which gives the same Patterson peaks (apart from the origin peak)? Derive an expression for the corresponding distribution $|F(u)|^2$ in reciprocal space. Repeat the considerations for an ideal gas of such molecules.

2. A diatomic molecule of fixed orientation (along x axis) vibrates about a fixed center of mass with frequency ν_0 . Sketch the correlation function $P(x, t)$ and the corresponding reciprocal space distribution $|F(u, \nu)|^2$.

3. Derive the Lorentz factor to be applied to X-ray diffraction from a gas and for the intensities of rings in a powder pattern (from crystallites in all possible orientations, each giving the same set of sharp peaks in reciprocal space).

This Page Intentionally Left Blank

Diffraction from crystals

6.1. Ideal crystals

Although, as is well known, real crystals contain many different types of imperfections and faults, including point defects, impurities, dislocations, stacking faults and so on, it is often possible to consider the main diffraction effects as coming from an ideally periodic average crystal, as we shall see in Chapter 7. The discussion of the kinematical diffraction from ideal periodic crystals forms the basis for the important field of crystal structure analysis and so merits some special attention here.

An ideal crystal is made up by the repetition in three dimensions of a unit cell containing one or more atoms. In general the unit cell is not rectangular. It is defined by three vectors a, b, c which have length a, b, c and angles between the axes α, β, γ . Hence, writing $\rho_0(r)$ for the contents of the unit cell and $s(r)$ for the shape function,

$$\rho(r) = \left[\rho_0(r) * \sum_l \sum_m \sum_n \delta\{r - (la + mb + nc)\} \right] s(r)$$

or

$$\rho(r) = \left[\rho_0(r) * \sum_l \sum_m \sum_n \delta(x-la, y-mb, z-nc) \right] s(r), \quad (1)$$

where x, y, z are coordinates with respect to the axes, in the directions of the unit vectors x, y, z parallel to a, b, c .

This is a generalization of equation (2.56) following (5.8). For the special case for which a, b, c are at right angles, we saw (equation (5.9)) that Fourier transform gives a reciprocal lattice of points with spacings a^{-1}, b^{-1}, c^{-1} i.e. that the reciprocal lattice is defined by vectors, a^*, b^*, c^* such that $|a^*| = a^{-1}$, $|b^*| = b^{-1}$, $|c^*| = c^{-1}$. For the more general case of non-rectangular axes, we must redefine the reciprocal lattice and put

$$\begin{aligned}
 a^* &= \frac{b \times c}{a(b \times c)} = \frac{b \times c}{V}, \\
 b^* &= \frac{c \times a}{V}, \\
 c^* &= \frac{a \times b}{V},
 \end{aligned}
 \tag{2}$$

where V is the unit cell volume, so that

$$a^* \cdot b = a^* \cdot c = b^* \cdot c = \dots = 0, \tag{3}$$

$$a \cdot a^* = b \cdot b^* = c \cdot c^* = 1. \tag{4}$$

Then Fourier transform of (1) gives

$$F(u) = F_0(u) \sum_h \sum_k \sum_l \delta\{u - (ha^* + kb^* + lc^*)\} * S(u)$$

or

$$F(u, v, w) = F_0(u) \sum_h \sum_k \sum_l \delta(u-ha^*, v-kb^*, w-lc^*) * S(u). \tag{5}$$

For a crystal having the dimensions A, B, C in the directions of the three axes, the function $S(u)$ has the same form as in equation (5.10). Alternatively, since the only values of $F_0(u)$ of interest are those at the reciprocal lattice points we may write

$$F(u) = \sum_h \sum_k \sum_l F_{hkl} \delta\{u - (ha^* + kb^* + lc^*)\} * S(u). \tag{6}$$

Then F_{hkl} is the "structure factor" or preferably, the "structure amplitude" for the h, k, l reciprocal lattice point and is given by

$$F_{hkl} = \int_0^a \int_0^b \int_0^c \rho(x, y, z) \exp\{2\pi i((hx/a) + (ky/b) + (lz/c))\} dx dy dz$$

or

$$F_h \equiv F_{hkl} = \int \rho(r) \exp\{2\pi i h \cdot r\} dr \tag{7}$$

where the integration is over the unit cell and h is the vector $ha^* + kb^* + lc^*$.

An alternative and often useful convention is to use fractional coordinates which we signify for the moment as X, Y, Z so that distances are measured in terms of unit cell parameters;

$$X = xa, \quad Y = yb, \quad Z = zc. \quad (8)$$

Then (7) becomes,

$$F_{hkl} = V \int_0^1 \int_0^1 \int_0^1 \rho(X, Y, Z) \exp\{2\pi i (hX + kY + lZ)\} dX dY dZ. \quad (9)$$

By inverse transform of (6) we derive the alternatives to (1);

$$\rho(x, y, z) = V^{-1} \sum_h \sum_k \sum_l F_{hkl} \exp\{-2\pi i ((hx/a) + (ky/b) + (lz/c))\} \cdot s(\mathbf{r}) \quad (10)$$

$$\rho(\mathbf{r}) = \sum_h F_h \exp\{-2\pi i \mathbf{h} \cdot \mathbf{r}\} \cdot s(\mathbf{r}), \quad (11)$$

or

$$\rho(X, Y, Z) = \sum_h \sum_k \sum_l F_{hkl} \exp\{-2\pi i (hX + kY + lZ)\} \cdot s(\mathbf{r}). \quad (12)$$

If the electron density in the unit cell is assumed to be the summation of the electron densities for individual atoms, we write

$$\rho_0(\mathbf{r}) = \sum_i \rho_i(\mathbf{r}) * \delta(\mathbf{r} - \mathbf{r}_i), \quad (13)$$

so that

$$F_{hkl} = \sum_i f_i \exp\{2\pi i (hX_i + kY_i + lZ_i)\}, \quad (14)$$

where X_i, Y_i, Z_i are the fractional coordinates of the atom at \mathbf{r}_i and f_i is the atomic scattering factor.

The distribution of scattering power in reciprocal space consists of a sharp peak around each reciprocal lattice point of the form $|S(\mathbf{u})|^2$ so we may write, to a good approximation

$$|F(\mathbf{u})|^2 = \sum_h \sum_k \sum_l |F_{hkl}|^2 \delta\{\mathbf{u} - (h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^*)\} * |S(\mathbf{u})|^2. \quad (15)$$

Inverse transformation of this then gives, following (12),

$$\begin{aligned}
 &P(X, Y, Z) \\
 &= V^{-1} \sum_h \sum_k \sum_l |F_{hkl}|^2 \exp\{-2\pi i (hx/a + ky/b + lz/c)\} [s(r) * s(-r)], \quad (16)
 \end{aligned}$$

$$\begin{aligned}
 &P(r) \\
 &= [\rho_0(r) * \rho_0(-r) * \sum_l \sum_m \sum_n \delta\{r - (la + mb + nc)\}] [s(r) * s(-r)]. \quad (17)
 \end{aligned}$$

Thus, apart from the gradual fall-off due to the shape convolution, the Patterson function is periodic, as foreseen in equation (5.17), and is made up by repetition of the self-convolution of the contents of the unit cell.

For convenience the shape convolution and its transform are often omitted. Then it must be understood that the periodic functions in real space and the delta functions in reciprocal space are operational abstractions which can be expanded into the more realistic descriptions of equations (15) to (17) when necessary.

6.2. Diffraction geometry

6.2.1. Laue and Bragg diffraction conditions

From equation (15) we see that for the ideal finite crystal the distribution of scattering power is a sharp peak, $|S(u)|^2$, around each point of the reciprocal lattice defined by the vectors a^* , b^* , c^* . Then a sharply defined diffracted beam is generated when the Ewald sphere cuts through one of these sharp peaks of scattering power. From our previous considerations of the Ewald sphere construction we see that the geometric condition to be satisfied is

$$q \equiv k - k_0 = h \equiv ha^* + kb^* + lc^*. \quad (18)$$

We may write this in terms of the projections of q on the real space axes as

$$q \cdot a = h, \quad q \cdot b = k, \quad q \cdot c = l, \quad (19)$$

which represent the well-known "Laue conditions" for diffraction.

The condition for diffraction may also be written in terms of the concept of planes of atoms in the crystal. The periodicity of the crystal ensures that sets of parallel planes may be drawn to pass through atom centers at regular

intervals. These sets of planes are denoted by the Miller indices hkl if, when one plane is drawn through an atom at the unit cell origin, the intercepts of the next plane of the set on the axes are $a/h, b/k, c/l$.

It is easy to see in the case of rectangular axes, that the perpendicular distance between planes of the hkl set is d_{hkl} where

$$1/d_{hkl}^2 = (h^2/a^2) + (k^2/b^2) + (l^2/c^2). \quad (20)$$

But this is exactly the square of the distance from the reciprocal lattice origin to the hkl reciprocal lattice point and the direction of the perpendicular to the hkl lattice planes is the direction from the origin to the hkl reciprocal lattice point. This relationship holds also for non-orthogonal axes, although then the relationship (20) becomes more complicated.

Hence the condition that a sharp diffraction beam should be generated is, from (19) and (20),

$$|q| = 1/d_{hkl},$$

or, since $|k_0| = |k| = 1/\lambda$, and the angle between k and k_0 is 2θ , we have

$$2d_{hkl} \sin\theta_{hkl} = \lambda, \quad (21)$$

which is Bragg's law. The condition that q should be perpendicular to the lattice planes is equivalent to Bragg's concept of a "reflection" in the optical sense from the planes, subject to (21). We refer for convenience to the strong well-defined diffracted beams given under these conditions as "Bragg reflections".

6.2.2. Shape transforms

The Bragg or Laue conditions refer to the Ewald sphere passing exactly through a reciprocal lattice point. Our more detailed discussion of diffraction conditions suggests that, for monochromatic incident plane waves, the diffracted intensity for this case would be given by a near-planar section of the distribution of scattering power, $|S(u)|^2$, given a weighting $|F_h|^2$. If, in practice, the divergence of the incident beam, finite acceptance angle of the detector and wavelength spread of the incident radiation together define a scattering volume, as described in the last chapter, which is very much larger in size than $|S(u)|^2$, then the observed intensity may correspond to a three-dimensional integration over the peak of scattering power and so be proportional to $|F_h|^2$. Alternatively a thin "scattering disc"

may be swept through the peak at a constant rate so that the observed intensity integrated over time is proportional to the integral over $|S(\mathbf{u})|^2 |F_h|^2$ and is proportional to $|F_h|^2$.

For X-rays and neutrons the peaks of scattering power are very sharp. The kinematical scattering approximation may apply for perfect crystal regions several thousand times the dimensions of the unit cell so that the dimensions of the $|S(\mathbf{u})|^2$ distribution are several thousand times smaller than the reciprocal lattice unit cell dimensions. Let us take as a representative value a halfwidth of $|S(\mathbf{u})|^2$ of $1/2000 \text{ \AA}^{-1}$. An incident beam divergence of 10^{-3} radians gives the Ewald sphere a thickness of about $1/2000 \text{ \AA}^{-1}$ for a reciprocal lattice point with $1/d_h = 0.5 \text{ \AA}^{-1}$ and the wavelength spread of the incident beam adds to this. Hence even for a single crystal region of this size and for favorable diffraction conditions it would not be feasible to see anything of the details of the function $|S(\mathbf{u})|^2$. Only an integrated intensity could be recorded.

However with readily available sources of radiation the intensity scattered by such a small region is too small to be observed. The crystals normally used for diffraction under near-kinematical conditions are imperfect and may be thought of as very large numbers of small regions of this order of size, or a "mosaic" of small crystals, having a spread of orientations of perhaps 10^{-2} to 10^{-3} radians. Then we may consider the Ewald sphere to be cutting an aggregate of a very large number of slightly rotated reciprocal lattice configurations or, equivalently, we may consider the reciprocal space distribution for a single average crystal cut by a further-broadened Ewald sphere. This makes it even more nearly certain that only an integration over the scattering power peak is seen and no information on individual $|S(\mathbf{u})|^2$ functions is attainable (Chapter 16).

6.2.3. Special cases for electron diffraction

For electron diffraction the situation is quite different. Crystals giving near-kinematical intensities are normally a few hundred \AA in size in at least the one direction parallel to the incident beam. Sources are sufficiently bright to allow diffraction from single crystals of this size to be observed readily and the monochromatization and collimation give a broadening of the Ewald sphere with an angular spread of as little as 10^{-5} radians. Thus for a reflection with $1/d_h = 0.5 \text{ \AA}^{-1}$ the extent of the shape transform function may be 10^{-2} \AA^{-1} or more while the thickness of the Ewald sphere may be as little as $5 \times 10^{-6} \text{ \AA}^{-1}$. Thus near-planar sections of the scattering power peak are frequently observed. Fig. 6.1 is a portion of a diffraction pattern from small needle-like crystals of ZnO (Rees and

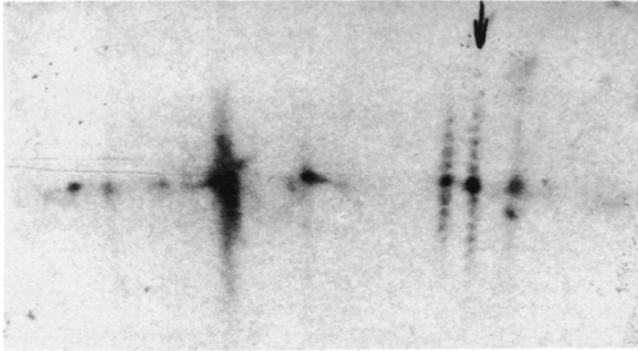


Fig. 6.1. Portion of a diffraction pattern (enlargement of spots on one ring) from small needle-like crystals of ZnO, showing shape-transform modulation of the scattering power (Rees and Spink [1950]).

Spink [1950]). The limitation of crystal size in the direction perpendicular to the beam gives rise to an extension of the peak of scattering power in the plane of the Ewald sphere. The modulation of the intensity corresponding to the $(\sin^2 x)/x^2$ form of $|S(u)|^2$ is clearly seen in the spots from several individual needle crystals. (The intensity variation is usually modified by dynamical effects but for these particular cases this is not very obvious.)

It follows that in electron diffraction special techniques of specimen preparation or of intensity recording are required in order to obtain integrated intensities proportional, in the kinematical approximation, to $|F(h)|^2$. For example, the techniques for crystal structure analysis by electron diffraction developed in the USSR (Pinsker [1953], Vainshtein [1964], Vainshtein et al. [1992]) depend mostly on the use of oriented polycrystalline specimens having a random distribution of orientations about one axis so that each reciprocal lattice spot is spread into an annular ring and the section of this by the Ewald sphere gives an integrated intensity. Single crystal patterns are often obtained from extended thin sheets of crystal, of the order of 100 Å thick but possibly microns in diameter. Inevitably these thin sheets are often bent. This again provides an integration over the scattering power maximum, although the observed intensities are proportional to $|F(h)|^2$ only if the bending is sufficiently uniform or sufficiently random to ensure that all crystal orientations are equally represented.

A further consequence of the difference in geometry between the X-ray and electron cases is the difference in the number of diffracted beams produced at any one time. For X-rays, even with the spread of the maxima of scattering

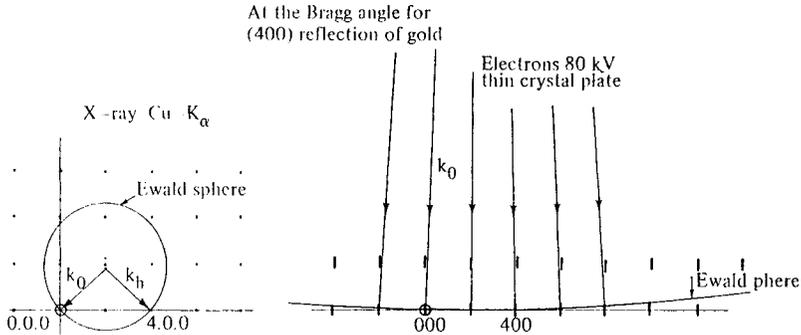


Fig. 6.2. Ewald sphere diagram for $\text{CuK}\alpha$ X-rays and for 80 keV electrons for a crystal of gold or aluminum when the Bragg condition is satisfied for the 400 reflection. The crystal dimensions in the beam direction are assumed to be 1000 Å for the X-rays and 50 Å for the electrons.

power or the Ewald sphere which we have discussed, the probability that a strong reflection should occur for any particular incident beam orientation is low for crystals with small unit cells. If a strong reflection does occur, it is unlikely that a second one will be generated. For electrons on the other hand, the Ewald sphere normally cuts a number of the extended scattering-power regions and for particular orientations the number of diffracted beams may be large. This is illustrated in Fig. 6.2 drawn approximately to scale for the diffraction of $\text{CuK}\alpha$ X-rays and 80 keV electrons from crystals of gold or aluminum with the Bragg reflection condition satisfied for the 400 reciprocal lattice point in each case. For the X-ray scattering the perfect crystal regions are assumed to be 1000 Å or more in size. For electrons the crystal is assumed to be a thin film 50 Å thick.

It is seen that for electrons, especially when the third dimension is taken into account, the number of simultaneously diffracted beams is quite large. If the voltage of the electron beam is increased and the wavelength is correspondingly reduced, the Ewald sphere becomes more nearly planar and the number of reflections for such an orientation increases rapidly, especially for the voltage range greater than about 200 keV for which relativistic effects become important.

If the thin crystals used in electron diffraction are bent, the rotation of the reciprocal lattice relative to the Ewald sphere ensures that many more diffracted beams are produced, giving the diffraction pattern the appearance of a

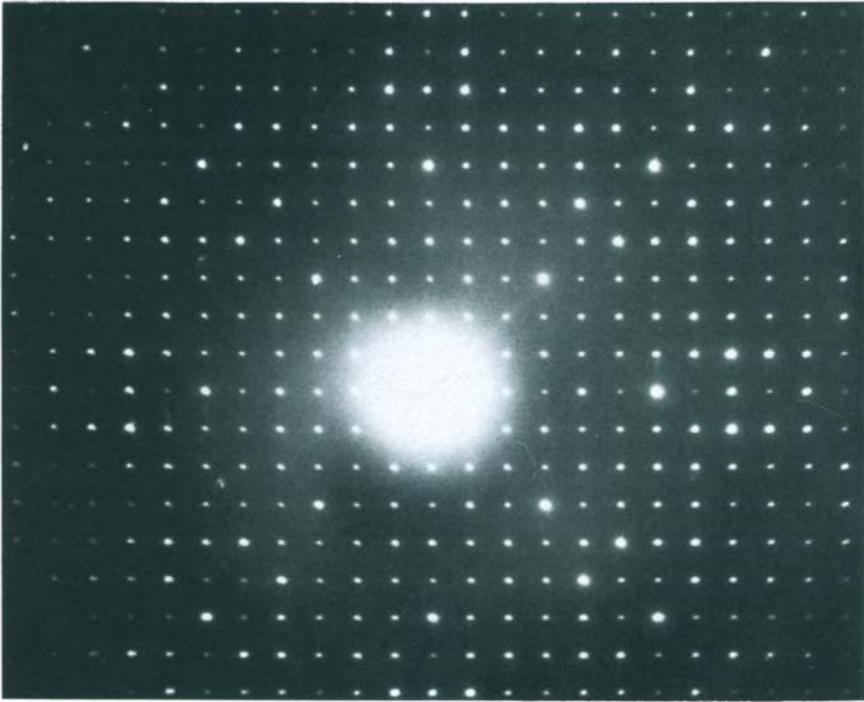


Fig. 6.3. Electron diffraction pattern from a thin, slightly bent crystal of Mo_5O_{14} with the incident beam nearly parallel to the c -axis.

complete section of the reciprocal lattice as in Fig. 6.3, a diffraction pattern obtained from a thin bent crystal with the incident beam approximately perpendicular to the $hk0$ reciprocal lattice plane.

Obviously this diffraction pattern contains information on the unit cell geometry and symmetry of the crystal. Also, provided that a kinematical approximation can be used, the relative intensities of the spots should allow values of $|F_h|^2$ to be derived as a basis for the study of crystal structures. The possibilities in this direction have been reviewed by Cowley [1967] and Dorset [1995].

6.3. Crystal structure analysis

6.3.1. The phase problem

The primary object of the analysis of the crystal structure of a substance not previously studied is to find

the positions of the centers of the atoms, as given by the maxima in the electron density function $\rho(\mathbf{r})$ and to identify the atoms present from the relative weights of the maxima. As secondary objectives to be achieved by use of more accurate data and more extended analysis, one may hope to find the distortions of the free-atom electron densities of the atoms due to ionization and bonding and also the mean-square displacements of the atoms from their average positions due to thermal vibration. Thus the aim is to determine $\rho(\mathbf{r})$ within the unit cell of the average, periodic lattice.

As we have seen in Chapter 5, we may write

$$\rho(x, y, z) = V^{-1} \sum_{\mathbf{h}} \sum_k \sum_l F_{hkl} \exp\{-2\pi i ((hx/a) + (ky/b) + (lz/c))\}, \quad (22)$$

or, from (12), in fractional coordinates,

$$\rho(X, Y, Z) = \sum_{\mathbf{h}} \sum_k \sum_l F_{hkl} \exp\{-2\pi i (hX + kY + lZ)\}. \quad (23)$$

However, the observable intensities give us $|F_{hkl}|^2$, not F_{hkl} . The phases of F_{hkl} (or $F_{\mathbf{h}}$) can not influence the intensities for kinematical scattering without absorption. In general $F_{\mathbf{h}}$ is complex and can be written $|F_{\mathbf{h}}| \exp\{i\eta_{\mathbf{h}}\}$ and it is the phase factor $\eta_{\mathbf{h}}$ which is lost. For a centrosymmetrical crystal $F_{\mathbf{h}}$ is real so that it is the choice between positive and negative signs which can not be made.

The question of overcoming this lack of the information essential for the derivation of $\rho(\mathbf{r})$ constitutes the phase problem of crystal structure analysis. In principle this phase problem may be solved in many ways since, for example, either dynamical diffraction effects or scattering with absorption may give scattering sensitive to the relative phases of the reflections and these effects are never completely absent. In practice, however, the phase problem remains as a serious hindrance to the derivation of electron density distributions and a great deal of ingenuity has been expended on the development of procedures to overcome this limitation.

The observable quantity in reciprocal space is $|F_{\mathbf{h}}|^2$ and from this the Patterson function $P(\mathbf{r})$ for the crystal (equation (17)) is directly derived. The deconvolution of $P(\mathbf{r})$ to give $\rho(\mathbf{r})$ is in general not possible. In order to proceed it is necessary to make use of the available knowledge of the form of $\rho(\mathbf{r})$ and any information about the structure which may be gained from other sources.

6.3.2. Supplementary information

It is known that $\rho(r)$ is a positive, real function having a small, near-constant value except for isolated peaks at the atom positions. The relative weights of the peaks depend on the atomic numbers of the atoms and the separations of the peaks are not less than some fairly well defined minimum value.

The number of each kind of atom contained in the unit cell may usually be deduced from the chemical analysis of the material, the unit cell dimensions (given by the geometry of the diffraction patterns) and the density of the crystal.

Information on the possible arrangements of atoms within the unit cell may be provided by consideration of the symmetry of the crystal structure. For each crystal, the atomic arrangement must conform with the symmetry elements of one of the 230 possible space groups as listed and described in the International Tables for Crystallography, Vol. A (Hahn, 1992). From our previous considerations it can be seen that a symmetry operation in real space involving a rotation of the crystal about an axis or reflection in a plane must be accompanied by the same symmetry operation in reciprocal space. Operations of a screw axis or glide plane involving translations in real space must involve multiplication by a phase factor in reciprocal space which may make the amplitudes of some points of reciprocal space equal to zero, giving "systematic absences" of some reflections. Thus a great deal of information on real space symmetry may be derived from observations on intensity distributions in reciprocal space. The important exception, as we have seen, is that the presence or absence of a center of symmetry can not be deduced directly from observation of diffraction intensities since $|F(\mathbf{u})|^2 = |F(-\mathbf{u})|^2$. As a consequence, only 58 of the space groups can be identified uniquely from kinematical diffraction data and altogether it is possible to distinguish only 122 sets of one or more space groups. In some cases the presence or absence of a center of symmetry may be indicated by non-diffraction measurements such as the observation of piezoelectricity.

For simple structures the available information on symmetry may be sufficient to locate all atoms in the unit cell. For more complicated structures it may limit the possible positions considerably. For example, if there are one or two atoms of a particular kind in a unit cell which contains a 3- or 4-fold rotation axis, those one or two atoms must lie on that axis. For a detailed discussion of such considerations see, for example, Lipson and Cochran [1966].

The problem of determining the remaining parameters needed to define the crystal structure may be considered in real space as the problem of finding the positions \mathbf{r}_i of the individual atoms: or it may be considered in reciprocal space

as the problem of assigning phases to the amplitudes $|F_{\mathbf{h}}|$. The main techniques which have been developed to assist with these problems have been described in detail in a number of books (e.g. Lipson and Cochran [1966], Ladd and Palmer [1993], Schwartz and Cohen [1977]) and so will not be treated extensively here. We mention only a few of the key ideas and methods related to our more general considerations.

6.4. Structure analysis methods

6.4.1. Trial and error

The earliest and most direct approach to finding the parameters not determined by symmetry is to postulate a structure and calculate intensities for comparison with observed intensities. From a set of atomic positions with fractional coordinates x_i, y_i, z_i and atomic scattering factors f_i assumed initially to be those of isolated atoms, the structure amplitude is calculated as

$$F_{\mathbf{hkl}} = \sum_i f_i \exp\{2\pi i(hx_i + ky_i + lz_i)\}. \quad (24)$$

Then $|F_{\mathbf{h}}|_{\text{calc}}^2$ is compared with $|F_{\mathbf{h}}|_{\text{obs}}^2$ derived from the intensities.

As a measure of agreement for the often large number of values involved it is usual to calculate the R -factor,

$$R = \sum_{\mathbf{h}} \frac{||F_{\mathbf{h}}| - F_{\mathbf{h}}|}{\sum_{\mathbf{h}} |F_{\mathbf{h}}|_{\text{calc}}}. \quad (25)$$

This factor is often multiplied by 100 and expressed as a percentage. Then changes in the postulated atomic coordinates are made in an effort to minimize R . For some purposes it is considered to be more appropriate to minimize R_1 defined by

$$R_1 = \sum_{\mathbf{h}} W_{\mathbf{h}} (|F_{\text{obs}}| - |F_{\text{calc}}|)^2, \quad (26)$$

where $W_{\mathbf{h}}$ is a weighting factor which is determined from estimates of the relative accuracy of the various experimental measurements.

Once the parameters of the structure have been determined with moderate accuracy systematic least-squares minimization procedures may be applied to refine them.

An alternative to this refinement in terms of the reciprocal space quantities is the real-space procedure of calculating successive Fourier maps, or contoured maps of

approximations to the electron density found by summation of the series

$$\rho_n(\mathbf{r}) = \sum_h |F_h|_{\text{obs}} S_{h,\text{calc}} \exp\{-2\pi i(\mathbf{h} \cdot \mathbf{r})\}, \quad (27)$$

where $S_{h,\text{calc}}$ is the sign of the structure amplitude calculated from (24) for a trial structure. With all the calculated signs correct $\rho_n(\mathbf{r})$ shows well-shaped symmetrical peaks at the atom positions and a flat background. Deviations from this suggest changes to be made in the trial structure.

6.4.2. Patterson function

The Patterson function for the periodic crystal structure may be calculated from the observed intensities as

$$P(\mathbf{r}) = \frac{1}{2} \sum_h |F_h|^2 \cos\{2\pi \mathbf{h} \cdot \mathbf{r}\}, \quad (28)$$

where the center of symmetry allows us to replace the exponential by the cosine. As we have seen in Chapter 5 in the case of the Generalized Patterson function, $P(\mathbf{r})$ gives an appropriately weighted mapping of the inter-atomic vectors. As indicated by equation (17), $P(\mathbf{r})$ may be considered as the periodic repetition of the Patterson function for the contents of one unit cell, $P_0(\mathbf{r})$. However this repetition may create some complication because it gives some overlapping of different parts of $P_0(\mathbf{r})$.

For relatively simple structures the recognition of particular interatomic vectors may allow a determination of the structure on a trial- and error-basis, but in general the Patterson contains too many poorly resolved peaks for this to be useful.

One approach to the interpretation of the Patterson is the so-called "image seeking" procedure introduced by Buerger [1959] based on the interpretation of the Patterson which is illustrated in Fig. 5.2. One seeks systematically for images of the structure which are repeated in the Patterson with the vector separations of the inverse structure.

6.4.3. Heavy-atom and isomorphous replacement methods

Of the many other techniques used, one class involves the use of specially favorable or specially prepared crystals. If one or more heavy atoms are incorporated in the unit cell, their contributions tend to have a dominating influence on the signs of the structure factors. Hence if the heavy-atom positions are found (for example by recognizing the heavy atom to heavy atom vector peaks in the Patterson) this will allow sufficient of the signs of the structure factors to be determined to allow a first model of the structure to be derived.

For some crystals it is possible to make an isomorphous replacement, which involves the substitution of one kind of an atom for another at some sites in the unit cell without appreciable displacement of the other atoms on other sites. Suppose that we write the scattering power for a centrosymmetric crystal containing atom type 1 as

$$I_1(u) = (F_1(u) + F_r(u))^2 = F_1^2 + 2F_1 F_r + F_r^2 \quad (29)$$

where F_1 is the contribution from the type 1 atoms and F_r is from the rest. Similarly when type 2 atoms are substituted for the type 1 atoms,

$$I_2(u) = (F_2(u) + F_r(u))^2 = F_2^2 + 2F_2 F_r + F_r^2. \quad (30)$$

Subtracting the two sets of intensities

$$I_1 - I_2 = F_1^2 - F_2^2 + 2(F_1 - F_2)F_r \quad (31)$$

so that, if the positions of the replacement atoms are known, F_1 and F_2 are known and F_r is determined in magnitude and sign. If the structure has no center of symmetry more than one substitution is required.

In a variation of this method no replacement of atoms is made but the atomic scattering factor for one type of atom is changed by changing the wavelength of the incident radiation to the vicinity of the absorption edge for the element.

6.4.4. Direct methods

The class of techniques known as "direct methods" are now the most widely used tools for solving structures having relatively small numbers of atoms in the asymmetric unit (up to about 150 atoms). Their extension to deal with larger structures, such as those of biological macromolecules, is being pursued but encounters severe difficulties. These techniques make use of algebraic equalities and inequalities between the structure amplitudes and are derived from the known properties of the electron density function such as its symmetry and the fact that $\rho(\mathbf{r})$ is a real positive function consisting of peaks of known shape and size. The important consideration is that these methods involve only the computerized manipulation of numerical data. Coupled with least-squares refinement procedures, they offer the possibility of almost automatic, fully computerized structure analysis.

Historically the first concept in this area was that of the Harker-Kasper [1948] inequalities, based on the Cauchy inequality relating to any sets of real or complex quantities, a_j and b_j :

$$|\sum_j a_j b_j|^2 \leq (\sum_j |a_j|^2) (\sum_j |b_j|^2). \quad (32)$$

This inequality is applied to the expressions for the unitary structure factors, defined by

$$U_h = \sum_j n_j \exp(2\pi i \mathbf{h} \cdot \mathbf{r}_j) \quad (33)$$

where $n_j = f_j / \sum_j f_j$ is the scattering factor normalized to remove the strong variation with scattering angle so that it corresponds approximately to the scattering from a point atom, and \mathbf{h} stands for the triple indices, h, k, l . Then for a structure with no symmetry, one derives $|U_h|^2 \leq 1$, which is trivial; but useful relations follow if symmetry elements are added. For example:

$$\begin{aligned} \text{For a center of symmetry; } & U_{2h}^2 \leq 0.5 + 0.5 U_{2h} \\ \text{For a 2-fold screw axis; } & |U_{h,k,l}|^2 \leq 0.5 + (-1)^k U_{2h,0,2l}. \end{aligned} \quad (34)$$

Thus, for a center of symmetry, if U_h is large enough, U_{2h} must be positive, and further definite phase assignments follow as further symmetry elements are added.

A further set of useful relationships follow when the Cauchy inequality is applied to $U_h \pm U_k$, where \mathbf{k} denotes a different set of hkl values, namely,

$$\begin{aligned} (|U_h| + |U_k|)^2 &\leq \{1 + s(\mathbf{h})s(\mathbf{k})s(\mathbf{h}+\mathbf{k})|U_{\mathbf{h}+\mathbf{k}}|\}. \\ \{1 + s(\mathbf{h})s(\mathbf{k})s(\mathbf{h}-\mathbf{k})|U_{\mathbf{h}-\mathbf{k}}|\}. & \end{aligned} \quad (35)$$

where $s(\mathbf{h})$ denotes the sign of the \mathbf{h} reflection. Then if the U_h are sufficiently large, either or both of the triple sign products must be positive. Given the signs of two of the unitary structure amplitudes involved, the sign of the other one is determined.

It is always possible to make a start on the use of this result because the signs of some structure amplitudes may always be assigned arbitrarily. The choice of these signs corresponds to the choice of the origin of the unit cell among the various centers of symmetry for a centrosymmetric structure. Guidance for the selection of the appropriate \mathbf{h} values for assignment of signs is given in the International Tables for Crystallography, Vol. B, (Giacovazzo, 1993). Further known signs, such as those derived from (34) may likewise be used to generate new signs.

In a different approach, Sayre [1952] developed an equality relating structure amplitudes. For a structure consisting of isolated, equal atoms, the structure amplitudes may be written

$$F_h = f_0 \sum_i \exp(2\pi i \mathbf{h} \cdot \mathbf{r}_i).$$

The square of the electron density distribution, $\rho^2(\mathbf{r})$, has peaks at the same positions and so, if it is Fourier transformed, one gets, similarly,

$$G_{\mathbf{h}} = (f_0 * f_0) \cdot \sum_i \exp(2\pi i \mathbf{h} \cdot \mathbf{r}_i) = [(f_0 * f_0) / f_0] \cdot F_{\mathbf{h}}$$

But the Fourier transform of $\rho^2(\mathbf{r})$ is also $F(\mathbf{u}) * F(\mathbf{u})$ so that

$$F_{\mathbf{h}} = [f_0 / (f_0 * f_0)] \cdot \sum_{\mathbf{k}} F_{\mathbf{k}} \cdot F_{\mathbf{h}-\mathbf{k}} \quad (36)$$

The sum over \mathbf{k} may well be dominated by the few terms involving pairs of large F values so that the the magnitude, or at least, the sign of $F_{\mathbf{h}}$ is probably given by summing over only a few such pairs.

The factor before the summation of (36) may be removed by using normalized structure amplitudes, $E_{\mathbf{h}}$, defined by

$$|E_{\mathbf{h}}|^2 = |F_{\mathbf{h}}|^2 / \epsilon \sum_i f_i^2$$

where ϵ is a weighting factor. Thus the sign of $E_{\mathbf{h}}$ is probably given by the sign of $\sum_{\mathbf{k}} E_{\mathbf{k}} E_{\mathbf{h}-\mathbf{k}}$ for particular \mathbf{k} values and the probability that this is so can be defined. Karle and Hauptman [1956] derived a general formula, the tangent formula, for the probable phases of triplets such as \mathbf{h} , \mathbf{k} , and $\mathbf{h}-\mathbf{k}$, and expressions for their probabilities which have become the basis for much of the present-day work with direct methods. For a general account of the origin and basis of the direct methods and the procedures used in their applications to crystal structure analysis, the reader is referred to Giacovazzo [1993] or Hauptman and Karle [1953]. Simpler but more limited descriptions are given, for example, by Ladd and Palmer [1993] or Schwartz and Cohen [1977].

6.5. Neutron diffraction structure analysis

6.5.1. Nuclear scattering

Compared with the X-ray case the main differences in the application of neutron diffraction for structure analysis from non-magnetic materials arise from the erratic variation of the scattering length in magnitude and sign with varying atomic number.

The light elements such as carbon, oxygen and even hydrogen scatter just as strongly as the heavy elements and so can be located with equal ease whereas for X-ray diffraction they are often undetectable. Particularly for hydrogen the X-ray scattering comes from a single, rather diffusely spread electron, or less if the hydrogen is partially ionized. For neutrons the scattering is from the

nucleus and is quite large and negative; hence the application of neutron diffraction for the important study of hydrogen positions in organic and biologically significant crystals.

The hydrogen positions are distinguished as negative peaks in the Fourier maps because of the negative b values. Similarly, negative peaks appear in the Patterson maps corresponding to vectors between atoms having positive and negative b values.

In practice hydrogen is not favored because it gives a high background of "incoherent" scattering (see Chapter 4). It is replaced when possible by deuterium which has a high positive b value and little background scattering.

A further limitation of X-ray diffraction is the difficulty of distinguishing between elements which are close together in the periodic table since the f values increase smoothly with atomic number. For a number of important cases the b values for neutrons differ greatly from one atom to the next and so allow them to be distinguished. For example for the transition elements we find $b_{Mn} = -0.36$, $b_{Fe} = 0.96$, $b_{Co} = 0.25$ and $b_{Ni} = 1.03$.

6.5.2. Magnetic scattering

For magnetic materials we must distinguish two atomic scattering factors for neutrons, the nuclear scattering amplitude b , independent of scattering angle, and the magnetic scattering amplitude p , depending on the distribution of unpaired electron spins and given by

$$p = (e^2\gamma/2mc^2) g J f, \quad (37)$$

where γ is the magnetic moment of the neutron, g is the Lande splitting factor, J is the spin-orbit quantum number and f is the form factor given by Fourier transform of the distribution of electrons having unpaired spins. But since the spin scattering involves the interactions of vector quantities we must also define a magnetic interaction vector q given by

$$q = \varepsilon(\varepsilon \cdot K) - K,$$

where ε is the unit vector perpendicular to the diffracting plane and K is the unit vector in the direction of the magnetic moment. Then $q = 0$ for ε parallel to K and $|q| = 1$ for ε perpendicular to K .

For an ordered array of spins in the lattice, the nuclear and magnetic scattering contribute independently to the diffracted intensity provided that the neutron beam is

unpolarized (Bacon [1975]) so that

$$|F(\mathbf{u})|^2 = |F_n(\mathbf{u})|^2 + |F_{\text{mag}}(\mathbf{u})|^2 \quad (38)$$

where

$$|F_n(\mathbf{u})|^2 = \left| \sum_i b_i \exp\{-M_i\} \exp\{2\pi i \mathbf{h} \cdot \mathbf{r}_i\} \right|^2 \quad (39)$$

$$|F_{\text{mag}}(\mathbf{u})|^2 = \sum_i \sum_j p_i q_i \cdot p_j q_j \exp\{-M_i - M_j\} \exp\{2\pi i \mathbf{h} \cdot (\mathbf{r}_i - \mathbf{r}_j)\}, \quad (40)$$

where we have inserted the Debye-Waller factors, $\exp\{-M_i\}$, arising from the thermal vibrations of the atoms (Chapters 7 and 12).

For a ferromagnetic material with all spins lined up parallel,

$$\begin{aligned} \mathbf{q}_i \cdot \mathbf{q}_j &= q^2 \\ &= \sin^2 \alpha, \end{aligned} \quad (41)$$

where α is the angle between \mathbf{K} and $\boldsymbol{\varepsilon}$.

For a simple anti-ferromagnetic case with spins on different sublattices of the structure aligned anti-parallel we have for some atom pairs that $\mathbf{K}_j = -\mathbf{K}_i$ so that $\mathbf{q}_i \cdot \mathbf{q}_j = -\sin^2 \alpha$. Then (40) may be written

$$|F_{\text{mag}}^2(\mathbf{u})|^2 = \sin^2 \alpha \sum_i \sum_j (\pm) p_i p_j \exp\{-M_i - M_j\} \exp\{2\pi i \mathbf{h} \cdot (\mathbf{r}_i - \mathbf{r}_j)\} \quad (42)$$

where the plus and minus signs refer to pairs of atoms with parallel and antiparallel spins. Correspondingly if the $\sin^2 \alpha$ factor is ignored, a Patterson function can be drawn with positive peaks for parallel pairs of spins and negative peaks for anti-parallel pairs.

For anti-ferromagnetic crystals the separation of like atoms, in the simplest case, into two sets, spin up and spin down, will almost inevitably lower the symmetry of the crystal and in many cases produces a unit cell size which is some multiple of that seen by X-rays. Then new, purely magnetic, "superlattice" reflections appear in the diffraction pattern. Gradually more and more complicated magnetic superlattice structures are being found with large repeat distances and spins inclined to each other at a variety of angles. For these, the simple formulation of (42) cannot be used.

6.6. Electron diffraction structure analysis

The atomic scattering factors for electrons increase smoothly with atomic number except at low scattering angles but not quite as rapidly as for X-rays. The difference is most apparent for hydrogen. The electron scattering depends on the potential field of the nucleus which is partially screened by the orbital electron. Partial ionization decreases the screening and increases the scattering factor. It has been estimated by Vainshtein [1964] that the ratio of the scattering by carbon and hydrogen is about 10 for X-rays but only 3 or 4 for electrons. However, in view of the ease of detection of hydrogen atoms by neutron diffraction the use of electron diffraction for these purposes is restricted to special cases for which neutron diffraction methods are not applicable.

One factor of interest is that the detection of hydrogen atoms is probably the one case in which the results should differ appreciably with the technique. X-ray diffraction should indicate the position of maximum electron density. Neutron diffraction should indicate the mean position of the nucleus. These positions may well be different when the atom is polarized as in an asymmetric bond. Then the peak of potential deduced from electron diffraction should be near to the position of the positive nucleus, but since the contribution of the electron cloud is negative, displacement of the electron distribution in one direction should displace the potential peak slightly in the opposite direction.

Determinations of the N-H bond length in NH_4Cl crystals give $0.94 \pm 0.03 \text{ \AA}$ by X-ray diffraction, $1.03 \pm 0.02 \text{ \AA}$ by neutron diffraction and $1.02 \pm 0.02 \text{ \AA}$ (Kuwabara [1959] or $1.04 \pm 0.02 \text{ \AA}$ (Avilov et al. [1973]) by electron diffraction.

Undoubtedly, the most important application of electron diffraction structure analysis is for the examination of crystals which are too small to be studied by any other method. In electron microscopes, the selected-area diffraction technique can be used to obtain single-crystal diffraction patterns from crystals from about 10 \AA up to several hundred \AA thick and having a lateral extent which is some reasonable fraction of the minimum practical size, which is about 1 \mu m for 100kV electrons and as small as 500 \AA for 500 to 1000kV electrons. By use of the very bright field-emission gun sources in scanning transmission electron microscopes it is possible to obtain very small electron probes at the specimen position and so to obtain diffraction patterns from regions 10 \AA or less in diameter. Then single-crystal patterns can be obtained from very small particles, such as the metal particles in supported catalysts, individual crystals of clay minerals, aerosols, colloids or other materials which give, at best, only very diffuse powder patterns with X-ray diffraction.

Because of the very strong dynamical diffraction effects occurring in electron diffraction patterns from most of the single crystals large enough to be readily visible in an electron microscope, the methods common for X-ray single-crystal structure analysis have been used very little until recently. An important group in the USSR carried out many structure analyses using mostly finely polycrystalline samples [see Vainshtein et al., 1992]. As will be discussed in Chapter 16, the averaging over a wide range of crystal orientations has the effect of reducing the dynamical diffraction effects or, within limits, of allowing a simple correction to be applied to produce almost kinematical intensities. A few structure analyses were attempted on the basis of spot patterns from severely bent or disordered crystals on the assumption that such defects also reduced the amount of dynamical diffraction [Cowley, 1967].

In recent years there has been rapid growth in the electron crystallography of organic materials [see Dorset, 1994, 1995]. Extensive thin single-crystal films of organic or biological materials, including biological membranes or other macro-molecular films only one or two molecules thick, have been prepared as for electron microscopy. Radiation damage of such materials prevents their direct imaging with high resolution, but diffraction information may be obtained from large areas, micrometers in diameter, with spots extending to angles corresponding to spacings of 1 or 2 Å. The diffraction patterns may be used as a basis for structure analysis directly, as for X-ray diffraction, or else to extend the information gained concerning the structure from relatively low-resolution electron images, as will be discussed in Chapter 13. For thin crystals with low-Z elements the kinematical approximation is usually a reasonable assumption. Estimates of the many-beam dynamical diffraction effects may be made using the computing methods outlined in Chapter 11.

Problems

1. A zinc single crystal has a hexagonal unit cell ($a = b \neq c$, $\alpha = \beta = 90^\circ$, $\gamma = 120^\circ$) with 3-fold symmetry about the c -axis. The zinc atoms have coordinates $0, 0, 0$ and $2/3, 1/3, 1/2$.

Find the axes and angles of the reciprocal lattice. Describe the form of the X-ray diffraction pattern recorded on a cylindrical film when the crystal is rotated about the c -axis which coincides with the axis of the cylindrical film. Find expressions for the intensities of the reflections in terms of the atomic scattering factors. Which reflections are "forbidden" i.e. have zero intensity as a result of space-group symmetries?

2. For a certain problem of structure analysis it is desirable to have a projection on the $x - y$ plane of the limited region of the unit cell lying between $z = 0.4$ and $z = 0.6$. What series should be summed to give this?

3. MnF_2 crystallizes with a tetragonal unit cell (SnO_2 -type) with $a = 4.865$, $c = 3.284 \text{ \AA}$: Space Group $D_{4h}^{14} - P_{42/mnm}$, $Z = 2$, with

Mn at $0, 0, 0; 1/2, 1/2, 1/2$
 F at $x, x, 0$ etc. with $x = 0.31$.

- What are the symmetry elements of the unit cell?
- What reflections are forbidden by symmetry?
- What series should be summed to give the electron density along a line through an Mn atom and a neighboring F atom in the $x - y$ plane?
- Sketch the form of the projection of the Patterson function on to the $x - z$ plane (z -axis is direction of c -axis of unit cell).

Note: Make use of International Tables of Crystallography, Vol. A (Hahn, 1992).

4. Below a certain critical temperature MnF_2 is anti-ferromagnetic with the spin of the Mn atom at $0, 0, 0$, directed along the positive c -axis and the spin of the atom at $1/2, 1/2, 1/2$, directed along the negative c -axis (see Bacon, p. 311).

- How will the symmetry and forbidden reflections for neutron diffraction differ from those for X-ray diffraction (c.f. problem 3)?
- Sketch the Patterson function for the projection on to the $x - z$ plane derived from neutron diffraction intensities including both nuclear and magnetic scattering.
- How do the various reflection intensities vary with temperature?

5. The use of an optical diffractometer to produce diffraction patterns from two-dimensional masks as an aid to visualizing processes in X-ray diffraction has been described by Taylor and Lipson [1964]. The masks may be made by drilling small holes in an opaque screen to represent atoms in the projections of molecular or crystal structures so that the diffraction pattern approximates to a section of reciprocal space. Alternatively, the holes may be regularly spaced to represent a section of reciprocal space. The areas of the holes are made proportional to $|F_{hkl}|^2$ values.

To what extent will the diffraction pattern of this mask represent the Patterson function of the crystal (a) near the center of the diffraction pattern, and (b) far from the center where the intensity is relatively small?

6. Two crystal structures may be assumed to be identical except that one atom per unit cell (which may be assumed to be at the origin) is different for the two cases. How will the two corresponding Patterson functions differ? How may the differences be used to determine the crystal structures unambiguously?

Diffraction from imperfect crystals

7.1. Formulation of the diffraction problem

7.1.1. Types of defects

The great variety of defects, imperfections, faults and other irregularities which may occur in real crystals provides a great variety of diffraction effects, including modifications of the sharp maxima about the reciprocal lattice points and also continuous distributions of intensity in the "background" regions between these maxima. These effects are of interest not only as an indication of the limitations of the ideal crystal model used for structure analysis purposes but also as one of the most powerful means available for the study of the nature of defects in crystals.

We distinguish two main classes of crystal defects: those for which it is possible to define an average periodic lattice and those for which this is not possible. This is an artificial and incomplete classification, but provides a useful starting point. The first class of defects includes mostly localized defects, including the single-atom or "point" defects such as vacancies, interstitial atoms, substitutions of atoms; also small clusters of point defects and the localized strain fields associated with point defects or clusters. For these cases the defect is surrounded by the three-dimensional bulk of crystal which defines an average periodicity, acting as a frame of reference and defining a basis from which deviations can be measured. The most commonly occurring example of a deviation of this type from an ideally periodic crystal lattice is provided by the thermal vibrations of atoms about their mean lattice positions.

In the second class of defects the simplest type of defect to appreciate is the planar fault or twin plane which cuts right across the crystal, giving the unit cells on one side a shift or change of orientation with respect to those on the other side. Then, while the shift or change of orientation between the two parts of the crystal may be readily defined, it is not feasible to refer both parts to one average lattice. For a progression of faults or twins occurring at more-or-less random intervals, certain definite geometric relationships may be maintained between various parts of the crystal, even though the full lattice periodicity is lost. Some well-defined sharp maxima of

scattering power occur in reciprocal space, but an average lattice is not relevant in that, however it is chosen, displacements from it equal to large fractions of the unit cell dimensions occur over a large proportion of the crystal.

In more complicated cases, faults may give progressive deviations from any locally defined regular lattice in two or three dimensions. As the frequency and variety of faults increases the distance over which any correlation between atom sites may be traced decreases and the crystallinity of the material decreases. In the limit, a liquid-like structure may be described in terms of a fault in the periodicity occurring every few unit cells in every direction. However, such extreme cases fall within the ill-defined boundary region between amorphous and polycrystalline materials. We confine our attention here to cases in which axial directions and principal periodicities may still be defined.

7.1.2. General diffraction formulation

Since for crystals with defects or faults it is not possible to use the special formulas for diffraction intensities developed in the last chapter for ideal crystals, it is necessary to revert to the earlier more general formulations of the diffraction problem. A common starting point is that of equation (5.5), from which the reciprocal space scattering power function is written in terms of the atomic scattering factors and positions; thus

$$F(\mathbf{u}) = \sum_i f_i \exp\{2\pi i \mathbf{u} \cdot \mathbf{r}_i\}. \quad (1)$$

$$J(\mathbf{u}) \equiv |F(\mathbf{u})|^2 = \sum_i \sum_j f_i f_j^* \exp\{2\pi i \mathbf{u} \cdot (\mathbf{r}_i - \mathbf{r}_j)\}. \quad (2)$$

It is not possible to insert into this expression the atomic scattering factors and positions of all atoms of the specimen. The intensity must be evaluated in terms of statistical relationships between these quantities.

One approach is to consider the average surroundings of an atom or of a particular type of atom. For example, if one atom of a statistically equivalent set is indicated by the subscript n , and there are N_n such atoms, we may take each atom in turn as the origin and find its average surroundings. Then

$$J(\mathbf{u}) = \sum_n N_n f_n \left\langle \sum_j f_{j-n} \exp\{2\pi i \mathbf{u} \cdot (\mathbf{r}_j - \mathbf{r}_n)\} \right\rangle. \quad (3)$$

where the brackets $\langle \rangle$ represent an averaging over all configurations of atoms about an atom of type n . For a monatomic solid, this reduces to

$$J(\mathbf{u}) = Nf^2 \sum_j \langle \exp\{2\pi i \mathbf{u} \cdot \mathbf{r}_j\} \rangle, \quad (4)$$

where \mathbf{r}_j is the position of an atom referred to an origin taken on another atom.

To be more general we should admit the possibility that the atomic positions should vary with time. Then, as we have seen before, for purely elastic scattering the diffracted amplitude is given by the time average of the structure so that

$$J_{el}(\mathbf{u}) = \left| \sum_i f_i \langle \exp\{2\pi i \mathbf{u} \cdot \mathbf{r}_i(t)\} \rangle_t \right|^2. \quad (5)$$

Usually for X-ray or electron diffraction the intensity measurement does not distinguish between purely elastic and inelastic scattering. The observed intensity is then a time average of the instantaneous intensities, or

$$\begin{aligned} \langle J_{tot}(\mathbf{u}, t) \rangle_t &= \left\langle \sum_i \sum_j f_i f_j \exp\{2\pi i \mathbf{u} \cdot (\mathbf{r}_i - \mathbf{r}_j)\} \right\rangle_t \\ &= \sum_i \sum_j f_i f_j \langle \exp 2\pi i \mathbf{u} \cdot (\mathbf{r}_i - \mathbf{r}_j) \rangle_t. \end{aligned} \quad (6)$$

In cases of practical significance the averaging over time can be considered completely equivalent to an averaging over space. Usually the region of the sample which can give coherent diffraction effects is limited in size to be very much smaller than the total illuminated size because of the limitations on coherence of the incident beam. The lateral coherence of the beam is usually limited to a few hundred Å by the convergence of the beam from a finite incoherent source. The coherence of the beam in the direction of propagation is limited by the spread in wavelengths to about 1 μm for X-rays and less for neutrons. Thus the total diffracted intensity from a sample may be regarded as the sum of intensities from a very large number of independent but statistically equivalent regions. This is the same as the sum of the intensities obtained from any one region at different times. In time any one region may take on all the possible configurations of the atoms present and for a large number of independent regions all possible configurations are represented at any one time.

7.2. Patterson function approach

7.2.1. Patterson with an average periodic structure

The discussion of the intensity expression in terms of the average surroundings of a given atom (eqs. (3) and (4))

suggests use of the Patterson function description and, in fact, this forms the basis for an alternative and very powerful method for studying the diffraction by imperfect crystals. We may apply directly the considerations of Ch. 5.

The two main classes of crystal imperfection which we have distinguished require somewhat different approaches and will be considered separately.

When an average periodic lattice may be defined we write

$$\rho(\mathbf{r}) = \langle \rho(\mathbf{r}) \rangle + \Delta\rho(\mathbf{r}), \quad (7)$$

where $\langle \rho(\mathbf{r}) \rangle$ is the electron density distribution for the average lattice, defined to be time independent and periodic (when we ignore shape-function limitations) and $\Delta\rho(\mathbf{r})$ represents the deviation from the average lattice, which is essentially non-periodic. This deviation from the average lattice may be time-dependent or not and the averaging process may be an averaging over time or over space. As suggested above, the two cases may be considered equivalent. By definition, for the averaging process involved, $\langle \Delta\rho \rangle = 0$.

From (7) we may write the Patterson function as

$$\begin{aligned} P(\mathbf{r}) &= \{ \langle \rho(\mathbf{r}) \rangle + \Delta\rho(\mathbf{r}) \} * \{ \langle \rho(-\mathbf{r}) \rangle + \Delta\rho(-\mathbf{r}) \} \\ &= \{ \langle \rho(\mathbf{r}) \rangle * \langle \rho(-\mathbf{r}) \rangle \} + \{ \langle \rho(\mathbf{r}) \rangle * \Delta\rho(-\mathbf{r}) \} \\ &\quad + \{ \Delta\rho(\mathbf{r}) * \Delta\rho(-\mathbf{r}) \} + \{ \Delta\rho(\mathbf{r}) * \langle \rho(-\mathbf{r}) \rangle \}. \end{aligned} \quad (8)$$

In terms of spatial averaging of time-independent functions, the second term may be written

$$\langle \rho_0(\mathbf{r}) \rangle * \sum_{\mathbf{r}} \delta(\mathbf{r} - \mathbf{R}_n) * \Delta\rho(-\mathbf{r}),$$

where \mathbf{R}_n is a lattice vector of the average lattice. Then the convolution of $\Delta\rho$ with $\sum \delta(\mathbf{r} - \mathbf{R}_n)$ represents the superposition of the function $\Delta\rho$ with all lattice vector shifts. There is thus an averaging of $\Delta\rho$ values at every point of the unit cell and by definition this is zero. The same applies to the convolution in the fourth term of (8). Hence

$$P(\mathbf{r}) = \{ \langle \rho(\mathbf{r}) \rangle * \langle \rho(-\mathbf{r}) \rangle \} + \{ \Delta\rho(\mathbf{r}) * \Delta\rho(-\mathbf{r}) \}. \quad (9)$$

The same argument holds if the averaging is over time rather than space to give the time-average of the instantaneous Patterson function $P(\mathbf{r}, 0)$. Fourier transform of (9) gives

$$J(\mathbf{u}) = |\bar{F}|^2 + |\Delta F|^2, \quad (10)$$

where \bar{F} and ΔF are the Fourier transforms of $\langle \rho(\mathbf{r}) \rangle$ and $\Delta\rho$.

Thus we have the general result that the total scattering power distribution is the sum of that for the average lattice and that for the deviations from the average lattice considered separately. Since $\langle \rho(r) \rangle$ is periodic, $|\bar{F}|^2$ consists of sharp peaks at the reciprocal lattice points only, giving the sharp Bragg reflections in diffraction patterns. Since $\Delta\rho$ is non-periodic, $\Delta\rho * \Delta\rho$ is non-periodic and decreases rapidly with distance from the origin. Hence $|\Delta F|^2$ represents a continuous distribution of scattering power between the reciprocal lattice points and so gives rise to diffuse scattering in diffraction patterns. It may be noted that for time dependent perturbations of an average periodic structure, the first terms of (9) and (10) represent scattering from the time average structure and so correspond to the purely elastic scattering while the second term represents the inelastic scattering.

7.2.2. Patterson function with no average structure

For the second class of crystal imperfections, an average periodic lattice can not be meaningfully defined and there can not be the same separation into sharp Bragg reflections and diffuse scattering. Then, in general, the Patterson function must be evaluated for the whole structure or the intensity expression (2) must be evaluated using any simplifications which seem appropriate for the particular case in hand.

For a large number of situations the structure may be described in terms of the repetition of one or more definite units of structure with irregular translations. For example the structure may be made up of identical planes of atoms which have an irregular interplanar spacing; or identical lines of atoms which are lined up parallel but have irregular spacings in the two directions at right angles to their lengths; or identical blocks of atoms separated from their neighbors by variable spacings in three dimensions. Then the electron density distribution may be written

$$\rho(r) = \rho_0(r) * d(r), \quad (11)$$

where the "distribution function", $d(r)$, is a set of delta functions representing the positions of reference points for each reproduction of the unit of structure $\rho_0(r)$. Then the Patterson is

$$P(r) = \rho_0(r) * \rho_0(-r) * d(r) * d(-r), \quad (12)$$

and the reciprocal space distribution of scattering power

$$J(u) = |F_0(u)|^2 |D(u)|^2. \quad (13)$$

Thus the scattering power function for the individual unit of structure may be calculated first and then multiplied by the function corresponding to the distribution function.

If there is more than one type of structural unit involved, we may write

$$\rho(r) = \sum_n \rho_n(r) * d_n(r). \quad (14)$$

Then the $P(r)$ and $J(u)$ expressions involve cross-product terms and are correspondingly more complicated.

In order to illustrate the treatment of diffraction problems for the two classes of crystal imperfections we have discussed, we proceed to work out a few simple examples using both the direct calculation of intensities from equation (2) and the calculation via the Patterson. This will then serve as a basis for the discussion of the more complicated problems to be met in the later chapters of Section IV.

7.3. Deviations from an average lattice

7.3.1. Random vacancies: no relaxation

As a first simple example we consider a monatomic crystal of simple structure. There are N atom sites but a number n of them, distributed at random, are vacant. We ignore any displacement of the atoms around the vacancies from their equilibrium lattice sites (relaxation).

Then the scattering power $J(u)$ can be evaluated from equation (2) by considering in turn the various possible vectors $r_i - r_j$.

For $i = j$ and $r_i - r_j = 0$, there are $N - n$ vectors for which $f_i = f_j = f$ and n vectors for which $f_i = f_j = 0$, so that the contribution is $(N - n)f^2$.

For any other vector length there is a probability $(N - n)/N$ of having an atom at i and $(N - n)/N$ of having an atom at j so that the contribution to $J(u)$ is

$$\left(\frac{N-n}{N}\right)^2 f^2 \exp\{2\pi i u \cdot (r_i - r_j)\}.$$

Then

$$\begin{aligned} J(u) &= (N - n)f^2 + \left(\frac{N-n}{N}\right)^2 f^2 \sum_{i \neq j} \sum_j \exp\{2\pi i u \cdot (r_i - r_j)\} \\ &= \left(\frac{N-n}{N}\right)^2 f^2 \sum_i \sum_j \exp\{2\pi i u \cdot (r_i - r_j)\} + (N - n)f^2 - \frac{(N-n)^2}{N} f^2. \end{aligned}$$

Here, we have removed the restriction $i \neq j$ on the summation by adding and subtracting a zero term. Then

$$J(u) = \left(\frac{N-n}{N}\right)^2 f^2 \sum_i \sum_j \exp\{2\pi i u \cdot (r_i - r_j)\} + \frac{n(N-n)}{N} f^2. \quad (15)$$

The first term of (15) represents the scattering power for a perfectly ordered lattice with no defects but with the average scattering factor of the atoms reduced in the ratio $(N-n)/N$. The second term is a continuous distribution of scattering power in reciprocal space, falling off with $|u|$ as f^2 and of strength proportional to the number of defects (if $n \ll N$). This is equal to the scattering from n independent, isolated atoms.

Alternatively we can follow the formulation represented by equations (7), (9) and (10). The average structure is periodic with a fraction $(N-n)/N$ of $\rho_0(r)$ at each lattice point. The corresponding contribution to the distribution $|\bar{F}|^2$ is that of the crystal with no vacancies, but with each sharp peak at a reciprocal lattice point reduced in weight by

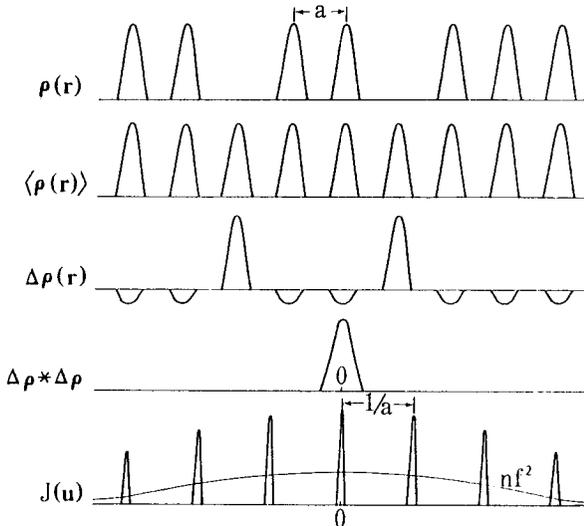


Fig. 7.1. One-dimensional diagrams representing the electron density distribution, $\rho(r)$ for a structure having random vacancies, the periodic average structure $\langle \rho(r) \rangle$, the deviations from the average structure $\Delta \rho(r)$, the Patterson function for this deviation function, and the scattering power distribution in reciprocal space showing sharp peaks plus diffuse scattering.

a factor $(N - n)^2/N^2$. The deviation from the average lattice is $(N - n) \rho_0(r)/N$ for each site where there is a vacancy and $-n\rho_0(r)/N$ for each site where there is no vacancy, as suggested in Fig. 7.1. Then the origin peak of the Patterson of $\Delta\rho$ is

$$\begin{aligned} \rho_0(r) * \rho_0(r) & \left[n \left(\frac{N-n}{N} \right)^2 + (N-n) \left(\frac{n}{N} \right)^2 \right] \\ & = \frac{n(N-n)}{N} \rho_0(r) * \rho_0(r). \end{aligned} \quad (16)$$

For any other Patterson peaks there are n contributions of weight

$$\frac{n}{N} \left\{ \left(\frac{N-n}{N} \right) \left(-\frac{n}{N} \right) + \left(\frac{n}{N} \right) \left(\frac{N-n}{N} \right) \right\},$$

when vacancy sites are taken as origin; and $N-n$ contributions of weight

$$\left(\frac{N-n}{N} \right) \left\{ \left(\frac{N-n}{N} \right) \left(-\frac{n}{N} \right) + \left(\frac{n}{N} \right) \left(\frac{N-n}{N} \right) \right\},$$

when an atom is taken at the origin. The factor in curly brackets in each case, representing a random weighting of peak heights according to the occupancies, is equal to zero.

Hence the defect Patterson consists of the origin peak only. The total scattering power distribution is then $(N - n)^2/N^2$ times that for a crystal with no vacancies plus, from (16),

$$\mathfrak{S}\{\Delta\rho * \Delta\rho\} = \frac{n(N-n)}{N} f^2,$$

as in equation (15).

7.3.2. Clustered vacancies

The particular advantage of the formulation of scattering via the Patterson is the possibility of obtaining an immediate appreciation of the form of the diffuse scattering for any particular type of defect and, in many cases, rapid quantitative estimates of the magnitude of the scattering power function.

We consider, for example, a case similar to (a) except that the distribution of vacancies is not random but shows some preference for the vacancies to clump together. To be specific we consider that the vacancies occur in pairs

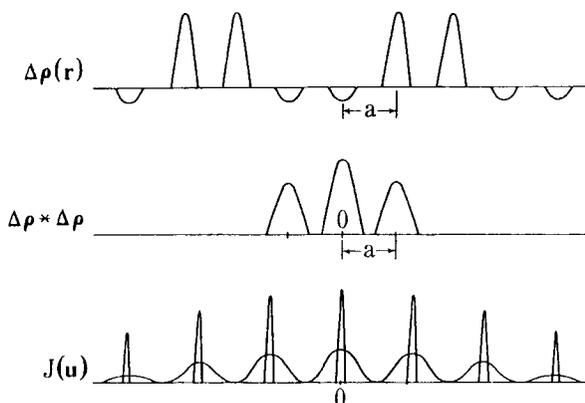


Fig. 7.2. The deviation from the average structure, the corresponding Patterson function and the scattering power distribution for a case such as in Fig. 7.1 when the vacancies occur in pairs.

parallel to one axis. For the same total number of vacancies, n , the average structure is exactly the same as for (a) so that $|\bar{F}|^2$ is the same.

The origin peak of $\Delta\rho * \Delta\rho$ is the same as for (a) and there is no correlation for interatomic vectors of greater than the nearest-neighbor distance but, as suggested in Fig. 7.2, there are peaks at nearest neighbor distances of weight $n(N-n)/2N$ since the probability of finding a vacancy at distance a from a given vacancy is half that of finding a vacancy at distance zero. The diffuse scattering is then modulated with a periodicity a^{-1} , thus;

$$|\Delta F(u)|^2 = \frac{n(N-n)}{N} f^2 (1 + \cos\{2\pi au\}). \quad (17)$$

If lines of more than two vacancies may occur in the one direction, the correlation will give peaks in the defect Patterson extending further from the origin and the modulation of the diffuse scattering function will be more sharply peaked around the reciprocal lattice positions. Similarly for a tendency for the formation of three-dimensional clumps of vacancies the Patterson shows correlations in three dimensions and the modulation of the diffuse part of the intensity appears in all directions.

With very little change, this simple theory may be applied to determine the scattering from crystals with random or clustered interstitial atoms or with occasional impurity atoms substituting for regular lattice atoms.

7.3.3. Lattice relaxation

We have neglected so far the effect of point defects on the surrounding atoms. For interstitials and, to a lesser extent, for vacancies, the neighboring atoms may be displaced from their average lattice sites by an appreciable fraction of the unit cell dimensions. This relaxation of the surrounding lattice may affect a large number of atoms and give rise to diffuse scattering effects which are more pronounced than the scattering from the defects themselves which we have considered above.

For simplicity we consider an idealized one-dimensional case in which n small interstitial atoms of negligible scattering factor are inserted at random in the lattice halfway between two atoms. These nearest-neighbor atoms are displaced a distance $a/4$ but no other atoms have appreciable displacement, as suggested by Fig. 7.3.

Then the average lattice has peaks of $(N - n) \rho_0(r)/N$ at each lattice site and subsidiary peaks of $n \rho_0(r)/N$ at distances $\pm a/4$ from the lattice sites. Hence $F(u)$ is modulated by a factor

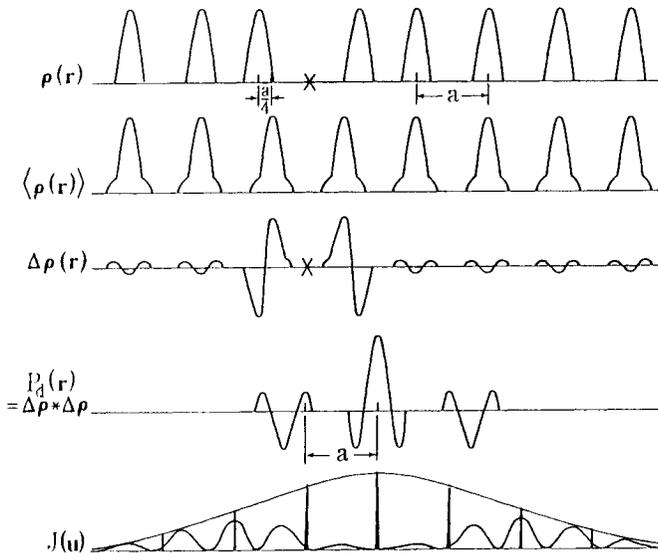


Fig. 7.3. The same set of functions as in Fig. 7.1 for the case of a one-dimensional structure in which a light interstitial atom at X displaces its nearest neighbors by a distance $a/4$.

$$\frac{N-n}{N} + 2\frac{n}{N} \cos 2\pi(au/4),$$

and the sharp peaks of scattering power in $|F(\mathbf{u})|^2$ are multiplied by

$$\left\{ 1 + 2\frac{n}{N-n} \cos 2\pi(au/4) \right\}^2 \left(\frac{N-n}{N} \right)^2.$$

By considering the vectors between positive and negative peaks of $\Delta(\mathbf{r})$ it is seen that the defect Patterson has a positive peak of relative weight 4 at the origin and negative peaks of relative weight 2 at $x = \pm a/4$. Then there are the same grouping of peaks, inverted, around $x = \pm 5a/4$. Thus the diffuse scattering distribution is of the form

$$\frac{n(N-n)}{N} f^2 \left(1 - \cos 2\pi \frac{au}{4} \right) \left\{ 1 - \cos 2\pi \left(\frac{5au}{4} \right) \right\}. \quad (18)$$

It is noted that, since the integrated value of $\Delta\rho * \Delta\rho$ is zero, the diffuse scattering function is zero at the reciprocal space origin. Apart from the f^2 fall-off, the maximum diffuse scattering comes at a position corresponding to the inverse of the displacements of the atoms.

This illustrates the general principle that diffuse scattering power arising from the displacement of atoms, without any change of scattering factors, is zero at the reciprocal lattice origin and increases with scattering angle. At the same time the sharp Bragg reflections are reduced by a factor which is unity for $|\mathbf{u}| = 0$ and decreases as $|\mathbf{u}|$ increases.

In more realistic models, the relaxation of atoms around a point defect is not limited to the nearest neighbor atoms but there are displacements of atoms decreasing gradually in three dimensions with distance from the center of dilation or contraction. Then the correlations in the defect Patterson extend to greater distances. The diffuse scattering power shows a steady overall increase with $|\mathbf{u}|$, apart from the fall-off with f^2 , and tends to have local maxima near to the reciprocal lattice point positions. The decrease of the sharp peaks with angle which is added to the f^2 decrease, can be expressed to a first approximation as $\exp\{-\beta u^2\}$ and so has roughly the same form as the Debye-Waller factor due to thermal vibrations (see also Chapter 12). This results because, when all the atom displacements are taken into account, the peaks of the average lattice $\langle \rho(\mathbf{r}) \rangle$ are spread out as if by convolution with something like a gaussian function.

7.3.4. Thermal vibrations - Einstein model

Finally, as an example involving a time-dependent perturbation of the lattice, we consider the case of a simple monatomic lattice in which all atoms independently vibrate about their mean lattice positions. In the harmonic approximation we may assume that the time average of the electron density function around each lattice point is given by spreading out the atom with a gaussian spread function.

In one dimension, for simplicity, we write

$$\langle \rho(x) \rangle_t = \sum_n \delta(x-na) * \rho_0(x) * (\pi b^2)^{-1/2} \exp\{-x^2/b^2\}. \quad (19)$$

Then

$$\bar{F} = F_0(u) \sum_h \delta(u-h/a) \exp\{-\pi^2 b^2 u^2\}.$$

and

$$|\bar{F}|^2 = |F_0(u)|^2 \exp\{-2\pi^2 b^2 u^2\} \sum_h \delta(u-h/a). \quad (20)$$

Thus if the root mean-square deviation of an atom from its lattice site is b , the intensities of the sharp reflections are decreased by the Debye-Waller factor $\exp\{-2\pi^2 b^2 u^2\}$.

As a slight variation of our former procedure, we now find the diffuse, inelastic scattering as the difference between the total scattering and the elastic scattering given by the average structure.

If all atoms are spread out by convolution with the gaussian function used in (19) and there is no correlation between the movements of different atoms, then the distribution of lengths of the interatomic vector between two atoms depends on the spread of positions of both atoms. Then the Patterson peaks are spread out by convolution by

$$\begin{aligned} & (\pi b^2)^{-1/2} \exp\{-x^2/b^2\} * (\pi b^2)^{-1/2} \exp\{-x^2/b^2\} \\ & = (2\pi b^2)^{-1/2} \exp\{-x^2/2b^2\}. \end{aligned} \quad (21)$$

This is true of all Patterson peaks except the peak at the origin since the vector from the center of an atom to itself is always zero (each atom "sees" itself at rest). Thus the total time-average of the instantaneous Patterson function may be written

$$\begin{aligned} P(x, 0) &= N\rho_0(x) * \rho_0(x) \\ &* \left[\delta(x) + \sum_{n \neq 0} \delta(x-na) * (2\pi b^2)^{-1/2} \exp\{-x^2/2b^2\} \right], \end{aligned} \quad (22)$$

and the scattering power is

$$\int J(u, v) dv = N|F_0|^2 \exp\{-2\pi^2 b^2 u^2\} \sum_h \delta(u - ha) + N|F_0|^2 [1 - \exp\{-2\pi^2 b^2 u^2\}]. \quad (23)$$

Here the first term is the elastic scattering from the average lattice as we found in equation (20). The second term is the diffuse inelastic scattering. The form of this is illustrated in Fig. 7.4.

As we have seen in the case of static displacements of atoms from their mean lattice positions, the effect of correlation between atom displacements will be to modulate the diffuse background scattering. If neighboring atoms tend to move in the same direction at the same time, as in acoustic-mode lattice vibrations, the diffuse scattering tends to peak at the positions of the reciprocal lattice points.

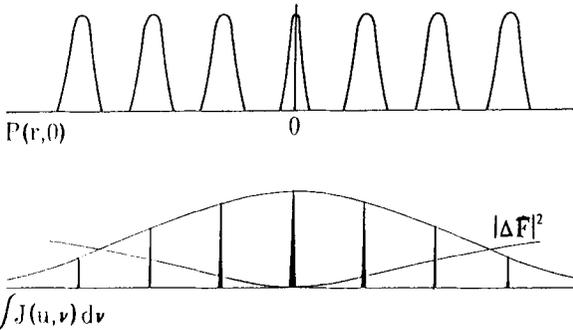


Fig. 7.4. One-dimensional diagrams illustrating the Patterson function and corresponding scattering distributions for uncorrelated atomic vibrations (Einstein model).

7.4. Imperfect crystals with no average lattice

7.4.1. Uneven separation of lattice planes

One of the few cases for which a non-repetitive structure can be simply defined and allows a simple analytical solution for the diffraction problem is that of parallel identical planes of atoms (or layers composed of several planar arrays of atoms) stacked in such a way that the spacing between them is not regular. This has some relevance in practice in that it may be considered a reasonable model for some crystals of clay minerals, for example, in which the two-dimensionally infinite sheets of atoms, consisting of tightly-bonded close-packed oxygen

arrays with metal atoms in octahedral or tetrahedral sites, have only weak bonding between them. Then a variation of the number or nature of the ions or molecules lying in the spaces between these sheets may modify the distances between the sheets without affecting their relative orientations.

Then, following (11), the structure may be described in terms of the electron density distribution of the individual sheets of atoms, $\rho_0(r)$ and a one-dimensional distribution function, $d(z)$, a set of delta functions giving the positions of the equivalent reference points within the sheets:

$$\rho(r) = \rho_0(r) * d(0, 0, z). \quad (24)$$

If we can assume $\rho_0(r)$ to be known, the problem is then to define $d(z)$ or $d(z) * d(z)$ and deduce $|D(u)|^2$ for use in (13).

The function $\rho_0(r)$ is considered to have periodicities a , b in the x and y directions but to be non-periodic and of limited extent (say 5 to 15 Å) in the z -direction. Then $|F_0(u)|^2$ will consist of a set of continuous lines parallel to the w -axis of reciprocal space with regular spacings a^* , b^* in the u and v directions. The variations of scattering power along these lines will depend on the relative positions of the atoms within the sheets.

To find a model for $d(z) * d(-z)$, we assume that the spacing between any two sheets of atoms has no influence on the spacing of any other sheets. Further we assume that the spacings between adjacent planes show a gaussian distribution about some mean value, c . Then the one-dimensional correlation function $d(z) * d(-z)$ has a form illustrated in Fig. 7.5.

For $z = 0$ we have a delta function of weight N corresponding to the zero distance of each reference point from itself. Around $z = \pm c$ there is a gaussian peak of half-width say, γ , corresponding to the distribution of nearest neighbor distance. Around $z = \pm 2c$ there will be a broader gaussian peak since for each position of the sheet which is nearest neighbor to a given sheet there will be a gaussian distribution of distances to the next, or second nearest neighbor sheet. The second nearest neighbor then has a distribution of positions, relative to the origin position, of

$$\begin{aligned} & (\pi\gamma^2)^{-1/2} \exp\{-z^2/\gamma^2\} * (\pi\gamma^2)^{-1/2} \exp\{-z^2/\gamma^2\} \\ & = (2\pi\gamma^2)^{-1/2} \exp\{-z^2/2\gamma^2\}. \end{aligned} \quad (25)$$

Similarly for the n th-nearest neighbor sheets to a given sheet, the spread of positions is given by an n -fold convolution of the nearest-neighbor distribution. Thus the

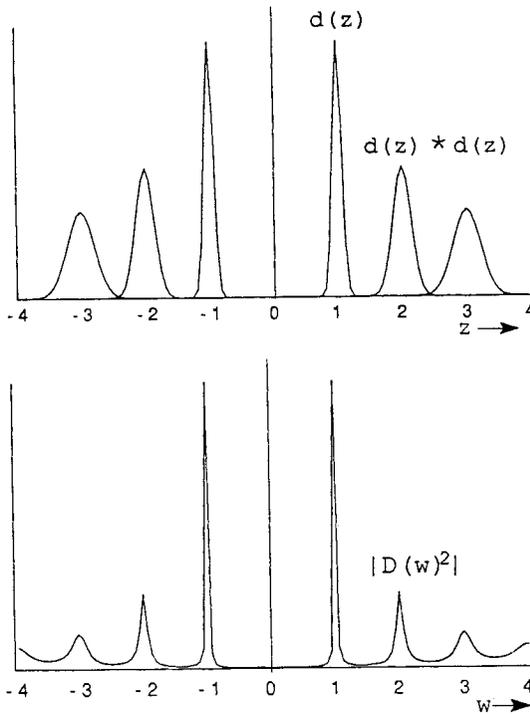


Fig. 7.5. The distribution function Patterson and its Fourier transform for the case of parallel planes of atoms having random variations of the inter-planar spacing.

Patterson of $d(z)$ may be written

$$d(z) * d(-z) = \sum_{n=-\infty}^{\infty} \delta(z - nc) * \frac{1}{(\pi|n|\gamma^2)^{1/2}} \exp\{-z^2 / |n|\gamma^2\}. \quad (26)$$

We note the need to insert $|n|$ in order to make the distribution symmetrical. Fourier transform of this gives

$$\begin{aligned} |D(w)|^2 &= \sum_n \exp\{-\pi^2 w^2 |n|\gamma^2\} \exp\{2\pi i w n c\} \\ &= \sum_{n=-\infty}^{\infty} \exp\{-(\pi^2 w^2 \gamma^2 \mp 2\pi i w c) |n|\}, \end{aligned} \quad (27)$$

where the minus sign refers to positive n .

Using the relationship $\sum_0^{\infty} x^n = (1 - x)^{-1}$, this becomes

$$\begin{aligned}
 |D(w)|^2 &= [1 - \exp\{-\pi^2 w^2 \gamma^2 + 2\pi i w c\}]^{-1} \\
 &\quad + [1 - \exp\{-\pi^2 w^2 \gamma^2 - 2\pi i w c\}]^{-1} - 1 \\
 &= \frac{1 - \exp\{-2\pi^2 w^2 \gamma^2\}}{1 + \exp\{-2\pi^2 w^2 \gamma^2\} - 2 \exp\{-\pi^2 w^2 \gamma^2\} \cos 2\pi w c}. \quad (28)
 \end{aligned}$$

Fig. 7.5 shows the form of this function. There is a delta function at $w = 0$, then maxima at $w = l/c$ for integral l . The heights of the maxima are given by

$$\left| D\left(\frac{l}{c}\right) \right|^2 = \frac{1}{l^2} \frac{2c^2}{\pi^2 \gamma^2} \quad \text{if} \quad l^2 \gamma^2 \ll c^2. \quad (29)$$

In between the maxima, for $w = (2l+1)/2c$, the minima have values which increase initially with w^2 . The widths of the maxima increase approximately as l^2 .

The total distribution of scattering power is then given by multiplying this function by the values of $|F_0(u)|^2$. It therefore consists of a set of regularly-spaced lines parallel to the w direction having sharp points on the $u-v$ plane and increasingly diffuse maxima as the distance from this plane increases.

This result gives an indication of what might be expected for more complicated cases for which the irregularity of spacing occurs in more than one dimension. Cases of one-dimensional disorder where the stacking disorder is introduced by translations in arbitrary directions are considered in Chapter 18. A two-dimensional equivalent of the above would be the case of long rod-like molecules packed together so that the ordering is near-perfect in the direction of the rods, with equivalent atoms in all rods lying in the same plane, but with an irregular spacing between the rods introduced by a random variation of side-groups of atoms attached to the molecules. An example in three dimensions would be provided by the packing of large molecules together into a lattice which is irregular because the presence of disordered side-groups or absorbed atoms gives a variation of the distance between molecules in all directions. In each case the maxima of scattering power in reciprocal space become progressively broader in the directions of the real-space irregularities.

7.4.2. Disordered orientations

When the disorder of the units of structure involves a relative rotation instead of, or in addition to, a relative displacement between neighbors, the use of a distribution function to simplify the Patterson is no longer possible and either the intensity expression (2) or the Patterson function must be derived from first principles.

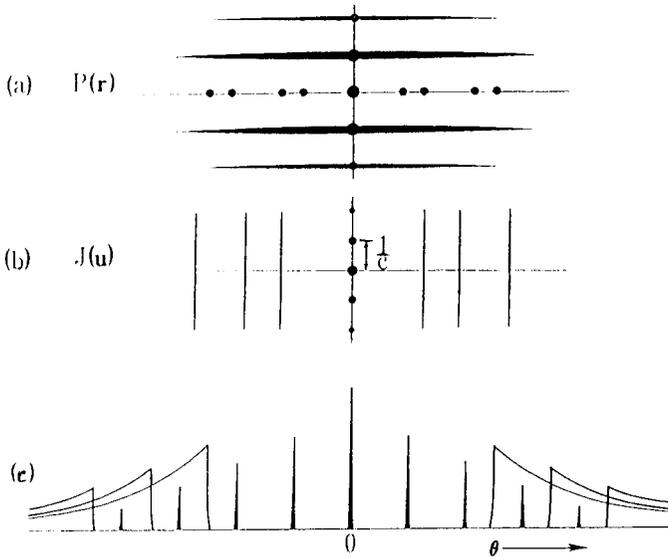


Fig. 7.6. (a) The Patterson function, and (b) the corresponding distribution in reciprocal space for a turbostratic structure of parallel, equally-spaced planes of atoms, having arbitrary relative orientations of the axes within the planes. The diagram (c) suggests the form of the powder pattern given by such a structure, consisting of sharp rings corresponding to the $(0\ 0\ l)$ reflections and broad asymmetrical peaks corresponding to the $(h\ k\ l)$ reflections with intensity a continuous function of the variable l , but with integral h, k values.

We give as an example of such situations the case of the "turbostratic" structure which has been used as a model for considering the diffraction from poorly-crystalline carbon black. Planar sheets of carbon atoms such as occur in the graphite structure are assumed to be stacked together with a constant distance, c , between the sheets as the only degree of regularity present. There is no relationship between the relative rotations or translations within the planes of adjacent sheets (Biscoe and Warren [1942]).

The Patterson function can be sketched immediately as in Fig. 7.6(a) where we show a planar section perpendicular to the plane of the sheets. In the plane through the origin of $P(r)$ parallel to the sheets, say the x - y plane, each sheet gives the two-dimensionally regular set of Patterson peaks corresponding to its hexagonal structure. But since the orientations in this plane are random, the Patterson for a single sheet will be rotated to give a set of sharp rings in the x - y plane. The vectors with z coordinates equal to nc , for $n \neq 0$, all have equivalent weight since the probability of an interatomic vector occurring is independent of the x, y

coordinates. Hence the Patterson contains a set of structureless planes of density at equal intervals, c , parallel to, but not including, the x - y plane through the origin. The relative weighting of these planes falls off with distance from the origin if the finite extent of the assemblies of sheets is taken into account.

In taking the Fourier transform to get $J(u)$ we divide $P(r)$ into two parts. A set of parallel structureless planes of spacing c , including one through the origin of $P(r)$, gives a set of sharp peaks of separation $1/c$ in the w direction in reciprocal space. Subtracting the plane through the origin of $P(r)$ gives a negative continuous line in the w direction through the origin of $J(u)$ but this is compensated by the general positive overall back-ground intensity arising from the origin peak of $P(r)$. The series of sharp rings in the x - y plane of $P(r)$ gives a set of concentric cylinders in reciprocal space with the w direction as axis. These show up as sets of parallel lines in the section of $J(u)$ drawn in Fig. 7.6(b).

Since the assemblies of parallel sheets may be assumed to be of limited extent, the sharp peaks and the cylinders of scattering power will be broadened by convolution with appropriate shape-transforms. If the distance between sheets is not quite regular, the sharp peaks on the central line of $J(u)$ will be progressively broadened in the way suggested by Fig. 7.5(b).

The appearance of a powder pattern obtained from a large number of randomly oriented assemblies of sheets may be derived by considering the $J(u)$ function to be rotated about the reciprocal lattice origin and then taking the intersection with the Ewald sphere. The result is suggested in Fig. 7.6(c). The sharp peaks give sharp rings in the powder pattern. The cylinders of scattering power give broadened rings with high intensity at the minimum scattering angle corresponding to the radius of the cylinder and intensity falling off to higher angles more and more slowly. These components, sketched separately in Fig. 7.6(c), are added together to give the total intensity.

The complete randomness we have assumed in the relative orientations and translations of adjacent sheets represents an over-simplification. Any correlations between sheets have the result of introducing some structure into the planes in $P(r)$ which do not pass through the origin with a consequent modulation of the scattering power in the cylinders of $J(u)$ and a modification of the simple form of the broad peaks in the powder pattern.

More complicated forms of disorder in orientation and relative translation in one, two or three dimensions arise frequently in poorly crystallized materials. We have attempted to demonstrate the value of the Patterson function approach in offering a rapid and usually adequate means for describing the state of order, for determining the resultant form of the reciprocal space distribution and so of the observable intensities. In Section IV we will apply these methods and

also the more conventional ones to the detailed study of several cases of diffraction from imperfectly ordered systems which are of particular interest.

Problems

1. Consider a one-dimensional array of atoms in the form of regularly spaced AB molecules where the separation of the A and B atoms in a molecule is one third of the separation of the centers of the molecules. At intervals of approximately T seconds, each molecule flips end-for-end (flipping time very short compared with T). There is no correlation in the flips of adjacent molecules. Find the elastic and inelastic scattering from this array and the energy changes which would be observed if sufficient energy resolution were available.

2. It is postulated that extra atoms may be introduced into the lattice of a face-centered cubic metal by the formation of a "split-interstitial"; e.g. the atom with fractional coordinates $1/2, 1/2, 0$ is replaced by two atoms having coordinates $1/2, 1/2, 1/4$ and $1/2, 1/2, -1/4$. In a crystal containing N atoms, a small number, n , of split interstitials is introduced, distributed at random in position and at random among the three symmetrically equivalent directions. How will this modify the sharp Bragg reflections and what diffuse scattering will be produced? (Ignore lattice relaxation effects.)

3. A complex oxide has a "tungsten-bronze" type of structure consisting for the most part of an almost close-packed array of oxygens with metal atoms in octahedral positions. Along the c -axis of the unit cell there is an open channel of about 4 Å diameter containing no oxygen atoms. In this channel, there are equivalent sites at $0, 0, 1/4$ and $0, 0, 3/4$ in which a heavy metal atom may sit. In each such channel (one per unit cell) a heavy atom sits at one or other of these two sites but not at both. The choice of site varies at random from one channel to the next. How does this random distribution affect the sharp Bragg reflections and what diffuse scattering is produced? How will the diffraction effects be affected if there is some tendency for ordering, such that, if a heavy atom sits at $0, 0, 1/4$ in one channel, this site will tend to be empty in neighboring channels?

4. Long straight identical molecules having a periodic structure along their lengths, are stacked together in a close-packed (hexagonal) array, but there are no correlations of the positions of the molecules in the direction of their lengths. Sketch the Patterson function for this array and deduce the form of the distribution of scattering power in reciprocal space. How are these functions in real and reciprocal space affected if there are irregular variations in the distances between molecules so that no average hexagonal stacking arrangement can be defined for more than a few repeat distances?

This Page Intentionally Left Blank

Section III

DYNAMICAL SCATTERING

This Page Intentionally Left Blank

Diffraction by perfect crystals

8.1. Multiple coherent scattering

Scattering from any three-dimensional object must, in principle, involve multiple scattering processes. Radiation scattered from one part of the object will pass through other parts of the object and will be scattered again.

We have seen that the kinematical, or single-scattering, approximation is very useful and has reasonable validity for a wide range of scattering experiments. Now, it is time to investigate the more general case of dynamical scattering in which the coherent interaction of multiply-scattered waves is taken into consideration.

For incoherent sources of finite dimensions and for an appreciable range of wavelengths, the observed intensities may be obtained by summing, incoherently, the intensities for individual point sources and single wavelengths. Hence we may start by considering perfectly coherent incident radiation for which the amplitudes of all multiply-scattered radiation are added.

If the atoms of the sample are spaced with complete regularity, the relative phases of the waves scattered by the different atoms are definable so that wave amplitudes may be added in a meaningful way. Any disorder in the atom positions or any imperfections of the crystal which prevent the definition of a regular periodic lattice introduces phase changes in the scattering from the various atoms. If this happens with sufficient frequency there is an averaging of the phase-sensitive components of the intensity and the more striking dynamical effects may be lost (see Chapter 16).

Hence as a first step we explore the nature of the dynamical scattering effects by considering the ideal case of a perfect crystal in coherent, monochromatic radiation. The strength of the diffracted waves produced in a crystal depends on the scattering cross-sections or atomic scattering factor of the atoms and the direction of the incident radiation with respect to the crystal axes. The importance of the multiple scattering effects depends on these factors and also on the dimensions of the sample. As may be expected, the variety of effects is great and there are considerable differences for the different radiations we have been discussing. However there is a common body of theory which allows us to appreciate the main types of phenomena to be

observed and we present this in the simplest form which serves these ends.

An important difference between the X-ray and electron diffraction cases comes from the fact that for X-rays polarization effects can not be ignored. With the scattering through large angles the polarization factor for each scattering process may vary from zero to unity. Multiple scattering sequences which would be equivalent for the low-angle electron cases must be clearly differentiated. However, since adequate accounts of these complications are given in the literature they will be ignored here except that, occasionally, we will quote some results for X-rays to show how they differ from the results of the simpler scalar-wave diffraction theory.

A second important difference for the different radiations arises from the relative strengths of the interactions with matter. For X-rays and neutrons, by the time the amplitude of the scattered wave has built up to the magnitude for which multiple scattering is important and the kinematical approximation fails, the radiation has traversed sufficient thickness of crystal for sharply-defined Bragg reflections to be produced and the probability that more than one Bragg reflected beam is produced at a time is very small. Then it is possible to make use of the assumption valid for most cases that only two beams need to be considered at any one time, the incident beam and a beam diffracted from one set of lattice planes.

For electrons, on the other hand, strong scattering may occur for passage of the radiation through only the first few atom-thicknesses of crystal, i.e. for a slice of crystal thin enough to be regarded as a two-dimensional phase-grating which may give rise to several tens or hundreds of diffracted beams simultaneously. In order to take account of the multiple coherent interactions of all these diffracted beams, an n -beam dynamical theory must be used. The fact that for particular orientations there is destructive interference which weakens all but two of the beams, means that for electron diffraction also the "two-beam approximation" has some relevance. We proceed to consider this approximation with the full understanding that, at least for the electron case, it represents an assumption which must be justified later by a more complete n -beam treatment.

8.2. Theoretical approaches

The theoretical approaches which have been used for the formulation of the dynamical theory of diffraction by crystals may be divided into two general classes; those based on the formulation of wave-mechanics as a differential

equation, the wave equation in a crystal lattice, and those based on the integral equation formulation. Approaches by means of quantum field theory for electron diffraction have been made by Ohtsuki and Yanagawa [1965, 1966] and a modern approach to X-ray diffraction theory has been given, for example, by Kuriyama [1970, 1972], but these will not be discussed here.

The integral equation methods follow from the ideas mentioned in Chapter 1. They may be considered to represent mathematically the progression of radiation through the crystal. An incident plane wave is successively scattered in the crystal and the multiply-scattered components are added up according to their relative amplitudes and phases to form the out-going waves. The use of the Born series, equations (1.17), (1.22) may be interpreted as considering scattering by successive volume elements. The incident wave (zero-order term) is scattered by each volume element of the crystal to give the singly-scattered wave amplitude (first-order term) which is scattered again by each volume element to give the doubly-scattered wave and so on. This approach was used by Fujiwara [1959] for electron diffraction. Although the convergence of the Born series is notoriously bad, Fujiwara was able to obtain series-solutions for the scattering from a crystal which allowed important general conclusions to be drawn including the nature of the modifications to the scattering theory required when relativistic effects for high energy incident electrons are included (Fujiwara [1961]).

For the particular case of medium or high energy electrons (or other small-angle scattering) it is possible to take advantage of the fact that the propagation of the wave is close to the forward direction only and consider the scattering by successive planes of infinitesimal thickness. This was the approach of Cowley and Moodie [1957] using the concept of transmission through an infinite number of two-dimensional objects as suggested in Part (5) of Chapter 3.

A somewhat similar approach, based on different initial assumption was the first treatment of the scattering of X-rays by crystals given by Darwin [1914] and the related method used for the calculation of electron microscope intensities given by Howie and Whelan [1961]. In these treatments individual planes of atoms are considered to diffract incident plane waves to give a set of diffracted beams i.e., the conditions for Fraunhofer diffraction, rather than Fresnel diffraction are assumed to be relevant in the inter-atomic distances. In the original Darwin treatment it was assumed that an incident plane wave would be reflected from a plane of atoms to give only a single diffracted beam. The justification of this in terms of expediency and reasonableness is clear, but since we know that a two-dimensional grating gives rise to many diffracted beams, a more complete

justification in terms of n -beam diffraction theory would seem appropriate. A more comprehensive and updated account of the Darwin approach to X-ray diffraction has been given by Borie [1966] and Warren [1969] and the approach to electron diffracton and microscopy is described by Hirsch et al. [1965].

The differential equation approach was used in the initial formulation of X-ray diffraction theory by Ewald [1916, 1917] and von Laue [1931] and in the first formulation of electron diffraction theory by Bethe [1928]. Recent accounts of this approach have been given by Zachariasen [1945], James [1950], Authier [1970], Pinsker [1978], and Batterman and Cole [1964] for X-rays and by Hirsch et al. [1965], Peng [1995] and Wang [1995] for electrons. The Bethe theory has been expressed in matrix form (Sturkey [1957]; Niehrs [1959a]; Fujimoto [1959]) and from this comes the scattering matrix method (Niehrs [1959b]; Sturkey [1962]) which has much in common with the integral equation methods in that progression of an electron wave through successive slices of crystal can be represented by repeated application of a scattering matrix. A further approach which has something of this dual character is that of Tournarie [1960, 1961]. An illuminating discussion of the nature and interrelation of the various approaches has been given by Goodman and Moodie [1974].

For our present purposes it is convenient to begin our discussion of dynamical effects by reference to the Bethe theory. Although referring to the scalar electron waves it provides, by analogy, an indication of the X-ray diffraction effects when allowances for the polarization effects is made. Also it allows a straightforward means for deriving the results for the simple two-beam case which is adequate for describing most dynamical effects observed with X-rays and neutrons and forms a reasonable first approximation for many of the electron diffraction phenomena. Readers familiar with matrix theory may prefer to use the scattering matrix approach, from which the two-beam approximation may be derived very simply as shown by Rez [1977] and outlined in Subsection 10.2.5.

8.3. Bethe theory

8.3.1. The dispersion equations

The Schrödinger equation (1.5) may be written in the form

$$\{\nabla^2 + K^2_0 + \vartheta(r)\} \psi(r) = 0 \quad (1)$$

where we have put

$$K_0^2 = 2meE/h^2; \quad \vartheta(r) = 2me\phi(r)/h^2 = 2K_0\sigma\phi(r).$$

Then K_0 is the wave number, or the modulus of \mathbf{K}_0 , the incident wave in vacuum, and $K_0 = 1/\lambda$.

If we now Fourier transform (1), the transform of the second differential, $\nabla^2\psi$, gives a term in $k^2\Psi(k)$ where \mathbf{k} is the wave vector of the wave in the crystal, so that we obtain

$$(K_0^2 - k^2) \Psi(k) + v(k) * \Psi(k) = 0. \quad (2)$$

If the potential distribution term $\vartheta(r)$ is periodic, the Fourier transform is given by

$$v(k) = \sum_{\mathbf{g}} v_{\mathbf{g}} \delta(k - \mathbf{g}), \quad (3)$$

where \mathbf{g} is a reciprocal lattice vector.

According to Bloch's theorem the wave field in the crystal must have the periodicity of the lattice and $\Psi(k)$ must therefore be of the form

$$\Psi(k) = \sum_{\mathbf{h}} \Psi_{\mathbf{h}} \delta(k - k_{\mathbf{h}}),$$

where the wave vector $k_{\mathbf{h}}$, corresponding to the reciprocal lattice point \mathbf{h} , is defined as

$$k_{\mathbf{h}} = k_0 + \mathbf{h}, \quad (4)$$

and k_0 is the wave vector corresponding to the reciprocal lattice origin. Then inserting (3) into (2) we obtain the relationship between the wave numbers $k_{\mathbf{h}}$ and wave amplitudes $\Psi_{\mathbf{h}}$, or $\Psi(\mathbf{h})$, for the set of diffracted waves in the crystal,

$$(K_0^2 + v_0 - k^2_{\mathbf{h}}) \Psi_{\mathbf{h}} + \sum'_{\mathbf{g}} v_{\mathbf{h}-\mathbf{g}} \Psi_{\mathbf{g}} = 0, \quad (5)$$

where the prime on the summation indicates that we have omitted the term for $\mathbf{g} = \mathbf{h}$, namely $v_0 \Psi_{\mathbf{h}}$, and included v_0 in the coefficients of $\Psi_{\mathbf{h}}$ in the first term. This is the fundamental reciprocal space equation for the waves in a crystal, known as the "dispersion equation" because it relates the wave numbers or momentums of the waves to their energies.

We may further simplify the notation by putting $K_0^2 + v_0 = k^2$. Then \mathbf{k} is the wave vector for the incident wave \mathbf{K}_0 after it has passed from vacuum ($v_0 = 0$) into a medium of constant potential v_0 equal to the average potential in crystal lattice.

The mean refractive index for an electron in the crystal is given by

$$n^2 = \kappa^2/K^2_0 = 1 + \varphi_0/E,$$

$$\text{or } n = 1 + \varphi_0/2E. \quad (6)$$

If, subject to all our reservations and restrictions on the use of the concept of absorption for electrons in crystals, we introduce absorption in a phenomenological way by making $\varphi(r)$ complex, then the refractive index n and the wave number K_h become complex. Similarly all the coefficients v_h and wave vectors k_h are complex. We expand our definitions to include these complications.

The equation (5) may be written as a matrix equation, thus:

$$\begin{pmatrix} \kappa^2 - k_0^2 & & v_{0h} & & v_{0g} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot \\ v_{h0} & \cdots & \kappa^2 - k_h^2 & \cdots & v_{hg} \\ \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot \\ v_{g0} & \cdots & v_{gh} & \cdots & \kappa^2 - k_g^2 \\ & & & & \cdot \\ & & & & \cdot \end{pmatrix} \begin{pmatrix} \Psi_0 \\ \cdot \\ \cdot \\ \cdot \\ \Psi_h \\ \cdot \\ \cdot \\ \cdot \\ \Psi_g \end{pmatrix} = 0 \quad (7)$$

where, for convenience, we have written $v_{hg} = v_{h \cdot g}$. In the absence of absorption the matrix is Hermitian since for a real potential $v_{hg} = v_{gh}^*$. If the crystal has a center of symmetry the matrix will be real symmetric since $v_{hg} = v_{gh}$. If absorption is present, a quantity $i\mu_0$ must be added to the diagonal elements and $i\mu_{hg}$ to the off-diagonal elements.

8.3.2. Solutions of the equations

The object is then to solve the set of non-linear equations (5) or the matrix equation (7) to obtain the wave vectors k_h and the Fourier coefficients of the wave functions, Ψ_h , for the waves in the crystal, subject to the boundary conditions. Since there is no limitation to the number of reciprocal lattice points, there will be, in principle, an infinite number of solutions and so an infinite number of wave vectors k_h^i and amplitudes Ψ_h^i for each reciprocal lattice point. Alternatively we may say that for the solution number i , there is a set of wave vectors, k_h^i and a set of amplitudes Ψ_h^i , one for each reciprocal lattice vector.

This set is known as the "Bloch wave" number i since it represents one of the solutions for a wave in the crystal which, according to Bloch's theorem must be of the form

$$\Psi^i(r) = \sum_h \Psi_h^i \exp\{ik_h^i \cdot r\}, \quad (8)$$

where k_h is given by (4).

8.3.3. Boundary conditions

For a crystal limited by a planar surface, the boundary conditions imply the equality of the tangential components of the wave vectors on the two sides of the boundary. If a plane wave, K_0 is incident on the crystal the projection on the surface of κ , the incident wave in the crystal, must be equal to the projection of K_0 . This is just Snell's law of refraction. Hence we may draw the diagram, Fig. 8.1, which is a reciprocal space representation with the crystal boundary (real space) sketched in in order to define the surface normals used to apply the boundary conditions. The vector κ is drawn from a point L, the "Laue point", to the reciprocal lattice origin, O, such that its projection on the surface is the same as that of K_0 . The i Bloch wave solution of the dispersion equations will give a set of vectors $k^i_0 + h$ drawn to the reciprocal lattice points, h , from a point L^i , distance ξ^i (the "anpassung") from L along the surface normal drawn through L. The diagram is drawn for the case of electron waves for which κ , $k^i_0 > K_0$. For X-rays K_0 is larger than κ .

If an Ewald sphere is drawn, centered on the point L and of radius $|\kappa|$ it will miss the reciprocal lattice points by an amount ζ_h (the "excitation error") measured along the Ewald sphere radius. Since in general the anpassung ξ^i is very small compared with $|\kappa|$, the wave vectors k^i_h are very nearly parallel to the corresponding Ewald sphere radii and we may write

$$k^i_h = \kappa - \zeta_h - \xi^i \cos\theta_h \quad (9)$$

where θ_h is the angle with the surface normal.

From the equations (5) or (7) it is clear that the solutions will give the k^i_h as functions of both the incident magnitude κ and the Fourier coefficients of potential v_{hg} . Hence from (9) we see that the anpassung, ξ^i , depends on both the incident beam magnitude and directions and the v_{hg} but the dependence on v_{hg} decreases as the diagonal terms of (7) become larger. As the incident beam direction changes and the surface normal through L and L^i changes its position relative to O in (Fig. 8.1), the locus of the point L is a sphere centered on O of radius κ . Each point L^i traces out a surface, known as the " i branch" of the dispersion surface.

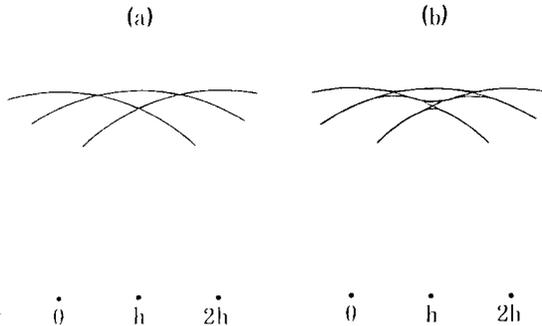


Fig. 8.2. The dispersion surface for waves in a crystal corresponding to the set of reciprocal lattice points $0, \mathbf{h}, 2\mathbf{h}\dots$ for (a) the kinematic case of very weak interaction and (b) for dynamical scattering with appreciable interactions.

For the general n -beam case the form of the dispersion surface is difficult to describe and in fact the solution of the equations (5) or (7) may not be made in general but only under simplifying assumptions chosen to give reasonable approximations to particular experimental conditions.

8.4. Two-beam approximation

8.4.1. Bloch waves and dispersion surfaces

Since for most important cases in X-ray and neutron diffraction and for selected cases in electron diffraction the maximum number of strong diffracted beams presented is two, it is a useful approximation to assume that only two wave amplitudes Ψ_0 and Ψ_h are not zero. It may be emphasized that this is not an approximation to a general solution in the normal sense. It is the solution of a different and simpler problem: the assumption of a universe in which only two waves can exist. Then the matrix equation (7) simplifies immediately to give

$$\begin{pmatrix} \kappa^2 - k_0^2 & v_{0h} \\ v_{h0} & \kappa^2 - k_h^2 \end{pmatrix} \begin{pmatrix} \Psi_0 \\ \Psi_h \end{pmatrix} = 0 \quad (10)$$

For a non-trivial solution the determinant of the matrix must be zero, giving, in general, four solutions for the wave vectors. However two of these solutions correspond to the back-scattering of electrons and may usually be ignored for high voltage transmission diffraction. An alternative derivation of the two beam approximation, starting from the scattering matrix formulation, is given in Subsection 10.2.5.

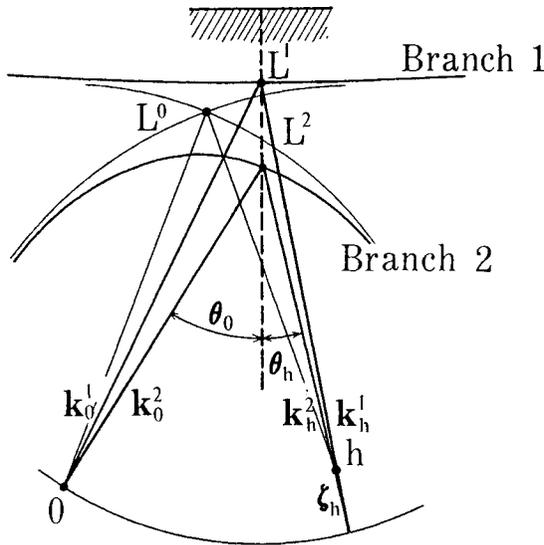


Fig. 8.3. The dispersion surface construction for the 2-beam case.

There are two forward scattering solutions with two Bloch waves $i = 1, 2$. For $i = 1$ there are wave amplitudes Ψ^1_0 and Ψ^1_h and wave vectors k^1_0, k^1_h and similarly for $i = 2$. The dispersion surface has two branches which approximate to spheres about the O and h reciprocal lattice points except near their line of intersection. This is all suggested in the diagram, Fig. 8.3, where we have numbered the branches of the dispersion surface in order of decreasing k^i_0 (Humphreys and Fisher [1971]). The dispersion surface section shown is symmetrical about the perpendicular bisector of the vector h . The spheres about O and h meet in the point L_0 which in three dimensions becomes a circle. Introducing the boundary conditions at an entrance surface defines the perpendicular through L which cuts the dispersion surface at the "tie-points" L^1 and L^2 .

If the h vector is parallel to the crystal surface and the angle of incidence is adjusted to make L coincide with L_0 then we have the simplest symmetrical case in which incident and diffracted beams make equal angles with the surface, the excitation error ζ_h is zero, the Bragg condition for the reflection is exactly satisfied and $|k^i_0| = |k^i_h|$.

Then the condition that the determinant of the matrix in (10) should be zero gives, for the case of no absorption,

$$\kappa^2 - k^2_h = (v_{0h} v_{h0})^{1/2} = |v_h|,$$

or, since κ and k_h differ by a relatively small amount

$$\kappa - k_h^i = \pm |v_h|/2\kappa. \quad (11)$$

Then the "anpassung" for this case becomes

$$\xi^i = \pm \frac{|v_h|}{2\kappa \cos\theta_h} \quad (12)$$

Thus the minimum separation of the two branches of the dispersion surface is proportional to $|v_h|$.

For this symmetrical diffraction condition, the two Bloch-wave solutions, corresponding to the two branches of the dispersion surface are equally excited. However, with increasing deviation from the Bragg condition it is seen that for one or other of the Bloch waves the wave vector k^i_0 becomes closer to κ , the incident wave vector without diffraction, while for the other k^i_0 deviates more and more from κ . It is to be expected that the value of k^i_0 involving less difference from κ will be favored especially as the strength of the diffraction effect decreases. Hence from Fig. 8.3, as the point L moves from left to right of the diagram, the greatest amplitude is in Bloch wave 1 at first and then switches over to Bloch wave 2 as the diffraction condition for the reflection h is passed.

In these considerations and in Fig. 8.3, we have taken into account only one of the two intersections of the surface normal with the dispersion surface. There is a further intersection diametrically opposite that shown. For many purposes this does not matter. The intersection shown is the one for which the incident beam in the crystal is approximately in the direction of the incident beam in vacuum and the diffracted beams are in the "forward" direction. However the other intersection can become important in the case of very long wavelength radiation or when the surface normal is rotated by about 90° with respect to our figure so that it becomes almost tangential to the dispersion surface as in the so-called "Bragg case" of diffraction from planes almost parallel to the surface.

8.4.2. Conduction electrons - energy representation

At this stage we may pause to note that the two-beam model we have used here is almost exactly the same as that used, with possibly less justification, for the consideration of the behavior of nearly-free conduction electrons in crystalline solids. In most text books on solid state physics the wave equation (1) is set up for an electron in a periodic lattice and the two-beam assumption is made immediately. The main difference from our treatment is that the aim is to establish the energy levels of the system rather than the

direction and amplitudes of diffracted beams. Hence the equation (10) is written in a form such as

$$\begin{pmatrix} \epsilon_{\kappa} - \epsilon_{k_0} & U_{-h} \\ U_h & \epsilon_{\kappa} - \epsilon_{k_0} \end{pmatrix} \begin{pmatrix} \Psi_0 \\ \Psi_h \end{pmatrix} = 0 \quad (13)$$

where ϵ_{κ} is the kinetic energy, equal to $\hbar^2 \kappa^2 / 2m$ and the solution gives the energies ϵ^i of the electrons. In particular it is found that when the Bragg condition is satisfied for a reflection h (when the electron wave vector extends to the Brillouin zone boundary, the plane bisecting the vector h at right angles) the electron energy has two values $\epsilon_0 \pm |U_h|$ where $U_h = cv_h$ volts and c is the appropriate constant. Then as the wave-vector of the conduction electron, K , passes through a Brillouin zone boundary, there is a discontinuity in its energy of amount $2|U_h|$. These results provide the basis for the "energy-band" picture widely used in solid-state physics and electronics.

8.4.3. X-ray diffraction; polarization

At this stage, also, we take note of the parallel development for X-ray diffraction, based on wave equations for the electric or magnetic field vectors derived from Maxwell's equations. From the equation equivalent to (10), the condition of zero value for the determinant is (Batterman and Cole [1964])

$$\begin{vmatrix} K_0^2(1 - \Gamma F_0) - k_0 \cdot k_0 & -k^2 P \Gamma F_h \\ -k^2 P \Gamma F_h & K_0^2(1 - \Gamma F_0) - k_h \cdot k_h \end{vmatrix} = 0 \quad (14)$$

where the value K_0 for the vacuum wave vector K_0 is modified inside the crystal by the factor $1 - \Gamma F_0$ where F_0 is the zero structure factor and $\Gamma = e^2 \lambda^2 / 4\pi \epsilon_0 m e^2 \cdot \pi V$. The polarization factor P is equal to unity for the σ polarization state (field vector E perpendicular to the plane of incidence) and equal to $\cos 2\theta$ for the π polarization state (E lying in the plane of incidence). Writing $k_0 \cdot k_0$ and $k_h \cdot k_h$ emphasizes that these quantities are complex.

8.5. The Laue (transmission) case

8.5.1. Electron diffraction for a thin crystal

We have so far considered only the system of waves set up when an incident beam enters through a planar surface into a semi-infinite periodic crystal field. Next we consider the

special cases which are suggestive of important real experimental conditions. For the relatively simple two-beam model, there are two situations for which a result can be obtained readily. The first is the "Laue case" of transmission (without any back scattering) through a perfect, parallel-sided crystal plate of infinite extent in two dimensions. The second, the "Bragg case" is for reflection from the planar surface of a semi-infinite crystal. By use of reasonable approximations it is possible to use the results for these idealized cases to discuss a wide variety of experimental situations.

The case of transmission through a thin parallel-sided plate with no back scattering, is treated by setting up the wave equation in the crystal in terms of the two Bloch waves and then applying the boundary conditions in appropriately simplified form at the two faces. Using equation (9), the equation (10) becomes

$$\begin{pmatrix} 2k\xi\cos\theta_0 & v_{0h} \\ v_{h0} & 2k\xi\cos\theta_h + 2k\zeta_h \end{pmatrix} \begin{pmatrix} \Psi_0 \\ \Psi_h \end{pmatrix} = 0, \quad (15)$$

where we have assumed all k vectors to be of approximately the same length and have retained only the essential differences between k values using (9). Solution of this equation gives the anpassung ξ as

$$\xi(2k\cos\theta_h) = -k\zeta_h \pm \left[k^2\zeta_h^2 + |v_{0h}|^2 \frac{\cos\theta_h}{\cos\theta_0} \right]^{1/2}, \quad (16)$$

and the reflection coefficient

$$C = \frac{\Psi_h}{\Psi_0} = \frac{\cos\theta_0}{v_{0h}\cos\theta_h} \left\{ -k\zeta_h \pm \left[k^2\zeta_h^2 + |v_{0h}|^2 \frac{\cos\theta_h}{\cos\theta_0} \right]^{1/2} \right\}. \quad (17)$$

The total wave in the crystal is given by the sum of the two Bloch waves

$$\psi(r) = \sum_{i=1,2} \alpha_i (\exp\{2\pi i k_i^0 \cdot r\} + C_i \exp\{2\pi i k_i^h \cdot r\}), \quad (18)$$

where C_1, C_2 are given by equation (17) with the + and - signs. Then since at the entrance surface there is unit amplitude in the k_0 direction and no amplitude in the diffracted direction, the boundary conditions imply that

$$\alpha_1 + \alpha_2 = 1$$

$$\alpha_1 C_1 + \alpha_2 C_2 = 0,$$

so that

$$\alpha_i = \frac{(-1)^i}{C_i} \left(\frac{C_1 C_2}{C_1 - C_2} \right). \quad (19)$$

At the exit surface, since there is assumed to be no wave reflected back into the crystal from any incident wave, each crystal wave is considered to pass directly into the vacuum unchanged except that the crystal wave vectors k_h^i become the vacuum wave vectors K_h . This adds no new boundary conditions. Combining (16) with the wave amplitudes found by solution of (10) then gives the crystal wave amplitudes after passage through a thickness H of crystal and, adding the $i = 1$ and 2 contributions to the amplitudes for the 0 and h beams in vacuum then gives the desired result. This result is conveniently expressed in terms of two new parameters, the deviation parameter, w , giving the deviation from the exact Bragg reflection condition, and the "extinction distance" ξ_h (not to be confused with the anpassung ξ^i) which is inversely proportional to the structure amplitude $|v_h|$.

These quantities are defined by reference to Fig. 8.4, which represents the portion of the dispersion surface close to L_0 . Since this region is normally very small in comparison to the radius of the two spheres which intersect there, the sections of the two spheres may be considered as straight lines and the section of the dispersion surface of interest is a hyperbola with these straight lines as asymptotes. The distance between branches of the dispersion surface in the direction of the surface normal is $2D$ given by

$$D = \pm(q^2 + t^2)^{1/2}, \quad (20)$$

where q by generalization of (12) is given by

$$q = \frac{(v_h v_{-h})^{1/2}}{2\kappa(\cos\theta_0 \cos\theta_h)^{1/2}} \quad (21)$$

and t is a measure of the deviation, $\Delta\theta$ from the Bragg angle, θ_B through the relationship

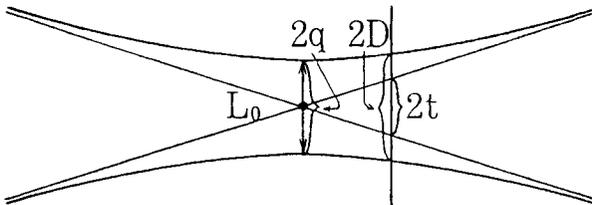


Fig. 8.4. Definition of the quantities q , D and t for two-beam diffraction.

$$\Delta\theta = \frac{2t\cos\theta_h}{\kappa\sin 2\theta_B}. \quad (22)$$

Then the parameters w and ξ_h are defined by

$$w = t/q; \quad \xi_h = 1/2q. \quad (23)$$

8.5.2. Small angle approximation

If we may make the small-angle approximations common in transmission electron diffraction and put $\cos\theta_0 \approx \cos\theta_h \approx 1$ and $\sin 2\theta_B \approx 2\sin\theta_B$, we obtain

$$\xi_h = \frac{\kappa}{(v_h v_{-h})^{1/2}} = \frac{1}{2\sigma\Phi_h} \quad (24)$$

$$w = \frac{\Delta\theta \cdot \kappa}{d_h (v_h v_{-h})^{1/2}} = \frac{\Delta\theta \cdot \xi_h}{d_h}, \quad (25)$$

or, since $d_h = |h|^{-1}$ and $\zeta_h = \Delta\theta \cdot |h|$,

$$w = \xi_h \cdot \zeta_h. \quad (26)$$

Then the constants obtained from the boundary conditions are

$$\begin{aligned} \alpha_1 &= \frac{1}{2} [1+w/(1+w^2)^{1/2}], \\ \alpha_2 &= \frac{1}{2} [1-w/(1+w^2)^{1/2}], \\ C_1 &= w - (1+w^2)^{1/2}, \\ C_2 &= w + (1+w^2)^{1/2}, \end{aligned} \quad (27)$$

and the intensities of the transmitted and diffracted beams are

$$\begin{aligned} I_0 &= |\alpha_1 \exp\{ik^1_0 \cdot r\} + \alpha_2 \exp\{ik^2_0 \cdot r\}|^2 \\ &= (1+w^2)^{-1} \left[w^2 + \cos^2 \left\{ \frac{\pi H(1+w^2)^{1/2}}{\xi_h} \right\} \right], \end{aligned} \quad (28)$$

$$\begin{aligned} I_h &= |\alpha_1 C_1 \exp\{ik^1_h \cdot r\} + \alpha_2 C_2 \exp\{ik^2_h \cdot r\}|^2 \\ &= (1+w^2)^{-1} \sin^2 \left\{ \frac{\pi H(1+w^2)^{1/2}}{\xi_h} \right\}, \end{aligned} \quad (29)$$

where in each case we have added the two components of the crystal wave which combine to form the vacuum wave. Thus the intensities of both the transmitted and diffracted beams oscillate with crystal thickness H . At the exact Bragg angle the periodicity is the extinction length, ξ_h .

8.5.3. Laue case with absorption

Here we have not excluded the absorption which would have the effect of making complex any quantity which involves ν_h . It is probably better to introduce the absorption coefficients μ_0 , μ_h explicitly and write these results in terms of real parameters w and ξ_h in which case (28) and (29) become

$$I_0 = \frac{1}{2} \exp\{-\mu_0 H\} \left[\left(1 + \frac{w^2}{1+w^2}\right) \cosh\left(\frac{\mu_h H}{(1+w^2)^{1/2}}\right) - \frac{2w}{(1+w^2)^{1/2}} \sinh\frac{\mu_h H}{(1+w^2)^{1/2}} + \frac{1}{(1+w^2)} \cos\left\{\frac{2\pi H(1+w^2)^{1/2}}{\xi_h}\right\} \right] \quad (30)$$

and

$$I_h = \frac{\exp\{-\mu_0 H\}}{2(1+w^2)} \left[\cosh\left(\frac{\mu_h H}{(1+w^2)^{1/2}}\right) - \cos\left\{\frac{2\pi H(1+w^2)^{1/2}}{\xi_h}\right\} \right]. \quad (31)$$

Thus the intensities of both beams are decreased by the average absorption effect. Apart from this the effect of absorption is to prevent a completely sinusoidal oscillation of intensities even for $w = 0$. There is a background, non-oscillatory term which becomes increasingly important for large crystal thicknesses. The significance of this term will be discussed in the following chapter.

8.6. Bethe potentials

In this original paper, Bethe [1928] took into account the fact that, especially in the electron diffraction case, the condition that only two beams should exist in the crystal is never fully satisfied. Some weak beams are always present, corresponding to reciprocal lattice points for which the excitation error is reasonably large but not large enough to extinguish the contribution entirely. For a particular class of reflections, the "systematic" set, the excitation errors are constant and the "systematic interactions" of the corresponding weak waves with the strong diffracted waves are always the same for any direction of the incident beam satisfying the Bragg angle for the reflection h . This systematic set consists

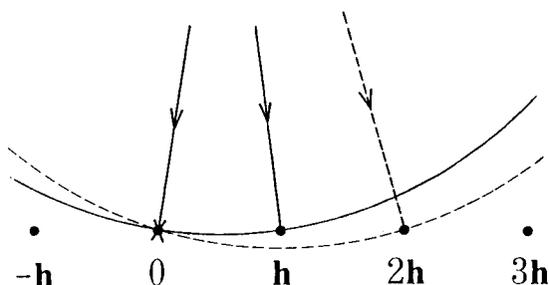


Fig. 8.5. The excitation of a systematic set of reflections. When the Bragg condition for one reflection is satisfied, the other reflections of the systematic set always have the same excitation errors.

of the set having integral or sub-integral multiples of the lattice vector h as suggested by Fig. 8.5.

The Bragg condition is satisfied for h for any orientation of the incident beam for which the Ewald sphere passes through h i.e., for any rotation of the plane of the figure about the $0-h$ line. For any such orientation, the excitation errors for a reflection nh will always be the same.

As the Ewald sphere is rotated about the $0-h$ line, it will pass through other "non-systematic" reflections at various irregular intervals. When the Ewald sphere is close to any non-systematic reflection then a definite n -beam situation will exist which must be handled accordingly. But if it is not too close to the reciprocal lattice point a weak beam will be produced which may be handled in much the same way as for the weak "systematic" set.

The amplitude of a weak beam g may be expressed approximately from (5) in terms of the amplitudes of the two strong beams only:

$$(K^2_0 + v_0 - k^2_g) \Psi_g + v_g \Psi_0 + v_{g-h} \Psi_h = 0$$

or

$$\Psi_g = -\frac{(v_g \Psi_0 + v_{g-h} \Psi_h)}{\kappa^2 - k_g^2}, \quad (32)$$

where we imply that the denominator must be large.

Then in reducing the set of equations (5) or (7) to give the two-beam case we include the contributions (32) to the summation in (5) and instead of (10) we obtain

$$\begin{pmatrix} \kappa^2_0 - k^2_0 & U_{0h} \\ U_{h0} & \kappa^2_h - k^2_h \end{pmatrix} \begin{pmatrix} \Psi_0 \\ \Psi_h \end{pmatrix} = 0 \quad (33)$$

where

$$\kappa_0^2 = K_0^2 + U_{00},$$

$$\kappa_h^2 = K_0^2 + U_{hh},$$

and

$$\begin{aligned} U_{00} &= v_0 - \sum_g'' \frac{v_g v_{-g}}{\kappa^2 - k_g^2}, \\ U_{hh} &= v_0 - \sum_g'' \frac{v_{g-h} v_{h-g}}{\kappa^2 - k_g^2}, \\ U_{0h} &= v_h - \sum_g'' \frac{v_g v_{h-g}}{\kappa^2 - k_g^2}, \\ U_{h0} &= v_{-h} - \sum_g'' \frac{v_{-g} v_{g-h}}{\kappa^2 - k_g^2}, \end{aligned} \quad (34)$$

and the double primes on the summations indicate that the values $g = 0, h$ are excluded.

Thus the effect of the weak beams is taken into account, to a first approximation by modifying the values for the potential Fourier coefficients v_0 and v_h in the two-beam solution. The modified potentials are often referred to as "Bethe potentials".

It is interesting to note that while, as we shall see later, the Bethe potentials give a good account of n -beam dynamical effects for some particular cases of experimental interest, the higher-order approximations generated by a repeated application of the process of (32) give no further improvement but in some cases, at least, give a much worse agreement with complete n -beam dynamical treatments.

It was pointed out by Miyake [1959] that in the limiting case of zero thickness of a crystal the use of the Bethe potentials gives the wrong result since the diffracted intensity given by (29) should be proportional to v^2_h as for kinematical scattering and not $|U_{0h}|^2$. This limiting case has been discussed in detail by Gjønnes [1962] who showed how a consistent treatment of weak beams may be derived.

8.7. The Bragg case

When an incident beam is diffracted from planes parallel or nearly parallel to a flat surface of a crystal so large that it may be considered semi-infinite, the diffracted beams

observed are those which reenter the vacuum on the same side of the crystal as the incident beam. Again we can obtain a relatively simple solution if we assume a two-beam case with only one strong reflected beam, but the results are quite different in form from the transmission case.

To treat this case we must return to the equation (15) and solve it for the given boundary conditions without making the simplifying approximations leading to equations (24) to (27), since the angles θ_0 and θ_h of the beams with the surface normal cannot be assumed small and equal. For high energy electrons these angles approach $\pi/2$ and the cosines usually are of opposite sign.

For reflection from a crystal surface, $\cos\theta_h$ is of opposite sign to $\cos\theta_0$ so that the square root term in (16) and (17) is imaginary for a range of ζ_h values between the limits

$$\zeta_h = \pm \frac{v_{nh}}{k} \left| \frac{\cos\theta_h}{\cos\theta_0} \right|^{1/2}. \tag{35}$$

Within this range the wave vectors in the crystal have imaginary values, corresponding to strongly damped, non-propagating waves. In the absence of absorption there is total reflection within this range with the reflectivity falling off rapidly on either side as suggested in Fig. 8.6. In the presence of absorption the reflectivity is less than 100 percent and is asymmetric, as indicated.

In terms of the dispersion surface representation, the crystal surface in the reflection case is almost perpendicular to the \mathbf{h} vector. The surface normal through the

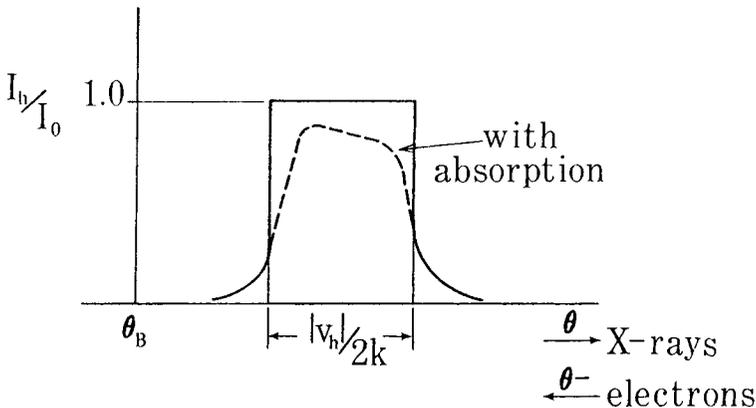


Fig. 8.6. The intensity of reflection as a function of the deviation from the Bragg angle for no absorption and with absorption (dotted line) for the Bragg case of reflection from a large perfect crystal.

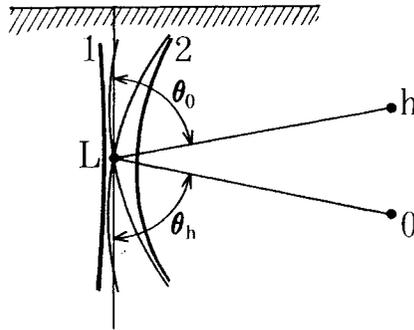


Fig. 8.7. The dispersion surface construction for the two beam Bragg case, when the surface normal may pass through the gap between the two branches.

Laue point L may cut the Branch 1 of the dispersion surface in two points, or it may cut Branch 2 in two points, or it may pass through the gap between the branches, thus giving imaginary components of the wave vectors corresponding to the exponentially damped waves in the crystal (Fig. 8.7). The angular range for total reflection is then given, for the symmetrical case $|\cos\theta_0| = |\cos\theta_h|$, by the width of the gap (35).

It is straightforward to include the Bethe potentials in the two-beam treatment as for the transmission electron diffraction case (see Pinsker [1964]). However it may well be that, as in the case of transmission through very thin crystals, the validity of the Bethe potentials is questionable when only a few layers of atoms are involved in the diffraction process, as is often the case for reflection electron diffraction.

For diffraction of high voltage (20 to 100 kV) electrons from surfaces the scattering angles are of the order of a few degrees. Consequently the angle made by the incident beam with the surface ($\pi/2 - \theta_0$) is of the order of one degree. The penetration of the beam into the crystal is severely limited by absorption or by the diffraction when a strong reflection is excited and may be only a few Å. Another important factor to be considered in this case is that, although the refractive index of the crystal may be only slightly greater than unity for electrons, the refraction effects are considerable for such small angles of incidence. Waves diffracted from planes parallel to the surface having spacings as small as 2 or 3 Å may suffer total internal reflection and be unable to leave the crystal. Diffracted waves reaching the crystal at slightly larger angles are refracted so that they are displaced by large distances in the diffraction patterns.

Because the diffraction angles are so small and the refraction effects relatively large, the reflection method of diffraction for high-energy electrons, RHEED, is extremely

sensitive to slight deviations from exact planarity of the surface as well as to the composition of the topmost layers of atoms of the crystal, (Menadue [1972]; Colella and Menadue [1972]) so that extreme care must be exercised in obtaining a flat clean surface if any attempt is to be made to relate the theoretical predictions to experimental observations.

A further complication is that even elementary considerations show that a simple two-beam approximation cannot be adequate for any practical experimental situation and hence a complete n -beam dynamical treatment is necessary (Colella [1972]). Similarly the use of low energy electrons (10 to 500 eV) with near-normal incidence on the surface (LEED) offers an essentially n -beam diffraction situation, with additional complications, as will be discussed in Chapter 19. An excellent summary of the theory used for LEED is given in the book by Pendry [1974].

As in other cases, however, the two beam approximation is adequate for most considerations of Bragg-case reflection of X-rays and neutrons from surfaces of large perfect crystals. The original result for a non-absorbing crystal predicting the region of total reflection, was obtained by Darwin [1914]. More complete discussions have been given by James [1948] and Batterman and Cole [1964]. The agreement between the two-beam theory and experimental measurements is excellent. For neutrons, see Bacon [1975].

This Page Intentionally Left Blank

Dynamical diffraction effects

9.1. Thickness fringes and rocking curves - electron diffraction

9.1.1. Intensity formulas

We have seen in Chapter 8 that when the usual small-angle approximations of high energy electron diffraction are applied in the two-beam approximation the intensities of the transmitted and diffracted beams from a flat parallel sided crystal plate may be written in simple form, especially for the case of no absorption and centro-symmetric crystals:

$$I_0 = (1+w^2)^{-1} \left[w^2 + \cos^2 \left\{ \frac{\pi H(1+w^2)^{1/2}}{\xi_h} \right\} \right], \quad (1)$$

$$I_h = 1 - I_0 = (1+w^2)^{-1} \left[\sin^2 \left\{ \frac{\pi H(1+w^2)^{1/2}}{\xi_h} \right\} \right], \quad (2)$$

where, under these special conditions, $\xi_h = \kappa/|v_h| = (2\sigma\Phi_h)^{-1}$ and $w = \xi_h \cdot \zeta_h = \Delta\theta \cdot \xi_h/d_h$, which is a measure of the deviation from the Bragg angle. If the incident beam is exactly at the Bragg angle, $w = 0$ and

$$I_h = 1 - I_0 = \sin^2 \{ \pi H |v_h| / \kappa \}. \quad (3)$$

This is the well-known pendellosung or pendulum solution of Ewald. The energy is passed back and forward between the incident and diffracted beams in much the same way as energy is exchanged between coupled pendulums. The coupling is provided in this case by the scattering from one beam to the other with a strength proportional to $|v_h|$.

9.1.2. Real space picture

An alternative picture of the diffraction process is provided by considering the wave-fields in real space in the crystal lattice. An incident plane wave ψ_0 and a diffracted wave ψ_h proceeding in directions inclined at $+\theta_B$ and $-\theta_B$ to the reflecting planes interfere to form a standing wave field having the periodicity of the lattice planes. But we have

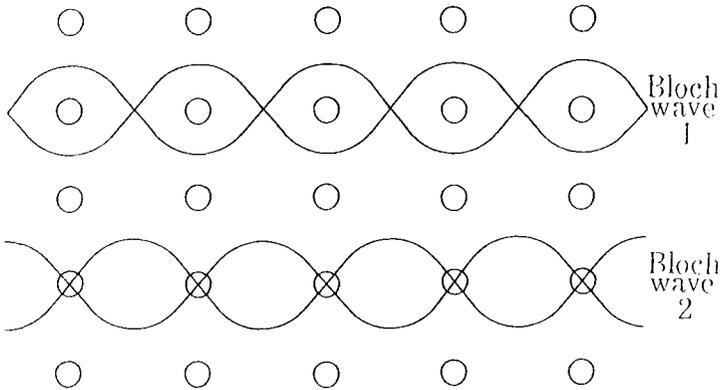


Fig. 9.1. Illustrating the relation of the two Bloch waves to the lattice planes for a two-beam approximation.

seen that under these conditions there are two solutions to the wave equation and two wave-fields or Bloch waves. These two Bloch waves correspond to the two standing wave solutions which have the full symmetry of the diffraction conditions in the lattice, namely one having nodes half-way between atomic planes and one with nodes on the atomic planes as suggested in Fig. 9.1. At the Bragg angle these two solutions are equally probable and have equal amplitudes. However, since the waves are in a potential field modulated by the lattice wave of amplitude φ_h , they experience not the average potential φ_0 but the modified potential $\varphi_0 \pm \varphi_h$ where the + sign refers to Bloch wave 1.

Thus the values of the refractive index and the wave number of the two Bloch waves differ (c.f. Fig. 8.3). After these two Bloch waves have passed through a thickness H of crystal they differ in phase by an amount proportional to $H(\varphi_h/E_0)$. Hence, when the two contributions to the diffracted wave in the crystal are added to form the diffracted wave in vacuum they may either reinforce or cancel each other, depending on the thickness. As the crystal thickness is increased the diffracted intensity then varies sinusoidally with a periodicity $H = \kappa/v_h = \xi_h$. The $\pi/2$ phase difference between transmitted and diffracted waves in each case ensures that when there is reinforcement for the diffracted beam there is cancellation for the transmitted beam, and the sum of the intensities of the two beams is constant.

Thus the pendulum solution may be looked upon as the result of double-refraction in the crystal, somewhat similar to, although different in principle from, the optical case of transmission of a polarized wave through an anisotropic crystal.

9.1.3. Rocking curves

The effect of a deviation from the Bragg angle ($w \neq 0$) is to reduce the average diffracted beam intensity by a factor $(1+w^2)^{-1}$ and to reduce the periodicity of the oscillations with thickness by multiplying ξ_h by $(1+w^2)^{-1/2}$.

For a fixed thickness the variation of diffracted intensity with deviation from the Bragg angle is best seen by putting $w = \xi_h \zeta_h$ as in (8.26) and writing (2) as

$$I_h = \left(\frac{\xi_h^{-2}}{\xi_h^{-2} + \zeta_h^2} \right) \sin^2 \{ \pi H (\xi_h^{-2} + \zeta_h^2)^{1/2} \}. \quad (4)$$

Then if ζ_h^2 is much greater than $\xi_h^{-2} = |v_h|^2/\kappa^2$, this is approximately equal to

$$I'_h = \frac{|v_h|^2}{\kappa^2} \pi^2 H^2 \frac{\sin^2(\pi H \zeta_h)}{(\pi H \zeta_h)^2} \quad (5)$$

which is exactly the expression for the kinematical diffraction from a parallel-sided plate of thickness H as a function of the excitation error ζ_h .

Thus the two-beam solution tends to the kinematical intensity formula if the strength of the interaction with the lattice, represented by $|v_h|$ goes to zero or if, for finite v_h , the excitation error is sufficiently large.

If the angle of incidence of the electron beam is varied and the diffracted intensity is recorded to give the "rocking curve" for the reflection, the deviation from the simple $(\sin^2 x)/x^2$ curve of the kinematical theory is most apparent near the exact Bragg condition. For $w = 0$, or $\zeta_h = 0$, the intensity oscillates sinusoidally as given by (3) as the thickness increases instead of increasing continuously with H^2 as given by (5).

9.1.4. Extinction contours

It is not convenient in practice to record a rocking curve by rotating a thin perfect crystal in an electron beam. Instead either of two methods may be used. Firstly the bright or dark-field images of a uniformly bent parallel-sided crystal plate may be observed in an electron microscope. For a plane parallel illuminating beam the angle of incidence of the beam with the lattice planes then varies with distance across the crystal as suggested by Fig. 9.2. The intensity of the transmitted and diffracted beams at any point of the exit

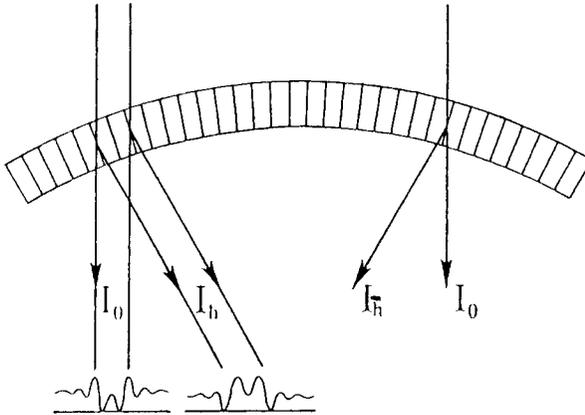


Fig. 9.2. The bright-field and dark-field images of a bent crystal show the variation of diffracted intensity as a function of angle of incidence.

surface depend on the excitation error for the lattice planes at that point and for uniform curvature of the crystal the excitation error varies almost linearly across the crystal. Then if the diffracted beams are imaged, the intensity in the resulting image shows the variation given by (4) as a function of distance across the crystal which takes the place of ζ_h . Imaging of the direct transmitted beams, I_0 , gives the bright-field image which in the present case with absorption neglected will show the intensity variation $1 - I_h(\zeta_h)$.

Thus the image of a bent crystal is traversed by "extinction contours" which are dark in bright field and bright in dark field and have the form of strong and subsidiary weak fringes marking the parts of the crystal which are in the correct orientation to give a diffracted beam.

An image of a crystal, bent in two dimensions, is shown in Fig. 9.3. The extinction contours tend to occur in sets of parallel lines corresponding to the sets of positive and negative higher orders of a strong reflection, i.e. the "systematic" sets such as $\dots -3h, -2h, -h, \dots h, 2h, 3h \dots$. The fact that the sets of parallel lines do not in general show the form of the simple contours suggested by the two beam theory and illustrated in Fig. 9.2 is partly a consequence of the omission of the effect of absorption from the two beam theory given above, and is due, in part, to the occurrence of n -beam dynamical diffraction effects which are specifically excluded from the two-beam treatment. Further evidence of n -beam dynamical scattering effects, in this case involving "non-systematic" interactions, is given at points where non-parallel extinction contours intersect. Here the intensities of the contours are by no means additive but may fluctuate wildly.

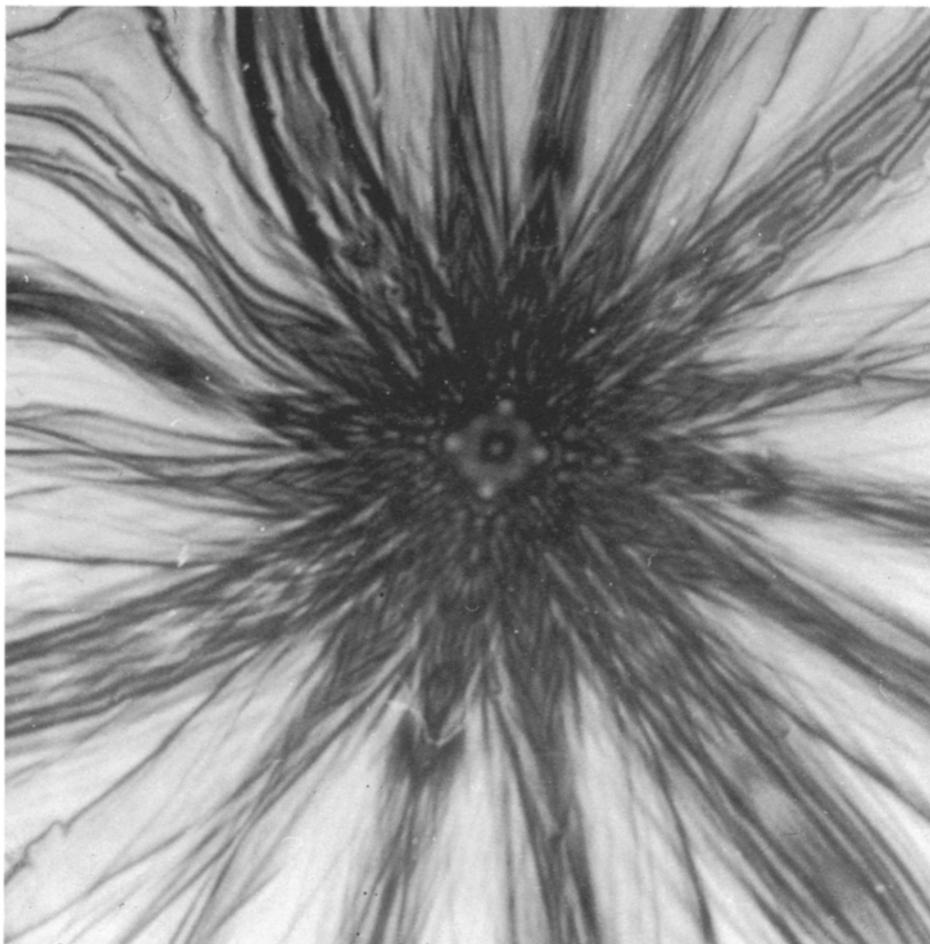


Fig. 9.3. Electron micrograph of a thin crystal of gold, showing extinction contours related to the bending of the crystal in two dimensions. The center of the "star" corresponds to the [100] direction. (After Berry et al. [1973]).

9.1.5. Convergent beam diffraction

A second means for observing the rocking curves for diffraction from a thin parallel sided crystal is by use of the convergent beam diffraction technique originally introduced by Kossel and Mollenstedt [1939]. Here an electron beam of finite aperture is focussed on the specimen. If the beam is defined by a circular aperture, each spot of the diffraction pattern is spread into a circular disc as suggested by Fig. 9.4(a). For each angle of incidence

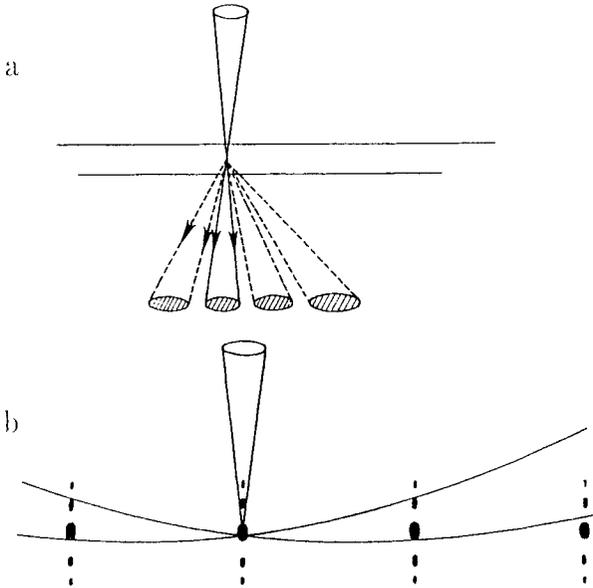


Fig. 9.4. The formation of a convergent-beam diffraction pattern illustrated (a) in the real-space configuration and (b) in reciprocal space where the Ewald sphere for each incident direction produces its own section of the scattering power distribution.

represented in the convergent beam the appropriate Ewald sphere may be drawn and the diffraction intensities found, as suggested in the (kinematical) reciprocal space diagram, Fig. 9.4(b). The variation of intensity across the central spot and diffracted spot discs in the diffraction pattern then gives the variation of intensities with angle of incidence. Fig. 9.5 gives a typical convergent beam diffraction pattern from a thin perfect crystal of MgO.

For a two-beam case, the incident and diffracted beam discs are crossed by parallel fringes of intensity given by equation (4) for I_h and $I_0 = 1 - I_h$. On this basis MacGillavry [1940] interpreted the convergent beam patterns obtained by Kossel and Mollenstedt from thin mica crystals and from the form of the intensity variations was able to deduce values of $|U_h|$ for the various reflections by use of equation (4) or the equivalent. The values she derived were, fortuitously, in excellent agreement with the values calculated from the structure of mica as then known.

The convergent beam diffraction technique has more recently been refined by Goodman and Lehmpfuhl [1964] and Cockayne et al. [1967]. As will be recounted later (Chapter 14), the technique has been widely developed as a means for

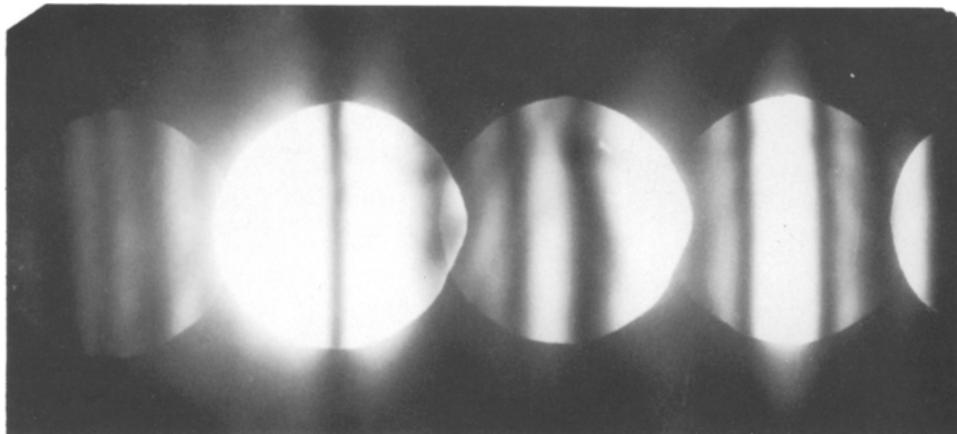


Fig. 9.5. Convergent beam diffraction pattern for a thin perfect crystal of MgO showing the variation of intensity with angle of incidence for the $2h, 0, 0$ reflections. Some perturbation of the fringes by non-systematic interactions is visible.

making accurate measurements of intensities, demonstrating that the two-beam results form only a first approximation to the n -beam diffraction situation and making use of the n -beam dynamical diffraction effects to derive highly accurate values for the Fourier coefficients ν_h or Φ_h .

9.1.6. Diffraction and imaging of crystal wedges

Diffraction from a wedge-shaped crystal can be treated, with reasonable approximations, by an extension of the treatment of transmission through a parallel-sided crystal. This has been done in detail by Kato [1952]. The boundary conditions for an exit face not parallel to the entrance face are imposed to give the wave entering the vacuum. From elementary principles we can say that since the crystal has two values for the effective refractive index corresponding to the potential coefficients $\nu_0 \pm \nu_h$, both incident and diffracted waves are refracted by the "prism" of crystal to give two waves coming out in slightly different directions. Then each spot in the diffraction pattern is split into a pair of two close spots. At the exit face of the crystal the two wave fields interfere to give a sinusoidal variation of intensity with thickness when viewed with either the transmitted or diffracted beams. Thus electron microscope images in either bright or dark field show a pattern of sinusoidal fringes across the image of the wedge, as in Fig. 9.6.

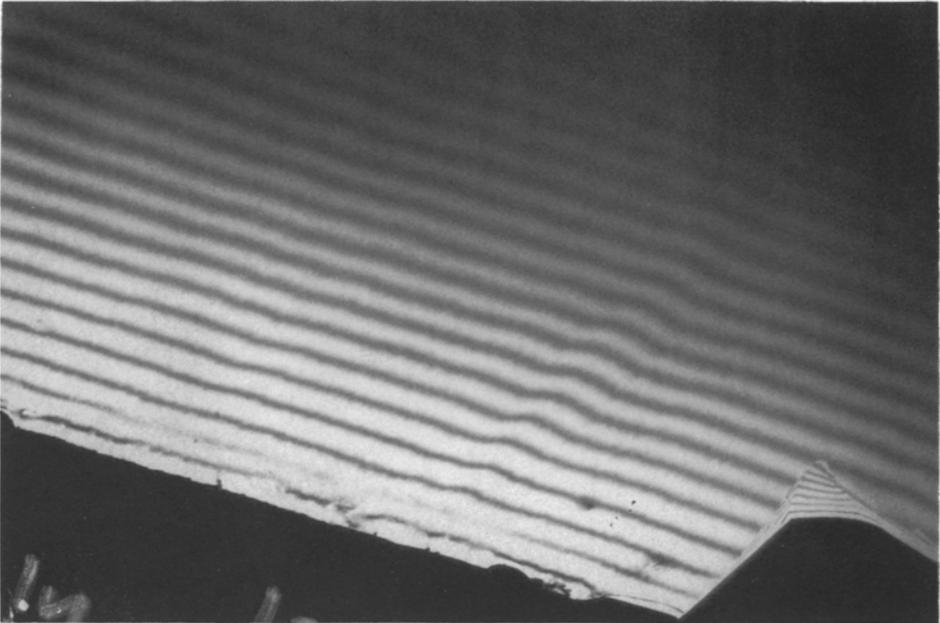


Fig. 9.6. Near-sinusoidal thickness fringes in the dark-field image of a silicon crystal wedge. The image is obtained using the $[111]$ reflection when the Bragg condition is satisfied for the $\{222\}$.

An alternative explanation of this effect which requires further justification but gives essentially the correct result, is based on the assumption that for each thickness of the wedge the intensities of the incident and diffracted beams are exactly the same as for a parallel-sided crystal of the same thickness. Then the variation of intensity across the bright or dark-field image of the wedge is given by the equation (1) or (2) and the distance in the image is related to the thickness H through the wedge angle and the orientation of the crystal with respect to the beam.

These thickness fringes were first observed in the images of MgO smoke crystals which form almost perfect cubes and so present up to six wedge-shaped regions to the incident beam (Heidenreich [1942]), (Kinder [1943]). The corresponding splitting of the diffraction spots due to the double-refraction effect was observed in diffraction patterns of MgO smoke by Cowley and Rees [1946, 1947] and Honjo [1947]. The six wedge-shaped regions of an MgO cube give rise to a star-shaped group of six radial pairs of spots surrounding the flat-crystal diffraction spot position. In recent years a more detailed study of the intensity distributions in images

of wedge crystals has shown that strong deviations from the simple sinusoidal intensity variations of the two-beam frequently occur. Accurate measurements of the intensity distributions have shown excellent agreement with n -beam dynamical calculations and have been used as a means for deriving values for the Fourier coefficients of potential v_h or Φ_h (for a review see Spence [1992]).

Similarly the more thorough study of the refraction fine structure of diffraction spots by Moliere and Wagenfeld [1958] and Lehmppuhl and Reissland [1968] have shown that a multiplicity of spots, rather than only two may be produced by a wedge. The latter authors recorded the splitting as a function of crystal orientation, or ζ_h , and so could map out the lengths of the various wave vectors in the crystal and then plot directly the form of the dispersion surface, showing a number of branches.

9.1.7. Absorption effects for wedges

With the introduction of absorption into the two beam solution the intensities of the incident and diffracted beams transmitted through a thin crystal are modified as indicated in equations (8.30) and (8.31). There is an overall loss of intensity of both beams due to the mean absorption coefficient μ_0 . Then a non-oscillating term is added which gives a background to the sinusoidal oscillations. For simplicity we consider the special case that the Bragg condition is exactly satisfied, i.e. $w = 0$. Then

$$I_0 = \frac{1}{2} \exp\{-\mu_0 H\} \left[\cosh \mu_h H + \cos \left(\frac{2\pi H}{\xi_h} \right) \right], \quad (6)$$

$$I_h = \frac{1}{2} \exp\{-\mu_0 H\} \left[\cosh \mu_h H - \cos \left(\frac{2\pi H}{\xi_h} \right) \right]. \quad (7)$$

These intensities are sketched as a function of thickness in Fig. 9.7 which thus represents the form to be expected for the thickness fringes in the images of a wedge crystal and is seen to give something like the intensity distribution of the fringes in Fig. 9.6. The first term in each of (6) and (7) gives the nonoscillating "center curve" of the profile as

$$\frac{1}{4} [\exp\{-(\mu_0 + \mu_h)H\} + \exp\{-(\mu_0 - \mu_h)H\}].$$

Of these two terms the first dies out more rapidly if μ_h has the same sign as μ_0 . Then for H large the center curve is $\exp\{-(\mu_0 - \mu_h)H\}$. This decreases more slowly with H than the

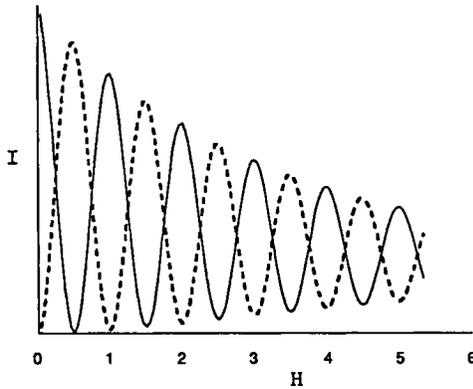


Fig. 9.7. Diagram suggesting the form of the intensity variation of incident and diffracted beams for a two-beam case with absorption and a wedge crystal of thickness H .

amplitude of the oscillating part which is proportional to $\exp\{-\mu_0 H\}$.

In principle then it is possible to deduce the values of μ_0 and μ_h from the form of the thickness fringes. Attempts to do this by Uyeda et al. [1965] gave some results but showed deviations from the two-beam formulas.

In terms of the picture of Bloch waves transmitted through the crystal, Fig. 9.1, it can be seen that, if the absorption process takes place when electrons pass close to the atoms, the two Bloch waves will be absorbed differently. Since the probability of finding an electron at a certain position is proportional to the square of the wave function modulus, the electrons of Bloch wave 1 are more likely to be in the vicinity of atoms and so are more strongly absorbed, with an absorption coefficient $\mu_0 + \mu_h$, while those of Bloch wave 2 spend most of their time between the planes of atoms and so are less absorbed, with an absorption coefficient $\mu_0 - \mu_h$. Then when they leave the crystal the contributions from the two Bloch waves to the diffracted beam do not have equal amplitude and so are not able to produce interference fringes of maximum contrast.

We saw before that the two Bloch waves have different refractive indices in the crystal and so give vacuum waves in slightly different directions from a wedge shaped crystal. With absorption, the diffraction spot is split into two components of unequal intensity. The wave refracted most, from Bloch wave 1, is also attenuated most. This was first observed and analysed by Honjo and Mihama [1954].

Absorption also has the effect of making the rocking curve for a crystal in the transmitted beam, i.e. the extinction contour for the bright field image of a bent crystal, no longer symmetric. This is readily seen from the presence of the asymmetric sinh term in equation (8.27). It follows also from the dispersion curve picture of Fig. 8.3. There we stated that at the Bragg angle both branches of the dispersion surface are equally excited, but for an incident beam such that L is to the left of L_0 , branch 1 predominates and on the other side of the Bragg angle branch 2 predominates. Since absorption makes the two branches non-equivalent, the intensities are different on the two sides of the Bragg angle.

The asymmetry is reversed when we go from the h to the \bar{h} reflection. Hence we see in the pattern of extinction contours, Fig. 9.3, that the region between two strong parallel, h and \bar{h} , fringes tends to be darker than the region outside of them.

9.2. Dynamical effects of X-ray and neutron diffraction

9.2.1. Techniques for X-ray diffraction

Apart from the polarization factor which appears in (8.14) and the inapplicability of the small angle approximation, it would appear that the dynamical effects for X-ray and neutron diffraction should be exactly the same as for electron diffraction and, in fact, thickness fringes have been observed for both X-rays and neutrons. However differences in the properties of the radiations and the conditions for the experimentally feasible observations introduce complications which are not involved in the electron diffraction situations.

For electrons we consider a small crystal bathed in incident radiation which can be so well collimated that it can be approximated by a plane wave. The thickness fringes which have periodicities of the order of hundreds or sometimes thousands of Å are viewed with an electron microscope. The diffraction angles are so small that the difference in path of incident and diffracted beams through the crystals can often be neglected. For X-rays the same possibilities for collimation of the beam and formation of the image by microscope methods are not available. The thickness fringe periodicities are of the order of hundreds of microns and so can be recorded directly on photographic film and then enlarged optically, but the specimens are of the order of millimeters in dimension and it is not practicable to produce well-collimated beams of reasonable

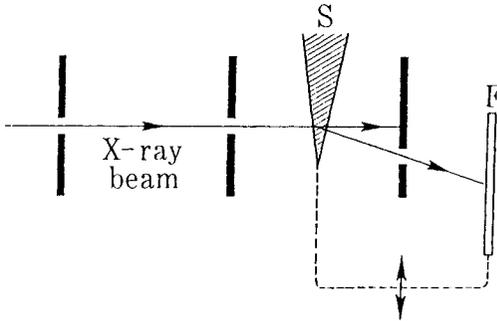


Fig. 9.8. The method used to obtain X-ray "projection topographs". The sample, S, in this case a wedge crystal, and the film F are rigidly coupled and move backward and forward together.

intensity over such areas. Two special experimental arrangements have been devised. The equipment introduced by Lang [1959] for obtaining X-ray "projection topographs" of crystals for the study of dislocations and other crystal defects may be employed. The principle of this is illustrated in Fig. 9.8. A set of fine slits define the X-ray beam incident on the sample S and the beam is diffracted at a given angle. The sample S and film F are coupled and moved back and forward uniformly together. The film then records the variation of diffracted intensity as a function of position on the specimen and for a wedge-shaped crystal this shows thickness fringes (Kato and Lang [1959]). In the second dimension, along the length of the slits, although the beam diverges from the source, this does not change the angle of incidence on the diffracting planes appreciably. Hence in this direction the diffraction conditions are uniform.

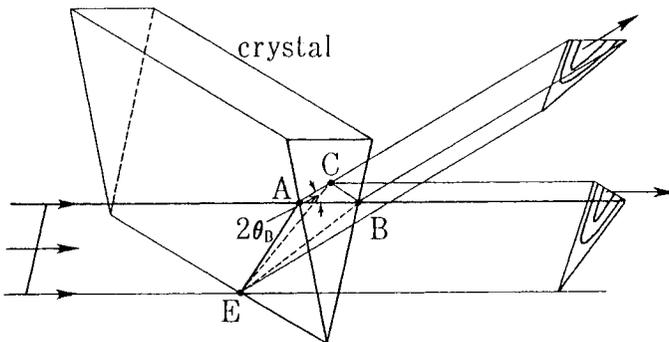


Fig. 9.9. The production of a "section topography" when a thin, flat collimated X-ray beam intersects a wedge-shaped crystal.

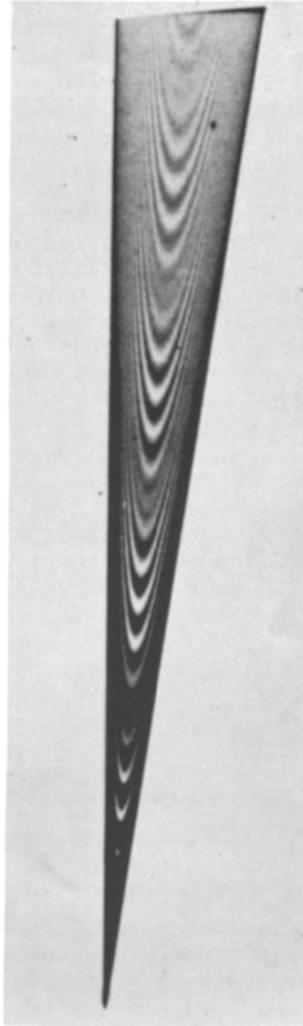


Fig. 9.10. Reproduction of a "section topograph" obtained from a silicon wedge using a 400 reflection by the method of Fig. 9.9. (From Kato [1969]: after T. Kajimura).

The resolution in the "topograph" depends on the source size in relation to the distances from source to specimen to film.

The second type of experiment introduced by Kato and Lang [1959] is that of the "section topograph" in which a thin flat, well-monochromated X-ray beam cuts across a wedge-shaped crystal at a considerable angle to the edge, as suggested in Fig. 9.9. The incident beam strikes the crystal along the line AE. The direct transmitted beam leaves the crystal along the line EB, but after being diffracted through an angle of $2\theta_B$ and back into the incident beam direction, the beam leaving the crystal in that direction comes from the whole of the triangle EBC. Similarly the diffracted beam comes from the whole of the triangle EBC. Hence a photographic film placed after the crystal shows two triangles of intensity. For the electron diffraction case one would expect these triangles to be crossed by equal thickness fringes parallel to the crystal edge. Instead they show the hook-shaped fringes which can be seen in the reproduction of a section topograph from a silicon wedge shown in Fig. 9.10.

9.2.2. Energy flow

The above results serve to emphasize the difference in the experimental situation for X-rays and electrons. Firstly, because the diffraction angles are large and the width of the beam is small compared with the crystal dimensions, the path of the beam through the crystal can be traced in detail. As can be seen in Fig. 9.10, for thin crystals the beam intensity is spread fairly evenly over the possible range of directions, k_0 to k_h , in the crystal. For thicker crystals the energy is concentrated in a particular direction. This is the direction of energy flow, given by the Poynting vector. It was shown by von Laue [1952] that the Poynting vector is

$$S = k_0 |E_0|^2 + k_h |E_h|^2, \quad (8)$$

where E_0 and E_h are the amplitudes of the electrical field vector in the crystal, equivalent to the scalar amplitudes ψ_0 and ψ_h .

Thus at the exact Bragg angle, when the amplitudes of incident and transmitted beams are equal, the direction of energy flow is half-way between k_0 and k_h i.e. it is parallel to the diffracting planes. Then for a thick crystal the path of the X-ray beam may be sketched as in Fig. 9.11. With increasing deviation from the Bragg angle the direction of energy flow reverts to that of the incident beam.

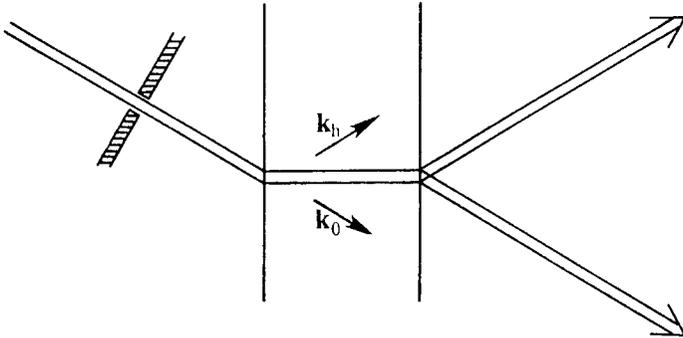


Fig. 9.11. Energy flow and diffracted beam configurations for a narrow X-ray beam incident on a thick perfect crystal.

9.2.3. Dispersion surface picture

A further difference from the electron case arises from the relative strengths of the interaction with the crystal. In terms of the dispersion surface representation we have seen that the deviation of the dispersion surface from the two intersecting spheres of Fig. 8.3 depends on this strength of the interaction since, for example, the anpassing at the Bragg angle is, from (8.12), equal to $|v_h| \cos \theta_h / 2\kappa$. For the Bragg case of reflection from the surface of a large crystal, for example, the angular width of rocking curve for the reflection is equal to twice this quantity. For X-rays this is of the order of 10^{-5} radians or less whereas for electrons it is of the order of 10^{-2} radians.

On the other hand the incident electron beam may be collimated to have an angular spread of 10^{-5} radians or less but for X-rays the divergence of the radiation from each point of the source gives a variation of incident angle across the illuminated region of the specimen (about $20 \mu\text{m}$ wide) of about 10^{-4} radians. Thus for electrons the plane wave approximation is good, but for X-rays one must consider a coherent spherical wave from each point of the source, with a variation of incident angle considerably greater than the angular width of the Bragg reflection. Then in the dispersion surface picture it is not possible to consider a single incident direction defining two tie-points on the two branches of the surface as in Fig. 8.3. Instead it must be considered that the whole region of the dispersion surface around L_0 is excited simultaneously and coherently. This situation was realized by Kato and Lang [1959] and Kato [1961] showed how to integrate over the spherical wave front and obtain expressions giving a reasonable account of the

features of the section topographs. The intensity of the thickness fringes obtained in the projection topographs is then derived by integration of the section topograph along equal-thickness lines.

A further complication of the X-ray diffraction situation is indicated by the periodic fading of the fringes in the section topograph, Fig. 9.10. This has been shown to be a result of polarization effects. The two components of the displacement vector of the electromagnetic field, polarized parallel and perpendicular to the diffraction plane, have different wave vectors and amplitudes in the crystal as a result of the factor P in the dispersion equations (see equation 8.14)). Hence two independent sets of fringes having slightly different periodicities are produced. The periodic modulation of the fringe contrast then results from the beating together of these two periodicities.

The spherical wave theory of Kato permits the interpretation of the X-ray thickness fringes with sufficient accuracy to allow them to be used for the determination of structure factors with high accuracy. Only in exceptional circumstances is there any complication arising from the limitations of the two-beam approximation since n -beam diffraction situations can be readily avoided. The main limitation of the method arises from the need to use crystals which are very nearly perfect, with no dislocations and no distortion. Such crystals can be obtained for silicon and germanium but for very few other substances. A summary of the methods and results has been given by Kato [1969] and Pinsker [1978].

9.2.4. Neutron diffraction

While neutron diffraction experiments may be devised to parallel those with X-ray diffraction, the most successful experiments with neutrons have been those of Shull [1968] which were designed to take advantage of the particular properties of the neutron source. The beam from a nuclear reactor is "white radiation" with a smooth distribution of intensity with wavelength. In the absence of the strong monochromatic "characteristic" lines of the X-ray spectrum it is wasteful of the rather scant supply of neutrons to make use of a strongly monochromated beam. Instead Shull used the wavelength as the variable for this experiment. With an experimental arrangement in which the beam paths were restricted to those of Fig. 9.11, the wavelength diffracted was varied by varying the angle of incidence. Then because the extinction length (8.21) is a function of wavelength, the transmitted intensity fluctuated as the number of extinction lengths in the crystal changed with crystal orientation. The results showed very good agreement with theory.

9.3. Borrmann effect

The anomalous transmission of X-rays through crystals, first observed by Borrmann [1941, 1950] is, at first sight, a surprising phenomenon: the transmission of a thick perfect crystal may increase by a very large factor when the crystal is rotated into the correct orientation for a strong Bragg reflection. This is in contrast with the thin crystal case for which the intensity of the transmitted beam is decreased when energy is taken from it to form a diffracted beam.

The phenomenon is a consequence of the effects of absorption on dynamical diffraction and can be understood to a first approximation by an extension of our discussion of absorption effects in 9.1.7. The two Bloch waves, as suggested by Fig. 9.1, have different absorption coefficients since for Bloch wave 2 the electrons pass between the rows of atoms while for Bloch wave 1 they are travelling mostly in the vicinity of the atoms and so have a higher probability of absorption. From equations (6) and (7) the oscillatory part of the transmitted beams in both the incident and diffracted directions is decreased by the exponential factor $\exp\{-\mu_0 H\}$ while the non-oscillatory part of both is composed of the two parts, having effective absorption coefficients $\mu_0 \pm \mu_h$. With increasing thickness the part with largest absorption coefficient dies away first, then the oscillatory part, so that for sufficiently large thicknesses only the part with absorption coefficient $\mu_0 - \mu_h$ remains. The intensities in the incident and diffracted beam directions are then the same. Provided that μ_h is an appreciable fraction of μ_0 , the intensity of each may easily exceed that of a beam in a non-diffracting orientation for which the absorption coefficient is μ_0 .

For X-rays the absorption process is highly localized since it comes mostly from the excitation of electrons from the inner shells of atoms. Thus the Fourier transform of the absorption function will die away very slowly with distance from the origin of reciprocal space, and the value corresponding to the diffracted beam direction μ_h may be very little less than that for the forward direction μ_0 .

The complications of the X-ray case by the vector nature of the amplitudes and the polarization effects modify the simple scalar result. The absorption coefficients for the two wave fields are given by (Batterman and Cole [1964])

$$\mu = \mu_0 \left(1 \pm P \frac{2|E_0||E_h|\mu_h}{E_0^2 + E_h^2 \mu_0} \right). \quad (9)$$

The two values of P for the two polarizations imply that there

are four portions of the beam having different absorption coefficients. The part with the smallest absorption coefficient may have an absorption coefficient which is very small compared with μ_0 and may give a beam traversing a large thickness of crystal and being almost perfectly polarized.

For example, for transmission of Cu $K\alpha$ radiation through a crystal of Ge of thickness 1 mm, the absorption without any diffraction gives an attenuation of $\exp\{-38\}$. At the Bragg condition for the (220) reflection the ratio μ_h/μ_0 is 0.95 and the attenuations on the four parts of the beam are

$$\begin{aligned} \sigma \text{ polarization, branch 2: } & \exp\{-1.9\}, \\ \pi \text{ polarization, branch 2: } & \exp\{-12.5\}, \\ \pi \text{ polarization, branch 1: } & \exp\{-63.5\}, \\ \sigma \text{ polarization, branch 1: } & \exp\{-74\}. \end{aligned}$$

Hence virtually all the transmitted energy comes from the one branch with σ polarization. The enhancement of transmission relative to the non-diffracting crystal is by a factor of $(1/4) \exp\{+36.1\}$.

Because of the smallness of the angular range over which X-ray reflections take place for perfect crystals, the Borrmann effect transmission gives X-ray beams which are very well collimated as well as being almost perfectly polarized. This has provided the incentive and basis for a considerable expansion of the possibilities for experiments involving precision measurements on near-perfect crystals and a means for a more complete study of the diffraction, absorption and scattering processes of X-rays.

For electron diffraction the absorption comes mostly from plasmon excitation, which contributes to μ_0 only, and thermal diffuse scattering which gives a μ_h value falling off moderating rapidly with scattering angle (Hall and Hirsch [1968]). The ratio of μ_h to μ_0 is usually small and the Borrmann effect not very striking.

One interesting consequence of the variation of absorption processes under Borrmann diffraction conditions is that changes in absorption in turn affect the intensity of any secondary radiation produced as a result of the absorption. Thus when electrons are strongly diffracted in a thick crystal there is a variation of the X-ray emission with angle of incidence (Duncumb [1962]; Miyake et al. [1968]). Likewise Knowles [1956] showed that for neutron diffraction the establishment of standing wave fields which reduced the density the neutrons in the vicinity of the nuclei of the atoms have the effect of decreasing the inelastic interaction of the neutrons with the nuclei which give rise to a γ -emission. For X-rays the corresponding observation made by Batterman [1962, 1964] is that the intensity of secondary

fluorescence emission from the atoms can be used as an indication of the electric field strength of the incident X-radiation on the atomic planes. A number of further observations of related nature involving Kikuchi- or Kossel-type patterns, especially with electrons, have been reported recently and will be discussed in Chapter 14.

Problems

1. A cube-shaped crystal of MgO is oriented to give a 220 reflection with the incident beam very nearly parallel to the body diagonal of the cube. Find the form and dimensions of the group of fine-structure spots formed around the normal diffraction spot position assuming a two-beam dynamical approximation and taking $\phi_0 = 13.5$ volts, $\phi_{220} = 5.0$ volts. How will the form of the group of spots vary with deviation from the Bragg condition? (See Cowley, Goodman and Rees [1957].)

2. Find the intensity distribution of the convergent beam diffraction pattern from a flat plate-shaped crystal of MgO set to give the 220 reflection when the thickness of the crystal is equal to (a) one half of the extinction distance and (b) twice the extinction distance. Assume the two beam approximation with no absorption.

3. Find how the intensity distributions in problem 2 will be modified in the presence of absorption if $\mu_{220}/\mu_0 = 0.2$. Find the sum of the transmitted and diffracted beam intensities and hence the energy lost by absorption as a function of incident angle. Hence, on the assumption that a constant fraction of the absorption gives rise to the emission of soft X-rays, find the variation of X-ray emission with orientation of the incident electron beam.

4. A beam of electrons of energy 50 keV strikes a flat surface of a perfect single crystal of MgO at a grazing incidence of a few degrees. Assuming a mean inner potential $\phi_0 = 13.5$ volts find an expression for the change of angle of the incident beam when it enters the crystal. If the surface is parallel to the 220 plane, find the orientation of the incident and diffracted beams, outside of the crystal, for the exact Bragg condition for 220 and 440 reflections. Assuming $\phi_{220} = 5.0$ volts and $\phi_{440} = 1.6$ volts find the angular width of the rocking curve for these reflections, assuming the two-beam approximation.

This Page Intentionally Left Blank

Extension to many beams

10.1. Dynamical n -beam diffraction

We have seen that for X-ray diffraction from a perfect crystal the angular width of a reflection is of the order of 10^{-5} radians. For a crystal of simple structure the angular separation of reflections is of the order of 10^{-1} radians. The occurrence of more than one reflection at a time is a very special case which, experimentally, must be sought with particular care. When three-beam or many-beam effects do occur they give some striking modifications of intensity, as seen in the Borrmann effect studies made with wide angle convergent beams (e.g. Borrmann and Hartwig [1965]). A theoretical study of the three-beam case has been made by Ewald and Heno [1968] (also Heno and Ewald [1968]).

For electron diffraction, as we have seen, the simultaneous appearance of more than one diffracted beam is the rule rather than the exception. If special care is taken to choose an appropriate orientation it is possible that as much as 99 percent of the energy of the Bragg beams may be contained in two strong beams for a thick crystal. On the other hand for thin crystals in principal orientations, such as that giving the diffraction pattern, Fig. 6.3 or Fig. 13.4(c), the number of beams occurring simultaneously may amount to several hundred. Because of the strong scattering of electrons by all but the lightest of atoms, the amplitudes of the scattered beams from even a thin crystal may be sufficiently strong to ensure that there is considerable multiple coherent scattering and hence we must deal with the dynamical theory for a very large number of diffracted beams. With the advent of high-voltage electron microscopes, with accelerating voltages up to 1 MeV or more, the occurrence of patterns having a very large number of diffracted beams has become even more common (Uyeda [1968]). The wave-length decreases even more rapidly with accelerating voltage than would be inferred from the non-relativistic relationship, since the relativistic form is

$$\lambda = \left\{ \frac{2m_0eE}{h^2} \left(1 + \frac{eE}{2m_0c^2} \right) \right\}^{-1/2}. \quad (1)$$

On the other hand the interaction constant σ is given by

$$\sigma = \frac{\pi}{E\lambda} \frac{2}{1+(1-\beta^2)^{1/2}} = \frac{2\pi m_0 e \lambda}{h^2} \left(1 + \frac{h^2}{m_0^2 c^2 \lambda^2} \right)^{1/2} \quad (2)$$

where $\beta = v/c$. As the voltage increases the value of σ decreases more and more slowly, tending to a limiting constant value.

With the decrease of wavelength the Ewald sphere becomes more nearly planar so that it may intersect a greater number of the extended reciprocal lattice points. At the same time the strength of interaction with atoms and hence the strength of dynamical interactions ceases to decrease. As a result, the number of interacting beams increases rapidly with accelerating voltage.

Hence the need for n -beam dynamical treatment arises most urgently for electrons and because of the variety of observed effects and theoretical complications, a number of different approaches have been applied. Most of these take advantage of the fact that for high voltage electrons the simplifying approximation of small angle scattering is possible. Hence in this chapter we discuss n -beam diffraction theory entirely from the point of view of high-voltage electrons.

One of the most powerful and effective approaches to the n -beam diffraction problem is the so-called physical optics approach, based on the type of approach to diffraction contained in our first few chapters. Since this has been expounded in print to only a limited extent we reserve a special chapter, the next, to its description. In the present chapter we treat other approaches, some more widely known and used, which follow more immediately from the treatment of the last two chapters.

The application of the Born series method by Fujiwara [1959] involves the evaluation of the series of equations (1.17) and (1.22), where the scattering potential is that of a flat plate-like crystal, with

$$\varphi(r) = c \sum_h v_h \exp\{2\pi i h \cdot r\}$$

within the limits $0 \leq z \leq H$.

This leads to the same type of general expressions for the n -beam diffraction amplitudes as will be derived by other methods. Because of the notoriously poor convergence of the Born series, this approach has not led directly to practical means for calculating intensities. Hence, while referring the reader to Fujiwara's interesting papers we do not offer any more detailed discussion.

When more than two beams occur in a crystal, the intensities in diffraction patterns and in images depend on both the magnitudes and relative phases of the structure amplitudes. This follows because in coherent interference effects it is the amplitudes, not the intensities, of the waves which are added. If the dynamical scattering process could be inverted, the structure amplitudes and hence the atom positions could be deduced directly and unambiguously from the diffraction data.

It seems that in principle it should be possible to invert the dynamical diffraction process. By using the variables of incident beam direction, specimen thickness and wavelength of the radiation it should be possible to obtain sufficient experimental data to over-determine the problem. So far no general solution has been found.

A major step in this direction, however, is the demonstration that for a centrosymmetric crystal, a complete inversion of the diffracted intensity distribution is possible for the three-beam case (Moodie [1979]). If convergent beam diffraction patterns are used, so that incident beam orientation is a variable, it may be shown that within the three beam pattern it is possible to distinguish what are essentially two-beam intensity distributions along particular lines and from the positions of these unique lines it is possible to derive the magnitudes of v_h , v_g , v_{h-g} and the sign of the product $v_h v_g v_{h-g}$.

Extensions of this type of solution to include more beams or structures not having a center of symmetry are difficult and may well involve the introduction of additional experimental variables.

10.2. Extension of Bethe theory - transmission

10.2.1. Matrix formulation

From the starting point of the Schrödinger equation for an electron in a periodic potential, equation (8.1), we derived the general matrix equation (8.7) relating the amplitudes Ψ_h and wave vectors k_h of the wave-field set up when an incident beam enters a crystal,

If we include N reciprocal lattice points there are $2N$ solutions given by the condition that the determinant of the matrix is zero. The second order terms on the diagonal of the $N \times N$ matrix create the difficulty. In general it can be shown that the $2N$ solutions are given by use of a $2N \times 2N$ matrix, having only first order terms, in an equation of the form

$$\det \begin{pmatrix} B - xI & -Q \\ I & 0 - xI \end{pmatrix} = 0$$

where B , Q , I and O are $N \times N$ matrices; I and O being the identity and null matrices (Collela [1972]; Moon [1972]).

The $2N$ eigenvalues correspond to waves travelling both forward and backward in relation to the incident beam. If as in the case of high-energy electron scattering we can make the assumption that back-scattering is negligible, N of the solutions can be ignored. The problem is simplified by reducing the diagonal terms to the first order by taking only the one value of $\kappa^2 - k_h^2$ which corresponds to a wave almost parallel to the incident beam, and making a small angle approximation.

The cosines of the diffracting angles are put equal to unity. Since in general the Fourier coefficients of the crystal potential, Φ_h , are very much smaller than the accelerating potential of the incident beam, the anpassung values ξ^i are very much less than κ . Hence, neglecting second order terms in ξ^i we may write, from (8.9),

$$k_h = \kappa - \zeta_h - \xi, \quad (3)$$

and

$$\kappa^2 - k_h^2 = 2\kappa\xi + 2\kappa\zeta_h - \zeta_h^2 \equiv -x + p_h, \quad (4)$$

where we have put

$$x = -2\kappa\xi \quad (5)$$

$$p_h = 2\kappa\zeta_h - \zeta_h^2, \quad (6)$$

so that the various values of x correspond to the values of ξ^i for the various Bloch waves while the values for p_h depend on the excitation errors for the various reciprocal lattice points and are the same for all Bloch waves.

The matrix equation (8.7) can then be re-written in the simplified, linear form

$$M\Psi = x\Psi \quad (7)$$

where

$$M = \begin{pmatrix} \cdots & p_0 & \cdots & v_{0h} & \cdots & v_{0g} & \cdots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ v_{h0} & \cdots & p_h & \cdots & v_{hg} & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ v_{g0} & \cdots & v_{gh} & \cdots & p_g & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} \quad (8)$$

Then the eigenvalues x^i give the values of ζ^i for the different Bloch waves. The eigen vectors Ψ^i have components $\Psi_0^i \dots \Psi_h^i \dots$, which are the amplitudes of the individual plane waves which make up the Bloch waves.

In the absence of absorption, the matrix M is Hermitian. For a centrosymmetric structure it is real, symmetric.

10.2.2. Small angle approximation

Frequently, because of the short wavelength of the electrons and consequent flatness of the Ewald sphere, the only reflections which are important are those corresponding to reciprocal lattice points in one plane almost perpendicular to the incident beam. For these conditions, as pointed out by Fujimoto [1959], the solution of the matrix equation is independent of the magnitude of κ but depends only on the projection of κ on the operative plane of the reciprocal lattice.

This is seen from the geometry of Fig. 10.1 for which the angles θ_0 and θ_h have been greatly exaggerated. To a reasonable approximation we may write

$$\zeta_h = (\kappa \sin\theta_h - \kappa \sin\theta_0) \sin\theta_h. \quad (9)$$

Then $p_h (= 2\kappa\zeta_h - \zeta_h^2)$ can be expressed entirely in terms of the quantities $\kappa \sin\theta_h$ and $\kappa \sin\theta_0$, which are projections on to the plane of the h reciprocal lattice points, and $\sin^2\theta_h$ which is negligible. Hence the matrix M is independent of the magnitude of κ and the amplitudes of the diffracted beams should be the same for different wavelengths and energies of the incident beam in this approximation, given equivalent orientations. This conclusion does not appear to have been adequately tested experimentally.

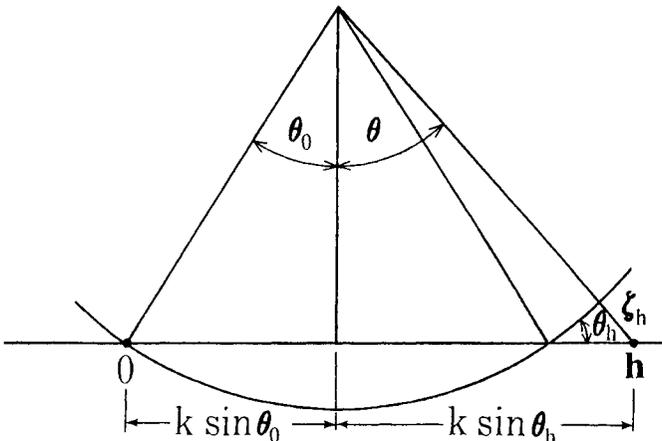


Fig. 10.1. Definition of the quantities involved in forward-scattering geometry.

10.2.3. Bloch waves and boundary conditions

We have seen that there are, in general, an infinite number of solutions of the equation (7). The solution number i gives the amplitudes Ψ_h^i of the plane waves which make up the Bloch wave number i . The intensities of the diffracted beams for any real experimental situation are determined by the distribution of energy between the various Bloch waves and this is determined by the boundary conditions. Thus to get the amplitude of each wave in the crystal and so the amplitudes of the vacuum waves emerging from the crystal we add the Bloch waves with a weighting factor α_i so that the amplitude of the wave corresponding to the h reciprocal lattice point in the Bloch wave i is written

$$u_h^i = \alpha_i \Psi_h^i. \quad (10)$$

The boundary conditions at the entrance surface ensure that the sum of the amplitudes of all beams is zero except in the incident beam direction, i.e.

$$\sum_i u_h^i = \sum_i \alpha_i \Psi_h^i = \delta_{h0} \quad (11)$$

where δ_{h0} is the Kronecker delta (= 1 for $h = 0$, and 0 for other h). Multiplying (11) by Ψ_h^{j*} and summing over h gives

$$\sum_h \sum_i \alpha_i \Psi_h^i \Psi_h^{j*} = \sum_h \delta_{h0} \Psi_h^{j*}. \quad (12)$$

Provided the Bloch waves are orthogonalized and normalised, $\sum_h \Psi_h^i \Psi_h^{j*} = 0$ unless $i=j$, and $\sum_h \Psi_h^j \Psi_h^{j*} = 1$. Then (12) becomes

$$\alpha_j \sum_h \Psi_h^j \Psi_h^{j*} = \Psi_0^{j*}$$

or

$$\alpha_j = \Psi_0^{j*} \quad (13)$$

and

$$u_h^i = \Psi_0^{i*} \Psi_h^i. \quad (14)$$

Thus the multiplier for the amplitudes of each Bloch wave is the complex conjugate of the zero Fourier coefficient of that Bloch wave.

For a parallel sided plate-shaped crystal, all waves in the crystal corresponding to the one reciprocal lattice point h are refracted into the same direction at the exit face to form the vacuum diffracted wave K_h . In the crystal the amplitude of this wave are

$$\Psi_h = \sum_i u^i_h \exp\{2\pi i k^i_h \cdot r\}. \tag{15}$$

Refraction at the exit face gives a change in length of the projection of the wave vector on the surface normal, as suggested in Fig. 10.2. In the small angle approximation, the difference in projection of κ and K_h is $\xi^i = x^i/2\kappa$, and the difference in projection of K_0 and κ is $v_0/2\kappa$. Then if the surface normal is unit vector z and $r \cdot z = z$, we have

$$\begin{aligned} k^i_h \cdot z &= \kappa \cdot z - \xi^i \\ \kappa \cdot z &= K_h \cdot z - v_0/2\kappa. \end{aligned} \tag{16}$$

Then, since the x and y components of k^i_h and K_h are equal, we can insert in (15),

$$\exp\{2\pi i k^i_h \cdot r\} = \exp\{2\pi i K_h \cdot r\} \exp\{\pi i v_0 z / \kappa\} \exp\{2\pi i \xi^i z\}. \tag{17}$$

Hence the vacuum wave function for the h reflection wave is

$$\psi^y_h(r) = \left[\sum_i \Psi_i^* \Psi^i_h \exp\left\{\frac{\pi i z x^i}{\kappa}\right\} \right] \exp\{2\pi i K_h \cdot r\}. \tag{18}$$

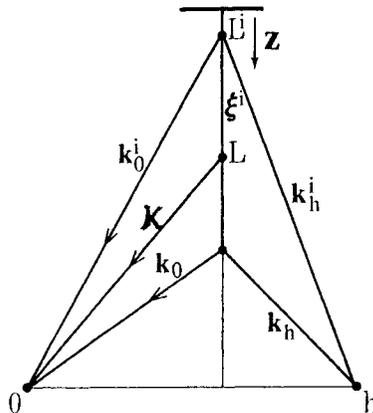


Fig. 10.2. The relationship of wave vectors in a crystal and in space for a two-beam case.

The exponential term $\exp\{\pi i v_0 z / \kappa\}$ is neglected because it is a common phase term for all beams. The intensity of the diffracted beam is then given by the square of the modulus of the summation in the square bracket. In order to calculate the amplitudes and intensities of the diffracted beams it is therefore necessary to find the eigen-values x^i and eigenvectors Ψ^i for the matrix equations (7) and (8).

As in Chapter 8 we may include absorption explicitly and deal with real quantities p_h and v_{hg} by adding the Fourier coefficients of the absorption function μ_0 and μ_h so that (8) becomes

$$M = \begin{pmatrix} \cdots p_0 \cdots v_{0h} \cdots v_{0g} \cdots \\ \cdots v_{h0} \cdots p_h \cdots v_{hg} \cdots \\ \cdots v_{g0} \cdots v_{gh} \cdots p_g \cdots \end{pmatrix} + i \begin{pmatrix} \cdots \mu_0 \cdots \mu_{0h} \cdots \mu_{0g} \cdots \\ \cdots \mu_{h0} \cdots \mu_0 \cdots \mu_{hg} \cdots \\ \cdots \mu_{g0} \cdots \mu_{gh} \cdots \mu_0 \cdots \end{pmatrix} \quad (19)$$

Then both of these matrices are Hermitian, but the sum of the two is general, complex.

10.2.4. The scattering matrix

Proceeding further, we now express the result in a rather different form, that of a scattering matrix. For a particular eigen-value, x^i , we have from (7)

$$M\Psi^i = x^i\Psi^i. \quad (20)$$

Since x^i is a simple number, applying the matrix to the right-hand side of (20) gives

$$M^2\Psi^i = (x^i)^2\Psi^i,$$

and, in general

$$M^n\Psi^i = (x^i)^n\Psi^i. \quad (21)$$

Then for the h component of the i Bloch wave

$$(x^i)^n\Psi^i_h = \sum_g \Psi^i_g (M^n)_{hg}. \quad (22)$$

where $(M^n)_{hg}$ is the component in the h column and g row of the matrix M^n .

In order to apply the boundary condition, we multiply both sides of (22) by Ψ^{i*}_0 and sum over i to give

$$\begin{aligned} \sum_i (x^i)^n\Psi^i_h\Psi^{i*}_0 &= \sum_i \sum_g \Psi^i_g\Psi^{i*}_0 (M^n)_{hg} \\ &= \sum_g \left(\sum_i \Psi^i_g\Psi^{i*}_0 \right) (M^n)_{hg} = (M^n)_{h0}, \end{aligned} \quad (23)$$

since the boundary condition (11) with $\alpha_i = \Psi_i^*{}_0$ gives

$$\sum_i \Psi_i^*{}_0 \Psi_i^h = \delta_{h0}.$$

Then, expanding the exponential in (18) we have

$$\begin{aligned} \Psi^v_h(r) &= \sum_i \Psi_i^*{}_0 \Psi_i^h \exp\left\{\frac{\pi izx^i}{\kappa}\right\} \exp\{2\pi i K_h \cdot r\} \\ &= \sum_i \Psi_i^*{}_0 \Psi_i^h \sum_n \frac{1}{n!} \left(\frac{\pi izx^i}{\kappa}\right)^n \exp\{2\pi i K_h \cdot r\} \\ &= \sum_n \frac{1}{n!} \left(\frac{\pi iz}{\kappa}\right)^n (M^n)_{h0} \exp\{2\pi i K_h \cdot r\}. \end{aligned} \tag{24}$$

By convention we write for matrices as for ordinary functions

$$\exp\{iM\} = I + iM - \frac{1}{2!}M^2 + \dots \tag{25}$$

and the matrix components are

$$[\exp\{iM\}]_{hg} = (I)_{hg} + i(M)_{hg} - \frac{1}{2!}(M^2)_{hg} + \dots \tag{26}$$

Hence, from (24), the amplitude of the h wave in vacuum, leaving the crystal of thickness z is

$$\left[\exp\left\{\frac{\pi iz}{\kappa} M\right\} \right]_{h0},$$

and in general we may write

$$\begin{pmatrix} \Psi_0 \\ \Psi_1 \\ \vdots \\ \Psi_h \\ \vdots \end{pmatrix} = \exp\left\{\frac{\pi iz M}{\kappa}\right\} \begin{pmatrix} \Psi_0^{(0)} \\ 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \end{pmatrix}, \tag{27}$$

where $\Psi_0^{(0)}$ is the incident wave amplitude and operation on the vector $\Psi^{(0)}$ with the exponential of the matrix gives the vector $\Psi^{(1)}$ which has components Ψ_h , the amplitudes of the

waves leaving the crystal. Thus propagation of a wave field through a crystal of thickness z is represented by the operation of the scattering matrix

$$S = \exp\left\{\frac{\pi izM}{\kappa}\right\}. \quad (28)$$

In the application of this concept to the calculation of dynamical intensities by Sturkey [1962], the scattering matrix is first calculated for a thin slice of crystal of thickness D . Then transmission through successive identical slices of crystal will be simulated by the repeated application of the scattering matrix so that we may write for n slices

$$\Psi^{(n)} = S^n \Psi^{(0)}$$

and

$$S^n = \exp\left\{\frac{\pi i D n M}{\kappa}\right\}. \quad (29)$$

For non-identical slices of crystal the same approach is valid if for each slice the operation on the wave field vector is made by means of the appropriate matrix. In this way it is possible to apply the formulation to problems of faults and imperfections or variations of structure of a crystal. A scattering matrix must be generated for each type of slice involved.

10.2.5. Derivation of the two-beam approximation

It has been shown by Rez [1977] and Goodman [private communication] that the two-beam approximation for transmission through a crystal plate follows very readily from the scattering matrix formulation. The wave function after traversing a distance z is given by equation (28) as

$$\begin{pmatrix} \Psi_0 \\ \Psi_h \end{pmatrix} = \exp\left\{\frac{\pi izM}{\kappa}\right\} \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \quad (30)$$

If v_h is real, i.e. for the centrosymmetric case without absorption, the matrix M is written, from (8),

$$M = \begin{pmatrix} -\frac{1}{2}p_h & v_h \\ v_h & \frac{1}{2}p_h \end{pmatrix} = -\frac{1}{2}p_h \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + v_h \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (31)$$

where we have subtracted out $\frac{1}{2}p_h$ which adds the same phase factor to each beam.

The matrices in the final part of (31) are Pauli spin matrices σ_1 and σ_2 for which $\sigma_1^2 = \sigma_2^2 = I$ and $\sigma_1 \cdot \sigma_2 = 0$ so that the terms in the expansion of the exponential, (25), take the form

$$\begin{aligned} M &= A + B = a\sigma_1 + b\sigma_2 \\ M^2 &= A^2 + B^2 = (a^2 + b^2) I \equiv r^2 I \\ M^3 &= (A^2 + B^2)(A + B) = M \cdot (a^2 + b^2) \\ M^4 &= (A^2 + B^2)^2 = r^4 \cdot I. \end{aligned}$$

Here we have put $r^2 = a^2 + b^2 = \frac{1}{4}p^2h + v^2h$. Summing the odd and even terms of the series separately gives

$$\exp\left\{\frac{\pi izM}{\kappa}\right\} = \cos\left(\frac{\pi zr}{\kappa}\right) \cdot I + ir^{-1} \sin\left(\frac{\pi zr}{\kappa}\right) \cdot M.$$

By putting $ph = 2\kappa\zeta_h$ (assuming $\zeta_h \ll \kappa$) and $r^2 = \kappa^2\zeta_h^2 + v^2h$, the wave amplitudes at the depth z are then given from (30) as

$$\begin{aligned} \Psi_0 &= \cos\left(\frac{\pi zr}{\kappa}\right) - i \frac{p_h}{2r} \sin\left(\frac{\pi zr}{\kappa}\right) \\ &= \cos\{\pi z (\zeta_h^2 + v^2h/\kappa^2)^{1/2}\} - i \frac{\zeta_h}{(\zeta_h^2 + v^2h/\kappa^2)^{1/2}} \frac{1}{2} \sin\{\pi z (\zeta_h^2 + v^2h/\kappa^2)^{1/2}\} \end{aligned}$$

$$\begin{aligned} \Psi_h &= v_h/r \sin\left(\frac{\pi zr}{\kappa}\right) \\ &= \frac{v_h/\kappa}{(\zeta_h^2 + v^2h/\kappa^2)} \sin\{\pi z (\zeta_h^2 + v^2h/\kappa^2)^{1/2}\}. \end{aligned}$$

Putting the crystal thickness $H = Z$ and $\kappa = 1/\lambda$ then makes this result identical with (8.28) and (8.29).

The case of v_h complex ($= v_h + iv_h^i$) can be treated by a simple extension, using the three Pauli spin matrices

$$\begin{aligned} M &= -p_h \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + v_h \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + v_h^i \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \\ &= a\sigma_1 + b\sigma_2 + c\sigma_3. \end{aligned}$$

10.3. The Darwin-type approach

The Darwin treatment of the X-ray diffraction by reflection from a face of a large perfect crystal (Darwin [1914]) involved the establishment of a transmission and reflection coefficient for each plane of atoms and then the summation of transmitted and diffracted amplitudes at each plane. The application to transmission electron diffraction of this type of theory was made by Howie and Whelan [1961] primarily for the purpose of finding the contrast in electron microscope images of defects. The amplitudes of the diffracted beams are considered to be continuous functions of the distance in the beam direction and are related by a set of differential equations. It is essentially a perfect-crystal theory for each slice of the crystal although variations in orientation of the diffracting planes from one slice to another may be included.

For the simple two beam case the relationship between transmitted and scattered amplitudes, Ψ_0 and Ψ_h , corresponding to incident and diffracted beam directions, for propagation in the z direction is given by

$$\begin{aligned}\frac{d\Psi_0}{dz} &= i\sigma\Phi_h\Psi_h, \\ \frac{d\Psi_h}{dz} &= i\sigma\Phi_h\Psi_0 + 2\pi i\zeta_h\Psi_h.\end{aligned}\tag{32}$$

In words, these equations state that the variation of the transmitted beam amplitude with thickness is given by scattering from the h beam with a phase change of $\pi/2$ and a scattering strength proportional to the Fourier coefficient of potential Φ_h . The diffracted beam amplitude is modified by scattering from the transmitted beam but also has a phase change because the wave vector k_h drawn to the reciprocal lattice point is not the same as that to the reciprocal lattice origin, k_0 , but differs from it by the excitation error, ζ_h .

Absorption may be taken into account by making the potential coefficients Φ_h complex so that in the equations, (32) the quantity $i\Phi_h$ is replaced by $i\Phi_h - \Phi'_h$ and there is an overall decrease in both beams due to the average absorption coefficient proportional to Φ'_0 . If one plane of the crystal is translated by a vector R relative to the host lattice then this will change the phase of the diffracted beam by an amount proportional to $h \cdot R$. Thus if the

displacement is a function of z , namely $R(z)$, the equations (32) with absorption become

$$\frac{d\Psi_0}{dz} = -\sigma\Phi'_0\Psi_0 + \sigma(i\Phi_h - \Phi'_h)\Psi_h,$$

$$\frac{d\Psi_h}{dz} = \sigma(i\Phi_h - \Phi'_h)\Psi_0 + \{-\sigma\Phi'_0 + 2\pi i(\zeta_h + \beta_h)\}\Psi_h, \quad (33)$$

where

$$\beta_h = \frac{d}{dz}(h \cdot R(z)). \quad (34)$$

The progressive changes of the amplitudes Ψ_0 and Ψ_h through the crystal may then be followed by integration of (33), i.e. by finding the successive modifications of phase and amplitude of Ψ_0 and Ψ_h together for infinitesimal increments of z .

Like the original Darwin approach, this two-beam version suffers from the difficulty that a slice of infinitesimal thickness will give a very large number of simultaneous beams and a two beam situation is only approached after transmission through a crystal thickness comparable to the extinction distance $2\sigma\Phi_h$ or more. Hence it is more appropriate in principle and, in the case of electrons, for practical considerations to make use of the n -beam form of (33) which may be expressed in matrix form as

$$\frac{d}{dz}\Psi = 2\pi i(A + \beta)\Psi \quad (35)$$

where Ψ is a column vector whose elements Ψ_h are the amplitudes of the diffracted waves, β is a diagonal matrix whose elements are given by (34) and A is a matrix whose elements are

$$A_{hh} = \zeta_h + i\sigma\Phi'_0/4\pi,$$

$$A_{hg} = \sigma(\Phi_{h-g} + i\Phi'_{h-g})/4\pi. \quad (36)$$

Then, from (35) the change in any amplitude depends on all other amplitudes, the strength of the interactions depending

on the potential coefficients Φ_{hg} and the excitation errors of all beams except the incident beam. The amplitudes of all the beams are integrated through the crystal simultaneously.

10.4. Special cases - beam reduction

In general the only way to gain an appreciation of the results of the dynamic interaction of a large number of beams in a crystal is by performing many detailed n -beam calculations for various crystals having a variety of thickness and orientations and attempting to analyse the results. There are special cases, however, for which the n -beam diffraction result may be understood by comparison with a simpler, related result for relatively few beams. These are cases of high symmetry in the diffraction pattern for which members of sets of beams are equivalent in that they have equal excitation errors and interact through equivalent structure factor values, Φ_{hg} . The way in which such sets of equivalent beams may be merged to give, for each set, a single representative beam, was demonstrated by Gjønnes by use of the integral equation representation (Gjønnes [1966]) and this approach was used by Fisher [1968]. A different approach through the matrix formulation of equation (10.8) was given by Fukuhara [1966].

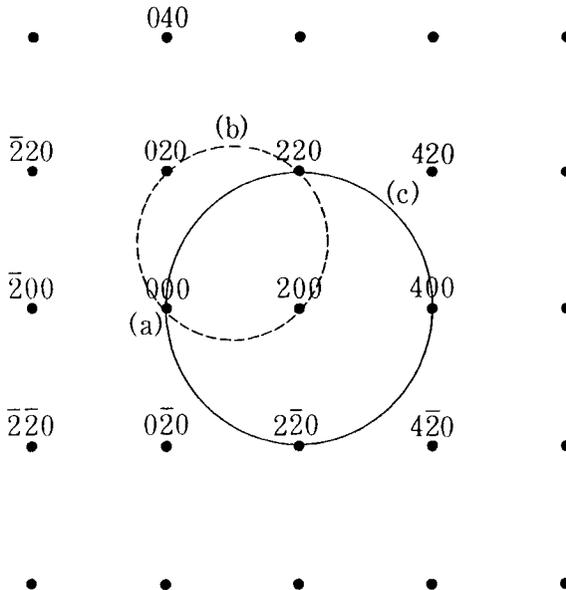


Fig. 10.3. Illustrating the interaction of the Ewald sphere with the hko plane of reflections for a F.C.C. crystal in three special cases where beam-reduction techniques can be used.

For purposes of illustration we refer to the particular case of an electron beam almost parallel to the cube axis of a face-centered cubic structure. The plane of $hk0$ reciprocal lattice spots near the Ewald sphere is sketched in Fig. 10.3.

For the situation (a), when the incident beam is exactly parallel to the c -axis, the four beams 200 , 020 , $\bar{2}00$ and $0\bar{2}0$ are exactly equivalent having the same excitation error, interacting between themselves through the Fourier coefficients of the types Φ_{220} and Φ_{400} , interacting with the zero beam through coefficients Φ_{200} , interacting with the set of 200 beams through Φ_{200} and Φ_{420} and so on. Thus the four 200 -type beams may be combined to give an equivalent single beam with definable interactions with the 000 beam, the representative beam of the 220 set, and so on.

If a 5-beam calculation is to be made, including only the 000 and the 200 set, this is immediately reduced to a 2-beam calculation. Similarly the problem is greatly reduced if more beams are included. Fisher [1968] calculated the intensities for this orientation (with no absorption) for a 49-beam case including reflections out to the 660 . By symmetry the number of effective beams to be considered is reduced to 10. The result of the calculation shown in Fig. 10.4 indicates the interesting situation that there is something resembling a two beam solution in that the zero beam oscillates out of phase with all the diffracted beams, varying together. This result may be influenced by the special circumstance that the excitation errors for the $hk0$ reflections are proportional to the square of $d^{-1}hk0$ and so to $h^2 + k^2$. Hence the phase changes introduced, in (36) for example, by the excitation errors are integral multiples of 2π when that for the 200 reflections is 2π . An investigation of such special cases by Fejes, Iijima and Cowley [1973] suggests, however, that the situation is more complicated and that the periodicity of the intensity variation with thickness depends on the presence of a particular relationship between scattering amplitudes and excitation errors.

A second special case indicated by (b) in Fig. 10.3 is for a tilt of the incident beam so that the Ewald sphere cuts exactly the 000 , 200 , 220 and 020 reciprocal lattice points. Then the four corresponding beams are equivalent in that they have the same excitation error and interact among themselves and with successive concentric sets of beams, through the same interaction potentials. By grouping these sets of reflections together the effective number of beams is again reduced. In this case the 000 beam is distinguished from the rest of the set when the boundary conditions for the entrance surface are applied but the beam-reduction is effective in simplifying the problem up to that stage.

Similarly the case (c) of Fig. 10.3 with the Ewald sphere

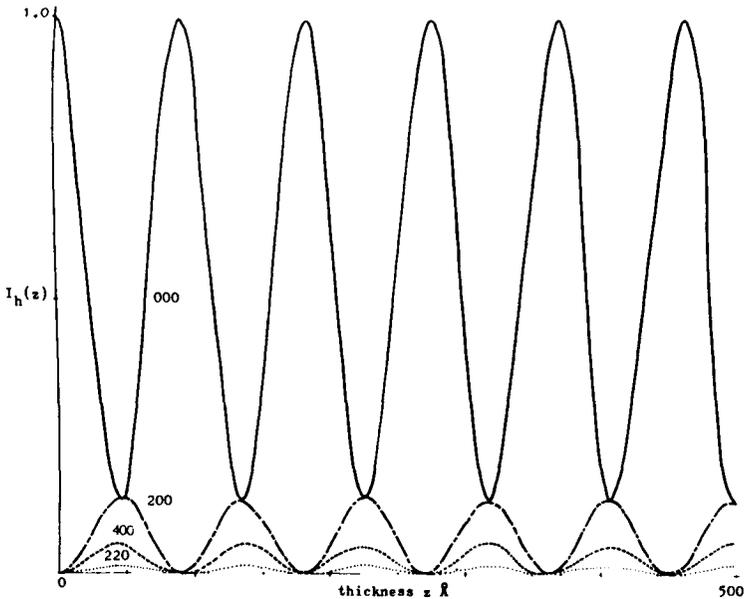


Fig. 10.4. Calculation of the intensities of hko reflections as a function of thickness for the incident beam exactly parallel to c -axis of a F.C.C. copper-gold alloy (disordered Cu_3Au) with no absorption (Fisher [1968]). 100 keV.

passing through the points 000 , 220 , 400 and $2\bar{2}0$ is reducible by symmetry for the infinite crystal formulation. Fukuhara [1966] gives a number of other special cases appropriate for square and hexagonal symmetry of the diffraction pattern.

10.5. Computing methods

In order to obtain a numerical solution to the matrix equation (7) subject to the boundary conditions, any of the standard programs available for modern digital computers may be used to diagonalize the matrix M . Programs based on the Jacobi and QR methods have been used, among others. The problem is greatly complicated in the presence of absorption, when the matrix elements are complex, unless simplifying assumptions based on the relative smallness of the imaginary components are made. However, even for the full problem of a non-centrosymmetric crystal with absorption the amount of

computation is not beyond the capabilities of a moderate-size computer for which a 50-beam problem, involving the diagonalization of a 50 X 50 matrix, would take only a few minutes (Fisher [1971]).

From the diagonalized matrix the eigen-values x^i and the eigenvector components Ψ^i_h may be deduced. Substituting in (18) then allows the diffracted amplitudes for all beams, h , to be derived very rapidly for any number of crystal thicknesses, z .

Since the computing time required increases very rapidly with the number of diffracted beams involved, it is of interest to determine the minimum number of beams which should be included in order to obtain a required degree of accuracy. This number will vary widely in practice, depending on the size of the unit cell projected in the beam direction (i.e. on the density of points in the plane of reciprocal space involved), on the wavelength of the electrons and the atomic numbers of the atoms present.

To give an indication on this point. Fisher [1968] made the calculation for the case (a) of Fig. 10.3 with the incident beam exactly parallel to the c -axis of a face centered cubic disordered copper-gold alloy. For this material the reciprocal lattice points in the plane of

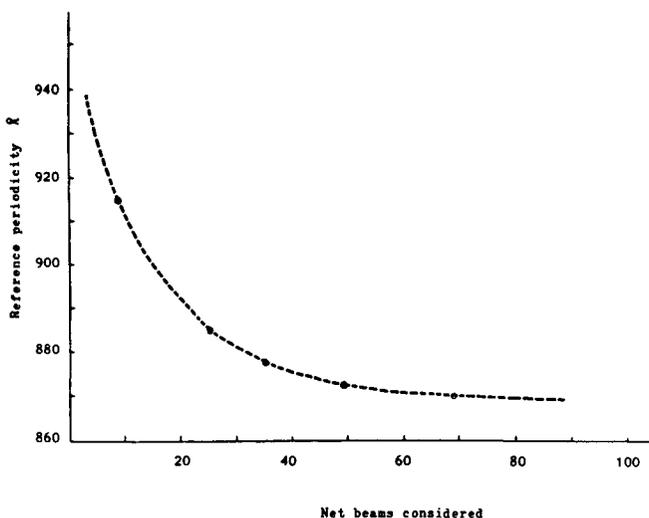


Fig. 10.5. Variation of the periodicity of the intensity oscillations of the beams for CdS, $\langle 120 \rangle$ orientation, 100 keV, as a function of the number of beams included in the calculation (Fisher [1968]).

interest are relatively widely spaced, being 0.5 \AA^{-1} apart, but the average atomic number is moderately high. The general features of the intensity distributions as shown in Fig. 10.4 were established for less than 10 beams but at this stage the relative intensities were badly wrong.

The periodicity of the oscillations of the incident beam intensity, the "extinction distance", was used as an indication of the accuracy of the computation and is plotted against the number of beams considered in Fig. 10.5 for CdS in $\langle 120 \rangle$ orientation. For the case of Fig. 10.4, from a value of about 128 \AA for the simplest, 5-beam case, the periodicity decreases to a limiting value of about 90 \AA . For an accuracy of 1 or 2 percent in predicting this periodicity it is necessary to include 40 or 50 beams.

These particular cases of high symmetry do not necessarily give a good indication for the general case, but experience with calculations for other orientations, such as that of Fig. 10.6, corresponding to case (c) of Fig. 10.3, shows that the requirements for accuracy of intensities are fairly well indicated by Fig. 10.5.

It is to be expected that for comparable accuracy the number of beams to be included should vary roughly as the square of the average unit cell dimension and in proportion to the atomic number. However such indications are necessarily very approximate. The only way to test whether, in any particular case, the number of beams taken is sufficient is to increase the number of beams by a considerable factor and ensure that such an increase makes no appreciable difference to the intensities of the beams of interest.

Some foreknowledge of the form of the diffraction pattern is, of course, desirable as a guide to the selection of the diffracted beams to include in the calculations. For example, for a crystal tilted away from a principle orientation the "laue circle" of spots given by the intersection of the Ewald sphere with the principle reciprocal lattice plane will be strong and should be included. Also care should be taken for complicated structures to include the often widely-spaced, strong reflections coming from a prominent sublattice of the structure.

As mentioned previously, Sturkey [1962] made use of the scattering matrix formulation, equation (27), to calculate diffracted beam intensities. The scattering matrix for a small crystal thickness H , raised to the n th power, gives the scattering from a crystal of thickness nH . In raising a matrix to higher powers errors accumulate rapidly and the requirements for accuracy in the initial thin-crystal matrix are increased. Considerable care is needed to ensure that the number of beams included and the accuracy of the calculations for the initial thin-crystal calculation are sufficient.

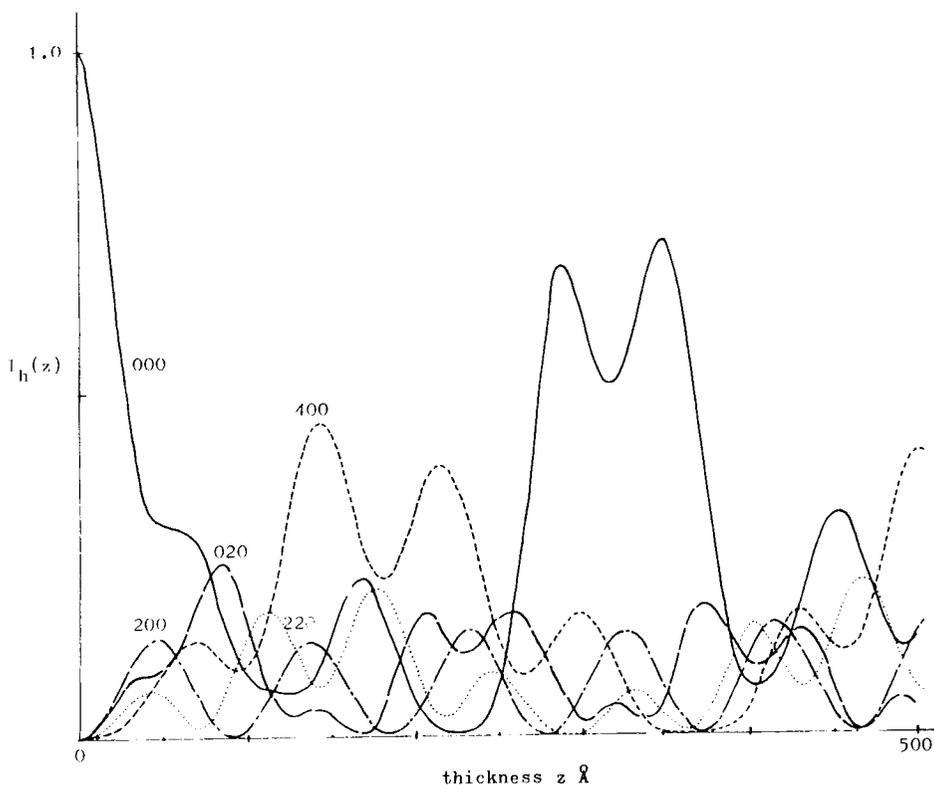


Fig. 10.6. Calculated intensities of hko reflections as a function of thickness for the incident beam tilted slightly away from the c -axis of the F.C.C. copper-gold alloy of Figs. 10.3 and 10.4 to give the Ewald sphere intersection indicated by (c) in Fig. 10.3.

Calculation based on the Howie-Whelan equations (35) involve the progressive integration of the amplitudes through the crystal.

10.6. Column approximation

The use of equations such as (35) to derive the form of electron microscope images or diffraction patterns from defects in crystals depends on the use of the "column approximation". This may be described by means of the diagram, Fig. 10.7. The amplitudes of the waves emerging from a crystal at a point P are assumed to be dependent only on the structure of a column of diameter D centered on P and extending through the crystal in the direction of energy propagation. Then if the contents of the column can be approximated by a function dependent on the z -coordinate only, the intensity at P will be the same as for a crystal of

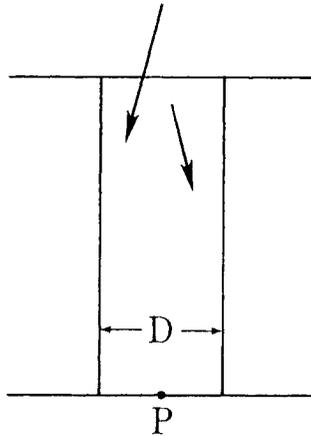


Fig. 10.7. Diagram illustrating assumptions of the "column approximation".

infinite lateral extent of structure having the same z -dependence, for which the intensity may be calculated by equations such as (35). The requirement is that the crystal structure should not vary appreciably across the width of the column. Then the intensity distribution across the exit face as seen in an electron microscope will give a representation of the changes of diffraction conditions with a resolution roughly equal to the column diameter.

The minimum diameter of the column may be estimated in a number of ways which give comparable results. For example in a near-two-beam case the energy is transferred from incident to diffracted beam or vice versa, in a distance $\xi_h/2$. Thus the lateral movement of the beam before its direction is changed is of the order of $\xi_h\theta_h$. For $\xi_h = 200 \text{ \AA}$ and $\theta_h = 10^{-2}$ the product is 2 \AA , but we should allow a factor of 2 or 3 greater than this to give the half width of the column as approximately 5 \AA .

Alternatively, if there is only weak scattering, the main contribution to the column width comes from the spread of the electron wave by Fresnel diffraction. As a measure of this spread we may take the condition

$$\exp\left\{\frac{ikx^2}{2R}\right\} = \exp\{i\pi\}$$

or

$$x = (R\lambda)^{1/2}.$$

Thus for $\lambda = 0.04 \text{ \AA}$ and a crystal thickness of 500 \AA we have that the spread of the wave due to Fresnel diffraction is indicated as $x = 4.5 \text{ \AA}$.

More complete and sophisticated discussions of the effective width of the column by Jouffrey and Taupin [1967] and Howie and Basinski [1968] give results of the same order of magnitude. Hence for electron microscopy with a resolution not better than 10 \AA and for perturbations of the crystal lattice which do not have strong variations over distances of 10 \AA the column approximation would not give rise to any serious error. The experimental observations under these conditions have been almost universally of a rough non-quantitative nature and within these limitations appear to give reasonable agreement with column approximation calculations.

A further application of the column approximation is to the intensities for a crystal for which the lateral variation is that of crystal thickness rather than crystal distortion. This is the basis for our simplified discussion of the origin of thickness fringes in electron micrographs in Chapter 9.

Problems

1. Show that the differential equations (10.32) are consistent with the two-beam form of the Bethe equations (8.10) without absorption and the solutions for a flat plate crystal (8.25) and (8.26). Similarly show that the n -beam form (10.33) is consistent with the dispersion equation (8.7).

2. Write down the form of the scattering matrix (10.29) appropriate for a very thin crystal in a first approximation. How is (10.30) then related to a kinematical approximation and to a phase-grating approximation? (See Sturkey [1962].)

This Page Intentionally Left Blank

Multi-slice approaches

11.1. Propagation of electrons in crystals

11.1.1. Transmission through thin slices

In Chapter 3 we considered transmission of an incident wave through two-dimensional phase- and amplitude-objects and the propagation between such objects, described in terms of Fresnel diffraction concepts. The effect on the incident wave of a two-dimensional object was represented by multiplication by a transmission function, $q_n(x,y)$ and transmission through space by convolution with a propagation function $p_n(x,y)$. The extension to the consideration of three-dimensional objects was indicated in equation (3.31) for real space and equation (3.32) for reciprocal space.

The relevance of this approach for a consideration of the scattering of electrons from crystals was indicated by the accumulation of evidence that for very thin single crystals there may be hundreds of simultaneous diffracted beams and the diffraction conditions are more closely related to those for a two dimensional grating than to those for an infinite periodic lattice which is the starting point of Bethe theory.

In the formulation of n -beam diffraction theory by Cowley and Moodie [1957] transmission of electrons through a sample is represented by transmission through a set of N two dimensional phase- and amplitude-objects separated by distances Δz . The total phase change and amplitude change of the electron wave in a slice of the specimen of thickness Δz is considered to take place on one plane. Then the propagation of the wave from one such plane to the next is by Fresnel diffraction in vacuum. It has been shown (Moodie [1972]) that in the limiting case that the thickness of the slice Δz goes to zero and the number of slices N goes to infinity in such a way that $N\Delta z = H$, where H is the specimen thickness, this form of description becomes a rigorous representation of the scattering process, completely consistent with the more conventional quantum mechanical descriptions. Several authors have provided derivations of the multislice formulation from established quantum mechanical bases such as that of Fineman (Jap and Glaeser [1978], Ishizuka and Uyeda [1977]).

An alternative formulation of the multislice approach is based on a specifically high-energy, forward scattering,

approximation to the Schrödinger wave equation as described in Section 11.6. This has been the basis for the practical real-space method of Van Dyck et al. [Van Dyck, 1983, 1985] for computing electron microscope image intensities and other methods recently explored. Initially, however, we follow the development of the "physical optics" approach of Cowley and Moodie which led to the multislice method for calculation of diffraction pattern and electron microscope image amplitudes and intensities which is currently in most common use.

The phase change of the electron wave transmitted through a thin slice depends on the value of the refractive index,

$$n = \left\{ 1 + \frac{\varphi(r)}{E_0} \left(1 + \frac{eE_0}{2m_0c^2} \right) \right\}^{1/2} \approx 1 + \frac{\varphi(r)}{2E}, \quad (1)$$

so that the phase difference with respect to a wave in vacuum for a slice thickness Δz at $z = z_n$ is

$$\left[\left(1 + \frac{\varphi}{2E} \right) - 1 \right] \Delta z \frac{2\pi}{\lambda} = \frac{\pi}{\lambda E} \Delta z \varphi(x, y, z_n). \quad (2)$$

We put $\pi/\lambda E$ equal to the interaction constant, σ , which varies with voltage as indicated by (10.2). We may also introduce an absorption function $\mu(r)$ depending on the circumstances of the experiment to be described.

Then the transmission function for the plane $z = z_n$ representing the slice may be written

$$q_n(x, y, z_n) = \exp\{-i\sigma\varphi(x, y, z_n)\Delta z - \mu(x, y, z_n)\Delta z\}. \quad (3)$$

The negative sign before the potential term is consistent with the use of the form $\exp\{i(\omega t - \mathbf{k} \cdot \mathbf{r})\}$ for a plane wave. The opposite sign was used by Cowley and Moodie [1957] who therefore unwittingly produced a theory of positron diffraction.

In the limit that the slice thickness goes to zero, we may write

$$q_n(x, y, z_n) = 1 - i\sigma\varphi(x, y, z_n)\Delta z - \mu(x, y, z_n)\Delta z, \quad (4)$$

and Fourier transforming in two dimensions gives

$$Q_n(u, v) = \delta(u, v) - i\sigma\Phi_n(u, v)\Delta z - M_n(u, v)\Delta z, \quad (5)$$

where

$$\Phi_n(u, v) = \int \Phi(u) \exp\{-2\pi i w z_n\} dw, \quad (6)$$

and

$$M_n(u, v) = \int M(u) \exp\{-2\pi i w z_n\} dw, \quad (7)$$

where $\Phi(u)$ and $M(u)$ are the Fourier transforms of $\varphi(r)$ and $\mu(r)$, respectively. For convenience we may use a "kinematical" structure amplitude

$$F(u) = \sigma\Phi(u) - iM(u) \quad (8)$$

so that

$$Q_n(u, v) = \delta(u, v) - iF_n(u, v)\Delta z. \quad (9)$$

The Fresnel diffraction of the wave on transmission through the distance Δz is given by convolution with the propagation function which, in the small angle approximation, is given by

$$p(x, y) = \frac{i}{R\lambda} \exp\{i\pi(x^2+y^2)/\lambda\Delta z\}, \quad (10)$$

or, in reciprocal space we multiply by

$$P(u, v) = \exp\{\pi i\lambda\Delta z(u^2+v^2)\}. \quad (11)$$

11.1.2. Three-dimensional objects

Transmission through a three-dimensional object is given by the limiting case, for N going to infinity, of equation (3.31) written in one dimension only;

$$\Psi(x) = q_N(x) \left[\dots [q_2(x) [q_1(x) [q_0(x) * p_0(x)] * p_1(x)] * p_2(x)] * \dots \right] * p_N(x), \quad (12)$$

where $q_N(x)$ multiplied by the N bracket represents the wave leaving the object and this wave is propagated by convolution with $p_N(x)$ to the plane of observation. If the radiation incident on the first slice of the object, which has transmission function $q_1(x)$, is a plane wave inclined at an angle α to the normal to the plane of the slices, the convolution $q_0 * p_0$ is replaced by

$$\exp\{2\pi i(\sin \alpha)x/\lambda\} \quad (13)$$

which multiplies $q_1(x)$.

It is convenient when dealing with diffraction by crystals to work with the reciprocal space expression given, in one dimension, by the limiting case of equation (3.32), namely,

$$\Psi(u) = [Q_N(u) * \dots [Q_2(u) * [Q_1(u) * Q_0(u) P_0(u)] P_1(u)] P_2(u) \dots] P_N(u). \quad (14)$$

For a plane wave incident, $Q_0(u)P_0(u)$ is replaced by $\delta(u + (\sin \alpha)/\lambda) \equiv \delta(u - u_0)$.

Both of the forms, (12) and (14), are appropriate to finding the wave function at or near the exit face of an object, and so are useful for finding the amplitude and intensity distributions in ideal electron microscope images. For this purpose the diffracted amplitudes from (14) are Fourier transformed.

The amplitudes of the diffracted beams from an object are given directly from (14) with the final phase factor $P_N(u)$ usually neglected, or may be obtained by Fourier transform of $\psi(x)$ from (12). Other forms for these expressions, involving some transmission functions in real space and some in reciprocal space may be generated for use in special cases by application of the relationships (3.33) to (3.37) (see Cowley and Moodie [1958]).

11.1.3. Diffraction by a crystal

In the case of diffraction by a crystal, the slice transmission functions $q_n(x, y)$ is periodic so that we may write

$$q_n(x, y) = \sum_h \sum_k Q_n(h, k) \exp\left\{-2\pi i\left(\frac{hx}{a} + \frac{ky}{b}\right)\right\} \quad (15)$$

or

$$Q_n(u, v) = \sum_h \sum_k Q_n(h, k) \delta\left(u - \frac{h}{a}, v - \frac{k}{b}\right). \quad (16)$$

From (5) and (9) we see that $Q_n(u, v)$ has a dominating delta function at the origin, plus, for all h, k (including $h, k = 0$),

$$Q_n(h, k) = -i\Delta z \cdot F_n(h, k), \quad (17)$$

and

$$F_n(h, k) = \sigma\Phi_n(h, k) - iM_n(h, k), \quad (18)$$

where

$$\Phi_n(h, k) = \sum_l \Phi_{hkl} \exp\left\{-2\pi il\frac{z_n}{c}\right\}, \quad (19)$$

$$M_n(h, k) = \sum_l M_{hkl} \exp\left\{-2\pi il\frac{z_n}{c}\right\}. \quad (20)$$

Measuring z from the entrance face of the crystal we have

$$z_n = (n-1)\Delta z \quad \text{and} \quad H = (N-1)\Delta z. \quad (21)$$

If now we return to the one-dimensional form of (14) for convenience, the bracket numbered 1, for an incident plane wave becomes $Q_1(u - u_0)$. The bracket 2 becomes

$$\begin{aligned}
 \left[\dots \right]_2 &= (-i\Delta z)^2 [F_2(u) * F_1(u - u_0) \exp\{\pi i \lambda \Delta z \cdot u^2\}] \\
 &= (-i\Delta z)^2 \sum_{h_1} \sum_{h_2} F_1(h_1) F_2(h_2) \\
 &\times \exp \left\{ \pi i \lambda \Delta z \left(u - \frac{h_2}{a} \right)^2 \right\} \delta \left(u - u_0 - \frac{h_1 + h_2}{a} \right). \quad (22)
 \end{aligned}$$

Here, for convenience, we have included the $h = 0$ term in the summation over h with a weight $i/\Delta z$ so that it has weight unity when multiplied by $(-i\Delta z)$. We see from (22) that for a beam diffracted twice corresponding to a reciprocal lattice point $h = h_1 + h_2$, the phase factor depends on the square of $\{u - (h_2/a)\}$, i.e. on the square of

$$u_0 + \frac{h_1 + h_2}{a} - \frac{h_2}{a} = \left(u_0 + \frac{h_1}{a} \right)$$

and so is not dependent on the diffraction indices in the same way as the total change in u as given by the δ -function. This, for the more general case of many slices, is the origin of much of the complication of the n -beam dynamical theory.

Similarly

$$\begin{aligned}
 [\dots]_3 &= (-i\Delta z)^3 \exp\{\pi i \lambda \Delta z_2 u^2\} \\
 &\times \sum_{h_1} \sum_{h_2} \sum_{h_3} F_1(h_1) F_2(h_2) F_3(h_3) \delta \left(u - u_0 - \frac{h_1 + h_2 + h_3}{a} \right) \\
 &\times \exp \left\{ \pi i \lambda \Delta z \left[\left(u_0 + \frac{h_1}{a} \right)^2 + \left(u_0 + \frac{h_1 + h_2}{a} \right)^2 \right] \right\}. \quad (23)
 \end{aligned}$$

Extending this to the case of N slices and making use of (21) we may write the final amplitude as

$$\begin{aligned}
 \Psi(u) &= (-i\Delta z)^N \exp\{\pi i \lambda H u^2\} \\
 &\times \sum_{h_1} \sum_{h_2} \dots \sum_{h_N} F_1(h_1) F_2(h_2) \dots F_N(h_N) \delta \left(u - \frac{h_1 + h_2 + \dots + h_N}{a} - u_0 \right) \\
 &\times \exp \left\{ 2\pi i \lambda \sum_{n=1}^N z_n \left[-\frac{u_0 h_N}{a} + \frac{h_N^2}{a^2} + \sum_{m=n+1}^N \frac{h_n h_m}{2a^2} \right] \right\}. \quad (24)
 \end{aligned}$$

The delta function specifies the direction of the h diffracted beam where $h = \sum_{n=1}^N h_n$. What we wish to find is the amplitude of the h beam for all scattering processes, i.e. for all combinations of $h_1, h_2 \dots h_N$ which lead to h .

The structure factors $F_n(h_n)$ or, more generally, $F_n(h_n, k_n)$, can all be expressed in terms of the one set of structure factors $F(h, k, l)$ for the crystal by use of the equations (18) to (20) which give

$$F_n(h, k) = \sum_l F(h, k, l) \exp\left\{2\pi i \frac{l z_n}{c}\right\}. \quad (25)$$

Hence each summation in (24) is replaced by summations over h_n, k_n and l_n and the exponential of (25) is added to complicate the exponentials of (24).

11.1.4. General expression in terms of excitation errors

In order to simplify the result we introduce the quantity

$$\zeta_n = -\frac{\lambda}{2} \left\{ \left(\frac{h^{(n)}}{a} \right)^2 + \left(\frac{k^{(n)}}{b} \right)^2 \right\} - \frac{l^{(n)}}{c} + \frac{\lambda}{2} \left(\frac{u_0 h^{(n)}}{a} + \frac{v_0 k^{(n)}}{b} \right). \quad (26)$$

This, in the small angle approximation, is exactly the excitation error for a reciprocal lattice point with coordinates

$$h^{(n)} = \sum_{r=1}^n h_r, \quad k^{(n)} = \sum_{r=1}^n k_r, \quad l^{(n)} = \sum_{r=1}^n l_r,$$

since the Ewald sphere is approximated by the paraboloid

$$w = -\frac{1}{2}\lambda(u^2 + v^2) + \frac{1}{2}\lambda(u_0 u + v_0 v). \quad (27)$$

The sign of ζ_n is chosen to conform with the convention that the excitation error should be positive for reciprocal lattice points inside the Ewald sphere, with the incident beam in the positive z direction. It is interesting to note that the Ewald sphere curvature, implied by the form (26) for the excitation error, is derived as a direct consequence of the form used for the propagation function and so represents the reciprocal space analog of Fresnel diffraction.

With the substitution (26) the general two-dimensional form of (24) becomes

$$\Psi(h, k) = (-i\Delta z)^N \exp\{\pi i \lambda H(u^2 + v^2)\}$$

$$\times \sum_{h_1} \sum_{k_1} \sum_{l_1} \sum_{h_2} \sum_{k_2} \sum_{l_2} \dots \sum_{h_{N-1}} \sum_{k_{N-1}} \sum_{l_{N-1}} \sum_l F(h_1, k_1, l_1) F(h_2, k_2, l_2)$$

$$\dots F\left(h - \sum_{n=1}^{N-1} h_n, k - \sum_{n=1}^{N-1} k_n, l - \sum_{n=1}^{N-1} l_n\right) \exp\left\{-2\pi i \left(H\zeta - \Delta z \sum_{n=1}^{N-1} \zeta_n\right)\right\}, \quad (28)$$

where ζ is the excitation error for the h, k, l reflection when $h = \sum_{n=1}^{N-1} h_n$, and so on. We note that while the structure amplitudes $F(h_n, k_n, l_n)$ refer to the n th scattering process only, the excitation errors ζ_n are cumulative and depend on the history of scatterings by all slices from 1 to n . This is the factor which makes it very difficult to simplify (28) further.

11.2. Multiple-scattering series

11.2.1. Zero-order scattering

If we now take the limit of N tending to infinity and Δz tending to zero, the number of summations in (28) becomes infinite. To proceed further we must systematize the choice of terms of the series. One way to do that is to note that, for each set of these summations over h_n, k_n, l_n , the zero term is very much larger than the rest since, from (17) and (18),

$$(-i\Delta z)F(0, 0, 0) = 1 - i\sigma\Delta z\Phi(0, 0, 0) - \Delta zM(0, 0, 0),$$

$$(-i\Delta z)F(h, k, l) = -i\sigma\Delta z\Phi(h, k, l) - \Delta zM(h, k, l) \quad (29)$$

Hence the largest single contribution to (28) comes where the $h, k, l = 0$ value is chosen from each set of summations. The next largest contributions correspond to the case when the zero term is chosen from all summations except one. These contributions give the singly-scattered radiation. By allowing successively two, three and more slices to give non-zero reflections we then generate the doubly, triply and multiply-scattered contributions and so obtain the equivalent of a Born series.

The zero order term of this series is

$$\Psi_0(h, k) = \lim_{N \rightarrow \infty} \{(-i\Delta z)^N F^N(0, 0, 0)\}$$

$$= \lim_{N \rightarrow \infty} [1 - i\sigma\Delta z\Phi(0, 0, 0) - \Delta zM(0, 0, 0)]^N$$

$$= \exp\{-i\sigma H\Phi(0, 0, 0) - HM(0, 0, 0)\}. \quad (30)$$

This corresponds to the transmission of the incident beam with a phase change appropriate to the mean inner potential and attenuation by the mean absorption coefficient.

11.2.2. Single scattering - kinematical approximation

The first-order term takes account of single scattering only. We may consider that the incident beam is transmitted without scattering through all slices but the m th. Then the summations of (28) are reduced by putting all $h_n, k_n, l_n = 0$ except that $h_m, k_m, l_m = h, k, l$. Then, from (26)

$$x = \begin{cases} 0 & \text{for } n < m, \\ \zeta & \text{for } n \geq m. \end{cases} \quad (31)$$

Then summing for all values of m , (28) gives us

$$\Psi_1(h, k) = (-i\Delta z) F^{N-1}(0, 0, 0) \sum_m \sum_l F(h, k, l) \exp\{-2\pi i z_m \zeta\}. \quad (32)$$

In the limit that N tends to infinity, the summation over m is replaced by an integral over z from 0 to H giving

$$\begin{aligned} \Psi_1(h, k) &= \exp\{-i\sigma H \Phi(0, 0, 0) - HM(0, 0, 0)\} \\ &\times (-i) \sum_l F(h, k, l) \exp\{-\pi i H \zeta\} [(\sin \pi H \zeta) / \pi \zeta]. \end{aligned} \quad (33)$$

Assuming the excitation error to be large for all l except $l = 0$ and that $\Phi(h, k, 0)$ and $M(h, k, 0)$ are real, the intensity of the single-scattered diffracted beam is

$$\begin{aligned} |\Psi_1(h, k)|^2 &= \exp\{-2HM(0, 0, 0)\} [\sigma^2 \Phi^2(h, k, 0) + M^2(h, k, 0)] \frac{\sin^2(\pi H \zeta)}{(\pi \zeta)^2}, \end{aligned} \quad (34)$$

and this is exactly the expression for kinematical scattering from a parallel-side plate-shaped crystal of thickness, H .

11.2.3. Multiple scattering

In much the same way it is possible to find the contribution to the amplitude when scattering, other than in the central beam, may take place in two, three or any number of slices. For example, the triple scattering term is found by considering that all $h_n, k_n, l_n = 0$ except that $h_m, k_m, l_m = h_1, k_1, l_1$, $h_p, k_p, l_p = h_2, k_2, l_2$ and $h_r, k_r, l_r = h - h_1 - h_2, k - k_1 - k_2, l - l_1 - l_2$.

Then

$$\zeta_n = \begin{cases} 0 & \text{for } 0 < n < m \\ \zeta_1 & \text{for } m \leq n < p \\ \zeta_2 & \text{for } p \leq n < r \\ \zeta & \text{for } r \leq n \leq N. \end{cases} \quad (35)$$

The term in the final exponent in (28) becomes

$$H\zeta - \Delta z \sum_{n=1}^{N-1} \zeta_n = z_1 \zeta_1 + z_2 (\zeta_2 - \zeta_1) + z_3 (\zeta - \zeta_2). \quad (36)$$

In the limiting case, the amplitude expression is integrated over z_3 from z_2 to H , over z_2 from z_1 to H and over z_1 from 0 to H , giving with some rearrangement of terms,

$$\begin{aligned} \Psi_3(h, k) &= \exp\{-i\sigma H\Phi(0, 0, 0) - HM(0, 0, 0)\} \\ &\times \sum_{h_1} \sum_{k_1} \sum_{l_1} \sum_{h_2} \sum_{k_2} \sum_{l_2} \sum_l F(h_1, k_1, l_1) F(h_2, k_2, l_2) F(h-h_1-h_2, k-k_1-k_2, \\ & l-l_1-l_2) \\ &\times \frac{\exp\{-2\pi i\zeta H\}}{(2\pi)^3} \left\{ \frac{\exp\{2\pi i\zeta_1 H\} - 1}{\zeta_1(\zeta_1 - \zeta_2)(\zeta_1 - \zeta)} + \frac{\exp\{2\pi i\zeta_2 H\} - 1}{\zeta_2(\zeta_2 - \zeta_1)(\zeta_2 - \zeta)} + \frac{\exp\{2\pi i\zeta H\} - 1}{\zeta(\zeta - \zeta_1)(\zeta - \zeta_2)} \right\}. \end{aligned} \quad (37)$$

Generalizing the result for n -times scattering gives

$$\begin{aligned} \Psi_n(h, k) &= \exp\{-i\sigma H\Phi(0, 0, 0) - HM(0, 0, 0)\} \\ &\times \sum_l \sum_{h_1} \sum_{k_1} \sum_{l_1} \dots \sum_{h_{n-1}} \sum_{k_{n-1}} \sum_{l_{n-1}} F(h_1, k_1, l_1) F(h_2, k_2, l_2) \\ &\dots F\left(h - \sum_{r=1}^{n-1} h_r, k - \sum_{r=1}^{n-1} k_r, l - \sum_{r=1}^{n-1} l_r\right) \frac{\exp\{-2\pi i\zeta H\}}{(2\pi)^n} \\ &\times \left[\frac{\exp\{2\pi i\zeta_1 H\} - 1}{\zeta_1(\zeta_1 - \zeta_2) \dots (\zeta_1 - \zeta_{n-1})(\zeta_1 - \zeta)} \right. \\ &+ \sum_{m=2}^{n-1} \frac{\exp\{2\pi i\zeta_m H\} - 1}{\zeta_m(\zeta_m - \zeta_1) \dots (\zeta_m - \zeta_{m-1})(\zeta_m - \zeta_{m+1}) \dots (\zeta_m - \zeta)} \\ &\left. + \frac{\exp\{2\pi i\zeta H\} - 1}{\zeta(\zeta - \zeta_1)(\zeta - \zeta_2) \dots (\zeta - \zeta_{n-1})} \right]. \end{aligned} \quad (38)$$

The total amplitude of the h, k , diffracted beam is then

$$\Psi(h, k) = \sum_{n=0}^{\infty} \Psi_n(h, k). \quad (39)$$

which is the Born Series type of expression for single, double and multiple scattering from planes in the crystal. The same result was obtained by different methods by Fujimoto [1959] and by Fujiwara [1959] with a little more generality in that the z components of the wave vectors for the individual reflections were included in the denominators to take account of the slight differences in direction of the corresponding waves.

As in the case of the more conventional Born Series, for scattering by volume elements, the convergence is slow when the scattering power or crystal thickness is too great for the kinematical single-scattering approximation to be used. Hence we seek other forms, more appropriate for situations in which the kinematical approximation can not be used.

11.3. General double-summation solution

11.3.1. General series solution

It was shown by Moodie (Cowley and Moodie [1962]) that the expression (39) could be rewritten in the form

$$\Psi(h) = \sum_{n=1}^{\infty} E_n(h) Z_n(h), \quad (40)$$

where we have used h to signify the pairs of indices h, k . The operator $E_n(h)$ represents all the operations of summation over indices and multiplication by F quantities in the expression (38). This operates on the functions $Z_n(h)$ which are purely geometric, depending on the wavelength, unit cell dimensions and crystal size but independent of scattering amplitudes. Expanding the exponentials of (38) and manipulating the terms gives

$$Z_n(h) = \sum_{r=0}^{\infty} \frac{(-2\pi i H)^{n+r}}{(n+r)!} h_r(\zeta, \zeta_1, \zeta_2 \dots \zeta_{n-1}), \quad (41)$$

where the functions $h_r(\dots)$ are the complete homogeneous symmetric polynomials of degree r i.e.

$$\begin{aligned} h_1(\zeta, \zeta_1, \zeta_2 \dots \zeta_{n-1}) &= \zeta + \zeta_1 + \zeta_2 + \dots + \zeta_{n-1}, \\ h_2(\zeta, \zeta_1, \zeta_2 \dots \zeta_{n-1}) &= \zeta^2 + \zeta_1^2 + \zeta_2^2 + \dots + \zeta \zeta_1 + \zeta \zeta_2 + \dots + \zeta_1 \zeta_2 + \dots \end{aligned} \quad (42)$$

Each polynomial contains, once only, each possible combination of variables, irrespective of ordering, which is of the prescribed power. The excitation errors, ξ_m , in these expressions are defined as in (26), and so depend on all indices for all scattering processes up to $n = m$.

Writing out (41) in detail then gives

$$\begin{aligned} Z_1(h) &= (-2\pi iH) + \frac{(-2\pi iH)^2}{2!} \zeta + \frac{(-2\pi iH)^3}{3!} \zeta^2 + \dots \\ Z_2(h) &= \frac{(-2\pi iH)^2}{2!} + \frac{(-2\pi iH)^3}{3!} (\zeta + \zeta_1) + \frac{(-2\pi iH)^4}{4!} (\zeta^2 + \zeta^2_1 + \zeta_1 \zeta) + \dots \\ Z_3(h) &= \frac{(-2\pi iH)^3}{3!} + \frac{(-2\pi iH)^4}{4!} (\zeta + \zeta_1 + \zeta_2) \\ &+ \frac{(-2\pi iH)^5}{5!} (\zeta^2 + \zeta^2_1 + \zeta^2_2 + \zeta \zeta_1 + \zeta \zeta_2 + \zeta_1 \zeta_2) + \dots \quad \text{etc.} \end{aligned} \quad (43)$$

Then adding the appropriate operators $E_n(h)$ we obtain a two-dimensional array of terms with n increasing vertically and r horizontally. This two-dimensional array can then be summed in a number of ways. For example, we may sum the rows horizontally. This gives for the first row

$$\begin{aligned} \Psi_1(h) &= \sum_l F(h, k, l) \left[\sum_r \frac{(-2\pi iH\zeta)^r}{r!} - 1 \right] \frac{1}{\pi\zeta} \\ &= (-i) \sum_l F(h, k, l) \exp\{-\pi iH\zeta\} \left(\frac{\sin \pi H\zeta}{\pi\zeta} \right), \end{aligned}$$

which is identical with (33) apart from the zero term. Similarly summing each row horizontally gives its the multiple-scattering series $\Psi_n(h)$ from which we started.

Fujimoto [1959] derived an expansion in terms of increasing powers of the crystal thickness H . This can be reproduced by selecting terms from (43) and summing along diagonal lines.

11.3.2. Phase grating approximation

If the terms of (43) are selected along the vertical lines, with the appropriate E operators, we obtain series for various r values. The first is

$$\begin{aligned}
 T_1(h) &= \sum_{n=1}^{\infty} \frac{\{-2\pi i H\}^n}{(2\pi)^n n!} \\
 &\times \sum_{h_1} \sum_{k_1} \sum_{l_1} \sum_{h_2} \sum_{k_2} \sum_{l_2} \dots \sum_{h_{n-1}} \sum_{k_{n-1}} \sum_{l_{n-1}} F(h_1, k_1, l_1) F(h_2, k_2, l_2) \\
 &\dots F\left(h - \sum_{p=1}^{n-1} h_p, k - \sum_{p=1}^{n-1} k_p, l - \sum_{p=1}^{n-1} l_p\right). \quad (44)
 \end{aligned}$$

The multiple summation represents the amplitude of the (h, k) term of the $(n - 1)$ fold self-convolution of

$$F(u) = \sum_h \sum_k \sum_l F(h, k, l) \delta\left(u - \frac{h}{a}, v - \frac{k}{b}, w - \frac{l}{c}\right).$$

Hence the set of diffracted beam amplitudes $T_1(h, k, 0)$ represented by the function

$$T_1(u, v) = \sum_h \sum_k \sum_l T_1(h, k, l) \delta\left(u - \frac{h}{a}, v - \frac{k}{b}\right)$$

is the Fourier transform of

$$\begin{aligned}
 \tau_1(x, y) &= \sum_{n=1}^{\infty} \frac{\{-2\pi i H\}^n}{(2\pi)^n n!} \{\sigma\varphi(x, y) - \mu(x, y)\}^n \\
 &= \exp\{-i\sigma H\varphi(x, y) - H\mu(x, y)\}, \quad (45)
 \end{aligned}$$

where $\varphi(x, y)$ and $\mu(x, y)$ are the values per unit thickness of the projections in the z direction of the potential and absorption functions $\varphi(\mathbf{r})$ and $\mu(\mathbf{r})$.

Thus the expression (44) given by summing the first vertical column from (43) gives the amplitudes of the diffracted beams when all excitation errors are assumed to be zero, as if the Ewald sphere were planar or the wavelength were zero. Thus it gives a high-energy approximation.

It is also the "phase-grating" approximation, given by assuming that the total scattering and absorption power of the crystal is concentrated on a single plane. More commonly the functions $\varphi(x, y)$ and $\mu(x, y)$ are taken to be the projections in the z direction of the potential distribution $\varphi(x, y, z)$ and the absorption function $\mu(x, y, z)$. The crystal is then replaced by a two-dimensional distribution having transmission function

$$q(x, y) = \exp\{-i\sigma\varphi(x, y) - \mu(x, y)\}. \quad (46)$$

This phase-grating approximation can be considered as the first term of a series, which has been referred to as the "phase-grating series", for which the successive terms involve the successive vertical lines of terms in (43) and may be expected to introduce the three-dimensional diffraction effects with increasing accuracy. However for terms of this series beyond the first, no convenient form of simplification has been found. In fact it is not easy to see that the higher terms can be very useful since they contain successively higher positive powers of the excitation errors which apparently emphasize the less important reflections; those for which the excitation errors are large.

Hence the general expression (40) forms a basis for a variety of series expansions each of which provides a relatively simple approximation to the general result valid in particular circumstances. The kinematical approximation is valid in the limits of the small structure factors for a given thickness. The phase-grating approximation is appropriate in the zero wavelength limit or in the limit of zero thickness.

The other simple approximation of proven usefulness, the two-beam approximation, is generated from the general expression by assuming that only two beams can exist; or that $\zeta_0 = 0$, ζ_h is small and ζ for any other reflection is very large. On this basis Cowley and Moodie [1957] reduced the expressions such as (38) to a set of terms which summed to give exactly the same amplitudes for the incident and diffracted beams as are given by Bethe theory.

11.4. Computing methods

11.4.1. "Slice method" calculations

On the basis of this "physical optics" approach to the diffraction problem, Goodman and Moodie [1965, 1974] (see also Self and O'Keefe [1988]) developed a computing technique which enables the intensities of diffraction patterns or electron microscope images to be calculated with arbitrary accuracy by integrating the effects on a wave transmitted through a crystal. It is, of course, not feasible to follow the analytical process of considering the limiting case of an infinite number of slices of zero thickness. Instead a finite number of slices is used and a rigid testing routine is established to ensure that the slice thickness is small enough to give no appreciable errors.

For a slice of finite thickness the approximation of equation (4), the single scattering approximation, can not be used since we have seen previously that multiple scattering may be important even for a single heavy atom. Then the transmission function for a slice must be written in terms of

the phase-grating approximation as

$$q_n(x, y) = \exp\{-i\sigma\varphi_n(x, y) - \mu_n(x, y)\}, \quad (47)$$

$$\begin{aligned} Q_n(u, v) &= \mathfrak{I}[q_n(x, y)] \\ &= \mathfrak{I}[\exp\{-\mu_n(x, y)\} \cos \sigma\varphi_n(x, y)] \\ &\quad - i\mathfrak{I}[\exp\{-\mu_n(x, y)\} \sin \sigma\varphi_n(x, y)]. \end{aligned} \quad (48)$$

The propagation function used for transfer of the wave from one plane to another, $p_n(x, y)$, can be written in the small angle approximation (10) for most purposes without appreciable error. Usually for periodic objects the reciprocal space form is preferable.

The total effect of a crystal on an incident wave is then given by the repetition ($N-1$) times of the unit process which converts a wave leaving the ($n-1$)th slice to the wave leaving the n slice, namely

$$\Psi_n(u, v) = \{\Psi_{n-1}(u, v) P_{n-1}(u, v)\} * Q_n(u, v), \quad (49)$$

or, for a periodic object,

$$\Psi_n(h, k) = \sum_{h'} \sum_{k'} \Psi_{n-1}(h', k') P_{n-1}(h', k') Q_n(h-h', k-k'). \quad (50)$$

The process of multiplication and addition of a set of (complex) numbers represented by (50) is rapidly performed on a digital computer. Then starting from a known incident wave $\Psi_0(u, v)$, repetition of this process gives the diffracted amplitudes after transmission through any number of crystal slices. The real space amplitude $\Psi_N(x, y)$ at the exit surface is obtained by summing the Fourier series with $\Psi_N(h, k)$ as coefficients and from this the in-focus microscope image is obtained. The out-of-focus images are given by multiplying the $\Psi_N(h, k)$ by the appropriate propagation function transform values $P_R(h, k)$.

It is convenient for the purposes of calculation if the $Q_n(h, k)$ and $P_n(h, k)$ values are the same for each slice. This is possible for a perfect crystal if the unit cell dimensions in the beam direction are sufficiently small, as is usually the case for relatively simple crystal structures. However it is not necessary for this to be the case. If the periodicity in the beam direction is large, the unit cell contents may be divided into a sequence of slices which is then repeated regularly in the calculations. Short-range or long-range changes in the periodicity or structure of the crystal which do not involve changes in the lateral periodicities may be simulated by varying the $Q_n(h, k)$ values, or the inter-slice

distances occurring in $P_n(h, k)$, for individual slices or for any number of slices. Thus the effects of a wide range of perturbations of the crystal structure may be included without any essential modification of the computer program. In conjunction with the use of a column approximation, the program may be used to calculate the effects on diffraction patterns or microscope images of localized perturbations of the crystal lattice (for example, dislocation images).

Within each column the lateral displacements of the atoms are assumed constant and represented by multiplying $Q_n(h, k)$ by a phase factor. Displacements in the beam direction could be included by varying the $P_n(h, k)$ values. Perturbations of a crystal lattice for which a column approximation is not adequate may be treated by special methods to be described later.

11.4.2. Steps in a computation

In order to clarify the nature of the calculations, we now set out the steps involved for the simplest case of the incident beam parallel to the c -axis of a centro-symmetric orthorhombic crystal, having no absorption. The c -axis dimension is taken as the slice thickness. The program is set up as follows:

- A. *Calculation of phase-grating amplitudes for slices.*
 - (a) The structure amplitudes $\Phi(h, k, 0)$ are calculated from the atomic scattering factors and atom positions.
 - (b) The unit cell projection $\varphi(x, y)$ is obtained by summing the Fourier series with $\Phi(h, k, 0)$ as coefficients.
 - (c) The real and imaginary parts of $\exp\{-i\sigma\varphi(x, y)\}$ are calculated.
 - (d) Fourier transform of $\cos\{\sigma\varphi(x, y)\}$ and $\sin\{\sigma\varphi(x, y)\}$ gives the real and imaginary parts of $Q(h, k)$.
- B. *The propagation function.*
 - (a) The excitation errors, ζ_h , are found for the required range of indices h, k . For this the paraboloidal approximation to the Ewald sphere is usually sufficient.
 - (b) The phase factors $P(h, k) = \exp\{2\pi i c \zeta(h, k)\}$ are calculated.
- C. *Iterative calculation.*
 - (a) For the first slice the amplitudes for the wave leaving the slice are $Q(h, k)$.
 - (b) For the second slice, $\Psi_2(h, k)$ values are calculated by performing the summation (50).
 - (c) The summation (50) is repeated for each successive slice.
- D. *Output.*
 - (a) The diffracted beam intensities as a function of crystal thickness are given by calculating $\Psi_n \Psi_n^*$ after each slice.

(b) In-focus images are found by summing the Fourier series to give $\Psi_n(x,y)$, then calculating $\Psi_n\Psi_n^*$.

(c) For out-of-focus and aberrated images: the effect of defocus is given by multiplying the amplitudes $\Psi_n(h,k)$ by $P_R(h,k) = \exp\{2\pi i R\zeta(h,k)\}$ before summing the Fourier series of (b). The effect of lens aberrations is simulated by adding a phase shift depending on higher orders of h/a and k/b .

In practice, the iteration involving the summation (50) is the most time-consuming part of the procedure. A considerable saving of computing time can be made by application of the "fast Fourier transform" (FFT) algorithm of Cooley and Tukey [1965] as described, for example, by Self and O'Keefe [1988]. The convolution operation of (50) is replaced by performing a Fourier transform, a multiplication and an inverse Fourier transform.

The addition of an absorption term makes the amplitudes $\Phi(h,k,0)$ complex, but since the remainder of the computation is made with complex numbers in any case, this introduces very little further complication. If the incident beam is not exactly parallel to a crystal axis a tilt with components α_x , α_y may be included by adding a translation by amounts $c \tan \alpha_x$ and $c \tan \alpha_y$ to the propagation function which then becomes

$$P_\alpha(h,k) = \exp\left\{\pi i \lambda c \left(\frac{h^2}{a^2} + \frac{k^2}{b^2}\right)\right\} \exp\left\{-2\pi i c \left(\frac{h}{a} \tan \alpha_x + \frac{k}{b} \tan \alpha_y\right)\right\}. \quad (51)$$

Alternatively it may be convenient to assume that the incident beam direction is unchanged but that successive slices of crystal are translated to give the equivalent change in diffraction conditions, so that

$$Q_n(h,k) = Q_1(h,k) \exp\left\{2\pi i \left[\frac{hc}{a}(n-1) \tan \alpha_x + \frac{kc}{b}(n-1) \tan \alpha_y\right]\right\}. \quad (52)$$

More exactly, the repeat distances a and b should be modified by factors $\cos \alpha_x$ and $\cos \alpha_y$ and the distance c should be slightly increased. These tilted beam approximations are suitable only for tilts of a few degrees. For large tilts different projections of the structure should be considered.

The paraboloidal approximation to the Ewald sphere implied by (51) is convenient but not essential. The true spherical form of the Ewald sphere was included in a rigorous formulation for the treatment of crystal tilts and of the effects of crystal boundaries inclined to the incident beam by Ishizuka [1982].

11.4.3. Possible errors

One of the most important approximations in this computing method is that of taking a finite slice thickness.

To test whether any assumed slice thickness is small enough, the calculation is repeated for smaller slice thicknesses and the results compared. In this way it has been found that for electrons of energy of the order of 100 keV and for moderately simple structures with moderately heavy atoms, no appreciable error results for slice thicknesses up to about 4 Å. For slice thicknesses of 8 to 10 Å errors of several percent are introduced, especially for weak beams. Slice thicknesses of 12 or 15 Å give serious errors so that only the strongest beams show even qualitatively correct behavior.

A further important possible source of error is the failure to take a sufficient number of diffracted beams into account. For matrix-method calculations we saw that the exclusion of significant beams distorts the intensity variations of those included since a false model is imposed on the crystal structure. For the slice-method calculations however, omission of some significant beams means that the interaction of these beams with other beams is excluded. Energy is scattered into these beams but can not be scattered back, and so is lost to the system.

11.4.4. Consistency tests

The test imposed to determine whether sufficient beams are included is then to sum the intensities of all diffracted beams considered at each stage of the calculations. In the absence of absorption this sum should be constant and equal to the incident beam intensity with an accuracy of better than one part in 10^5 or more. As in the matrix calculations, it is found that for a two-dimensional diffraction pattern from a simple face-centered cubic metal, for example, at least 50 and preferably more beams should be included, and the number should be increased for substances having larger unit cells or for higher electron accelerating voltages.

The requirement of conservation of energy is part of a more general consistency requirement which should be met for each phase grating slice. In the absence of absorption

$$|q(x, y)|^2 = |\exp\{i\sigma\phi(x, y)\}|^2 = 1. \quad (53)$$

In reciprocal space the corresponding relationship is

$$\sum_{h'} \sum_{k'} Q(h', k') Q(h+h', k+k') = \begin{cases} 1 & \text{if } h, k = 0 \\ 0 & \text{otherwise.} \end{cases} \quad (54)$$

Sets of $Q(h, k)$ values may be considered satisfactory if this relationship is satisfied with an accuracy of, typically, one part in a million. This is necessary in order that the errors will not accumulate to a significant extent in calculations involving the repeated use of these values.

11.4.5. *Computing times*

The rapid advance in the capabilities of computers has meant that the multi-slice calculations involving many thousands of "beams", thought to be impracticable only ten or fifteen years ago, can now be carried out using even small, desk-top or personal, computers in reasonable times. The progress made in computing techniques can favour either the multi-slice methods or the matrix techniques mentioned in the previous Chapter so that one or the other may be more convenient for particular purposes at various points in time.

A comparison of the multi-slice and matrix methods of calculation was made, for example, by Self et al. [1983]. For a computation involving N beams, the computation time for the matrix methods is proportional to N^2 , whereas for the multi-slice method using the FFT the computing time is proportional to $N \log_2 N$. The multi-slice method is much faster when N exceeds 16. Another factor of importance is the size of the computer memory that is required. The computer space required increases with N^2 for the matrix method but with N for the multi-slice. For a maximum available storage space of 0.5 million bytes (128K words), the multi-slice can be run for an array size of 16384 elements, whereas for the matrix method only 150 beams can be used.

In general, for perfect crystals having relatively small unit cells, such as metals or semiconductors, or where a column approximation can be made for the imaging of defects in such materials, the matrix methods may well be preferable and are often preferred in that they give a more intuitively satisfying account of the diffraction process. For materials with large unit cells and especially for the diffraction or high-resolution imaging of defects when the method of periodic continuation is used, as outlined in the next Section, the multi-slice method has important advantages and may be readily applied when matrix methods are not feasible.

It has been shown by Qin and Urban [1990] that for calculations of diffraction patterns involving higher-order Laue zones, special care must be taken with the multi-slice method in order to avoid errors due to an improper representation of the structure in the incident beam direction. For this reason the use of matrix methods may be preferable for such computations.

11.5. **Intensities from non-periodic objects**

Calculations of dynamical scattering from non-periodic objects, such as defects in crystals or small particles or molecules, have nearly all been made by use of the column approximation as discussed in Chapter 10. For each column of the sample, the calculation is made using one of the methods described in the last chapter or the slice method, which

allow the modifications of the structure or displacements of unit cells to be included.

The principle limitation of this technique is that it is concerned only with the modifications of the amplitudes of the set of diffracted beams, h , which correspond to the reciprocal lattice points for the perfect structure. As we have seen in Chapter 7, diffraction effects from crystal imperfections, as from general non-periodic objects, are not confined to this discrete set of beams. Much of the information on the nature of the defect or on the atomic configurations in a non-crystalline assembly of atoms, is contained in the continuous distribution of back-ground scattering in the diffraction pattern. This kinematical result will be equally relevant for the phase-grating scattering from each of the thin slices considered in the dynamical scattering formulation. Hence the diffuse scattering must be taken into account for any realistic consideration of the diffraction effects or images for any but very special kinds of deviations from perfect crystal periodicity.

Since it is not feasible to deal with continuous scattering functions in computer calculations, the distributions $Q_n(u, v)$ corresponding to non-periodic transmission functions $q_n(x, y)$ must be sampled at closely spaced points and so are replaced by the sets of values

$$\sum_H \sum_K Q_n(H, K) \delta\left(u - \frac{H}{A}, v - \frac{K}{B}\right), \quad (55)$$

where the intervals $1/A$ and $1/B$ are made small enough to allow a reasonable representation of the fluctuations of interest in $Q_n(u, v)$. The equivalent operation in real space is to consider that the region of interest in $q_n(x, y)$ is reproduced at regular intervals A, B thus

$$q_n(x, y) * \sum_m \sum_p \delta(x - mA, y - pB). \quad (56)$$

Thus the crystal defect, or other arrangement of atoms of interest, is repeated regularly to form a superlattice having the diffraction pattern (55). Then the intensities of the diffraction pattern or image for this superlattice may be calculated using the perfect-crystal slice-method computer programs. The variation of scattering in the z direction is included by varying the contents of the successive slices.

Such a calculation may involve a very large number of diffracted beams corresponding to the very large unit cell size for the superlattice which must normally be assumed. The calculation may often be brought within reasonable bounds by use of special assumptions such as the consideration of a

one-dimensional case. For example, it may be considered that the "core" of a dislocation, the region in which atom displacements are large, is of the order of 20 Å in diameter. If it is desired to calculate the image of such a dislocation with a resolution of 3 Å the diffraction pattern must be sampled out to $u = 0.33 \text{ \AA}^{-1}$. The 20 Å dimension of the core implies detailed structure in the diffraction pattern on a scale of 0.05 \AA^{-1} . Then the diffuse scattering may be sampled at intervals of 0.01 \AA^{-1} . The dynamical diffraction calculation is then made with $2 \times 33.3 = 67$ beams in one dimension. That means that, in real space, we assume parallel dislocations to exist at intervals of 100 Å. In the imaging process each dislocation is imaged separately. There is no over-lapping of waves scattered from individual dislocations and no interference of their images unless the image is obtained far out of focus. Simple Fresnel diffraction theory gives this out-of-focus distance as approximately $A^2/2\lambda$ or, in this case, about 10 μm .

Calculations of this sort, involving the imposition of an artificial superlattice periodicity, have been carried out for dislocations by Fejes [1973] and for a completely non-periodic object chosen to simulate a rod-like protein molecule, negatively stained to increase the electron microscope contrast (Grinton and Cowley [1971]) with results which will be discussed in a later chapter. Calculations involving 12,000 beams have been used by O'Keefe and Iijima [1978] to simulate images of defects in oxide crystals, using a superlattice unit cell approximately 30 Å X 30 Å in size. Calculations of similar, or greater, size may be required when an incident beam of small diameter is used for convergent beam electron diffraction, as discussed in Chapters 9 and 15, or for scanning transmission electron microscopy, as described in Chapter 13. Then the incident beam wave function is non-periodic and a super-cell must be chosen such that the incident beam amplitude is essentially zero at its boundaries.

The application of the multi-slice method for the calculation of RHEED patterns or REM images, to be described in Chapter 19, may involve the assumption of a much larger size of artificial unit cell in at least one dimension, up to about 100 X 30 Å or more, with a correspondingly greater number of beams.

11.6. Real-space formulations

11.6.1. High-energy approximation

For high energy electrons the Schrödinger wave equation (1.5) may be written in the form

$$\nabla^2 \psi(\mathbf{r}) + 4\pi^2 k^2 \psi(\mathbf{r}) + V(\mathbf{r}) \cdot \psi(\mathbf{r}) = 0 \quad (57)$$

where $V(\mathbf{r})$ has been put equal to $(8\pi^2 m e / h^2) \phi(\mathbf{r})$. Since the potential term $V(\mathbf{r})$ is very much smaller than k^2 , the effect of the potential can be regarded as a perturbation on the plane incident wave, or

$$\psi(\mathbf{r}) = \exp\{2\pi i \mathbf{k} \cdot \mathbf{r}\} \cdot \phi(\mathbf{r}). \quad (58)$$

Substitution of (58) in (57) is readily shown to give the wave equation for $\phi(\mathbf{r})$ as

$$\nabla^2 \phi(\mathbf{r}) + 4\pi i \mathbf{k} \cdot \nabla \phi(\mathbf{r}) + V(\mathbf{r}) \cdot \phi(\mathbf{r}) = 0. \quad (59)$$

Because it is possible to make a small-angle approximation, it may be assumed that for the components of \mathbf{k} , $k_z \gg k_{x,y}$ and in the high-energy approximation, it may be assumed that $\phi(\mathbf{r})$ varies slowly with z so that $\partial^2 \phi / \partial z^2$ may be neglected. Then (60) may be rearranged to give

$$\frac{\partial \phi}{\partial z} = \left(\frac{i \nabla_{xy}^2}{4\pi k_z} - \frac{k_{xy}}{k_z} \nabla_{xy} \right) \phi(\mathbf{r}) + \frac{iV(\mathbf{r})}{4\pi k_z}$$

or,

$$\frac{\partial \phi}{\partial z} = \frac{i}{4\pi k_z} [\Delta + V(\mathbf{r})] \cdot \phi(\mathbf{r}) \quad (60)$$

where $\Delta = \nabla_{xy}^2 + 4\pi i k_{xy} \nabla_{xy}$

$$= \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + 4\pi i k_x \frac{\partial}{\partial x} + 4\pi i k_y \frac{\partial}{\partial y}$$

This provides the basis for a step-by-step integration of the wave function in the z -direction through a series of x - y slices. The V term represents the change in the wavefunction due to the potential and the Δ term represents the spread of the wave by Fresnel diffraction. Thus if the latter term is neglected, the equation becomes

$$\frac{\partial \phi(\mathbf{r})}{\partial z} = \frac{i}{4\pi k_z} V(\mathbf{r}) \cdot \phi(\mathbf{r})$$

for which the solution is

$$\phi(\mathbf{r}) = \exp\left\{ \frac{i}{4\pi k_z} \int_0^T V(\mathbf{r}) dz \right\}$$

$$\text{or, } \psi(\mathbf{r}) = \exp\{2\pi i \mathbf{k} \cdot \mathbf{r}\} \cdot \exp\{-i\sigma\phi(\mathbf{r})\}$$

which is just the phase object approximation of (4.14).

11.6.2. Useful approximations

From the equation (60) it is possible to derive the various formulations of the scattering equation which are of value for various purposes as a basis for the calculation of diffracted or image intensities (see Van Dyck [1983, 1985]). For example, for the case of transmission through a perfect crystal, the potential distribution is described by the Fourier series,

$$V(\mathbf{r}) = \sum_{\mathbf{g}} V_{\mathbf{g}} \exp\{2\pi i \mathbf{g} \cdot \mathbf{r}\}$$

and the reduced wave-function can be written

$$\varphi(\mathbf{r}) = \sum_{\mathbf{g}} \varphi_{\mathbf{g}}(z) \exp\{2\pi i \mathbf{g} \cdot \mathbf{r}\} \quad (62)$$

i.e., the Fourier coefficients vary with z . Substituting this in the wave-equation (60) gives

$$\frac{d\varphi_{\mathbf{g}}(z)}{dz} = 2\pi i \zeta_{\mathbf{g}} \varphi_{\mathbf{g}}(z) + \sum_{\mathbf{h}} V_{\mathbf{g}-\mathbf{h}} \varphi_{\mathbf{h}}(z) \quad (63)$$

where we have introduced the excitation error $\zeta_{\mathbf{g}} = [k^2 - (\mathbf{k} + \mathbf{g})^2] / 2k_{\mathbf{g}}$.

This is equivalent to the general Howie-Whelan equation, (10.35), for the case of a perfect crystal. It can be expressed as a matrix equation and follows from the scattering matrix formulation of Sturkey [1962] or the semi-reciprocal formulation of Tournarie [1960, 1961] (see Van Dyck [1985]). It can be used as the basis for a column-approximation for the discussion of the imaging of crystals seen in an axial orientation (Van Dyck et al., [1989]).

For the calculation of images of defects in crystals, these equations can be used with various levels of approximation. If only the first-order differentials of (60) are included, one obtains the equivalent of the Takagi-Taupin approach to the calculation of scattering by defects in crystals (see Anstis and Cockayne [1979]) in which, in effect, the Ewald sphere is approximated by a tangential plane around each reciprocal lattice point. When the second-order differentials are also included and only a limited number of 'beams', or Fourier coefficients, are used, one obtains the Howie and Basinski [1968] method which is useful for the calculation of electron microscope images of small-unit-cell materials (Anstis and Cockayne [1979]) and has been applied successfully by Anstis [1989, 1994] for the calculation of images of defects in reflection electron microscopy.

11.6.3. A real-space basis for computing

For the calculation of electron microscope images of non-periodic objects, such as defects or edges of crystals or isolated small objects, the methods of Section 11.5 can serve very well in many cases but it may sometimes be preferable to use a real-space formulation, as suggested by Van Dyck [1983]. This is based on an integral equation form of the modified wave equation (60) analogous to the Green's function form (1.17) which serves as the basis for the Born series expansion (1.22).

The integral equation giving the change in the wave-function for a small change, ε , in the z-coordinate can be written

$$\varphi(xy\varepsilon) = \varphi(xy0) + \frac{i}{4\pi k_z} \int_0^\varepsilon [\Delta + V(xyz)] \varphi(xyz) dz \quad (64)$$

As in the case of the Born series, this can be written as an iteration;

$$\begin{aligned} \varphi(xy\varepsilon) = & \left[1 + \frac{i}{4\pi k_z} \int_0^\varepsilon [\Delta + V(xyz)] dz \right. \\ & \left. + \left(\frac{i}{4\pi k_z} \right)^2 \int_0^\varepsilon [\Delta + V(xyz)] \int_0^z [\Delta + V(xyz')] dz' dz + \dots \right] \varphi(xy0) \end{aligned} \quad (65)$$

For convenience in the computing, this series is approximated by a product of exponentials which reproduce it to at least the second order:

$$\varphi(xy\varepsilon) = \exp\left\{\frac{i\lambda}{8\pi} V_P(1+\delta)\right\} \exp\left\{\frac{i\lambda}{4\pi} \Delta\varepsilon\right\} \exp\left\{\frac{i\lambda}{8\pi} V_P(1-\delta)\right\} \varphi(xy0) \quad (66)$$

where

$$\begin{aligned} V_P &= \int_0^\varepsilon V(xyz) dz \\ \delta &= \{ \langle z(xy) \rangle - \varepsilon/2 \} / (\varepsilon/2) \\ \langle z(xy) \rangle &= \int_0^{\varepsilon z} \frac{V(xyz)}{V_P(xy)} dz \end{aligned}$$

so that $\langle z(xy) \rangle$ is the center of the potential of the slice. Then δ is the relative deviation of the center of the potential from the geometric center of the slice.

The calculations are made by iterating the expression (66), slice by slice through the thickness of the sample. A detailed account of the success of this method of calculation and a comparison with other methods is given by Van Dyck [1985].

Problem

Derive the differential equation formulation of eqs. (10.35) and (10.36) from the slice formulation, e.g. equation (11.49).

Section IV

APPLICATIONS TO SELECTED TOPICS

This Page Intentionally Left Blank

Diffuse scattering and absorption effects

12.1. Thermal diffuse scattering

12.1.1. Phonons and vibrational waves

Apart from being scattered by the average periodic structure of a crystal, the incident radiation may interact with the crystal in a number of ways. The interaction may be described within the framework of the quantum mechanical theory. A quantum of radiation interacts to create or annihilate an excitation of the crystal such as a phonon, plasmon or exciton with consequent loss or gain of energy. The result of the interaction is, in general, the production of diffuse inelastically scattered radiation in the background of the diffraction pattern with an associated loss of intensity from the sharp Bragg reflections from the average periodic structure.

The main object of studying these effects is to investigate the nature and properties of the crystal excitations. A secondary, but still important aim, is to find the diffuse scattering and absorption effects due to these processes with sufficient accuracy so that they may be subtracted out from measurements of diffraction intensities required for crystallographic purposes.

So long as we stay within the kinematical or first Born approximation we may describe the scattering due to these excitations by use of the generalized Patterson function which we introduced in Chapter 7 and which we prefer to use because it forms a natural extension of our other considerations of kinematical scattering.

In Chapter 7 we used the simple Einstein model of thermal motion of atoms as an example for the use of the generalized Patterson function following Van Hove [1954]. Here we apply the same method to introduce a rather more realistic discussion of the atomic vibrations in crystals which constitute the most universal and important of crystal excitations.

The assumption of the Einstein model is that the motions of all atoms are independent. At the other extreme it is customary in rather more sophisticated treatments to consider the motion to be strongly correlated. The displacements of the atoms are considered to be given by the sums of displacements corresponding to the individual lattice waves which can occur in the crystal with a relationship between wavelength

(or wave vector k , or momentum $\hbar k$) and frequency ω (or energy $\hbar\omega$) given by the dispersion relationship for the material. The simple model is that each individual lattice wave may be represented by a plane wave with correlation of atom displacements extending throughout the crystal and so, effectively, to infinity in space and time. This is an oversimplified picture since the range of correlations of atom displacements is limited in both space and time. One may say that a phonon has a finite extent in space and a finite lifetime, being not much more than one wavelength or one period, respectively, for high frequency phonons in many cases. But we stick to the simpler picture of infinite waves since this gives results which are usually not too far wrong.

12.1.2. Scattering for a longitudinal wave

For a longitudinal wave travelling in the x -direction, the displacements of the atoms from their mean positions may be written

$$\Delta = A \cos 2\pi(lx - \nu_l t) \quad (1)$$

where $l = 1/\Lambda$ and Λ is the wavelength and ν_l the frequency of the lattice wave. The electron density function for an atom displaced from the average lattice position at $x = X$ is

$$\begin{aligned} \varphi_0(x, t) &= \rho_0(X + \Delta, t) \\ &= \rho_0(X) + \Delta(X, t) \left. \frac{d\rho_0}{dx} \right|_{x=X} + \frac{\Delta^2(X, t)}{2!} \left. \frac{\partial^2 \rho_0}{\partial x^2} \right|_{x=X} + \dots \end{aligned} \quad (2)$$

The corresponding generalized Patterson function is then

$$P(r, t) = \left[\rho_0 * \sum_n \delta(x - x_n) \right] * \left[\rho_0(-r) * \sum_m \delta(x + x_m) \right]$$

or we can write

$$P(r, t) = \bar{\rho} * \bar{\rho} + (\bar{\rho} - \rho) * (\bar{\rho} - \rho), \quad (3)$$

where the convolution is with respect to x and t .

Here we have separated the Patterson into the Patterson of the spatially periodic, time invariant, averaged electron density $\bar{\rho}$ and the Patterson function for the deviations from this average, in analogy with equation (7.9).

From (2) it is readily seen that in averaging over space or time, $\langle \Delta(t) \rangle = 0$ so that

$$\bar{\rho}_0 = \rho_0(X) + \frac{1}{2} \langle \Delta^2 \rangle \frac{\partial^2 \rho_0}{\partial x^2} + \dots$$

Fourier transforming gives

$$\begin{aligned}\bar{F}(u) &= F_0(u) \left[1 - \frac{4\pi^2}{2} u^2 \langle \Delta^2 \rangle + \dots \right] \\ &\approx F_0(u) \exp\{-2\pi^2 \langle \Delta^2 \rangle u^2\},\end{aligned}$$

and the average intensity function in reciprocal space is

$$\bar{F} \cdot \bar{F}^* = (F_0(u))^2 \exp\{-4\pi^2 \langle \Delta^2 \rangle u^2\}, \quad (4)$$

where the exponential is the contribution to the Debye-Waller factor for this wave.

12.1.3. Diffuse scattering component

The second part of the Patterson, equation (3), giving the diffuse scattering is, to a first approximation for a small Δ ,

$$\begin{aligned}\rho'_0(x) * \sum_n \delta(x-na) A \cos 2\pi(lna - v_l t) * \rho'_0(-x) \\ * \sum_m \delta(x+ma) A \cos 2\pi(lma - v_l t),\end{aligned}$$

where the prime on $\rho'_0(x)$ denotes differentiation. Carrying out the convolution with respect to both x and t , we get

$$\begin{aligned}\Delta\rho * \Delta\rho &= \sum_n \sum_m A^2 \cos 2\pi\{l(n-m)a - v_l t\} \\ &\quad \times [\rho'_0(x) * \rho'_0(x) * \delta\{x - (n-m)a\}] \\ &= NA^2 \sum_n [\rho'_0(x) * \rho'_0(-x) * \delta(x-na)] \cos 2\pi(lna - v_l t).\end{aligned} \quad (5)$$

Hence the intensity distribution function in reciprocal space is

$$I_{th}(u) = A^2 \{2\pi u F(u)\}^2 \sum_h \delta\left(u - \frac{h}{a}, v\right) * \delta(u \pm l, v \pm v_l). \quad (6)$$

The two functions (5) and (6) are sketched in Fig. 12.1(a), (b) and (c). In (a) and (b) the functions $\rho'_0(x) * \rho'_0(-x)$ and $4\pi^2 u^2 F^2(u)$ have not been included in the diagram. The width of the lines parallel to the t axis represent the magnitudes (positive and negative) modulating the function $\rho'_0(x) * \rho'_0(-x)$ sketched on the left-hand side. The periodicity is obviously $1/v_l$ in the t direction. The other axis of the

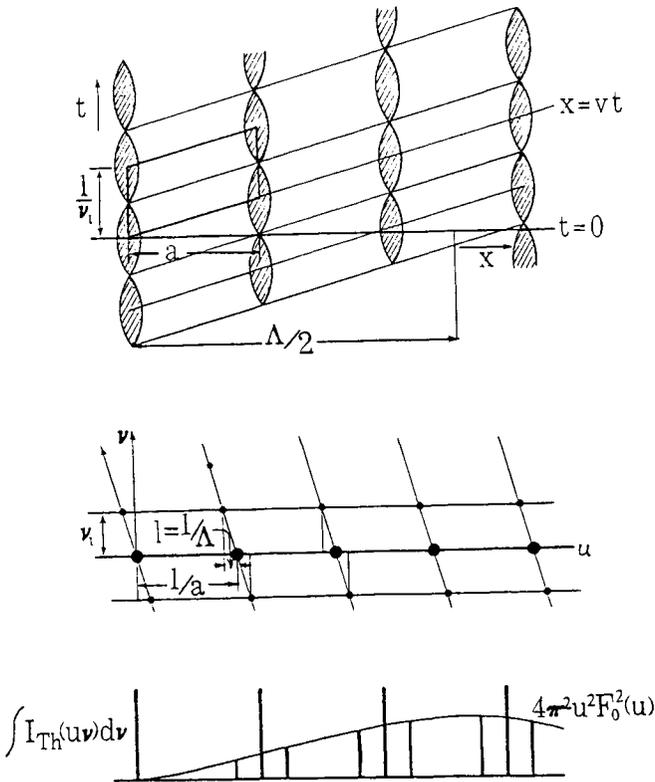


Fig. 12.1. (a) The Patterson function for the deviation from the time-averaged structure as a function of one spatial coordinate, x , and time, t , for the case of a longitudinal wave travelling in the x direction. (b) The Fourier transform of the function (a) giving the variation of scattering power with u and the change of frequency, v . (c) The intensity distribution resulting from (b) when the measurement does not distinguish differences of frequency.

"unit cell" of the distribution is the constant-phase line $x = vt$ where v is the velocity of the lattice wave.

Correspondingly in reciprocal space, Fig. 12.1(b), one axis of the reciprocal lattice unit cell is the u direction, with a periodicity $1/a$, and the other axis, the v axis, is perpendicular to the $x = vt$ line, with a periodicity corresponding to $v = v_l$. The average Patterson gives the strong scattering points at $u = h/a, v = 0$. The Patterson for the deviation from the average lattice adds to these the satellite spots for $u = \pm l, v = \pm v_l$.

From Fig. 12.1(b) we deduce that the purely elastic scattering $I(u, 0)$, the $v = 0$ section of $I(u, v)$, contains only the peaks corresponding to the average periodic lattice. The lattice wave introduces reflections with gain or loss of energy $h\nu_l$.

If, as in the case of X-ray or electron diffraction, no energy resolution is possible, what is measured is the total scattering, or the projection $\int I(u, v) dv$. From Fig. 12.1(c), this includes the strong diffraction spots due to the average lattice plus the weak satellite peaks due to the lattice wave. The former have intensity given by (4). The latter have intensity proportional to the square of the wave amplitude A , and also to $4\pi^2 u^2 F_0^2(u)$. This function is zero at the origin of reciprocal space, increases to a maximum usually for u of the order of 0.5 to 1 \AA^{-1} and then decreases slowly.

Each thermal wave occurring in the lattice gives a pair of satellite peaks around each reciprocal lattice point, with a displacement inversely proportional to the wavelength and the relative intensity depending on its amplitude squared, and so on the energy of the wave ($=h\nu_l$).

12.1.4. Dispersion curves

The relationship of the wave energy or frequency to the inverse wavelength, which is described by the dispersion curves, is linear for long wavelengths, the slope of the lines being given in the long wavelength limit by the elastic constants of the material. Hence the intensity of the satellite peaks is inversely proportional to the square of the distance from the reciprocal lattice points. Then for a large number of possible wave with values of inverse wavelengths equally distributed, the satellites merge into a diffuse intensity, peaking sharply at reciprocal lattice positions.

If the intensity distribution can be determined as a function of both the reciprocal distance, u , or momentum change and the change in frequency ν , or energy of the incident radiation, then the form of the dispersion curve may be derived. This may be done with neutron diffraction since the energy of the incident thermal neutron is of the order of 0.02 eV while the thermal waves have energies of the same order. The changes of energy of the incident neutrons are relatively large and can be deduced by using a diffracting crystal to analyse the energy (or wavelength) distribution of the scattered neutrons. For X-rays or electrons incident with energies in the range of tens of kilovolts or higher, the changes of energy due to the thermal waves are in general not detectable. On the assumption that the dispersion curves are linear in the long-wavelength range, measurements of the diffuse intensity allow the derivation of the elastic constants (Warren [1969]).

12.1.5. Three-dimensional generalizations

We have treated only the simplest case of longitudinal waves in one dimension. The three-dimensional case is further complicated by the existence of both longitudinal and transverse waves, each with different dispersion curves, and by the presence of both "optical" and "acoustic" modes of vibration. Thus the vector displacement Δ may be at any angle with the diffraction vector u and we must consider the scalar product $u \cdot \Delta$. For a treatment including these complications we refer the reader to the various detailed discussions in the literature, including those by Hoppe [1956, 1964], Laval [1958], Warren [1969] for the X-ray diffraction case and Bacon [1975] for the description of neutron diffraction methods and measurements. See also Willis and Pryor [1975].

We have limited our discussion also to the first-order approximation valid for small displacements, Δ , only. Taking into account the higher order terms of the series (2) includes the effects of multiple scattering by single lattice waves. When more than one lattice wave must be present at a time it is necessary for completeness to consider also the successive scattering by different waves.

These effects are of importance in practice for accurate diffraction studies and particularly for studies of thermal diffuse scattering by crystals at high temperatures. It is easy to see how our treatment may be extended to include them. They have been discussed using other approaches by Paskin [1958, 1959], Borie [1970] and Wang [1995].

For the general three-dimension case the expression (4) for the Debye-Waller factor remains valid if $\langle \Delta^2 \rangle$ is interpreted as the mean-square displacement of the atom in the direction of the diffraction vector. In general this value is not isotropic and varies with the type of atom and its environment. In accurate crystal structure analysis, three parameters defining a vibration ellipsoid are normally refined for each non-equivalent atom.

12.2. Static atom displacements

12.2.1. Relaxation around point defects

Diffuse scattering in diffraction patterns may arise from an irregular arrangement of atoms on the lattice sites so that there is only short range ordering of the site occupancies (Chapter 17) or from disordered arrangement of vacancies, interstitials or impurity atoms. The most important diffraction effect produced by small defects in crystals often comes from the strain field produced in the

surrounding crystal. For an atom vacancy, an interstitial atom or pair of atoms, or an impurity atom, the neighboring atoms may be displaced from their average-lattice sites by appreciable fractions of the interatomic distance. The atom displacement usually decreases fairly slowly with distance from the point defect so that a very large number of atoms may be affected. Then the diffraction effect due to the atom displacements may well be much greater than the scattering by the impurity, interstitial or missing atom itself.

Simple models for the diffraction by vacancies and by the relaxation of atoms around a vacancy were discussed in Chapter 7 as examples for the use of the generalized Patterson function. We now treat this type of problem in a different context, providing a basis for the understanding of experimentally important situations.

For the static displacement of atoms there is no essential difference from the dynamical case of thermal vibrations except that, since there is no time dependence, there is elastic scattering only and the practical distinction between the capabilities of X-rays or electrons and neutrons is removed. The generalized Patterson function is considered as a function of the three spatial coordinates only. As we shall see, there is diffuse scattering rather similar to the thermal diffuse and a factor is applied to the sharp Bragg reflections which is similar to the Debye-Waller factor.

It would be possible to resolve the static displacements into plane wave Fourier components and proceed as in the last section. Instead we illustrate another method of approach to diffuse scattering problems and, at the same time, provide an introduction to the treatment to be used later (Chapter 17) in the discussion of short-range order diffuse scattering.

12.2.2. Diffraction intensities for displaced atoms

For simplicity we assume the crystal lattice to be primitive with one atom site per unit cell. Then the structure amplitude is

$$F(\mathbf{u}) = \sum_i f_i \exp\{2\pi i \mathbf{u} \cdot (\mathbf{R}_i + \Delta_i)\}$$

and

$$F(\mathbf{u}) F^*(\mathbf{u}) = \sum_i \sum_j f_i f_j \exp\{2\pi i \mathbf{u} \cdot (\mathbf{R}_i - \mathbf{R}_j)\} \exp\{2\pi i \mathbf{u} \cdot (\Delta_i - \Delta_j)\}, \quad (7)$$

where the \mathbf{R}_i vectors are lattice vectors for the average periodic lattice and Δ_i is the displacement of the atom from the lattice site defined by \mathbf{R}_i .

For simplicity also we consider that the lattice sites are occupied by two types of atom, A and B, present in the crystal in fractional proportions m_A and m_B defined so that $m_A + m_B = 1$. The atom B is considered as impurities (or vacancies) so that m_B is small and the B atoms may be assumed as a first approximation, to occur at random.

We now specify the occupancy of the sites by means of the occupational parameters (Flinn [1956])

$$\sigma_i^A = \begin{cases} 1 & \text{for an A onsite } i, \\ 0 & \text{for a B onsite } i, \end{cases}$$

$$\sigma_i^B = \begin{cases} 1 & \text{for a B onsite } i, \\ 0 & \text{for an A onsite } i. \end{cases} \quad (8)$$

If an isolated atom B is larger than the average atom size we would expect its neighboring atoms to be displaced outwards, away from it. We assume that we can describe these displacements in terms of a "displacement field" which is a function of the distances between lattice sites (which approximate to the distance between atoms). Assuming that the displacement caused by a B atom does not depend on the nature of the atom being displaced we can write the displacement of an atom on the j site due to the B atom on the i site as Δ_{ij}^B . If a A atom is considered to also give rise to a displacement field, the reasonable requirement that an average atom should provide no displacement gives the relationship

$$m_A \Delta_{ij}^A + m_B \Delta_{ij}^B = 0, \quad (9)$$

which implies that displacements due to A and B atoms are collinear and oppositely directed.

The displacement of the atom at i from its average lattice site defined by R_i is assumed to be given by the sum of the displacements due to all other atoms, i.e.

$$\Delta_i = \sum_k \{ \sigma_k^A \Delta_{ki}^A + \sigma_k^B \Delta_{ki}^B \}, \quad (10)$$

and, by definition

$$\langle \Delta_i \rangle = 0, \quad (11)$$

where the angular bracket represents an averaging over all lattice sites of the average, periodic lattice.

12.2.3. *The Bragg peaks*

The sharp Bragg peaks correspond to diffraction from the average structure for which, from (7)

$$\begin{aligned}\bar{F}(u) &= \left\langle \sum_i f_i \exp\{2\pi i u \cdot (R_i + \Delta_i)\} \right\rangle \\ &= \sum_i \exp\{2\pi i u \cdot R_i\} \langle f_i \exp\{2\pi i u \cdot \Delta_i\} \rangle.\end{aligned}\quad (12)$$

The average effective atomic scattering factor included as the last term of (12) is

$$\bar{f} = \langle (\sigma^A_i f_A + \sigma^B_i f_B) \exp\{2\pi i u \cdot \Delta_i\} \rangle.\quad (13)$$

If it can be assumed that the atomic displacements are small so that $u \cdot \Delta_i \ll 1$, the exponential can be expanded in a power series to give, using (10),

$$\begin{aligned}\bar{f} &= f_A \left[\langle \sigma^A_i \rangle + 2\pi i \sum_k \{ \langle \sigma^A_i \sigma^A_k \rangle (u \cdot \Delta^A_{ki}) + \langle \sigma^A_i \sigma^B_k \rangle (u \cdot \Delta^B_{ki}) \} \right. \\ &\quad - 2\pi^2 \sum_k \sum_l \{ \langle \sigma^A_i \sigma^A_k \sigma^A_l \rangle (u \cdot \Delta^A_{ki}) (u \cdot \Delta^A_{li}) \\ &\quad \left. + \{ \langle \sigma^A_i \sigma^A_k \sigma^B_l \rangle (u \cdot \Delta^A_{ki}) (u \cdot \Delta^B_{li}) + \dots \} + \dots \right] + f_B \left[\langle \sigma^B_i \rangle + \dots \right]\end{aligned}\quad (14)$$

From the definition of σ^A_i we have $\langle \sigma^A_i \rangle = m_A$. If the B atoms are distributed at random in the lattice we have, for example,

$$\begin{aligned}\langle \sigma^A_i \sigma^B_k \rangle &= \langle \sigma^A_i \rangle \langle \sigma^B_k \rangle = m_A m_B, \\ \langle \sigma^A_i \sigma^B_k \sigma^B_l \rangle &= m_A m_B^2.\end{aligned}\quad (15)$$

Then in (14) the first order terms in $(u \cdot \Delta)$ are zero because of the relationship (9) and the second order term is similarly zero unless $k = l$ so that we obtain

$$\begin{aligned}\bar{f} &= f_A \left[m_A - 2\pi^2 m_A \sum_k \{ m_A (u \cdot \Delta^A_{ki})^2 + m_B (u \cdot \Delta^B_{ki})^2 \} + \dots \right] \\ &\quad + f_B \left[m_B - 2\pi^2 m_B \sum_k \{ m_A (u \cdot \Delta^A_{ki})^2 + m_B (u \cdot \Delta^B_{ki})^2 \} + \dots \right]\end{aligned}$$

or

$$\begin{aligned} \bar{f} &\approx (m_A f_A + m_B f_B) \exp \left\{ -2\pi^2 \frac{m_B}{m_A} \sum_k (u \cdot \Delta^{B_{ki}})^2 \right\} \\ &\approx (m_A f_A + m_B f_B) \exp \{ -2\pi^2 (u \cdot \Delta_i)^2 \}. \end{aligned} \quad (16)$$

Thus, at least to the second order approximation, the effect of the atom displacements on the Bragg peaks is to multiply the structure amplitudes by an exponential factor which is of exactly the same form as the Debye-Waller factor due to thermal motion. The fact that this pseudo Debye-Waller factor is the same for both types of atom is a result of our assumption that the displacement fields act on all atoms equally.

Detection of this factor depends on separation from the thermal Debye-Waller factor by making use of either its lack of dependence on temperature or its linear dependence on impurity concentration as evidenced by the first part of (16).

12.2.4. The diffuse scattering

To calculate the diffuse scattering, we may evaluate the total scattering intensity by use of equation (7) and then subtract out the Bragg scattering terms. The double summation in (7) can be interpreted as indicating that we may take each site j in turn as origin and consider the scattering amplitude and relative phases for atoms separated from it by vectors $r_i - r_j$. Thus (7) may be written, for real f values,

$$|F(u)|^2 = N \sum_i (f_0 f_i \exp\{2\pi i u \cdot R_i\} \cdot \exp\{2\pi i u \cdot (\Delta_i - \Delta_0)\}). \quad (17)$$

Then making the same sort of substitutions and expanding the exponential as in (13) and (14)

$$\begin{aligned} |F(u)|^2 &= N \sum_i \exp\{2\pi i u \cdot R_i\} \left\{ (\sigma^{A_0} f_A + \sigma^{B_0} f_B) (\sigma^A_i f_A + \sigma^B_i f_B) \right. \\ &\quad \times \left[1 - 2\pi i \sum_k \{ \sigma^A_k (u \cdot \Delta^A_{k0}) + \sigma^B_k (u \cdot \Delta^B_{k0}) \} + \dots \right] \\ &\quad \left. \times \left[1 - 2\pi i \sum_k \{ \sigma^A_k (u \cdot \Delta^A_{ki}) + \sigma^B_k (u \cdot \Delta^B_{ki}) \} + \dots \right] \right\}. \end{aligned} \quad (18)$$

We isolate first the case $i = 0$ and use the relations $\sigma^{A_0} \sigma^{A_0} = \sigma^{A_0}$ and $\sigma^{A_0} \sigma^{B_0} = 0$ to get the contribution to the diffuse part as

$$N(m_A f_A^2 + m_B f_B^2) - N(m_A f_A + m_B f_B)^2 = N m_A m_B (f_A - f_B)^2. \quad (19)$$

Continuing with the assumption $i \neq 0$, we see that the zero order term in $(\mathbf{u} \cdot \Delta)$ in (18) is part of the Bragg scattering. In the first order term the contributions for $k \neq 0$, i are zero from (10) and (11). For $k = 0$ and i this is not so since $\Delta_{ii}^B = \Delta_{00}^B = 0$. Then using (9) to express the result in terms of Δ_{ij}^B only, we obtain

$$2N(m_A f_A + m_B f_B) (f_B - f_A) m_B 2\pi i \sum_i (\mathbf{u} \cdot \Delta_{0i}^B) \exp\{2\pi i \mathbf{u} \cdot \mathbf{R}_i\} \\ = 2\pi N m_B (m_A f_A + m_B f_B) (f_B - f_A) \sum_i (\mathbf{u} \cdot \Delta_{0i}^B) \sin 2\pi (\mathbf{u} \cdot \mathbf{R}_i). \quad (20)$$

The second order terms of the series in the square brackets of (18) and the second order term obtained by multiplying these series give summations of the two sets of indices indicated by k and l . Again using (9) or (11) we can set all terms equal to zero unless $k = l$, when the expressions simplify again for $k = 0$, or else for the special cases $k = 0$, $l = 1$ or $k = 1$, $l = 0$. Then we derive the second order contribution in the form

$$4\pi^2 N \frac{m_B}{m_A} (m_A f_A + m_B f_B)^2 \sum_i \left\{ \sum_{k \neq 0, i} (\mathbf{u} \cdot \Delta_{k0}^B) (\mathbf{u} \cdot \Delta_{ki}^B) \right\} \cos 2\pi \mathbf{u} \cdot \mathbf{R}_i \\ - 4\pi^2 N \left[\frac{m_B}{m_A} (m_A f_A + m_B f_B) (m_B f_A + m_A f_B) + m_B^2 (f_A - f_B)^2 \right] \\ \times \sum_i (\mathbf{u} \cdot \Delta_{i0}^B)^2 \cos 2\pi \mathbf{u} \cdot \mathbf{R}_i. \quad (21)$$

The forms of these contributions to the diffuse scattering are suggested in Fig. 12.2. The term (19) sketched in 12.2(a) is symmetric about the reciprocal space origin and falls off monotonically. The first order term in $(\mathbf{u} \cdot \Delta)$, given by (20) is symmetric about the origin but anti-symmetric about each reciprocal lattice point. Then the second order term (21) gives a contribution symmetric about each reciprocal lattice point. When these are added together we see that there is no peak around the origin, but around each reciprocal lattice point there is a diffuse peak with its maximum somewhat displaced from the reciprocal lattice point because of the influence of the asymmetrical contribution from (20). The direction of this displacement depends on the relative signs of $(f_B - f_A)$ and $(\mathbf{u} \cdot \Delta_{0i}^B)$. If the "impurity" atom B is larger in size than the host atom A and also has a larger scattering factor, these two quantities are of the

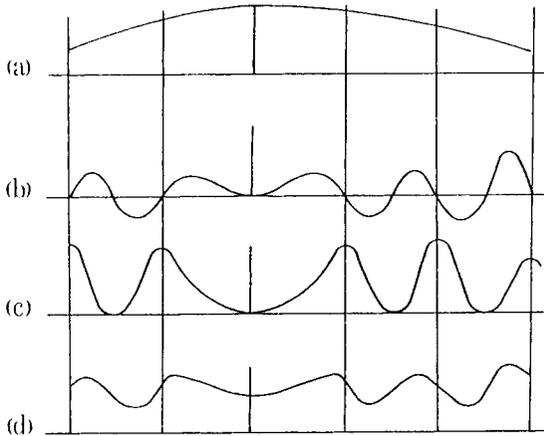


Fig. 12.2. The contributions to the intensity distribution of the diffuse scattering due to size-effect displacements of atoms in the presence of impurities or defects: (a) is the "Laue monatomic" scattering, (b) is the term asymmetric about the reciprocal lattice points and (c) is the part symmetric about the reciprocal lattice points: (d) is the summation of these components.

same sign and the displacement of the diffuse peaks is towards the origin, but for a small, heavy impurity the displacement tends to be outwards.

In order to calculate the distributions of diffuse scattering in detail, the exact form of the displacement vectors Δ_{0i}^B must be known, but there are few materials for which these vectors are known with any confidence. A reasonable first approximation, used by Huang [1947] and Borie [1957, 1959, 1961] is to make use of the formula derived for the macroscopic case of a center of dilation in a uniform isotropic solid, giving

$$\Delta_{0i}^B = \frac{c_B r_{0i}}{|r_{0i}|^3}. \quad (22)$$

Then some of the summations in the above expressions may be performed analytically. For example

$$\sum_k (u \cdot \Delta_{ki}^B) = c_B^2 \sum_k \frac{(u \cdot r_{ki})^2}{|r_{ki}|^6}, \quad (23)$$

and for a simple face-centered cubic lattice (Born and Misra [1940]) this is equal to $c_B^2 \cdot 33.7 \cdot |u|^2 / a_0^4$.

These solutions assuming the strain field to have the asymptotic form given by continuum elastic theory give reasonable results for the scattering very close to the Bragg peaks, but are less appropriate for the description of atom displacements near the defect and for the diffuse scattering. Calculations of the displacements of the near neighbors of a point defect in solid argon by Kanzaki [1957] and computer simulations of the surroundings of point defects in copper by Tewordt [1958] have shown behaviour very different from this. Along some directions, such as cube axes for argon, the displacements may actually change sign with increasing distance from the defect. Flocken and Hardy [1970] have deduced that the asymptotic solution is valid only for distances from the defect greater than several times the unit cell dimensions.

Calculations of the diffuse scattering distributions far from Bragg reflections given by various point defects have been made (e.g. Benedek and Ho [1973]) and observations of this diffuse scattering for X-ray diffraction from large concentrations of point defects has been reported by Haubold [1974]. Results of investigations of the structure and agglomeration of interstitials in metals irradiated at low temperatures have been reported by Ehrhart [1978].

Meanwhile a general approach to the scattering from defects which allows large displacements of atoms to be treated has been formulated by Krivoglaz [1969] and this has been further developed by many others.

12.3. Electron excitations

12.3.1. Inelastic X-ray scattering

For both X-rays and electrons an important source of diffuse background scattering and of absorption of energy from the sharp Bragg reflections is the inelastic scattering of the incident radiation by electrons in the crystal. The addition of real and imaginary parts to the X-ray atomic scattering factor as a result of excitation of electrons of the inner electron shells has been discussed in Chapter 4. The imaginary part of the scattering factors provides an absorption coefficient which may be very large for incident wavelengths just smaller than the absorption edges, i.e., when the incident quanta have just enough energy to eject an electron from one of the inner shells. Diffuse background is then produced in the diffraction pattern by the generation of characteristic radiation from the atoms of the specimen.

The collision of X-ray quanta with nearly-free electrons in the specimen gives the well-known Compton scattering. In

Chapter 5 we considered the one electron case as an example of the use of the generalized Patterson function $P(\mathbf{r}, t)$, giving correlations in both space and time. The total scattering from one electron is 1 electron unit. The elastic scattering is

$$\Im \int P(\mathbf{r}, t) d\mathbf{r} = |F(\mathbf{u}, 0)|^2 = |f_e(\mathbf{u})|^2,$$

so that the inelastic scattering, given by the difference, is

$$I_{inel} = 1 - |f_e(\mathbf{u})|^2. \quad (24)$$

For an atom containing many electrons, the total time-average electron density distribution $\rho(\mathbf{r})$ is the sum of the electron density functions for all electrons so that, Fourier transforming,

$$\begin{aligned} I_{el} &= \left| \sum_k f_k(\mathbf{u}) \right|^2 \\ &= \sum_k |f_k|^2 + \sum_{j \neq k} f_j f_k^*. \end{aligned} \quad (25)$$

A first approximation to the inelastic scattering is given by assuming all electrons to scatter independently and write, from (24),

$$I_{inel} = \sum_k \{1 - |f_k(\mathbf{u})|^2\}.$$

However at this stage we must take into account the quantum-mechanical properties of electrons in atoms including the effects of the Pauli Exclusion Principle (James [1948]). Then the simple classical generalized Patterson is no longer sufficient. The result obtained is

$$I_{inel} = \sum_k \{1 - |f_k|^2\} - \sum_{j \neq k} |f_{jk}|^2 \quad (26)$$

where f_{jk} is the transition matrix element involving exchange.

The Compton scattering intensity is seen to be practically zero at small scattering angles for which the f_k for each electron is near unity. Then it increases with increasing angle until, particularly for light elements, it may be comparable with the elastic scattering given by near-amorphous materials (James [1948]).

12.3.2. Electron excitation by electrons - plasmons

Incident electron beams may excite electrons from the inner shells of atoms. This, of course, provides the means for X-ray generation in X-ray tubes. Apart from the characteristic X-rays and white radiation, or "brehmstrahlung",

this may give rise to the emission of "Auger" electrons having energies characteristic of the atoms. While both the characteristic X-rays and Auger electrons provide important means for analysis of the chemical constitution of materials, neither contributes appreciably to background in electron diffraction patterns or to the absorption coefficients of electrons in solids. In these respects the excitation of the outer-shell or valence electrons is much more important.

For some metals, such as Al and Mg, the predominant electron excitation is the excitation of collective oscillations, or plasmons. The sea of nearly free electrons may be set vibrating with a characteristic plasmon frequency by the impact of a fast incident electron. The energy of the plasmon and the energy loss of the incident electron are typically of the order of 10 to 20 eV. The wavelength of the plasmon oscillation is normally many unit cells and the angle of scattering of the incident electron is very small compared with Bragg angles for crystalline reflections.

The theory of the excitation of plasmons has been given by Pines and Bohm (see Pines [1955, 1964]; Ferrell, [1956]). Ritchie [1957] introduced the concept of the surface plasmon, generated at the boundaries of solids, lower in energy and important for diffraction from very small crystals. A survey of the theory and experimental results has been given by Raether [1980].

The characteristic frequency for bulk plasmons is given by elementary considerations as $\omega_p = (ne^2/m\epsilon_0)^{1/2}$. The slight dependence of the frequency on k is given by the dispersion relationship,

$$\omega(k) = \omega_p \left(1 + \frac{3}{5} \left(\frac{E_F}{\hbar\omega_p} \right) \frac{\hbar}{m\omega_p} k^2 + \dots \right) \quad (27)$$

where E_F is the Fermi energy.

The intensity distribution of the electrons which have suffered one plasmon loss is given by the differential cross-section relationship of Ferrell [1956] as

$$\sigma_B(\theta) = \frac{\theta_E}{2\pi a_0 n} \frac{1}{\theta^2 + \theta_E^2} G^{-1}(\theta) \quad (28)$$

where $\theta_E = \Delta E/2E_0$, a_0 is the Bohr radius, n is the electron density and $G^{-1}(\theta)$ is a function which is unity for small angles but falls to zero at $\theta = \theta_c$, the cut-off angle given by

$$\theta_c = 0.74 (E_F/E_0)^{1/2}.$$

The scattering intensity for surface plasmons depends on the thickness of the specimen foil or size of the specimen

particles (Fujimoto and Komaki [1968]) but for the limiting case of a thick foil is

$$\sigma_s(\theta) = \frac{e^2}{\pi \hbar v} \frac{2}{1 + \varepsilon} \frac{\theta \cdot \theta_E}{(\theta^2 + \theta_E^2)^2} f \quad (29)$$

where v is the incident electron velocity and f is a geometric factor. For a foil of thickness D there are two values for the surface plasmon frequency,

$$\omega_{\pm}(k) = (1 + \varepsilon)^{-1/2} \omega_p (1 \pm \exp\{-kD\})^{1/2}, \quad (30)$$

where ε is the dielectric constant. Experiments have verified the existence of these ω_+ and ω_- frequencies for very thin films.

From (28) it is seen that for bulk plasmons the half width of the angular distribution is given by $\theta = \theta_E$. For 50-100 kV electrons and characteristic energy losses ΔE of 10 to 20 eV this half width is of the order of 10^{-4} radians. This may be compared with the scattering angle of about 10^{-2} radians for the first Bragg reflection from simple metals. For surface plasmons the distribution is even narrower. Thus plasmon scattering may be considered as essentially small angle scattering, tending to broaden diffraction spots slightly. With increasing specimen thickness multiple plasmon losses occur with increasing frequency and the angular distribution is progressively broadened.

The path length for plasmon scattering is typically of the order of 1000 Å. For Al, for example, Spence and Spargo [1970] found a value of 740 ± 40 Å for 58.5 kV. For very thin films surface plasmons predominate but usually these losses become less important for thicknesses of the order of 100 Å.

It is only for relatively few metals, including Al and Mg that well defined plasmon losses dominate the energy loss spectrum. For most metals, semiconductors and insulators, the energy-loss spectrum is complicated, with occasional peaks which can be assigned to plasmon excitation but with other strong, sharp or diffuse bands usually described as being due to single electron excitations although the associations with particular excitation processes of the crystal electrons is at best rather vague.

12.3.3. Single-electron excitations

For nearly-free, or valence, electrons in the crystal the excitations may be considered in terms of the energy-band picture as corresponding to transitions from one energy level to another within the same band (intra-band transitions, possible when the band is not filled, as in metals) or to a level within a higher unfilled band (an interband transition).

In general it may be said that the wave function for an inelastically scattered electron derives from the wave equation

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + H_{nm} - E_n \right) \psi_n = -\sum_{m \neq n} H_{mn} \psi_m, \quad (31)$$

where

$$H_{nm}(r) = \int a_n^*(r_1 \dots r_N) H(r, r_1 \dots r_N) a_m(r_1 \dots r_N) dr_1 \dots dr_N. \quad (32)$$

Here $a_n(r_1 \dots r_N)$ is the wave function for the electrons in the crystal having positions $r_1 \dots r_N$ and $H(r, r_1 \dots r_N)$ is the interaction energy of the incident electron (position r) with the crystal electrons.

For electrons in a crystal the interaction energy must show the periodicity of the crystal lattice and it was shown by Howie [1963] that

$$H_{nm}(r) = \exp\{-2\pi i(q_{nm} \cdot r)\} \sum_g H_g^{nm} \exp\{2\pi i g \cdot r\} \quad (33)$$

where q_{nm} is the wave vector for the excitation in the crystal and g is the reciprocal lattice vector.

Assuming a simple Coulomb interaction of the electrons the interaction energy is of the form $H(r) = \sum_n e^2 / (r - r_n)$, which allows the integration over r in (32) to be carried out directly. Then, following Cundy et al. [1966] the kinematical amplitude of the scattering involving a single electron excitation from a state m to a state n is proportional to

$$H_g^{nm} = \frac{e^2}{\pi V \epsilon(q-g, \omega) |q-g|^2} \int a_n^*(r) \exp\{2\pi i(q-g) \cdot r\} a_m(r) dr, \quad (34)$$

where $\epsilon(q-g, \omega)$ is the value of the dielectric constant for a change of momentum $\hbar(q-g)$ and a change of energy $\hbar\omega$; V is the volume of normalization. As a special case of this we note that H_g^{nn} terms relate to elastic scattering. In

particular H_g^{00} is the scattering for all electrons remaining in the ground state, and this depends on the Fourier transform of $a_0^*(r)a_0(r)$ which is just the electron density distribution $\rho(r)$ for the ground state of the crystal. Taking the g^{-2} term into account relates this to the Fourier transform of the unperturbed potential, $\phi_0(r)$.

For the inelastic processes the amplitude of scattering is given by the Fourier transform of the product of wave functions $a_n^* a_0$. For nearly free electrons both of these

functions may be Bloch waves of crystal electrons and therefore periodic. The diffracted amplitudes are then sharply peaked at the Bragg reflection positions, just as for elastic scattering. However, in contrast to the elastic scattering case, only the outer electron shells are involved. Hence the effective scattering distribution has a broader spread around the lattice points: the effective atomic scattering factor falls off more rapidly with angle than for elastic scattering. This is the effect observed by Kuwabara in measurements of the relative intensities of diffraction rings obtained from thin metal and other films when an energy analyser was used to select electrons having only particular values of energy loss (Kuwabara and Cowley [1973]).

The greatest contribution of diffuse scattering intensity due to electron excitations comes from the $g = 0$ terms of equations such as (34). For these the intensity is proportional to $|q|^{-2}$ and so falls in much the same way as for plasmon scattering, with a comparable half-width (Humphreys and Whelan [1969].) The width of the scattering distribution can be taken as inversely dependent on the localization of the inelastic scattering process. The inelastic scattering is localized to a region of width approximately $\lambda E/\Delta E$ for an energy loss of ΔE . The effect of the localization or delocalization of the process on the contrast of electron microscope images of defects in crystals has been discussed, for example, by Craven et al. [1978] and Spence [1992].

12.4. Dynamical effects in diffuse scattering

12.4.1. Scattering and re-scattering

Under the conditions of diffraction for which dynamical diffraction of the sharp Bragg reflections is important, the dynamical diffraction also affects the diffuse scattering intensity. Firstly it must be taken into account that the incident beam is not the only strong beam in the crystal. Each of the diffracted beams also acts as a source of diffuse scattering. Secondly the diffusely scattered radiation undergoes diffraction as it passes through the crystal. Diffuse scattering in two directions separated by twice the Bragg angle for a Bragg reflection may be coupled dynamically, giving rise, among other things, to the appearance of Kossel or Kikuchi lines (see Chapter 14). Finally, diffusely scattered radiation may be diffusely scattered a second time or more times so that for a thick crystal the observed diffuse intensity may be the sum of many multiply scattered components, all modified by dynamical interactions through Bragg reflections.

For X-ray diffraction in a perfect crystal a two beam dynamical theory is normally sufficient. In the case of

thermal diffuse scattering for example (see O'Connor [1967]), both the incident and diffracted beams can be considered to generate diffuse scattering in proportion to their intensities. In general the diffusely scattered radiation is transmitted through the crystal with the average absorption coefficient. But where it meets a diffracting plane at the Bragg angle it is diffracted giving rise to sharp Kossel or Kikuchi lines.

For electron diffraction the situation is normally complicated by n -beam diffraction effects, although some useful results were obtained for thermal diffuse scattering by Takagi [1958] for a two beam case and by Fujimoto and Kainuma [1963], Fujimoto and Howie [1966] and Ishida [1970] who extended this approach and included other types of diffuse scattering. A useful approach to the general n -beam dynamical treatment was made by Gjønnes [1965, 1966] and applied by Gjønnes and Watanabe [1966] for cases involving relatively few beams and by Fisher [1965] to the case of short-range order diffuse scattering. In extension of this approach, Cowley and Pogany [1968] gave a general theory and outlined computational methods which were subsequently used by Doyle [1969, 1971] for detailed calculations on thermal diffuse and plasmon scattering and by Cowley and Murray [1968] for consideration of short-range order scattering.

If a crystal is considered to be divided into many thin slices almost perpendicular to the incident beam, the total single-diffuse scattering can be considered as the sum of contributions from each slice taken separately. For diffuse scattering from a particular slice the crystal can be con-

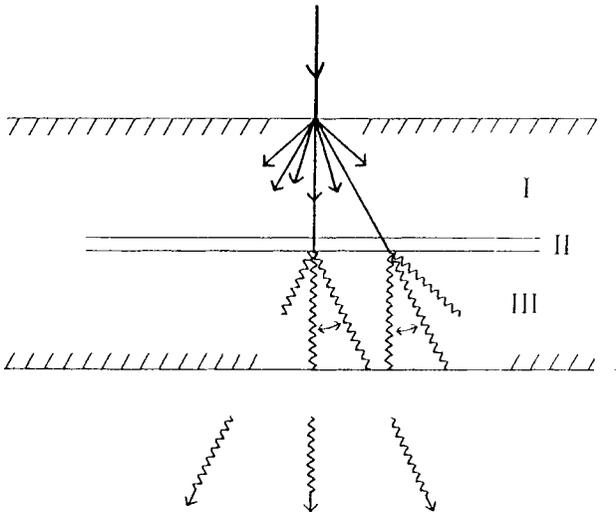


Fig. 12.3. Diagram representing the scheme for dynamical electron diffraction calculation of diffuse scattering.

sidered as divided into three parts as suggested in Fig. 12.3. The incident beam enters region I and undergoes n -beam dynamical diffraction as for a perfect crystal. Region II is the slice in which each Bragg beam gives rise to diffuse scattering. Then in region III the diffusely scattered radiation undergoes n -beam dynamical interactions through the Bragg reflections.

12.4.2. Coherent and incoherent scattering

Even though the electron wave may have lost energy or otherwise have been rendered incoherent with the elastically scattered beam it maintains coherence, or the ability to interfere, with itself. As suggested in Fig. 12.4, if the diffuse scattering process corresponds to a change q in the scattering vector, n -beam dynamical diffraction takes place between the points $h + q$, where h is a reciprocal lattice vector, with interactions depending on the structure amplitudes $\Phi(h_1 - h_2)$ and the appropriate excitation errors. The calculation for region III must be made for each vector within the Brillouin zone (or basic reciprocal lattice unit cell).

If the amplitude of diffuse scattering from the m slice of the crystal is $\Psi_m(u, v)$ the total single-scattering diffuse intensity is given by adding either the amplitudes or intensities of diffuse scattering from each slice, depending on whether there is correlation between the diffuse scattering processes of separate slices or not. Thus in the case of plasmon diffuse scattering it may be assumed that the plasmon wave extends through the whole thickness of a crystal. This provides a definite relationship between the amplitudes of scattering for all slices and the diffuse scattering intensity is given by

$$I_{\text{coh}}(u, v) = \left| \sum_m \Psi_m(u, v) \right|^2. \quad (35)$$

If the diffuse scattering comes from a random distribution

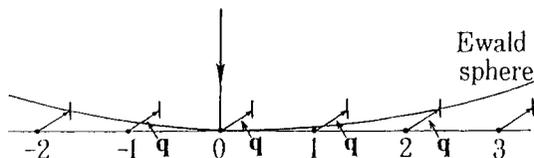


Fig. 12.4. The intersection of the Ewald sphere with the regions of scattering power corresponding to a phonon having momentum q .

of point defects there is no definite phase relationship between the scattering from successive slices and when the average is taken over a large number of defects the diffuse intensities are added incoherently so that

$$I_{\text{incoh}}(u, v) = \sum_m |\Psi_m(u, v)|^2. \quad (36)$$

There are intermediate cases for which there is a phase relationship in the diffuse scattering over a limited distance. This is the case for some phonon scattering for which it is suggested, for example by neutron diffraction observations, that the correlation in the movements of the atoms extends over perhaps 10 to 100 Å. Then the amplitudes of diffuse scattering must be added over a limited range. The total intensity may then be approximated by an expression such as

$$I(u, v) = \sum_m \left| \frac{1}{n} \sum_{p=m-n/2}^{m+n/2} \Psi_p(u, v) \right|^2. \quad (37)$$

The addition of amplitudes as in (35) was the basis for the detailed n -beam calculations of Doyle [1971] of the intensity of plasmon scattering given by thin crystals of Al oriented so that only the hhh systematic set of reflections are excited. In this way he showed agreement with the experimental observation that when the strong 111 reflection is excited the diffuse scattering tends to be excluded from the region between the strong 000 and 111 beams (see also Howie [1963]). Also he derived the form of the thickness fringes obtained with the 000 and 111 beams for electrons which had suffered a single plasmon energy loss (15 eV), in agreement with experiment.

The summation of intensities as in (36) may be expected to be appropriate for the diffuse scattering well away from the strong Bragg reflections since this is produced mostly by scattering involving short wavelength phonons, small defects, or the excitation of electrons tightly bound to atoms. For each of these cases the excitation is limited to a small region of the crystal.

The assumption of the addition of intensities was used by Gjønnes [1966] and has since been extended (Gjønnes and Høier [1971]; Høier [1973]) to provide a general description of dynamical diffuse scattering in terms of a six-dimensional real-space distribution.

The case of an intermediate range of correlation of the excitation giving diffuse scattering, treated by use of (33), was the basis for the study of thermal diffuse scattering in thin crystal of gold, carried out by Doyle [1969]. One result was that the need to average at each point over the diffuse scattering due to a variety of phonons having the same

components of q perpendicular to the beam has the effect of greatly reducing the dependence of the diffuse intensity on phonon correlation length, so that all assumptions as to correlation length tended to give much the same result as the use of the incoherent addition of intensity given by (36). For this case the oscillations of intensity with thickness, characteristic of the Bragg beams, are almost eliminated. This is in agreement with the experimental observation that dark-field images of wedge-shaped crystals obtained from the pseudo-elastic diffuse scattering (mostly thermal) show thickness fringes of very low contrast (Watanabe [1966]; Castaing [1966]; Cundy, Metherall and Whelan [1966]).

The calculations by Doyle also showed the generation of Kikuchi lines and bands in the thermal diffuse scattering, including the "zero-order" Kikuchi line, often seen in the center of a Kikuchi band, which can only arise from n -beam dynamical scattering (Kainuma and Kogiso [1968]).

12.4.3. Analysis of diffuse scattering

The possibility of obtaining useful information on crystal defects, disorders or perturbations from diffuse scattering in electron diffraction patterns has been considered by several authors. There are obvious limitations because of the formation of Kikuchi line patterns in any diffuse-scattering distribution, but in practice these effects can be largely eliminated by averaging over a small range of angles of incidence (or crystal orientation) since the Kikuchi lines are strongly orientation-dependent. The initial calculations by Fisher [1965] suggested that the intensity of short-range order diffuse scattering in electron diffraction patterns from Cu-Au alloys could be related to the kinematical scattering intensities by a smoothly-varying "dynamical factor". However, it was found that the "size-effect" modification of the diffuse scattering from such alloys could be eliminated by the strong two-dimensional dynamical interactions occurring near principle orientations (see Chapter 16). This, and related observations on thermal diffuse scattering, plus arguments based on phase-grating approximations led Cowley [1965] to propose that the dynamical effects may prove useful in their ability to distinguish between different sources of diffuse scattering. Strong dynamical interactions tend to suppress diffuse scattering due to displacements of atoms but not the diffuse scattering due to atom replacements (short-range ordering, vacancies, substitutional impurities, etc.).

The subsequent analysis of dynamical diffuse scattering by Gjønnes and Høier [1971] has provided a basis for a more complete interpretation of experimental observations, particularly for relatively thick crystals for which the parts of intensity expressions which oscillate with thickness

may be ignored. The application to the study of the defect distribution in non-stoichiometric vanadium oxide is given by Anderson, Gjønnes and Taftø [1974].

A fundamental assumption made in most of the treatments of dynamical diffuse scattering is difficult to justify. This is the assumption that each layer of the structure gives diffuse scattering which is the same as that given by averaging over the whole structure. Since, in the scheme of fig. 12.3, the contribution to the exit wave at any point (and hence the image intensity or the contribution to the diffraction pattern) depends only on the atom positions in a rather narrow column through the crystal, it is assumed that the displacements or replacements of the few atoms of each layer of a column give scattering equal to that from a statistical averaging of the whole crystal. Thus, to calculate the diffraction pattern the averaging is applied to the structure before dynamical scattering is considered, rather than being applied to the exit wave function.

A method for avoiding this assumption has been developed by Cowley and Fields [1979] (also Fields and Cowley [1978]) and appropriate computing methods have been outlined. The concept of a dynamical factor or more appropriately, a "multiplicative dynamical scattering function", is still valid within the approximations of single and weak scattering for the deviations from the average structure. An evaluation of the errors inherent in the assumption made in earlier treatments must rely on detailed calculations for particular cases but these have yet to be made. A recent review of dynamical effects in diffuse scattering has been given by Gjønnes [1993].

12.5. Absorption effects

12.5.1. *The nature of absorption parameters*

We will emphasize repeatedly that the effects to be considered as constituting "absorption" are strongly dependent on the nature of the experimental observation being made. A number of elastic and inelastic scattering processes take place when a beam of radiation interacts with matter. The degree to which the scattered radiation is included in the experimental measurement determines whether a particular scattering process contributes to the measured intensity directly or through the application of an absorption function. For example in neutron diffraction with energy analysis, measurement of the sharp Bragg reflections from a crystal excludes thermal diffuse scattering. The loss of energy from the incident beam and Bragg reflections to give the thermal diffuse intensity provides a small absorption effect.

In the case of X-ray and electron diffraction, much of the thermal diffuse scattering may be included in some

measurements of the Bragg reflection intensity. Then this does not provide any absorption effect. However if the experiment is made in such a way as to separate the sharp Bragg peaks from the attendant thermal diffuse, then an absorption function must be applied in the calculation of the Bragg reflection intensity.

For X-ray or neutron diffraction the absorption function due to thermal diffuse scattering is very small since it enters first in second-order scattering terms and so, unlike the Debye-Waller factor, is negligible under kinematical scattering conditions. Under dynamical scattering conditions for X-rays the probability of double diffuse scattering with appreciable amplitude is also negligible. But, as we shall see, under the conditions for dynamical diffraction of electrons the absorption coefficients due to thermal diffuse scattering may be important.

The influence of the experimental conditions of the measurement of intensities on the nature of the absorption function is probably most severe in electron microscopy of crystals, where the objective aperture transmits a combination of the direct beam with inelastically and elastically scattered beams both in the Bragg spots and the diffuse background. Then these various components are influenced differently by the lens aberrations. This case will be discussed in the next chapter. Here we confine our attention to the case of most universal significance; the absorption functions relevant to the sharp elastic Bragg reflections given by a crystal.

12.5.2. Absorption of X-rays and neutrons

For X-rays and neutrons the major absorption effect is one which usually provides no contribution to the diffraction pattern. Incident X-rays may excite inner-shell electrons from the specimen atoms, losing most of their energy in the process. The characteristic radiation emitted by the excited atoms is normally filtered out. As has been discussed in Chapter 4, the atomic scattering factors for the specimen atoms are thereby made complex by the addition of a real and an imaginary part: $f = f_0 + f' + if''$. The imaginary part is associated with absorption. For the incident beam direction for example, the scattered radiation is added $\pi/2$ out of the phase to give an amplitude, in electron units, $\Psi_0 + if(0)$. Hence $f''(0)$ is subtracted from Ψ_0 and so reduces the incident intensity.

For neutron scattering the absorption component of the scattering factor comes from the imaginary part of equation (4.27). This arises from inelastic interactions of the neutron with the nucleus and results in capture of the neutron by the nucleus with excitation of the nucleus and often emission of secondary radiations which are not normally detected in the diffraction experiment.

Of the other scattering processes which subtract energy from the incident and Bragg beams and so contribute absorption effects, the most important for X-rays are the Compton and thermal diffuse scatterings. The relative contributions from these sources to the average attenuation coefficients for X-rays in crystals have been calculated and compared with experiment by DeMarco and Suortti [1971]. They find that for a variety of elements and X-ray wavelengths, these effects contribute 1 to 3 percent of the absorption coefficient due to inner electron excitation.

12.5.3. "Absorption" for electrons

For high energy incident electron beams the situation is different in that the processes contributing to the diffuse elastic and inelastic scattering observed in the diffraction patterns are the major contributors to the absorption effects. These are the excitation of plasmons, single electron excitations and phonons, with appreciable contributions from short-range order or defect scattering in particular cases.

By an extension of the Schrödinger equation formulation to include excited states of the scattering atoms as in (31), Yoshioka [1957] showed that the effect of inelastic scattering on the elastic scattering amplitudes could be represented by the addition of an imaginary component to the scattering potential and so to the structure factors for centrosymmetric crystals. A number of authors have subsequently estimated or derived from experiment the contributions to this imaginary, absorption component due to the various scattering processes.

The added imaginary parts of the structure factor can be regarded as the Fourier coefficients of a three-dimensional "absorption function", $\mu(\mathbf{r})$. In general, $\mathfrak{S}\{\Phi(\mathbf{u}) + i\Phi'(\mathbf{u})\} = \varphi(\mathbf{r}) + i\varphi'(\mathbf{r})$.

For a thin object we may then write in the phase object approximation for the transmitted wave

$$\begin{aligned}\psi(x,y) &= \exp\{-i\sigma[\varphi(x,y) - i\varphi'(x,y)]\} \\ &= \exp\{-i\sigma\varphi(x,y) - \mu(x,y)\},\end{aligned}\tag{38}$$

where the absorption function $\mu(x,y)$ is the projection in the beam direction of the absorption function $\mu(\mathbf{r})$.

We may use this phase object formulation as a basis for an approximate treatment of the general consideration of absorption in dynamical diffraction because, in the slice methods used for developing dynamical scattering, the basic form of the interaction is introduced in this way.

We deduce, for example, the form of the absorption coefficients to be applied to sharp Bragg reflections when deviations from the perfectly periodic structure give rise to

diffuse scattering. As in Chapter 7 we write the projection of the potential distribution as

$$\varphi(x, y) = \bar{\varphi}(x, y) + \Delta\varphi(x, y),$$

where $\bar{\varphi}(x, y)$ is the average periodic structure. For a thin phase object the transmission function is

$$q(x, y) = \exp\{-i\sigma[\bar{\varphi}(x, y) - \Delta\varphi(x, y)]\}.$$

The sharp Bragg reflections are then given by the Fourier transform of

$$\langle \exp\{-i\sigma[\bar{\varphi}(x, y) + \Delta\varphi(x, y)]\} \rangle$$

where the brackets $\langle \rangle$ indicate a periodic ensemble average over all unit cells. Since $\bar{\varphi}$ is periodic, this may be written

$$\exp\{-i\sigma\bar{\varphi}(x, y)\} \{1 - i\sigma\langle \Delta\varphi \rangle - \frac{1}{2}\sigma^2\langle \Delta\varphi^2 \rangle + \dots\} \quad (39)$$

By definition, $\langle \Delta\varphi \rangle$ is zero. Hence to a first approximation this is

$$\exp\{-i\sigma\bar{\varphi}(x, y)\} \exp\{-\frac{1}{2}\sigma^2\langle \Delta\varphi^2 \rangle\} = \exp\{-i\sigma\bar{\varphi}(x, y) - \mu(x, y)\}, \quad (40)$$

where $\mu(x, y) = \frac{1}{2}\sigma^2\langle \Delta\varphi^2 \rangle$.

Thus the effective absorption coefficient is proportional to the mean square deviation from the average potential. The effective value of the structure factor to be included in n -beam dynamical calculations is $\bar{\Phi}(u, v) - iM(u, v)$ where $\bar{\Phi}(u, v)$ is the Fourier transform of $\bar{\varphi}(x, y)$ and $M(u, v)$ is the Fourier transform of $\sigma^{-1}\mu(x, y)$.

12.5.4. Absorption due to thermal vibrations

For the case of thermal vibrations of the lattice, $\bar{\varphi}(x, y)$ includes the averaging over all displaced atoms so that $\bar{\Phi}(u, v)$ includes the Debye-Waller factor. For each atom the contribution to $\Delta\varphi$ is written using a Taylor series as

$$\varphi_0(x + \varepsilon) = \varphi_0(x) + \varepsilon \frac{\partial\varphi}{\partial x} + \dots$$

where ε is the displacement from the mean position. Then the absorption function is

$$M(u) = \langle \varepsilon^2 4\pi^2 \{uf(u) * uf(u)\} \rangle. \quad (41)$$

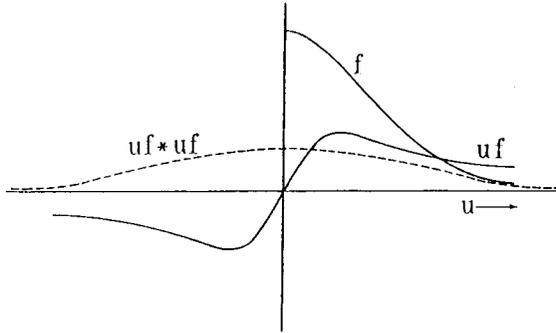


Fig. 12.5. Sketch suggesting the relation between the functions $f(u)$, $uf(u)$ and the self-convolution of $uf(u)$, where $f(u)$ is an atomic scattering factor.

In Fig. 12.5, $uf(u)$ and $uf(u) * uf(u)$ are compared with $f(u)$. It is seen that the imaginary component of the structure factor, $M(u, v)$, in general fall off much more slowly than the elastic part $\Phi(u, v)$ with $|u|$, at least for the angles of scattering usually considered.

The values for the absorption coefficients calculated from (41) are in good agreement with those derived by Hall

Table 12.1

Calculated absorption coefficients due to plasmon scattering (P), thermal diffuse scattering (TDS) and single electron excitations (SE) for 111 systematic reflections of Al at 40 keV (from Doyle [1970]).

Units: $\times 10^{-3} \text{ \AA}^{-1}$				
g	μ^P	μ^{TDS}	μ^{SE}	μ^{total}
000	1.448	0.667	0.235	2.350
111		0.591	0.053	0.644
222		0.440	0.047	0.487
333		0.288	0.038	0.326
444		0.165	0.029	0.194
555		0.081	0.020	0.101
666		0.032	0.013	0.045
777		0.007	0.008	0.015
888		-0.003	0.004	0.001

and Hirsch [1965] using a rather different set of approximations and the Einstein model for thermal vibrations and by Hall [1965] using a many-phonon Debye model. Doyle [1970] derived the values for Al (111) reflections for 40 kV electrons at 300 K as given in Table 12.1.

For static displacements of atoms due to defects or impurities, the same derivation as for thermal vibrations is appropriate. It can usually be assumed that the density of defects is sufficiently small to make this contribution negligible (see Hall et al. [1966]) but the possibility should be born in mind that situations may arise when this is not so, as in cases where the incident electron beam produces intense irradiation damage.

12.5.5. Absorption from electron excitations

Although the argument of equations (38) to (41) may not be strictly appropriate for quantum processes, it does suggest the form of the absorption coefficients to be expected for the important cases. For plasmon excitations the deviations from the average potential have a wavelength of hundreds of Å with little or no modulation with the periodicity of the lattice. Correspondingly the contribution to $M(\mathbf{u})$ is confined to the $|\mathbf{u}| = 0$ Bragg peak. This, like more rigorous theories, gives a uniform absorption, $\mu^P_0 (= M^P(0))$ inversely proportional to the mean free path for plasmon excitation and given by Ferrell [1956] as

$$\mu^P_0 = \frac{1}{\Lambda} = \frac{\theta_E}{a_0} \int_0^1 \frac{xG^{-1}(x)}{x^2 + (\theta_E/\theta_c)^2} dx, \quad (42)$$

where the notation is as for (28) except that the cut-off angle θ_c is equal to $0.74 (E_F/E)^{1/2}$. A similar term due to surface plasmons is relatively small for the thick specimens for which absorption terms become important.

Single electron excitations involving inner shells are highly localized, while those involving outer electron shells will be influenced by the periodicity of the lattice. Hence we might expect the associated part of $\mu(x,y)$ to show the lattice periodicity and the Fourier transform to give $M(\mathbf{u})$ values for all Bragg reflections. The values given in Table 12.1 for μ^{SE}_h (values of $M(h)$) by Doyle were obtained by use of a simple approximation due to Heidenreich [1962]. More sophisticated treatments by Whelan [1965], Ohtsuki [1967], Pogany [1968], Cundy et al. [1969] and Humphreys and Whelan [1969] give much the same sort of result.

12.5.6. Values of absorption coefficients

The figures of Table 12.1 and other calculations show that for the zero beam the ratio of the imaginary part of the structure factor due to absorption to the real part is usually of the order of 0.05 for light elements. The ratio is about 0.03 for inner reflections, increases with increasing scattering angle and then falls away rapidly for larger angles. However the very small or negative values for μ_h due to thermal diffuse at large angles are probably not reliable and presumably are subject to large corrections due to neglect of higher order terms in the expansion of (39) and other effects.

For heavier atoms the ratios of the imaginary to the real parts of the scattering factor are larger, being about 0.10 for the inner reflections of gold crystals, for example. Data on a number of elements is given by Humphreys and Hirsch [1968] and Radi [1970].

A number of further complications may become important when greater accuracy of calculation of absorption coefficients is required. For example, the assumption that any absorption coefficient derived as in (40) for a thin slice can be used in thick-crystal calculations (or equivalent assumptions in other treatments) implies that the deviation from the average periodic lattice is uncorrelated for successive slices. This is not usually the case. Arguments may be made which suggest that the error may not be too great but an adequate treatment taking into account the considerable extent of phonon or lattice defects would be a very complicated business and the effect may not be describable in terms of a simple absorption function picture.

Other sources of absorption may be important for particular samples. It was shown by Cowley and Murray [1968] that short-range ordering of atoms in binary solid solutions, giving rise to short-range-order diffuse scattering, may produce appreciable absorption coefficients in some cases (see Chapter 17).

We have been concerned so far with the one idealized case of the effect of diffuse scattering on sharp Bragg reflections. For comparison with experiment it must be taken into account that a measurement of a Bragg reflection maximum usually includes some thermal diffuse scattering and, unless energy filtering is used, most of the plasmon scattering. For different experimental conditions, the appropriate absorption coefficients are different. This is important for the understanding of observations in electron microscopes when the aperture sizes and aberrations of the lenses have a strong influence on the apparent absorption effects.

This Page Intentionally Left Blank

Electron microscope imaging

13.1. Electron microscopes

13.1.1. Conventional transmission electron microscopes

In the most commonly used form of transmission electron microscope, a thin specimen is illuminated by a near-parallel electron beam of energy between about 100keV and 1MeV and a series of electromagnetic lenses is used to form a highly magnified image of the transmitted electrons. Most commercial instruments operate at 100 to 400keV. The high voltage supply is stabilized to 1 part in 10^5 or better to give sufficient chromatic coherence to allow the assumption of a monochromatic beam to be adequate for most purposes. The effective source diameter for the electron beam is of the order of a few micrometers when a thermally emitting pointed W or LaB₆ tip is used but, increasingly, sharp field-emission tips are employed to give very bright effective sources of diameter as small as 40Å. The illumination of the specimen is controlled by two or more condenser lenses. The beam divergence at the specimen level may be made as small as 10^{-6} radians but, for the high intensities of illumination required for high magnifications, it may approach 10^{-2} radians, especially if the specimen is immersed in the magnetic field of the objective lens to the extent that the pre-specimen field acts as a short-focal-length condenser lens.

The electromagnetic lenses may have focal lengths as small as 1mm or less. The initial stage of magnification by the objective lens is the critical stage for the determination of the resolution and contrast of the image and it is the aberrations and aperture of this lens which must be considered in detail. Two or more additional lenses, the relatively weak intermediate lenses and strong projector lens serve to provide the desirable range of magnifications, from zero up to perhaps 1,000,000 times.

While an electromagnetic lens differs in a fundamental way from a light-optical glass lens because of the vector nature of the magnetic fields, the differences which arise in the imaging properties, such as the rotation and rotational aberrations of the image, are usually minimized by the electron-optical design. For the paraxial imaging properties with which we are most concerned, the considerations and essential theory are much the same as for light optics.

The most important aberration of the objective lens is the third-order spherical aberration which can not be reduced

beyond a certain value for magnetic fields produced by configured iron pole-pieces of cylindrical symmetry. Attempts to reduce this aberration by use of combinations of multi-pole lenses have so far been unsuccessful. Lenses of focal length 1 to 3mm usually have spherical aberration constants of a few mm, limiting the possible resolution to a few Å. However the so-called incoherent limitations to the resolution, including instabilities of the high voltage supply or of the lens current supplies, or mechanical instabilities, may often be significant. Astigmatism of the objective lens, arising from deviations from exact cylindrical symmetry of pole-pieces or from other distortions of the field, is normally corrected in operation by adjustment of the magnetic field by use of electromagnetic stigmator coils.

The magnified images of the specimen may be recorded on photographic plates which have a desirable, linear response to high voltage electrons. Increasingly, two-dimensional CCD detector devices are being introduced to provide quantitative, digitized recording of intensities which may be fed into on-line or off-line data-processing systems. Television-style image-intensifier and display systems are used for convenient visual display.

One very useful feature of the electron microscope optics is the possibility of switching rapidly from the image to the diffraction pattern of a selected specimen area, brought about because the focal lengths of the lenses may be varied simply and rapidly by changing the currents in the exciting coils. Fig. 13.1(a) is a geometric-optics representation of the action of a simple 3-lens magnifying system of objective, intermediate and projector lenses. The Fraunhofer diffraction pattern of the object is formed in the back-focal plane of the objective lens, as in Fig. 3.2. In Fig. 13.1(b), the focal length of the intermediate lens has been increased so that, instead of the image plane, the back-focal plane of the objective is imaged at the object plane of the projector lens and so is projected, magnified, on the fluorescent viewing screen or photographic plate.

If a small aperture is placed in the image plane of the objective lens it has the effect of blocking out all but a small part of the image in the magnifying mode of Fig. 13.1(a). When the intermediate lens is switched to the diffraction mode of Fig. 13.1(b), only those electrons passing through this aperture contribute to the diffraction pattern. This provides the possibility of obtaining a "selected area electron diffraction" (SAED) pattern from any chosen region of the specimen. The SAED method is of great value, especially for crystalline specimens, in allowing a correlation of the image contrast with the diffraction conditions. In practice, the minimum diameter of the area which can be usefully selected in this way is limited by the spherical aberration of the objective lens to about $0.5\mu\text{m}$ at 100keV and perhaps 500Å at 1MeV.

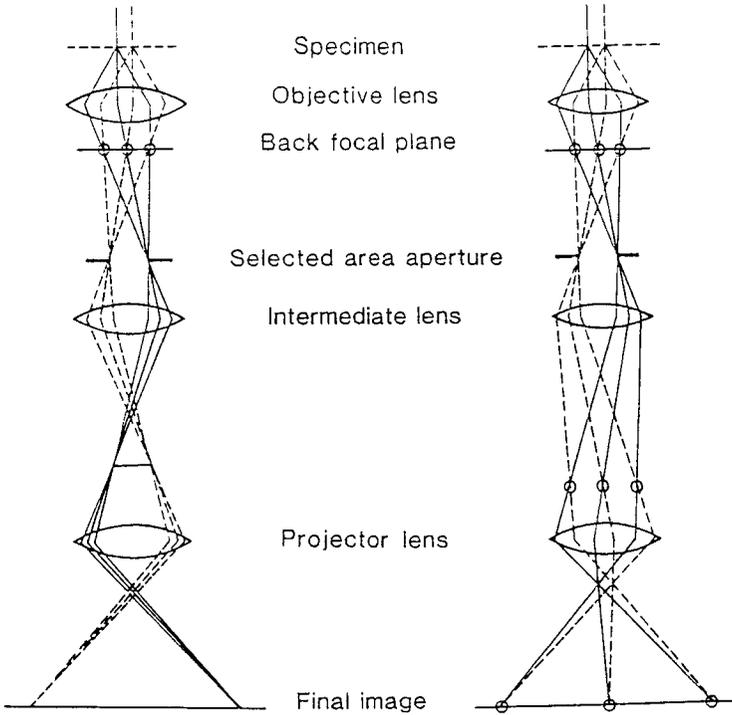


Fig. 13.1. The ray paths in an electron microscope used (a) to produce a high magnification image and (b) to produce a diffraction pattern of a selected area of the specimen.

More detailed accounts of the components and operation of electron microscopes are to be found in various books such as those of Reimer [1993] or Spence [1988]. For a full discussion of the principles and design of electron-optical systems see Hawkes and Kasper, Vols. 1 and 2 [1993].

13.1.2. Scanning transmission electron microscopes

An alternative to the conventional "fixed-beam" transmission electron microscope is the scanning transmission electron microscopy (STEM) instrument illustrated in principle by Fig. 13.2. Here the short-focus objective-type lens is used to form an electron probe of small diameter at the specimen level by demagnification of a small, bright electron source, preferably a cold field-emission tip. The fine probe is scanned across the specimen in a television-type raster and a selected part of the transmitted beam is detected to form the signal which is then displayed on a cathode ray tube to form a magnified image of the object.

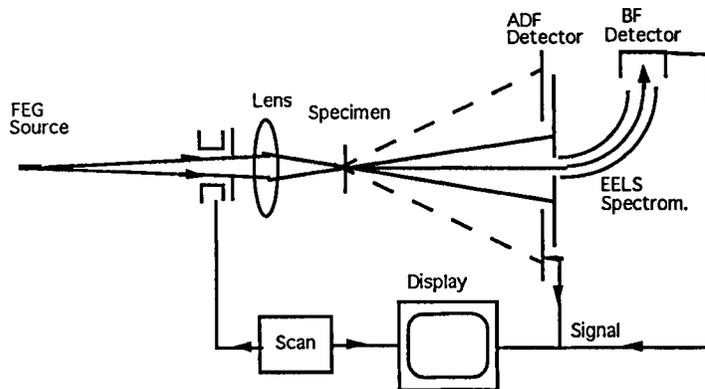


Fig. 13.2. Diagram suggesting the mode of operation of a scanning transmission electron microscope.

A straight-forward application of the reciprocity theorem (Cowley [1969]) shows that, for equivalent lenses, apertures and geometry, the image contrast and resolution for STEM should be the same as for conventional transmission electron microscopy (TEM). This was first shown experimentally by Crewe and Wall [1970] who observed the familiar types of contrast resulting from Fresnel and Fraunhofer diffraction with a STEM instrument operating at about 20kV. However the particular features of the practical design of STEM instruments give them particular advantages and disadvantages relative to TEM instruments so that they have special virtues for particular types of applications and serve to complement the TEM instruments in many respects.

Many of the special capabilities of STEM arise from the ease with which detector configurations can be chosen so that particular portions of the transmitted electron beam array may be detected to form the image. On any plane following the specimen, a convergent beam diffraction pattern is formed. Selection of any part of the central spot, corresponding to directly transmitted electrons, gives a bright-field image. Selection of any one diffracted beam, or any portion of the diffraction pattern outside of the central beam, gives the corresponding dark-field image. In the first high-resolution STEM instruments, built by Crewe and associates (see e.g. Crewe and Wall [1970]) an annular detector was used to detect all electrons scattered outside the incident beam cone, as suggested in Fig. 13.2, to give annular dark-field (ADF) images. These were the first electron microscope images to show clearly the individual heavy atoms sitting on a light-atom support. A number of other special forms of detector have since been used to provide particular imaging characteristics.

The image signal in STEM consists of a time-varying voltage or current which may readily be manipulated electronically to provide on-line control of the image intensity or contrast. Several detectors may be used simultaneously and their signals may be added, subtracted or multiplied to give images emphasizing particular specimen characteristics. Also the signals may readily be recorded for subsequent off-line computer processing for various forms of image enhancement or other manipulation.

If the scanning beam is stopped at any position within the STEM image, the convergent beam diffraction pattern from the corresponding point of the specimen is produced on the detector plane and may be recorded by use of a suitable two-dimensional detector system such as a fluorescent screen viewed by a TV camera or a CCD detector array. The diffraction pattern produced can come from an area of diameter comparable with the image resolution. Patterns have been recorded, at TV rates of 30 per second, from regions less than 3\AA in diameter as the beam is scanned slowly over a specimen area (Cowley et al. [1984]), but recording patterns from 10\AA diameter or larger areas is often preferred.

For thick specimens, STEM has the advantage over the fixed-beam TEM instruments that the resolution is not degraded by the effects of the chromatic aberration of the objective lens on the electrons which have suffered multiple inelastic scattering and hence have a large energy spread (Sellar and Cowley [1973]). Also it is possible to take advantage of the flexibility of the STEM detection system to obtain the greatest possible contrast and signal strength for any particular specimen thickness.

Because in the STEM instrument a very small, intense electron probe strikes the specimen, it is possible to combine the STEM imaging and diffraction modes with microanalysis of very small areas by detecting energy losses due to inelastic scattering processes or the intensities of the secondary radiations produced. An energy analyser is normally provided, as suggested in Fig.3.2. From any specimen area, as small as 10\AA or less in diameter, an energy loss spectrum can be obtained to reveal the plasmon or single-electron excitation processes discussed in Section 12.3. The energy losses due to the excitation of inner electron shells are characteristic of the elements so that the electron energy loss spectroscopy (EELS) signals provide the means for chemical microchemical analysis. If the particular energy loss characteristic of a particular element is selected, the signal produced may be used to form a scanning image which maps the distribution of that element in the specimen. An alternative form of microanalysis is provided by detection of the characteristic X-rays emitted when the specimen is irradiated using energy-dispersive

spectrometry (EDS). If a particular x-ray emission peak is selected, a map of the distribution of the corresponding element is obtained as a scanning image.

The combination of the microanalytical techniques with the STEM imaging has proven to be of great value in materials science and in some areas of biological investigations. The same combination may be achieved, although with more limited spatial resolution, by adding STEM attachments and suitable detectors to conventional TEM instruments. The term Analytical Electron Microscopy (AEM) has become current in recognition of the important expansion of the power of electron-optical techniques which is involved (Hren et al. [1979]).

In addition, for some STEM instruments, and especially for those incorporating ultra-high vacuum systems for the preparation and examination of samples with atomically clean surfaces, valuable signals can be obtained by the detection of low-energy secondary electrons or of Auger electrons. Both of these signals give information concerning the top-most few layers of atoms on a specimen surface and the Auger signals, in particular, are characteristic of the elements present (Hembree and Venables [1992]). The possibilities introduced for obtaining structural and compositional information on solid surfaces with spatial resolutions of the order of 10 Å or better represent an important advance for many areas of surface science.

13.2. Image formation

Because both images and diffraction patterns are readily observed in electron microscopes and their interrelationships are of great importance for many studies of crystalline materials, the most useful approach to the theory of TEM imaging is through the Abbe theory as outlined in Chapter 3. For the usual range of accelerating voltages the diffraction angles are small, in the range of 10^{-2} radians or less, so that the small-angle approximations are adequate for most applications and we will take advantage of their attendant simplicity and clarity.

If the effect of a thin object on the incident wave can be represented by multiplication by a transmission function $q(x, y)$, an incident wave of amplitude unity gives an exit wave $q(x, y)$. The amplitude distribution in the back-focal plane of an ideal objective lens is given by Fourier transform as $Q(u, v)$, where $u = x/f\lambda$ and $v = y/f\lambda$. The amplitude on the image plane is given by further Fourier transform as $q(-x/M, -y/M)$ where M is the magnification. Usually, for convenience, we ignore the factor $-M$ and refer the amplitude back to the object plane so that the amplitude for an ideally perfect lens is written as $q(x, y)$ and the image intensity is $I(x, y) = |q(x, y)|^2$.

Then the effects of the aperture limitation, defocus and aberrations of the objective lens are introduced by modifying the amplitude and phase of the wave on the back-focal plane by multiplication by the transfer function of the lens, $T(u, v)$. The effect of the objective aperture alone is to multiply $Q(u, v)$ by the aperture function $A(u, v)$ which is unity for $u^2 + v^2 < u_0^2$. The image is then smeared out by convolution with $J_1(2\pi u_0 r) / \pi r$, as in equation (3.5).

A defocus of the lens by an amount Δ is represented as in equation (3.20) by multiplying $Q(u, v)$ by the phase factor $\exp\{\pi i \Delta \lambda (u^2 + v^2)\}$. The third-order spherical aberration introduces a fourth-order term so that the most important limitations on the imaging process are introduced by multiplying the amplitude distribution in the back-focal plane by the transfer function,

$$T(u, v) = A(u, v) \cdot \exp\{i\chi(u, v)\}$$

where
$$\chi(u, v) = \pi \Delta \lambda (u^2 + v^2) + \pi C_s \lambda^3 (u^2 + v^2)^2 / 2. \quad (1)$$

The amplitude distribution of the image is then

$$\psi(x, y) = q(x, y) * t(x, y) \quad (2)$$

where $t(x, y)$ is the Fourier transform of $T(u, v)$ and is known as the spread function because the convolution by $t(x, y)$ spreads out the image amplitude, $q(x, y)$, to describe the loss of resolution. The image intensity distribution is then given by

$$I(x, y) = \psi \psi^* = |q(x, y) * t(x, y)|^2. \quad (3)$$

Since both $q(x, y)$ and $t(x, y)$ are complex functions, it is difficult to appreciate the effects of the convolution on the image except for a few simple cases which will be discussed later.

The above description is for purely coherent imaging which is appropriate since, for the usual values of the incident beam convergence, the coherence width of the incident beam is of the order of hundreds or thousands of Å, very much greater than the resolution limit. However in practice the resolution of the microscope is often limited by incoherent factors such as chromatic aberration effects resulting from the energy spread of the incident beam due to the thermionic emission process or high voltage fluctuations or the fluctuations of the objective lens current which vary the focal length of the lens. Also there may be mechanical or electrical instabilities of the microscope which tend to smear out the image. These effects may be represented by convolution of the intensity function (3) by incoherent spread functions.

As a rough guide to the effects of the incoherent factors on the image resolution, it may be assumed that the image spreads may be added in quadrature. If the image of a point object resulting from the coherent imaging as in (3) may be approximated by a Gaussian of half-width d_0 , and if the chromatic aberration and instabilities give Gaussian spreads of d_c and d_i , respectively, then the resulting image may be considered to have a resolution given by d where

$$d^2 = d_0^2 + d_c^2 + d_i^2. \quad (4)$$

In many modern electron microscopes the incoherent factors limiting the resolution have only minor significance. We consider only the coherent imaging case represented by (3) and introduce correction factors for the incoherent effects where necessary.

13.3. Contrast for thin specimens

13.3.1. The phase-object approximation

High resolution electron microscopy usually implies the use of thin specimens since for thick specimens the amount of superimposed fine-scale detail becomes too great and also the resolution is degraded by multiple elastic and inelastic scattering processes. For very thin specimens it is possible to use the thin phase-object approximation, as described in Section 4.2.3. Then the transmission function for the specimen may be written

$$q(x, y) = \exp\{-i\sigma\phi(x, y) - \mu(x, y)\}. \quad (5)$$

The use of the two-dimensional projection functions, $\phi(x, y)$ and $\mu(x, y)$, depends on the assumption that the spread of the electron wave in the sample due to Fresnel diffraction can be ignored. Since the first Fresnel fringe width is given by the square root of $R\lambda$ for a distance R , we may take as a rough guide that the maximum thickness for which the Fresnel diffraction spread of the wave may be ignored for a resolution d is given by

$$H_{max.} = d^2/2\lambda., \quad (6)$$

which is equal to 1200Å for 10Å resolution and 100Å for 3Å resolution for 100keV electrons but may be less than 40Å for current high-resolution imaging.

The absorption function, $\mu(x, y)$, arising from inelastic scattering effects as described in Section 4.2.5, is generally small enough to be neglected as a first approximation. Without this factor, the transmission function (5) is that of a pure phase object. Then if the effect of the

spherical aberration can be neglected, for the in-focus image $T(u, v) = 1$ and $t(x, y)$ is a delta function so that the image intensity given by (3) becomes $|q(x, y)|^2$ which is unity, i.e. the image shows no contrast. It is a common observation that for thin specimens the image contrast is a minimum near the in-focus position. The image contrast increases with defocus and reverses from over-focus to under-focus.

This situation is familiar from optical microscopy of thin biological objects which usually can be approximated as phase-objects. For the traditional light microscopy, the imaging process is close to being incoherent so that, although defocus gives increased image contrast, it also erases all the fine image detail. The most effective way of obtaining good contrast in-focus is by use of phase-shifting methods as in Zernike phase-contrast microscopy. The Zernike method has been applied in electron microscopy with some success (see Unwin [1971]) but is experimentally inconvenient and little used. Most electron microscopy of thin objects depends on obtaining contrast from defocus or from aperture limitation and these contrast mechanisms can be envisaged following the discussions of Sections 3.4.2 and 3.4.3.

From Section 3.4.2, if the absorption term is neglected as being very small and the effects of the spherical aberration is ignored, the intensity for a slightly defocussed image may be written

$$I(x, y) = 1 - (\Delta\lambda/2\pi)\sigma(d^2\phi/dx^2 + d^2\phi/dy^2).$$

The second-differential term in this expression is the two-dimensional form of $\nabla^2\phi$. From Poisson's equation, $\nabla^2\phi = -\rho(xy)/\epsilon_0\epsilon$ where $\rho(x, y)$ is the projection of the charge density distribution, it follows that

$$I(x, y) = 1 + (\Delta\lambda\sigma/2\pi\epsilon_0\epsilon)\rho(x, y). \quad (7)$$

This is the "projected charge density" approximation. The charge density includes both positive and negative charges and so for perfect imaging of stationary atoms the image of an atom would show a sharp positive peak for the nucleus surrounded by a broader negative peak for the electron cloud. However thermal motion of the atom smears the peaks so that a single broadened peak is seen which is dark for under-focus (D negative) and bright for over-focus. The application of this approximation for the interpretation of the images of thin crystals was explored by Moodie and Warble [1967] and Allpress et al. [1972]. It has an important range of validity but does not apply for the higher u, v values where, from (1), the spherical aberration term dominates the phase factor.

13.3.2. The weak-phase-object approximation

For sufficiently thin films containing only light atoms it is possible to make the further simplifying assumption that $\sigma\phi(x,y) \ll 1$. For non-crystalline specimens, or in general when atoms are not resolved, $\phi(x,y)$ is the projection of the mean inner potential and so can be written $\Phi_0 H$, where H is the thickness of the object. Then the condition becomes $\sigma\Phi_0 H \ll 1$. This implies that, for 100keV electrons and a mean inner potential of 10 volts, the specimen thickness must be much less than 100Å. But this would be the requirement if the phase changes in the specimen were compared with that for zero thickness as in the case of an isolated small particle. More usually, the specimen takes the form of a reasonably uniform thin film. A constant phase change can be ignored since it represents only a change of origin of the phase measurement. Then the thickness H can be interpreted as the variation in thickness from an average value or, taking into account compositional variations, $\Phi_0 H$ can be regarded as the deviation of the product from the average value. Then the weak-phase-object approximation (WPOA) can be applied to much thicker specimens, such as biological materials of thickness up to about 200Å or more.

For the WPOA, the expression (5) can be written, ignoring the absorption term and all powers of $\sigma\phi$ except the first, as

$$q(x,y) = 1 - i\sigma\phi(x,y) \quad (8)$$

so that the amplitude in the back-focal plane is, ignoring the aperture function,

$$\begin{aligned} \Psi(u,v) &= [\delta(u,v) - i\sigma\Phi(u,v)].T(u,v) \\ &= \delta(u,v) + \sigma\Phi(u,v). \sin\chi(u,v) - i\sigma\Phi(u,v). \cos\chi(u,v), \end{aligned} \quad (9)$$

so that the image amplitude is $1 + \sigma\phi(x,y)*s(x,y) - i\sigma\phi(x,y)*c(x,y)$, where $s(x,y)$ is the Fourier transform of $\sin\chi(u,v)$ and $c(x,y)$ is the Fourier transform of $\cos\chi(u,v)$, so that, ignoring higher powers of $\sigma\phi(x,y)$, the image intensity is given by

$$I(x,y) = 1 + 2\sigma\phi(x,y)*s(x,y). \quad (10)$$

Thus, if the values of the defocus and aberration coefficient are such that $s(x,y)$ represents a single sharp (negative) peak, the intensity distribution of the image shows a dark spot for a single atom. The image gives a direct representation of the projected potential smeared out as if by a real spread function, $s(x,y)$. The image resolution depends on the width of this spread function and so on the form of the $\sin\chi$ function.

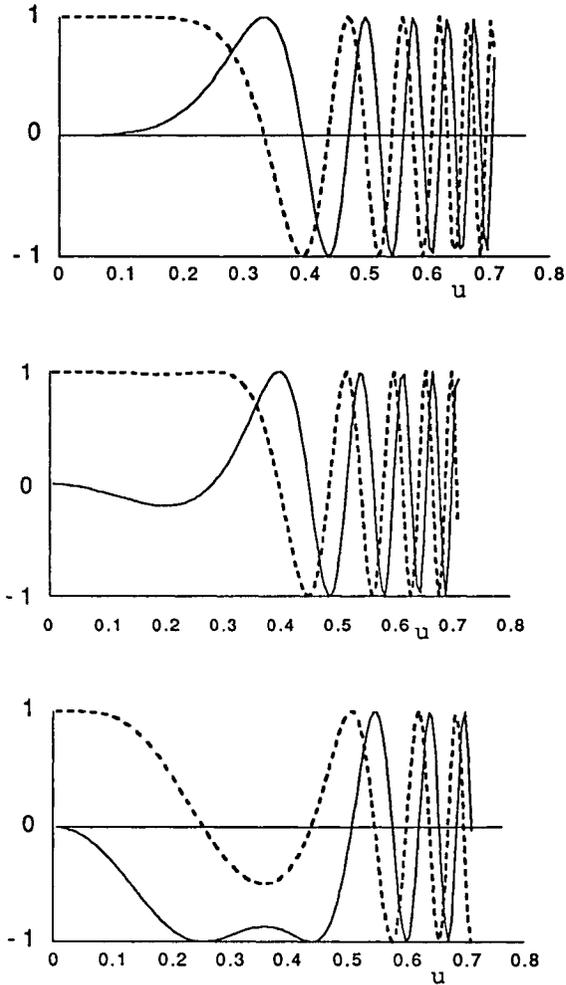


Fig. 13.3. The functions $\sin\chi(u)$ (continuous line) and $\cos\chi(u)$ (dotted line) for 300keV electrons and $C_S = 1.0\text{mm}$ from equation (13.1). The defocus values, Δ are (a) 0, (b) -160\AA and (c) 510\AA (the optimum defocus).

The form of the $\sin\chi(u)$ function for a single radial variable u , (and also the $\cos\chi(u)$), is shown in Fig. 13.3 for various defocus values and a typical C_S value. It is seen that for all the defocus values, $\sin\chi$ is zero for $u = 0$ and is small for small u values. Hence the contribution of all Fourier coefficients in the inner part of the diffraction pattern, corresponding to large periodicities in real space,

is very small so that slow variations of $\phi(x,y)$ are not seen. Ideally, in order to get the best representation of the projected potential distribution of the object, $\sin\chi$ should be close to +1 or -1 for as large a range of u as possible. This is so for a certain optimum defocus value, the so-called Scherzer defocus (Scherzer [1949]). For this value of defocus $\sin\chi$ comes close to -1 for a considerable range of u before going to zero. For greater values of u the value of $\chi(u)$ increase rapidly, as seen from (1) so that $\sin\chi(u)$ oscillates more and more rapidly from +1 to -1. It is customary to use the objective aperture function $A(u)$ to cut off all contributions to the image beyond the first zero point of $\sin\chi(u)$ and consider that the information in the diffraction pattern out to this zero point is properly represented in the image so that this cut-off point represents the resolution limit for the imaging.

The optimum defocus value is obtained by specifying that $d\chi/du = 0$ for χ approximately equal to $-2\pi/3$ so that $\sin\chi$ goes to -1, increases to -0.87 and then goes to -1 again before going to zero. This happens for an optimum defocus value

$$\Delta = -(4C_S\lambda/3)^{1/2} \quad (11)$$

The optimum aperture size is then,

$$u_{\max.} = 1.51 C_S^{-1/4} \lambda^{-3/4} \quad (12)$$

and the least resolvable distance in the image is given by the inverse of this as

$$\Delta x = 0.66(C_S\lambda^3)^{1/4}, \quad (13)$$

which gives $\Delta x = 3.5\text{\AA}$ for 100keV electrons with $C_S = 1.8\text{mm}$; or 1.7\AA for 400keV or 1.2\AA for 1MeV electrons with $C_S = 1\text{mm}$.

The WPOA has been used extensively as a basis for interpretation of images of relatively thin biological specimens including thin sections and macromolecules (proteins, viruses, etc.). For most such samples, the resolution obtained is much poorer than that indicated by (13), largely because observation at very high magnifications is precluded by the severe radiation damage of the specimen by the incident electron beam but also because a small objective aperture is commonly used to enhance the image contrast. On the basis of image interpretation in terms of a simple absorption function, such as is implied by (10), three dimensional reconstructions have been made from multiple images of small objects taken with different orientations of the incident beam, using computerized Fourier transforms of the images to build up a complete three-dimensional reciprocal space model (see deRosier and Klug [1968], Vainshtein [1993]).

The use of a small objective aperture would seem from the above considerations to be counter-productive when used with the WPOA since the value of $\sin\chi$ is very small for low u values. However it may be considered that the effect of the small objective aperture is to prevent most of the scattered radiation from contributing to the image and so to introduce an effective absorption function. On the same basis as for the consideration of the absorption functions derived in the case of inelastic scattering or scattering by disorder or defects in crystals, it was shown by Grinton and Cowley [1971] that, if the defocus is not too great, the appropriate absorption function can be written $\exp\{-(\sigma^2/2) \langle (\Delta\phi)^2 \rangle\}$, where $\langle (\Delta\phi)^2 \rangle$ is the mean square deviation from the value of $\Delta\phi$ averaged over the resolution distance as determined by the small aperture. When an absorption function $\mu(x,y)$ is added, as in (5), the image intensity function (10) is modified to become

$$I(x,y) = 1 + 2\sigma\phi(x,y) * s(x,y) + 2\mu(x,y) * c(x,y), \quad (14)$$

and since $s(x,y)$ is very small for low-angle scattering, the intensity is dominated by the last term. As can be seen from Fig. 13.3, the value of $\cos\chi(u)$ is very close to unity for a large range of defocus for small u values and $c(x,y)$ is relatively large. Thus, by introducing an absorption function into $q(x,y)$ we have redefined $\phi(x,y)$ to be the average over the resolvable distances and have eliminated the need to include the aperture limitation in the transfer function.

13.3.3. Failure of the weak-phase-object approximation

The range of validity of the WPOA has been explored by comparing its predictions with the results of calculations made using one of the formulations of the complete many-beam dynamical diffraction theory, as outlined in Chapters 10 and 11. This was done for various organic and biological crystals by Jap and Glaeser [1980] and for thin crystals containing heavier atoms by Lynch et al. [1972].

At the limits of its range of validity, the WPOA may fail in one of two ways, or both. For light-atom materials the failure may arise because of the assumption of a transmission function for the specimen, i.e., the effect of the specimen on the incident wave can not be represented by multiplication by a function $q(x,y)$. This is the case when the spread of the electron wave within the specimen by Fresnel diffraction becomes important or, what is equivalent, when the curvature of the Ewald sphere in reciprocal space affects the diffraction amplitudes. Then a three-dimensional kinematical diffraction description is more appropriate. For heavier-atom materials, the limitation of the WPOA is more likely to come because the assumption that $\sigma\phi(x,y) \ll 1$ is no

longer valid. Then higher-order terms in the expansion of the exponential in (5) must be included. Some impression of the effects of the failure of the WPOA in the latter case may be obtained by including the second-order terms of this expansion (Cowley [1988]). It is seen that the expressions for the image intensity become much more complicated than in (10) and involve both $s(x,y)$ and $c(x,y)$, so that it is much more difficult to obtain an intuitive idea of the variation of contrast with specimen thickness or other parameters.

13.3.4. Dark-field images

For dark-field images, obtained when the contribution to the image from the directly transmitted incident beam is removed, it has sometimes been assumed that the image intensity is given simply by the amount of scattering matter present at each point of the object. This is a reasonable approximation for some low-resolution images but can not be valid for high resolution images of thin specimens. We demonstrate this in the case of a weak phase object.

In the ideal case that only the directly transmitted beam is removed, the effect is to subtract the $\delta(u,v)$ function from the expression (9) so that the amplitude in the back-focal plane is

$$\Psi(u,v) = \sigma\Phi(u,v) \cdot \cos\chi(u,v) - \sigma\Phi(0,0) - i\sigma\Phi(u,v) \cdot \sin\chi(u,v). \quad (15)$$

where the scattering in the forward direction, $\sigma\Phi(0,0)$ has been subtracted because this is removed when the central incident beam is removed. In order to obtain the most directly interpretable image, it is desirable to make both $\cos\chi$ and $\sin\chi$ constant over as large a range of u as possible. From Fig. 13.3 it is seen that this is best done for a small negative defocus for which $\cos\chi$ is close to unity and $\sin\chi$ is close to zero over a large range. Then the image intensity is given by

$$I_{DF} = \sigma^2 [\phi(x,y) - \phi_0]^2, \quad (16)$$

and so depends on the square of the deviation of the projected potential from its average. Thus a small bright spot in the image may represent either a concentration of atoms or a hole in the specimen. A sinusoidal variation of projected potential produces a sinusoidal intensity variation having half the specimen periodicity.

As a matter of experimental convenience, the idealised case in which only the transmitted beam is stopped is rarely attempted. More commonly, an aperture is used to select some portion of the diffraction pattern which is then centered on the axis of the objective lens and used to form the image.

Then the dark-field image can not be expressed by a simple formula such as (16), but detailed calculations are required for each particular case. When the WPOA fails, the calculations become correspondingly more complicated.

13.4. The imaging of crystals

13.4.1. *Imaging of thin crystals: structure images*

For crystals, as for other materials, the phase object approximation may be used for sufficiently small thicknesses. The most interesting cases are those for which the incident electron beam is exactly parallel to a principle axis of the crystal so that the projected potential distribution is a multiple of the projected potential distribution of one unit cell and, with sufficient resolution, the crystal structure can be visualized directly. In this case, the variation of the projected potential, and consequently the variation of the phase change of the electron wave is much greater than for amorphous materials because the beam may pass through a whole row of atoms or else through none. Hence it is to be expected that the WPOA will fail for much smaller specimen thicknesses. Nevertheless, images were found to give good representations of the arrangements of atoms in thin crystals for the defocus values which are close to those which are optimum for weak phase objects.

Uyeda et al. [1970, 1972] obtained images of thin crystals of Cu-hexadecachlorophthalocyanine showing the shapes of the individual molecules in projection. Iijima [1971] obtained even more striking images showing the arrangement of the metal atoms in thin crystals of $Ti_2Nb_{10}O_{29}$ viewed along the short b -axis of the orthorhombic unit cell, with an under-focus close to 900\AA , the optimum defocus for WPOA imaging. As can be seen by comparing the image of Fig. 13.4(c) with the diagram of the structure, as determined by X-ray diffraction, in Fig. 13.4(b), there is an excellent correlation between the dark spots in the image and the metal atom positions. At the position where there is one heavy atom per unit cell in projection there is a grey spot. Where two atoms in a unit cell are too close together to be resolved, there is a darker spot. Fig. 13.4(d) is an image of the same material taken more recently, at 350keV , (Smith [1991]) with an instrument having a resolution of better than 2\AA so that the close pairs of metal atoms are clearly separated.

For such images, the conditions for the WPOA are clearly not satisfied. Although it is not easy to determine the crystal thickness at each point of the specimen, experience suggests that the thin portions near to the crystal edge may be about 20\AA thick, but the image gives a good representation of the structure for thicknesses up to 100\AA or more and these

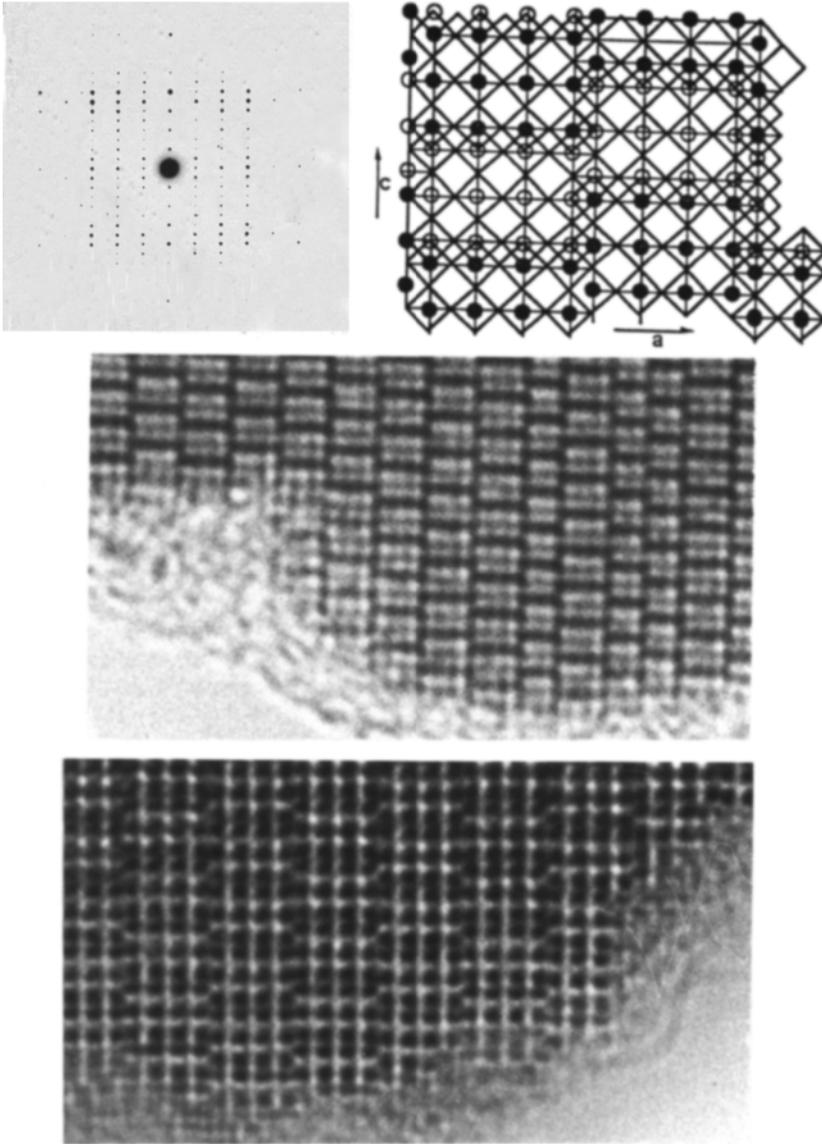


Fig. 13.4. A thin crystal of $\text{Ti}_2\text{Nb}_{10}\text{O}_{29}$, viewed along the short b -axis with $a = 28.5\text{\AA}$, $c = 20.5\text{\AA}$. (a) Diffraction pattern, (b) Diagram of the structure with two layers of metal atoms (dark and light) surrounded by oxygen octahedra. (c) Image at 100keV with resolution approx. 3\AA [Iijima [1971]]. (d) Image at 350keV with resolution less than 2\AA (after Smith [1989]).

favorable imaging conditions apply for both the perfect crystal regions and for regions of crystal defects. For greater thicknesses, there are major variations of contrast so that it is rarely possible to deduce the atom arrangement from the image with any degree of confidence for even the most favorable cases.

On the basis of such favorable initial results, the studies of crystal structures and crystal defects by direct high resolution imaging has been extended to a wide range of inorganic materials. With the relatively poor resolution of the first microscopes used for this purpose (3 to 3.5Å) the structures and defects could be determined for many of the oxide systems based on the juxtaposition of metal-oxygen octahedra as in the ReO_3 -type structures and for many minerals. As the resolution of the electron microscopes improved to better than 2Å, the more densely packed structures such as alloys, semiconductors, sulfides and a much wider range of oxides and minerals became accessible as evidenced, for example, by Buseck et al. [1988] or Spence [1988].

The images of thin crystals which could be directly interpreted in terms of atomic positions became known as "structure images". However it became evident, particularly as the resolution was improved and more complicated structures were imaged with increasing emphasis on the quantitative recording and analysis of image intensities, that considerable care must be exercised in the interpretation of images. Except in the case of very thin, light-atom crystals, the many-beam diffraction effects are present and can complicate the images even if there appears to be a direct correlation between the contrast and the more prominent atom positions. The only reliable way to assure that the interpretation is correct is to make image simulations for the proposed structure models using calculations based on one of the formulations of many-beam dynamical theory, as discussed in Chapters 10 and 11 and Section 13.4.2.

Arguments can be made as to why the images of crystals, viewed along the directions of principle axes, show clear indications of atom positions even when the WPOA clearly fails. A projection approximation, or a "column approximation" (Section 10.6) can be made if the spread of the electron waves by Fresnel diffraction does not exceed the resolution limit of the microscope. Then a column of atoms in the beam direction gives a localized variation of phase of the wave exiting the crystal. The maximum phase change, relative to the positions with no atoms, may be several times π , so that intensity of the image point is given by a complicated formula involving both $s(x,y)$ and $c(x,y)$, but for considerable thicknesses the image of the atom column is a dark spot. Detailed calculations using full many-beam theory (Fejes [1973]) confirm that the contrast for such a column of

relatively heavy atoms increases with thickness at first linearly, as might be expected from the WPOA, and then more slowly until it varies very little with thickness. This is evidently so for such images as Fig. 13.4(c) and (d), obtained from crystal wedges with uniformly increasing thicknesses.

Another factor which is of importance for crystals of moderately heavy or heavy atoms is that the incoming electrons tend to be channelled along the rows of atoms parallel to the incident beam. In a classical particle-picture, it may be considered that a row of atoms forms a linear potential well so that an electron beam entering the opening to the channel tends to be confined by the potential walls and does not spread sideways in the crystal. In the wave-mechanical treatment it may be considered, as mentioned in Chapter 11, that in the solution of the wave-equation for the two dimensions perpendicular to the incident beam direction, there may be bound states in the potential well of an atom row and these represent a propagation of the waves preferentially along the row direction. Hence images showing atom positions as, possibly, either dark or bright spots depending on the thickness and defocus, may be obtained even for relatively thick crystals (see Amelinckx et al. [1993]).

From such considerations it is evident that the image contrast for rows of atoms parallel to the incident beam is very little influenced by the periodicity or otherwise of the arrangements of the atoms in directions perpendicular to the beam. Defects in a crystal which do not disturb the arrangements of the atoms along rows parallel to the beam are then imaged just as clearly as regions of perfect crystal. This is illustrated in Fig. 13.4(c) where some departures from periodicity of the crystal lattice are evident and also in Fig. 13.5 which shows the arrangement of the atoms in a

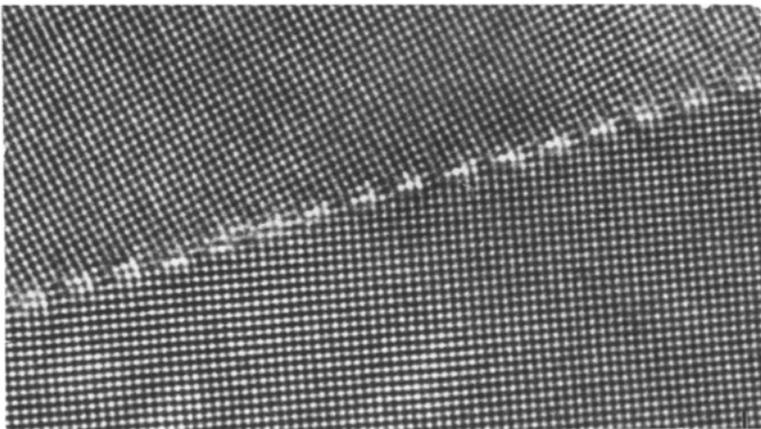


Fig. 13.5 Atomic resolution electron micrograph showing an asymmetric $\langle 001 \rangle$ tilt grain boundary in nickel oxide near $\Sigma 13$ (510). (Merkle and Smith [1987])

tilt grain boundary and Fig. 13.6 which shows the reconstruction on the surface of a thin gold crystal viewed in profile. However it must be recognized that the projection or channelling approximation is at best qualitative and has limitations. Especially for interfaces and for reliable, quantitative image interpretations for defects, it is necessary to compare the observed images with computed simulations made for model structures. The calculations for the non-periodic structures of defects or interfaces can be made by using a large unit cell and the assumption of periodic continuation, as described in Section 11.5 and illustrated in Fig. 13.6 (insert).

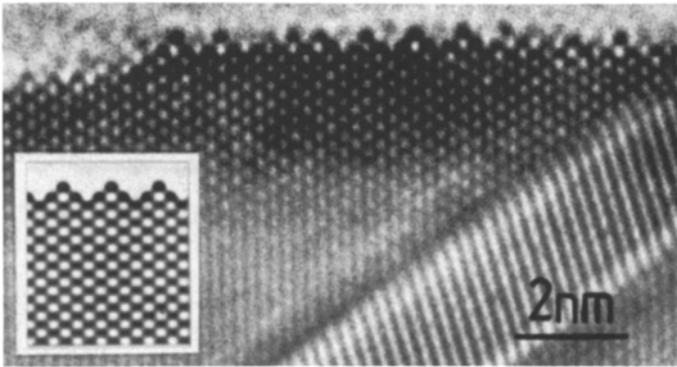


Fig. 13.6 High resolution 'profile' image of a (110) surface on the edge of a thin gold crystal showing a 2x1 reconstruction of the surface. The insert is the image calculated by use of the periodic-continuation multi-slice method. (Marks and Smith [1983])

13.4.2. Calculation of images of crystals: envelope functions

To examine the imaging process in detail and to provide simulated images for comparison with the observations, computer calculations are usually made by use of the multi-slice technique of Section 11.5 (see Goodman and Moodie [1974], Self and O'Keefe [1988]), although, especially for crystals of relatively simple structure such as metals or semiconductors, the matrix methods of Chapter 10 are sometimes to be preferred and the real-space calculations mentioned in Section 10.7 are sometimes preferred for non-periodic samples or for other special purposes. The multislice calculation gives the complex amplitudes for all the "beams", i.e., for the Fourier coefficients of the periodic wave function at the exit face of the crystal. To introduce the effects of defocus and aberrations of the objective lens, these amplitudes are multiplied by the appropriate values of the transfer function,

$T(u, v)$. Then the Fourier series are summed to give the amplitude distribution and hence the intensity distribution of the image. Such calculations were made by Lynch and O'Keefe [1972] and with further developments by Van Dyck [1978] and Ishizuka [1980].

Especially for images obtained with the best possible resolution, the incoherent imaging factors can not be ignored. The most significant of such factors are usually the finite convergence of the incident beam, which in the case of thermal-emission sources can imply a finite, ideally incoherent source, and the variation of the defocus value for the objective lens due to the energy spread of the incident beam, fluctuations of the high-voltage supply and fluctuations of the objective lens current. It was shown, for example by Anstis and O'Keefe [1976] and Ishizuka [1980], that to a good approximation these effects may be incorporated by multiplying the transfer function of the lens by "envelope functions" which have the effect of damping out the higher-order Fourier coefficients, provided that the WPOA is valid. Thus the effect of a spread of defocus values with a standard deviation D , the transfer function is multiplied by $\exp\{-\pi^2 D^2 \lambda^2 u^4\}$. For a spread of incident beam angles up to $u = u_c$, where u is the magnitude of the two-dimensional vector, u , the appropriate envelope function is $2J_1(q)/q$, where the vector q is given by $[2\pi u_c \lambda (Du + \lambda(\lambda C_S - i\pi D^2)u^3)]$. These envelope functions are derived with the assumption that the central beam of the diffraction pattern is very much stronger than any diffracted beam, as is the case for the WPOA and for some particular cases of thicker crystals. They take into account only the variations of the phase differences between the zero and other Fourier coefficients of the image wave. Hence these expressions can not be used when the diffracted beams, other than the zero beam, become strong. Then the effects of interference between different diffracted beams make important contributions to the image amplitudes and the effects of the beam spread and defocus spread on the relative phases of the diffracted beams must be considered. As a better approximation the above envelope functions are replaced by much more complicated "second-order" envelope functions involving terms in $u_i - u_j$ or all possible pairs of diffracted beams, i and j (Self and O'Keefe [1988]). For greater accuracy, especially for large ranges of defocus values or for large convergence angles, it may be necessary to make calculations for a large number of incident beam angles and defocus values and add the resulting intensity distributions.

With increasing use of field-emission guns in electron microscopes, the effects of convergence of the incident beam need to be reconsidered. While the effects of a range of defocus values should be treated as before by the incoherent addition of intensities, the effects of beam convergence must be treated by taking into account the coherence of the incident radiation. If the image amplitudes are calculated

for a variety of incident beam directions, these amplitudes must be added together before squaring to obtain the image intensities. Alternatively, if the convergence of the coherent incident beam is such as to create an incident beam probe of small diameter, the incident beam amplitude at the specimen is described as a small non-periodic object and the calculation of its interaction with the specimen may be made by using the assumption of a large artificially periodic unit cell and periodic continuation, as in the case of calculations for diffraction from non-periodic objects.

Excellent agreement has now been shown between calculations and observations for a great number of images, and for through-focus series of images, of crystals of known and unknown structure. Also many calculations have been made for crystal defects, edges and interfaces using the assumption of periodic continuation. An example of such results is shown in Fig. 13.6 which shows an image and simulation for the edge of a thin gold crystal. (Marks [1984]). The artificial superlattice unit cell was twice the size shown. The number of beams used for such calculations may be 10,000 or more. The rapid development of computer capabilities has meant that calculations of this magnitude no longer represent major difficulties, and calculations with much larger numbers of beams have been made in recent years as in the case of the simulation of images of crystal surfaces and interfaces and also for the calculation of reflection high-energy electron diffraction patterns and reflection electron microscopy images by the methods to be described in Chapter 19.

13.4.3. Imaging of crystals with inelastic scattering

Usually, the imaging of crystals is considered to involve only the elastically scattered electrons. The effects of inelastic scattering are included only in terms of an absorption effect, applied to the elastic scattering, as discussed in Sec. 12.5. However it is obvious that inelastically scattered electrons are also transmitted through the objective aperture and contribute to the image intensity.

For plasmon and single-electron excitations, with energy losses in the range of 5 to 50eV, the scattering is mostly through small angles, of the order of 10^{-3} rad or less. For thin crystals, it may be considered that for each energy-loss process, since the excitation process is delocalized and the energy-loss electrons are incoherent with the elastic scattering, the imaging is given to a good approximation by considering a plane-wave of slightly lower energy, or increased wavelength, being scattered by the crystal and imaged with the corresponding transfer function. Since the transfer function depends on the electron energy most strongly through the variation of the focal length and hence the defocus value, the main effect of a particular energy

loss is to give an out-of-focus image, with no appreciable effect on the resolution. However, when the total broad distribution of energy losses is taken into account, the effect produces an image smeared out by the chromatic aberration effect as a background to the elastic-electron image. The importance of this background image in practice for crystals in the customary range of thickness may be appreciable, as pointed out by Boothroyd and Stobbs [1988]. It may be noted that for scanning transmission electron microscopy, since the objective lens comes before the specimen, the effects of these energy losses are negligible and the energy-loss electrons may be considered to give the same image intensities and resolution as the zero-loss electrons.

For thermal diffuse scattering the case is quite different because the energy losses involved are negligibly small and the scattering is to large angles (see Section 7.3 and 12.1). An appropriate absorption function is applied to the elastic scattering. The contribution to the image from the thermal diffuse scattering is generally small but increases as a larger objective aperture is used to obtain higher resolution.

In general, for a time-dependent scattering system, giving an exit-wave $\psi(\mathbf{r}, t)$, the image intensity is given by

$$I(\mathbf{r}) = \langle |\psi(\mathbf{r}, t) * t(\mathbf{r})|^2 \rangle \quad (17)$$

where the angular brackets denote a time average. Then writing $\psi(\mathbf{r}, t)$ as an average value plus a deviation from the average,

$$\psi(\mathbf{r}, t) = \psi_0(\mathbf{r}) + \Delta\psi(\mathbf{r}, t) \quad (18)$$

the equation (17) becomes

$$I(\mathbf{r}) = |\psi_0(\mathbf{r}) * t_0(\mathbf{r})|^2 + \sum_m \langle |\Delta\psi_m(\mathbf{r}, t) * t_m(\mathbf{r})|^2 \rangle \quad (19)$$

where the subscript m distinguishes the various energy-loss processes.

For thermal diffuse scattering, only one m value is involved. The absorption function to be applied to the elastic scattering is given in the phase-object approximation, as in Section 12.5, as

$$\mu(x) \approx \frac{1}{2} \sigma^2 \langle \Delta\phi^2 \rangle = \frac{1}{2} \langle \epsilon^2 \rangle \left(\frac{\partial \phi}{\partial x} \right)^2 \quad (20)$$

where ϵ is the atom displacement. The contribution to the image intensity from the last part of (19) is then

$$\Delta I(x) = \sigma^2 \sum_i \left\langle \varepsilon_i^2(t) \right\rangle \left| \frac{\partial \phi_i(x - x_i)}{\partial x} * t(x) \right|^2 \quad (21)$$

where the sum over i is the sum over the individual atoms or projections of rows of atoms in the structure. Thus for low resolution, the convolution with a broad $t(x)$ reduces the last term to zero. But for high resolution, when $t(x)$ approaches a δ -function, the image contrast is given by the square of the differential of the projected potential. Thus the atom peaks in the image are somewhat sharper than for the elastic-scattering image.

For thicker crystals this simple result can no longer apply since multiple scattering involving both the elastic and thermal diffuse scattering must occur. However, the above results related to a thin phase object can be taken as a guide to the effects at least for the case where the crystal is in a principle orientation so that the incident electrons are channelled along the rows of atoms and each atomic row can be considered to be imaged separately (see Cowley [1988]). A detailed discussion of the theory and observation of imaging with inelastically scattered electrons is given by Wang [1995].

13.4.4. Lattice fringe imaging

As the resolution of electron microscopes improved, the first observations showing of the periodicities of crystals were those of Menter [1956] who showed patterns of sinusoidal fringes produced by the interference of the transmitted beam with one strong diffracted beam for crystal lattice spacings of about 10 Å. For many years such fringe patterns were the most common indicators of crystallinity, especially for metals and semiconductors having relatively small unit cells. Deviations from the fringe periodicities were used for the detection and characterization of crystal defects. Such fringe patterns are still useful for lattice spacings beyond the resolution limit of the microscope or when the resolution is intentionally restricted as is sometimes the case for imaging accompanying microanalysis. The fringe patterns may be obtained under the diffraction conditions of the classical two-beam case or for any situation where two diffracted beams are predominant or are isolated by use of the objective aperture.

When three or more strong non-collinear, diffracted beams contribute to the image, two or more sets of intersecting fringes are produced. The two-dimensionally periodic image produced is then often similar to a structure-image of a thin crystal, such as discussed in the previous Section. However unless the diffraction conditions are appropriate for structure imaging as in the case of a thin crystal in axial orientation, there is no clear relationship between the image

intensities and the crystal structure. Maxima and minima of intensity can not be interpreted in terms of atom positions. We refer to such fringe images, in one or two dimensions as "lattice fringe images". Their nature and lack of relationship to the structure images may be best illustrated initially in terms of the two-beam approximation.

For the ideal two-beam situation, the amplitudes of the transmitted and diffracted beams at the exit face of a plane-parallel crystal of thickness H , are given from equations (8.17) to (8.27) for the symmetric case, $\theta_0 = \theta_h$, as

$$\Psi_0 = \exp\{-iH\sigma\Phi_0 - H\mu_0 + i\zeta_h H\} \cdot [\cos WH - (i\pi\zeta_h/W)\sin WH], \quad (22)$$

$$\Psi_h = \exp\{-iH\sigma\Phi_0 - H\mu_0 - i\zeta_h H\} \cdot [2\pi i\sigma\Phi_h (\sin WH/W)]. \quad (23)$$

where we have put $\Phi_0 = \Phi(000)$, $\mu_0 = \mu(000)$, $W = 2\pi\sigma\Phi_h(1+w^2)^{1/2}$ and $w = \xi_h\zeta_h$ with $\xi_h = 1/2\sigma\Phi_h$.

The intensity distribution for the in-focus fringes is given by

$$I = \psi\psi^* \quad \text{where } \psi(x) = \Psi_0 + \Psi_h \exp\{2\pi i h x/a\} \quad (24)$$

Then

$$I(x) = \exp\{-2\mu_0 H\} \{1 + \{2w/(1+w^2)\} \cdot \sin^2\{2\pi\sigma\Phi_h H(1+w^2)^{1/2}\} \cdot \cos(2\pi h x/a) + \{1/(1+w^2)^{1/2}\} \cdot \sin\{4\pi\sigma\Phi_h H(1+w^2)^{1/2}\} \cdot \sin(2\pi h x/a)\} \quad (25)$$

At the Bragg angle, $w = 0$, this reduces to

$$I(x) = \exp\{-2\mu_0 H\} \cdot [1 + \sin(4\pi\sigma\Phi_h H) \cdot \sin(2\pi h x/a)]. \quad (26)$$

or, taking the absorption into account explicitly,

$$I(x) = \exp\{-2\mu_0 H\} \cdot [1 - \sin(4\pi\sigma\Phi_h H) \cdot \cos 2\pi(hx/a - \alpha)] \quad (27)$$

where $2\pi\alpha = \tan^{-1}(\sigma\Phi_h/\mu_h)$.

If we assume that $\mu_h \ll \sigma\Phi_h$, the phase angle $2\pi\alpha$ is approximately $\pi/2$ so that the image shows sinusoidal fringes with maximum intensity at one quarter of the distance between the atomic planes. The contrast of the fringes varies sinusoidally with the thickness, H . When $4\pi\sigma\Phi_h H$ becomes equal to π , i.e. for the thickness equal to half the extinction distance, ζ_h , the fringe contrast goes to zero and then reverses.

The more general situations, involving deviations from the exact Bragg condition, were explored in detail by Hashimoto et al. [1961]. They showed that for varying ζ_h and H , the fringes are no longer parallel to the lattice planes. On occasion this can give the false impression that crystal defects are present when, in fact, there is just a local variation in crystal tilt, as in a strain field, or in thickness. From (25) it is seen, for example, that for w

large the intensity distribution is given by

$$I(x) = 1 + (2/w) \sin \zeta_h H \cdot \sin 2\pi(hx/a + \zeta_h H/2\pi), \quad (28)$$

so that a phase factor is introduced that shifts the fringes by an amount depending on both ζ_h and H .

For the two-beam symmetrical case, the zero and diffracted beams make equal angles with the lens axis. The phase factor added by the transfer function is the same for each and disappears when intensities are considered. Hence there is no change of the fringes with defocus. The situation is somewhat similar for the symmetrical three-beam case when the h and $-h$ reflections have the same excitation error and hence the same phase factor due to the transfer function. The image amplitude can then be written, for a centrosymmetrical crystal,

$$\begin{aligned} \psi(x) &= \Psi_0 + \Psi_h \exp\{2\pi hx/a\} \cdot \exp\{i\chi(h/a)\} + \Psi_h \exp\{-2\pi hx/a\} \cdot \exp\{i\chi(h/a)\} \\ &= \Psi_0 + 2\Psi_h \exp\{i\chi(h/a)\} \cdot \cos(2\pi hx/a) \end{aligned}$$

so that the intensity is

$$\begin{aligned} I(x) &= |\Psi_0|^2 + 4|\Psi_h|^2 \cos^2(2\pi hx/a) \\ &+ 2\text{Re}[\Psi_0^* \cdot \Psi_h \exp\{i\chi(h/a)\} \cdot \cos(2\pi hx/a)]. \end{aligned} \quad (29)$$

where Re denotes the real part. There is thus a component of the fringe intensity which has the periodicity a and varies sinusoidally with the defocus, with an origin point for the sine wave dependent on the spherical aberration constant, but with no lateral shift. There is also a component of periodicity $a/2$ which is independent of the transfer function phase factor and hence independent of defocus.

Thus for both the symmetrical two-beam case and the symmetrical three-beam case, there are components of the fringe pattern which are not changed as the defocus is changed. These components are not affected by the spread of defocus values due to the beam energy spread or variations of the objective lens excitation. Hence the envelope function applied to the transfer function for the WPOA case is irrelevant here. Similar arguments can be made for the envelope function due to convergence of the incident radiation. As a result it is a common observation that fringe patterns may be observed with spacings which are much smaller than the resolution limit of the electron microscope as defined for non-periodic objects. Thus fringes corresponding to spacings of less than 1 Å have been observed with microscopes having a nominal resolution of about 3 Å.

Similar arguments may be made for more complicated cases of two-dimensionally periodic objects such as centro-

symmetrical crystals viewed along a principle axis. The interference effects between diffracted beams having the same excitation errors are insensitive to defocus variations and small variations of incident beam direction and so are not reduced by the WPOA-type envelope functions. For this reason it was possible for Hashimoto et al. [1977] to obtain images of thin gold crystals, viewed along the [100] axis, in which there were clearly defined spots at the positions of the rows of gold atoms, 2\AA apart, and showing a fine-structure suggestive of 1\AA resolution or better, even though the the microscope used had a resolution as defined by the Scherzer limit, equation (13), of 3\AA or more. Also, with a similar instrument, Izui et al. [1977] produced pictures of silicon in the [110] direction showing an apparent resolution of the rows of silicon atoms 1.3\AA apart.

Such results suggest that it may be possible to obtain information regarding perfect crystals in principal orientations with a greatly improved resolution relative to what can be done for general specimens. However it is clear from the complications of the situation that the images obtained in this way cannot be interpreted in any direct intuitive manner. The intensities can be calculated for model structures if experimental conditions, such as the crystal thickness and orientation and the instrumental parameters, are defined with very high accuracy but cannot be expected to bear any direct relationship to the potential distribution in the crystal. For images of crystal defects, the information concerning the departures from the crystal periodicity is contained in the diffuse scattering between the Bragg spots in the diffraction pattern so that special relations between the phases of the transfer function can not apply. It has been suggested by Spence et al. [1988] that information concerning defects with effective resolution well beyond the Scherzer limit may be derived, but only for special cases.

13.4.5. Crystal imaging without lattice resolution

If an objective aperture is inserted so that only the central spot or only one diffracted spot is transmitted to form the image, or if the microscope resolution is insufficient to record the interference of diffracted beams, the image of an ideally perfect thin crystal shows no contrast. For crystals having defects or variations of thickness or orientation, the bright-field image shows the variations of the transmitted beam intensity and the dark-field image formed with a particular diffraction spot shows the variation of the diffracted intensity with position. By use of the many-beam dynamical diffraction calculations, described above, the images may be calculated using the column approximation in cases where the diffraction conditions vary slowly with position or using the methods

involving periodic continuation (Section 11.5) for rapidly varying deformations, as in the case of the core regions of dislocations or edges or interfaces of crystals almost parallel to the incident beam. However for most cases it is preferable to obtain images under conditions which make the image interpretation relatively simple, for example by tilting the crystal so that the two-beam approximation is reasonably good.

An account of the dependence of the transmitted and diffracted beam intensities on crystal thickness and tilt for the two-beam case has been given in Chapter 9. This provides a useful basis for the interpretation of thickness fringes and extinction contours in bright- and dark-field images in most cases. The contrast observed for lattice defects is found by use of the column approximation as described in Chapter 10. The progressive changes of lattice displacement in each column are fed into the equations (10.32), (10.33) or, for many-beam situations, (10.35) in order to map out the variations of contrast due to the local strain fields or discontinuities. Systematic procedures for carrying out the computations have been given by Head et al. [1973].

A great deal of experience and empirical knowledge concerning the imaging of extended crystal defects has been accumulated on the basis of two-beam, and sometimes many-beam, dynamical calculations, and this has been of immense value for the understanding of the physical properties of solids in terms of the distributions and interactions of dislocations and faults. This has been the subject of numerous extended treatments in books, such as those of Hirsch et al. [1977], Thomas and Goringe [1979] and Jouffrey et al. [1983], and review articles such as Cowley [1991], Humphreys and Bithell [1992] and Amelinckx and van Dyck [1993]. We do not attempt to include any of this material in this volume except that some results on extended defects are mentioned in Chapter 18.

13.5. Image contrast in scanning transmission electron microscopy

13.5.1. STEM imaging of thin crystals

For those STEM imaging modes which may be directly related to TEM imaging modes by use of the reciprocity relationship it is not necessary to derive expressions for the image contrast. This applies for bright-field imaging with a small detector angle or for dark-field imaging when a small detector picks out one diffracted beam direction. However for the STEM imaging modes, which are of particular interest because there is no convenient equivalent TEM mode, it is desirable to treat the STEM cases separately. Most such modes

take advantage of the possibility of using specially configured detectors in the plane of the diffraction pattern following the specimen. We therefore derive an expression for the intensity distribution in the detector plane, multiply this by a detector function, $D(u)$, and integrate to give the signal intensity for each incident beam position on the specimen.

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, $T(u)$. For convenience we consider only one-dimensional functions. Then the wave amplitude incident on the specimen is given by Fourier transform of $T(u)$ as the spread function $t(x)$, which is multiplied by the transmission function of the specimen, $q(x)$. 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

$$I_X(u) = |Q(u) * T(u) \exp\{2\pi i u X\}|^2. \quad (30)$$

If the detector has a sensitivity function, $D(u)$, the signal detected as a function of the incident beam position is

$$S(X) = \int D(u).I_X(u) du. \quad (31)$$

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 (31) then gives,

$$\begin{aligned} S(X) = & \int D(u).A(u) du + \int D(u).|P(u) * T(u) \exp\{2\pi i u X\}|^2 du \\ & - \int D(u).T^*(u) \exp\{-2\pi i u X\}.[P(u) * T(u) \exp\{2\pi i u X\}] du \\ & - c.c. \end{aligned} \quad (32)$$

If one makes use of the relation,

$$Q(u) * T(u) \exp\{2\pi i u X\} = \exp\{2\pi i u X\} [q(X) * t(X) \exp\{-2\pi i u X\}],$$

the equation (32) can also be written as

$$\begin{aligned} S(X) = & \int D(u).A(u) du + \int D(u).|p(X) * t(X) \exp\{-2\pi i u X\}|^2 du \\ & - p(X) * \{t(X) \int D(u).T^*(u). \exp\{-2\pi i u X\} du\}. \\ & - c.c. \end{aligned} \quad (33)$$

In the expressions on the right hand side of (32) and (33), 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 $p(X)$ and are 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 (33) reduces to

$$S(X) = |q(X) * t(X)|^2, \quad (34)$$

which is identical with the expression for bright-field CTEM with a parallel beam illumination. The specimen transmission function is smeared-out by the spread function and its modulus is squared. In the weak-phase-object approximation, the intensity of the image is given simply as $I(X) = 1 + 2\sigma\phi(X) * s(X)$, where $s(X)$ is the Fourier transform of $\sin\chi(u)$, the imaginary part of the transfer function, so a direct smeared-out representation is given of the projected potential.

For an annular detector which collects all the electrons scattered outside the central beam spot, the product $D(u) \cdot T(u)$ is zero so that only the second term remains in (32) and (33). In general, this term is not easily evaluated. For convenience the approximation is often 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

$$\begin{aligned} S(X) &= \int |P(u) * T(u) \exp\{2\pi i u X\}|^2 = \int |p(X) \cdot t(x-X)|^2 dx \\ &= |p(X)|^2 * |t(X)|^2, \end{aligned} \quad (35)$$

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 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 (Cowley [1976]).

A particular advantage of the ADF imaging mode is that, for the assumption of (35), the intensity of the image peak for a single heavy atom is approximately proportional to $Z^{3/2}$. Hence heavy atoms are readily distinguished from light atoms such as those in a supporting carbon film by a so-called Z -contrast. This mode has proved to be very effective, for example, for the detection of small metal particles embedded in a light-atom support in supported metal catalyst samples.

If $D(u)$ represents an axial, circular detector, smaller than the central beam disc, the intensity distribution in the image can be evaluated by computer calculations (Cowley and Au [1978]). However some indication of the effect can be seen from the form of the last term in (33). 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 contrast 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 (32) or (33), for $D(u) = A(u)$, becomes equal to the negative of (35), so that a bright-field image is given with the same resolution as the ADF dark-field image although with relatively poor contrast because of the constant background intensity from the first term of (32) or (33) (Liu and Cowley [1993]).

In order to improve the bright-field resolution even further, and 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 (33) represents a peaked function much sharper than $t(X)$. It was shown by Cowley [1993], for example, that if $D(u)$ represents a thin annulus, with inside and outside diameters differing by only about 10 percent 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 by Rose [1977] and Kunath and Gross [1985].

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

$$\Psi(u) = T(u) \exp\{2\pi i u X\} * \sum F_h \delta(u - h/a).$$

From this may be derived the intensity distribution which 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 discs, patterns of interference fringes appear. For the moment, however we confine our attention to the mid-points of the regions of overlap. The mid-point of the region of overlap of the discs due to the h and $h+1$ reflections comes at $u = +1/2a$ relative to the h reflection and at $u = -1/2a$ relative to the $h+1$ reflection, so that the intensity at that point is

$$\begin{aligned} |\Psi(u+1/2a)|^2 &= |\Phi_h \exp\{i\chi(+1/2a) + \pi iX/a\} \\ &\quad + \Phi_{h+1} \exp\{i\chi(-1/2a) - \pi iX/a\}|^2 \\ &= |\Phi_h|^2 + |\Phi_{h+1}|^2 + |\Phi_h| \cdot |\Phi_{h+1}| \cos\{2\pi X/a + \alpha\}, \end{aligned} \quad (36)$$

where α 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 (Spence and Cowley [1978]). The reciprocity relation shows that this situation is equivalent to the symmetrical two-beam case for TEM.

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 Φ_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 (Hoppe [1969], Nathan [1976]). The practical difficulties which have prevented the realization of this scheme include the difficulty of obtaining sufficiently thin single crystals of non-trivial structure which are sufficiently resistant to radiation damage by the incident beam.

13.5.2. STEM imaging of thicker crystals

For thicker crystals (thickness greater than 2-5 nm for 100 keV electrons) a simple transmission function, such as the phase-object approximation, can not 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 multi-slice methods. 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 multi-slice calculation is made for an incident beam amplitude described by the spread function, $t(x,y)$. Since $t(x,y)$ is non-periodic, 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.

These methods are well suited for the calculation of the dark-field images obtained with an annular detector (ADF images) and, in particular, for the high-angle annular dark-field (HAADF) images which have been found to be valuable for special purposes. In the usual ADF images, the diffracted beams from a micro-crystalline light-atom support are detected and can obscure the scattering from small heavy-atom clumps. For HAADF imaging (Howie [1979]), the inner diameter of the annular detector is made so large that the contribution of diffracted beams is negligible and only high-angle diffuse scattering is recorded. In Chapters 7 and 12 it was shown that, whereas for elastic scattering the intensities fall off roughly in proportion to the square of the atomic scattering factors, $f(u)$, the thermal diffuse scattering, even in the first-order approximation, depends on $u^2 f^2(u)$, and so falls off much more slowly with scattering angle. Hence the HAADF detector collects mostly thermal diffuse scattering (Wang and Cowley [1990], Zu et al. [1990]). The intensities of the thermal diffuse scattering may be calculated by the methods outlined in Section 12.4.

Recently it has been shown that HAADF mode has important advantages for the high-resolution imaging of interfaces in relatively thick crystal samples (Pennycook and Jesson [1991]). The Z-contrast shows clear indications of differences in composition across the interface. Because the diffuse scattering detected is averaged over the whole of a row of atoms in the beam direction, the image does not show the strong dependences on crystal thickness or orientation that confuse elastic images under strong dynamical diffraction conditions. As a first approximation it may be assumed that the image intensities depend only on the numbers and Z-values of the atoms present, although more detailed calculation (Hillyard and Silcox [1993]) show that departures from this desirable condition may be significant.

13.6. Electron holography

It has been shown above that, at least for weak phase objects, the limitation of resolution of electron microscopes comes from the increasingly large phase changes introduced in

the transfer function for higher scattering angles by the fourth-order term in (1) due to the third-order spherical aberration. From the expression (13) for the optimum "Scherzer" resolution, it is seen that the resolution may be improved by decreasing C_s or λ .

The wavelength may be decreased by using higher accelerating voltages and, in practice, resolution limits approaching 1Å have been achieved by use of electron microscopes operating at up to 1.2 MeV. However limitations arise, apart from the expense of such large machines, in that it becomes increasingly difficult to obtain sufficient electrical and mechanical stability as the voltage is increased. Also the radiation damage of the specimen becomes excessive, even for the usually radiation-resistant specimens such as metals, because the high-voltage electrons can eject atoms from their lattice sites through "knock-on" processes.

The alternative of reducing the spherical aberration constant has been exploited as far as possible but a decrease of C_s is difficult to engineer and is, in any case, not very effective because the coefficient C_s occurs to the one-quarter power in (13). There is a fundamental limit to the value of C_s which can be achieved using magnetic fields generated by lens pole-pieces of cylindrical symmetry. Attempts to correct the lens aberrations by use of combinations of multi-pole lenses (Crewe [1980], Rose [1995]) have as yet met with only limited success in practice.

The concept of electron holography was introduced by Gabor [1948, 1949] as a means for overcoming this resolution limitation, as mentioned in Section 3.6. Gabor proposed that the relative phases of the components of the image wave could be determined by observing the effects of interference with a known reference wave, and then the phases could be corrected for the perturbations caused by the spherical aberration in a process of image reconstruction. Gabor's original proposal was to use a very small, point-like source of electrons, such as is now available in dedicated STEM instruments, and record the greatly magnified shadow-image, or "hologram". He suggested that the reconstruction be made using an optical lens system with appropriate aberration correction (see Fig. 3.4). For the practical realization of this scheme, it was found convenient to digitize the hologram and perform the reconstruction by computer manipulation (Lin and Cowley [1986]). The advantage of the computerized reconstruction is that the real and the imaginary parts of the image function may be derived as separate functions. In later applications of holography, this separate derivation of the real and imaginary parts, or of the amplitude and phase terms, serves a very important function, adding greatly to the power of the holographic process.

The absence of a suitably bright small source of electrons, provided by a FEG, prevented the immediate testing of Gabor's ideas with electrons, but it was soon realized that a plane

parallel incident wave could serve equally well. The out-of-focus image of a thin, weakly-scattering object in a CTEM instrument contains all interference effects between the transmitted wave and the waves scattered by the object, spread by Fresnel diffraction. Holographic reconstructions were made from such images by Haine and Mulvey [1952] and others, although with no obvious improvement of resolution. As with the original Gabor scheme, the desired reconstructed image was accompanied by an unwanted, defocussed conjugate image (3.30). Various, more elaborate schemes have been evolved in order to remove this unwanted background and give a unique, uncluttered reconstruction of the object transmission function.

For a thin object with transmission function which we write as $q(x) = 1 - p(x)$, in order to distinguish the strong transmitted wave, the image intensity is given by

$$|\psi(x)|^2 = |1 - p(x) * t(x)|^2 \quad (37)$$

and for $p(x)$ small, the Fourier transform of this is

$$I(u) \approx \delta(u) - P(u) \cdot T(u) - P^*(u) \cdot T^*(u). \quad (38)$$

Ignoring the aperture function, $|T(u)|^2 = 1$, so that if (37) is multiplied by $T^*(u)$ we obtain

$$I(u) \cdot T^*(u) = \delta(u) - P(u) - P^*(u) \cdot T^{*2}(u). \quad (39)$$

Then the inverse Fourier transform of (39) gives the desired, aberration-free transmission function, $q(x)$, plus the doubly-defocussed and aberrated conjugate image.

For a weak phase object, the expression (38) is given more simply by the Fourier transform of (10) as

$$I(u) = \delta(u) + 2\sigma\Phi(u) \cdot \sin\chi(u) \quad (40)$$

Then multiplying by $(\sin\chi(u))^{-1}$ and taking the inverse Fourier transform gives $1 + \sigma\phi(x)$, as desired. However the function $(\sin\chi(u))^{-1}$ has infinite values for $\sin\chi(u) = 0$. The information for these u -values is lost. To avoid this difficulty a series of images with different defocus values, i.e., a through-focus series, can be used so that information lost from one image is retrieved from another image. This idea was exploited by Thon and Willasch (see Stroke et al. [1974]) to reconstruct images of thin films of carbon with an apparent gain of resolution.

Most thin electron microscope specimens are not weak-phase objects and, for them, more elaborate schemes involving through-focus series of images have been evolved. Thus Kirkland et al. [1985] developed a method based on a maximum a-posteriori recursive algorithm for image reconstruction and obtained considerable improvements in resolution of images of thin crystals of chlorine-substituted copper phthalocyanine

obtained with the Kyoto 500keV microscope. A further successful method is that of Van Dyck and Op de Beeck [1990] who proposed a "focus-variation method" making use of through-focus sets of images with very small focus changes so that expressions could be used that are derived from the differential of the wave function with respect to the defocus. One obvious difficulty with such methods is that because a number of images is required, the exposure of the sample to the electron beam and hence its radiation damage is increased and the requirements for specimen stability are more exacting.

The most widely used form of electron holography, however, makes use of the off-axis configuration suggested by Leith and Upatnieks [1962] (see Section 3.5) in which the reference wave is one which has not passed through the object but has passed only through vacuum. This scheme has the great advantage that the desired image wave may be readily separated from the unwanted conjugate image wave. Also there is no requirement that the specimen should be very thin and scatter weakly.

For electron holography in a CTEM instrument, the arrangement is as suggested in Fig. 13.7(a). Close to the image plane of the objective lens there is inserted an electrostatic biprism, as suggested by Möllenstedt and Düker [1956]. This consists of a very thin wire, charged positively, so that it deflects two parts of the image to overlap. In the absence of a specimen, the overlap of the image waves with a relative tilt produces an extensive pattern of sinusoidal interference fringes. If one of the overlapping waves comes from the vacuum outside of a specimen and the other is modified in phase and amplitude by transmission through a specimen area, the changes in phase in the specimen result in lateral shifts of the interference fringes and changes in amplitude result in changes of the image intensity. The modulated pattern of fringes then constitutes the hologram.

In the reconstruction process, suggested in Fig. 13.7(b), a diffraction pattern of the hologram, or a numerical Fourier transform of a digitized form of the hologram, is produced. In the absence of a specimen, the Fourier transform of the sinusoidal fringes gives a central delta function, $\delta(u)$, and two diffraction spots, $\delta(u-u_0)$ and $\delta(u+u_0)$, where u_0 is the spatial frequency of the fringes. With a specimen present, the usual diffraction pattern of the specimen appears around the $u = 0$ position and patterns corresponding to the Fourier transforms of the image wave function, $\psi(x)$, and its complex conjugate, $\psi^*(x)$, appear around the $u = -u_0$ and $u = u_0$ positions. Separating out the distribution around the $u = -u_0$ position by means of a suitable aperture and Fourier transforming then gives $\psi(x)$, free of any conjugate image. Since the phases as well as the amplitudes of the image wave function are derived, corrections can be made for the phase changes due to the lens aberrations, and hence the object transmission function, $q(x)$, can be deduced.

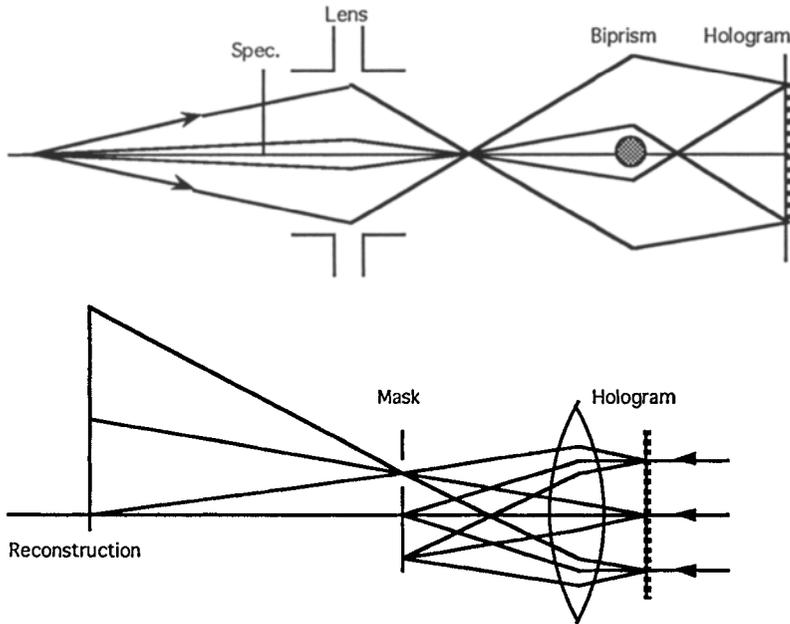


Fig. 13.7 Scheme for off-axis holography in a CTEM instrument. (a) An electrostatic biprism deflects the image wave of the specimen and the reference wave to give a modulated fringe pattern, the hologram. (b) One side-band of the diffraction pattern of the hologram is chosen to give the reconstructed specimen wave.

Mathematically, we may assume that a tilted plane wave, $\exp\{2\pi i u_0 x\}$ is added to the image wave so that the hologram intensity is

$$I(x) = |q(x) * t(x) + \exp\{2\pi i u_0 x\}|^2 \\ = |q(x) * t(x)|^2 + 1 + \{q(x) * t(x)\} \exp\{2\pi i u_0 x\} + c.c. \quad (41)$$

Fourier transforming gives

$$J(u) = \delta(u) + Q(u) \cdot T(u) * Q^*(-u) \cdot T^*(-u) \\ + Q(u) \cdot T(u) * \delta(u+u_0) + Q^*(-u) \cdot T^*(-u) * \delta(u-u_0). \quad (42)$$

Selecting out the term centered on the direction defined by $\delta(u+u_0)$ by use of an aperture, multiplying by $T^*(u)$ and Fourier transforming then gives $q(x)$ directly, with no aberrations and no conjugate image.

The essential requirement for this process is that the Fourier transform of the desired image, centered on $\delta(u+u_0)$, should be well separated from the diffraction pattern of the object, represented by the second term of (42). This implies

that the maximum radius of $Q(u)$, which defines the attainable resolution of the reconstructed image, should be no more than about one third of the distance u_0 , defined by the fringe periodicity. Thus the fringe periodicity must be less than about one third of the least resolvable distance which is sought in the reconstructed image. Since, in digitizing the hologram for computer manipulation, the fringe intensities should be sampled at intervals of about one quarter of the periodicity, the number of pixels required in the digitized hologram must be large. Recording the hologram with a 1028×1028 CCD camera allows a maximum image dimension of only about 100 times the least resolvable distance.

The implementation of this scheme for the enhancement of resolution of images of thin crystals has been described, for example, by Lichte [1991]. As pointed out by this author, one of the main difficulties encountered in this process, or in any process of holographic resolution enhancement, is that in order to make corrections for the aberrations of the imaging system, those aberrations, including the defocus, must be known with an accuracy greatly in excess of that required for the interpretation of normal images. This represents a severe challenge to the experimenter, since the values of C_s and defocus are normally determined by analysis of the images and the accuracy with which they can be derived is consequently limited by the image resolution. One method for overcoming this dilemma is to make use of some fore-knowledge of the nature of the specimen: for example, that it is a thin single crystal in a principal orientation, so that the optimum image must consist of individual isolated peaks corresponding to rows of atoms parallel to the incident beam (Gribelyuk & Cowley [1993]).

An important area of application of the off-axis CTEM holography, not requiring resolution enhancement, is for the visualization and measurement of magnetic and electrical fields in thin specimens. This is possible because the image reconstruction process can give separately the amplitude and phase changes induced by the specimen. Magnetic and electric fields are, in effect, pure phase objects. The phase change can be written as

$$\Delta\phi = (2\pi/h) \int mvd\mathbf{l} - (2\pi e/h) \iint \mathbf{B} \, d\mathbf{S}, \quad (43)$$

where \mathbf{v} is the electron velocity, \mathbf{B} is the magnetic induction, and \mathbf{S} is the area enclosed by the two electron paths which meet to interfere in the hologram. The first term on the right gives the phase change due to electrical fields, including the inner-potential field of the solid specimen, and the second term gives the phase change due to the magnetic field present.

It has been shown, particularly by Tonomura and colleagues (Tonomura [1992]) that graphic displays of the magnetic field distributions showing, for example, the

structures of magnetic domains in thin ferromagnetic films, and accurate measurements of magnetizations in small regions can be obtained with excellent spatial resolutions.

Many other forms of electron holography may be envisaged. More than twenty possible forms have been described (Cowley [1992]; Ru et al. [1992]). For each form involving a CTEM instrument, there is a form, equivalent by the reciprocity relationship, involving a STEM instrument. For each on-line form there is an off-axis form involving the use of a biprism, and there are both bright-field and dark-field forms.

If an electrostatic biprism is inserted close to the gun in a STEM instrument, the effect is to produce two mutually coherent virtual sources so that two coherent probes are formed at the specimen level as suggested in Fig. 13.8(a). Interference between these probes produces a set of fringes

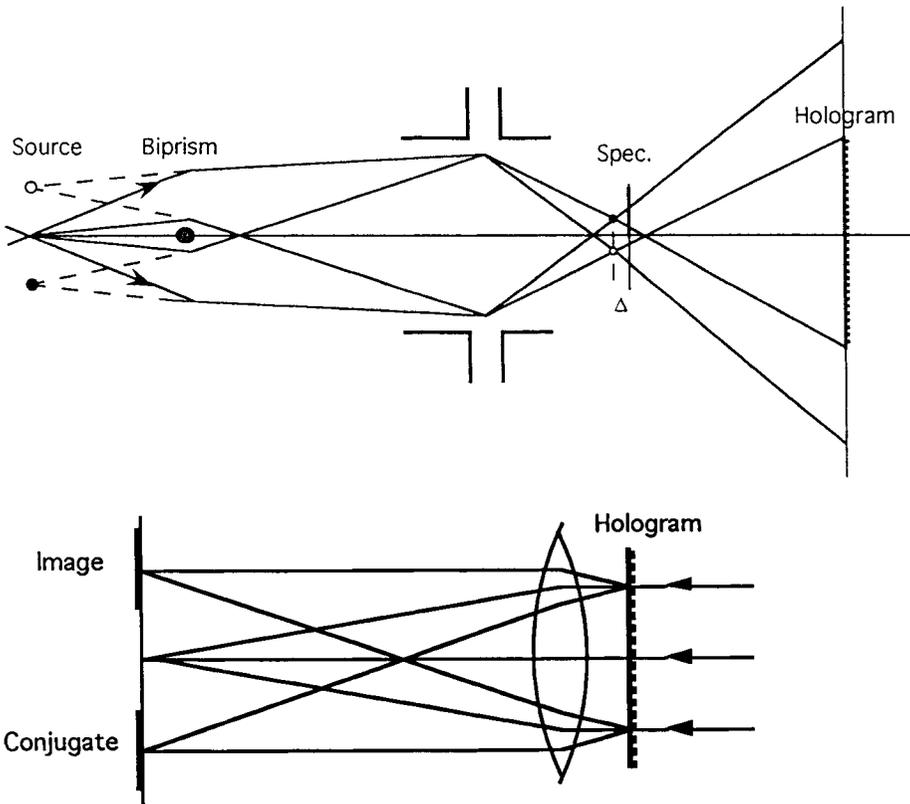


Fig. 13.8 Scheme for off-axis holography in a STEM instrument. (a) A biprism in the illumination system gives two coherent cross-overs in the specimen plane producing a modulated fringe pattern in the diffraction plane. (b) A reconstructed image wave is produced in the back-focal plane when the hologram is illuminated with a plane wave.

in the plane of observation. If one of the probes passes through the specimen and the other one through vacuum the fringe pattern is perturbed. The situation is similar to that for off-axis CTEM holography except that the hologram is formed in the diffraction plane rather than the image plane. Reconstruction from the hologram is made as suggested by Fig. 13.8(b) or by the equivalent computer manipulation of a digitized hologram. A single Fourier transform of the hologram gives three images, one of which is the desired image wave reconstruction.

Mathematically, the wave leaving the specimen may be described, for a separation A of the incident probes, as

$$\Psi(x) = q(x).t(x) + t(x-A),$$

so that the intensity distribution in the plane of observation is

$$I(u) = |Q(u)*T(u) + T(u).exp\{2\pi iAu\}|^2 \\ = |Q(u)*T(u)|^2 + 1 + T^*(u)exp\{-2\pi iAu\}.[Q(u)*T(u)] + c.c. \quad (44)$$

and the Fourier transform of this gives the reconstruction,

$$J(x) = [1 + \{q(x).t(x) * q^*(x).t^*(x)\}] * \delta(x) \\ + t^*(x)*\{q(x).t(x)\} * \delta(x+A) + t(x)*\{q^*(x).t^*(x)\} * \delta(x-A), \quad (45)$$

Separating out the term around $\delta(x+A)$ and multiplying by $t(x)$ gives the function $q(x).t(x)$ for the region defined by $|t(x)|^2$, i.e., for the region of the incident beam intensity distribution. The function $q(x)$ itself is obtained by Fourier transforming, multiplying by $T^*(u)$, and inverse transforming.

When the defocus is small, the image area defined by $|t(x)|^2$ is very small. Then many such small image areas must be processed as the incident beam is scanned over the specimen in order to obtain a reasonable picture area. A thorough analysis of this STEM form of holography has been made by Gribelyuk and Cowley [1992]. As with other forms of off-axis holography, the improvement of resolution which can be achieved in practice is limited by such things as the number of picture elements which can be handled by the available CCD cameras used to record the hologram as well as the accuracies with which the parameters of defocus and spherical aberration can be determined. However, with existing equipment, resolution enhancements by a factor of about 1.6 should be possible.

For large defocus values, the intensity distribution on the detector plane of the STEM instrument resembles a magnified image, rather than a diffraction pattern, the hologram is similar to that for off-axis CTEM holography, and reconstruction is made in much the same way. Although this form of far-out-of-focus off-axis holography is not

particularly suited for high-resolution reconstructions, it can be very effective for the study of the magnetic fields in thin films and small particles with resolutions approaching 10Å (Mankos et al. [1994]).

13.7. Combining high-resolution imaging with diffraction

For the study of the structure of matter by use of an electron microscope, it is customary to consider either the Fraunhofer diffraction pattern, as formed in the back-focal plane of the objective lens, or else the image, formed in the image plane. For other planes in the imaging system, the diffraction and imaging information is combined. In principle, it should be possible to derive much more information if the intensities on these other planes could be interpreted since for coherent imaging conditions these intensities depend on the relative phases of the image waves. In the previous section, the use of holographic reconstructions to derive both the amplitude and phase functions of the wave function at the object exit plane was discussed. In Chapter 15 we will discuss the topic of the so-called ptychography in which phase information in or close to the diffraction plane is derived for all points of the image. Here we consider the more widely explored possibilities for treating only the image plane and the back-focal plane but taking into account the complementary nature of the information gained from these two sources.

Among the first efforts in this direction was the generation of the Gerchberg-Saxton algorithm (see Saxton [1978]) by which the information from the diffraction pattern and image of the same specimen region could be combined in an iterative process. Applications of this have been limited, partly because of the practical difficulty of ensuring that the diffraction information came from exactly the imaged region. Greater success has been achieved when the specimens were limited to thin crystals for which the weak-phase object approximation represented a reasonable approximation.

Unwin and Henderson [1975] considered the problem of determining the structure of the purple membrane of *halobacterium halobium* built of a single layer of molecules. As with most biological materials, the maximum magnification, and consequently the best resolution, which could be obtained in electron microscope images was severely limited by the irradiation damage caused by the incident electron beam. Considerable improvement of the effective resolution could be achieved by taking advantage of the periodic nature of the image intensity. For a very low dose of incident electrons, the images showed only what was apparently random noise, but the optical diffraction pattern obtained from about 10,000 unit cells revealed the presence of the periodicity. Then comparison with the phase-contrast transfer function allowed the relative phases of the reflections to be deduced. When a

selected-area electron diffraction pattern was recorded from a large area of the crystal with insufficient irradiation per unit area to cause appreciable radiation damage, diffraction spots were observed to extend in reciprocal space to a radius corresponding to lattice spacings of 5Å or less.

The diffraction data provided the more accurate values of the intensities, and hence the amplitudes, of the diffracted beams. The image intensities gave information on the relative phases of the diffracted beams but only very approximate values for the relative amplitudes. By combining the information from the two sources, a structure could be derived with an effective resolution of about 7Å, sufficient to define the form of the main components of the molecule, including the arrangement of the alpha helices oriented perpendicular to the plane of the film.

More recently, further extensions of this approach have proved very successful for a number of organic and biological thin crystal films (see Dorset [1995]) and even, to some extent, for thin crystals of inorganic materials (Fu et al. [1994]). Typically, the information in the diffraction pattern extends much further out in reciprocal space than the radius corresponding to the image resolution limit. Several methods have been explored for extending the phase information from that of the high-resolution image to the range of lattice spacings corresponding to the limits of the diffraction pattern. In one approach, Sayre's equation has been used, as described in Section 6.4.4. The structure factors of strong inner reflections, having phases derived from the image, may dominate the summations of equation (6.36) to give probable phases for the outer reflections. Another approach has made use of the maximum-entropy concept (Fryer and Gilmore [1992]) in which the most probable structure consistent with the available data is determined.

In the case of inorganic materials the image resolution may not be limited by radiation effects to the same extent as for organic and biological materials. The question of the limits of validity of the weak-phase-object approximation may well arise but, given sufficiently thin, weakly scattering samples, the phase-extension methods making use of diffraction intensities may provide an improvement of the resolution for crystalline specimens beyond the present limits of high-resolution microscopy, between 1 and 2Å, to the limit of about 0.5Å set by the extent of the diffraction pattern.

Problems

1. From the expressions (13.14) and (13.15), derive the variation of the contrast and the position of the two-beam lattice fringes as a function of crystal thickness and excitation error for the case of no absorption. (c.f. Hashimoto et al. [1961]).

2. Inelastic scattering effects tend to decrease the contrast of thickness fringes. Phonon scattering gives little contrast and the plasmon contribution, although showing strong dynamical effects, is defocussed by the chromatic aberration of the lens. Assuming an average energy loss of 20 eV, a chromatic aberration constant C_C of 2 mm and a loss of resolution given by $\Delta r = C_C \alpha \Delta E / E$, ($E = 100$ kV) find the loss of contrast of thickness fringes of 200 Å periodicity given by the 200 reflection from MgO. How does this affect the derivation of absorption coefficients from thickness fringe intensity variations? What other factors cause errors in absorption coefficients derived in this way?

3. A perfect crystal has flat faces and uniform thickness except that there is a small rectangular protuberance 20 Å high on one face. Assuming 100 keV electrons, a mean inner potential Φ_0 of 20 volts and that the amplitude for the h reflection is 10 volts, find the contrast in the image due to the protuberance, when

(a) no strong diffracted beams occur and the microscope is suitably de-focussed;

(b) the Bragg angle for the reflection h is exactly satisfied? (Use a column approximation and find the bright- and dark-field images assuming that the lattice spacing is not resolved.)

K-line patterns and channelling

14.1. Kossel lines

14.1.1. Geometry of Kossel lines

Diffraction patterns from single crystals produced by divergent radiation and consisting of continuous sets of lines have been observed with X-rays, electrons and other radiations under a great variety of experimental conditions. Kossel lines are formed in diffraction patterns produced by X-rays generated within a single crystal. The somewhat similar "Kikuchi lines" are produced when electrons diffusely and inelastically scattered within a crystal are diffracted by the crystal. These and all the other similar types of patterns to be discussed in this chapter we refer to collectively as "K-line patterns".

In the initial observations by Kossel et al. [1934] and the earlier rather questionable observations of Clarke and Duane [1922], a single crystal was made the anticathode in an X-ray tube. The characteristic X-rays excited by the incident electron beam were then diffracted in the crystal to produce the Kossel line pattern on a photographic plate. A pattern recorded by Voges [1936] from a copper single crystal is reproduced in James [1950]. Similar patterns were obtained by Borrmann [1936] using X-rays rather than electrons to excite the fluorescent X-rays in the crystal. The patterns in these cases were observed on the same side of the crystal as the incident electron or X-ray beam but a transmission geometry is also possible using a thin crystal.

The geometry of the line patterns can be understood by considering the diffraction of radiation emitted by a point source in a crystal as in Fig. 14.1. For a particular set of planes, radiation from the point P makes the Bragg angle θ_h for a particular direction of emission giving the transmitted beam T_h and diffracted beam D_h . The sum of all directions of T_h , where intensity is removed by the diffraction process is given by a cone of half angle $\pi/2 - \theta_h$ and axis perpendicular to the diffracting planes. The sum of all diffracted beams D_h forms a similar cone of directions having increased intensity with oppositely directed axis. The intersections of these cones with a flat photographic plate or film form conic sections: ellipses, parabolas or hyperbolas. An account of the geometry of the patterns has been given by Tixier and Wache [1970].

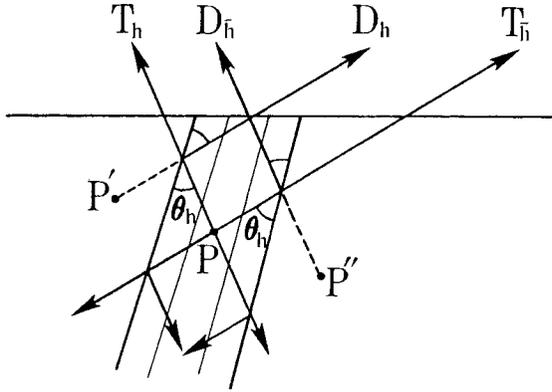


Fig. 14.1. Diagram illustrating the formation of Kossel line patterns with radiation produced at a point P in a crystal.

However from Fig. 14.1 we see that there is a reflection \bar{h} , from the opposite side of the atomic planes also, giving a transmitted beam $T_{\bar{h}}$ parallel to D_h and a diffracted beam $D_{\bar{h}}$ parallel to T_h , and the excess or deficit of intensity observed due to diffraction from the \bar{h} planes tend to be cancelled out. In fact, if X-rays are generated at all points in the crystal with equal intensity, we see that the intensity excess due to D_h is cancelled directly by transmitted radiation from a point such as P' . According to simple kinematical diffraction ideas or a two beam dynamical treatment without absorption, the sum of transmitted and diffracted beams in any direction is always the same and the Kossel lines show no contrast. To predict anything like the observed black-white contrast of the lines we must use a dynamical theory with absorption.

14.1.2. Dynamical theory of Kossel intensities

At first sight the dynamical theory for a point source in a crystal emitting a spherical wave seems formidable. However von Laue [1935] showed how to simplify the problem by use of the reciprocity theorem (Chapter 1). From this theorem, the amplitude at a point A far distant from a crystal due to a point source B within the crystal is equal to the amplitude at the same position B within the crystal due to a point source at A.

Thus the problem of calculating the Kossel pattern is just that of calculating the wave field in the crystal due to a plane incident wave. Since the Kossel pattern is given by the independent emission of characteristic radiation from atoms distributed throughout a volume of crystal, the

intensity of the wave-field in the crystal due to a plane incident wave is summed for all positions of emitting atoms. This can be done making use of the theory of diffraction as developed by von Laue [1935], and later by Zachariassen [1945], Hirsch [1952], and Kato [1955], among others.

The polarizability of the crystal may be written

$$\Psi(r) = \Psi'(r) + i\Psi''(r), \quad (1)$$

where the imaginary part is added to take account of absorption. The real and imaginary parts are periodic and have Fourier coefficients denoted by Ψ'_h and Ψ''_h respectively. For a crystal having more than one type of atom these are not necessarily similar because the ratio of absorption factors for the various atoms may be very different from the ratio of their elastic scattering factors. The distribution of atoms emitting the characteristic radiation is different again and may be denoted by

$$\xi(r) = \sum_h \xi_h \exp(-2\pi i h \cdot r). \quad (2)$$

The expression for the Kossel line intensity K_α as a function of the deviation from the Bragg angle, $\alpha = 2(\theta_B - \theta)$ $\sin 2\theta_B$, is complicated. It is simplified if we consider that the Kossel lines are generated in a crystal which is thin compared to X-ray extinction distances and also if we consider the symmetric case that h and \bar{h} reflections make equal angles with the exit surface. Then, following Cowley [1964] we obtain

$$K_\alpha \propto \left[\xi_0 + \frac{K^2 a_0 \Psi'_h \Psi''_h}{2(K^2 \Psi'^2_h + \alpha^2/4)} (\xi_0 \alpha/2 + K \Psi'_h \xi_h) \right]. \quad (3)$$

Here K is the polarization factor equal to 1 for the normal component and $|\cos 2\theta|$ for the parallel component and a_0 is the maximum value of a thickness parameter $a = \pi t_0 / \lambda \gamma_0$, where t_0 is the actual thickness of crystal and γ_0 is the direction cosine of the incident beam relative to the surface normal.

Thus the intensity is made up of a term symmetric in α , depending on the product $\Psi'_h \xi_h$ and an antisymmetric term with α in the numerator, proportional to the product $\Psi'_h \Psi''_h$. The sum of these terms as suggested in Fig. 14.2 shows a black-white contrast. The antisymmetric part of this contrast is reversed if $\Psi'_h \Psi''_h$ changes sign and so can be used as an indicator of the relative positions within the unit cell of the atoms which are emitting the characteristic radiation giving the Kossel lines. For crystals containing only one

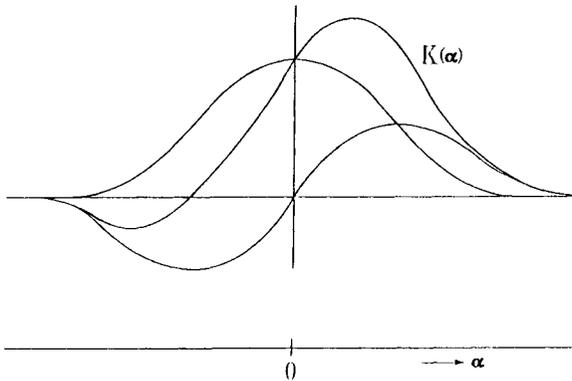


Fig. 14.2. The intensity distribution K_α across a Kossel line shown as the sum of symmetric and antisymmetric parts.

type of atom Ψ'_h , Ψ''_h and ξ_h are of the same sign and the asymmetry is such that the white line on a photographic print (high intensity) are always on the same side of the line, namely on the outside of the cones of diffracted beams.

From equation (3) it is seen that the half width of the symmetric part of the intensity peak is given by $\alpha = 2K\Psi'_h$, which gives values of the order of 10^{-5} radians. The separation of black and white (defect and excess intensity) lines is much the same i.e. of the order of 10^{-3} mm on a film placed 10 cm from the crystal.

14.1.3. Kossel lines with limited resolution

For many observations such as those of Borrmann [1936] the lack of resolution results in an integration over the intensities of the black-white pairs to give a single line which may be either defect or excess, depending on geometric factors. But for many other observations the lines show black-white contrast with an angular separation of the maximum and minimum of 10^{-3} to 10^{-4} radians. The contrast then is not explainable purely in terms of this idealized two-beam dynamical theory. We must add a more realistic consideration of the geometry of the experiments including the finite width of the incident electron beam.

Referring to Fig. 14.1, suppose that the electron beam which excites the characteristic X-rays is centered on P but does not extend to P' or P''. Then using an argument based on simple geometric optics ray diagrams (justifiable because the crystal dimensions involved are sufficient to define directions of propagation and beam locations with the necessary accuracy) we see that the diffracted beam D_h due to the radiation from P is not compensated by a reduced transmission

of radiation from P' . The reduced transmission of the beam $T\bar{h}$ from P is not compensated by a diffracted beam of radiation from P'' . Hence there is a dark line where $T\bar{h}$ intercepts a photographic plate and a white line due to $D\bar{h}$ separated by a distance which depends on the separation of P' and P'' from P . As a refinement of this kinematical treatment we may take into account the attenuation of both transmitted and diffracted beams by absorption and by diffraction processes (the "extinction" effect). Then the relative intensities of these beams depend on the distances they travel through the crystal which are not, in general, equal.

If we then take into account the finite width of the incident electron beam and the generation of X-rays from the volume of crystal defined by the penetration and spread of the electron beam, we see that we get a separation and contrast of the black-white pairs of lines which depends in a rather complicated way on the dimensions of this region of X-ray generation, the dimensions of the crystal, the Bragg angle, the absorption of the X-rays in the crystal and the angle between the diffracting planes and the crystal surface.

We can see why some of the clearest Kossel line patterns have been obtained by use of the configuration suggested in Fig. 14.3(a), where an electron beam is focussed to a small spot on one side of a thin crystal and the Kossel pattern is observed in the transmitted X-rays. From the geometry of this case it is seen that the white (excess intensity) line tends to be on the outside of the cones of radiation and so on the side of the curve on the photographic plate away from the center of curvature. The contrast then depends very little on the positions of the emitting atoms in the unit cell.

The diagram 14.3(a) is appropriate for a near-kinematic case of diffraction from an imperfect crystal. For a thick perfect crystal, Fig. 14.3(b) is more appropriate.

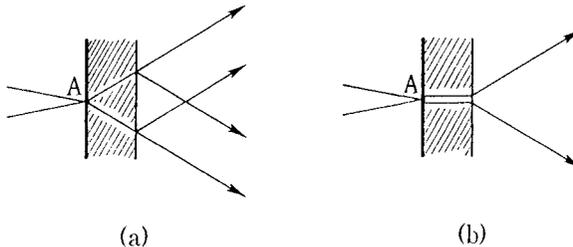


Fig. 14.3. Production of a Kossel-line pattern from a source of radiation, A, outside the crystal for the cases of (a) a thin crystal and (b) a thick crystal.

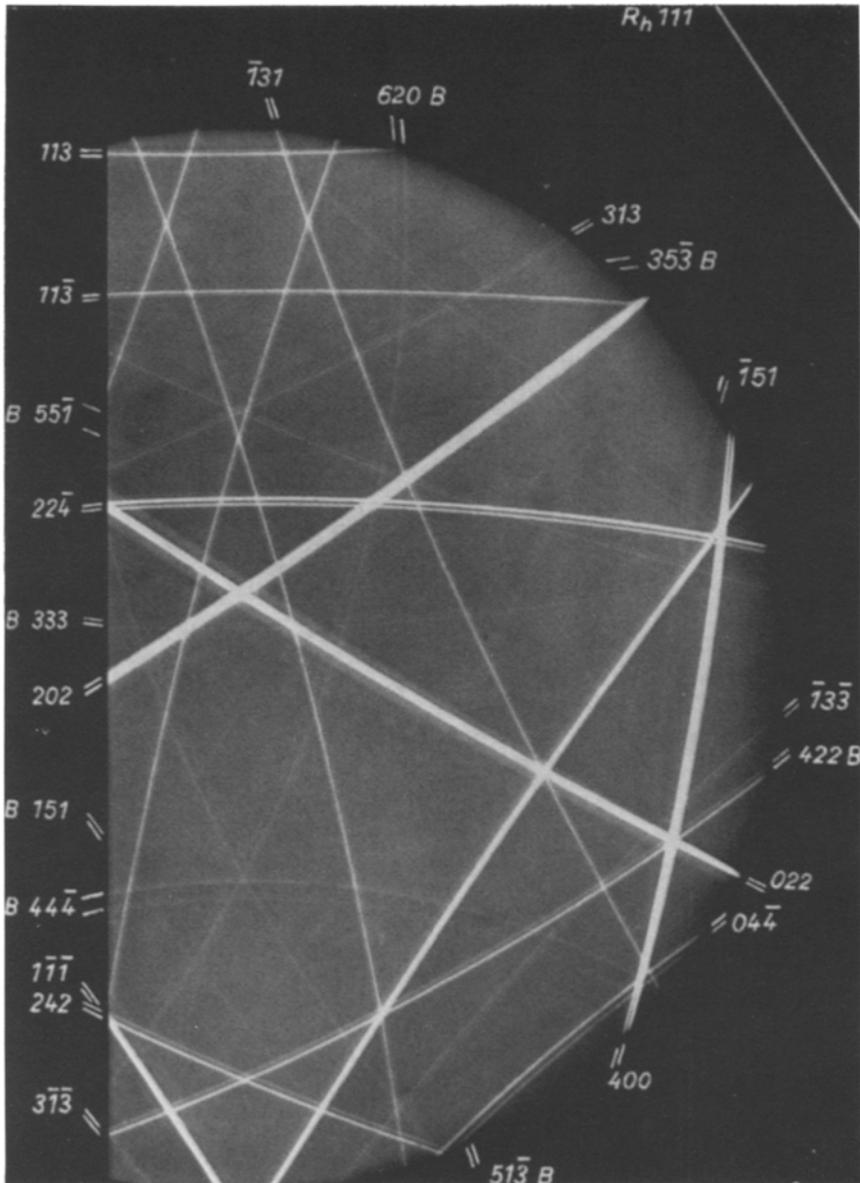


Fig. 14.4. Kossel line pattern from a crystal of germanium, thickness 0.23 mm, produced by placing a source of $\text{CuK}\alpha$ radiation 3 cm from the crystal. A number of transmission lines and one reflection line, $1\bar{1}\bar{1}$, are visible. (From Borrmann [1964].)

As we saw in discussion of the Borrmann effect (Chapter 9) the energy is conveyed through the crystal preferentially along strongly diffracting planes. The transmitted and diffracted beams are equally strong, narrow intensity peaks. Thus in both directions making angles $\pm\theta_h$ with the diffracting planes sharply defined bright lines are produced. The background to the pattern is very low because of the much higher absorption coefficient in nondiffracting directions. A Kossel line pattern produced by a source of radiation considerably removed from a near-perfect crystal of germanium is shown in Fig. 14.4.

Where Kossel lines intersect, two diffracted beams occur simultaneously. For the perfect crystal cases the intensities of the lines are not additive but are given by the complicated 3-beam X-ray dynamical theory (Ewald and Heno [1968]). In particular for the thick-crystal case of Borrmann effect transmission the 3-beam situation can give an even smaller absorption coefficient than the 2-beam case so that bright spots are seen at the intersection points (Borrmann and Hartwig [1965]; Balter et al. [1971]).

14.2. Kikuchi lines

The line patterns first observed by Kikuchi [1928] in the diffuse background of electron diffraction patterns from single crystals are of much the same form as Kossel lines except that, because of the much smaller wavelengths and diffraction angles in the case of electrons, the cones of diffracted rays are very shallow and intersect a photographic plate in what appear to be straight lines (see Fig. 14.5). The other obvious difference is that, while in Kossel lines the $+\theta_h$ and $-\theta_h$ lines are equivalent, each with a close black-white contrast, in the Kikuchi pattern one of the $+\theta_h$ and $-\theta_h$ lines is white and the other black, with the black line always nearest to the central beam of the diffraction pattern. For those symmetrical cases where the $+\theta_h$ and $-\theta_h$ lines are almost the same distance from the central beam, there is often a more complicated "Kikuchi band" structure with an excess or defect of intensity over the whole area between the line pairs and an asymmetry of intensity about the actual line positions.

The diffuse scattering in single-crystal electron diffraction patterns is made up of pseudo-elastic scattering due to thermal vibrations of the atoms, atomic disorder or crystal defects plus inelastic scattering due to electron excitations. For thick crystals it becomes largely multiple diffuse scattering with a broader distribution in angle and in energy. Since the scattering processes render the electron beams incoherent with the incident beam and with each other the diffuse scattering can be considered to be generated

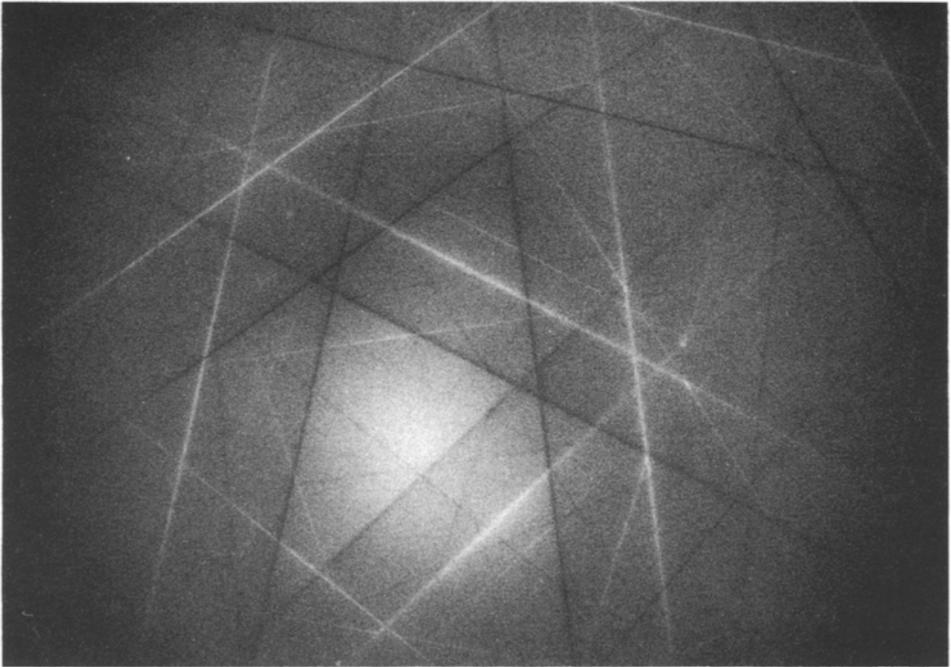


Fig. 14.5. Kikuchi line pattern obtained by transmission of 483 keV electrons through a crystal of aluminum. (From Beauvillain [1970].)

within the crystal. However in contrast to the case of Kossel lines where the emission of X-rays is isotropic, the scattering is strongly peaked in the forward direction of the incident beam. Then a simple kinematical treatment gives a reasonable account of the line-pair intensities. For the case sketched in Fig. 14.6 for example, there is more scattered intensity in the $-\theta_h$ than in the $+\theta_h$ direction. Hence in the $-\theta_h$ direction there is more energy lost by the Bragg reflection than is gained from the beam diffracted from the $+\theta_h$ direction. There is a defect of intensity (black line) in the $-\theta_h$ direction and an excess in the $+\theta_h$ direction (white line).

This simple treatment fails to predict any contrast for the symmetrical case of the Kikuchi bands. For these cases a two-beam dynamical theory with absorption gives an asymmetry of the lines as in the X-ray case. For electrons the magnitude of the structure factors is much greater because of the much stronger interaction of the radiation with atoms. For strong reflections the width of the reflections or the distance between maxima and minima of the asymmetric part of the intensity profile is an appreciable fraction of the distance between $+\theta_h$ and $-\theta_h$ lines and the long tails on the intensity profiles can add up to give a band of excess or

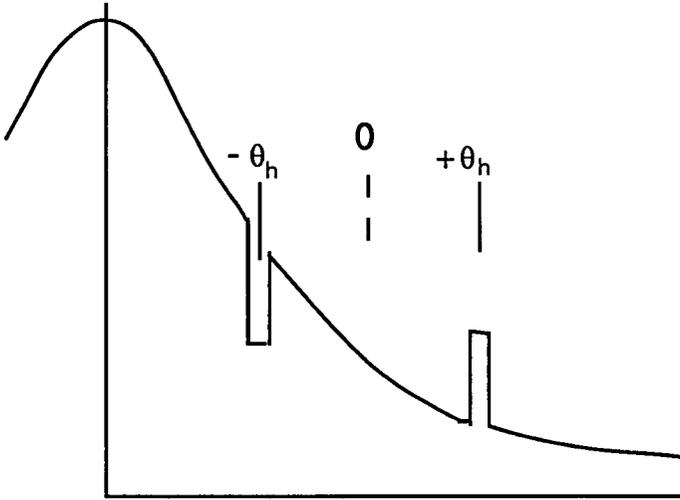


Fig. 14.6. Diagram illustrating the formation of a black-white pair of Kikuchi lines in an electron diffraction pattern.

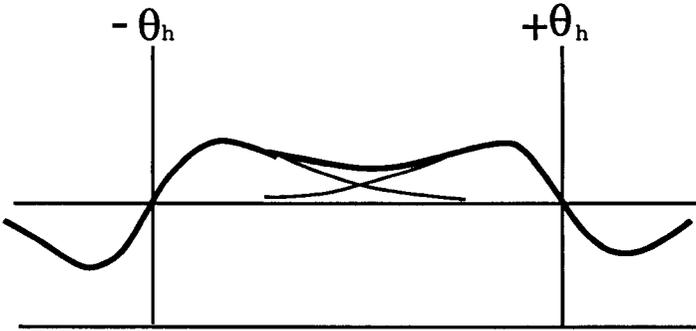


Fig. 14.7. Diagram suggesting the origin of a Kikuchi band with excess intensity between the lines.

defect intensity as suggested in Fig. 14.7. However for a complete account of the intensity profiles an n -beam treatment is necessary, particularly for higher voltage electrons (>100 keV) for which the intensity maximum may actually occur in the middle of the band rather than near a Bragg angle, and the whole intensity distribution due to a systematic row of reflections may be quite complicated. Then when bands or lines intersect, very complicated configurations may result from strong non-systematic dynamical interactions, as seen in the extinction lines of Fig. 9.3.

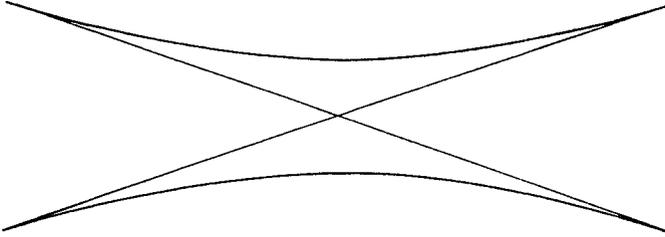


Fig. 14.8. The interaction of intersecting Kikuchi lines due to 3-beam dynamical interactions.

The simplest case of dynamical interaction, and the first n -beam dynamical diffraction effect to be investigated (Shinohara [1932]) is given by the crossing of two well defined isolated Kikuchi lines. The 3-beam interactions in this case may cause the lines to avoid the point of intersection, turning the intersection of two straight lines into the two branches of a hyperbola as sketched in Fig. 14.8.

The complete dynamical calculation of the intensities of Kikuchi patterns is complicated both for formulation and for computation and has been attempted only with simplifying assumptions of very restricted validity such as the assumption of single inelastic scattering only. The problem has been formulated using an extension of the Schrödinger equation method by Pfister [1953] and Kainuma [1953] and extended by Fujimoto and Kainuma [1963] and Okamoto et al. [1971]. A somewhat different, and conceptually simpler approach was initiated by Gjønnes [1966]. On the basis of the ideas of Gjønnes, Doyle has made detailed calculations of both thermal diffuse scattering (Doyle [1969]) and plasmon scattering (Doyle [1971]) using the three-stage calculations outlined in Chapter 12 and showing the generation of Kikuchi lines at least in the thermal diffuse scattering for thin crystals of gold. His results showed clearly the zero-order line, the line appearing in the center of a set of parallel lines due to a set of systematic reflections (Kainuma and Kogiso [1968]). This is a specifically n -beam phenomenon.

From the early days of electron diffraction until the present day the Kikuchi line patterns have been used as a testing ground for electron diffraction theory, as in the studies of n -beam diffraction effects (Shinohara [1932]; Kambe [1957]), the observation and theoretical treatment of the failure of Friedel's Law under conditions of dynamical scattering (Miyake and Uyeda [1955]; Kohra [1954]) and the relativistic variations of n -beam dynamical effects (see Chapter 15). These studies have been made without reference to the fact that the Kikuchi pattern arises from multiple

inelastic and incoherent scattering processes. However, so long as quantitative intensity values are not being considered it seems to be sufficient to take into account that all electrons in the crystal, scattered by any number of processes, have almost the same wavelengths and undergo much the same n -beam dynamical scattering process, so that consideration of elastic scattering of electrons from a single source is sufficient.

14.3. External sources of divergent radiation

For both X-rays and electrons the same sorts of K-line patterns can be generated using sources of divergent radiation outside of the crystal. A small source of X-rays may be formed in a thin metal foil by focussing a beam of electrons on it. Then if the foil forms the wall of a vacuum system a thin crystal of any material may be placed against the foil and a K-line pattern is generated by transmission with the geometry of Fig. 14.3. Or if a crystal, necessarily of larger size, is held a few cm away from the X-ray source a K-line pattern is obtained by back reflection. Patterns of this sort, generated by the use of a converging cone of X-rays with angle restricted to about 60° were obtained first by Seemann [1919, 1926].

In general X-ray line patterns given by an external source, as in Fig. 14.4, are usually known as "pseudo-Kossel patterns". They may be obtained from crystals in which no useful characteristic radiation can be generated or crystals which may deteriorate in vacuum or under electron beam irradiation. For example, Lonsdale [1947] made extensive studies of organic crystals in this way.

For a source of radiation which is not integral with the crystal lattice giving the diffraction effects there is no dependence of the diffracted intensities on the source atom position. The intensity is no longer calculated by use of Laue's method employing reciprocity. The X-ray sources are usually sufficiently far from the diffracting crystal to allow the incident radiation from each point source to be considered as the sum of plane waves for which the intensities are calculated separately and added, using the geometric considerations for either kinematic scattering (Fig. 14.3(a)) or dynamical perfect crystal scattering (Fig. 14.3(b)) or something between these limiting cases.

An interesting observation by Lonsdale [1947] and others is that the contrast of the pseudo-Kossel line patterns is much better for imperfect crystals than for near-perfect ones. This can be understood intuitively on the basis of several combinations of poorly-defined ideas. It may be said, for example, that for perfect crystals the width of the black-white lines is as determined from the two-beam dynamical theory and so is so small that the contrast is lost

when the line is broadened by finite source size effects. Alternatively one could say that for near-perfect crystals there is very strong absorption of the transmitted and diffracted beams by Bragg reflection (extinction) and this reduces the contrast of the diffraction effects.

On the other hand, if a crystal is too imperfect, the range of angular orientations of the individual small regions of the crystal causes a spreading of the Kossel line intensity profile which comes from the geometric effects previously described and a consequent loss of contrast. Hence for best pseudo-Kossel patterns it is necessary to use a crystal having just the right amount of imperfection.

In recent years the development of the electron probe micro-analyser has provided an almost ideal means for the observation of Kossel and pseudo-Kossel lines. In this instrument an intense beam of electrons is focussed to form a spot a few μm in diameter on a specimen to produce fluorescent X-rays which may be analysed to show the composition of the small irradiated area. But by using this small intense X-ray source for the production of K-line diffraction patterns some excellent results have been obtained (Castaing and Guinier [1951]).

For electrons, small sources of radiation outside the crystal can be obtained readily by using a lens to focus an electron beam to a small cross-over. In this way Kossel and Möllenstedt [1939] produced convergent beam patterns (see Chapter 9 and Figs. 9.4 and 9.5) in which the round discs of intensity produced for each diffracted spot are crossed by dark or white lines. These differ from Kikuchi lines (which, incidentally, are produced in the background inelastic scattering of convergent beam patterns) in that there is not the same overlapping of transmitted and diffracted beams.

Such overlapping occurs if the angle of divergence of the incident radiation is increased until the incident beam cone is great enough to include the diffracted beam directions for many reflections. Then the line patterns are somewhat similar to pseudo-Kossel patterns in that within limits the intensity distribution of an incident radiation is uniform. The simple kinematical explanation of contrast for Kikuchi lines based on the rapid fall-off of intensity with angle does not apply.

The kinematical type of explanation of Kossel line contrast as in Fig. 14.3(a) applies to some extent but for electrons the directions of the diffracted beams are not defined with such precision and the dimensions of the crystals are very much smaller. Thus the black-white contrast of the K-lines is seen if the pattern very close to the crystal is greatly magnified by use of an electron microscope, as in the observations of Dupuoy and Beauvillian [1970]. But if the pattern is observed at a large distance (20 to 50 cm) the spread of the diffracted beams due to their natural half-width is sufficient to smear out this black-white structure. The contrast is then due purely to dynamical

interference effects which give a fine structure which has a greater angular width.

A further complication arises if the incident electron beam is not focussed exactly on the crystal. Then any bending, faulting or defects of the crystal will modify the K-line patterns. It was shown by Smith and Cowley [1971] that bending of the crystal causes the separation of a K-line into one strong black line and one white line with a separation proportional to the amount of defocus and the curvature.

14.4. Information from K-line patterns

The useful information concerning crystalline specimens which has been derived from K-line patterns has come almost entirely from considerations of their geometry. While there is obviously some relationship between the contrast of K-lines and the structure amplitudes for reflections from the planes involved, very little has been done toward quantitative interpretation of observed contrast, partly because there are complications due to dynamical diffraction effects and dependence on the experimental conditions, and partly because K-line patterns are most often obtained from simple substances of well-known structure.

The form of the intensity distribution around the intersection of two K-lines depends on the relative phases of the reflections from the two sets of planes involved. Hence, as pointed out by Fues [1949] and others, observations of all available K-line intersections should in principle provide an unambiguous assignment of phases of structure amplitudes and hence a solution to the phase problem of crystal structure analysis. However for a number of reasons, mostly relating to experimental difficulties, no practical use has been made of this idea.

The geometry of K-lines obviously reflects the orientation of the crystal and shows the symmetry of the equivalent sets of atomic planes. When the pattern is produced by a fine electron beam giving X-rays or scattered electrons from a small region of a crystalline sample, it offers a sensitive indication of changes of orientation or of the degree of perfection of an imperfect crystal.

A further application of importance is the derivation of accurate lattice parameter values from observations of the details of the K-line geometry, preferably in regions where n -beam dynamical diffraction effects are small. The initial idea of Lonsdale [1947] for the use of pseudo-Kossel lines was elegant. Occasionally in K-line patterns three lines meet at a point. This is sometimes the inevitable result of crystal symmetry, as when the three planes concerned belong to one cone and the coincidence is independent of wavelength. For other cases it results because the line of intersection of two K-line cones lies on the reflection cone for a third set of planes for one particular wavelength.

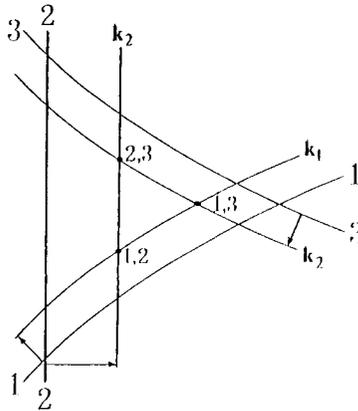


Fig. 14.9. The intersection of three Kossel lines to form a triangle. Lines 1, 2 and 3 are intersections of three crystal planes with the film. Lines K_1 , K_2 and K_3 are the corresponding Kossel lines for one particular wavelength.

For example in the sketch of Fig. 14.9 the lines marked 1, 2 and 3 are the intercepts of three sets of planes on a film. Diffraction with a particular wave-length λ gives the K-lines K_1 , K_2 and K_3 . As the wavelengths and hence the diffraction angles increase the points of intersection of K_1 , K_2 and K_3 move closer together until for a particular wavelength, λ_c , they coincide. Then the observation of a coincidence provides a unique relationship between the angles between the three planes, their interplanar spacings and the wavelength λ_c . If the crystal is cubic and the three sets of planes can be indexed, the one remaining unknown, the lattice constant, a_0 , can be found in terms of the wavelength with an accuracy which Lonsdale estimated as about $\pm 5 \times 10^{-5}$ Å. For a unit cell with more than one parameter to be found, a greater number of coincidences must be observed.

The beauty of this method lies in its simplicity. The accuracy of the results does not depend at all on the geometry of the experimental arrangement. The shape, bending and distance from the crystal of the recording film is arbitrary. The main difficulty lies in the need for the accidental occurrence of at least one 3-fold coincidence for each crystal.

This difficulty is largely overcome by an extension of the method used by Schwarzenberger [1959] and Mackay [1966]. This takes advantage of the fact that pseudo-Kossel line patterns often include K-lines from the $K_{\alpha 1}$, $K_{\alpha 2}$ and $K\beta$ characteristic X-ray emission lines. Hence instead of the set of three intersecting lines as in Fig. 14.9, there are three parallel sets. If the critical wave-length λ_c for exact coincidence of the intersections comes close to these

characteristic wavelengths it may be found with considerable accuracy by simple linear interpolation or extrapolation.

Other related methods for obtaining accurate lattice constants from Kossel or pseudo-Kossel patterns have been summarized by Tixier and Wache [1970]. In a similar way, the intersections of Kikuchi lines in electron diffraction patterns have been used to establish a relation between the lattice constant and the wavelength and so determine one in terms of the other. The ease with which the electron wavelength can be changed by changing the accelerating voltage makes the coincidence method of Lonsdale easy to apply (Høier [1969]).

14.5. Channelling

In recent years considerable interest has been aroused by the observation that fast charged particles show preferential penetration through crystals when the direction of incidence is almost parallel to prominent planes or axes of the lattice. The particles concerned range from protons to the ions of elements with quite high atomic number. Their energies may be as high as many MeV or as low as a few kilovolts. In the preferred directions of transmission through the crystal the number of particles transmitted is normally greater and the average energy loss is less.

Those concerned with the initial observations of this type were accustomed to thinking in terms of collisions of particles. The explanations evolved and developed (J. Lindhard [1965], Lervig et al. [1967]) to explain the observations invoked a picture of the incident particles travelling along channels in the crystals between planes or rows of atoms, oscillating from side to side of the channels as if reflected back and forth by the potential walls created by the atoms of the crystal. For reviews and discussion of channelling phenomena see Datz et al. [1967] and Chadderton [1970, 1973].

From the first, the similarity of channelling patterns recorded for these fast heavy particles with the Kikuchi line patterns given by electrons was apparent and, with time, the observations showed analogies with the whole variety of K-line effects which we have discussed in this chapter. For example, channelling patterns observed on a fluorescent screen when a near-parallel beam of protons is transmitted through a thin single crystal are very much reminiscent of transmission Kikuchi line and band patterns.

Other observations have been made by firing a beam of fast ions at a thin single crystal and measuring the total transmitted intensity as a function of incident angle. A simple application of the reciprocity theorem suggests that this is equivalent to a wide-angle convergent beam experiment in which a beam of high convergence angle strikes the

specimen and the intensity transmitted is observed as a function of the angle of observation. There are similarities also in the production of secondary radiations as will be discussed in the next section.

Various attempts have been made to describe channelling in terms of diffraction phenomena. The observation of anomalous transmission in lattice plane directions is reminiscent of the Borrmann effect. But a moment's reflection shows that the two-beam dynamical theory usually used to discuss the Borrmann effect even for electrons, is quite out of place here. For protons the wavelength is approximately $1/40$ that for electrons of the same energy. At the same time the strength of the elastic interaction with matter, indicated by the value for $\sigma = \pi/\lambda E$, is approximately 40 times as great and the strength of the inelastic scattering is relatively much greater than this. Hence for proton diffraction the thickness of crystal over which coherent diffraction takes place is tens of Å, the number of simultaneous reflections is very great and the Ewald sphere is very nearly planar. Under these circumstances the phase-grating approximation including absorption is very good indeed and of sufficient accuracy to deal with any conceivable observation of diffraction with protons or heavier ions.

The observations which have been reported are for crystal thicknesses very much greater than the range for elastic scattering. The channelling process essentially involves a large number of incoherent inelastic scattering processes. Hence a wave diffraction treatment of channelling would necessarily involve the consideration of elastic scattering processes, treated by use of a phase- and amplitude-grating approximation, combined with multiple inelastic processes.

Also there are important considerations of geometry to be considered. For protons having a wavelength of the order of 5×10^{-4} Å the diffraction angles for simple metals would be around 10^{-4} radians. The available sources of protons are not sufficiently bright to allow beams of divergence less than 10^{-4} radians to be obtained with appreciable intensity. Hence for protons or heavier charged particles diffraction effects are usually not observed.

This point may be restated in terms of the lateral coherence of the incident beam (Cowley [1968]). Diffraction from a crystal lattice is observable only if the incident radiation is coherent in the direction perpendicular to the beam over a distance larger than the inter-planar spacing. The lateral coherence is approximately equal to λ/α where α is the beam divergence. For the proton beams usually employed this is a small fraction of 1 Å.

The lateral coherence of a particle beam may also be considered as a measure of the "size" of an individual particle. Since quantum mechanically it must be said that a

quantum of radiation can only interfere with itself, radiation scattered by two atoms can interfere only if the wave packet which represents the particle overlaps the two atoms at the same time.

Hence for most channelling observations it may be assumed that incident particles are small compared with the distance between lattice planes. Hence a description of channelling in terms of a series of collisions of the incident particles with individual atoms, resulting in preferred transmission by bouncing of the particles along channels in the crystal, is fully justified. The treatment of the phenomena in terms of waves is equally valid and possible in principle, but would be complicated without simplifying assumptions.

The case of very high voltage electrons is an intermediate one since the effects observed are essentially diffraction phenomena, describable in terms of n -beam dynamical diffraction but a semi-classical particle approach gives a reasonable approximation in some cases. A further special consideration is that for electrons the preferred minimum-potential paths through the crystal are close to rows or planes of atoms whereas for positive particles the preferred paths lie well away from the atoms. For electrons there is thus a minimum transmission near principal directions, although small maxima have been observed in the middle of these minima (Kumm et al. [1972]).

The interpretation of these small maxima in terms of the particle channelling model is that electrons travel along paths confined to the potential troughs around rows or planes of atoms, oscillating about the potential minima.

In the wave mechanical treatment, one can separate out the component of the momentum in the forward direction (almost identical for all waves) and then consider the components of the momentum at right angles interacting with, effectively, a two-dimensionally modulated potential (Howie [1966]). The preferential transmission of electrons near atom positions is correlated with the formation of Bloch waves having maxima there (Berry and Mount [1972]). These Bloch waves can be identified with bound states of the electrons in the potential minima, describable in terms of the tight-binding approximation to solution of the wave equation (Kambe et al. [1974]). This type of approach offers a potentially useful means for the interpretation of observations in the high voltage microscopy of crystals.

This type of approach has been further extended by Fujimoto and associates (Fujimoto [1977]; Sumida et al. [1977]) who have used it to characterize a large range of high voltage observations of diffraction patterns and images. The Kikuchi and Kossel patterns given in the range of electron energies up to 2 MeV for monatomic crystals of simple structure show characteristic changes at voltages which vary systematically with the atomic number of the elements present and the radius of the channel through the crystal between the atoms.

14.6. Secondary radiations

In most cases the intensity variations in K-line patterns can be considered to result from variations of the effective absorption coefficient for a dynamical diffraction condition. The absorption process frequently involves the emission of secondary radiation and the intensity of the secondary radiation then fluctuates with the incident beam direction.

A striking example of this may be seen in scanning electron micrographs of single crystal faces as first seen by Coates [1967] and discussed by Booker et al. [1967]. (See also Booker [1970].) The scan of the beam is normally arranged so that the beam passes through the center point of the final lens. Then as the beam sweeps over the surface the angle of incidence on the crystal surface changes. When the beam is near the Bragg angle for a strong reflection the effective absorption coefficient is decreased and the primary beam penetrates further into the crystal. There is a reduction of the number of secondary electrons produced near enough to the surface to allow them to escape and be detected. Hence the image formed by displaying the signal obtained by collecting secondary electrons shows a pattern of lines which is very close to being the inverse of a wide angle convergent beam diffraction pattern obtained by transmission through a single crystal or, at one stage removed, a Kikuchi line pattern.

The intensity of secondary electrons emitted depends on the way in which the primary electrons penetrate the crystal and the way in which they are scattered back towards the surface. A rough theory in terms of the behaviour of individual Bloch waves was proposed by Humphreys et al. [1972] and refined by Humphreys et al. [1973] to give reasonable agreement with experimental observations. A point of particular interest is the possibility that crystal defects at the surfaces of bulk crystals might be detectable through their perturbation of this scattering process. The excitation of the characteristic X-rays of particular atoms within a crystal irradiated by high-energy electrons or other particles or by X-rays forms the basis for various widely-used techniques of compositional analysis. When the crystal is near-perfect so that the diffraction of the incident radiation approximates to an ideal dynamical diffraction (preferably a two-beam diffraction situation), the observation of the emitted characteristic radiation can also give the positions of the emitting atoms relative to the crystal lattice planes. The intensity of the primary radiation, and hence the intensity of the emitted radiation at the position of any atom in the crystal depends on how the nodes or antinodes of the wave-field established in the crystal are located with respect to the diffracting planes and this varies with the angle of incidence of the primary radiation. As suggested by Cowley [1964], observation of the emitted intensity for a rocking of the crystal through the

Bragg angle for several different diffracting planes may provide a basis for the three dimensional location of minority atoms or impurity atoms.

For the Bragg case of diffraction from the surface of a large crystal in reflection, Batterman [1964, 1969] showed that by comparing the observed and theoretical variation of emitted characteristic X-rays with incident angle through the range of strong diffraction by various lattice planes, the positions of emitting atoms could be located with high accuracy. Developments of this idea have been used for the locations of particular atoms relative to crystal surfaces, as will be described in Chapter 19.

The somewhat analogous effect in the case of the transmission of high-energy electrons through single crystals forms the basis for the ALCHEMI (Atom Location by Channelling-Enhanced Microanalysis) developed by Tafto and Spence [1982], which is widely applied in electron microscopes for the analysis of minerals and other inorganic materials. (For a review see Spence [1992]). In general the variation of emitted radiation from the main lattice atoms has a strong anti-symmetrical component as the incident beam angle is rocked through the Bragg angle. An immediate indication as to whether an impurity atom is interstitial or substitutional is given by the observation of whether or not its emitted radiation has the same asymmetry as that of the host atom. More quantitative measurements allow a more exact location of the emitting atoms. Impurity atoms with concentrations as low as 0.01 percent have been located in this way.

Secondary radiations in the form of X-rays, γ -rays or charged particles produced by nuclear reactions may be detected as a means to locate atoms when high-energy charged particles are channelled by single-crystal lattice planes or along rows of atoms. The channelled particles normally travel in between the lattice planes so that impurity atoms lying within the planes of atoms are shielded, but those between the lattice planes are exposed to the incident radiation. For example, Cairns and Nelson [1968] detected the K_{α} radiation of Si and the L radiation of Sb when a proton beam was incident on a silicon single crystal doped with 1 percent of antimony. The fact that the two radiations detected gave the same dip in intensity when the protons were channelled confirmed that the Sb atoms form a substitutional solid solution. For this case it was predicted that one impurity atom in 10^6 could be detected and located.

At the other end of the energy scale, there are the techniques for structure analysis depending on the observation of diffraction intensity distributions for low-energy electrons emitted or incoherently scattered from particular types of atoms. Because the pathlength for low-energy electrons is very small, these methods are of value for the analysis of the top few monolayers of atoms on crystal surfaces and are described in that context in Chapter 19.

This Page Intentionally Left Blank

Applications of dynamical effects in single crystals

15.1. Dependence of dynamical effects on crystal parameters

As can be seen from the discussions of dynamical scattering effects in Chapters 8 to 11, the intensity of a diffracted beam given by a near-perfect single crystal may be a strongly-varying function of the structure amplitude for the reflection, the crystal thickness in the beam direction, the orientation of the crystal relative to the incident beam and the form and magnitude or frequency of recurrence of deviations from the crystal periodicity. Also, less directly, the intensity depends on the other conditions of the scattering process including the temperature and the presence of absorption or inelastic scattering processes. It follows that observations on dynamical diffraction intensities may be used to obtain measures of any of these quantities or effects with high precision provided that the others are sufficiently well controlled. Recently a number of techniques have been devised whereby dynamical effects are used to obtain data of value in a variety of areas of science and technology.

The angular width of a diffracted X-ray beam from a thick crystal is proportional to the structure amplitude and is typically of the order of a few seconds of arc or 10^{-5} radians. Hence very small deviations in orientation of the lattice planes may cause large changes of intensity. It has been known for many years that a slight strain applied to a thick crystal, such as caused by the thermal gradient due to a finger placed near the crystal, will change the diffracted intensity by some orders of magnitude. The methods of X-ray topography, developed by Lang [1959], have been used to record the strain fields due to isolated dislocations or impurity aggregates in near-perfect crystals. The apparatus used is as sketched in Fig. 9.8. A review of recent results obtained with this technique has been given by Lang [1970], and Authier [1970] has reviewed the theory of contrast of the images.

15.2. X-ray interferometry

Measurements of high precision have been made possible by the development of X-ray interferometers by Bonse and Hart [1965, 1970]. A number of variants of these instruments have been

developed employing both transmission and reflection of the X-ray beam from near-perfect crystals of silicon or germanium. The first and simplest is that sketched in Fig. 15.1. Three near-perfect thick crystals are arranged to be exactly parallel. The transmitted and diffracted beams given by the first crystal are diffracted again by the second crystal to give beams which come together at the third crystal. Interference between these beams produces a wave field modulated with the periodicity of the lattice planes. If the lattice of the third crystal is an exact continuation of the lattice of the first two the intensity maxima of the wave field will coincide with the gaps between the planes of atoms and there will be maximum transmission. A lateral shift of either the third crystal, or the modulations of the wave field, by half the interplanar spacing will give the condition for maximum absorption in the third crystal and hence minimum transmission. Hence the intensity transmitted by this device is extremely sensitive to any variation in orientation or lattice constant of the third crystal or any change of phase of either X-ray beam, resulting for example from the insertion of a thin object in the position indicated in Fig. 15.1(a).

The problem of maintaining three crystals parallel and in register with a lateral precision of a few Å and an angular precision of better than about 10^{-6} radians has been overcome by cutting the three crystals and a massive connecting base from one large near-perfect single crystal as suggested by Fig. 15.1(b). The lateral dimensions of the crystals and the width of the X-ray beams used may be several cm and mm, respectively.

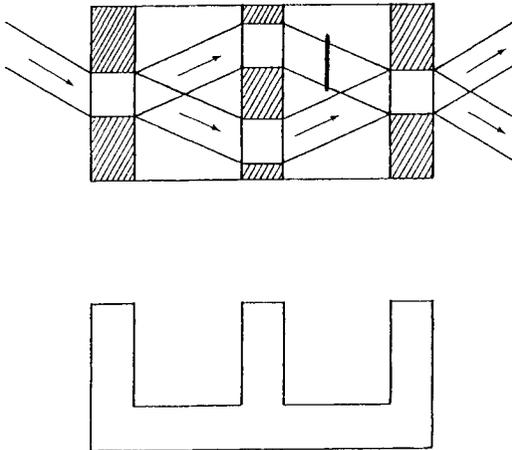


Fig. 15.1. The configuration of an X-ray interferometer formed by cutting three parallel connected slabs from a large near-perfect silicon crystal.

Instruments of this sort may be used to measure the thickness or refractive index for X-rays of any object inserted in one ray path of the interferometer. The sensitivity of the transmission to the alignment of the third crystal means that dilations of the lattice, $\Delta d/d$, as small as 10^{-8} and angular rotations as small as 10^{-8} radians can be detected. Intrinsic strains due to variations of impurity content of a crystal by 1 part in 10^7 or due to individual dislocations or other defects may be seen; or the device may be used to detect or measure strains due to external causes such as thermal gradients or elastic deformations.

Methods have been devised for combining an X-ray interferometer with an optical interferometer so that both radiations are used to measure the same displacement of a crystal. By this means the X-ray wavelengths and crystal lattice parameters may be referred to the fundamental standards of length, defined in terms of the wavelengths of particular optical spectral lines. This promises an improvement in the absolute accuracy of these quantities by several orders of magnitude.

15.3. *n*-beam and 2-beam dynamical diffraction

Before proceeding to discuss the methods used for deriving information from dynamical electron diffraction effects it may be as well to discuss in more detail the applicability of the two-beam approximation. For X-ray or neutron diffraction, as we have seen, this approximation is good for almost all cases and it is only necessary to ensure that the geometry of the experiment is not such that a third beam of appreciable amplitude is introduced. The assumption of 2-beam conditions was the only practical basis for the initial attempts to derive information from electron diffraction dynamical effects and persisted until it became abundantly clear that even in the most favorable of cases the neglect of *n*-beam interactions was leading to errors which were serious for the levels of accuracy being sought.

With the advent of detailed *n*-beam computations of intensities it was shown that the variation of intensities with crystal thickness and orientation could be as complicated as suggested, for example by Fig. 10.5 and 10.6. On the other hand it was confirmed that a relatively simple variation of intensity with thickness could be achieved in two types of special circumstance: one when a two-beam situation is carefully selected for a strong inner reflection; the other when the incident beam is very close to being parallel to a principal axis of the crystal.

Even for a two-beam orientation, with the Bragg condition for a strong inner reflection satisfied and large excitation errors for all other strong reflections, it is recognized that for a very thin crystal a multitude of

diffracted beams appear. Then the manner of the convergence to the two-beam situation with increasing thickness is of interest. Initial n -beam calculations assuming no absorption suggested that the relative intensities of all beams were established with the first "extinction distance" i.e. the principal periodicity shown by the incident beam. Beyond that, all beams maintained their relative intensities, averaged over a few extinction distances, for all thicknesses.

This conclusion must be modified slightly in the light of the results of calculations including absorption, such as those of Fisher [1969, 1972]. In order to avoid scaling difficulties and to make the result clearer, Fisher eliminated the effect of the overall decrease in intensity due to the average absorption effects and plotted the intensities of each beam as a percentage of the sum of the intensities in all beams for each thickness. Typical results are shown in Figs. 15.2 and 15.3. In the case of the 200 reflection from aluminum for 100 keV electrons with a tilt away from the $\langle 001 \rangle$ orientation which ensures that very little effect of non-systematic interactions is involved, the percentage of the energy contained in the two main beams is about 99 at a thickness of one extinction distance. With

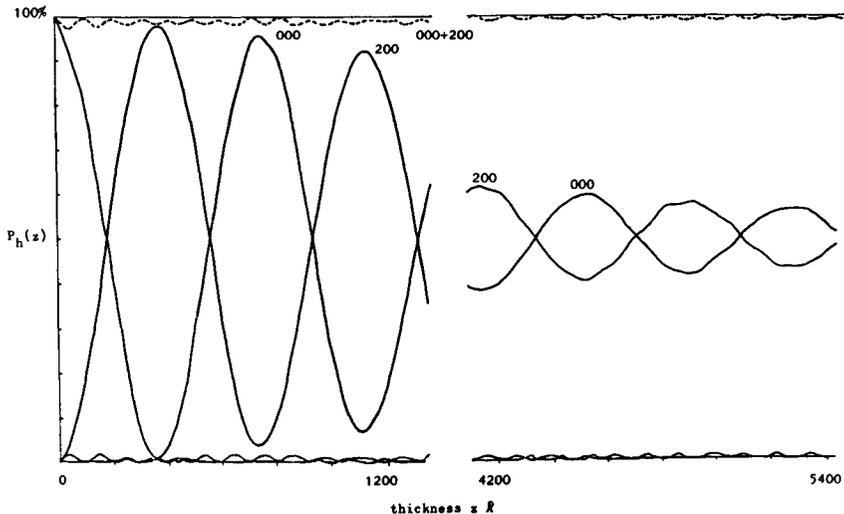


Fig. 15.2. Variation of the relative intensities of the diffracted beams with thickness, calculated for an aluminum crystal set for Bragg reflection of the (200) reflection with systematic interactions only, with 100 keV electrons. The intensities are given as percentages of the sum of all beam intensities for each thickness. (Fisher [1969])

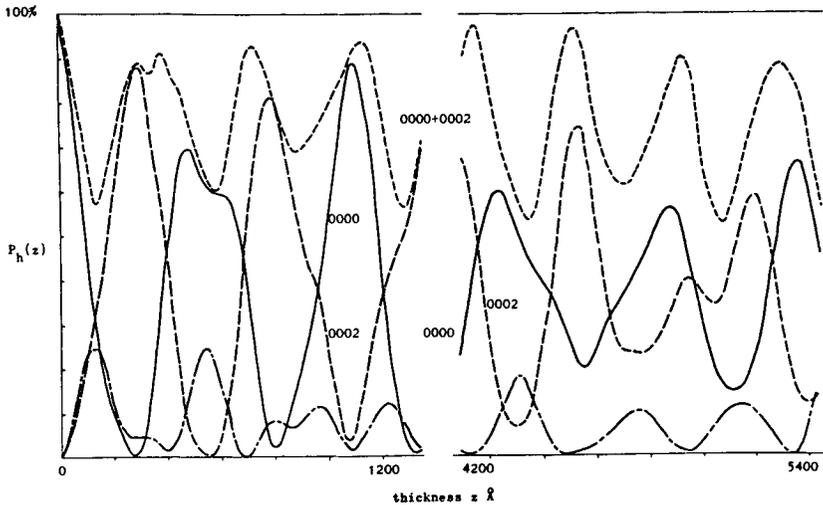


Fig. 15.3. Plot similar to 15.2 but for a crystal of CdS in $\langle 120 \rangle$ orientation with Bragg condition satisfied for (0002) reflection and systematic interactions only. (Fisher [1969])

increasing thickness there is a gradual increase in the percentage energy in the two main beams, to 99.67 percent for very large thicknesses. Thus the weak beams lose energy slowly. In the thickness range of a few thousand Å which represents the normal limit of observability, their contribution decreases by a factor of about one half.

With increasing thickness it is seen that the two strong beams share the energy almost equally but oscillate less and less. The absorption coefficient related to the intensity oscillations is greater than the overall absorption coefficient. This behavior is characteristic of the Borrmann effect, although for electrons this effect is by no means as striking as for X-rays.

For heavier atoms or larger unit cells or for orientations giving more non-systematic interactions the percentage of energy contained in the weak beams is greater. For example in the case of gold with the Bragg condition satisfied for the (200) reflection and systematic interactions only, the percentage energy in the weak beams decreases from about 13 percent at 200 Å to 4.2 percent for thicknesses greater than about 5000 Å. For the non-centrosymmetric CdS structure with the (002) reflection satisfied, the two-beam condition is again less well defined, with considerable deviations from a sinusoidal variation of intensities of the main beams and about 25 percent of the energy in weak beams for large thicknesses (Fig. 15.3).

We may conclude that within a few extinction distances something approaching a two beam situation can be set up. For favorable orientations of simple crystal structures the two main beams may have all but a few percent of the energy. The average combined strength of the weak beams then decreases slowly with thickness until, for a few thousand Å, which is beyond the usual range of observations for 100 keV electrons, it may be as small as one percent of the total energy. The weak beams continue to oscillate while the oscillations of the main beams become relatively less. This limiting situation may be upset by crystal imperfections. Calculations by Goodman [1968] show that a single fault may renew strong oscillations of the main beams and increase the weak beams considerably.

Apart from the decrease in energy of the two main beams, the presence of weak beams also has the effect of modifying the apparent extinction distance as measured by the periodicity of the oscillations. For the strong inner reflections from simple materials the effect is to decrease the extinction distance by from 5 to 10 percent (Hirsch et al. [1965]; Goodman and Lehmpfuhl [1967]).

It is now fully appreciated that the importance of n -beam interactions increases with increasing accelerating voltage. Observations on the periodicities of thickness fringes have been used to illustrate this dependence. For example, Mazel and Ayroles [1968] have shown theoretically and experimentally that the extinction distance for the 111 reflection of aluminum differs from the two-beam value by about 10 percent at 50 keV, 16 percent at 500 keV and 22 percent at 1200 keV.

For reasons which are not so well understood as in the two-beam case, a relatively simple variation of diffracted beam intensity with thickness may occur in the case of an incident beam very nearly parallel to a principal axis of the crystal. The calculations of Fisher [1969] for a beam parallel to the [001] axis of a copper-gold crystal, reproduced in Fig. 10.4, show that for the case of no absorption, all beams oscillate sinusoidally with the same periodicity. All diffracted beam intensities oscillate in phase with each other but out of phase with the incident beam.

In the presence of absorption and for other structures this relationship of the periodicities and phases is not so exact. This is shown for example by the calculations made by P.L. Fejes in relation to studies on the repetition with thickness of the lattice images obtained with the beam parallel to the b -axis of a crystal of TiNb_2O_6 (Fejes, Iijima and Cowley [1973]) (see Chapter 13).

15.4. Accurate determinations of structure amplitudes

The strong dependence of diffraction intensities on crystal

thickness and orientation under conditions of dynamical diffraction imply that if the crystal geometric parameters can be determined with sufficient precision, it may be possible to deduce the amplitudes and phases of the contributing reflections with high accuracy. This has been found to be possible for both X-rays and electrons using a variety of techniques. In particular, highly accurate values have been found for the inner reflections from a number of relatively simple materials with small unit cells. These reflections are those most strongly affected by the detailed distribution of electron density or potential around the atoms, so that the determination of their amplitudes has provided the most accurate information on ionization and bonding in crystals, as well as thermal motions, for comparison with theoretical models. For X-ray diffraction, the number of specimen materials that can be used is limited by the requirement for defect-free single crystals of mm size or greater. For high-voltage electrons, the possible effects of radiation damage and a vacuum environment can pose limitations but, because the perfect-crystal region required may be as small as a micrometer or less, a wide range of materials are available for measurement.

Valuable reviews of the electron diffraction techniques for accurate determination of structure amplitudes and of the relationship to determinations by X-ray diffraction have been given in the review articles of Gjønnes [1993] and Spence [1992] and the book of Spence and Zuo [1992]. Here we provide only a brief description of the methods to outline their basic ideas and limitations.

15.4.1. *Measurements of thickness fringes*

For strong reflections obtained under two-beam conditions from a near-perfect crystal which has the form of a wedge, the intensities of transmitted and diffracted beams show damped sinusoidal oscillations with thickness, giving thickness fringes parallel to the edge of the wedge, as described in Chapter 9. From the relationship, equation (9.1) and (9.2), the structure amplitude can be derived from the periodicity of the fringes if the wedge angle and the deviation from the Bragg angle are known accurately.

For X-rays the periodicities of the oscillations of intensity with thickness are of the order of hundreds of microns. Measurements may be made on small-angle wedges, cut from near-perfect crystals, by use of the X-ray topography technique illustrated in Fig. 9.8 (Kato and Lang [1959]). However for X-rays greater accuracy, reliability and convenience can be obtained by use of the section topographs (Chapter 9). A summary of experiments of this type is given by Kato [1969]. It is estimated that accuracies of determination of absolute values of structure amplitudes obtained in this way may be about 1 percent but relative values may be

obtained to about 0.1 percent. Still greater accuracy is claimed for measurements made by use of an X-ray interferometer. Hart and Milne [1970] estimate an accuracy of about 0.2 percent for their measurement of the structure factor for the (220) reflection of silicon made by observing the interference fringes arising when a small nondiffracting gap separates two thick perfect parallel regions of single crystal.

For electrons the observations of thickness fringes are made in bright and dark-field electron microscope images. In principle, measurements could be made on wedge-shaped crystals only a few thousand Å in size such as the small cubic crystals of MgO smoke. However there are practical difficulties of determining the orientation of such crystals with high accuracy and of maintaining the orientation constant. The most accurate measurements made so far have utilized wedge-shaped edges of large crystals which can be held rigidly. Then orientations can be determined accurately from Kikuchi line patterns obtained from a relatively large area of the crystal.

Many-beam dynamical effects are inevitably present and important for the electron diffraction case, although for relative simplicity of observation and computation the orientation is usually chosen to give systematic interactions only. Within this limitation the orientation may be selected to give maximum sensitivity for each parameter to be refined.

It may be noted that when strong n -beam dynamical interactions take place the thickness fringes no longer have the simple sinusoidal variation of the two-beam case. The combination of structure amplitudes and the excitation errors involved provide a unique form for the intensity variation with thickness which can be recognized independently of the scale of the image. Hence there is no need for an accurate determination of the crystal wedge angle or the magnification of the electron microscope.

This was found to be the case in the studies of the silicon (hhh) reflections by Pollard [1970, 1971] who used crystal wedges formed by fracture of large crystals. The wedges usually had good flat faces but the wedge angles varied greatly.

Typical results are shown in Fig. 15.4 in which part (a) is the measured intensity of elastically scattered electrons in the (222) reflection from a crystal set at the Bragg angle for (222). The measurement was made by electron-counting using the computerized EMMIE system (Holmes et al. [1970]) with energy filtering to remove inelastic scattering and an accelerating voltage of 65 keV. Part (b) is the theoretical intensity curve computed for refined values of the structure amplitudes. The calculations were made using 11 beams for the systematic reflections with careful investigation of the possible errors due to the effects of non-systematic interactions, inelastic scattering and finite beam convergence.

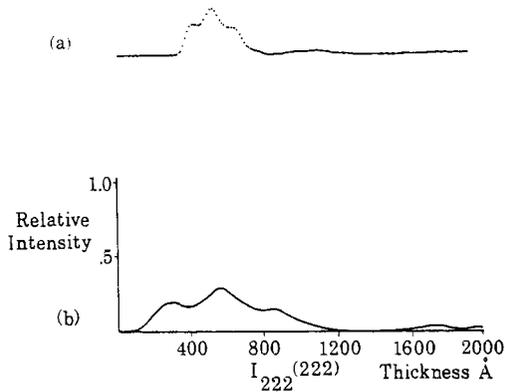


Fig. 15.4. Variation of intensity with thickness for the (222) reflection from a silicon crystal set at the (222) Bragg condition. (a) Measurement of intensity of a dark-field electron microscope image of a crystal wedge. (b) Calculated curve.

From these, and other curves obtained for other orientations, the structure amplitudes found were 5.02 volts and -0.105 volts for the (111) and (222) reflections respectively, with a conservatively estimated limit of error of about 5 percent. For the (222) reflection the structure amplitude would be zero for spherically symmetrical atoms. The value obtained is a measure of the covalent tetrahedral bonding present. It is in excellent agreement with the value of 0.107 volts deduced by Dawson [1967] from an analysis of X-ray diffraction results.

More recent applications of this method to the silicon reflections have been made by Ando et al. [1974]. See also Ichimiya and Uyeda [1977].

15.4.2 Structure amplitudes from rocking curves

A further set of experimental methods for deducing structure amplitudes depends on the measurement of diffraction intensities as a function of the angle of incidence of the incident beam. For X-ray diffraction the experiment giving the most direct and accurate data on structure amplitudes seems to be that of measuring the width of the Bragg-case reflection from the surface of a large perfect crystal under strict two-beam conditions (Kikuta et al. [1970], Kikuta [1971]). The theoretical curve (see Chapter 9) has a width proportional to $|F_h|$. This is broadened in practice by the angular width of the incident beam but if this incident beam is provided by a double-crystal spectrometer using asymmetric reflections from perfect crystals (Kohra and Kikuta [1968]) the angular width may be as small as 0.10° and accurate corrections for this can be made.

In this way the width of the 422 reflection from silicon was measured to be 2.95" from which the value of the atomic scattering factor was found with an estimated error of about 0.3 percent which could be improved upon by slight modifications of the equipment. Improvements of accuracy and versatility of the method have resulted from the use of a triple-crystal arrangement (Nakayama et al. [1971]).

In electron diffraction experiments the situation is complicated by the presence of n -beam dynamical diffraction effects. In addition, for the Bragg case of reflection from the surface of a large crystal, difficulties arise for 20 to 100 keV electrons because, with small Bragg angles and small angles of incidence (1 or 2 degrees with the surface), the intensities and widths of the reflections are extremely sensitive to deviations from planarity of the surface and to modifications of the crystal structure in the surface layers of atoms (Colella and Menadue [1972]). However for transmission through thin perfect crystals the correlation of experimental observations with theoretical calculations can be made with much greater confidence. Rocking curves for crystals of silicon 1200 and 2700 Å thick have been made by Kreutle and MeyerEhmsen [1969] using 71 kV electrons. By comparison with theoretical curves obtained by including the interactions of 6 or 18 systematic reflections, they deduced structure amplitude values for the (111) and (222) reflections (5.16 and 0.10 volts) in reasonable agreement with results from other methods.

15.4.3. Convergent beam electron diffraction methods

A number of accurate determinations of structure amplitudes have been made by use of the convergent beam diffraction method (Chapter 9) which is equivalent to a rocking curve method except that instead of measuring the diffracted intensities as a function of angle of rotation of a crystal for a parallel incident beam, one has a range of incident angles in the incident beam and the corresponding variation of diffracted beam intensity appears in the profile of the broadened diffraction spots. This method was developed and applied to the study of the structure amplitudes of MgO by Goodman and Lehmpfuhl [1967].

The extent of the agreement between theoretical and experimental intensities for the $h00$ systematics of MgO is suggested in Fig. 15.5. With this technique an area of perfect crystal of diameter as small as 200 Å has been used. It should be possible to find an area of perfect crystal of this size for almost any stable material. When a field-emission gun is used to produce a very small bright source of electrons, convergent beam patterns can be obtained from regions of diameter as small as 1 nm or less and such patterns can be recorded in a few milliseconds (Cowley [1992]).

Recently, Zuo et al. [1988] made measurements of the structure amplitudes for GaAs with high accuracy. After converting these values to the equivalent X-ray amplitudes, they were able to make a comparison of the electron density distribution with that calculated for neutral atoms and derive a charge-density difference map showing in detail the redistribution of electrons in the structure due to ionization and bonding. The charge in the interatomic bond was found to be 0.071 ± 0.045 electrons, in excellent agreement with the estimate of 0.08 electrons found from the pseudo-potential calculations of Cohen [1981].

The procedures for such measurements can be summarized as follows. The objective aperture size is adjusted to make the convergent beam disks as large as possible without overlapping. The crystal orientation is chosen by observing the K-line pattern to ensure that only systematic interactions, associated with a single line of reciprocal lattice points through the origin, are present and interference from non-systematic reflections is minimized. Only two reflections of a systematic set can show their Bragg maxima within the diffraction disks at any one time, so a set of crystal tilts is used to maximize the information on pairs of adjacent reflections.

The accelerating voltage and crystal orientation may be determined from the positions of HOLZ lines in the central disk (Fitzgerald and Johnson [1984]). Absorption parameters may be derived by matching the asymmetry of the diffraction lines within the central disk with many-beam calculations (Hirsch et al. [1977]). The crystal thickness may be derived from the near-kinematical parts of the rocking curves for large excitation error or weak reflections (eq. (9.5)).

The background in the diffraction pattern due to inelastic scattering must be subtracted. To remove the scattering due to plasmon or single-electron excitation, with energy losses greater than a few eV, an energy filter may be used. The thermal diffuse scattering may possibly be calculated theoretically or else determined from experimental measurements of the scattering outside the diffraction disks or measurements of the scattering around spots in focussed diffraction patterns from regions of thickness as close as possible to that of the region giving the CBED pattern.

15.4.4. The use of critical voltages

With increasing electron accelerating voltage the n -beam dynamical effects in electron diffraction become increasingly important as a result of the relativistic increase of mass. This is seen from the wave equation (8.1) where the potential $\phi(r)$ is multiplied by $2me/h^2$ to give $v(r)$ and the Fourier coefficients which enter into the dynamical equations. For a voltage E ,

$$v_h(E) = (1 + eEm_0c^2)v_h(0) \equiv Rv_h(0). \quad (1)$$

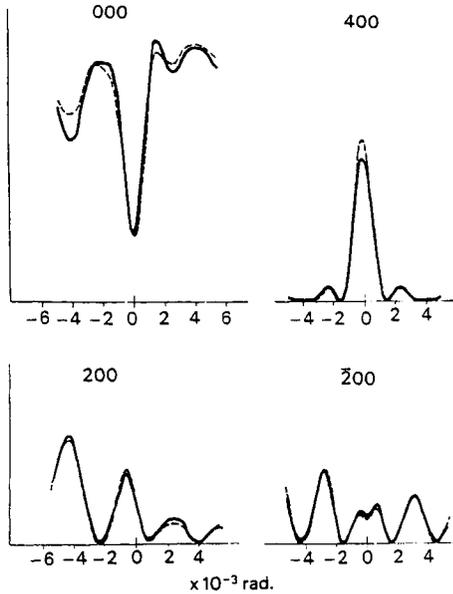


Fig. 15.5. Experimental (broken lines) and calculated (full lines) intensity distributions across spots in the convergent beam diffraction pattern from a thin MgO crystal, set at the Bragg angle for the 400 reflection with systematic interactions only (Goodman and Lehmpfuhl [1967]).

The relativistic factor R is approximately 1.2 for 100 keV and 2 for 500 keV since $m_0c^2 = 510.9$ keV.

From the Born series type of development, equations (1.16) or (1.17), it is seen that R enters as second, third and higher powers for the double, triple and multiple scattering terms which therefore become increasingly important with increasing voltage. If these terms are not all of the same sign there may be some voltage for which the intensity of a particular reflection goes to zero. This has been observed to happen in the relatively simple cases involving only the interactions of a systematic set of reflections for some uncomplicated crystal structures and has been used to obtain relationship between structure amplitudes and so accurate values for particular structure amplitudes.

Nagata and Fukuhara [1967] and more recently, Fisher et al. [1970] observed the vanishing of the extinction contour corresponding to a second order reflection at a certain critical voltage E_C . The corresponding disappearance of the second- or higher-order Kikuchi lines at the critical voltage in diffraction patterns from thick crystals was reported by Uyeda [1968] and by Watanabe et al. [1968] and similar effects in convergent beam diffraction patterns were observed by Bell [1971] and more recently by Imeson et al. [1979]. A review of the technique and results is given by Fox and Fisher [1988].

The original explanation of the effect by Uyeda [1968] and Watanabe et al. [1968] was made in terms of the Bethe potentials (see Chapter 8). The second Bethe approximation gives

$$v'_h(E) = R \left[v_h(0) - R \sum_{g \neq 0, h} \frac{v_g(0)v_{h-g}(0)}{\kappa^2 - k_g^2} \right] \quad (2)$$

where κ is the wave number for the average crystal potential and k_g is that for the g reflection. The denominator is thus proportional to the excitation error ζ_g . If the Bragg condition for the h reflection is satisfied, this becomes simply $g(h-g)/d_1^2$ where d_1 is the spacing for the first order reflection.

For a second order reflection $h = 2$ and the main contribution to the sum is for $g = 1$, $h - g = 1$. This is a relatively large term for simple structures so that for many metals and compounds having small unit cells $v'_h(E)$ becomes zero for a value of R corresponding to a voltage in the range 0 to 1000 keV.

Putting the bracket of (2) equal to zero for an experimentally determined value of R gives a relationship between structure amplitudes. The interest is usually in the determination of the structure amplitudes of the first order reflections since these are most affected by the ionization and bonding of atoms in crystals. It can be assumed that the structure amplitudes for outer reflections are given to a good approximation by the theoretical values for isolated atoms. The relationships derived from (2) have been used to find values for the first order structure amplitudes for a number of metals by Watanabe et al. [1969].

The use of the Bethe second approximation for the interpretation of the critical voltages in this way has been compared with the results of complete n -beam calculations (Watanabe et al. [1969]) and shown to give an accuracy of better than 0.5 percent in the critical voltage (and hence an even smaller error in v_1) for many substances.

The extinction phenomenon may also be described in terms of the Bloch-wave formulation by recognizing that at the disappearance voltage two branches of the dispersion surface touch, i.e., the eigenvalues of the matrix in equation (10.8) are degenerate (Lally et al. [1972]). Analogous degeneracies in the eigenvalues occur for axial orientations, giving critical phenomena such as reversals of symmetry. These effects may occur at lower voltages than for the systematic cases and so may be experimentally more accessible (Matsuhata and Steeds [1987]).

The critical voltage method is capable of giving the most accurate values of structure amplitudes attainable with electron diffraction techniques, with errors as small as 0.4%

(Hewat and Humphreys [1974]) which correspond to errors of about 0.1% in the corresponding X-ray diffraction values. The values of the structure amplitudes derived are dependent on the Debye-Waller factor, the long-range order parameter for solid solutions, the concentrations of impurity atoms and so on and applications of the critical voltage method have been made to determine all of these quantities. A discussion of the values of Debye-Waller factors to be applied in dynamical scattering of X-rays and electrons has been given by Spence [1992].

15.4.5. Intersecting K-lines

The regions of intersection of two K-lines (Kossel lines for X-rays or either Kikuchi lines or the corresponding lines in CBED patterns for electrons) are regions where three-beam dynamical diffraction effects become obvious. The perturbation of the two-beam intensities and positions of the lines by the introduction of a third beam is known as the Renninger effect in X-ray diffraction and it has been shown that the observed effects may form the basis for the derivation of the relative phases of the structure amplitudes (see Ewald and Heno [1968], Colella [1974], Post [1979]).

In Kikuchi patterns or within the disks or background of CBED patterns, the intersecting pairs of lines are seen to be replaced by two nonintersecting hyperbolas, as illustrated in Fig. 14.8. In rocking-curve experiments, Lehmpfuhl and Reissland [1968] were able to plot out the form of the intersecting regions and show that they represented sections of the dispersion surface. Gjønnes and Høier [1971] recognized that the intersecting K-lines represented sections of the dispersion surface in the same way. For the two-beam case it is known (Fig. 8.3) that the separation of the two hyperbolic branches of the dispersion surface is proportional to the structure amplitude of the diffracting beam. For the three-beam case, involving the reflections h and g as well as the zero beam, the separation of the two branches of the hyperbola depends on an effective potential involving the amplitudes and relative phases of the three beams h , g and $g-h$. The relative phases enter through a three-phase invariant, $\phi_h + \phi_{g-h} - \phi_g$ (see Spence [1992]). Thus values of the structure amplitudes, as well as information on their relative phases may be derived from measurements of the geometry at the intersection points.

The information on three-phase invariants, derived from three-beam diffraction situations, applies equally well for the centrosymmetric cases, for which all structure amplitudes are real, and for the non-centrosymmetric cases. On this basis Spence and Zuo [1992] have derived procedures for the accurate determination of the relative phases of structure amplitudes from observations of CBED patterns and applied these procedures to several non-centrosymmetric systems.

15.5. The determination of crystal symmetries

As mentioned in Chapters 5 and 6, diffraction experiments for which the kinematical approximation applies can give information regarding some of the symmetry elements of crystal structure but are severely limited in that they cannot reveal symmetry elements whose effect is to be found in the relative phases, rather than the amplitudes of reflections. This limitation is expressed by Friedel's law and the most obvious example is that the presence or absence of a center of symmetry cannot be determined.

As pointed out in the previous Section, the observation of three-beam situations, as in the intersections of K-lines with X-rays of electron diffraction may give information on three-phase invariants and hence on individual relative phases for reflections in favorable cases. The difficulty of obtaining suitable perfect crystal specimens has restricted the use of such methods for the solution of the phase problem in X-ray diffraction structure analysis. Electron diffraction information on relative phases, and so on the crystal symmetries that are not immediately apparent from kinematical diffraction data, may be obtained by direct inspection of the symmetries of convergent beam electron diffraction patterns and the systematic methods developed for the associated determination of space-group symmetries have been widely and usefully applied.

The initial observations of the failure of Friedel's law in reflection electron diffraction patterns from large crystals were reported and discussed by Miyake and Uyeda [1950] and treated theoretically using three-beam dynamical diffraction theory by Kohra [1954] and Miyake and Uyeda [1955].

In early transmission as well as reflection experiments, it appeared that the information on symmetry was less than for kinematical diffraction since the "forbidden" reflections associated with the presence of screw axes and glide planes often appeared strongly. Their presence was attributed to multiple scattering. It was not until more carefully controlled experiments with well-characterized single-crystal specimens became possible that the conditions under which symmetry-indicating observations could be made in a consistent and reliable manner. The current use of CBED patterns, obtained mostly in axial orientations, may be summarized briefly. More detailed accounts of the techniques, the equipment and the systematic methods for symmetry determination may be found in the review articles of Goodman [1993] and Eades [1992] and in the profusely illustrated books of Tanaka and Terauchi [1985], Tanaka, Terauchi and Kaneyama [1988] and Tanaka and Terauchi [1994]. An axial CBED pattern, such as illustrated in Fig. 15.6, with the incident beam axis parallel to, for example, the crystal *c*-axis, includes the zero-beam disk and the array of

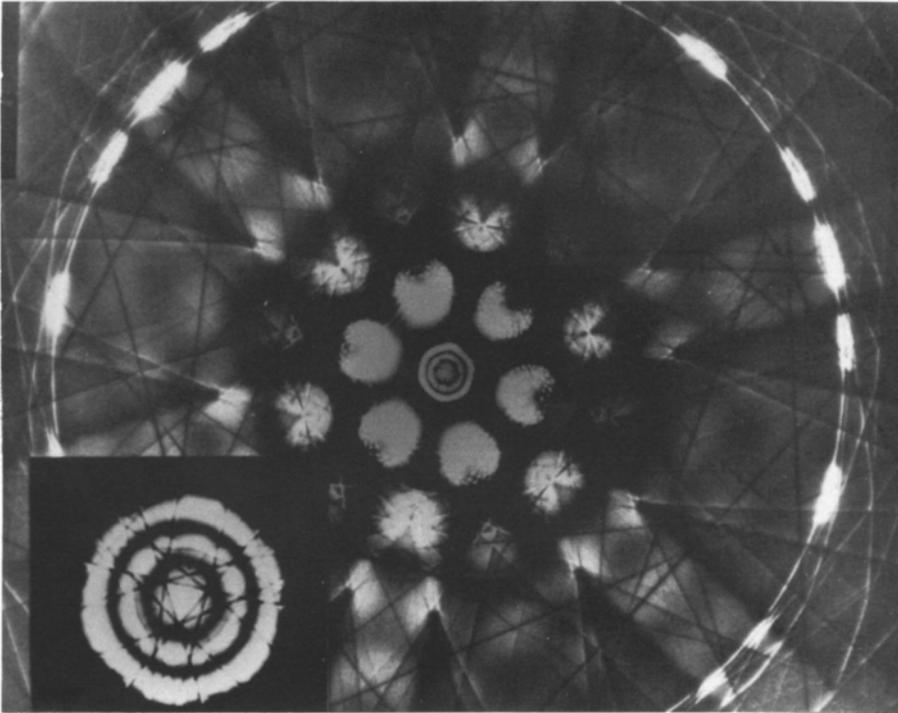


Fig. 15.6 Convergent beam electron diffraction patterns from a silicon crystal in [111] orientation showing the zero-layer spots, the K-line pattern and first layer spots. The insert is an enlargement of the 000 beam crossed by upper-layer lines. (Courtesy, Dr. R.W. Carpenter).

zero-layer disks, corresponding to $hk0$ reflections of the zero-order Laue zone (ZOLZ), from which the geometry and the symmetry of the projection of the structure in the c -axis direction may be deduced. The intensity variations due to ZOLZ reflections are characteristically rather broad and diffuse. In contrast, there are sharp and usually weak lines, crossing the zero disk and the $hk0$ disks, given by diffraction by upper-layer-line, i.e. hkl , reflections, also referred to as HOLZ or higher-order Laue zone reflections. The symmetry of the arrangement of these lines gives information on the three-dimensional symmetry of the crystal. The symmetry elements of the diffraction pattern can be classified as those within the zero disk due to the projection or to the HOLZ lines, and the those within the whole pattern due to the projection or to the HOLZ lines.

There are 31 possible symmetries of an axial CBED pattern, known as the "diffraction groups" (Buxton et al. [1976]). From observations of these symmetries for one or more axial orientations it is possible to deduce the point group of the crystal. Tabulations of the diffraction symmetries which facilitate the deduction of point groups are found in the general references given above.

The determination of space groups depends, as in the case of kinematical diffraction, on the determination of screw axes and glide planes through the observation of systematic absences. Under dynamical diffraction conditions, however, the extinction of kinematically forbidden reflections occurs only in the case of two-fold screw axes and glide planes. It was shown by Gjønnes and Moodie [1965] that these symmetry elements give rise to dark bands, the GM lines, running through the zero-layer disks of odd-order reflections, either on the line through the zero-beam direction or at right-angles to this, as seen in Fig. 15.7. The combination of the knowledge from these observations, in conjunction with the point-group determination, allows unique identification of 181 of the 230 space groups.

Other symmetry elements giving kinematical extinctions, such as the other screw axes, may often be inferred from near-extinctions in the corresponding CBED positions and confirmed, for example, by obtaining patterns from very thin crystal regions. This allows the identification of a further 27 space groups. Of the remaining space groups, 18 consist of 9 enantiomorphic pairs and for these the "handedness" may be inferred from calculations for particular off-axis orientations (see Goodman and Johnson [1977]). Thus the ambiguities arising in space-group determination from kinematical diffraction data are almost entirely avoided.

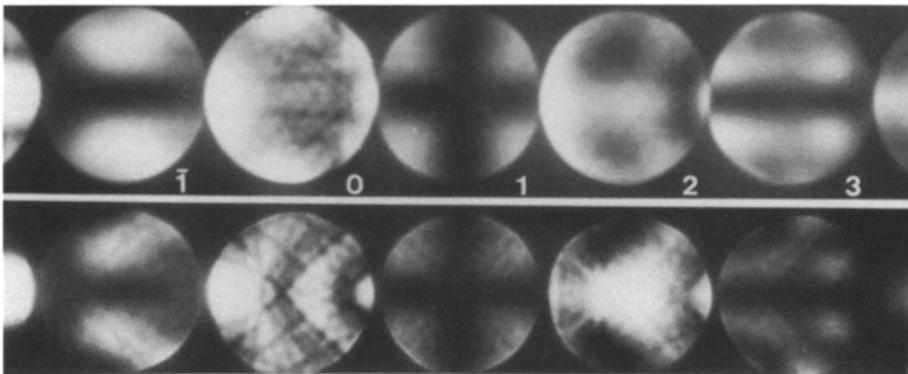


Fig. 15.7 CBED patterns showing $h00$ spots from (a) thin and (b) thick crystals of FeS_2 . The black G-M lines are visible in odd-order reflections. (After Tanaka and Terauchi [1985]).

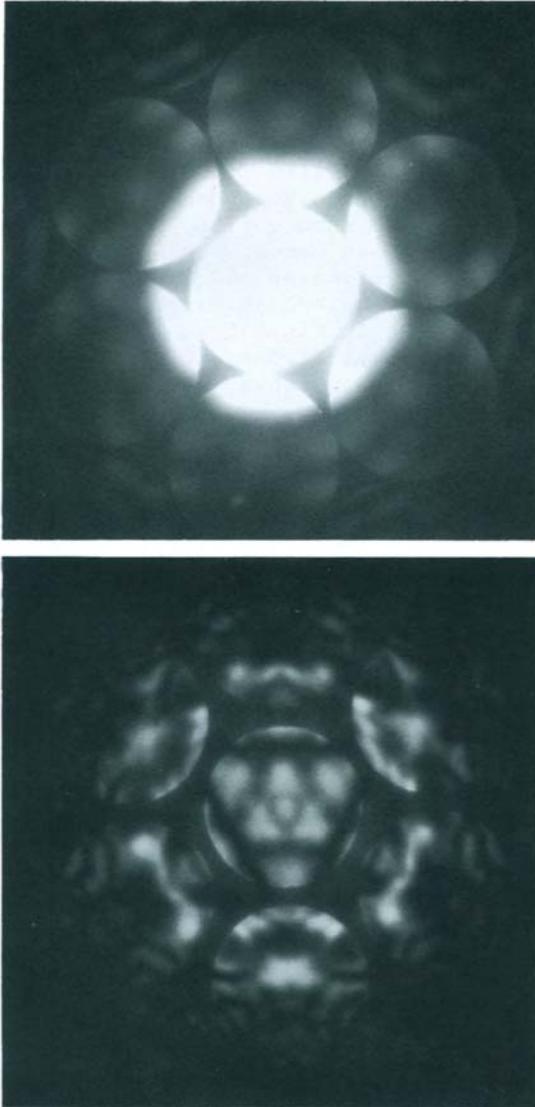


Fig. 15.8. Convergent beam diffraction patterns from graphite crystals with the center of the incident beam parallel to the c -axis, showing (a) the hexagonal symmetry for a perfect crystal and (b) the trigonal symmetry corresponding to a crystal having one or more stacking faults.

An important point to be made in this connection is that the symmetry and intensities of a transmission electron diffraction pattern reflect the symmetries of the crystal and not the symmetries of the unit cell. Thus for graphite with an incident beam in the *c*-axis direction the projection of the structure is hexagonal and most diffraction patterns (Fig. 15.8(a)) show a hexagonal symmetry. But if there is a stacking fault in the graphite crystal (but not in the middle of it) such that a rhombohedral stacking sequence occurs (...ABABCACAC...), the symmetry of the projection of the crystal is trigonal and a diffraction pattern having trigonal symmetry can result (Fig. 15.8(b)) (Johnson [1972]).

A further example is provided if the periodicity, or symmetry, of individual layers of atoms perpendicular to the beam is not that of the projection of the unit cell. Then, if the crystal contains a non-integral number of unit cells, the symmetry or periodicity of the projection of the crystal is that of the individual layers. For example, gold crystals, viewed in the [111] direction, show a strong periodicity of intensities of the fractional order reflections (such as $4/3$, $2/3$, $-2/3$), as individual layers of gold atoms are added beyond an integral number of unit cells (Lynch [1971]). Also intensities for MoO_3 crystals show a strongly varying contribution of the "forbidden" odd-order reflections with the additions of fractions of the unit cell contents on the surface of crystals (Goodman and Moodie [1974]). This suggests important possibilities for the study of surface steps and defects on crystals by the use of dark-field electron microscopy from these reflections as discussed in Chapter 19.

15.6. Coherent convergent-beam electron diffraction

In the above discussion of CBED the usual assumption has been made that the source of the convergent beam is effectively incoherent, i.e. that the intensities of the diffraction patterns for all angles of incidence are added and no interference effects are present due to the coherent interaction of electron beams coming from different directions. This assumption can be justified for the patterns produced in electron microscopes having thermal-emission electron sources but may fail for the case that the very small bright sources produced by field-emission guns are used, as in scanning transmission electron microscopes and some recent TEM instruments. With field-emission guns the effective source sizes may be as small as about 40 Å in diameter, the coherence width of the radiation at the level of the aperture which defines the convergence of the incident beam on the specimen may be greater than the aperture diameter, and it is a good approximation to assume that the

incident convergent beam is fully coherent. Then effects may be visible as a result of interference between electron waves scattered for different angles of incidence. Discussions of the appearance and applications of the resultant diffraction patterns have been given by Cowley [1992] and Spence and Zuo [1992].

For ideally perfect thin single crystals, the diffraction spots for each incident beam direction are delta-functions so that no overlap of spots for different direction of incidence occurs if there is no overlapping of CBED spot discs. When the spot discs do overlap, however, there can be interference between contributions to the overlap region from two different incident beam directions giving different diffraction spots. In the simplest case of the overlap of neighboring spots, interference fringes appear in the region of overlap, as shown very elegantly by Vine et al. [1992] and Tanaka et al. [1994]. The fringe positions depend on the magnitudes and phases of the structure amplitudes of the corresponding reflections and also on the position of the incident beam spot relative to the unit cell origin. If a detector is placed in the regions of overlap and the incident beam is scanned over the specimen a STEM image showing lattice-fringes is obtained, as described in Section 13.5.

For larger angles of convergence of a coherent incident beam, many diffraction spots may overlap and interfere. This corresponds to the situation that the beam focussed on the specimen gives an electron probe of diameter much smaller than the periodicities of the projection of the unit cell in the beam direction. Probe diameters of 2 Å or less are currently feasible. The intensity distribution of the diffraction pattern no longer reflects the unit cell periodicity, but changes as the probe is moved around within the unit cell (Cowley [1981a]). The symmetry of the pattern depends on the symmetry of the array of atoms illuminated by the beam. In principle, the symmetry elements of the unit cell could be investigated individually and structure analyses could be made of individual groups of atoms.

Using this approach, local defects in crystals may be investigated. In the study of planar defects in diamond crystals, Cowley et al. [1984] showed the change of symmetry and intensity distribution of coherent CBED patterns as an incident beam of diameter 3 Å was scanned across a defect. Beams of diameter about 10 Å have been shown to give characteristic streaking and distortions of the diffraction patterns when focussed on crystal edges (Cowley [1981b], Pan et al. [1989]) or internal discontinuities such as faults or out-of-phase domain boundaries (Zhu et al. [1985]). All such experiments may be regarded as special cases of in-line electron holography, as discussed in Section 13.6, and could be treated as such.

Mosaic crystals and polycrystalline materials

16.1. General

For the application of kinematical diffraction theory it is necessary to make the assumption that crystals are either very small or else are "ideally imperfect". On the other hand dynamical theory has been developed for ideally perfect crystals, with extensions to the treatment of crystal imperfections which become increasingly complicated and intractable as the departure from idealized situations increases. In between the limiting cases which may be approached by these relatively simple theories lies the vast majority of situations encountered in practice. The materials normally available for study by diffraction methods are often far from ideal with respect to either set of approximation. They may have a very complicated array of both extended and localized defects which are neither random nor isotropic in their distribution. The range of orientations of the crystal lattice may be very small or very large and the changes in orientation may take place discontinuously on well-defined planes (grain boundaries) or may be gradual (involving distortions of the lattice).

A general diffraction theory capable of taking all these factors into account is scarcely feasible and the detailed data on form and distribution of the crystal defects and discontinuities, which would be needed for such a theory, is rarely available.

To deal with the variety of experimentally important situations a number of convenient idealized models have been introduced and useful theoretical approximations have been made. The acceptance of these models and approximations have depended more on considerations of practical utility than on their accuracy in representing the physical situation or on theoretical rigor.

In many cases the possibility of producing a relatively simple model depends on the averaging of intensities which is appropriate when the coherence of the incident radiation and the resolution of the diffraction pattern are limited. This is the basis, for example, of the assumption that the intensities given by crystallites having different orientations may be added incoherently.

16.2. Mosaic crystals

16.2.1. The mosaic crystal model

The earliest model introduced to explain the kinematical nature of the diffraction from macroscopic crystals was the "mosaic crystal" model of Darwin [1914]. It was assumed that a crystal is made up of small perfect-crystal blocks, small enough so that the diffraction from each individually is given by a kinematical approximation and having a random distribution in orientation over a range of angles which is much greater than the angular width of reflection from an individual crystal.

This model has been retained as a useful one even after detailed studies of the structure of crystals by electron microscopy and other methods have shown that its relationship to the actual state of affairs is often quite remote. In some materials it is seen that most dislocations have aggregated to form dislocation networks, forming small-angle grain boundaries separating regions of crystal which are only slightly distorted. For many metals there are tangles of dislocations and other defects giving statistical accumulations of distortions and a variation of orientation of the lattice which is more or less random but continuous. In other materials stacking disorders, micro-twins, impurities and other defects provide much the same effect. Frequently there is a strong anisotropy since the defects occur preferentially on particular crystallographic planes or in particular directions.

The diffuse scattering due to the defects and the modification of the sharp Bragg reflection intensities by pseudo-temperature effects have been treated in other chapters. Here we consider only the Bragg intensities in relationship to the kinematical and dynamical approximations.

16.2.2. Kinematical integrated intensities

For a mosaic crystal we can consider the distribution of scattering power around all reciprocal lattice points to be broadened equally by the average of the shape transforms for the individual crystallites. In addition the distribution of orientations of the crystallites spreads the scattering power maximum in two dimensions over the surface of a sphere centered at the reciprocal space origin.

For most X-ray diffraction work the measurement is made of the scattering power integrated in three dimensions over the whole of the scattering power maximum, giving the sum of the intensities from all individual crystallites. For a crystal rotated with angular velocity ω through the Bragg position, this integrated intensity is then (Warren [1969]):

$$E = \frac{I_0}{\omega} \left(\frac{e^4}{m^2 c^4} \right) \frac{\lambda^3 V |F_T|^2}{\Omega^2} \left(\frac{1 + \cos^2 2\theta}{2 \sin 2\theta} \right), \quad (1)$$

where I_0 is the incident intensity, V is the volume of the sample, F_T is the structure amplitude modified by the temperature factor, $(1 + \cos^2 2\theta)/2$ is the polarization factor and $\sin^{-1} 2\theta$ is the Lorentz factor.

In the case of electron diffraction, intensities approaching those for kinematical scattering from single crystals are given only by very thin crystal sheets which are normally very much greater in lateral dimension than in thickness. The most common origin of misorientation between different parts of the crystal is bending of the crystal by rotation about axes lying approximately parallel to the sheet. The scattering power around reciprocal lattice points is greatly elongated in the direction nearly parallel to the incident beam by the small crystal dimension. This extension may be represented by convolution of the scattering power distribution by $\pi^{1/2} C \exp\{-\pi^2 C^2 w^2\}$, where C is the average crystal thickness and w is the appropriate reciprocal space coordinate.

If the mean angular spread of orientations is α , there is an additional extension of the scattering power spikes represented approximately by convolution with

$$(\pi^{1/2} d_h / \alpha) \exp\{-\pi^2 d_h^2 w^2 / \alpha^2\},$$

where d_h is the lattice spacing for the h reflection, giving the distribution suggested in Fig. 16.1. The fact that the misorientation gives the spikes a finite curvature and width may be ignored as a first approximation.

The intensity of the diffraction spot integrated over the photographic plate is then given by the intersection of the Ewald sphere with the corresponding spike as

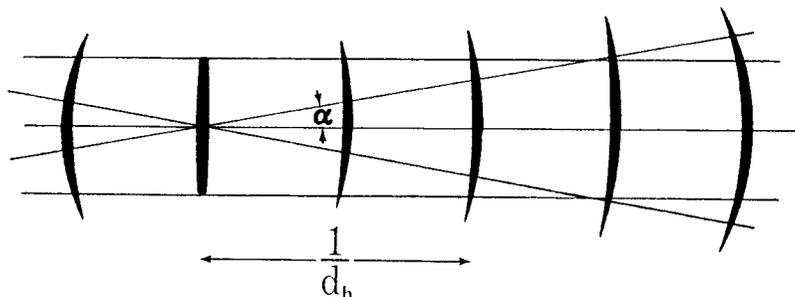


Fig. 16.1. The elongation and bending of the "shape-transform" spikes around reciprocal lattice points for a thin crystal, bent to give an angular spread of orientations, α .

$$I_h = I_0 \frac{\sigma^2 |\Phi_h|^2}{\Omega^2} \frac{\pi^{1/2} C d_h V}{(C^2 \alpha^2 + d_h^2)^{1/2}} \exp \left\{ - \frac{\pi^2 C^2 d_h^2 w^2}{C^2 \alpha^2 + d_h^2} \right\} \quad (2)$$

with w equal to the excitation error ζ_h for the h reflection. Here Φ_h is the Fourier coefficient of the potential distribution. For high voltage electrons and a large angular range of orientations it may be assumed as a first approximation that the exponent is zero and the expression becomes

$$I_h = I_0 \frac{\sigma^2 |\Phi_h|^2}{\Omega^2} \frac{\pi^{1/2} d_h V}{\alpha}. \quad (3)$$

If the exponent in (2) is not zero, the effect will be similar to that of an additional temperature factor on the intensities.

16.2.3. Extinction effects

It was recognized by Darwin [1914] that the approximation of kinematical scattering from an ideal mosaic crystal could give appreciable errors for a variety of reasons. He distinguished two types of "extinction" effect which could act to reduce the intensity below the kinematical value.

Primary extinction is said to be present if individual perfect crystal blocks are too large to allow the assumption of kinematical scattering. Under the two-beam diffraction assumption this approximation fails for appreciable values of $F_h H$ for X-rays or $\sigma \Phi_h H$ for electrons where H is a measure of the crystal dimensions. Thus for a given crystal size one would expect the stronger reflections to be affected most and the weaker reflections to be most nearly kinematical. Under the n -beam diffraction conditions of electron diffraction, of course, this does not apply.

Even in the absence of primary extinction, there may be an effect known as secondary extinction associated with the repeated diffraction of an incident beam by several separate mosaic blocks. The intensity incident upon one mosaic block may be attenuated by diffraction in several previous blocks having almost the same orientation. This effect will be most pronounced for inner reflections, for which the angular range of reflection from a small volume is greater, for reflections having large structure amplitudes and for samples having small mosaic spread or large dimensions.

Secondary extinction has been found particularly important for neutron diffraction for which the mosaic blocks are usually much too small to give primary extinction effects, but relatively large samples are normally used in order to get useful diffraction intensities from relatively weak sources.

Since in the crystals used for X-ray structure analysis, one form or other of extinction is nearly always present, a practical method for making corrections for extinction is of great value. Initial treatments of the scattering problem, designed to give useful results by making appropriate approximations were given by Zachariasen [1968, 1969], Cooper and Rouse [1970] and Becker and Coppens [1974] and a more elaborate treatment was made by Kato [1976, 1979]. The difficulty of treating the coherent interactions of diffracted waves in an imperfect and inhomogeneous crystal is avoided by the assumption that the averaging out of the phase-sensitive terms can be done before, rather than after, the consideration of the interactions of the beams. Then a set of differential equations analogous to (10.32) can be used for the intensities, rather than the amplitudes of the beams.

Following Sabine [1992], the flow of radiation in the crystal may be described by the equations

$$\begin{aligned}\frac{\partial P_i}{\partial t_i} &= \tau P_i + \sigma P_f, \\ \frac{\partial P_f}{\partial t_f} &= \tau P_f + \sigma P_i,\end{aligned}\tag{4}$$

where P_i and P_f are the radiation current densities in the initial and final beams and t_i and t_f are the distances measured along these beam directions. The coupling constant σ is the cross section per unit volume for scattering by the Bragg reflection and τ , which is always negative, is the cross section per unit volume for removal of radiation from the beams. If a linear absorption coefficient μ describes the only significant removal of energy, $\tau = -(\mu + \sigma)$.

The equations (4) have an analytical solution for simple crystal geometries. For example, for the Laue case of transmission through a plane crystal slab, for which the path length of the diffracted beam is D ,

$$P_f / P_i^0 = \frac{1}{2} \exp(-\mu D) [1 - \exp(-2\sigma D)]$$

The quantity σ is determined from the assumptions regarding the crystal imperfections, expressed in terms of the intensity and angular distribution of the local scattering. For X-ray or neutron scattering (but not for electron scattering), it is a reasonable approximation to express σ in terms of the kinematical approximation for the integrated intensity and angular distribution (preferably in Lorentzian form following Sabine [1988]) for the coherently-scattering crystal region. Expressions can then be obtained in terms of these variables which give either primary or secondary extinction as limiting cases.

16.2.4. Dynamical electron diffraction intensities

A treatment of the dynamical scattering of electrons by imperfect crystals on the same basis as this "extinction" treatment for X-ray or neutron diffraction is scarcely feasible because strong dynamical effects take place within crystal regions much smaller than the usual "mosaic block" size and because the number of mosaic blocks along the path of an electron beam through a "single crystal" sample is rarely greater than one or two.

For the most part the spot patterns given by imperfect, bent or mosaic crystals must be considered as the sum of dynamical diffraction patterns from individual near-perfect crystals. In some cases where non-parallel crystals overlap, there will be the effects of double diffraction as each diffracted beam of the first crystal acts as a primary beam for the second crystal, giving diffraction spots whose positions are defined by the sums of the operative diffraction vectors for the individual crystals. When there is no simple relationship between the orientations of the two crystals the directions of the doubly diffracted beams prevent any coherent interactions with the singly diffracted beams. This is the case termed "secondary elastic scattering" (Cowley, Rees and Spink [1951]) to distinguish it from coherent multiple (dynamical) scattering.

In general electron diffraction spot patterns from large areas of thin crystal films may be considered as representing the averaging of dynamical diffraction intensities over a range of thickness and orientation.

Attempts which have been made to find analytical expressions for the electron diffraction intensities from a crystal averaged over orientation of thickness have not produced any very useful result (Kogiso and Kainuma [1968], Cowley [1969]). The averaging of intensities, computed using n -beam diffraction, over thickness is readily performed but averaging over angle of incidence is laborious. The indications are that when the intensities in a diffraction pattern are averaged in either or both of these ways, they may show a pattern of spots which is distinctive and characteristic of the crystal structure but not readily related to the kinematical diffraction pattern. In averaging over thickness for a perfect crystal in a principal orientation, the characteristic absences corresponding to some symmetry properties of the crystal are maintained (Chapter 14) but this is not necessarily the case for imperfect crystals or when there is an averaging over orientation.

16.3. Polycrystalline material

16.3.1. Idealized models

The ideal polycrystalline material contains a very large number of small, independently-diffracting crystallites

having complete randomness in orientation about at least one axis. It may consist of fine powder, either loose or compacted, or a continuous solid having closely-spaced large-angle grain boundaries, as in a fine-grained piece of metal.

If the orientations of the crystallites are completely random, with no preferred direction for any crystal axis, the diffracted beams form continuous cones of radiation with the incident beam as axis and a half-angle equal to $2\theta_h$ where θ_h is the Bragg angle for the h reflection. The intersection of these cones of radiation with cylindrical films in the case of X-ray diffraction, or with flat plates in the case of electron diffraction, give the well-known powder patterns of continuous lines or rings.

A comprehensive survey of the techniques and applications of powder diffraction techniques and applications for X-ray and neutron diffraction is given in the collection of papers edited by Bish and Post [1989]. These papers cover the use of synchrotron X-radiation and the time-of flight methods for thermal neutrons from spallation sources. The use of electron powder diffraction patterns has been surveyed by Cowley [1992] and by Vainstein et al. [1992].

Particularly in the case of electron diffraction from thin layers of polycrystalline materials, preferred orientations of the crystals may result from the particular conditions of specimen preparation. If the crystallites have a thin, plate-like habit, they may tend to lie flat on a supporting film when the specimen is formed by settling from suspension, by growth on a flat substrate or by a mechanical spreading action. Then, ideally, the one crystallographic axis perpendicular to the plane of the plates has a strongly preferred orientation, but the orientations about this axis are completely random. This occurs frequently for electron diffraction specimens consisting of very small thin crystals supported on a thin supporting film of carbon or other light-element material.

To illustrate the form of the diffraction patterns produced by such a sample we may consider the case of crystals for which the a and b axis lie in a plane which is preferentially oriented parallel to a supporting surface, so that there is preferred orientation of the c^* axis perpendicular to the support. In reciprocal space, the randomness of orientation spreads the scattering power maximum around each reciprocal lattice point into a continuous circle about the c^* axis, as suggested in Fig. 16.2. This circle is a sharp line if the preferred orientation of the c^* axis is exact but is spread into a ring of finite width by any spread of orientations.

For convenience we consider an orthorhombic cell for which c^* and c are parallel. Then if an incident beam of fast electrons is parallel to the c axis, the Ewald sphere is tangential and close to the $a^* - b^*$ plane and parallel to the

rings of scattering power for the $hk0$ reflections. The diffraction pattern therefore consists of continuous circles but includes those for $hk0$ reflections only.

If the incident beam is tilted at an angle ϕ to the c axis, as suggested in Fig. 16.2(a), the near-planar Ewald sphere intersects the rings of scattering power in a set of short arcs. The $hk0$ arcs lie on a line through the origin. Other parallel lines of arcs are given for the hkl arcs, with all reflections having the same l value giving arcs lying on the same line as suggest in Fig. 16.2(b). These patterns, which resemble the rotating-crystal patterns familiar in X-ray diffraction, have been used extensively for purposes of structure analysis by the Soviet school (Vainshtein [1964]) who have named them "oblique texture patterns".

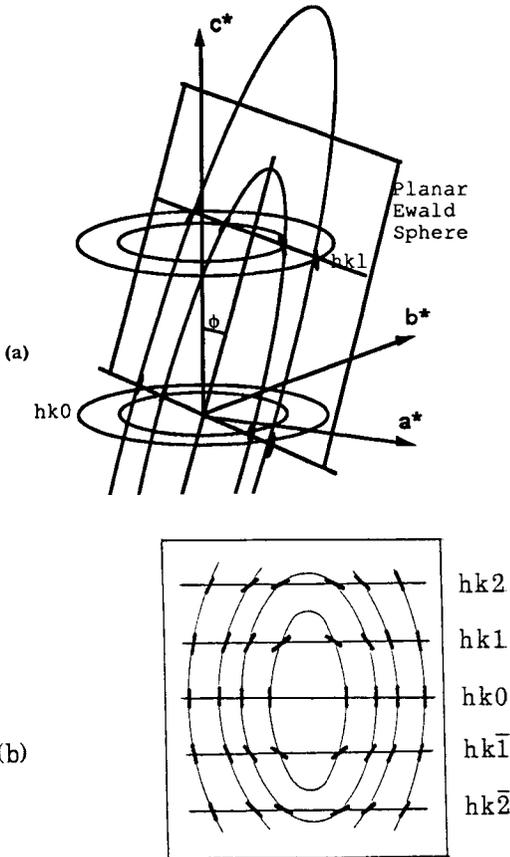


Fig. 16.2. The production of an "oblique texture" arc pattern when the Ewald sphere intersects the set of rings generated in reciprocal space by the reciprocal lattice points of crystals randomly oriented except that their c -axes are all parallel.

A similar type of pattern may be given by fine needle-shaped crystals for which the needle axis is given a preferred orientation, for example by drawing or extruding of a sample into a thin rod, or also by some special growth process. In this case it is normally a real-space axis (the c axis rather than c^*) which has the preferred orientation. The incident beam is then usually almost perpendicular to the fiber axis and gives a diffraction pattern which is the same as a conventional rotating-crystal pattern.

16.3.2. Kinematical diffraction intensities

From the expression (1) for the integrated intensity from a single small crystal we may derive the expression for the total intensity in a diffraction cone given by a random powder as (Warren [1964])

$$P = I_0 \left(\frac{e^4}{m^2 c^4} \right) \frac{V \lambda^3 m F_T^2}{4 \Omega^2} \left(\frac{1 + \cos^2 2\theta}{2 \sin \theta} \right). \quad (5)$$

Hence m is the multiplicity factor, giving the number of equivalent reflections, having the same Bragg angle, which contribute to the ring intensity.

The measurement usually made is that of the intensity per unit length of a diffraction ring, given by dividing P of (5) by the circumference of the ring,

$$P' = P / 2\pi R \sin 2\theta, \quad (6)$$

when R is the distance from specimen to film.

For oblique texture patterns we may similarly distinguish between the total intensity of an arc and the intensity per unit length at the center of the arc, which is often easier to measure. For the total intensity within the arc we may write (Vainshtein [1964])

$$I_h = I_0 \frac{\sigma^2 |\Phi_h|^2}{\Omega^2} \frac{VR\lambda m}{2\pi r' \sin \phi} \quad (7)$$

where r' is the radius of the ring of scattering power for the h reflection.

If the angular spread of the arc is α and r is the radial distance of the arc from the center of the pattern, the local intensity per unit length is,

$$I'_h = I_h / r\alpha. \quad (8)$$

16.3.3. Line profile analysis

For a small near-perfect single crystal the distribution

of scattering power in reciprocal space around each reciprocal lattice point is given by the square of the Fourier transform of the shape function of the crystal. If the crystal is bent or distorted or if there are many such crystals which are closely parallel with some distribution of orientation or of lattice constants, the reciprocal space distribution is modified in a characteristic way, as illustrated for a particular case by Fig. 16.1. A detailed investigation of the distribution of scattering power in reciprocal space should therefore allow a great deal of information to be derived about the crystal sizes, the spread of orientations and also the spread of unit cell dimensions.

For many materials of industrial or scientific significance it is not possible to work with single crystals. This is the case particularly for X-ray diffraction studies of microcrystalline materials, such as metals which have been subjected to various degrees of cold working. Then only powder patterns can be obtained and the only information on the form of the reciprocal space distributions around the reciprocal lattice points of the small crystallites is the statistically averaged data contained in the intensity profiles of the diffraction rings.

The use of powder patterns suffers, on the one hand, because on averaging over all orientations of the Ewald sphere, the scattering function in three dimensions is reduced to a one-dimensional function. On the other hand, the intensities and dimensions of the diffraction features can be measured with considerable accuracy in powder patterns and a detailed analysis is possible in terms of a limited number of well-chosen parameters.

An elegant method introduced by Warren and Averbach [1950, 1952] allows the contributions of crystal size and strain to be distinguished through analysis of the Fourier transforms of sets of powder pattern lines. Since that time, the development of the line-profile analysis techniques has been extensive and a considerable literature on the subject has accumulated (see Howard and Preston [1989]).

In order to derive information concerning the specimen from the details of the line intensity profiles, it is necessary to take into account the spread of the powder lines due to instrumental factors such as the spread of wavelengths, finite source size and slit widths and other geometric factors. Initially a simple Gaussian line shape was assumed but, although reasonable for neutron diffraction, this proved inadequate for the better resolution attainable with X-ray diffraction. Similarly a simple Lorentzian shape gave a poor fit and more complicated functions such the convolution of a Gaussian and a Lorentzian, the Voigt function, have proved more realistic (see Howard and Preston [1989]). Given an accurate representation of the line intensity profile for the ideal case it is possible to assume that the observed line intensity profile is given by convolution of this ideal intensity profile with spread functions due to the

particular specimen characteristics under investigation. The line profile is measured and digitised for small increments of 2θ and the spread function due to the specimen is obtained by deconvolution. In this way it is possible to study the average size and size distribution of the crystallites in the sample, the effects of strain and of faulting or other forms of crystal defects and the effects of incomplete long-range atomic ordering for those cases for which single-crystal studies are not possible or else not relevant for the samples being considered.

16.3.4. Rietveld refinements

Major developments in the scope and power of powder diffraction techniques followed from the demonstration by Rietveld [1967, 1969] that the intensities of a powder pattern, measured by a step-scan with small increments of 2θ through the whole range of scattering angles, could be utilized for the refinement of structural information. A major defect of the powder methods for structural studies had been that, except for relatively simple structures with small unit cells, the number of reflections contributing to the one-dimensional intensity distribution tended to be so large that there was a great deal of superposition of non-equivalent reflections and overlapping of diffraction peaks. The Rietveld concept was that each intensity measurement of the step scan could be considered as a separate experimental observation and the parameters defining the specimen structure could be refined by comparing experimental and calculated values, using a least-squares refinement procedure. Discussion of the recent developments and typical results of the Rietveld method are given by Post and Bish [1989] and Willis and Albinati [1992].

The quantity that is minimized in the least-squares refinement procedure is

$$R = \sum_i w_i (I_{iO} - I_{iC})^2, \quad (9)$$

where I_{iO} is the observed intensity and I_{iC} is the calculated intensity for step i of the scan and w_i is a weighting factor determined by the counting statistics or other consideration.

The refinement may be done taking into account only those measurements for the regions where the intensity is appreciably different from the background, or else a "whole-pattern" refinement may be made including all background measurements. Measurements within an intensity peak may be considered as having contributions from a number of Bragg reflections. Thus the intensity at the point i may be written as

$$I_{iC} = I_{ib} + \sum_k G_{ik} I_k, \quad (10)$$

where I_{ib} is background intensity, G_{ik} is a normalized peak profile function and I_k is intensity for a Bragg reflection. For neutron diffraction, which was the original application of the technique by Rietveld, the peak profile function could be approximated by a Gaussian. For the later-developing applications to X-ray diffraction powder patterns, the Voigt function mentioned in the previous Section, or more complicated functions, have been used. The background intensity can usually be fitted with a suitable polynomial function.

The refinement by minimization of the quantity (9) can be made only if a reasonably good model for the structure is available but such a model can usually be obtained from preliminary analysis of the pattern. One area of application, for example, is for the refinement of the determination of unit cell dimensions when the usual methods, which rely on the resolution of individual lines, can not be applied. Refinements to give five-figure accuracy are possible.

Valuable determinations of crystal structures have been made using the Rietveld procedures for many cases where the material can be made, or found in nature, only in the form of fine powders so that single-crystal techniques are not available. Initial models for the structure may frequently be derived from structures of analogous compounds or from high resolution electron microscopy. The number of parameters which may be varied to fit the experimental data may be of the order of 50 and these may be positional parameters, thermal vibration parameters or parameters defining partial or fractional occupations of sites. Mixtures of phases may be analysed and in some cases the presence of small amounts of unexpected impurities may be revealed. Difficulties that may complicate the process included the presence of preferred orientations of some crystallites and differences in crystallite dimensions for some phases.

16.3.5. Dynamical diffraction intensities

For X-ray and neutron diffraction, extinction effects are present in somewhat modified form for powder samples. Primary extinction becomes important if the crystals are too large in comparison with the structure amplitudes. Secondary extinction effects may occur if the sample size is too large although, for random orientations of the crystals, the incident beam is weakened by all possible diffraction processes at once. The effect is the same for all reflections and resembles a uniform absorption effect.

For electron diffraction it is necessary to average the dynamic diffraction intensities over angle of incidence. This was done initially for the pure two-beam case without absorption by Blackman [1939].

If we start from the expression for the intensity of the diffracted beam I_h in the Laue case, equation (8.29), we can replace integration over angle by integration over the

variable w and use the relationship

$$\int_{-\infty}^{\infty} \frac{\sin^2[A(1+w^2)^{1/2}]}{1+w^2} dw = \pi A \int_0^A J_0(2x) dx. \quad (11)$$

In this case

$$A \equiv A_h = v_h \lambda H / 4\pi = \sigma H \Phi_h. \quad (12)$$

Hence we deduce that the ratio of observed dynamical to kinematical intensities should be

$$I_{\text{dyn}}/I_{\text{kin}} = A_h^{-1} \int_0^{A_h} J_0(2x) dx. \quad (13)$$

Then for small value of A_h , the integral is proportional to A_h and the kinematical result is the limiting case. For large values of A_h , the integral tends to its limiting value of 0.5 and since I_{kin} is proportional to $|A_h|^2$, we see that the dynamical intensity is proportional to $|A_h|$ and the ratio (11) tends to zero through slight oscillations.

The variation of the ratio (11) with A_h is the "Blackman curve", plotted in Fig. 16.3 together with experimental

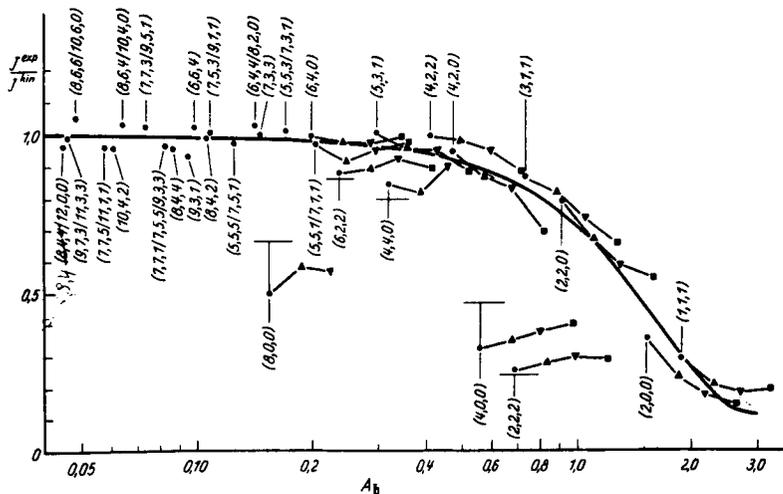


Fig. 16.3. The Blackman curve (Blackman [1939]) for the ratio of dynamical to kinematical intensities for a ring pattern as a function of $A = \sigma H \Phi_h$. The experimental points are those obtained by Horstmann and Meyer [1965], from measurements on ring patterns from aluminum films at various voltages. The short horizontal lines indicate values calculated using the Bethe potentials, equation (12). (After Horstmann and Meyer [1965].)

measurements of the relative intensities measured from ring patterns given by thin aluminum films (Horstmann and Meyer [1962, 1965]). The experimental measurements were made for a number of different reflections (different Φ_h), for various values of the average thickness, H , and for various electron wavelengths (and hence σ values). It is seen that in general the agreement with the Blackman curve is quite good. Comparable results have been obtained for ring patterns of other metals (Wedel [1963]; Kuwabara [1961, 1962]) and various inorganic salts (Kuwabara [1967]).

In each case a few of the reflections showed large deviations from the Blackman curve, as in the case of the (400) and (222) reflections of Fig. 16.3. These are the higher orders of strong inner reflections and for simple structures they show uniformly low intensity values.

For the most part it appears that the strong n -beam dynamical interactions which occur for many orientations of small crystals are averaged out and have little effect on the powder pattern intensities but, as might be anticipated, the "systematic" interactions remain since for the higher and lower orders of a given reflection the excitation error is the same for all orientations when the given reflection is excited.

Some account of this effect is given by use of the Bethe potentials (8.34). As Bethe [1928] suggested, the effect of weak beams may be included into a two-beam theory by replacing the potential coefficients by, for example,

$$U_{0,h} = v_h - \sum_g \frac{v_g v_{h-g}}{\kappa^2 - k_g^2}, \quad (14)$$

and it is these modified potential coefficients which enter into the expressions for the diffracted intensities.

For a reflection $2h$, which is the second order of a strong inner reflection h , the most important contribution to the second term of (12) is for $g = h$ so that

$$U_{0,2h} \approx v_{2h} - \{v_h^2 / (\kappa^2 - k_h^2)\}. \quad (15)$$

Since the denominator is positive, the effective scattering factor is reduced.

Applying these corrections gives reasonable agreement with the experimentally observed intensities for the case of Fig. 16.3. However, as mentioned in 8.6 the Bethe potentials are, in principle, not valid for very thin crystals. In the phase-object approximation, which is valid for very thin crystals, the ratio of the first and second order terms (from (11.44)) in the expansion of the structure amplitude is proportional to the thickness. Hence some modified form of the Bethe potentials must be used (Horstmann and Meyer [1965]; Gjønnes [1962]).

The expression (11) has been used as the basis for a sort of "extinction correction" in a great deal of structure analysis work based on ring or arc patterns from polycrystalline materials (Vainshtein [1964]). It has been assumed that the relationship of the integrated intensity of a reflection to the structure amplitude $|\Phi_h|$ varies from a second order to a first-order dependence with increasing strength of the dynamical effects and that for a particular diffraction pattern the appropriate fractional order dependence can be determined and used as a basis for the interpretation of the intensities. The use of Bethe potentials to apply corrections for some particular intensities offers some improvement on this somewhat oversimplified concept for crystals which are not too thin.

16.3.6. *n*-beam diffraction effects

It has been argued with some justification that the structures of crystals which are of interest for structural investigations are not usually simple and may have many atoms in a relatively large unit cell. Then the absolute values of the structure amplitudes are less and the dynamical effects correspondingly weaker than for the metals or other crystals of simple structure normally used to test the dynamical theory expressions. On the other hand, for substances having relatively large unit cells the density of reciprocal lattice points is high, the number of simultaneous reflections for any orientation of a crystal is large and the *n*-beam dynamical effects may be important even when the individual reflections are relatively weak. It is therefore important for structure analysis applications that the effects of *n*-beam diffraction effects on ring and arc patterns should be evaluated.

In the absence of any convenient analytical method of integrating *n*-beam dynamical intensities over orientation and thickness, the only feasible procedure appears to be the very laborious one of calculating intensities for a sufficiently finely spaced set of orientations. An attempt to do this was made by Turner and Cowley [1969] who made *n*-beam calculations of intensities of arc patterns for thin BiOCl crystals and for the substance AgTlSe₂ for which the structure analysis had been performed by Imamov and Pinsker [1965]. The experimental measurements had been analysed on the basis of the two-beam dynamical approximation and it was concluded from this that the intensities were purely kinematical. However, the *n*-beam dynamical calculations showed that there was probably sufficient dynamical modification of the intensities to have the effect of introducing appreciable errors in the details of the structure deduced.

Later Imamov et al. [1976] made systematic tests to reveal the presence of *n*-beam dynamical interactions. The zero

layer-line intensities of oblique texture patterns of PbSe and Bi₂Se₃ were measured for various angles of tilt of the axis of preferred orientation and thus for various strengths of *n*-beam interaction with upper layer line reflections. For some particular tilts the zero-layer line intensities were modified by 40 percent or more. A more detailed study with close coordination of the theoretical calculations and the experimental measurements, performed for a variety of materials, would be of considerable value.

Ordering of atoms in crystals

17.1. The nature and description of disordered states

The disordered arrangement of different kinds of atoms or molecules on the lattice sites of a crystal is of interest to a wide variety of scientists for a very varied set of reasons. For the theoretical physicist it represents one example of the three-dimensional lattice ordering problem, similar to the ordering of spins in a ferromagnet. The analysis of the statistical mechanics of ordering starts with the idealized Ising model and gets little further except for approximate or asymptotic solutions (Brout [1965]).

From this point of view, the interest is in the simplest possible ordering systems, such as that of β -brass, in which Cu and Zn atoms occupy the two sites of a body-centered cubic (B.C.C.) lattice in a disordered way above the critical temperature and with partial ordering over large distances (many unit cells) below T_C (see Fig. 17.1). From the completely ordered state, the equilibrium state at 0 K, the disordering increases more and more rapidly with temperature until it becomes catastrophic at T_C , giving ideally, a second-order phase transformation.

For the metallurgist ordering is an important parameter in determining the physical properties of materials such as the electrical and thermal resistivity, the hardness, ductility and so on. From the high-temperature disordered state the ordering may proceed in various ways. At one extreme, unlike atoms may tend to alternate, with each atom tending to surround itself by unlike neighbors and the system moves towards an ordered superlattice. At the other extreme atoms of the same kind may tend to clump together, leading to a segregation of phases at lower temperature. Various intermediate situations provide materials of great commercial significance (Cohen [1968]).

High temperature chemists are concerned usually with the ordering of only a fraction of the ions in a crystal. For example the anions may have a disordered distribution among the possible sites defined by an ordered cation lattice, as in the case of many complex oxide or sulfide phases. The interest in the nature and amount of ordering derives from its relationship to the non-stoichiometry and phase transitions of such structures (Anderson [1969]).

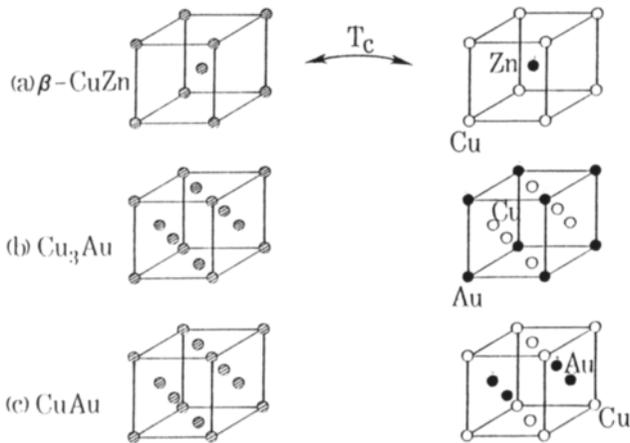


Fig. 17.1. The disordered and ordered structures of (a) β -brass, (b) Cu_3Au and (c) CuAu .

Crystallographers involved with structure analysis of crystals have, in the past, tended to regard disorder chiefly as a hindrance to their purpose and have avoided dealing with disordered systems. Recently the increasing evidence for the widespread occurrence of partially or imperfectly ordered materials and the increasing importance of such materials in many areas of science has forced them to consider the implications for structure analysis very seriously. The disorder may be of many types: one or more types of atom may be disordered, or replaced by vacancies, on the lattice sites. There may be static or dynamic displacements of atoms sometimes correlated over large distances to form partially or fully ordered superstructures, commensurate or incommensurate with the basic sublattice. Molecules or parts of molecules may show disordered or partially or fully correlated rotations or distortions. Accounts of many such phenomena are included in the conference report edited by Cowley et al. [1979].

For diffraction physicists, disorder represents an interesting example of diffraction from an imperfect crystal coming within the first of the two main classes discussed in Chapter 7. We treat it as such here, adding some comments on the additional considerations involved with the practical complications of dynamical scattering and the possible implications of the combination of scattering theory with statistical mechanics. For simplicity we confine our attention to simple binary alloy solid solutions, composed of A and B atoms in fractional proportions m_A and m_B . The

disordered alloys are assumed to have simple structures such as B.C.C. as for β -CuZn or F.C.C. as for Cu-Au alloys (Fig. 17.1a, b). The extension to systems with more than two types of atoms and of more complicated structure follows with elaboration of the algebra but little conceptual innovation.

17.2. Order parameters

17.2.1. Short-range order

The state of order of the crystal is defined by order parameters which define correlations between the occupancy of sites and so specify, for example, the probability of finding a B atom at a given distance (vectorial) away from an A atom. This probability is written P^{AB}_{ij} where i and j signify the sites specified by the vectors r_i and r_j from some origin.

We start by defining occupation parameters (Flinn [1956])

$$\begin{aligned}\sigma_i^A &= \begin{cases} 1 & \text{for an A at site } i, \\ 0 & \text{for a B at site } i, \end{cases} \\ \sigma_i^B &= \begin{cases} 1 & \text{for a B at site } i, \\ 0 & \text{for an A at site } i. \end{cases}\end{aligned}\quad (1)$$

Since these parameters are interconnected we may replace both by a single parameter

$$\sigma_i = \begin{cases} m_B & \text{for an A at } i \\ -m_A & \text{for a B at } i, \end{cases}\quad (2)$$

$$\begin{aligned}\sigma_i &= \sigma_i^A - m_A = m_B - \sigma_i^B, \\ \langle \sigma_i \rangle &= 0,\end{aligned}\quad (3)$$

where the $\langle \rangle$ brackets are used to denote an average over all sites.

The probability P^{AB}_{ij} is readily seen to be given by

$$\begin{aligned}P^{AB}_{ij} &= \langle \sigma_i^A \sigma_j^B \rangle \\ &= m_A m_B - \langle \sigma_i \sigma_j \rangle.\end{aligned}\quad (4)$$

Similarly

$$P^{AA}_{ij} = \langle \sigma_i^A \sigma_j^A \rangle = m_A^2 + \langle \sigma_i \sigma_j \rangle.\quad (5)$$

Then the average $\langle \sigma_i \sigma_j \rangle$ is the pair-correlation parameter, giving the difference between the probability for a particular pair of atoms occurring at a distance $r_i - r_j$ apart, and the probability for a completely random array.

In the same way we may define a three-atom correlation parameter $\langle \sigma_i \sigma_j \sigma_k \rangle$ derived from the probability of occupation of three sites by particular types of atom. For example,

$$\begin{aligned} P^{ABB}_{ijk} &= \langle \sigma_i^A \sigma_j^B \sigma_k^B \rangle \\ &= m_A m_B^2 + m_A \langle \sigma_j \sigma_k \rangle - m_B \langle \sigma_i \sigma_k \rangle - m_B \langle \sigma_i \sigma_j \rangle - \langle \sigma_i \sigma_j \sigma_k \rangle. \end{aligned} \quad (6)$$

Manipulation of the σ_i operators is aided by use of the relationship (Shirley and Wilkins [1972])

$$(\sigma_i)^n = A_n + B_n \sigma_i,$$

where

$$\begin{aligned} A_n &= \{m_A m_B^n + (-1)^n m_B^n m_A\}, \\ B_n &= m_B^n - (-m_A)^n. \end{aligned} \quad (7)$$

This expression allows us to assign values to correlation parameters when two or more of the sites coincide: for example,

$$\begin{aligned} \langle \sigma_i \sigma_i \rangle &= \langle \sigma_i^2 \rangle = m_A m_B, \\ \langle \sigma_i \sigma_j \rangle &= (m_B - m_A) \langle \sigma_i \sigma_j \rangle, \end{aligned} \quad (8)$$

and so on.

The two-atom correlation parameters, or "Warren short-range order-parameters", commonly used are α_{ij} , given by

$$\langle \sigma_i \sigma_j \rangle = m_A m_B \alpha_{ij}. \quad (9)$$

The whole set of correlation parameters, for all numbers of atoms, can be considered to specify the state of order of the system. Usually, however, only two-atom correlations are considered. The values of $\langle \sigma_i \sigma_j \rangle$ specify the degree to which the neighbors of one sort of atom tend to be preferably of the same sort or of the opposite sort. If $\langle \sigma_i \sigma_j \rangle$ is positive for all short vector lengths $r_i - r_j (=r_{ij})$, it is seen from (5) that like atoms tend to clump together. If $\langle \sigma_i \sigma_j \rangle$ alternates positive and negative with increasing $|r_{ij}|$ values, being negative for nearest neighbors, there is a tendency for the two sorts of atom to form an ordered superlattice.

Above the critical temperature for ordering, the values of $\langle \sigma_i \sigma_j \rangle$ decrease rapidly with r_{ij} , typically falling off by an order of magnitude in a few unit cell lengths. This indicates a state of "short-range order" (s.r.o.). The transition to long-range order (l.r.o.) as the crystal is cooled through T_C implies that the range of the correlation

between the occupation of lattice sites extends effectively to infinity. The values of the two-atom correlation parameters decrease over short distances but then tend to limiting, finite values as a regular superlattice is formed with particular atoms predominating on particular sites of the superlattice unit cell.

17.2.2. Long-range order

The limiting values of the correlation parameters are not the same for vectors from a given lattice site to the different sites of the superlattice unit cell. For example for the F.C.C. lattice of a Cu_3Au type structure, the sites given by the vectors $h_1 a_1 + h_2 a_2 + h_3 a_3$ have one limiting value, say s_1 , if a_1, a_2, a_3 are the unit cell axial vectors and h_1, h_2, h_3 are integers, while for vectors $h_1 a_1 + h_2 a_2 + h_3 a_3$ plus $(a_1 + a_2)/2$ or $(a_2 + a_3)/2$ or $(a_3 + a_1)/2$, the limiting value will be equal to $s_2 \neq s_1$.

It is customary to assume a fixed ratio of s_1 to s_2 , namely $s_1 = -3s_2$ for this case, although this involves an additional assumption of uniform composition (Cowley [1965a]). With this assumption, the s_1 and s_2 values are simply related to the square of the traditional Bragg-Williams l.r.o. parameter S , defined as a measure of the fraction, r_α , of atoms sitting on their "correct" lattice sites. For the case of Cu_3Au structure (Cowley [1950])

$$S = \frac{3}{4} \left(\frac{r_\alpha - m_A}{1 - m_A} \right) + \frac{1}{4} \left(\frac{r_\beta - m_B}{1 - m_B} \right) \quad (10)$$

and

$$s_1 = \lim_{r_{ij} \rightarrow \infty} \left\{ \frac{\langle \sigma_i \sigma_j \rangle}{m_A m_B} \right\} = \frac{16}{3} m_A m_B S^2. \quad (11)$$

It is usually assumed for convenience that a state of l.r.o. involves correlations over infinite distances. In practice the correlations do not extend beyond the limits set by grain boundaries, crystal defects, or the anti-phase boundaries to be described later. Hence the correlation functions actually decrease with a half-width of the average size of the perfect single crystal regions.

17.3. Patterson function

If, instead of site occupancies, we consider correlations in terms of specific continuous functions which may imply the site occupancy, such as the electron density function $\rho(r)$ (or nuclear density function or potential) we may recall that correlations may be described by the Patterson function

$$P(r) = \rho(r) * \rho(-r). \quad (12)$$

If $\rho(r)$ is considered to consist of symmetrical contributions from each atom, $\rho_A(r)$ or $\rho_B(r)$, the Patterson becomes

$$\begin{aligned} \rho(r) * \rho(-r) &= \sum_i (\sigma^A_i \rho_A + \sigma^B_i \rho_B) * \delta(r - r_i) * \sum_j (\sigma^A_j \rho_A + \sigma^B_j \rho_B) * \delta(r + r_j) \\ &= \sum_i \sum_j [\sigma^A_i \sigma^A_j (\rho_A * \rho_A) + \sigma^B_i \sigma^B_j (\rho_B * \rho_B) + 2\sigma^A_i \sigma^B_j (\rho_A * \rho_B)] \\ &\quad * \delta(r - r_i + r_j). \end{aligned} \quad (13)$$

This is expressed in terms of σ_i and σ_j by use of the equation (3). The double summation may be replaced by the number of atoms, N , times the average surroundings of a particular site. Putting $r_j = 0$ in (13) then gives, using (3),

$$\begin{aligned} \rho(r) * \rho(-r) &= N \sum_i [(\rho_A * \rho_A) (m^2_A + \langle \sigma_0 \sigma_i \rangle) + (\rho_B * \rho_B) (m^2_B + \langle \sigma_0 \sigma_i \rangle) \\ &\quad + 2(\rho_A * \rho_B) (m_A m_B - \langle \sigma_0 \sigma_i \rangle)] * \delta(r - r_i) \\ &= N \sum_i [(m_A \rho_A + m_B \rho_B) * (m_A \rho_A + m_B \rho_B) \\ &\quad + \langle \sigma_0 \sigma_i \rangle \{(\rho_A - \rho_B) * (\rho_A - \rho_B)\}] * \delta(r - r_i). \end{aligned} \quad (14)$$

This is in the form of the Patterson of the average disordered lattice plus the Patterson of the deviations from the average lattice since, from (4), $\langle \sigma_0 \sigma_i \rangle$ represents the deviation from the average probability for site-pair occupancies.

17.4. Size effects

In making the transition from (13) to (14) however, we have made the important assumption that the set of r_i vectors is the same when referred to any member of the set i.e. that the vectors are the vectors R_i between lattice points of a perfect crystal lattice. This is so only in the limiting case that the two types of atom may be interchanged without any perturbation of atom positions, i.e. are of exactly the same size. Only in rare cases is this true, even to within one percent. For example for Cu-Au alloys although the effective radius of atoms in alloys is not necessarily that in the monatomic crystals, the difference of about 10 percent between the lattice constants for Cu and Au (3.65 and 4.07 Å resp.) suggests a considerable difference in effective atom size in the alloys.

If a single large atom is inserted into a lattice of smaller atoms one would expect it to displace the neighboring atoms or, more precisely, the atom has a displacement field, displacing all atoms around it by amounts and in directions depending on their vector separations from it. A smaller atom than the average may be expected to have a displacement field of opposite sign.

The extension of this concept to the case of binary alloys in which every atom is either larger or smaller than the average is not obvious. As a first approximation it may be assumed that each atom produces its own displacement field, acting on all other atoms and depending only on their vectorial separations from it, not on their natures. With a further stage of approximation we can assume that the displacement of one atom by another depends not on the actual vector separation, $r_i - r_j$, but on the separation of the nearest lattice points, which we may write as $R_i - R_j$. Then we assume that the total displacement of any one atom from the lattice point of the average lattice for the crystal is given by the sum of the displacements due to all other atoms. The vector displacement of the atom from the i lattice point is thus:

$$\Delta_i = \sum_k (\sigma_k^A \Delta_{ik}^A + \sigma_k^B \Delta_{ik}^B). \quad (15)$$

The displacement Δ_{ik}^A is for an A atom at k and Δ_{ik}^B is for a B atom at k . It may further be assumed that for an average atom at k the displacement would be zero i.e.

$$m_A \Delta_{ik}^A + m_B \Delta_{ik}^B = 0. \quad (16)$$

Then the displacements due to A and B atoms are collinear and

$$\Delta_{ik}^B = -(m_A/m_B) \Delta_{ik}^A$$

and (15) becomes

$$\Delta_i = \sum_k \sigma_k \Delta_{ik} \quad (17)$$

where

$$\Delta_{ik} = \Delta_{ik}^A / m_B.$$

Hence the delta function in (13) is replaced by

$$\delta \left(r - R_i + R_j - \sum_k \sigma_k (\Delta_{ik} - \Delta_{jk}) \right),$$

and when the averaging process is performed, instead of (14) we obtain

$$\begin{aligned}
 P(r) = & N \sum_i \left[(m_A \rho_A + m_B \rho_B) * (m_A \rho_A + m_B \rho_B) \right. \\
 & * \left. \left\langle \delta \left\{ r - R_i - \sum_k \sigma_k (\Delta_{ik} - \Delta_{jk}) \right\} \right\rangle \right. \\
 & + \left. \left\langle (m_A \rho_A + m_B \rho_B) * (\rho_A - \rho_B) (\sigma_i + \sigma_0) * \delta \left\{ r - R_i - \sum_k \sigma_k (\Delta_{ik} - \Delta_{jk}) \right\} \right\rangle \right. \\
 & \left. + \left. \left\langle (\rho_A - \rho_B) * (\rho_A - \rho_B) (\sigma_0 \sigma_i) * \delta \left\{ r - R_i - \sum_k \sigma_k (\Delta_{ik} - \Delta_{jk}) \right\} \right\rangle \right]. \quad (18)
 \end{aligned}$$

Thus for the Patterson of the average lattice, the first term of (18), there is a spread of the peaks around the average lattice vector positions, R_i . For the contributions from the deviations from the average lattice, the Patterson peaks are likewise spread but the displacements given by individual atom-pair vectors from the R_i positions are correlated with the site occupancies so that there can be displacements of the mean positions of the peaks as well as an order-dependent broadening of the peaks.

As indicated earlier, this simplified formulation of the Patterson function rests on a number of assumptions which are not necessarily justified, but it can serve to illustrate the type of diffraction effects to be observed.

17.5. Kinematical diffraction

17.5.1. Diffraction with ordering only

As a basis for obtaining the intensities in diffraction patterns under the assumptions of kinematical diffraction conditions we obtain the scattering power distributions in reciprocal space by Fourier transform of the Patterson function.

In the absence of size effect, Fourier transform of (14) gives simply,

$$\begin{aligned}
 I(u) = & N(m_A f_A + m_B f_B)^2 \sum_i \exp\{2\pi i u \cdot R_i\} \\
 & + N(f_A - f_B)^2 \sum_i \langle \sigma_0 \sigma_i \rangle \exp\{2\pi i u \cdot R_i\}, \quad (19)
 \end{aligned}$$

where f_A and f_B are the atomic scattering factors for the radiation used and may be assumed to include the temperature factors such as $\exp\{-M_A\}$. The first term of (19) gives the delta-function peaks at the reciprocal lattice points for the average lattice which produce the so called "fundamental" diffraction peaks which are independent of ordering. The second term, due to the deviations from the average lattice,

is order dependent, giving rise to diffuse scattering and is written separately as I_d .

In the limiting, high temperature case of complete randomness all $\langle \sigma_0 \sigma_i \rangle$ values are zero except that, from (8), $\langle \sigma_0 \sigma_0 \rangle = m_A m_B$ and the second term of (19) becomes

$$I_d = N m_A m_B (f_A - f_B)^2, \quad (20)$$

which gives rise to a uniform background scattering, falling off smoothly with scattering angle roughly in proportion to f^2 .

Correlation of occupancy of neighboring sites gives a modulation of this background. For a tendency for clustering of like atoms, all $\langle \sigma_0 \sigma_i \rangle$ values tend to be of the same sign so that maxima in I_d appear around the reciprocal lattice points for the average structure. If there is a tendency for ordering into a superlattice, with an alternation of the two

Table 1
Measured values of i (for the i th neighbor shell)
given by Chen et al. [1979] for Cu_3Au

Shell number	Typical coordinates	Perfect order	$T=396$ ($^{\circ}\text{C}$)	α_i $T=420$ ($^{\circ}\text{C}$)	$T=685$ ($^{\circ}\text{C}$)
i	lmn	$\alpha_i(T=0)$			
1	110	-1/3	-0.176	-0.125	-0.130
2	200	1	+0.214	0.154	0.106
3	211	-1/3	+0.005	0.023	0.032
4	220	1	+0.062	0.049	0.019
5	310	-1/3	-0.079	-0.078	-0.066
6	222	1	+0.022	0.010	-0.009
7	321	-1/3	-0.010	-0.012	-0.002
8	400	1	+0.073	0.071	0.029
9	330	-1/3	-0.030	-0.022	-0.000
10	411	-1/3	+0.026	0.018	0.007
11	420	1	+0.034	0.024	0.000

kinds of atom on lattice sites, $\langle \sigma_0 \sigma_i \rangle$ are negative for nearest neighbors, positive for second nearest neighbors and so on. Table 1 gives the values of the order parameters, α_{0i} ($= (16/3) \langle \sigma_0 \sigma_i \rangle$), for Cu_3Au at three temperatures above T_C (394°C), given by Chen et al. [1979] and Fig. 17.2 shows the corresponding reciprocal space distribution $I_d(u)$.

Broad maxima of scattering power appear at the 100, 110 and similar reciprocal lattice points which are the positions of delta function maxima for the ordered Cu_3Au lattice, Fig. 17.1(b) but not for the average lattice i.e. at the positions of the sharp superlattice reflections of the ordered state.

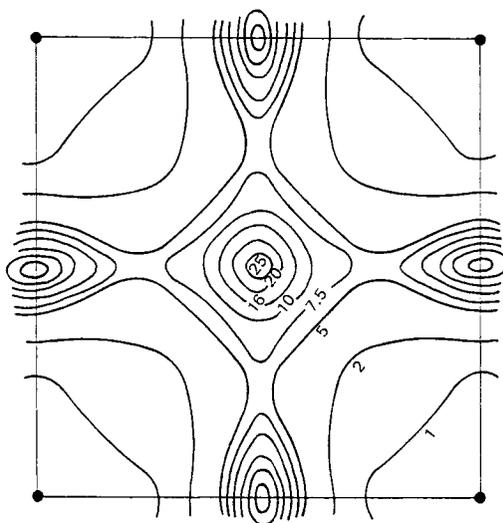


Fig. 17.2. Contour map of diffuse scattering intensity due to short-range order in the $hk0$ reciprocal lattice plane for a disordered crystal of Cu_3Au (after Cowley [1950b]).

Below the critical temperature, T_C , the values of $\langle \sigma_0 \sigma_i \rangle$ tend to constant limiting values $m_A m_B s_i$ as R_i becomes large. Then we may separate out the contributions from these limiting values and write

$$I_d(\mathbf{u}) = N m_A m_B (f_A - f_B)^2 \sum_i s_i \exp\{2\pi i \mathbf{u} \cdot \mathbf{R}_i\} + N (f_A - f_B)^2 \sum_i \{\langle \sigma_0 \sigma_i \rangle - m_A m_B s_i\} \exp\{2\pi i \mathbf{u} \cdot \mathbf{R}_i\}. \quad (21)$$

The first part of this expression comes from a periodic structure and gives the sharp superlattice peaks. The second part includes the differences of the $\langle \sigma_0 \sigma_i \rangle$ from their limiting values, which fall off rapidly with distance and so give rise to some residual diffuse scattering. In the limiting case of zero temperature the equilibrium structure would be perfectly ordered. Both $\langle \sigma_0 \sigma_i \rangle$ and $m_A m_B s_i$ would tend to the values for the perfect superlattice structure and the second term of (21) would vanish.

The integrated intensity of the superlattice reflections are proportional to the s_i values and so, from (11) to S^2 . Hence the Bragg-Williams l.r.o. parameter S may be found (apart from an ambiguity of sign) from measurements of intensity of superlattice reflections. In this way accurate values have been obtained for the l.r.o. parameter for β -brass and compared with the prediction of the Ising model and various approximate, although more realistic theories (Chipman and Walker [1972]).

From the general expression (19) it is seen that the values of the correlation coefficients may be derived from the diffuse scattering by Fourier transform:

$$\langle \sigma_0 \sigma_i \rangle = \int \frac{I_d(u)}{N(f_A - f_B)^2} \exp\{-2\pi i u \cdot R_i\} du, \quad (22)$$

where the integral is taken over one unit cell in reciprocal space. In this way the values of α_i in Table 1 and similar results were obtained from single-crystal X-ray diffraction measurements, after correction of the data for Compton scattering, thermal diffuse scattering, instrumental background and the size effects discussed below.

17.5.2. Diffraction with ordering and size effects

When size effects are present, the Patterson function, to a first approximation, is given by (18). The reciprocal space scattering power given by Fourier transform, is then written,

$$\begin{aligned} I(u) &= N(m_A f_A + m_B f_B)^2 \sum_i \exp\{2\pi i u \cdot R_i\} \\ &\times \langle \exp\left\{2\pi i u \cdot \sum_k \sigma_k (\Delta_{ik} - \Delta_{jk})\right\} \rangle + N(m_A f_A + m_B f_B) (f_A - f_B) \\ &\times \sum_i \exp\{2\pi i u \cdot R_i\} \langle (\sigma_i + \sigma_0) \exp\left\{2\pi i u \cdot \sum_k \sigma_k (\Delta_{ik} - \Delta_{jk})\right\} \rangle \\ &+ N(f_A - f_B)^2 \sum_i \exp\{2\pi i u \cdot R_i\} \langle \sigma_0 \sigma_i \exp\left\{2\pi i u \cdot \sum_k \sigma_k (\Delta_{ik} - \Delta_{jk})\right\} \rangle. \end{aligned} \quad (23)$$

The averages in this expression are not easy to evaluate. One method for finding the scattering would be to do the averaging in the real space of the Patterson function, and for each Patterson function peak find the displacement from the lattice site and the spread function which broadens it. Then the Fourier transform is evaluated in detail.

The alternative, which has been most widely used, is to assume that all displacement terms Δ_{ik} are small and to expand the exponentials in (23) in power series. It is necessary to take into account at least the second order terms in order to include qualitatively all significant diffraction effects.

From the first term in (23) we obtain for the fundamental reflections,

$$\begin{aligned} I(u) &= N(m_A f_A + m_B f_B)^2 \sum_i \exp\{2\pi i u \cdot R_i\} \\ &\times \left[1 - 2\pi^2 \sum_k \sum_l \langle \sigma_k \sigma_l \rangle \{u \cdot (\Delta_{ik} - \Delta_{0k})\} \{u \cdot (\Delta_{il} - \Delta_{0l})\} + \dots \right]. \end{aligned} \quad (24)$$

From the last bracket we may separate out the terms independent of the vector R_i . Since by symmetry the summations over k and l of $(u \cdot \Delta_{ik})(u \cdot \Delta_{il})$ must be the same as the summations of $(u \cdot \Delta_{0k})(u \cdot \Delta_{0l})$, these are

$$\left[1 - 4\pi^2 \sum_k \sum_l \langle \sigma_k \sigma_l \rangle (u \cdot \Delta_{0k})(u \cdot \Delta_{0l}) + \dots \right]. \quad (25)$$

This has the form of an order-dependent effective Debye-Waller factor, $\exp \{-M'\}$. In the absence of any s.r.o. it becomes

$$1 - 4\pi^2 \sum_k (u \cdot \Delta_{0k})^2 + \dots, \quad (26)$$

and the second part is just the sum of the squares of all displacements of atoms from the lattice sites. With s.r.o. present this effective Debye-Waller factor in general comes closer to unity, which must be its value for a completely ordered lattice.

The terms of the last bracket of (24) which are dependent on the vectors R_i , such as

$$2\pi^2 \sum_k \sum_l \langle \sigma_k \sigma_l \rangle (u \cdot \Delta_{ik})(u \cdot \Delta_{0l}) \quad (27)$$

are centrosymmetric and decreasing more-or-less uniformly with $|R_i|$ and so give rise to a diffuse scattering around the fundamental reciprocal lattice points, the Huang scattering (Huang [1947]).

For the second term of (23), the average is rewritten

$$\begin{aligned} \langle \dots \rangle &= 0 + 2\pi i \sum_k \{ \langle \sigma_i \sigma_k \rangle + \langle \sigma_0 \sigma_k \rangle \} \{ u \cdot (\Delta_{ik} - \Delta_{0k}) \} \\ &- 2\pi^2 \sum_k \sum_l \{ \langle \sigma_i \sigma_k \sigma_l \rangle + \langle \sigma_0 \sigma_k \sigma_l \rangle \} \{ u \cdot (\Delta_{ik} - \Delta_{0k}) \} \{ u \cdot (\Delta_{il} - \Delta_{0l}) \} - i \dots \end{aligned} \quad (28)$$

and from the third term we obtain

$$\begin{aligned} \langle \dots \rangle &= \langle \sigma_0 \sigma_i \rangle + 2\pi i \sum_k \langle \sigma_0 \sigma_i \sigma_k \rangle \{ u \cdot (\Delta_{ik} - \Delta_{0k}) \} \\ &- 2\pi^2 \sum_k \sum_l \langle \sigma_0 \sigma_i \sigma_k \sigma_l \rangle \{ u \cdot (\Delta_{ik} - \Delta_{0k}) \} \{ u \cdot (\Delta_{il} - \Delta_{0l}) \} + \dots \end{aligned} \quad (29)$$

Apart from the s.r.o. diffuse scattering of equation (19) given by the first term in (29), we now have other terms dependent on both correlation parameters and pair-wise displacement vectors. The antisymmetric terms containing $(2\pi i)$ give rise to antisymmetric diffuse scattering around each fundamental reciprocal lattice peak except the origin, giving a displacement of the apparent center of the Huang scattering. Also they provide antisymmetric contributions to the diffuse scattering around the s.r.o. diffuse maxima and this produces an apparent displacement of the s.r.o. diffuse peaks by an amount increasing with $|\mathbf{u}|$. This effect was first noticed and interpreted as due to size-effect displacements by Warren et al. [1951].

The next terms in (28) and (29) with double summations over k and l may be resolved in much the same way as the similar, simpler term in (24) into the equivalent of a Debye-Waller factor on the s.r.o. diffuse scattering and further symmetrical diffuse scattering terms.

It may be noted that three-atom correlation parameters enter into both the symmetrical and antisymmetrical contributions to the diffuse scattering at the first stage beyond the elementary s.r.o. and size-effect contributions. With expansion of the exponentials to higher powers of $(\mathbf{u} \cdot \Delta)$, terms containing all multiple-atom correlation parameters are introduced (Cowley [1968]). The complications introduced with consideration of size effect prevent the straight-forward derivation of correlation parameters as in (22).

The size effect perturbation of the s.r.o. scattering is small when the displacements Δ_{ik} are small and the range of \mathbf{u} values is restricted to the first one or two reciprocal lattice unit cells about the origin. Under these conditions it has been shown by Borie and Sparks [1971] and Gragg [1970] that the size effect contributions can be separated by making use of the differences in their symmetry characteristic from those of the s.r.o. scattering. In this way diffraction data containing size effects can be used in the determination of both the $\langle \sigma_0 \sigma_i \rangle$ and parameters characterizing the atom displacements. The methods used have been summarized by Bardhan and Cohen [1976] and Chen et al. [1979].

As a function of the reciprocal space coordinates u, v, w , the diffuse scattering intensity including up to second order terms in interatomic displacements, can be written

$$I_D(uvw) = N m_A m_B |f_A - f_B|^2 [I_{SRO} + uQ_x + vQ_y + wQ_z + u^2R_x + v^2R_y + w^2R_z + uvS_{xy} + vwS_{yz} + wuS_{zx}].$$

Here $m_A m_B I_{SRO}$ is the last part of (19), the scattering due to short range order in the absence of size effects. The quantities Q, R, S are introduced by the size effects and can

be expressed in terms of parameters such as $\gamma_i^x, \delta_i^x, \epsilon_i^x$, which depend on combinations of the short range order parameters α_i or $\langle \sigma_0 \sigma_i \rangle$, the displacement parameters x_{0i}^{AA}, y_{0i}^{AB} , etc. which are the x, y components of the displacements of an atom at r_i due to an atom at the origin, and also the parameters of the type $\langle x_{0i}^A y_{0i}^B \rangle$ which measures the correlation of the y displacement of a B atom at r_i and the x displacement of an A atom at r_0 . The parameters γ, δ, ϵ also depend on the ratios of scattering amplitudes $\eta = f_A / |f_A - f_B|$ and $\zeta = f_B / |f_A - f_B|$.

The quantity Q_x is written

$$Q_x = \sum_i \gamma_i^x \sin(2\pi u x_i) \cos(2\pi v y_i) \cos(2\pi w z_i)$$

where x_i, y_i, z_i are the components of r_i in terms of unit cell dimensions, and so represents contributions to the diffuse scattering which are antisymmetric about the reciprocal lattice points in the x -direction, contained in our expressions (28) and (29). The function R_x gives scattering centrosymmetric about the reciprocal lattice points and so describes the Huang scattering, while S_{xy} gives terms antisymmetric in both x and y directions.

If the statistical symmetry of the disordered system is used, so that for a cubic system the x, y, z and u, v, w dependencies can be assumed to be equivalent, it is possible to derive the functions I_{SRO}, Q_x, R_x and S_{xy} from linear combinations of measurements of the observed diffuse intensity from symmetry-related points in reciprocal space. Then if it can be assumed that the scattering amplitude ratios η and ζ are constants, independent of the reciprocal space coordinates, the values of the parameters α (or $\langle \sigma_0 \sigma_i \rangle$), γ, δ and ϵ may be derived by Fourier transform, as described by Schwartz and Cohen [1977].

The assumption that η and ζ are constants is obviously good for neutron diffraction. For X-ray diffraction, the approximation may be poor and may lead to serious errors in some cases, especially for clustering of atoms in the presence of appreciable size effects.

An alternative method of analysis has been suggested by Tibballs [1975] and developed into a practical procedure by Georgopolous and Cohen [1977]. In this, η and ζ are separated from the Q, R, S functions and inserted explicitly. The short range order and size effect parameters may then be obtained directly by solving a large array of linear relationships by least squares methods. This technique works very well for X-ray measurements for which varying values of η and ζ provide sufficiently large numbers of independent linear relationships. It does not work for neutron diffraction.

These methods of analysis have not been extended beyond the terms of second order in displacement parameters or to include higher order correlation parameters. To do so would complicate the analysis enormously. No systematic method has been devised to evaluate correlation parameters such as $\langle \sigma_0 \sigma_i \sigma_j \rangle$ or parameters involving correlations of greater numbers of atoms. It is not clear to what extent terms containing these parameters and occurring in equations (24) to (29) do affect the intensities in any practical cases or to what extent omission of these terms modifies the values of the parameters derived from treatments limited to second order terms. Clearly such terms become more important for larger values of $|u|$.

The values of the size effect parameters derived experimentally from the diffraction methods are as yet insufficient to provide very clear pictures of the displacement fields around atoms. Some earlier calculations of the size effects to be expected were made using a simple model for the displacements, namely that the displacements are radial, falling off with the inverse square of the distance from the origin as in the macroscopic case of the perturbation of an isotropic solid by a center of dilation (Borie [1957]). All available experimental evidence and approximate theoretical models suggest that the displacement fields of point defects are not isotropic and vary in a complicated way with distance and direction from the defect, and the effective displacement fields to be used for a disordered alloy should no doubt be similar.

The amount of experimental data which could be gathered should, in principle, be sufficient to allow derivation of considerably more than the s.r.o. parameters. To assist this process the scattering power $I(u)$ could be measured for the whole range of u and also other factors could be varied. Variations of composition of alloys could be useful since, for the 50:50 composition for example, the three-atom correlation parameters $\langle \sigma_0 \sigma_i \sigma_k \rangle$ are all zero. The relative values of the structure factors f_A and f_B are different for different radiations. Thus for X-rays f_{Cu} and f_{Au} are widely different but f_{Cu} and f_{Zn} are almost the same. For neutron diffraction the situation is reversed. The terms of (23) containing $(f_A - f_B)$ almost disappear for neutron diffraction from Cu-Au alloys or for X-ray diffraction from β -CuZn.

On the other hand for some alloys containing atoms with negative scattering lengths it is possible to adjust the composition or isotopic abundances in such a way that for neutron diffraction the average scattering factor $(m_A f_A + m_B f_B)$ is zero, as in the case of Cu-Ni alloys studied by Moser et al. [1968]. For such "null-matrix" alloys, the fundamental reflections and their attendant thermal diffuse and Huang scattering are zero and only the third part of (23) remains.

17.6. Relationship with ordering energies

The tendency for either like or unlike atoms to come together, giving either clustering or ordering towards a superlattice is a consequence of the balance of the interaction energies between atoms. It is customary to make the approximation, at least for alloys, that the energy for a configuration of atoms is the sum of interactions between pairs of atoms. The configurational energy is then written

$$U = N \{ P^{AA}_{ij} V^{AA}_{ij} + P^{BB}_{ij} V^{BB}_{ij} + P^{AB}_{ij} V^{AB}_{ij} + P^{BA}_{ij} V^{BA}_{ij} \}, \quad (30)$$

where V^{AB}_{ij} is the interaction energy for an A atom at the position i and a B atom at the position j .

By use of the expressions (4) and (5) we obtain

$$U = \frac{1}{2} \sum_i \sum_j (m_A^2 V^{AA}_{ij} + 2m_A m_B V^{AB}_{ij} + m_B^2 V^{BB}_{ij}) + \sum_i \sum_j \langle \sigma_i \sigma_j \rangle V_{ij} \quad (31)$$

where

$$V_{ij} = \frac{1}{2} \{ (V^{AA}_{ij} + V^{BB}_{ij} - 2V^{AB}_{ij}) \},$$

which is the increase in energy when like atom pairs replace unlike atom pairs. The first part of (31) is order-independent and can be ignored.

The problem of finding the correlation parameters as a function of temperature for given values of V_{ij} is an unsolved problem of statistical mechanics, equivalent in its simplest form to the three-dimensional Ising model problem. However for temperatures above the critical ordering temperature T_C , several approximate solutions are available (Brout [1965]) and from these can be obtained the relationship (Clapp and Moss [1968])

$$\alpha(k) = \frac{G_2(T)}{1 + G_1(T)V(k)}, \quad (32)$$

where we have used the reciprocal space functions

$$\alpha(k) = \sum_i \alpha_{0i} \exp\{ik \cdot r_{0i}\}, \quad (33)$$

$$V(k) = \sum_i V_{0i} \exp\{ik \cdot r_{0i}\}, \quad (34)$$

The function $G_1(T)$ is approximately proportional to $(T/T_C)^{-1}$ and $G_2(T)$ is almost a constant, being only weakly dependent on T . From (19) it is seen that $\alpha(k)$ is directly proportional to the component of the diffuse scattering intensity due to short-range ordering, and so is an

observable quantity.

From (32) $V(\mathbf{k})$ and hence the interaction energies may be obtained directly from the diffuse scattering measurements. For reasons of practical convenience the determinations of interaction energies to date have relied on the real-space relations between α_{ij} and V_{ij} , equivalent to (32).

One immediate consequence of (32) is that the maxima of the diffuse scattering will occur at the points where $V(\mathbf{k})$ has minima. Hence for various assumptions of the nature of the interatomic energy functions and so the relative values for the V_{ij} , the positions of the diffuse scattering maxima may be predicted and hence the type of ordered structure which will tend to form in the alloy may be deduced. This aspect of the situation has been explored by Clapp and Moss [1968] who found interesting correlations with the ordered structures of real alloy systems.

Determinations of V_{ij} values from observed diffraction intensities have been used, in particular, to investigate the contributions of conduction electrons to the configuration energy of alloys. It has been shown that minima of $V(\mathbf{k})$ and hence maxima of $\alpha(\mathbf{k})$ may occur for \mathbf{k} values corresponding to \mathbf{k} vectors between flat areas of the Fermi surface for the alloy. Hence the form of the Fermi surface may strongly influence the form of the diffuse scattering and so the type of superlattice which tends to be formed. The relationship with Fermi surfaces is discussed by Cowley and Wilkins [1972] and a more general discussion including an account of the formation of long-period (10 to 40 Å) out-of phase domain superlattices in relation to long-range oscillatory potentials and speculations on the situation for non-metals is given by Cowley [1971].

17.7. Dynamical scattering from disordered crystals

17.7.1. Dynamical effects in diffuse scattering

In passing through the crystal both the sharp (fundamental or superlattice) reflections and the diffuse scattering are subject to further scattering. For the sharp reflections this is coherent dynamical scattering of the usual sort but involving an absorption coefficient because energy is lost from the sharp reflections to the diffuse background.

For the diffuse scattering the strongest dynamical effects are dynamical interactions of strength $F_{\mathbf{h}}$ between diffuse amplitudes separated by vectors \mathbf{h} , where \mathbf{h} and $F_{\mathbf{h}}$ refer to the fundamental reflections and their reciprocal lattice points. Such interactions should give rise to Kossel lines or related effects in X-ray diffraction patterns but

because the crystals used commonly are imperfect with a relatively large mosaic spread, these lines have been observed only very weakly and the dynamical effects are usually ignored.

For electron diffraction, however, dynamical diffraction effects are inevitably strong and can not be ignored. Since electron diffraction patterns are being increasingly used in studies of disordered alloys and superlattice formation because of the relative ease by which observations can be made, it is important to gain at least some approximate indication of the extent to which dynamical effects might modify the configurations and relative intensities of the kinematical diffuse scattering. This has been done by Fisher [1965] and Cowley and Murray [1968].

17.7.2. Calculations of diffuse scattering

Fisher made detailed calculations of diffuse scattering intensities for thin crystals of disordered copper-gold alloys using a method formulated by Gjønnes [1962] and developed into a general n -beam treatment by Cowley and Pogany [1968]. Considering first-order diffuse scattering only, the total diffuse scattering is taken as the sum of the diffuse intensities produced by each thin slice of the crystal separately. The range of correlation of atomic positions may be considered to be small so that diffuse scattering from separate slices is incoherent and intensities, not amplitudes, are added.

For a slice of thickness Δz at a depth z in the crystal of thickness H , as in Fig. 12.3, it is considered that an incident beam is first diffracted by the average lattice in the region 0 to z giving a set of fundamental beam amplitudes $\Psi_z(h)$. Each one of these beams is then scattered in the region of thickness giving fundamental reflections plus diffuse scattering. The diffuse scattering is given by the planar section, $w = 0$, of the distribution $I_d(u)$ of diffuse scattering power given as in (19). Then in the final section of the crystal, from z to H , all parts of the diffuse scattering undergo dynamical interaction through the fundamental Bragg reflection. A diffuse scattering beam in the direction u', v' for example interacts with all beams $h + u', k + v'$ where h, k are reflection indices for the average lattice. This gives the diffuse scattering intensity from the slice at position z and this is integrated over z from 0 to H to give the total diffuse scattering.

The results of these calculations for two-dimensional diffraction patterns are of considerable interest. It appears that, in general, the ratio of dynamical to kinematical intensities for pure s.r.o. diffuse scattering may be represented by a smoothly varying function increasing with

distance from the origin. Hence it is to be expected that the positions and shapes of the diffuse scattering maxima are not affected by dynamical effects (unless for a near-perfect crystal, a strong Kikuchi line is generated, passing through a maximum) but there is an over-all modification of relative intensities from one region of the pattern to the next. The more complete treatment by Fields and Cowley [1978] leads to the same conclusion.

On the other hand the calculations showed that if the size-effect displacement of s.r.o. peaks is included in the scattering from the individual slices, this displacement may be eliminated almost completely when strong two-dimensional dynamical scattering takes place. This result is in accord with the speculation by Cowley [1965], based on crude arguments, that strong dynamical scattering could eliminate contributions to diffuse intensities due to the displacements of atoms, but would not affect contributions due to the interchange, or variation of scattering power, of atoms. It is also in agreement with experimental observations. For example Fig. 17.3 shows the intensity distribution along the $h00$ line in reciprocal space for disordered CuAu_3 obtained by X-ray diffraction (Batterman [1957]) and by electron diffraction (Watanabe and Fisher [1965]). The size-effect displacement of the diffuse s.r.o. peak is considerable for X-ray diffraction but absent for the electron diffraction case. However if electron diffraction observations are made on a crystal tilted in such a way as to avoid strong dynamical interactions for particular diffuse peaks, then the size effect displacement of these peaks is visible.

17.7.3. Strong scattering and multi-atom correlations

One defect of these considerations which could, in principle, be serious is the approximation made that the scattering from a slice of the crystal should be kinematical even when the slice thickness is sufficiently large to allow the assumption that the correlation of atom positions does not extend from one slice to the next. For a heavy-atom alloy such as those of Cu and Au, even the difference term ($f_A - f_B$) is equivalent to scattering by a medium weight atom for which the kinematical approximation may be expected to fail for a few atoms thickness.

A re-formulation of the problem in terms of a phase-grating approximation for a single slice has been made by Cowley and Murray [1968]. When the potential distribution in a slice is projected, the maxima of projected potential vary with the numbers of atoms of either kind in the rows of atoms in the incident beam direction. Putting these maxima into the complex exponential of the phase grating transmission function, the scattering is no longer a linear function of the numbers and kinds of atoms. The scattered amplitudes depend on the probabilities of occurrence of, say, lines of

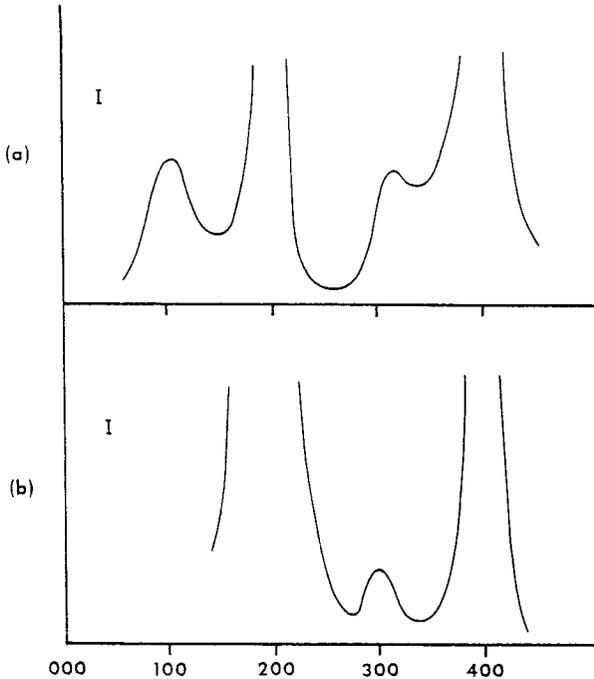


Fig. 17.3. Diffuse scattering intensity measured along the $h00$ line in reciprocal space for a disordered crystal of CuAu_3 . (a) X-ray measurement (Batterman [1957]) showing size-effect displacement of the 300 diffuse scattering peaks. (b) Electron diffraction measurement (Watanabe and Fisher [1965]) showing no displacement of the 300 peak.

three or four gold atoms. The sharp fundamental reflections are modified by a pseudo-temperature factor (see Chapter 12) which, like the diffuse scattering intensity, depends on the values of particular many-atom correlation parameters.

Preliminary calculations indicate however that, apart from an over-all decrease of diffuse scattering intensities, the influence on the diffraction patterns is not great. For the possible range of values of the many-atom correlation parameters the differences in diffuse scattering intensities are scarcely measurable with the present experimental systems, especially when complicated by dynamical scattering effects. Pending the improvement of measuring techniques and more detailed calculations for representative cases, it appears that the simple kinematical approximation for slice scattering may be trusted to give qualitatively correct results.

17.7.4. High resolution imaging of disordered crystals

From the general considerations of Chapter 13 it may be concluded that, if a thin crystal of a disordered alloy is

viewed in an axial orientation in an electron microscope having sufficient resolution, the image shows the columns of atoms parallel to the incident beam as distinct dark spots having intensities depending on the number and proportion of the atomic species in the columns. If an optical diffraction pattern is obtained from such an image, or if the image is digitized and the Fourier transform is obtained by use of a computer, a distribution of diffuse scattering is obtained which resembles the distribution of scattering power in the corresponding section of reciprocal space such as would be seen under kinematical scattering conditions (e.g., Chevalier and Hytch [1993]). The relationship of image intensity to the number and type of atoms present in a column, however, is expected to be non-linear and although the type of local ordering may often be deduced, it can not be expected that quantitative information on short-range order coefficients can be derived in this way.

Calculations of image intensities for simple models of disordered binary alloy crystals have been made using many-beam dynamical calculations (Buxton et al. [1978], Van Dyck et al. [1989], Tanaka and Cowley [1987]). It was confirmed that the intensities for the image spots representing rows of atoms do depend on the numbers of atoms in the rows of the two kinds, although in a non-linear fashion. There is some dependence on the ordering of the atoms within the rows but this dependence is weak and the possibility of using this dependence to characterize the local ordering in three dimensions appears remote.

17.8. Out-of-phase domains

17.8.1. Ordered out-of-phase superlattices

In many binary-alloy and other systems, there are, in addition to the simple ordered structures we have been considering, other superlattices having much larger unit cells, generated by the periodic occurrence of shift faults in the ordered structure. For the Cu_3Au structure, Fig. 17.1(b), for example, the Au atom could equally well be located preferentially on any one of the four equivalent sites of the F.C.C. unit cell. Thus there are four "variants" of the structure related by vector shifts having coordinates $(1/2, 1/2, 0)$, $(0, 1/2, 1/2)$ etc. A superlattice may be formed by a periodic alternation of two (or more) of these variants in one, two or three dimensions. (See Sato and Toth [1963]; Cowley [1971].)

The best known and most thoroughly investigated of these structures is the CuAu II structure, formed by alternation in one dimension of five unit cells of each of the two variants of the CuAu I structure illustrated in Fig. 17.1(c). One variant is as shown, with a plane of Cu atoms through the origin of the unit cell. In the other variant the Cu and Au

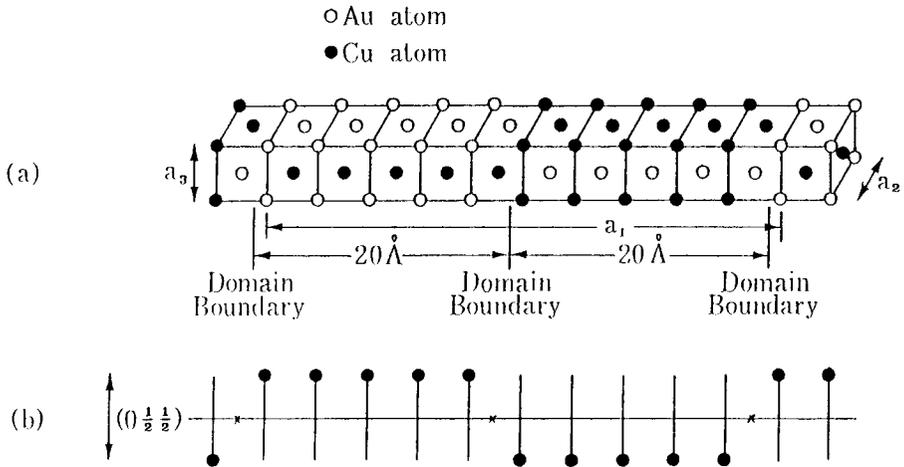


Fig. 17.4. (a) Diagram of the periodic out-of-phase domain structure of the Cu-Au-II superlattice. (b) The distribution function of one point per unit cell used in the derivation of equation (17.35).

atoms are interchanged. The resultant superlattice is orthorhombic with $a = 3.96$, $b = M \times 3.97$, $c = 3.68 \text{ \AA}$, as illustrated by Fig. 17.4(a). Here M is the number of unit cells in half the long-period repeat unit, taken as $M = 5.0$. As a first approximation the structure can be described as the ordered CuAu I unit cell contents convoluted by a distribution function which has one point per unit cell and a shift of $(1/2, 0, 1/2)$ every five unit cells, as suggested in Fig. 17.4(b).

The Fourier transform of this distribution function can be written

$$\sum_k e^{2\pi i 110k} [1 + 2\cos 2\pi k + 2\cos 2\pi 2k] 2\cos \pi(U/2 + 5k) \quad (35)$$

where $U = h + l$ and h, k, l refer to the CuAu I unit cell. The first term gives sharp peaks at intervals of $k/10$. For $h + l$ even, there are maxima only for k integral. For $h + l$ odd there are maxima for k an odd multiple of $1/10$. Hence each non-fundamental reciprocal lattice point of the CuAu I structure is replaced by two strong maxima, separated by $1/5$ of b^* plus weaker subsidiary maxima. Since the axis of the out-of-phase superlattice may occur in any of the cube axis directions, the total diffraction pattern produced is the sum of a number of equivalent distributions and the electron diffraction pattern for a beam in a (100) direction is as illustrated in Fig. 17.5, with a characteristic group of spots replacing each CuAu I superlattice spot.

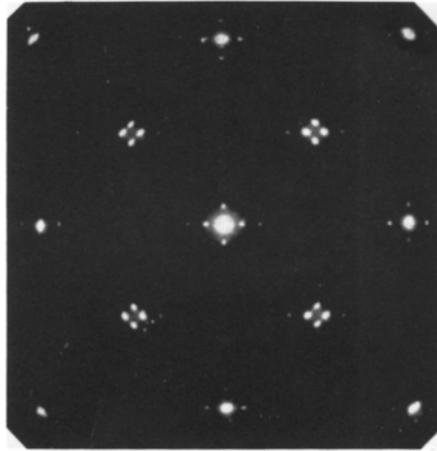


Fig. 17.5. Electron diffraction pattern from CuAu II in [100] orientation showing the characteristic groups of superlattice spots (after Glossop and Pashley [1959]).

The periodicities of the out-of-phase superlattices for alloys have been attributed to the occurrence of long-range oscillatory pseudo-potentials, associated with the restriction of the conduction electron energies and momenta at the Fermi surface (Sato and Toth [1963]; Tachiki and Teramoto [1966]). It has been demonstrated very clearly (Hashimoto and Ogawa [1970]) that, if the electron-atom ratio in an alloy is varied by the addition of substitutional impurity atoms of different valency, the periodicity of the long-period superlattice varies in such a way that the energy of the electrons at the Fermi surface is minimized.

17.8.2. Out-of-phase domains in disordered alloys

When an alloy is heated so that long-range ordering disappears, it is to be expected that the long-range oscillatory potential describing the electron-energy terms remains virtually unchanged. Hence there is a tendency in the short-range ordered state for a correlation of atom positions, related to this long-range oscillatory potential through the relationship (17.32), to persist.

In fact minima of $V(k)$ and maxima of the diffuse intensity, proportional to $\alpha(k)$, do exist around the superlattice reciprocal lattice positions. The diffuse maxima given by many alloys having only short-range order do appear to be split in much the same way as the sharp maxima for ordered alloys.

This splitting was first observed by Raether [1952] using electron diffraction, investigated by Watanabe and

Fisher [1965] and finally observed using X-ray diffraction by Moss [1965] when the X-ray experimental methods were refined to provide sufficient resolution.

Since the presence of this splitting in the diffuse diffraction maxima implies that correlations between atom site occupancies exist over distances of, perhaps, 20 to 40 Å, it is inconvenient to describe the state of short-range order in terms of correlation parameters. The number of order coefficients required would be very large. As an alternative it has become a common practise to describe the state of short-range order in terms of an assembly of microdomains, separated by out-of-phase boundaries and similar in many cases, but not always, to the type of anti-phase domains which are present in the alloys having long-range order. Computer simulation methods (Gehlen and Cohen [1965]) have allowed the correlation between micro-domain structures and short-range order coefficients to be investigated and visualized in a very illuminating manner.

This model for short-range order is extremely useful but has, perhaps, been interpreted rather too literally by investigators who attempt to image the microdomains in short range ordered alloys by obtaining dark-field electron micrographs, using only the diffuse maxima of the diffraction patterns to obtain the image. Some of the difficulties and pitfalls of this technique have been pointed out by Cowley [1973].

17.8.3. Modulated structures

The results that we have described for the relatively simple binary alloy systems in the previous two Subsections are paralleled by observations made recently on many types of materials. Accounts of many experimental observations and theoretical descriptions of ordered or disordered superstructure formation in which the superstructure periodicity may or may not be commensurate with the sublattice unit cell, have been given in the report of the conference on Modulated Structures (Cowley et al. [1979]). The examples are drawn from the fields of solid-state structural chemistry, mineralogy, metallurgy and solid-state physics and the theoretical treatments are correspondingly diverse.

Further reviews of modulated and intergrowths structures of minerals have been given by Buseck and Cowley [1983] and Buseck and Veblen [1988]. The diffraction effects associated with modulated structures and examples for several classes of compounds have been given by Amelinckx and Van Dyck [1993].

In the case of the out-of-phase domain superstructures of Subsection 17.8.1, the long period repeat distance is usually not an integral multiple of the basic unit cell size: i.e. M is not an integer. Correspondingly the superstructure reflections in patterns such as Fig. 17.5 do not form a regular subdivision of the sublattice spot separations.

It is not to be concluded that the domain boundaries such as are illustrated in Fig. 17.4 occur regularly at intervals of, say, 4.7 rather than 5.0 unit cells. Rather, it seems that for $M = 4.7$ the domain boundaries maintain the same form and occur at intervals of either 4 or 5 unit cells, with a random distribution of the 4 and 5 cell spacings occurring with relative frequencies such that the average spacing is 4.7 unit cells. It has been shown by Fujiwara [1957] that this arrangement can give sharp spots at intervals $u = 1/4.7a$ in reciprocal space, plus some weak diffuse scattering.

Thus the structure may be regarded as a locally disordered sequence of commensurate superstructure units with a statistically long-range ordered, incommensurate superstructure. The long-range order with the $M = 4.7$ periodicity is presumably induced by a corresponding periodicity in the long-range oscillatory pseudo-potential mentioned in Subsection 17.8.1.

There are many similar cases among the nonstoichiometric oxides and minerals in which the formation of a superstructure involves a distortion of the unit cell or a displacement or replacement of atoms which can occur only at definite sites within the unit cell so that a long period incommensurate structure is made up of locally commensurate units. On the other hand, many examples can be found for which there is apparently a progressive distortion of atom groups within a structure, forming an incommensurate long range periodicity which appears to be quite independent of the subcell periodicity. These include materials such as K_2SeO_4 in which it is shown by neutron diffraction (Iizumi et al. [1977]) that the incommensurate superstructure appears as the limiting, zero frequency case of a soft-mode optical phonon which can be related to the elastic properties of the material. Also a wide range of both organic and inorganic materials show incommensurate and commensurate long range periodicities which are related to their striking one-dimensional or two-dimensional electrical conduction properties. In some cases incommensurate-to-commensurate (lock-in) transitions occur with changes of temperature. The concept of charge-density waves, with associated lattice distortions or atomic ordering, has been invoked to account for many observations of diffuse maxima or sharp spots seen clearly in electron diffraction patterns or investigated by X-ray diffraction.

Changes of sign or relative magnitudes of the short-range order parameters can lead to a segregation of the individual types of atom rather than an ordering towards superlattice formation. The operation of a long-range oscillatory pseudo-potential in these cases can give rise to metastable periodic variations of composition as in the spinodal decomposition of alloys. The additional closely

spaced sharp or diffuse spots in diffraction patterns then appear around the strong fundamental lattice reflections and so are often difficult to detect. Similar phenomena, often with periodic changes of unit cell dimensions and symmetry, occur in minerals such as the feldspars giving rise to superstructure spots in diffraction patterns and a mottling of the contrast in electron micrographs.

Problems

1. Find the limiting values of the order parameters α_{0i} and $\langle \sigma_0 \sigma_i \rangle$ for the perfectly ordered lattices of Cu_3Au and CuAu . How do the values of these order parameters vary with temperature and with the vector length, $|r_{0i}|$?

2. Given that the energy term for nearest neighbors is positive and the ratios $V_2/V_1 = -0.5$, $V_3/V_1 = 0.2$, where V_2 and V_3 refer to second- and third-nearest neighbor atomic pairs, find the minima in the values of $V(k)$ and hence the positions of diffuse scattering maxima for the Cu_3Au structure and the type of ordered lattice which will tend to form. Do this also for $V_2/V_1 = +0.75$, $V_3/V_1 = +0.2$.

Extended defects

18.1. Introduction

A vast literature exists concerning the extended defects which occur in most types of crystals. They are of fundamental importance for the consideration of the physical and chemical properties of solids and of great technological significance. The simplest and best-known are the stacking faults, twins and the various forms of dislocation. To these one can add defects clusters, impurity aggregates, segregated concentrations of particular atoms as in G-P zones, coherent and incoherent precipitates, vacancy clusters, voids, ordering nuclei, and so on.

Here we plan to do little more than illustrate the methods by which such defects can be studied by diffraction methods, and will confine our considerations to stacking faults and dislocations.

There have been two main avenues of approach to the study of these faults. Historically the first, and still important, is the observation of the streaking or diffuse scattering in diffraction patterns. The classical example is that in close-packed structures the sequence of stacking of the hexagonal close-packed planes of atoms may not follow the regular two-plane periodicity of the hexagonal close-packed structure or the regular three-plane periodicity of the face-centered cubic structure but may show faults in either type of sequence, or there may even be an almost complete randomness in switching from one type of sequence to another. The effect in reciprocal space is to produce continuous lines of scattering power, perpendicular to the close-packed planes and passing through some reciprocal lattice points. The initial analysis was on stacking faults in hexagonal cobalt (Wilson [1942]) but other examples followed and analogous effects were found to exist for a very wide range of metallic and non-metallic materials (Guinier [1963]; Warren [1969]).

The analysis of the nature of the faults and their distributions from the diffraction patterns, assuming kinematical scattering, is necessarily in terms of statistical averages over very large numbers of faults. Starting from the diffraction intensities one can derive, and then attempt to interpret, the generalized Patterson function (Chapter 7). The means for doing this have included the construction of a "probability tree" (see Warren [1969]) or

the more formally elegant matrix methods developed for example by Hendricks and Teller [1942] and Kakinoki and Komura [1951, 1952] and somewhat different approaches by Jagodzinski [1949] and by Michalski [1988]. We have chosen to present here a slightly different method developed in terms of the probabilities of occurrence of various types of fault in the structure.

With the extensive development of electron microscopy in the 1950s, the study of individual defects in thin crystalline films became possible. Fault planes gave the appearance of bands of fringes. Dislocations appeared as dark or dark-and-light lines. The basis for interpretation of the observations was essentially the dynamical theory of electron diffraction and an extensive body of experience on both the configurations of the defects and the rules for the interpretation of the images was rapidly assembled (see Hirsch et al. [1965]). Equivalent observations of defects in near-perfect crystals by X-ray diffraction under dynamical scattering conditions followed a few years later (Lang [1958, 1959]; Kato and Lang [1959]), and the appropriate X-ray diffraction theory was developed on the initial work of Kato [1960, 1961]. More recently, more exact treatments in terms of n -beam dynamical theories have been developed for electron diffraction and for all radiations the difficult task has been tackled of providing an adequate dynamical theory for imperfect crystals (e.g. Kato [1973]; Kuriyama [1973]). We will follow these developments in outline only.

18.2. Stacking faults - statistical, kinematical theory

18.2.1. Patterson method for a simple case

There are many substances for which the ideal crystal structure can be considered to be built up by the regular superposition of identical layers, each layer being one unit cell thick. In practise the regular superposition is occasionally interrupted by a fault which gives a displacement of one layer, and all subsequent layers, relative to the previous layers. If these faults are not too numerous, we may assume as a first approximation that they occur at random. We suppose that there is a probability, α , that a fault occurs in which the displacement is defined by the vector s .

If the vector s is not parallel to the plane of the layers, there must in general be a subtraction or addition of atoms to a layer at the fault to maintain approximately the same density of material. We defer consideration of this point until later.

Since the structure can be described in terms of an electron density distribution

$$\rho(r) = \rho_0(r) * d(r),$$

where $\rho_0(r)$ is the electron density of one layer and $d(r)$ is some distribution function, the generalized Patterson function can be written, as in (7.12) as

$$P(r) = \rho_0(r) * \rho_0(-r) * D(r),$$

where $D(r)$ is a distribution function Patterson describing the probability that if one layer is centered at the origin, another layer will be centered at r . After n layers the probability that a number, m , of faults should have occurred is given by the Poisson distribution function as follows:

$$\text{Probability of 0 faults} = \exp\{-\alpha n\},$$

$$\text{Probability of 1 fault} = \alpha n \exp\{-\alpha n\},$$

$$\text{Probability of } m \text{ faults} = \frac{(\alpha n)^m}{m!} \exp\{-\alpha n\}.$$

The vectors between the origins of two layers separated by n normal translations plus m faults will be $na + ms$. The generalized Patterson function is then

$$P(r) = N\rho_0(r) * \rho_0(-r) * \left[\sum_{n=1}^{\infty} \sum_{m=0}^{\infty} \delta(r-na-ms) \frac{(\alpha n)^m}{m!} \exp\{-\alpha n\} + \sum_{n=1}^{\infty} \sum_{m=0}^{\infty} \delta(r+na+ms) \frac{(\alpha n)^m}{m!} \exp\{-\alpha n\} + \delta(r) \right]. \quad (1)$$

Fourier transforming gives the intensity distribution

$$I(u) = |F(u)|^2 \left[\sum_{n=1}^{\infty} \sum_{m=0}^{\infty} \exp\{2\pi i u \cdot an\} \exp\{2\pi i mu \cdot s\} \frac{(\alpha n)^m}{m!} \exp\{-\alpha n\} + \sum_{n=1}^{\infty} \sum_{m=0}^{\infty} \exp\{-2\pi i u \cdot an\} \exp\{-2\pi i mu \cdot s\} \frac{(\alpha n)^m}{m!} \exp\{-\alpha n\} + 1 \right]. \quad (2)$$

The summation over m gives $\exp\{n\alpha \exp[2\pi i u \cdot s]\}$ and the summation over n is made using the relation $\sum_{n=0}^{\infty} x^n = (1-x)^{-1}$ so that

$$\begin{aligned}
 I(u) &= |F(u)|^2 [(1 - \exp\{2\pi i u \cdot a - \alpha + \alpha \exp(2\pi i u \cdot s)\})^{-1} \\
 &+ (1 - \exp\{-2\pi i u \cdot a - \alpha + \alpha \exp(-2\pi i u \cdot s)\})^{-1} - 1] = |F(u)|^2 \\
 &\times \left[\frac{1 - \exp[2\alpha(\cos 2\pi u \cdot s) - 1]}{1 + \exp[2\alpha(\cos 2\pi u \cdot s) - 1] - 2 \exp\{\alpha(\cos 2\pi u \cdot s - 1) \cos(2\pi u \cdot a + \alpha \sin 2\pi u \cdot s)\}} \right] \quad (3)
 \end{aligned}$$

If α is small it can be seen readily that this function has fairly sharp maxima which are displaced from the reciprocal lattice points of the unfaulted crystal if s is not parallel to the layers. The positions of the maxima are given by

$$u \cdot a - h/a = -(2\pi)^{-1} \alpha \sin 2\pi u \cdot s. \quad (4)$$

These maxima are of height $[\alpha(\cos 2\pi u \cdot s - 1)]^{-1} |F(u)|^2$ and of half width $(\alpha/\pi)(\cos 2\pi u \cdot s - 1)$.

As an example we may quote the case of magnesium fluorogermanate (Bless et al. [1972]) in which the structure can be considered as made up of four layers of metal atoms per unit cell within the close-packed oxygen-fluorine structure. The layers are perpendicular to the c -axis. The presence of fluorine is associated with faults where one of the four layers is omitted. The component of the s vector in the c -axis direction is then $-c/4$. For the $00l$ spots, the 001 spot is shifted by an amount $+\alpha/2\pi$ away from the origin and has a half width α/π , the 002 spot is unshifted but has a half width $2\alpha/\pi$, the 003 spot is shifted $-\alpha/2\pi$ towards the origin and has width α/π while the 004 spot is sharp and unshifted. This can be seen to be the case in the electron diffraction pattern, Fig. 18.1, from which it can be deduced readily that $\alpha = 0.2$.

In all cases such as this when s is not parallel to the layers, the presence of a fault implies the addition or subtraction of part of a layer. In order to treat such cases adequately it is necessary to use a somewhat different, more general approach (Cowley [1976a]) although, as we will see, it is often possible to reduce the more general result to a relatively simple form in many cases of practical significance, especially if the experimental data is not strictly quantitative.

18.2.2. A general treatment

We consider the crystal to be made up of an arbitrary number of different types of layers. The i th type of layer has an electron density distribution $\rho_i(\mathbf{r})$ and if such layers were stacked regularly the translation vector would be \mathbf{R}_i . If

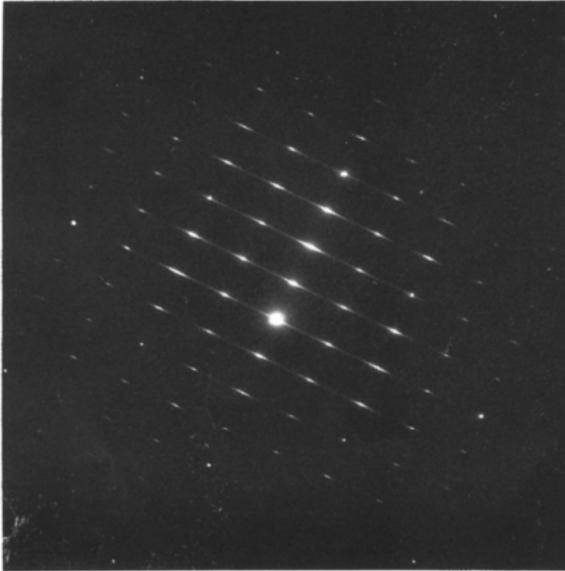


Fig. 18.1. Electron diffraction pattern from a crystal of magnesium fluoro-germanate showing streaking of spots along the c^* -direction due to planar faults in the crystal (courtesy of P. Kunzmann).

a fault occurs to change the layer from $\rho_i(\mathbf{r})$ to $\rho_j(\mathbf{r})$ which has a translation vector \mathbf{R}_j there is the addition of an electron density $\Delta_{ij}(\mathbf{r})$ and the vector \mathbf{R}_i is modified to $\mathbf{R}_i + \mathbf{S}_{ij}$. The probability of such a fault is α_{ij} .

The assumption that such faults occur at random with this probability is not as restrictive as it may at first appear. If particular sequences of planes occur commonly, each of these sequences may be regarded as a separate type of layer. Hence, the preference for clumps of layers of particular types may be included in the description.

The probability, g_i , that the i th type of layer should occur at any particular position is given by equating the number of transitions to and from the i -type:

$$\sum_j g_j \alpha_{ji} = g_i \sum_j \alpha_{ij}. \quad (5)$$

The generalized Patterson function can then be written as a series of terms corresponding to 0, 1, 2... interlayer vectors \mathbf{R}_n . For the interlayer vector of zero length,

$$P_0(r) = N \sum_i g_i [(1-A_i) \{\rho_i(r) * \rho_i(-r)\} + \sum_j \alpha_{ij} (\rho_i(r) + \Delta_{ij}(r)) * (\rho_i(-r) + \Delta_{ij}(-r))] \quad (6)$$

Here we have put $A_i = \sum_j \alpha_{ij}$ so that $(1 - A_i)$ is the probability that an i layer will not be modified by a fault. The second term in the square bracket comes from layers modified by faults.

The contribution from (6) to the observed intensity is given by Fourier transforming as

$$I_0/N = \sum_i g_i (1-A_i) |F_i|^2 + \sum_j \alpha_{ij} |F_i + G_{ij}|^2 \quad (7)$$

where F_i and G_{ij} are the Fourier transforms of ρ_i and Δ_{ij} .

The contribution to the intensity due to terms with a single interlayer vector R_i to which there may be added S_{ij} if a fault occurs, is given by the same sort of reasoning as

$$I_1/N = \sum_i g_i [(1-A_i) F_i^* \{F_i + \sum_j \alpha_{ij} G_{ij}\}] \exp\{2\pi i u R_i\} + \sum_i g_i \sum_j \alpha_{ij} [(F_i^* + G_{ij}^*) \cdot \{F_j + \sum_k \alpha_{jk} G_{jk}\}] \exp\{2\pi i u (R_i + S_{ij})\}, \quad (8)$$

which represents the sum of terms for which there is no fault plus the sum of terms for which a fault occurs between neighboring layers.

To simplify this and subsequent expressions we define $F'_i \equiv F_i + \sum_j \alpha_{ij} G_{ij}$ which represents the average structure amplitude for a layer, as modified by the possibility of faults.

For subsequent terms we can simplify the expressions further by writing

$$B_i \equiv (1-A_i) F'_i + \sum_j \alpha_{ij} F'_j \exp\{2\pi i u \cdot S_{ij}\}$$

which represents the average neighbor of a given layer, including the possibility of a fault or no faults. For vectors between second nearest neighbors, the contribution to the intensity is then

$$I_2/N = \sum_i g_i F_i^* B_i \exp\{2\pi i u \cdot 2R_i\} + \sum_i g_i \sum_j \alpha_{ij} (F_i^* + G_{ij}^*) B_j \exp\{2\pi i u \cdot (R_i + R_j + S_{ij})\}. \quad (9)$$

If we take, from each of these terms, the contribution from the cases where no fault occurs we obtain the series

$$1 + (1-A_i) \exp\{2\pi i u R_i\} + (1-A_i)^2 \exp\{2\pi i u \cdot 2R_i\} + \dots$$

which sums to give

$$[1 - (1-A_i) \exp\{2\pi i u \cdot R_i\}]^{-1}.$$

For vectors $-R_i$ we obtain terms $I_{-n} = I_n^*$. Then summing for all interplanar vectors we obtain the general expression:

$$\begin{aligned} I/N = & \sum_i g_i (1-A_i) |F_i|^2 + \sum_i g_i \sum_j \alpha_{ij} |F_i + G_{ij}|^2 \\ & + \sum_i g_i (1-A_i) F_i^* \exp\{2\pi i u \cdot R_i\} \\ & \times \left[F_i' + \frac{\exp\{2\pi i u \cdot R_i\}}{1 - (1-A_i) \exp\{2\pi i u \cdot R_i\}} \right. \\ & \times \left[B_i + \sum_j \alpha_{ij} \frac{\exp\{2\pi i u \cdot (R_j + S_{ij})\}}{1 - (1-A_j) \exp\{2\pi i u \cdot R_j\}} [B_j \right. \\ & \left. + \sum_k \alpha_{jk} \frac{\exp\{2\pi i u \cdot (R_k + S_{jk})\}}{1 - (1-A_k) \exp\{2\pi i u \cdot R_k\}} [B_k + \dots]] \right] \\ & \left. + \sum_i g_i \sum_j \alpha_{ij} (F_i^* + G_{ij}^*) \exp\{2\pi i u \cdot (R_i + S_{ij})\} \right. \\ & \left. \times \left[F_j' + \frac{\exp\{2\pi i u R_j\}}{1 - (1-A_j) \exp\{2\pi i u \cdot R_j\}} [B_j + \dots] \right] + \text{c.c.} \right] \quad (10) \end{aligned}$$

The first two terms represent intralayer contributions. The next two terms represent the contributions from interlayer terms starting on nonfault layers and layers with faults respectively for positive vectors. The initial parts of these two terms are different but from the bracket starting with B_j they are identical. The complex conjugate gives the terms for negative vectors.

If the probabilities for faults are low only a few of the brackets need be considered. From the nature of the denominators it is seen that maxima of intensity tend to occur whenever $u \cdot R_i$ is close to an integer, i.e. around the reciprocal lattice points for an ordered stacking of each of the individual types of layers.

The general expression (10) can be used to treat cases of considerable complexity which have scarcely been attempted

in the past. It is particularly useful when fault probabilities are low so that the nested series of terms can be terminated relatively quickly or when assumptions are possible which allow analytical summations to be made but the use of this formulation may be tedious for general cases with large fault probabilities. It was pointed out by Treacy et al. [1991] that the nested series of equation (10) may be described in terms of recursion relationships and expressed by the operation of a probability matrix. Then the intensity expression may be written as a matrix equation which may be used conveniently for a wide range of problems. A FORTRAN program, DIFFAX, has been written to allow the convenient application for the calculation of X-ray or neutron powder patterns or electron diffraction single-crystal patterns for such complicated faulted systems as those found in zeolites (Treacy et al. [1991, 1993]).

A number of familiar cases of stacking fault diffraction effects, however, may be treated by simplifying equation (10) so that algebraic expressions may be derived from which the intensity distributions are readily deduced. A particularly simple case is that when there is only one type of layer and a probability α that a fault occurs with a displacement vector S but no modification of the layer structure at the fault. Then (10) becomes

$$I/N = \frac{F^*}{1 - (1 - \alpha) \exp(2\pi i u \cdot R)} \left[F + \frac{\alpha \exp(2\pi i u \cdot (R + S))}{1 - (1 - \alpha) \exp(2\pi i u \cdot R)} \right. \\ \left. \cdot \left[F + \frac{\alpha \exp(2\pi i u \cdot (R + S))}{1 - (1 - \alpha) \exp(2\pi i u \cdot R)} [F + \dots] \right] \right] + \text{c.c.} - |F|^2.$$

The expression in the square brackets may be summed as a geometric series to give

$$I/N = \frac{|F|^2 \alpha (1 - \alpha) (1 - \cos 2\pi u \cdot S)}{1 - \alpha + \alpha^2 - (1 - \alpha) \cos 2\pi u \cdot R - \alpha \cos 2\pi u \cdot (R + S) + \alpha (1 - \alpha) \cos 2\pi u \cdot S}. \quad (11)$$

This expression appears different from (3) but gives an intensity distribution which is almost identical for small α . The intensity maxima are close to the reciprocal lattice points for which $u \cdot R$ is an integer, unless $u \cdot S$ is close to an integer. The maxima are actually displaced from the reciprocal lattice points by an amount ϵ given by

$$\epsilon = -\alpha \sin(2\pi u \cdot S) / \{1 - \alpha(1 - \cos 2\pi u \cdot S)\}$$

and the maximum intensity is

$$I/N \approx 2|F|^2 (1 + \alpha) / [\alpha(1 - \cos 2\pi u \cdot S)].$$

For the case of the magnesium fluoro-germanate discussed above and illustrated by Fig. 18.1, roughly one quarter of the layer content is omitted when there is a fault so that it is necessary to consider the slightly greater complication that $C_{ij} \neq 0$. This modifies the treatment only slightly. The result is an expression for which the denominator is exactly the same as for (11) but the numerator becomes

$$\begin{aligned} & \alpha(1-\alpha) [|F|^2 (1-\cos 2\pi u \cdot S) \\ & + (\operatorname{Re} F * G) \{1 + \cos 2\pi u \cdot (R + S) - \cos 2\pi u \cdot R - \cos 2\pi u \cdot S\} \\ & - (\operatorname{Im} F * G) \{ \sin 2\pi u \cdot R + (1-2\alpha) \sin 2\pi u \cdot S \\ & - \sin 2\pi u \cdot (R + S) \} \\ & + |G|^2 (1 - \cos 2\pi u \cdot R)], \end{aligned} \quad (12)$$

Here Re and Im indicate the real and imaginary parts of the product $F * G$.

In eq. (12), G appears to occur with much the same weighting as F even though the faults are much less frequent than the unfaulted layers. However it is seen that for x small the terms involving G are very small in the neighborhood of the intensity maxima which are close to the reciprocal lattice points given by integer values of $u \cdot R$. The nature of the faults then influences the intensity distribution very little except in the low-intensity regions well away from the intensity maxima (see Cowley [1976]). The assumption that $G = 0$ may often be used to simplify the calculations for α small, especially if the aim of the investigation can be satisfied by a determination of α and S from the position and width of the intensity maxima.

This treatment of stacking faults is most useful if the assumption can be made of random faults. In many cases however, faults do tend to cluster together or else to avoid each other. If the clustering is well defined these cases can often be treated by considering particular groupings of layers, possibly with particular sequences of faults, as the layers of different types in the above equations. For other cases, particularly with $G = 0$, it is possible to consider combinations of two fault probabilities α and γ giving an intensity distribution

$$I(u) = aI_{\alpha}(u) + cI_{\gamma}(u) \quad (13)$$

with $a + c = 1$ and $a\alpha + c\gamma > 0$. The probability of no fault in n layers then takes the form

$$a(1-\alpha)^n + c(1-\gamma)^n.$$

If $\gamma > \alpha$ and c is positive, the faults tend to cluster together. If c is negative with $\gamma > \alpha$, the faults will tend to avoid each other.

18.2.3. Faults in close-packed structures

The close-packing of atoms in simple structures is conventionally described in terms of the sequence of positions of the origins of hexagonal two-dimensional layers of atoms as illustrated in Fig. 18.2. Relative to an origin layer labelled A, the neighboring layers may take the positions B or C. The hexagonal close-packed (HCP) sequence is the ABABAB... or ACACAC... or equivalent, and the face-centered cubic sequence is ABCABC... or ACBACB... or equivalent. Convenient diagrams for the sequences and their faults are obtained by considering sections of the structure on (110) planes as in Fig. 18.2b; and axes a and c can be drawn within this plane. Then $a = 3^{1/2}a_0$, where a_0 is the hexagonal layer periodicity and c may be chosen equal to the

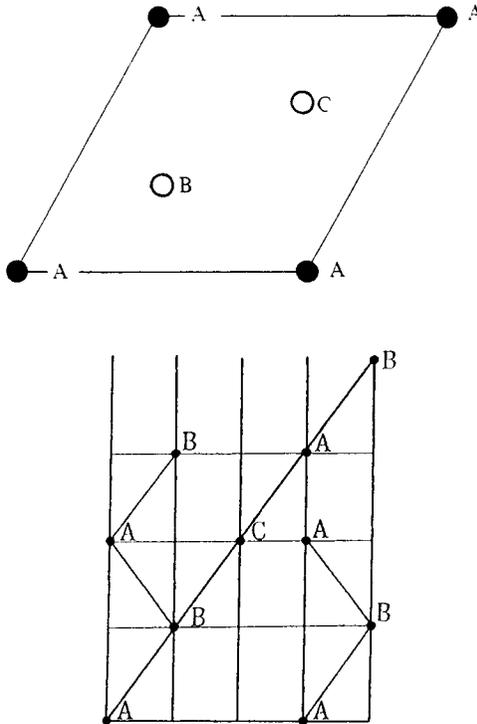


Fig. 18.2. The stacking sequences of close-packed structures. (a) The labelling of alternate stacking positions for two-dimensional close-packed layers. (b) Stacking sequences illustrated by the atom positions in (110) planes.

H.C.P. c -axis, the F.C.C. three-layer repeat distance or, for convenience, the thickness of one layer.

In order to describe the faults in HCP structures according to the scheme given above, we take a c -axis equal to the HCP c -axis so that $R = c$. Then growth faults, giving sequences such as ABABCBCBCB... can be described in terms of two types of fault occurring with equal probability, one being a fault at the $z = 0$ layer and the other a fault occurring at the $z = 1/2$ layer; the first gives a shift $s = c/2 + a/3$ and one plane of atoms is added; the second gives a shift $-s$ and one plane of atoms is subtracted.

Then, neglecting initially the added and subtracted planes,

$$I(u) = \sum_N \exp\{2\pi i u \cdot R N\} |F_0(u)|^2$$

$$\times [(1-2\alpha)^N + \alpha(\exp\{2\pi i u \cdot s\} + \exp\{-2\pi i u \cdot s\}) C_1^N (1-2\alpha)^{N-1} + \dots$$

$$+ \alpha^n (2\cos 2\pi u \cdot s)^n C_n^N (1-2\alpha)^{N-n} + \dots]. \quad (14)$$

Summing over N then gives:

$$\frac{I(u)}{|F_0(u)|^2} = \sum_{n=0}^{\infty} \left[\frac{\exp\{2\pi i u \cdot R n\} \alpha^n (2\cos 2\pi u \cdot s)^n}{[1 - (1-2\alpha)\exp\{2\pi i u \cdot R\}]^{n+1}} + \text{c.c.} \right] - 1$$

$$= \frac{1 - [1 - 2\alpha(1 - \cos 2\pi u \cdot s)]^2}{1 + [1 - 2\alpha(1 - \cos 2\pi u \cdot s)]^2 - 2[1 - 2\alpha(1 - \cos 2\pi u \cdot s)]\cos 2\pi u \cdot R}. \quad (15)$$

It is readily confirmed that this result gives the same intensity distribution as derived by other methods (Warren [1969]). The hl spots of our two dimensional unit cell are sharp for $h = 3n$. For $h \neq 3n$ there are continuous lines of intensity in the c^* direction with maxima for l integral, the maxima for l odd being 9 times the height and one third of the width of those for l even.

For the F.C.C. structure we take an oblique two dimensional unit cell within the (110) plane, one axis being the a axis and the other being the vector from an A atom in one layer to a B or C atom position in the next layer. Thus we have two "layer" types ρ_1 and ρ_2 differentiated by repetition vectors R_1 and R_2 . A growth fault in an F.C.C. structure is then one for which R_1 changes to R_2 or vice versa, with $s = 0$. This case has been treated by Cowley [1976a].

For "displacement" or "deformation" faults of the F.C.C. lattice, if the structure is characterized by a repetition vector R_1 , a fault gives a shift $s = a/3$ and for R_2 the shift is $-s = -a/3$, but the type of structure is not

changed from one variant to the other. The various other types of fault can be described similarly and treated by the appropriate simplification of the general formulation given above. For example, the case of frequent twinning which sometimes occurs for minerals such as the feldspars, has been treated in this way by Cowley and Au [1978].

18.3. Dynamical diffraction by stacking faults

The possibility of interpreting the images of individual stacking faults, appearing in electron micrographs, has provided a great expansion in the knowledge of the form and variety of stacking faults plus, more importantly, the interactions of faults with other planar faults, dislocations, grain boundaries and so on. It is this latter type of information which is almost completely inaccessible if one is limited to the diffraction studies of statistical distributions which we have discussed above.

The interpretation of fault images normally involves the use of the column approximation, Fig. 10.7. For a fault plane intersecting a thin foil sample, we may consider diffraction in the perfect crystal region above the fault, a translation of the lattice by a shift vector s at the fault, approximated by a shift perpendicular to the column, and then diffraction in the perfect crystal region following the fault. The calculation of amplitudes for the perfect crystal regions may be made by any of the available n -beam dynamical treatments and the shift modulates the structure amplitudes for the subsequent part of the crystal by a factor $\exp\{2\pi i h \cdot s\}$.

The most familiar features of fault images may be derived simply from the 2-beam dynamical theory of Chapters 8 and 10. According to the Bloch-wave formulation, a beam incident at the Bragg angle generates the two Bloch waves $\psi^{(1)}$ and $\psi^{(2)}$ for which the components of the wave vector in the direction of propagation are $k_0 \pm \xi$ where $\xi = |v_h|/2k$. The two Bloch waves propagate with refractive indices $n = 1 + (\Phi_0 \pm \Phi_h)/2E$, and absorption coefficients $\mu_0 \pm \mu_h$. After passing through a thickness z of crystal the waves are out of phase by an amount $2\xi z$ which is equal to 2π for $z = 2h^2\lambda/\pi me\Phi_h \equiv \xi_h$, the extinction distance. This progressive phase difference leads to the appearance of the sinusoidal thickness fringes for a wedge-shaped crystal. The Bloch waves may be written

$$\begin{aligned}\psi^{(1)} &= 2^{-1/2} [\exp\{2\pi i k^{(1)} \cdot r\} - \exp\{2\pi i (k^{(1)} + h) \cdot r\}], \\ \psi^{(2)} &= 2^{-1/2} [\exp\{2\pi i k^{(2)} \cdot r\} + \exp\{2\pi i (k^{(2)} + h) \cdot r\}].\end{aligned}\quad (16)$$

At the position of a stacking fault each of these Bloch waves of the initial part of the crystal acts as an incident plane

wave and each generates two Bloch waves in the second part of the crystal:

$$\psi^{(1)} \rightarrow \psi^{11} + \psi^{12},$$

$$\psi^{(2)} \rightarrow \psi^{21} + \psi^{22},$$

From the diagram, Fig. 9.1, it is clear that if $h \cdot s = 1/2$, the wave which had nodes on the atom planes now has nodes between them and vice versa. Hence the roles of the two Bloch waves are interchanged. The phase difference in a crystal of thickness H with a fault at z is then

$$(k_1 - k_2)z + (k_2 - k_1)(H - z) = (k_2 - k_1)H + 2(k_1 - k_2)z. \quad (17)$$

Thus the variation of the phase change with z is twice as great as with H . Hence fringes occur in the image as for a wedge crystal but with a spacing half as great.

For a thick crystal in which absorption is important we may use the argument of Hashimoto et al. [1960] based on Fig. 18.3. If the fault is at the bottom of the crystal, as on the left side, the wave $\psi^{(1)}$ is strongly attenuated in the top part of the crystal so that at the fault there is only the

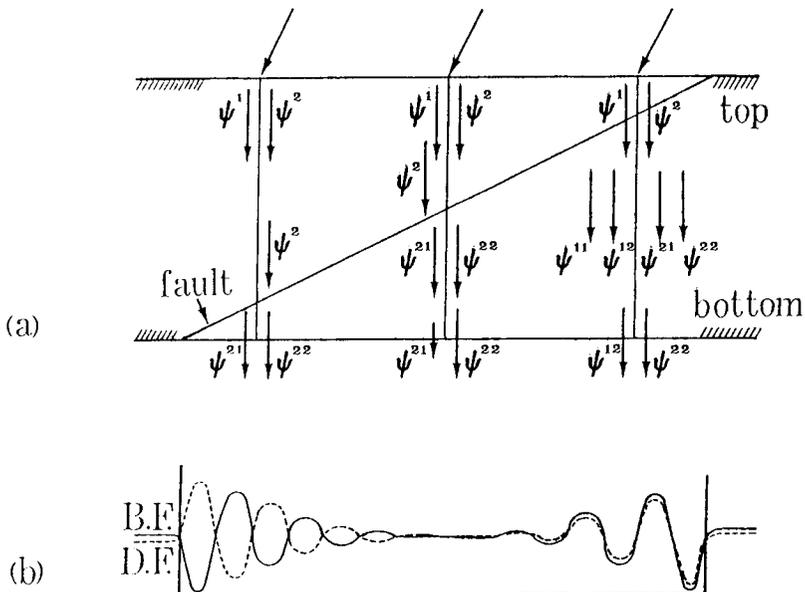


Fig. 18.3. (a) Illustration of the relative contributions from various Bloch waves to the contrast of the image of a stacking fault in a thick absorbing crystal. (b) The dark-field and bright field intensity profiles generated under the conditions of (a).

wave $\psi^{(2)}$, giving ψ^{21} and ψ^{22} which interfere with almost equal amplitude at the exit face to give strong fringes in bright field and dark field, as for a thin wedge-shaped crystal.

For the fault in the middle of the crystal, the Bloch wave $\psi^{(2)}$ is strongest at the fault and gives ψ^{21} and ψ^{22} , but in the second half of the crystal ψ^{21} is more strongly attenuated. Then the image is given by the interference of ψ^{22} with a much weaker ψ^{21} and the fringes have very low contrast.

For the fault at the top of the crystal all four waves ψ^{11} , ψ^{12} , ψ^{21} and ψ^{22} are generated at the fault but ψ^{11} and ψ^{21} are more strongly attenuated so that the image is given by interference of equally strong ψ^{12} and ψ^{22} , giving strong fringes. However, it can be seen from (16) that for this case the phase of the diffracted wave component relative to the incident wave component of the Bloch wave is opposite for ψ^{12} and ψ^{22} whereas the phases are the same for ψ^{21} and ψ^{22} which were important for the fault at the bottom of the crystal. Hence for the dark field image, the fringes are out of phase with the bright-field fringes for the fault at the bottom but in phase for the fault at the top, as suggested by Fig. 18.3(b). The top and bottom of the foil may thereby be distinguished.

For deviations from the Bragg angle the difference between $k^{(1)}$ and $k^{(2)}$ increases and the initial amplitudes of the Bloch waves cease to be equal. The fringes therefore become weaker and more closely spaced.

Deviations from the ideal two beam conditions add more Bloch waves and render the fringe pattern more complicated. Interpretation then depends on complete n -beam calculations.

18.4. Dislocations

18.4.1. Diffraction effects

Since the usual diffraction and imaging methods are not sensitive to the detailed configuration of atoms around the core of a dislocation, it is usually sufficient to assume the simple classical models of dislocation strain fields based on macroscopic elastic theory. Often the considerations are limited even further by the assumption of isotropic elastic properties for the material.

For a screw dislocation the Burgers vector \mathbf{b} is parallel to the dislocation line. The displacements of the atoms are in the direction of \mathbf{b} and decrease in inverse proportion to the distance from the dislocation line. The spacings of the planes of atoms parallel to the dislocation line are assumed to be unaffected.

The pure edge dislocation can be envisaged as the edge of an additional half-plane of atoms. The Burgers vector b is perpendicular to the extra half-plane and so perpendicular to the dislocation line. Planes of atoms perpendicular to the dislocation line maintain their normal spacings. Within these planes, the displacements of the atoms have components R_1 parallel to b and R_2 perpendicular to b where

$$R_1 = \frac{b}{2\pi} \left[\Phi + \frac{\sin 2\Phi}{4(1-\nu)} \right]$$

$$R_2 = -\frac{b}{2\pi} \left[\frac{1-2\nu}{2(1-\nu)} \ln|r| + \frac{\cos 2\Phi}{4(1-\nu)} \right]. \quad (18)$$

Here Φ is the angle measured from the direction perpendicular to the extra half-plane, r is the distance from the dislocation line, and ν is the Poisson's ratio for the material.

The problem of the kinematical diffraction from a needle-shaped crystal having an axial screw dislocation was worked out by Wilson [1952] who showed that the reciprocal lattice points were broadened into discs perpendicular to the dislocation axis, assumed to be the c -axis. The width of these discs then increased with $|b|l$ where b is the Burgers vector and l the appropriate index. The reciprocal lattice maxima for $l = 0$ were unaffected by the dislocation. The equivalent results for a pure edge and mixed dislocation have also been obtained (see Krivoglaz and Ryaboshapka [1963]).

The possibility of observing the diffraction effects due to individual dislocations by using X-ray diffraction is remote since the volume of material appreciably affected by the presence of one dislocation is much too small to give measurable intensities. In most materials containing large numbers of dislocations the orientations of the dislocations may be more or less random or else the dislocations may be segregated into dislocation networks forming small-angle grain boundaries. The diffraction problem then comes within the province of diffraction by a mosaic crystal or by a crystal having internal strain. In each case the effect on the intensities can be evaluated statistically (see Chapter 16). However, in some naturally occurring grain boundaries and in grain boundaries between specially-prepared, superimposed crystals where the lattices on the two sides are related by a simple rotation, periodic arrays of parallel dislocations may be formed. Electron diffraction patterns (Guan and Sass [1973]; Sass and Baluffi [1976]) and X-ray diffraction patterns (Gaudig and Sass [1979]) from such arrays show superlattice reflections corresponding to the

dislocation periodicity. It is important, especially in the electron diffraction case to differentiate between extra spots owing to a dislocation network and similar spots which could be produced by double diffraction from the two superimposed crystals. The patterns from dislocation networks, however, do have distinctive features, and agreement has been found between observed patterns and those calculated on this basis (see Sass [1980]).

It now appears quite feasible to obtain electron diffraction patterns from thin needle crystals containing screw dislocations (Cowley [1954]) or from small regions of thin crystals containing individual dislocations of any type (Cockayne et al. [1967]; Cowley [1970]). However, experimental difficulties and the uncertainties of interpretation of intensities strongly affected by dynamical diffraction have so far discouraged any detailed studies by these methods.

18.4.2. The imaging of dislocations

Some indication of the form of the contrast in electron microscope images of dislocations can be derived from arguments based on the column approximation. For columns passing through the dislocation core the planes of atoms are displaced as at a stacking fault except that the displacement takes place over a distance of tens or hundreds of Angstroms. Therefore the projected dislocation line may be expected to show oscillatory contrast similar to that of the stacking fault fringes.

Away from the dislocation line the main effect of the dislocation is to tilt the lattice planes towards or away from the Bragg angle, the tilt being in opposite directions on the two sides. Therefore, except when the undistorted crystal is at the exact Bragg angle, the contrast may be expected to be asymmetric across the dislocation line image.

These deductions are in accord with experimental observations for some cases, but the observed images show a great amount of detail, in general, depending on the diffraction conditions, the Burgers vector and the elastic constants of the material. A system for the rapid computer-generation of theoretical images for various values of these parameters has been evolved by Head [1967]. The method is normally used with a 2-beam dynamical theory approximation but extension to n -beam is possible. By this means images of dislocations, stacking faults or other defects may be calculated for all possible combinations of parameters for a particular system. Then comparison with observed images allows a unique identification of the form of the defect (Humble 1970; Head et al. [1973]).

For the simplest 2-beam theory for isotropic material the lack of distortion of the spacings of some planes gives the result that for $\mathbf{g} \cdot \mathbf{b} = 0$ the contrast of a dislocation line image will be zero, where \mathbf{g} is the diffraction vector.

This simple relationship has been used widely for identification of dislocations, but, as shown by detailed calculations, it may not be valid for anisotropic materials or under n -beam conditions, and so must be used with care.

18.4.3. Averaging over angles of incidence

In electron microscopy the incident beam is normally made convergent at the specimen level by the focussing action of the condenser lenses and the fore-field of the objective lens, but the angle of convergence rarely exceeds 10^{-3} radians and is usually less than the angular width for the occurrence of a strong reflection. Hence the assumption of a plane parallel incident beam does not give any serious errors.

For scanning transmission electron microscopy it is often convenient to use a wide-angle detector to collect a large proportion of the scattered radiation. By application of the Reciprocity relationship (Cowley [1969]) this is seen to be equivalent to the use of a very large angle of incidence for conventional transmission electron microscopy. The image contrast is then given by averaging over a large range of incident beam directions. The effect is a strong reduction of the contrast of stacking fault fringes. For dislocation images the overall contrast tends to be somewhat reduced, the oscillatory component tends to vanish and the dislocation image tends to be uniformly dark (in a positive print of a bright field image) (Booker et al. [1974]).

The images of dislocations in X-ray topographs represent an extreme case of averaging over incident angle. While the incident beam is reasonably well collimated by normal X-ray diffraction standards, its convergence angle tends to be several orders of magnitude greater than the angular widths of reflections from perfect crystal regions (typically 10^{-5} radians). The dislocation images are almost uniform black lines with very little indication of any oscillatory contrast or asymmetric profiles (Lang [1959]). It is possible, of course, to obtain well collimated incident X-ray beams; for example by using a beam reflected or transmitted from a thick near-perfect crystal. Then the full range of dynamical contrast effects can be observed as in electron microscopy. Fig. 18.4 is an X-ray topograph of a silicon crystal showing thickness fringes and dislocation images with some oscillatory contrast.

18.4.4. n -beam diffraction effects

If the column approximation is assumed, calculations of diffraction amplitudes or image intensities can be made by suitable modification of the computer programs used for perfect single crystals. The Howie and Whelan [1961] equations given in Chapter 10 have been used by a number of people. The methods of Chapter 11 may be applied equally well

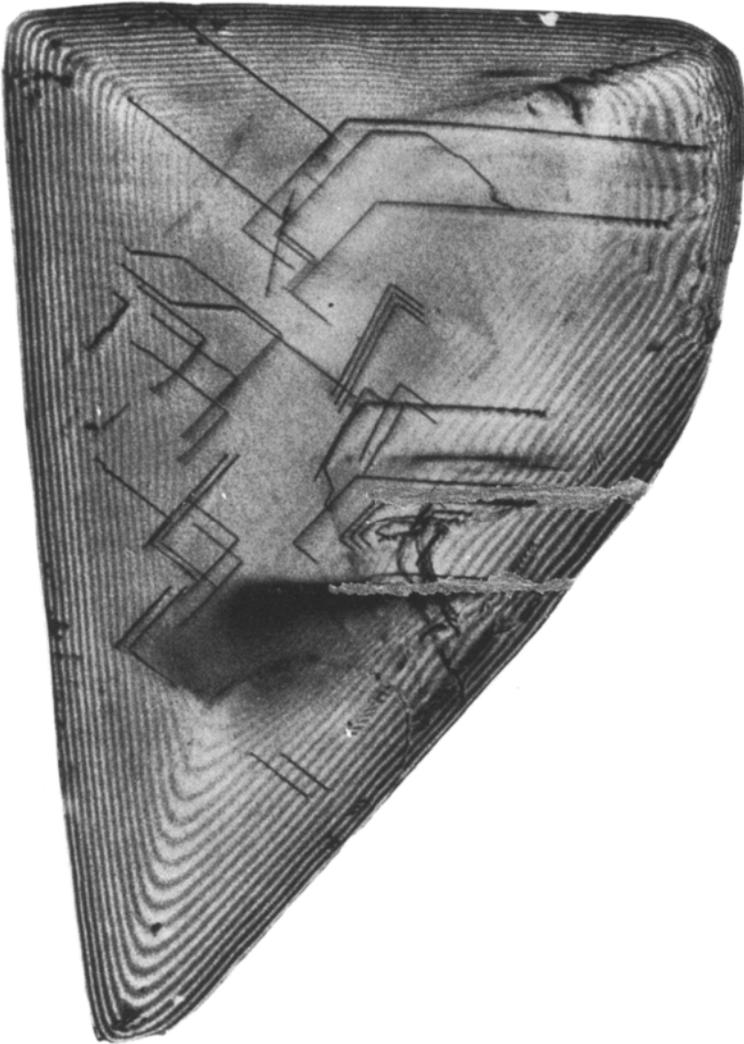


Fig. 18.4. Equal thickness fringes and dislocation images in an X-ray topograph of a pyramidal silicon crystal of maximum diameter 1.5 mm. Obtained using $\text{AgK}\alpha$ radiation and a (111)-type reflection. (From Borrmann [1964].)

by replacing the constant values of the Fourier coefficients of the potential distributions of the various slices by values which are functions of the depth in the crystal. The values of these Fourier coefficients for the various slices are then generated by a sub-routine from the assumed form of the perturbation of the structure.

If the column approximation is not valid, as in the calculation of high resolution (2 to 3 Å) images of the distorted crystal structure, it is possible to use the technique mentioned at the end of Chapter 11 whereby the dislocation or other defect is assumed to occur periodically and the calculation is made as for a structure having a large unit cell, using a very large number of diffracted beams.

One important outcome of the n -beam approach is the weak-beam dark-field imaging method of Cockayne et al. [1969]. These authors showed that if, instead of a strongly excited inner reflection, a weak reflection corresponding to a reciprocal lattice point far from the Ewald sphere is used to form a dark-field image, the images of dislocations produced are much sharper. Widths of 10 to 20 Å are common, as compared with the 100 to 200 Å normally observed. Consequently a wealth of fine detail on dislocation separations and interactions has been made available.

Cockayne has pointed out that a simple pseudo-kinematic description gives reasonable agreement with the observations. It may be argued that for a reciprocal lattice point far from the Ewald sphere appreciable intensity is generated only when the lattice planes are tilted through a relatively large angle and this occurs only near the dislocation core. However the resolution of the images is now such that a description in terms of lattice planes is becoming inadequate for the core region of a dislocation and the calculation of the intensity must, in any case, involve n -beam dynamical calculations.

The methods available for the calculation of high-resolution images of defects in crystals have been discussed by Rez [1978] and by Anstis and Cockayne [1979]. The column approximation is valid if the crystal is sufficiently thin. The limiting thickness depends on the accuracy of the results required but may be estimated roughly, as mentioned in Section 10.6, from the requirement that the lateral spread of the waves owing to Fresnel diffraction should be small compared with the resolution limit. Another way of looking at the column approximation is to note that only the finite set of sharp Bragg reflections is considered and no account is taken of the fact that defects give rise to diffuse scattering around or between the Bragg peaks. Thus the assumption is made that the excitation error for the scattering by the defects is the same as that for the neighboring Bragg reflection so that the Ewald sphere is approximated by a set of planar patches perpendicular to the

beam and centered on the reciprocal lattice points. This approximation could clearly be good even for large thicknesses if the diffuse scattering occurs only in the regions very close to the Bragg reflections, i.e. if the defects involve only slow variations of the lattice spacings.

A better approximation is that introduced by Takagi [1962] and Jouffrey and Taupin [1967]. This introduces a variation of excitation error which is linear with the distance of the scattering from a Bragg reflection. Thus the Ewald sphere is, in effect, approximated by a set of planar patches, which are tangential to the sphere, around each reciprocal lattice point. This is an obvious improvement over the column approximation but fails when the diffuse scattering from the defects is far from the Bragg reflections, i.e. when the variations of atom spacings become appreciable within the distance separating the lattice planes.

The method of Howie and Basinski [1968] goes one step further in including second-order as well as the first-order terms and so gives an accurate representation of the Ewald sphere for all diffuse scattering. It can be used to represent the diffraction and imaging from defects containing quite large local distortions of the crystal structure. Calculations using this method are laborious if the scattering far from Bragg reflections must be included, but these are feasible with reasonably localized scattering such as is given by strain fields extending quite close to dislocation cores. The method of periodic continuation mentioned in Section 11.5 and Subsection 13.4.2 may best be applied when the distortions of the crystal structure occur over only a small region or where the distortion can be divided into small distinct regions which may be calculated separately. It can deal with arbitrarily large displacements of atoms or with disorder in the occupancy of lattice sites but is not appropriate for extended strain fields. Hence it is complimentary to the above methods which treat successive levels of deviation from a perfectly periodic structure.

The recent improvement of electron microscopes to give resolutions which are routinely better than 2 \AA , and approaching 1 \AA in some cases, has permitted the imaging of the local arrangements of atoms in defects in metals and semiconductors in favorable cases. Single images give, at best, two dimensional projections of the atomic arrangements. The complications arising from dynamical diffraction may be minimized by the channelling effects in images obtained in orientations parallel to principal axes; but dynamical diffraction and experimental factors such as radiation damage make it difficult to use stereoscopic techniques for three-dimensional imaging. Thus the most successful applications of high resolution electron microscopy have been for defects such as dislocations, planar faults or interfaces parallel to the incident beam.

Krivanek et al. [1977] showed an image of a grain boundary in germanium, viewed in the [110] direction, in which the positions of all atom rows parallel to the beam could be deduced and in more recent studies details of atom arrangements in grain boundaries have been determined with increasing clarity. End-on images of straight-line dislocations have been obtained (e.g. by Bourret et al. [1979, 1982]) and detailed comparisons with models of the dislocations have been made using many-beam dynamical calculations with periodic-continuation assumptions. Summaries of the current capabilities of the techniques and examples of recent results have been given by Spence [1988], Amelinckx and Van Dyck [1993] and the collection of papers edited by Bradley and King [1993].

This Page Intentionally Left Blank

Diffraction from surfaces

19.1. Introduction

In recent years the proliferation of diffraction techniques developed specifically for the study of the structures of surfaces has added greatly to both the fundamental scientific basis and the range of practical exploitations of surface science. The degree of crystalline order, the deviations of the surface atom positions from those for a continuation of the bulk structure, the rearrangement of atomic bond configurations, the presence and arrangements of foreign atoms on the surface and the modifications of the electron energy levels are all of essential importance for the determination of the physical properties of small particles and thin films and the chemical reactivities of solids in gaseous or liquid environments. The few top-most layers of atoms on the surfaces of solids are enormously influential in determining the way in which solids interact with their surroundings.

As in the case of three-dimensional solids, the diffraction methods that we will be considering may be divided into two categories. For the essentially two-dimensional structures of surfaces, there are the purely diffraction methods, in which an averaging is taken over many millions of "unit cells", assumed to be equivalent, or described in terms of statistical distributions for which a preferably kinematical, or if necessary, a dynamical diffraction theory can be applied to the whole assembly. Then there are the methods of diffraction from very small individual regions, usually combined with imaging techniques, possible as in the case of electron beams for which a very high intensity of incident radiation may be concentrated in a very small area, and strong lenses are available to provide images of high magnification.

Of the purely diffraction techniques, low-energy electron diffraction (LEED) has for many years been the dominant method, its great advantage being that, because of the very small penetration of low-energy electrons into solids even at near-normal incidence, the diffraction information is obtained from only the top-most few layers of atoms. The major difficulty with LEED is that, because of the strong interaction of the electrons with even single layers of atoms, the diffraction is heavily dynamical and the theory needed for the interpretation of the diffraction intensities

is complicated and non-intuitive. The possibility of using X-ray and neutron diffraction methods for surface studies, with a kinematical basis for the interpretations of the diffracted intensities, has arisen more recently, especially with the development of synchrotron radiation sources for X-rays and the availability of high-flux neutron sources. An essential difficulty for these radiations is that, because the penetration of the radiation into the bulk is great, of the order of tens or hundreds of interatomic distances, the diffracted amplitude from the essential few top layers of atoms on a surface is liable to be hidden by the scattering from the bulk of a solid. Special geometries must be used.

Our description of surface diffraction techniques will follow a roughly geometric classification, consistent with the properties of the various radiations. First we consider the case of transmission through very thin films of solids, for which the total scattering from the two surfaces of a film may be expected to be appreciable relative to the scattering from the bulk: then the scattering of radiation from the surface of a crystalline solid with the incident beam at grazing incidence to limit its penetration into the bulk, and then the diffraction with near-normal incidence and small penetration, as in LEED. Finally some techniques depending on the emission of radiation from surface atoms are mentioned.

The information regarding the surface layers of atoms which is desired may be categorized roughly as follows. For pure, clean crystalline surfaces, surface relaxation implies a rearrangement of the surface atoms so that their positions differ from those for an extension of the bulk structure. In accordance with the local symmetry of the atomic arrangements, the surface atom displacements may, or may not, be perpendicular to the surface. Surface reconstructions imply a rearrangement of the surface atoms, in response to their modified environment, so that the symmetries and periodicities of the surface layers differ from those of the bulk: this frequently involves the formation of surface superlattice structures. The chemical composition of the surface layers may or may not be identical with the bulk composition.

Layers of foreign atoms may be added to a surface by segregation from the bulk, by physical absorption or by chemical reaction. The nature of the foreign atoms and their positions relative to the bulk structure need to be determined. As in the case of bulk structures, the degree of ordering of the surface layers may vary over a wide range, from almost perfect ordering to liquid-like disorder, and the diffraction methods may be used to characterize the degree of order in terms of order parameters.

For kinematical diffraction conditions, it may be considered that the two surface layers on a thin film give sufficient scattering to be detectable if they constitute more than a few percent of the scattering matter i.e. if the

"bulk" material has a thickness of no more than a few nm. For X-rays and neutrons, the scattering from such thin films is so small that the use of single-layer thin films is scarcely feasible. Specimens having usable scattering power could be built up by superposition of many parallel thin films, but it is more practicable to make an assembly of randomly oriented thin films and then obtain what is essentially a powder pattern. This was done, for example, by Nielsen et al. [1987] who studied argon absorbed on graphite to investigate the melting of its two-dimensionally ordered distribution using synchrotron X-rays. The powder pattern from such an assembly could be measured with high precision. The intensity distribution is given by an averaging over all orientations of the square of the Fourier transform of the structure of the two-dimensionally periodic layer lattice. It is possible that the methods of Rejtveid refinement (Section 16.3.4) could be applied to patterns of this sort to extract good three-dimensional data on the positions of absorbed atoms.

19.2. Surface imaging and diffraction with electrons in transmission

19.2.1. Phase-contrast imaging

Under high-resolution conditions, when a number of diffracted beams from the bulk of a specimen are included in the objective aperture of an electron microscope, the periodic image of the crystal lattice dominates the micrograph and it is very difficult or impossible to detect small modifications of the surface structure such as surface steps or projections or surface layers of structure differing from that of the bulk. However, if the strong diffraction spots from the bulk crystal are prevented from contributing to the image, the scattering from the surface structure may contribute a detectable image contrast. This may be done for crystals in principal orientations, at the expense of image resolution, by using a small objective aperture to exclude the bulk reflections or, with no loss of resolution, by tilting the crystal so that no reflections from the bulk structure appear within a larger objective aperture. Under these circumstances, the bulk of the crystal acts as a medium of uniform refractive index, adding only a constant phase and amplitude change to the image wave. Then the scattering from the atom layers on each surface can be treated as scattering from isolated layers of atoms and can be described by the weak-phase-object approximation.

In this way, Moodie and Warble [1967] were able to observe steps on the surfaces of small MgO crystals and Iijima [1977] showed clear images of steps, one atom high, on the surface of thin graphite crystals, with individual

tungsten atoms sitting on the steps. For thicker crystals when there are strong diffracted beams excited, the intensity of the bright-field image may vary strongly with thickness. Then the line image of a step, given by phase-contrast, may be associated with an appreciable intensity change across the step, as in the case of the imaging of single-atom steps on silicon by Iijima [1981] or on MgO plates by Lehmpfuhl and Warble [1986].

One limitation on the imaging of surface structure in this way is that the image contrast is reduced by the inelastic scattering from the bulk of the crystal which is strongly peaked around the incident beam position. Bulk-plasmon scattering increases with the crystal thickness but there is also surface-plasmon scattering which does not depend on thickness and may be important for even very thin crystals.

19.2.2. Crystal terminations and ordered surface superlattices

High energy electrons are scattered sufficiently strongly by only one or two layers of atoms to allow diffraction spots to be observed from ordered surface layers which are only a few nm in extent, laterally, unless these diffraction spots are obscured by strong scattering from the bulk of the crystal. In favorable cases, the surface diffraction spots may be selected to form images of the corresponding surface structures. Thus detailed studies of ordered surface structures may be made, sometimes with imaging to show their distributions. A major difficulty for such experiments is that, because of the forward-scattering geometry, the electron diffraction pattern is derived mostly from the projection of the structure in the incident beam direction and so is insensitive to atom displacements perpendicular to the plane of the thin film. However, important information may follow from the fact that the diffraction intensities reflect the symmetry and structure of the projection in the beam direction.

It was pointed out in Section 15.8 that the symmetry and periodicity of the projected structure of a thin crystal made by the successive stacking of close-packed layers of metal atoms depends on the number of atom layers in the stack. A single layer of gold atoms, for example, is hexagonal with a periodicity equal to the interatomic distance, 2.88 Å, as in the A layer of Fig. 18.2, giving a hexagonal diffraction pattern with spots at a radius of $1/2.50 \text{ \AA}^{-1}$. A second layer, added at the B position of Fig. 18.2, gives a projected unit cell of the same dimensions and so a diffraction pattern with the same dimensions although different intensities. The addition of a third layer, at the C position of Fig. 18.2, however, gives a projected structure which is hexagonal with

a unit cell $\sqrt{3}$ times smaller and rotated by 30° with respect to the single-layer cell. The corresponding diffraction pattern has spots in a hexagonal array at $1/1.44 \text{ \AA}^{-1}$, the familiar pattern of {220} spots for a FCC structure viewed in the [111] direction. The single-layer spots would have indices such as $2/3, 4/3, 2/3$ for a FCC cell.

As further layers of atoms are added, the single-layer spots appear for $2n+1$ layers but vanish for $3n$ layers, so long as the kinematical approximation and the projection approximation apply. Many-beam dynamical calculations such as those of Lynch [1971] show that, in practice, the intensities for the single-layer spots are much weaker for $3n$ layers than for $3n+1$ layers, but not zero. Thus if high-resolution dark-field transmission electron micrographs are obtained from thin films of gold using any one of the single-layer spots, the image appears bright for $3n+1$ layers and dark for $3n$ layers. The dark-light boundaries correspond to single-atom steps on one or other of the film surfaces (Cherns [1974]).

In much the same way Krakow [1979] was able to image the structure of the surface atom layer on the (100) face of a thin gold crystal. The single surface layer of an FCC structure has the periodicity of $a_0/\sqrt{2}$, whereas the projection of the bulk structure has the periodicity $a_0/2$. The former periodicity was clearly revealed by an off-line image processing technique.

One way of interpreting the diffraction spots from the termination layers of crystals is to consider that the bulk crystal, with an integral number of unit cells, gives the scattering maxima at the reciprocal lattice points (with associated shape-transform extensions) and the extra surface layers give scattering with continuous lines in the reciprocal space, perpendicular to the surface, which then cut the Ewald sphere to give the extra spots.

Alternatively, one might consider, in the kinematical approximation, that the shape transform, extending from all reciprocal lattice points, gives an intensity which oscillates, depending on the distance of the reciprocal lattice point from the Ewald sphere, as given by equation (2.40). For a flat Ewald sphere parallel to one of the principal reciprocal lattice planes, the distance of the reciprocal lattice points from the Ewald sphere is a multiple of the inverse thickness for an integral number of unit cells, giving zero intensity (or maximum intensity if the point is on the plane) but for thicknesses involving fractional unit cells, the shape-transform value is not zero and a corresponding spot appears.

Such a consideration of the shape-transform provides a ready explanation for the appearance of the dark-field image of Fig. 19.1. This image was obtained by Takayanagi [1981] from thin layers of Ag evaporated on the surface of a crystal of molybdenite. The Ewald sphere was tilted to avoid the

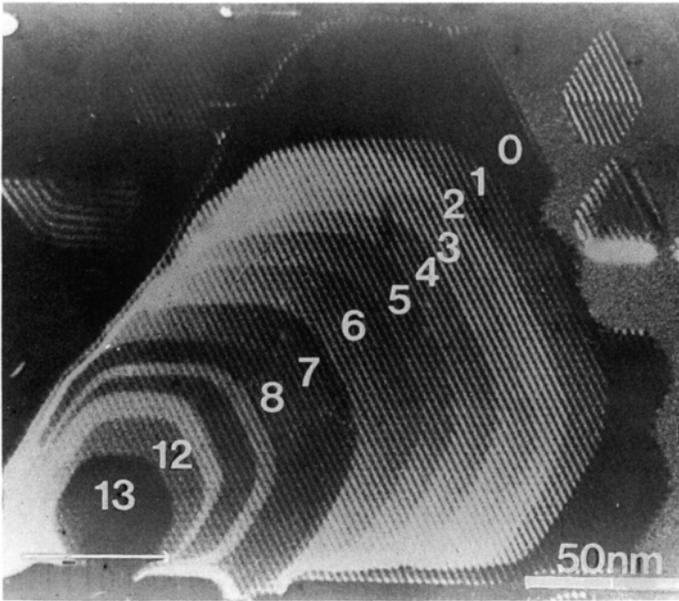


Fig. 19.1. Dark-field TEM image from thin layers of Ag deposited on a molybdenite crystal. The numerals 1 to 13, on the regions of different contrast, indicate the numbers of monolayers of Ag present. (from Takayanagi [1981]).

strong zero-layer reflections from the molybdenite. The dark-field image was obtained by selecting the diffraction spot given when the Ewald sphere cut the extended shape transform from the silver layers. For a silver layer of thickness N times the single-layer thickness c , the intensity of the diffraction spot is proportional to $(\sin\pi Ncw/\sin\pi cw)^2$, where w is the distance from the reciprocal lattice point in the direction perpendicular to the plane of the layer. The intensity of the diffraction spot, and of the corresponding image area, varies sinusoidally with N . Thus it can be seen in Fig. 19.1 that the areas marked with the integers 0 to 13 correspond to regions having the indicated number of monolayers of Ag atoms. The fine vertical fringes are moiré fringes, given by interference between the silver reflections and an adjacent molybdenite reflection.

19.2.3. Structure analysis of surface superlattices

When super-structures form in the surface layers of single-crystal samples, with two-dimensional unit cell

dimensions which are multiples of the bulk cell dimensions, the corresponding superlattice spots appear weakly in the transmission diffraction pattern. A regularly spaced subset of these spots coincides with the strong spots from the bulk crystal and so are inaccessible. The remainder of the spots however come only from the few monolayers on the surface.

It was shown by Spence [1983] that, to a good approximation, the intensities of these spots can be interpreted on the basis of a simple kinematical approximation and so can be used as the basis for structure analyses of the surface superstructure. A simple argument favoring this conclusion is made by considering that the exit wave from the bulk crystal, $\Psi_B(xy)$, is multiplied by the transmission function of the surface layer on the exit side, considered as a weak phase-object, $1-i\sigma\Phi_S(xy)$, so that the diffraction pattern amplitude is

$$\begin{aligned} I(uv) &= |\Psi_B(uv) * \{\delta(uv) - i\sigma\Phi_S(uv)\}|^2 \\ &= |\Psi_B(uv)|^2 + \sigma^2 |\Psi_B(uv) * \Phi_S(uv)|^2. \end{aligned} \quad (1)$$

If, in the diffraction pattern from the bulk, the zero beam has a much greater amplitude than any of the diffraction spots, the second term of (1) approximates to the diffraction pattern of the surface layers as if isolated in space. This happens when the bulk crystal is thin and usually also for some particular larger thicknesses.

Takayanagi [1984] observed the superlattice spots from the 7x7 reconstruction on the Si(111) face in transmission patterns of thin crystals of silicon heated in the ultra-high vacuum environment of a special electron microscope. The existence of this surface structure was well known from LEED observations and from imaging in the pioneering work on scanning tunnelling microscopy (STM) by Binnig et al. [1982] but the complication of the structure had defied analysis. The complete structure analysis was made by Takayanagi et al. [1985] revealing the so-called DAS model of the surface layers which was later shown to be consistent with the LEED and STM observations. Later structure analyses of surface structures have included the analysis of the Si(100)-2x1 structure by Jayaram et al. [1993] for which the dynamical scattering effects were taken fully into account.

19.2.4. Crystal profile imaging

It was shown by Marks and Smith [1983] that valuable information could be obtained on the structures of the surfaces of crystals by viewing the edges of small particles or thin films in profile, with the incident beam of a high-resolution electron microscope parallel to a crystal face of width 10 to 100 Å. For the plan-view imaging techniques, mentioned above, the resolution in the image is usually

restricted by the necessity of excluding the strong bulk reflections, and the information on atom positions is restricted to the two dimensions parallel to the surface. For the profile imaging method, in contrast, the full resolution capabilities of the modern high-resolution microscopes may be applied. The information on atom positions includes that on the shifts of atoms perpendicular to the surface as well as their positions in one dimension parallel to the surface. The other dimension parallel to the surface may sometimes be visualized by rotating the specimen about the normal to its face. The image of Fig. 13.6, for example, shows clearly the atom positions in the reconstructed surface of a gold crystal. A summary of the results on surface structures of noble metal small particles and crystal edges is given by Marks [1986]. Applications to studies of surfaces of oxides and semiconductors have been reviewed by Smith et al. [1991].

An important point to note for the interpretation of profile images is that the positions of the dark spots in images such as Fig. 13.6 do not necessarily correspond to the positions of the rows of atoms seen in projection. The scattering from the heavy-metal crystals is strongly dynamical. For axial orientations, the wave fields along rows of atoms in the bulk reflect the symmetry of the surroundings of the atom rows so that the intensity maxima or minima are confined to the high-symmetry positions. But these considerations do not apply to surface atom rows for which the surroundings are asymmetric. It is necessary that detailed dynamical diffraction calculations are made, using the method of periodic continuation to take into account the non-periodic nature of the specimen. In this way it was shown by Marks [1984], for example, that the true expansion of the surface layer of gold atoms, expressed as a percentage of the bulk inter-planar distance, could differ from the expansion suggested by the image by about 5% for properly focussed images. The accuracy with which projected atom positions may be determined is of the order of 0.1 Å or better.

19.3. Reflection from surfaces: grazing incidence

19.3.1. Kinematical approximation: X-rays and neutrons

For X-rays and neutrons, and also for high-energy electrons, a method commonly used in order to emphasize the scattering from surface layers relative to the scattering from the bulk of a sample is to bring the incident beam in at a small glancing angle to a flat surface. For high-energy electrons, the reflection high-energy electron diffraction (RHEED) technique has been widely used since the early days of electron diffraction (Nishikawa and Kikuchi [1928]) for qualitative studies of surfaces but, because the scattering is strongly dynamical, the quantitative interpretation of RHEED patterns has not been possible until quite recently.

For X-rays and neutrons, however, the kinematical approximation is usually valid, at least for the relatively weak scattering from surface layers, and quantitative interpretations have been possible from the start. For these radiations, the applications to surface studies had to await the availability of very intense, well-collimated beams, such as those from synchrotron radiation sources for X-rays and from high-power nuclear reactors or intense spallation pulsed sources for neutrons. An account of recent applications of X-ray and neutron techniques is given by Zabel and Robinson [1992]. The use of synchrotron X-rays for surface studies is described by Robinson [1990].

The reciprocal space for a crystal terminated at a flat surface, and the regions of reciprocal space studied by various techniques, are indicated in Fig. 19.2. The reciprocal space coordinates, u, v , are taken to be in the plane of the surface and the coordinate w is perpendicular and parallel to the z -axis. Through each bulk reciprocal lattice point, including the origin, there is a continuous line perpendicular to the surface, known as the crystal truncation rod and arising from the discontinuity at the surface (Robinson [1986]). Its form may be derived by considering that the crystal structure is cut off by

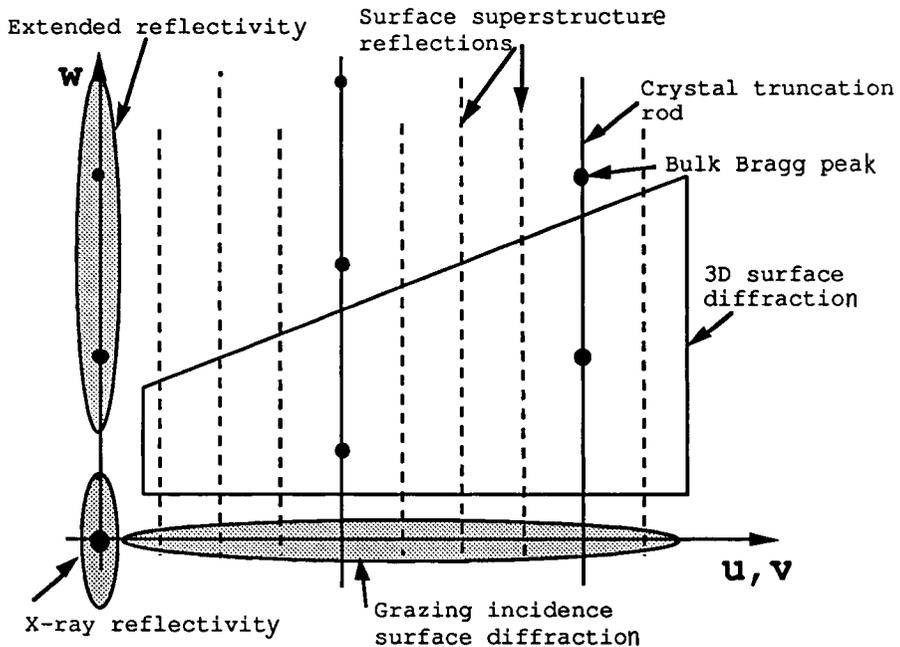


Fig. 19.2. Diagram suggesting the features and regions of reciprocal space which are accessible by surface scattering of X-rays and neutrons. (after Robinson [1990]).

multiplying by a step function, $s(z)$ which is 1 for $z < 0$ and 0 for $z > 0$. Then the Fourier transform of the crystal, the weighted reciprocal lattice, is convoluted with the Fourier transform of $s(z)$, namely $S(w) = \{\delta(w) + i/\pi w\}/2$. Thus each reciprocal lattice point is extended in a streak in the w -direction with intensity falling off in proportion to the square of the distance.

The intensity distribution along the truncation rods is modified by any deviation of the surface structure from the ideal case of termination of the bulk structure by a planar discontinuity. The truncation rod through the origin point gives the "X-ray reflectivity", obtained by varying the glancing angle of incidence and the angle of the detected, scattered beam simultaneously. Fourier transform of the X-ray reflectivity function gives the auto-correlation of the projection of the structure on the surface normal and so is used for the detection and study of surface roughness, surface reconstructions and adsorbed layers on surfaces.

The intensities in the crystal truncation rods around all the reciprocal lattice points are given, at least for simple structures, by the convolution of the X-ray reflectivity function with the reciprocal lattice intensity maxima for the crystal (You [1992]). It was pointed out by Harada [1993] that the intensity distributions for a microscopically rough surface may be described using a formulation which is similar to that used for the description of short-range-order scattering for binary alloy systems (see Section 17.5).

In addition to the lines through the bulk reciprocal lattice points, there are lines or rods of scattering perpendicular to the surface due to incomplete unit cells at the crystal surface or to surface reconstructions or superlattices. When both the incident beam and diffracted beam make grazing angles with the surface, the intensities on these rods close to the $w = 0$ positions are measured. From these intensities it is possible to deduce the projections of the surface structures on the crystal surface, as in the case of the high-energy transmission electron diffraction, but with the great added advantage that the interpretations may be made on the basis of the simple kinematical approximation. The intensity distributions along the rods, found using larger incident or diffracted angles, provide information on the third dimension of the structures. Studies on a number of structures of various types are given in Zabel and Robinson [1992]. One important restriction of the X-ray technique, which is even more important for neutrons, is the requirement that in order to obtain sufficiently strong signals, especially for grazing angles of incidence, the crystal surface must be flat and uniform over quite large areas. However the rapid development of special surface-preparation techniques has made an increasing number of systems available for study.

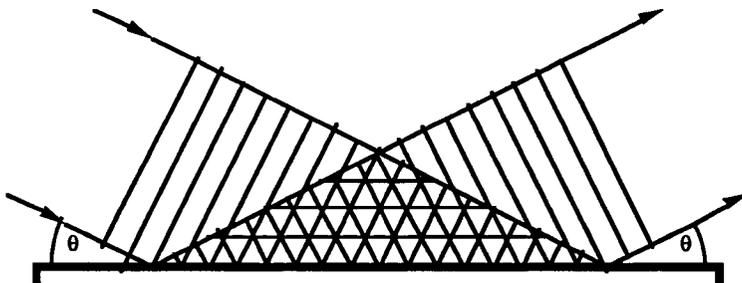


Fig. 19.3. The standing-wave field established by interference of waves incident and reflected from a surface.

As is the case for studies of bulk materials, the special characteristics of neutron scattering provide important possibilities for neutron surface diffraction studies. The magnetic structures of surfaces, or of thin films, of ferromagnetic or antiferromagnetic materials may be studied. Also neutrons are particularly effective in the study of light-atom materials, such as thin organic or biological films on flat solid surfaces or even on liquid surfaces.

19.3.2. Standing wave techniques

In the reflection geometry, the incident beam and the specular reflected beam overlap in the space outside the crystal and interfere to form a standing wave pattern, as suggested by Fig. 19.3. For angle of incidence θ , the periodicity of the standing wave is given by $D = \lambda / (2\sin\theta)$.

If the effect of refraction at the surface is neglected, $D = d_{hkl}$ for diffraction at the Bragg angle from the hkl planes parallel to the surface. For high-energy electrons, the refractive index is slightly greater than unity so that the angle of incidence is less than the Bragg angle for the surface planes so that $D > d_{hkl}$. However there is strong specular reflection for very small grazing angles of incidence, much less than the angle of incidence for any diffracted beam, so that D may be made very large.

For X-rays and neutrons, the refractive index is normally slightly less than unity so that for angles of incidence less than the critical angle, θ_c , (usually 1 to 10 mrad.) there is total external reflection. The periodicity of the standing wave field outside the crystal may thus take any value greater than $\lambda / (2\sin\theta_c)$ which is, for example, 80 Å for a gold crystal and 200 Å for silicon. For diffraction from the crystal lattice planes, the incident angle is relatively large and the effect of the deviation of the refractive index from unity has very little effect so that the periodicity of

the standing wave pattern is very close to d_{hkl} , both inside and outside the crystal.

The first experiments showing the existence of the standing wave field made use of a specular Bragg reflection. Batterman [1964, 1969] measured the intensity of the fluorescent X-radiation from lattice atoms or from impurity atoms as a function of the angle of incidence. The phase of the reflected wave and hence the positions of the nodes and antinodes of the standing wave field vary strongly as the angle of incidence passes through the range of total reflection around the Bragg angle (see Fig. 8.6). The positions of the emitting atoms, relative to the lattice planes, may therefore be deduced from the form of the dependence of the fluorescent radiation intensity on the angle of incidence. The applications to studies of surface structure were initiated by Cowan et al. [1980] who detected the variation of intensity of the fluorescent $K\alpha$ radiation from a layer of bromine atoms on the surface of a silicon crystal as the angle of incidence of $MoK\alpha$ radiation was varied through the silicon (220) reflection angle.

The demonstration of the existence of the standing wave field outside the crystal in the case of total external reflection and applications to surface structure problems were initiated by Bedzyk [1988] (also Bedzyk et al. [1990]). For example, the fluorescent radiation from Zn atoms embedded in the top layer of a Langmuir-Blodgett multilayer of arachadate on an Au surface was shown to vary sinusoidally with the angle of incidence of X-rays of energy chosen to coincide with the Zn absorption edge, allowing the height of the Zn layer above the surface to be determined as 218 Å (Bedzyk et al. [1989]).

19.3.3. RHEED and REM

Reflection high-energy electron diffraction (RHEED) has been widely used for surface studies for many years and is the subject of numerous reviews such as that of Yagi [1993] and the collection of papers in Larsen and Dobson [1988]. RHEED patterns are usually obtained when a beam of electrons of energy 20 to 200keV impinges on the flat face of a crystal at a glancing angle of a few degrees, usually in electron microscopes. Patterns are also obtained, especially in ultra-high vacuum, surface research, equipment in the energy range of 2 to 20keV and with correspondingly larger angles of incidence. The technique is then sometimes referred to as medium-energy electron diffraction (MEED).

A typical RHEED pattern from a flat single-crystal face is shown in Fig. 19.4 and Fig. 19.5 is a diagram of the corresponding geometry in reciprocal space. If we consider the incident beam to be close to the c -axis of a crystal and the c -axis lies in the surface plane, the Ewald sphere cuts

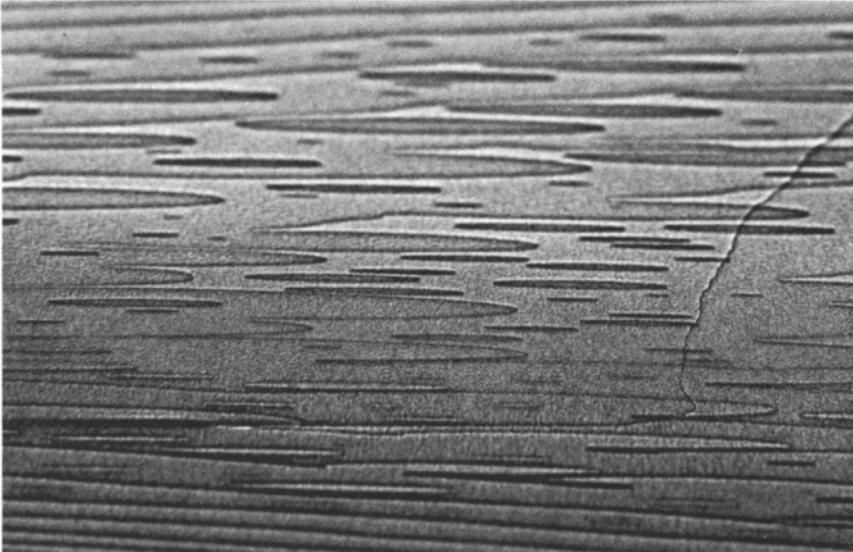


Fig. 19.4. Reflection high-energy electron diffraction (RHEED) pattern from the (110) surface of a GaAs crystal showing diffraction spots, K-lines, parabolas and a circle. (courtesy of N. Yao).

the $hk0$ reciprocal lattice plane to give a strong pattern of zero-Laue-zone (ZOLZ) diffraction spots close to the "shadow edge", the limit of the diffraction pattern given by the shadow of the specimen surface. The strong spots form on a circle passing through the incident beam position which is sometimes indicated when some of the broad incident beam passes the edge of the specimen. Concentric with this ZOLZ circle of spots, there are higher-angle circles of spots formed as the Ewald sphere cuts the successive hkl planes of reciprocal lattice points (see Fig. 19.5).

In the background of the pattern there is a complicated pattern of Kikuchi lines and Kikuchi bands, generated when electrons diffusely scattered in the crystal by thermal diffuse scattering and other inelastic scattering processes are further scattered by the crystal lattice (see Section 14.2). The diffuse scattering is strong because the average distance travelled by the incident electrons in the crystal at grazing incidence may be an appreciable fraction of $1\mu\text{m}$. Measurements have indicated that, for both the Bragg spots and the Kikuchi lines, from 40 to 80% of the electrons may have

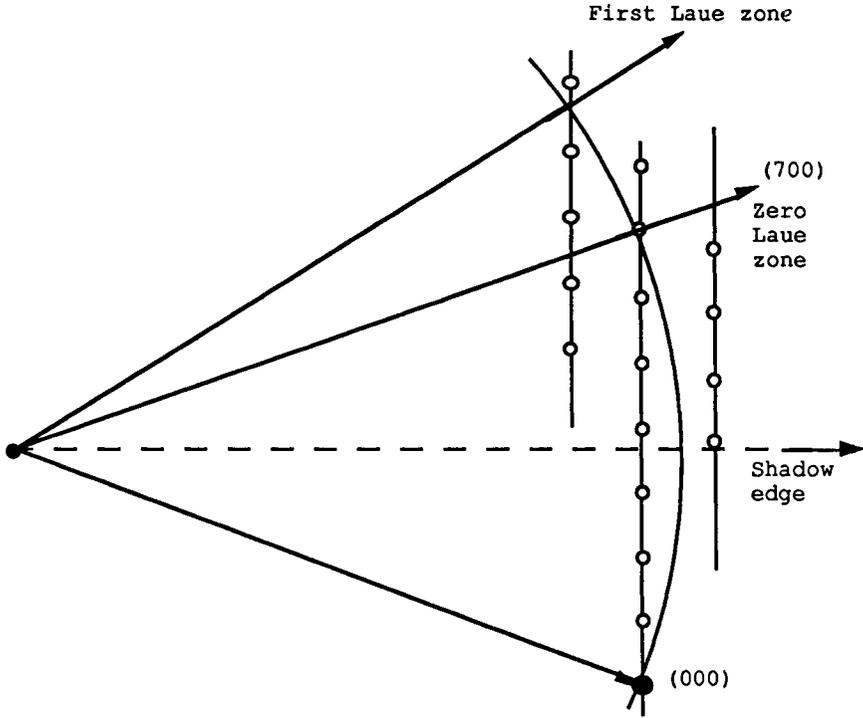


Fig. 19.5. Reciprocal space diagram showing the intersection of the Ewald sphere with lines perpendicular to a crystal surface through reciprocal lattice points to produce the spots in a RHEED pattern.

energy losses of more than a few eV.

In addition to the straight Kikuchi lines and bands, the patterns usually show sets of continuous parabolas and circles such as are prominent in Fig. 19.4. These curves may be regarded as the asymptotes of sets of Kikuchi lines generated by diffractions corresponding to closely spaced lines or planes of reciprocal lattice points. The circles, for example correspond to diffraction by the hkl spots of the successive Laue zone, or, in real space, diffraction by the lines of atoms almost parallel to the incident beam. It was pointed out by Emslie [1934] that the circles are not actually tangential to the Kikuchi lines but are displaced

slightly towards the shadow edge. This effect he attributed to a channelling of the diffusely scattered electrons along the potential "tubes" created by lines of atoms parallel to the surface. Similarly the intense parabolas having their axes parallel to the crystal surface may be attributed to diffraction of the diffusely scattered electrons by reflections corresponding to lines of reciprocal lattice points perpendicular to the surface. The continuous curvature of the parabolas and their separation from the near-tangential Kikuchi lines is attributed to a channelling of electrons along planes of atoms parallel to the surface (Gajdardziska and Cowley [1991]).

The phenomenon of surface channelling of the electrons, with a strong concentration of electrons in the one or two layers of atoms at the surface of the crystal, has been confirmed also for elastically scattered electrons, both by observation and by detailed calculation of the wave-fields in the crystal (e.g. Martin and Meyer-Ehmsen [1985], Wang et al. [1987]). The application of EELS analysis to the surface-channelled electrons offers the possibility of a sensitive means for compositional analysis of the top one or two layers of atoms on a crystal surface (Wang and Cowley [1988]). The channelling is usually associated with the so-called surface resonance condition which occurs when a strong diffracted beam is generated which is directed almost parallel to the crystal surface. When this occurs, both the specular beam and the whole RHEED diffraction pattern show a strong enhancement of intensity.

All features of RHEED patterns are displaced by the effects of refraction at the crystal surface. In terms of the small angles made by the electron beams with the surface, θ_0 outside the crystal and θ_1 inside the crystal, Snell's Law may be written $\cos\theta_0/\cos\theta_1 = n = 1 + \Phi_0/2E$, where Φ_0 is the mean inner potential and E is the accelerating voltage, or, in the small-angle approximation, $\theta_1^2 - \theta_0^2 = \Phi_0/2E$. Since the right hand side of this expression is usually of the order of 10^{-4} , the difference between θ_0 and θ_1 may be as great as 10^{-2} which is comparable with the Bragg angle for prominent lattice planes. Thus all spots and lines in the RHEED pattern are displaced by an amount which increases rapidly towards the shadow edge. Some diffracted beams which, in the absence of refraction, would occur close to the shadow edge may be trapped within the crystal by total internal reflection and so may contribute to the surface resonance effect.

In a way which is analogous to dark-field imaging in transmission electron microscopy of crystals, the RHEED diffraction spots may be directed along the axis of an electron microscope objective lens and used to form an image of the surface. In this technique of reflection electron microscopy (REM), images with resolutions of 1nm or better can be obtained which are highly sensitive to any

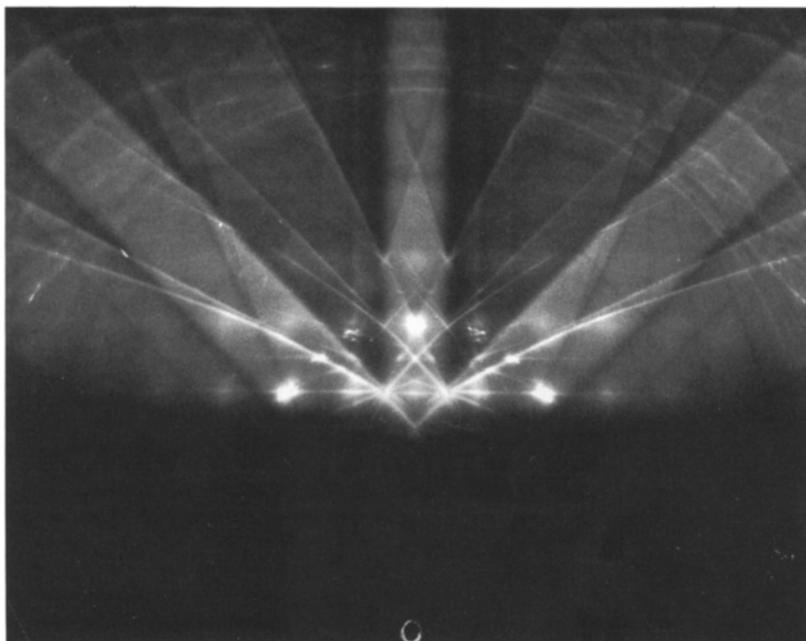


Fig. 19.6. Reflection electron microscopy (REM) image of the 111 surface of a Pt crystal showing steps on the surface, one atom high. The thin horizontal ellipses are foreshortened images of circular, atom-high, hollows or protrusions. A black-white pair of streaks (bottom left) indicates a dislocation which travelled across the crystal leaving a slip-trace step. (courtesy of T. Hsu).

perturbation of the surface which affects the diffracted beam intensity. Very small changes of orientation, as in the strain fields around lattice defects, are clearly visible and surface steps only one atom high give high contrast as seen in Fig. 19.6. Surface monolayers of altered structure, as in the case of surface reconstructions or absorbed layers of foreign atoms, are clearly visible. One limitation of the REM method is that, because of the small angle of grazing incidence, the images are heavily foreshortened and the high resolution is obtained only in the one direction in the surface, perpendicular to the incident beam. Reviews of the principles and applications of REM have been given, for example, by Yagi [1987], Hsu [1992] and Cowley [1995]. Surface imaging in the somewhat analogous scanning mode (SREM) is described by Liu and Cowley [1993].

The theoretical description of RHEED and REM, needed for the quantitative interpretation of the observations, necessarily involves many-beam dynamical diffraction simulations which are more complicated than for the transmission case

since diffracted beams travelling both into and out of the surface must be considered. The original electron diffraction theory of Bethe [1928] can be applied in principle for the case of a perfect crystal cut off by a planar boundary, and was used in few-beam approximations by Colella [1972] for RHEED and by Shuman [1977] for the REM images of strain fields of crystal defects such as emerging dislocations.

The more complete current theoretical treatments may be divided into two types. In one, the crystal potential distribution is divided into thin slices parallel to the surface. Then the propagation of the electron wave into and out of the crystal is considered in terms of its modification from one layer to the next (Maksym and Beeby [1984], Ichimiya [1983], Tong et al. [1988], Peng [1995]). This approach works very well for crystal surfaces which have perfect periodicity in the two dimensions parallel to the surface and relatively small surface unit cell dimensions. As the unit cell dimensions become larger, the number of beams involved and the size of the computation increases rapidly. For non-periodic objects such as disordered surface structures or faults such as surface steps or islands, it is necessary to assume a large unit cell with the assumption of periodic continuation, as discussed in Section 11.5 for the transmission case. For RHEED or REM the dimension of the artificial unit cell in the beam direction must be very large because of the small glancing angle of incidence. After a surface step, for example, the electron wave must travel along the surface for a distance of the order of 500 Å before a new equilibrium wave-function is established. Then the number of sampling points in reciprocal space may appear quite excessive. Even for this case, however, some successful calculations have been made by McCoy and Maksym [1994].

In the alternative approach, the crystal is divided into thin slices perpendicular to the crystal surface. The transmission through these slices, and propagation of the wave along the crystal face is then considered as for the case of transmission, as outlined in Sections 11.1 to 11.5. In this case the structure of each slice is essentially non-periodic since it consists partly of crystal and partly of vacuum. The periodic super-cell assumes a periodic array of crystals interspersed with vacuum spaces and an incident beam which has an amplitude at the entrance plane of zero in the crystal regions and unity in the vacuum spaces, as suggested by Fig. 19.7 (see Cowley and Warburton [1967], Peng and Cowley [1986]). A very large number of beams must be taken to represent the structure of the large super-cell even for a perfect crystal terminated by a planar face, but very little extra complication is introduced by modifications of the structure perpendicular to the surface or by variations of the structure in the beam direction due to surface steps, faults of the bulk crystal or local variations of surface

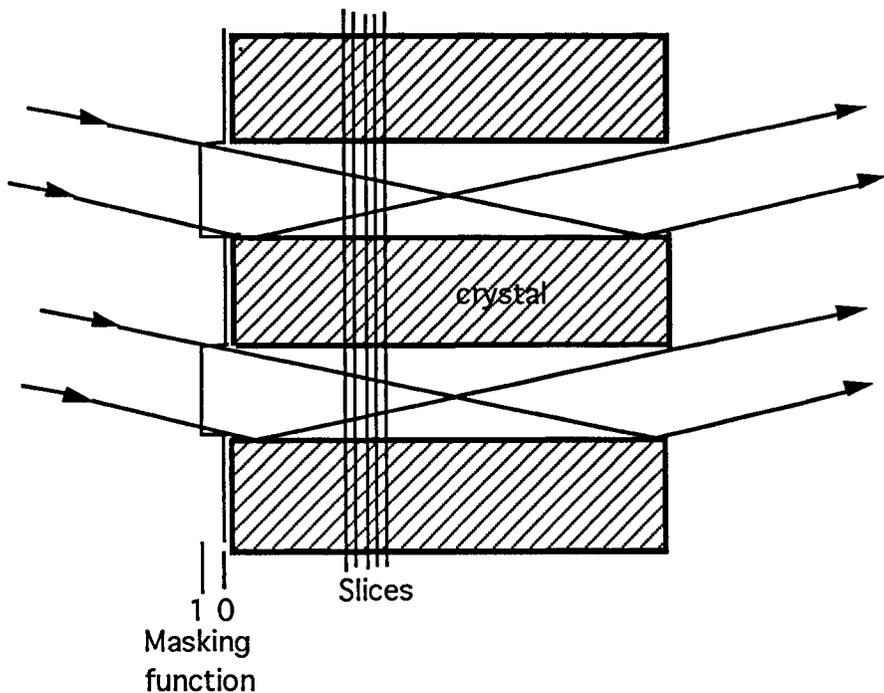


Fig. 19.7. Scheme for calculating RHEED and REM intensities using a multi-slice formulation and the assumption of a large artificial periodicity perpendicular to the crystal face.

composition. This approach has been further refined and applied to the calculation of REM images of surface steps by Ma and Marks [1992] and Anstis and Gan [1994] have made an analogous approach based on the methods outlined in Section 11.6.

The fact that RHEED patterns contain a large proportion of inelastically scattered electrons adds to the difficulty of simulating the experimental observations. The inelastic scattering may, with difficulty, be incorporated in the calculations (Wang and Lu [1988]). As an alternative, it is probably preferable to eliminate the inelastically scattered electrons experimentally by use of an energy filter in the electron microscope and include only their effect on the elastically scattered electrons, introduced by use of appropriate absorption functions.

19.4. Reflection at normal incidence: LEED

In the early work of Davisson and Germer [1927] a beam of electrons of energy less than 100eV was directed perpendicular to the surface of a single crystal and the intensities of electrons back-reflected in various directions were measured as a function of voltage. Subsequent developments of the technique of low-energy electron diffraction (LEED), incorporating the advanced methods for preparing specimens and making observations on specimens in ultra-high vacuum, are well summarized in the reviews by Somorjai [1981] and Van Hove [1981] and the books by Pendry [1974] and Van Hove and Tong [1985].

The effectiveness of the LEED technique for the study of surfaces depends on the fact that for electrons in the energy range of 10 to 1000eV, the intensities of electrons elastically scattered from a surface at normal incidence depend on the structure of only the top few layers of atoms on the surface. The scattering is strongly dynamical in character, but an idea of the essential geometry for the scattering may be obtained from the simple kinematical picture such as that of Fig. 19.8. Because only a few layers of atoms are involved, the reciprocal space configuration for a crystal surface is that of a set of rods perpendicular to the surface and modulated in scattering power by the variations of structure, perpendicular to the surface, in the topmost surface layers. These rods are cut by an Ewald sphere of relatively small radius to give the observed intensities of the back-scattered diffracted beams.

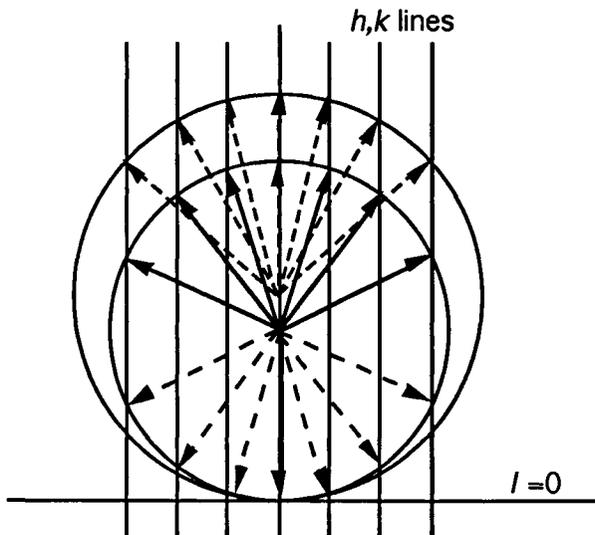


Fig. 19.8. Reciprocal space diagram for the production of LEED spots by intersection of the Ewald sphere with h,k lines, shown for two different wavelengths.

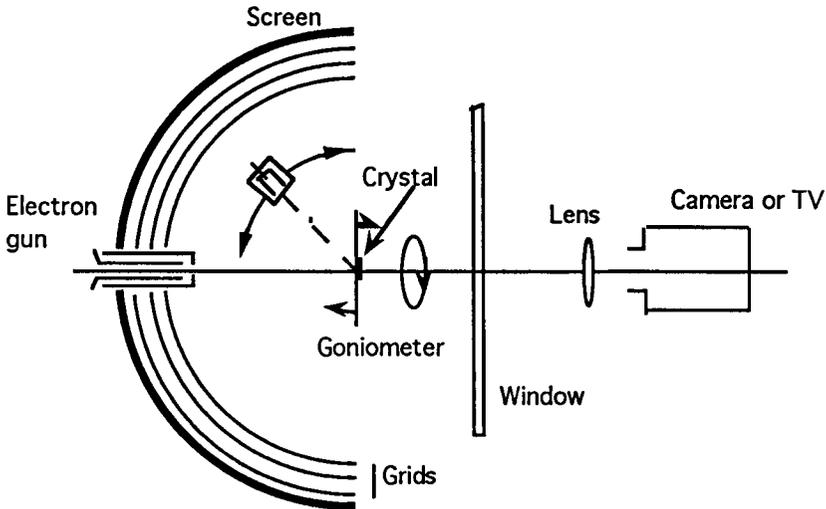


Fig. 19.9. Diagram of a LEED system. The electron beam from the gun is back-scattered by the crystal. Diffracted beams are energy-filtered and accelerated by concentric screens to a fluorescent screen which is photographed through a window.

In the ultra-high vacuum environment of the specimen, it is often more convenient to explore the intensity distributions along the reciprocal space rods by varying the voltage of the incident beam rather than by manipulating the angles of incidence and the detector angles. The patterns of diffraction spots are commonly observed on a hemispherical fluorescent screen, centered on the specimen and observed through a vacuum window in the system (see Fig. 19.9). Concentric fine-mesh hemispherical screens serve to filter out the inelastically scattered electrons and accelerate the diffracted electrons in order to enhance the visibility of the pattern on the fluorescent screen. The measured intensities are normally plotted against incident electron voltage (I - V curves). For sharply-terminated perfect crystals, these curves sometimes show maxima close to the positions corresponding to the reciprocal lattice points of the bulk structure, displaced by the effects of refraction by the mean inner-potential.

After the demonstration by McRae [1966] that multiple-scattering theory was necessary for the interpretation of LEED intensities, new formulations of dynamical diffraction theory, appropriate for LEED, were introduced by Beeby [1968], Pendry [1971, 1974], Jepsen et al. [1972] and others and standard computer programs became available (Van Hove and

Tong [1979]). Good agreement between observed and calculated intensities have become commonplace. The usual approach to calculation of diffracted intensities involves the development of transmission and reflection matrices for individual layers of the structure, parallel to the surface. For the scattering within a layer it is not feasible to apply a simple transmission function, as in the case of high-energy electrons. The scattering from a single atom must be developed in terms of spherical waves and multiple scattering between atoms within a layer must be taken into account.

The application of LEED has provided essential information on a wide range of topics of surface science. The most accurate structure determinations are probably those for pure clean metal crystal surfaces for which, in some cases, the only modification of the bulk structure is a slight contraction or expansion of the bond lengths in the surface layers of atoms. For other metal surfaces, and also, notably, for semiconductor surfaces, a wide variety of surface reconstructions and surface superlattices has been observed and analysed with high accuracy. Adsorbed layers of gases or layers of deposited atoms, ranging from small fractions of a monolayer to many atoms in thickness, have been analysed. Layers of adsorbed molecules and their transitions and reactions have been studied with consequent inferences for the understanding of the mechanisms of catalytic reactions. A summary of the results obtained by LEED in the various areas of surface crystallography is given by MacLaren et al. [1987].

19.5. Diffraction of emitted electrons

In recent years several techniques have been developed for the study of the structure of surface layers by the use of low energy electrons which originate within the layers themselves. These may be electrons emitted from the atoms, as in photoemission, or electrons derived from incident electron beams but made incoherent with the incident electrons through inelastic scattering processes or by diffuse scattering. To some extent, the methods may be considered analogous to the observation of Kossel patterns for X-rays generated in a crystal or of Kikuchi patterns for high-energy electrons, but the importance for surface studies arises because the emitted electrons must be considered as scattered by the atoms surrounding their source rather than as being diffracted by the crystal lattice planes. Measurements of the angular and energy distributions of the emitted electrons can be used to derive the average configurations of the atoms surrounding the source points.

For the interpretation of these diffraction experiments it is often effective to make use of the concept of a reference wave scattered from some particular type of atom, in much the same way as heavy atoms or atoms giving anomalous

scattering are used as reference atoms in the process of X-ray structure analysis (Section 6.4). The analogy with Gabor's concept of using a reference wave in imaging as a basis for holography has led to the unfortunate practice of referring to these methods of electron diffraction as "holography", implying that they provide images of the local arrangements of atoms. It should be emphasized that these are purely diffraction techniques and lead to an imaging of atom positions only in the same sense that X-ray diffraction does. The "images" formed by computation from the observed diffraction intensities represent the average environments of a very large number of reference points. It has been suggested (Cowley [1993]) that the alternative designation of "holographic diffraction" may be more acceptable.

In the case of photo-emitted electrons, generated by incident X-rays or ultraviolet radiation, the electron wave emitted by the excited atoms may be regarded as a reference wave. This interferes with the electron waves scattered from the atoms in the immediate neighborhood of the source atom. If the electrons detected are those inelastically scattered from an incident electron beam the technique is known as Kikuchi electron diffraction (KED), by analogy with the high-energy electron case, and the electrons involved are selected by energy analysis.

The electron scattering from a partial monolayer of adsorbed atoms on the surface may give a diffuse LEED pattern (DLEED) when the adsorbed atoms are distributed in a disordered way among the many equivalent sites on the supporting atom layer. The scattering may be described as a two-dimensional analogue of the diffuse scattering of X-rays or electrons from a three-dimensionally disordered occupancy of specific sites in a crystal lattice by impurity or substitutional atoms (Chapter 17). Because the scattering of low-energy electrons is strongly dynamical, the wave from a single adsorbed atom can not be described by a simple kinematical scattering factor. Instead one may consider the difference, $\Delta\Psi(\mathbf{r})$, between the wave-function generated by multiple scattering by the adsorbed atom and all its neighbors and the wave function for the same site with no adsorbed atom. Then the diffuse scattering comes from the total wave function as $\Delta\Psi(\mathbf{r}) * \sum_i \sigma_i \delta(\mathbf{r} - \mathbf{r}_i)$ for all the sites \mathbf{r}_i where σ_i is an occupation operator similar to that defined in Section 17.2. Then the diffuse scattering intensity is given by Fourier transform as $|\Delta\Psi(\mathbf{u})|^2 \cdot N \sum_i \langle \sigma_0 \sigma_i \rangle \exp\{2\pi i \mathbf{u} \cdot \mathbf{r}_i\}$. The summation term gives the distribution of diffuse scattering periodic in reciprocal space with the periodicity of the substrate structure and the square of $\Delta\Psi$ gives the over-all modulation of this diffuse scattering intensity which depends on the environment of the reference atom (c.f. Section 17.7) and so can be separately interpreted.

Attempts to interpret the distributions of scattered intensity by direct Fourier transform showed limited success, partly because of the strong deviations from kinematical scattering and partly because of the ambiguities resulting because the Fourier transform of the intensity distribution gives the auto-correlation (Patterson) function of the wave function. Effective methods have been devised, however, for making use of the other available variable, the electron energy, to overcome these difficulties sufficiently to allow derivation of the relative average positions of atoms surrounding the reference atoms with a high degree of confidence. The techniques and representative results are well described in the group of articles by Tong et al. [1994], Heinz and Wedler [1994] and Wei [1994].

This Page Intentionally Left Blank

REFERENCES

- Allpress, J.G., Elizabeth Hewat, A.F. Moodie and J.V. Sanders, 1972, *Acta Cryst.* A28, 528. [Ch. 13].
- Amelinckx, S. and D. Van Dyck, 1993, in *Electron Diffraction Techniques, Vol. 2*, Ed. J.M. Cowley, (Oxford Univ. Press, Oxford). [Ch. 13,17,18].
- Anderson, J.S., 1969, *Bull. Soc. Chim. France*, 2203. [Ch. 17].
- Andersson, B., J.K. Gjønnes and J. Taftø, 1974, *Acta Cryst.* A30, 216. [Ch. 12].
- Ando, Y., A. Ichimiya and R. Uyeda, 1974, *Acta Cryst.* A30, 600. [Ch. 15].
- Anstis, G.R. and M.A. O'Keefe, 1976, In: *Proc. 34th annual meet. electron microscopy*, Ed. G.W. Bailey (Claitors, Baton Rouge) p. 480. [Ch. 13].
- Anstis, G.R. and D.J.H. Cockayne, 1979, *Acta Cryst.* A35, 511. [Ch. 11,18].
- Anstis, G.R., 1989, in *Electron Microscope Diffraction and Images*, Eds. W. Krakow and M.A. O'Keefe, (Minerals, Metals and Materials Soc., Warrendale, PA). [Ch. 11].
- Anstis, G.R. and X.S. Gan, 1994, *Surface Sci.*, 314, L919. [Ch. 11,19].
- Arsac, J., 1966, *Fourier transforms and the theory of distributions* (Prentice Hall, Englewood-Cliffs, N.J.). [Ch. 2].
- Authier, A., 1970, in: *Modern diffraction and imaging techniques in materials science*, Eds. S. Amelinckx et al. (North-Holland, Amsterdam). [Ch. 15].
- Authier, A., 1970, in: *Advances in structural research by diffraction methods*, Eds. R. Brill and R. Mason, Vol. 3 (Pergamon, Oxford) p. 1. [Ch. 8].
- Avilov, A.S., R.M. Imamov, R.K. Karakhanyan and Z.G. Pinsker, 1973, *Kristallografiya* 18, 49. [Ch. 6].
- Bachrach, R.E., Ed., 1992, *Synchrotron Radiation Research; Advances in Surface and Interface Science*, (Plenum Press, NY). [Ch. 4].
- Bacon, G.E., 1975, *Neutron diffraction*, 3rd edition (Clarendon Press, Oxford). [Ch. 4,6,8,12].
- Balter, S., R. Feldman and B. Post, 1971, *Phys. Rev. Letters* 27, 307. [Ch. 14].
- Bardhan, P. and J.B. Cohen, 1976, *Acta Cryst.* A32, 597. [Ch. 17].
- Bartell, L.S., 1975, *J. Chem. Phys.* 63, 3750. [Ch. 4].
- Batterman, B.W., 1957, *J. Appl. Phys.* 28, 556. [Ch. 17].
- Batterman, B.W., 1962, *Appl. Phys. Letters* 1, 68. [Ch. 9].
- Batterman, B.W., 1964, *Phys. Rev.* A133, 759. [Ch. 9,14,19].
- Batterman, B.W. and H. Cole, 1964, *Rev. Mod. Phys.* 36, 681. [Ch. 8,9].
- Batterman, B.W., 1969, *Phys. Rev. Letters* 22, 703. [Ch. 14,19].
- Beauvillain, J., 1970, *J. de Microscopie* 9, 455. [Ch. 14].
- Becker, P.J. and P. Coppens, 1974, *Acta Cryst.* A30, 129. [Ch. 16].
- Bedzyk, M.J., 1988, *Nucl. Inst. Meth.* A266, 679. [Ch. 19].

- Bedzyk, M.J., G.M. Bommarito, J.S. Shildkraut, 1989, *Phys. Rev. Letters*, 62, 1376. [Ch. 19].
- Bedzyk, M.J., G.M. Bommarito, M. Caffrey and T.L. Penner, 1990, *Science*, 248, 52. [Ch. 19].
- Beeby, J.L., 1968, *J. Phys.* C1, 82. [Ch. 19].
- Bell, W., 1971, *Proc. 29th Annual EMSA Meeting*, p. 184. [Ch. 15].
- Benedek, R. and P.S. Ho, 1973, *J. Phys. F: Metal Physics* 3, 1285. [Ch. 12].
- Berry, M.V., B.F. Buxton and A.M. Ozorio de Almeda, 1973, *Radiat. EH.* 20, 1. [Ch. 9].
- Berry, M.V. and K.E. Mount, 1972, *Rep. Progr. Phys.* 35, 315. [Ch. 14].
- Bethe, H.A., 1928, *Ann. Physik* 87, 55. [Ch. 1,8,16,19].
- Binnig, G., H. Rohrer, Ch. Gerber and E. Weibel, 1982, *Phys. Rev. Letters* 49, 57. [Ch. 19].
- Biscoe, J. and B.E. Warren, 1942, *J. Appl. Phys.* 13, 364. [Ch. 7].
- Bish, D.L. and J.E. Post, Eds., 1989, *Reviews in Mineralogy, Vol. 20. Modern Powder Diffraction.* (Mineralogical Soc. America). [Ch. 16].
- Blackman, M., 1939, *Proc. Roy. Soc. Lond.* A173, 68. [Ch. 16].
- Bless, P.W., R.B. vonDreele, E. Kostiner and R.E. Hughes, 1972, *J. Solid State Chem.* 4, 262. [Ch. 18].
- Bonse, U. and M. Hart, 1965, *Appl. Phys. Letters* 6, 155. [Ch. 15].
- Bonse, U. and M. Hart, 1970, *Physics Today*, Aug., p. 26. [Ch. 15].
- Booker, G.R., A.M.B. Shaw, M.J. Whelan and P.B. Hirsch, 1967, *Phil. Mag.* 16, 1185. [Ch. 14].
- Booker, G.R., 1970, in: *Modern diffraction and imaging techniques in materials science*, Eds. S. Amelinckx et al. (North-Holland, Amsterdam). [Ch. 14].
- Booker, G.R., D.C. Joy, J.P. Spencer, H. Graf von Harrach and M.N. Thompson, 1974, in: *Scanning electron microscopy, 1974*, *Proc. 7th annual scanning electron microscopy symposium*, Eds. Om Johari and Irene Corvin (I.I.T. Research Institute, Chicago) p. 225. [Ch. 18].
- Boothroyd, C.B. and W.M. Stobbs, 1988, *Ultramicros.*, 26, 361. [Ch. 13].
- Borie, B., 1957, *Acta Cryst.* 10, 89. [Ch. 12,17].
- Borie, B., 1959, *Acta Cryst.* 12, 280. [Ch. 12].
- Borie, B., 1961, *Acta Cryst.* 14, 472. [Ch. 12].
- Borie, B., 1966, *Acta Cryst.* 21, 470. [Ch. 8].
- Borie, B., 1970, *Acta Cryst.* A26, 533. [Ch. 12].
- Borie, B. and C.J. Sparks, 1971, *Acta Cryst.* A27, 198. [Ch. 17].
- Born, M. and R.D. Misra, 1940, *Proc. Camb. Phil. Soc.* 36, 466. [Ch. 12].
- Born, M. and E. Wolf, 1989, *Principles of optics*, 6th edition (Pergamon Press, London). [Ch. 1,3].
- Borrmann, G., 1936, *Ann. d. Phys.* 27, 669. [Ch. 14].
- Borrmann, G., 1941, *Phys. Zeit.* 43, 157. [Ch. 9].
- Borrmann, G., 1950, *Zeit. f. Phys.* 127, 297. [Ch. 9].
- Borrmann, G., 1964, *Z. f. Kristallogr.* 120, 143. [Ch. 14,18].
- Borrmann, G. and W. Hartwig, 1965, *Z. f. Kristallogr.* 121, 6 and 401. [Ch. 10,14].

- Bourret, A., A. Renault and G.R. Anstis, 1978-1977, *Chem. Scripta* 14, 207. [Ch. 18].
- Bourret, A., J. Desseaux and A. Renault, 1982, *Phil. Mag.* A45, 1. [Ch. 18].
- Bracewell, R.N., 1986, *The Fourier Transform and its Applications, 2nd Edit., Revised* (McGraw Hill, New York). [Ch. 2].
- Bradley, S.A. and W.E. King, Eds., 1993, *Ultramicros.* 51, 1-347. [Ch. 18].
- Brout, R., 1965, *Phase transitions* (Benjamin, New York). [Ch. 17].
- Buerger, M.J., 1959, *Vector space* (Wiley, New York). [Ch. 6].
- Bunyan, P., 1963, *Proc. Phys. Soc.* 82, 1051. [Ch. 4].
- Buseck, P.R. and J.M. Cowley, 1983, *Amer. Min.* 68, 18. [Ch. 17].
- Buseck, P.R., J.M. Cowley and L. Eyring, 1988, *High Resolution Transmission Electron Microscopy and Associated Techniques* (Oxford Univ. Press, Oxford). [Ch. 13].
- Buseck, P.R. and D.R. Veblen, 1988, in *High Resolution Transmission Electron Microscopy and Associated Techniques*, Eds. P.R. Buseck, J.M. Cowley and L. Eyring, (Oxford Univ. Press, Oxford). [Ch. 17].
- Buxton, B.F., J.A. Eades, J.W. Steeds and G.M. Rackham, 1976, *Phil. Trans. Roy. Soc. London* 281, 171. [Ch. 15].
- Buxton, B.F., J.E. Loveluck and J.W. Steeds, 1978, *Phil. Mag.* A38, 259. [Ch. 17].
- Cairns, J.A. and R.S. Nelsen, 1968, *Phys. Letters* 27A, 15. [Ch. 14].
- Castaing, R. and A. Guinier, 1951, *C.R. Acad. Sci., Paris* 232, 1948. [Ch. 14].
- Castaing, R., 1966, in: *Electron microscopy 1966*, Sixth Internat. Conf. on Electron Microscopy, Kyoto, Vol. 1 (Maruzen Co., Tokyo) p. 63. [Ch. 12].
- Chadderton, L.T., 1970, *J. Appl. Cryst.* 3, 429. [Ch. 14].
- Chadderton, L.T., 1973, in: *Channelling: theory, observation and application*, Ed. D.V. Morgan (Wiley, London) p. 287. [Ch. 14].
- Chen, H., R.J. Comstock and J.B. Cohen, 1979, *Ann. Rev. Mater. Sci.* 9, 51. [Ch. 17].
- Cherns, D., 1974, *Phil. Mag.* 30, 549. [Ch. 19].
- Chevalier, J.-P. and M.J. Hytch, 1993, *Ultramicros.* 52, 253 [Ch. 17].
- Chipman, D.R. and C.B. Walker, 1972, *Phys. Rev.* B5, 3823. [Ch. 17].
- Clapp, P.C. and S.C. Moss, 1968, *Phys. Rev.* 171, 754. [Ch. 17].
- Clarke, G.L. and W. Duane, 1922, *Proc. Nat. Acad. Sci.* 8, 90. [Ch. 14].
- Coates, D.G., 1967, *Phil. Mag.* 16, 1179. [Ch. 14].
- Cockayne, D.J.H., P. Goodman, J.C. Mills and A.F. Moodie, 1967, *Rev. Sci. Inst.* 38, 1093. [Ch. 9,18].
- Cockayne, D.J.H., I.L.F. Ray and M.J. Whelan, 1969, *Phil. Mag.* 20, 1265. [Ch. 18].
- Cohen, J.B., 1968, *Recent developments concerning the order-disorder transformation, 1968 Seminar on phase transformations*, Detroit, Mich. [Ch. 17].

- Cohen, M.L., 1981, *Structure and Bonding in Crystals*, (Academic Press, New York). [Ch. 15].
- Colella, R., 1972, *Acta Cryst.* A28, 11. [Ch. 8,10,19].
- Colella, R. and J.F. Menadue, 1972, *Acta Cryst.* A28, 16. [Ch. 8,15].
- Colella, R., 1974, *Acta Cryst.* A30, 413. [Ch. 15].
- Cooley, J.W. and J.W. Tukey, 1965, *Math. Computation*, 19, 297. [Ch. 11].
- Cooper, M.J. and K.D. Rouse, 1970, *Acta Cryst.* A26, 213. [Ch. 16].
- Coppens, P., 1992, *Synchrotron Radiation Crystallography*, (Academic Press, London, San Diego). [Ch. 4].
- Cowan, P.L., M.F. Golovchenko and M.F. Robbins, 1980, *Phys. Rev. Letters* 44, 1680. [Ch. 19].
- Cowley, J.M. and A.L.G. Rees, 1946, *Nature* 158, 550. [Ch. 9].
- Cowley, J.M., A.L.G. Rees, 1947, *Proc. Phys. Soc.* 59, 283. [Ch. 9].
- Cowley, J.M., 1950a, *Phys. Rev.* 77, 669. [Ch. 17].
- Cowley, J.M., 1950b, *J. Appl. Phys.* 21, 24. [Ch. 17].
- Cowley, J.M., A.L.G. Rees and J.A. Spink, 1951, *Proc. Phys. Soc.* 64, 609. [Ch. 16].
- Cowley, J.M., 1954, *J. Electrochem. Soc.* 101, 277. [Ch. 18].
- Cowley, J.M. and A.F. Moodie, 1957, *Acta Cryst.* 10, 609. [Ch. 8,11].
- Cowley, J.M. and A.F. Moodie, 1957a, *Proc. Phys. Soc.* B70, 486. [Ch. 1].
- Cowley, J.M. and A.F. Moodie, 1957b, *Proc. Phys. Soc.* B70, 497. [Ch. 1].
- Cowley, J.M. and A.F. Moodie, 1957c, *Proc. Phys. Soc.* B70, 505. [Ch. 1].
- Cowley, J.M., P. Goodman & A.L.G. Rees, 1957, *Acta Cryst.* 10, 19. [Ch. 9].
- Cowley, J.M. and A.F. Moodie, 1958, *Proc. Phys. Soc.* 71, 533. [Ch. 11].
- Cowley, J.M., A.F. Moodie, 1960, *Proc. Phys. Soc.* 76, 378. [Ch. 1].
- Cowley, J.M. and A.F. Moodie, 1962, *J. Phys. Soc. Japan.* 17, B-II, 86. [Ch. 11].
- Cowley, J.M., 1964, *Acta Cryst.* 17, 33. [Ch. 14].
- Cowley, J.M., 1965a, *Phys. Rev.* 138, A1384. [Ch. 17].
- Cowley, J.M., 1965b, *Proc. Int. Conf. on Electron diffraction and the nature of defects in crystals*, Melbourne (Australian Acad. of Science) Paper J-5. [Ch. 12,17].
- Cowley, J.M., 1967, *Crystal Structure Determination by Electron Diffraction*, in *Progress in Materials Science*, Vol. 13, No. 6. (Pergamon Press, Oxford). [Ch. 6].
- Cowley, J.M. and P.M. Warburton, 1967, in *The Structure and Chemistry of Solid Surfaces*, Ed. G. Somorjai, (Wiley and Sons, New York). [Ch. 19].
- Cowley, J.M., 1968, *Phys. Letters* 26A, 623. [Ch. 14].
- Cowley, J.M. and R.J. Murray, 1968, *Acta Cryst.* A24, 329. [Ch. 12,17].
- Cowley, J.M. and A.P. Pogany, 1968, *Acta Cryst.* A24, 109. [Ch. 12,17].

- Cowley, J.M., 1969, Appl. Phys. Letters 15, 58. [Ch. 13].
- Cowley, J.M., 1969, Z. f. Angew. Phys. 27, 149. [Ch. 16].
- Cowley, J.M., 1969, Appl. Phys. Letters 15, 58. [Ch. 18].
- Cowley, J.M., 1970, J. Appl. Cryst. 3, 49. [Ch. 18].
- Cowley, J.M., 1971, in: *Advances in high-temperature chemistry*, Vol. 3 (Academic Press, New York). [Ch. 17].
- Cowley, J.M. and S.W. Wilkins, 1972, in: *Interatomic potentials and simulation of lattice defects*, Eds. P.C. Gehlen, J.R. Beeler and R.I. Jaffee (Plenum Press, NY) p. 265. [Ch. 17].
- Cowley, J.M., 1973, Acta Cryst. A29, 537. [Ch. 17].
- Cowley, J.M., 1976a, Acta Cryst. A32, 83. [Ch. 18].
- Cowley, J.M., 1976b, Acta Cryst. A32, 88. [Ch. 18].
- Cowley, J.M., 1976, Ultramicros. 2, 3. [Ch. 13].
- Cowley, J.M. and A.Y. Au, 1978, Acta Cryst A34, 738. [Ch. 13,18].
- Cowley, J.M., J.B. Cohen, M.B. Salamon and B.J. Wuensch, Eds., 1979, *Modulated structures-1979 (Kailua Kona, Hawaii)* (American Inst. Physics, New York). [Ch. 17].
- Cowley, J.M. and P.M. Fields, 1979, Acta Cryst A35, 28. [Ch. 12].
- Cowley, J.M., 1981a, Soviet Physics, Crystallog. 26, 549. [Ch. 15].
- Cowley, J.M., 1981b, Ultramicros. 7, 19. [Ch. 15].
- Cowley, J.M., M. Osman and P. Humble, 1984, Ultramicros. 15, 311. [Ch. 13,15].
- Cowley, J.M., 1988, in *High Resolution Electron Microscopy and Associated Techniques*, Eds., P.R. Buseck, J.M. Cowley and L. Eyring (Oxford Univ. Press, Oxford). [Ch. 13].
- Cowley, J.M., 1991, in *Physical Methods of Chemistry*, Vol. 4, Microscopy, Eds. B.W. Rossiter and J.F. Hamilton, (John Wiley and Sons, New York). [Ch. 13].
- Cowley, J.M., 1992, Ultramicros. 41, 335. [Ch. 13].
- Cowley, J.M., 1992, in *Electron Diffraction Techniques*, Vol. 1, Ed. J.M. Cowley, (Oxford Univ. Press, Oxford). [Ch. 1,9,15].
- Cowley, J.M., 1992, in *International Tables for Crystallography*, Vol. C, Ed. A.J.C. Wilson, (Kluwer Academic Publ., Dordrecht). [Ch. 16].
- Cowley, J.M., 1993, Ultramicros. 49, 4. [Ch. 13].
- Cowley, J.M., 1994, Surface Sci. 298, 336. [Ch. 19].
- Cowley, J.M., 1995 in *Handbook of Microscopy*, Vol. I, Eds. S. Amelinckx, D. Van Dyck, J.F. Van Landuyt and G. Van Tendeloo, (VCH Verlag., Weinheim). [Ch. 19].
- Craven, A.J., J.M. Gibson, A. Howie and D.R. Spalding, 1978, Philos. Mag. A38, 519. [Ch. 12].
- Craven, B.M., 1987, *Symposium of Neutron Diffraction; Univ. Texas, Austin*, (Amer. Inst. Physics, New York). [Ch. 4,6].
- Crewe, A.V. and J. Wall, 1970, Optik. 30, 461. [Ch. 13].
- Crewe, A.V., 1980, Optik 60, 271. [Ch. 13].
- Cundy, S.L., A.J.F. Metherell and M.J. Whelan, 1966, in: *Electron Microscopy 1966*, Sixth Internat. Conf. on Electron Microscopy, Kyoto, Vol. 1 (Maruzen Co., Tokyo) p. 87. [Ch. 12].
- Cundy, S.L., A. Howie and U. Valdre, 1969, Phil. Mag. 20, 147. [Ch. 12].

- Darwin, C.G., 1914, *Phil. Mag.* 27, 315, 675. [Ch. 1,8,10,16].
- Datz, S., C. Erginsoy, G. Liefried and H.O. Lutz, 1967, *Ann. Rev. Nucl. Sci.* 17, 129. [Ch. 14].
- Davison, C.J. and L.H. Germer, 1927, *Nature*, 119, 558. [Ch. 19].
- Dawson, B., 1967, *Proc. Roy. Soc. A298*, 255, 379. [Ch. 15].
- Dawson, B., P. Goodman, A.W.S. Johnson, D.F. Lynch and A.F. Moodie, 1974, *Acta Cryst.* A30, 297. [Ch. 4].
- De Marco, J.J. and P. Suortti, 1971, *Phys. Rev.* B4, 1028. [Ch. 12].
- de Rosier, D.J. and A. Klug, 1968, *Nature* 217, 130. [Ch. 13].
- Ditchburn, R.W., 1976, *Light*, Vol. 1, 3rd edition (Blackie & Sons, London). [Ch. 1].
- Dorset, D.L., 1994, *Advances Electronics Electron Phys.* 88, 111. [Ch. 6].
- Dorset, D., Ed., 1995, *Structural Electron Crystallography*, (Plenum Press, New York). [Ch. 6,13].
- Doyle, P.A. and P.S. Turner, 1968, *Acta Cryst.* A24, 390. [Ch. 4].
- Doyle, P.A., 1969, Ph.D. Thesis, University of Melbourne. [Ch. 4].
- Doyle, P.A., 1969, *Acta Cryst.* A25, 569. [Ch. 12,14].
- Doyle, P.A., 1970, *Acta Cryst.* A26, 133. [Ch. 12].
- Doyle, P.A., 1971, *Acta Cryst.* A27, 109. [Ch. 12,14].
- Duffieux, P.M., 1946, *L'integral de Fourier et ses applications à l'optique* (Privately printed, Besançon). [Ch. 1,3].
- Duncumb, P., 1962, *Phil. Mag.* 7, 2101. [Ch. 9].
- Dupuoy, G. and J. Beauvillain, 1970, in: *Microscopic electronique*, 1970, Vol. II, 207. [Ch. 14].
- Eades, J.A., 1992, in *Electron Diffraction Techniques*, Vol. 1, Ed. J.M. Cowley, (Oxford Univ. Press, Oxford). [Ch. 15].
- Ehrhart, P., 1978, *J. Nucl. Mater.* 69-70, 200. [Ch. 12].
- Erdeyli, A., 1954, *Tables of integral transforms*, Vol. 1, Bateman Mathematical Project (McGraw-Hill, New York). [Ch. 2].
- Emslie, A.G., 1934, *Phys. Rev.* 45, 43. [Ch. 19].
- Ewald, P.P., 1916, *Ann. Physik.* 49, 1, 117. [Ch. 1,8].
- Ewald, P.P., 1917, *Ann. Physik.* 54, 519. [Ch. 8].
- Ewald, P.P. and Y. Heno, 1968, *Acta Cryst.* A24, 5. [Ch. 10,14,15].
- Fejes, P., 1973, Ph.D. Thesis, Arizona State University. [Ch. 11,13].
- Fejes, P.L., Sumio Iijima and J.M. Cowley, 1973, *Acta Cryst.* A29, 710. [Ch. 10,15].
- Fellgett, P.B. and E.H. Linfoot, 1955, *Phil. Trans. Roy. Soc.* A247, 369. [Ch. 3].
- Ferrell, R.A., 1956, *Phys. Rev.* 101, 554. [Ch. 12].
- Fields, P.M. and J.M. Cowley, 1978, *Acta Cryst.* A34, 103. [Ch. 12,17].
- Fisher, P.M.J., 1965, *Proc. Int. Conf. on Electron diffraction and the nature of defects in crystals* (Australian Acad. of Science, Melbourne) paper IH-4. [Ch. 12].
- Fisher, P.M.J., 1968, *Jap. J. Appl. Phys.* 7, 191. [Ch. 10].
- Fisher, P.M.J., 1969, Ph.D. Thesis, Univ. of Melbourne. [Ch. 15].

- Fisher, P.M.J., 1972, private communication. [Ch. 15,17].
- Fisher, R.M., J.S. Lally, C.J. Humphreys and A.J.E. Metherell, in: *Microscopie Electronique*, 1970, 1, 107. [Ch. 15].
- Fitzgerald, J.D. and A.W.S. Johnson, 1984, *Ultramicros.* 12, 231. [Ch. 15].
- Flinn, P.A., 1956, *Phys. Rev.* 104, 350. [Ch. 12,17].
- Flocken, J.W. and J.R. Hardy, 1970, *Phys. Rev. B1*, 2447. [Ch. 12].
- Fox, A.G. and R.M. Fisher, 1988, *Phil. Mag.* B57, 197. [Ch. 15].
- Fryer, J.R. and C.J. Gilmore, 1992, *Trans. Amer. Crystallogr. Assn.* 28, 57. [Ch. 13].
- Fu, Z.G., D.X. Huang, F.H. Li, J.Q. Li, Z.X. Zhao, T.Z. Cheng and H. Fan, 1994, *Ultramicros.* 54, 229. [Ch. 13].
- Fues, E., 1949, *Z. f. Phys.* 125, 531. [Ch. 14].
- Fujimoto, F., 1959, *J. Phys. Soc. Japan* 14, 1558. [Ch. 8,10,11].
- Fujimoto, F. and Y. Kainuma, 1963, *J. Phys. Soc. Japan* 18, 1792. [Ch. 12,14].
- Fujimoto, F. and A. Howie, 1966, *Phil. Mag.* 13, 1131. [Ch. 12].
- Fujimoto, F. and K. Komaki, 1968, *J. Phys. Soc. Japan* 25, 1679. [Ch. 12].
- Fujimoto, F., 1977, in: *High voltage electron microscopy 1977*, Eds. T. Imura and H. Hashimoto (Japanese Soc. Electron Microscopy, Tokyo) p. 271. [Ch. 14].
- Fujiwara, K., 1957, *J. Phys. Soc. Japan* 12, 7. [Ch. 17].
- Fujiwara, K., 1959, *J. Phys. Soc. Japan* 14, 1513. [Ch. 8,10,11].
- Fujiwara, K., 1961, *J. Phys. Soc. Japan* 16, 2226. [Ch. 8].
- Fukuhara, A., 1966, *J. Phys. Soc. Japan* 21, 2645. [Ch. 10].
- Gabor, D., 1948, *Nature*, 161, 777. [Ch. 3,13].
- Gabor, D., 1949, *Proc. Roy. Soc., London*, A197, 454. [Ch. 3,13].
- Gajdardziska-Josifovska, M. and J.M. Cowley, 1991, *Acta Cryst.* A47, 74. [Ch. 19].
- Gaudig, W. and S.L. Sass, 1979, *Philos. Mag.* A39, 725. [Ch. 18].
- Gehlen, P.C. and J.B. Cohen, 1965, *Phys. Rev.* 139, A844. [Ch. 17].
- Geogopolous, P. and J.B. Cohen, 1977, *J. de Physique C-7*, Suppl. 12, 191. [Ch. 17].
- Giacovazzo, C., 1993, in *International Tables for Crystallography*, Vol. B, Ed. U. Shmueli. (Kluwer Academic Publ., Dordrecht). [Ch. 6].
- Gjønnnes, J.K., 1962, *Acta Cryst.* 15, 703. [Ch. 8,16].
- Gjønnnes, J.K., 1962, *J. Phys. Soc. Japan* 17, Suppl. BII, 137. [Ch. 17].
- Gjønnnes, J.K., 1964, *Acta Cryst.* 17, 1075. [Ch. 4].
- Gjønnnes, J.K., 1965, *Proc. Int. Conf. on Electron diffraction and the nature of defects in crystals*, Melbourne (Australian Acad. of Science) paper IH-2. [Ch. 12].
- Gjønnnes, J.K. and A.F. Moodie, 1965, *Acta Cryst.* 19, 65. [Ch. 15].
- Gjønnnes, J.K., 1966, *Acta Cryst.* 20, 240. [Ch. 10,12,14].
- Gjønnnes, J.K. and D. Watanabe, 1966, *Acta Cryst.* 21, 297. [Ch. 12].
- Gjønnnes, J.K. and R. Høier, 1971, *Acta Cryst.* A27, 166. [Ch. 12].
- Gjønnnes, J.K. and R. Høier, 1971, *Acta Cryst.* A27, 313. [Ch. 15].

- Gjønnnes, J.K., 1993 in *Electron Diffraction Techniques*, Vol. 2, Ed. J.M. Cowley (Oxford Univ. Press, Oxford). [Ch. 12].
- Glauber, R. and V. Schomaker, 1953, *Phys. Rev.* 89, 667. [Ch. 4].
- Glossop, A.B. and D.W. Pashley, 1959, *Proc. Roy. Soc. A250*, 132 [Ch. 17].
- Goodman, J.W., 1968, *Introduction to fourier optics* (McGraw Hill, New York). [Ch. 1].
- Goodman, P., G. Lehmpfuhl, 1964, *Z. Naturforsch.* 19a, 818. [Ch. 9].
- Goodman, P. and A.F. Moodie, 1965, in: *Proc. Int. Conf. on Electron diffraction and the nature of defects in crystals*, Melbourne (Australian Acad. of Science). [Ch. 11].
- Goodman, P. and G. Lehmpfuhl, 1967, *Acta Cryst.* 22, 14. [Ch. 15].
- Goodman, P., 1968, *Acta Cryst.* A24, 400. [Ch. 15].
- Goodman, P. and A.F. Moodie, 1974, *Acta Cryst.* A30, 280. [Ch. 8,11,13,15].
- Goodman, P., A.W.S. Johnson, 1977, *Acta Cryst.* A33, 997. [Ch. 15].
- Goodman, P., 1993, in *International Tables for Crystallography, Vol.B*, Ed. U. Shmueli, (Kluwer Academic Publ., Dordrecht). [Ch. 15].
- Gragg, J.E., 1970, Ph.D. Thesis, Northwestern Univ. [Ch. 17].
- Gribelyuk, M.A. and J.M. Cowley, 1992, *Ultramicros.*, 45, 103. [Ch. 13].
- Gribelyuk, M.A. and J.M. Cowley, 1993, *Ultramicros.*, 50, 29. [Ch. 13].
- Grinton, G.R. and J.M. Cowley, 1971, *Optik.* 34, 221. [Ch. 11,13].
- Guan, D.Y. and S.L. Sass, 1973, *Philos. Mag.* 26, 1211. [Ch. 18].
- Guinier, A., 1963, *X-ray diffraction in crystals, imperfect crystals and amorphous bodies* (W.H. Freeman and Co., San Francisco and London). [Ch. 5,18].
- Hahn, T., Ed., 1992, *International Tables for Crystallography, Vol. A; Space Group Symmetry*, (Kluwer Academic Publ., Dordrecht). [Ch. 6].
- Haine, M.E. and T. Mulvey, 1952, *J. Optical Soc. Amer.* 42, 763. [Ch. 13].
- Hall, C.R., 1965, *Phil. Mag.* 12, 815. [Ch. 12].
- Hall, C.R. and P.B. Hirsch, 1965, *Proc. Roy. Soc. A286*, 158. [Ch. 12].
- Hall, C.R., P.B. Hirsch and G.R. Booker, 1966, *Phil. Mag.* 14, 979. [Ch. 12].
- Hall, C.R. and P.B. Hirsch, 1968, *Phil. Mag.* 18, 115. [Ch. 9].
- Harada, J., 1993, *Ultramicros.*, 52, 233. [Ch. 19].
- Harburn, G., C.A. Taylor and T.R. Welberry, 1975, *Atlas of optical transforms* (Cornell Univ. Press, New York). [Ch. 2].
- Hargittai, I., 1992, *Gas Electron Diffraction*, in *Electron Diffraction Techniques, Vol. 1*, Ed. J. M. Cowley, (Oxford Univ. Press, Oxford). [Ch. 5].
- Hargittai, I. and M. Hargittai, Eds., 1988, *Stereochemical Applications of Gas-Phase Electron Diffraction. Part A; The Electron Diffraction Technique. Part B; Structural Information from Selected Classes of Compounds.* (VCH, New York). [Ch. 5].

- Harker, D. and J.S. Kasper, 1948, *Acta Cryst.* 1, 70. [Ch. 6].
- Hart, M. and A.D. Milne, 1970, *Acta Cryst.* A26, 223. [Ch. 15].
- Hashimoto, H., A. Howie and M.J. Whelan, 1960, *Phil. Mag.* 5, 967. [Ch. 18].
- Hashimoto, H., M. Mannami and T. Naiki, 1961, *Phil. Trans. Roy. Soc.* 253, 459 and 490. [Ch. 13].
- Hashimoto, H., H. Endoh, T. Tanji, A. Ono and E. Watanabe, 1977, *J. Phys. Soc. Japan* 42, 1073. [Ch. 13].
- Hashimoto, S. and S. Ogawa, 1970, *J. Phys. Soc. Japan* 29, 710. [Ch. 17].
- Haubold, H.G., 1974, *Proc. Int. discussion meeting on Studies of lattice distortions and local atomic arrangements by X-ray, neutron and electron diffraction, Jülich, April-May, 1974.* [Ch. 12].
- Hauptman, H. and J. Karle, 1953, *Solution of the Phase Problem. I. The Centrosymmetric Crystal.* (A.C.A. Monograph, No.3). [Ch. 6].
- Hawkes, P.W. and E. Kasper, 1989, *Principles of Electron Optics, Vol.1, Basic geometric optics, Vol.2, Applied geometric optics.* (Academic Press, London). [Ch. 13].
- Hawkes, P.W. and E. Kasper, 1994, *Principles of Electron Optics, Vol. 3, Wave Optics* (Academic Press, London). [Ch. 13].
- Head, A.K., 1967, *Australian J. Phys.* 20, 557. [Ch. 18].
- Head, A.K., P. Humble, L.M. Clarebrough, A.J. Morton and C.T. Forwood, 1973, *Computed electron micrographs and defect identification* (North-Holland, Amsterdam). [Ch. 18].
- Heidenreich, R.D., 1942, *Phys. Rev.* 62, 291. [Ch. 9].
- Heidenreich, R.D., 1962, *J. Appl. Phys.* 33, 2321. [Ch. 12].
- Heinz, K. and H. Wedler, 1994, *Surface Rev. Letters*, 1, 319. [Ch. 19].
- Helmholtz, H.V., 1886, *Crelles Journ.* 100, 213. [Ch. 1].
- Hembree, G.G. and J.A. Venables, 1992, *Ultramicros.*, 49, 109. [Ch. 13].
- Heno, Y. and P.P. Ewald, 1968, *Acta Cryst.* A24, 16. [Ch. 10].
- Hendricks, S. and E. Teller, 1942, *J. Chem. Phys.* 10, 147. [Ch. 18].
- Hewat, E.A. and C.J. Humphreys, 1974, in *High Voltage Electron Microscopy*, (Academic Press, New York). [Ch. 15].
- Hillyard, S. and J. Silcox, 1993, *Ultramicros.* 52, 325. [Ch. 13].
- Hirsch, P.B., 1952, *Acta Cryst.* 5, 176. [Ch. 14].
- Hirsch, P.B., A. Howie, R.B. Nicholson, D.W. Pashley and M.J. Whelan, 1965, *Electron microscopy of thin crystals* (Butterworth and Co., London). [Ch. 8,13,15,18].
- Hoerni, J.A. and J.A. Ibers, 1953, *Phys. Rev.* 91, 1182. [Ch. 4].
- Hoerni, J.A., 1956, *Phys. Rev.* 102, 1534. [Ch. 4].
- Høier, R., 1969, *Acta Cryst.* A25, 516. [Ch. 14].
- Høier, R., 1973, *Acta Cryst.* A29, 663. [Ch. 12].
- Holmes, R.J., I.E. Pollard and C.J. Ryan, 1970, *J. Appl. Cryst.* 3, 200. [Ch. 15].
- Honjo, G., 1947, *J. Phys. Soc. Japan.* 2, 133. [Ch. 9].
- Honjo, G. and K. Mihama, 1954, *J. Phys. Soc. Japan* 9, 184. [Ch. 9].

- Hopkins, H.H., 1950, *Wave theory of aberrations* (Oxford University Press). [Ch. 3].
- Hopkins, H.H., 1953, Proc. Roy. Soc. A217, 408. [Ch. 3].
- Hoppe, W., 1956, Z. f. Krist. 107, 406. [Ch. 12].
- Hoppe, W., 1964, in: *Advances in structural research by diffraction methods*, Ed. Brill (Interscience Publ.) p. 90. [Ch. 12].
- Hoppe, W., 1969, Acta Cryst. A25, 495. [Ch. 13].
- Horstmann, M. and G. Meyer, 1962, Acta Cryst. 15, 271. [Ch. 16].
- Horstmann, M. and G. Meyer, 1965, Zeit. f. Phys. 182, 380. [Ch. 16].
- Hosemann, R. and R.N. Baggchi, 1962, *Direct analysis of diffraction by matter* (North-Holland, Amsterdam). [Ch. 5].
- Hove, L. van, 1954, Phys. Rev. 95, 249. [Ch. 12].
- Howard, S.A. and K.D. Preston, 1989, in *Reviews of Mineralogy*, Vol. 20, Eds. D.L. Bish and J.E. Post, (Mineralogical Soc., America). [Ch. 16].
- Howie, A. and M.J. Whelan, 1961, Proc. Roy. Soc. A263, 217. [Ch. 8,10,18].
- Howie, A., 1963, Proc. Roy. Soc. A271, 268. [Ch. 12].
- Howie, A., 1966, Phil. Mag. 14, 223. [Ch. 14].
- Howie, A. and Z.S. Basinski, 1968, Phil. Mag. 17, 1039. [Ch. 10,11,18].
- Howie, A., 1979, J. Microscopy, 117, 11. [Ch. 13].
- Hren, J.J., J.I. Goldstein and D.C. Joy, Eds., 1979, *Introduction to analytical electron microscopy* (Plenum Press, New York). [Ch. 13].
- Hsu, T., 1992, Micros. Res. Tech. 20, 318. [Ch. 19].
- Huang, K., 1947, Proc. Roy. Soc. A190, 102. [Ch. 12,17].
- Humble, P., 1970, in: *Modern diffraction and imaging techniques in materials science*, Eds. S. Amelinckx et al. (North-Holland, Amsterdam) p. 99. [Ch. 18].
- Humphreys, C.J. and P.B. Hirsch 1968, Phil. Mag. 18, 115 [Ch. 12].
- Humphreys, C.J. and M.J. Whelan 1969, Phil. Mag. 20, 165 [Ch. 12].
- Humphreys, C.J. and R.M. Fisher, 1971, Acta Cryst. A27, 42. [Ch. 8].
- Humphreys, C.J., J.P. Spencer, R.J. Woolf, D.C. Joy, J.M. Titchmarsh and G.R. Booker, 1972, in: *Scanning electron microscopy 1972* (I.I.T. Research Institute, Chicago). [Ch. 14].
- Humphreys, C.J., R. Sandstrom and J.P. Spencer, 1973, in: *Scanning electron microscopy 1973* (I.I.T. Research Institute, Chicago). [Ch. 14].
- Humphreys, C.J. and E.G. Bithel, 1992, in *Electron Diffraction Techniques, Vol.1*, Ed. J.M. Cowley, (Oxford Univ. Press, Oxford). [Ch. 13].
- Ichimiya, A. and R. Uyeda, 1977, Z. Naturforsch. 32a, 269. [Ch. 15].
- Ichimiya, A., 1983, Jpn. J. Appl. Phys. 22, 176. [Ch. 19].
- Iijima, S., 1971, J. Appl. Phys. 42, 5891. [Ch. 13].
- Iijima, S., 1977, Optik 47, 37. [Ch. 19].
- Iijima, S., 1981, Ultramicros. 6, 41. [Ch. 19].

- Iizumi, M., J.D. Axe, G. Shirane and K. Shimaoka, 1977, *Phys. Rev.* B15, 4392. [Ch. 17].
- Imamov, R.M. and Z.G. Pinsker, 1965, *Soviet Physics-Crystallography* 10, 148. [Ch. 16].
- Imamov, R.M., V. Pannkhorst, A.S. Avilov and Z.G. Pinsker, 1976, *Kristallografiya* 21, 364. [Ch. 16].
- Imeson, D., J.R. Sellar and C.J. Humphreys, 1979, *Proc. 37th ann. meet. Electron Microscopy Soc. Amer.*, Ed. G.W. Bailey (Claitor's, Raton Rouge). [Ch. 15].
- Ishida, K., 1970, *J. Phys. Soc. Japan* 28, 450. [Ch. 12].
- Ishizuka, K. and N. Uyeda, 1977, *Acta Cryst.* A33, 740. [Ch. 11].
- Ishizuka, K., 1980, *Ultramicroscopy* 5, 55. [Ch. 13].
- Ishizuka, K., 1982, *Acta Cryst.* A38, 773. [Ch. 11].
- Izui, K., S. Furuno and H. Otsu, 1977, *J. Elect. Micro.* 26, 129. [Ch. 13].
- Jagodzinski, H., 1949, *Acta Cryst.* 2, 201, 208 and 298. [Ch. 18].
- James, R.W., 1948, *The optical principles of the diffraction on X-rays* (G. Bell and Sons, London). [Ch. 8,12,14].
- Jap, B.K. and R.M. Glaeser, 1978, *Acta Cryst.* A34, 94. [Ch. 11].
- Jap, B.K. and R.M. Glaeser, 1980, *Acta Cryst.* A36, 57. [Ch. 13].
- Jayaram, G., P. Xu and L.D. Marks, 1993, *Phys. Rev. Letters* 71, 3489. [Ch. 19].
- Jepson, D.W., P.M. Marcus & F. Jona, 1972, *Phys. Rev.* B5, 3933. [Ch. 19].
- Johnson, A.W.S., 1972, *Acta Cryst.* A28, 89. [Ch. 15].
- Jouffrey, B. and D. Taupin, 1967, *Phil. Mag.* 15, 507. [Ch. 10].
- Jouffrey, B. and D. Taupin, 1967, *Phil. Mag.* 16, 703. [Ch. 18].
- Jouffrey, B., A. Bourret and C. Colliex, 1983, *Cours de L'Ecole de Microscopie Electronique en Science des Materiaux*, (Editions du CNRS, Paris). [Ch. 13].
- Kainuma, Y., 1953, *J. Phys. Soc. Japan* 8, 685. [Ch. 14].
- Kainuma, Y. and M. Kogiso, 1968, *Acta Cryst.* A24, 81. [Ch. 12,14].
- Kakinoki, J. and Y. Komura, 1951, *J. Inst. Poly Tech., Osaka City Uni.* 2, 1. [Ch. 18].
- Kakinoki, J. and Y. Komura, 1952, *J. Inst. Poly Tech., Osaka City Uni.* 3B, 1 and 35. [Ch. 18].
- Kakinoki, J. and Y. Komura, 1952, *J. Phys. Soc. Japan* 7, 30. [Ch. 18].
- Kambe, K., 1957, *J. Phys. Soc. Japan* 12, 13, 25. [Ch. 14].
- Kambe, K., G. Lehmppfuhl and F. Fujimoto, 1974, *Z. f. Naturforsch.* 29a, 1034. [Ch. 14].
- Kanzaki, H., 1957, *J. Phys. Chem. Solids* 2, 107. [Ch. 12].
- Karle, J. and H. Hauptman, 1956, *Acta Cryst.* 9, 635. [Ch. 6].
- Kato, N., 1952, *J. Phys. Soc. Japan* 7, 397 and 406. [Ch. 9].
- Kato, N., 1955, *J. Phys. Soc. Japan* 10, 46. [Ch. 14].
- Kato, N. and A.R. Lang, 1959, *Acta Cryst.* 12, 787. [Ch. 9,15,18].
- Kato, N., 1960, *Z. f. Naturforsch.* 15a, 369. [Ch. 18].
- Kato, N., 1961, *Acta Cryst.* 14, 526 and 627. [Ch. 9,18].
- Kato, N., 1969, *Acta Cryst.* A25, 119. [Ch. 9,15].

- Kato, N., 1973, Z. f. Naturforsch. 28a, 604. [Ch. 18].
Kato, N., 1976, Acta Cryst. A32, 543, 458. [Ch. 16].
Kato, N., 1979, Acta Cryst. A35, 9. [Ch. 16].
Kikuchi, S., 1928, Proc. Jap. Acad. Sci. 4, 271, 275, 354, 475. [Ch. 14].
Kikuta, S., T. Matsushita and K. Kohra, 1970, Phys. Letters 33A, 151. [Ch. 15].
Kikuta, S., 1971, Phys. Stat. Sol.(b) 45, 333. [Ch. 15].
Kinder, E., 1943, Naturwiss. 31, 149. [Ch. 9].
Kirkland, E.J., B.M. Siegel, N. Uyeda and Y. Fujiyoshi, 1985, Ultramicros. 17, 87. [Ch. 13].
Knowles, J.W., 1956, Acta Cryst. 9, 61. [Ch. 9].
Kogiso, M. and Y. Kainuma, 1968, J. Phys. Soc. Japan 25, 498. [Ch. 16].
Kohra, K., 1954, J. Phys. Soc. Japan 9, 690. [Ch. 14,15].
Kohra, K. and S. Kikuta, 1968, Acta Cryst. A24, 200. [Ch. 15].
Kossel, W., V. Loeck and H. Voges, 1934, Z. f. Phys. 94, 139. [Ch. 14].
Kossel, W. and G. Mollenstedt, 1939, Ann. Phys. 36, 113. [Ch. 9,14].
Krakow, W., 1979, Ultramicros. 4, 55. [Ch. 19].
Kreutle, M. and G. Meyer-Ehmsen, 1969, Phys. Stat. Sol. 35, K17. [Ch. 15].
Krivaneck, O., S. Isoda and K. Kobayashi, 1977, Phil. Mag. 36, 931. [Ch. 18].
Krivoglaz, M.A. and K.P. Ryaboshapka, 1963, Fiz. Metal. i Metalloved. 15, 18. [Ch. 18].
Krivoglaz, M.A., 1969, *Theory of X-ray and thermal neutron scattering by real crystals* (Plenum Press, New York). [Ch. 12].
Kumm, H., F. Bell, R. Sizmann, H.J. Kreiner and D. Harder, 1972, Rad. Effects 12, 52. [Ch. 14].
Kunath, W. and H. Gross, 1985, Ultramicros. 16, 349. [Ch. 13].
Kuriyama, M., 1970, Acta Cryst. A26, 56 and 667. [Ch. 8].
Kuriyama, M., 1972, Acta Cryst. A28, 588. [Ch. 8].
Kuriyama, M., 1973, Z. f. Naturforsch. 28a, 622. [Ch. 18].
Kuwabara, S., 1959, J. Phys. Soc. Japan 14, 1205. [Ch. 6].
Kuwabara, S., 1961, J. Phys. Soc. Japan 16, 2226. [Ch. 16].
Kuwabara, S., 1962, J. Phys. Soc. Japan 17, 1414. [Ch. 16].
Kuwabara, S., 1967, J. Phys. Soc. Japan 22, 1245. [Ch. 16].
Kuwabara, S. & J.M. Cowley, 1973, J. Phys. Soc. Japan 34, 1575. [Ch. 12].
- Ladd, M.F.C. and R.A. Palmer, 1993, *Structure determination by X-ray crystallography*, 3rd Edition (Plenum, New York). [Ch. 6].
Lally, J.S., C.J. Humphreys, A.J.F. Metherell and R.M. Fisher, 1972, Phil. Mag. 25 321. [Ch. 15].
Lang, A., 1958, J. Appl. Phys. 29, 597. [Ch. 18].
Lang, A., 1959, Acta Cryst. 12, 249. [Ch. 9,15,18].
Lang, A., 1970, in *Modern diffraction and imaging techniques in materials science*, Eds. S. Amelinckx et al. (North-Holland, Amsterdam). [Ch. 15].

- Larsen, P.K. and P.J. Dobson, Eds., 1988, *Reflection High Energy Electron Diffraction and Reflection Electron Imaging of Surfaces*, (Plenum Press, New York). [Ch. 19].
- Laue, M. von, 1931, *Ergeb. Exakt. Natur.* 10, 133. [Ch. 1,8].
- Laue, M. von, 1935, *Ann. Phys. Lpz.* 23, 705. [Ch. 1,14].
- Laue, M. von, 1952, *Acta Cryst.* 5, 619. [Ch. 9].
- Laval, J., 1958, *Rev. Mod. Phys.* 30, 222. [Ch. 12].
- Lehmpfuhl, G. and A. Reissland, 1968, *Z.f. Naturforsch.* 23a, 544. [Ch. 9,15].
- Lehmpfuhl, G. and C. Warble, 1986, *Ultramicros.* 19, 135. [Ch. 19].
- Leith, E.N. and J. Upatnieks, 1962, *J. Optic. Soc. Amer.* 52, 1123. [Ch. 3,13].
- Lervig, P., J. Lindhard & V. Nielsen, 1967, *Nucl. Phys.* A96, 481. [Ch. 14].
- Lichte, H., 1991, *Adv. Optical Electron Micros.* 12, 25. [Ch. 13].
- Lighthill, M.J., 1960, *Fourier analysis and generalized functions* (Cambridge University Press, Cambridge, UK). [Ch. 2].
- Lin, J.A. and J.M. Cowley, 1986, *Ultramicros.* 19, 179. [Ch. 3,13].
- Lindhard, J., 1965, *Mat. Fys. Medd. Dan. Vid. Selsk.* 34, 1. [Ch. 14].
- Linfoot, E.H., 1955, *Recent advances in optics* (Oxford University Press). [Ch. 3].
- Lipson, H. and W. Cochran, 1966, *The determination of crystal structures*, 3rd revised edition (G. Bell and Sons, London). [Ch. 6].
- Lipson, S.G. and H. Lipson, 1969, *Optical physics* (Cambridge University Press, Cambridge, UK). [Ch. 1].
- Liu, J. and J.M. Cowley, 1993, *Ultramicros.* 52, 335. [Ch. 13,19].
- Lonsdale, K., 1947, *Phil. Trans. Roy. Soc.* A240, 219. [Ch. 14].
- Lynch, D., 1971, *Acta Cryst.* A27, 399. [Ch. 11].
- Lynch, D.F. and M. O'Keefe, 1973, *Acta Cryst.* A28, 536. [Ch. 13].
- Ma, Y. and L.D. Marks, 1992, *Micros. Res. Tech.* 20, 371. [Ch. 19].
- MacGillavry, C.H., 1940, *Physica* 7, 329. [Ch. 9].
- Mackay, K.J.H., 1966, *Proc. IV Congr. X-ray optics and microanalysis* (Hermann, Paris) p. 544. [Ch. 14].
- MacLaren, J.M., J.B. Pendry, P.J. Rous, D.K. Saldin, G.A. Somorjai, M.A. Van Hove and D.D. Vvendsensky, 1987, *Surface Crystallography Information Service. A Handbook of Surface Structures*, (D. Reidel Publishing Co., Dordrecht). [Ch. 19].
- Majkrzak, C.F. and J.L. Wood, Eds., 1992, *Neutron Optical Devices and Applications*, *Proc. S.P.I.E.*, Vol. 1738. [Ch. 4].
- Maksym, P.A. and J.L. Beeby, 1984, *Surface Sci.* 140, 77. [Ch. 19].
- Mankos, M., M.R. Scheinfein and J.M. Cowley, 1994, *J. Appl. Phys.* 75, 7418. [Ch. 13].
- Margaritondo, G., 1988, *Introduction to Synchrotron Radiation*, (Oxford Univ. Press, New York). [Ch. 4].
- Marks, L.D. and D.J. Smith, 1983, *Nature* 303, 316. [Ch. 19].
- Marks, L.D., 1984, *Surface Sci.* 143, 495. [Ch. 19].
- Marks, L.D., 1986, *High-Resolution Electron Microscopy of Surfaces*, in *Topics in Current Physics*, Vol. 41, Eds., W. Schommers and P. von Blanckenhagen (Springer-Verlag, Berlin). [Ch. 19].

- Martin, H. and G. Meyer-Ehmsen, 1985, *Surf. Sci.* 151, 570 [Ch. 19].
- Matsuhata, H. and J.W. Steeds, 1987, *Phil. Mag.* 55, 39. [Ch. 15].
- Mazel, A. and R. Ayroles, 1968, *J. de Microscopie* 7, 793. [Ch. 15].
- McCoy, J.M. and P.A. Maksym, 1994, *Surface Sci.* 310, 217. [Ch. 19].
- McRae, E.G., 1966, *J. Chem. Phys.* 45, 3268. [Ch. 19].
- Menadue, J.F., 1972, *Acta Cryst.* A28, 1. [Ch. 8].
- Menter, J.W., 1956, *Proc. Roy. Soc.* A236, 119. [Ch. 13].
- Michalski, E., 1988, *Acta Cryst.* A44, 640. [Ch. 18].
- Miyake, S. and R. Uyeda, 1955, *Acta Cryst.* 8, 335. [Ch. 14,15].
- Miyake, S., 1959, *J. Phys. Soc. Japan* 14, 1347. [Ch. 8].
- Miyake, S., K. Hayakawa and R. Milda, 1968, *Acta Cryst.* A24, 182. [Ch. 9].
- Moliere, K. and H. Wagenfeld, 1958, *Z. Krist.* 110, 3. [Ch. 9].
- Möllenstedt, G. and H. Düker, 1956, *Zeit. f. Phys.* 145, 377. [Ch. 13].
- Moodie, A.F. and C.E. Warble, 1967, *Phil. Mag.* 16, 891. [Ch. 13,19].
- Moodie, A.F., 1972, *Dynamical n-beam theory of electron diffraction*, in: *Encyclopaedia dictionary of phys.* Suppl. 4, Ed. A. Thewlis. [Ch. 3,11].
- Moodie, A.F., 1978-1979, *Chem. Scripta* 14, 21. [Ch. 10].
- Moon, A.R., 1972, *Zeit. f. Naturforsch.* 27b, 390. [Ch. 10].
- Morse, P.M., 1930, *Phys. Rev.* 35, 1310. [Ch. 8].
- Moser, B., D.T. Keating and S.C. Moss, 1968, *Phys. Rev.* 175, 868. [Ch. 17].
- Moss, S.C., 1964, *J. Appl. Phys.* 35, 3547. [Ch. 17].
- Moss, S.C., 1965, in: *Local atomic arrangements studied by X-ray diffraction*, Eds. J.B. Cohen and J.E. Hilliard (Gordon and Breach, New York). [Ch. 17].
- Nagata, F. and A. Fukuhara, 1967, *Jap. J. Appl. Phys.* 6, 1233. [Ch. 15].
- Nakayama, K., S. Kikuta and K. Kohra, 1971, *Phys. Letters* 37A, 29. [Ch. 15].
- Nathan, R. 1976, in *Digital Processing of Biomedical Images*, Eds. K. Preston and M. Onoe (Plenum, New York). [Ch. 13].
- Niehurs, H., 1959a, *Z. f. Phys.* 156, 446. [Ch. 8].
- Niehurs, H., 1959b, *Z. f. Naturforsch.* 149, 504. [Ch. 8].
- Nielsen, M., J. Als-Nielsen, J. Bohr, J.P. McTague, D.E. Moncton and P.W. Stephens, 1987, *Phys. Rev.* B35, 1419. [Ch. 19].
- Nishikawa, S. and S. Kikichi, 1928, *Proc. Imp. Jpn. Acad.* 4, 475, and *Nature* 121, 1019. [Ch. 19].
- Nussbaum, A. and R.A. Phillips, 1976, *Contemporary optics for scientists and engineers* (Prentice Hall, New Jersey). [Ch. 1].
- O'Connor, D.A., 1967, *Proc. Phys. Soc.* 91, 917. [Ch. 12].
- Ohtsuki, Y.H. and S. Yanagawa, 1965, *Phys. Letters* 14, 186 [Ch. 8].
- Ohtsuki, Y.H. and S. Yanagawa, 1966, *J. Phys. Soc. Japan* 21, 326. [Ch. 8].

- Ohtsuki, Y.H., 1967, *Phys. Letters* A24, 691. [Ch. 12].
- Okamoto, K., T. Ichinokawa and Y. Ohtsuki, 1971, *J. Phys. Soc. Japan* 30, 1690. [Ch. 14].
- O'Keefe, M.A. and S. Iijima, 1978, in: *Electron Microscopy, 1978*, Vol. I, Ed. J.M. Sturgess (Microscopical Soc., Toronto, Canada) p. 282. [Ch. 11].
- Pan, M., J.M. Cowley and J. Barry, 1989, *Ultramicros.* 30, 385. [Ch. 15].
- Paskin, A., 1958, *Acta Cryst.* 11, 165. [Ch. 12].
- Paskin, A., 1959, *Acta Cryst.* 12, 290. [Ch. 12].
- Pendry, J.B., 1971, *J. Phys.* C4, 2514. [Ch. 19].
- Pendry, J.B., 1974, *Low Energy Electron Diffraction* (Academic Press, London). [Ch. 8,19].
- Peng, L.-M. and J.M. Cowley, 1986, *Acta Cryst.* A42, 552. [Ch. 19].
- Peng, L.-M., 1995, *Advances in Imaging and Electron Physics*, 90, 206 [Ch. 8,19].
- Pennycook, S.J. and D.E. Jesson, 1991, *Ultramicros.* 37, 4. [Ch. 13].
- Petykiewicz, J., 1992, *Wave Optics* (Kluwer Academic Publ., Dordrecht) [Ch. 1].
- Pfister, H., 1953, *Ann. Phys.* 11, 239. [Ch. 14].
- Pines, D., 1955, in: *Solid state physics* (Academic Press, New York). [Ch. 12].
- Pines, D., 1956, *Rev. Mod. Phys.* 28, 184. [Ch. 12].
- Pines, D., 1964, *Elementary excitations in solids* (Benjamin, New York). [Ch. 12].
- Pinsker, Z.G., 1964, *Structure analysis by electron diffraction*, Translated from the Russian (Butterworths, London). [Ch. 6,8].
- Pinsker, Z.G., 1978, *Dynamical theory of X-ray scattering in ideal Crystals* (Springer-Verlag, Berlin) (Russian Edition, 1974). [Ch. 8,9].
- Pogany, A.P., 1968, Ph.D. Thesis, Univ. of Melbourne. [Ch. 12].
- Pogany, A.P. and P.S. Turner, 1968, *Acta Cryst.* A24, 103. [Ch. 1].
- Pollard, I.E., 1970, Ph.D. Thesis, Univ. of Melbourne. [Ch. 15].
- Pollard, I.E., 1971, private communication. [Ch. 15].
- Post, B., 1979, *Acta Cryst.* A35, 17 [Ch. 15].
- Post, J.E. and D.L. Bish, 1989, in *Reviews of Mineralogy, Vol. 20*, Eds. D.L. Bish and J.E. Post (Mineral. Soc. America). [Ch. 16].
- Qin, L.C. and K. Urban, 1990, *Ultramicros.* 33, 159. [Ch. 11].
- Radi, G., 1970, *Acta Cryst.* A26, 41. [Ch. 12].
- Raether, H., 1952, *Angew. Phys.* 4, 53. [Ch. 17].
- Raether, H.A., 1980, *Excitation of plasmons and interband transitions by electrons* (Springer-Verlag, Berlin). [Ch. 12].
- Rayleigh, Lord, 1881, *Phil. Mag.* 11, 196. [Ch. 1].
- Rees, A.L.G. and J.A. Spink, 1950, *Acta Cryst.* 3, 316. [Ch. 6].
- Reimer, L., 1993, *Transmission Electron Microscopy*, 3rd Edition. (Springer-Verlag, Berlin). [Ch. 13].
- Rez, P., 1977, Ph.D. Thesis (Oxford Univ., U.K.). [Ch. 8,10].

- Rez, P., 1978, in: *Electron Diffraction 1927-77*, Eds. P.J. Dobson, J.B. Pendry & C.J. Humphreys (Institute of Physics, Bristol) p. 61. [Ch. 18].
- Rietveld, H.M., 1967, *Acta Cryst.* 22, 151. [Ch. 16].
- Rietveld, H.M., 1969, *J. Appl. Cryst.* 2, 65. [Ch. 16].
- Ritchie, R.H., 1957, *Phys. Rev.* 106, 874. [Ch. 12].
- Robinson, I.K., 1986, *Phys. Rev.* B33, 3830. [Ch. 19].
- Robinson, I.K., 1990, in *Handbook of Synchrotron Radiation, Vol. III*, Eds. D.E. Moncton and G.S. Brown, (North-Holland, Amsterdam). [Ch. 19].
- Rose, H., 1977, *Ultramicros.* 2, 251. [Ch. 13].
- Rose, H., 1995, *Ultramicros.* 56, 11. [Ch. 13].
- Ru, Q., J. Endo, T. Tanji and A. Tonomura, 1992, *Optik* 92, 51. [Ch. 13].
- Sabine, T.M., 1988, *Acta Cryst.* A44, 368. [Ch. 16].
- Sabine, T.M., 1992, in *International Tables for Crystallography, Vol. C*, Ed. A.J.C. Wilson. (Kluwer Academic Publ., Dordrecht). [Ch. 16].
- Sass, S.L. and R.W. Baluffi, 1976, *Phil. Mag.* 33, 103. [Ch. 18].
- Sass, S.L., 1980, *J. Appl. Cryst.* 13, 109. [Ch. 18].
- Sato, H. and R.S. Toth, 1963, in: *Alloying behaviour and effects in concentrated solid solutions*, Ed. T.B. Massalski (Gordon and Breach, New York). [Ch. 17].
- Saxton, W.O., 1978, *Computer Techniques of Image Processing in Electron Microscopy*, (Academic Press, New York). [Ch. 13].
- Sayre, D., 1952, *Acta Cryst.* 5, 60. [Ch. 6].
- Scherzer, O., 1949, *J. Appl. Phys.* 20, 20. [Ch. 13].
- Schomaker, V. and R. Glauber, 1952, *Nature* 170, 290. [Ch. 4].
- Schwartz, L.H. and J.B. Cohen, 1977, *Diffraction from Materials*, (Academic Press, New York). [Ch. 6,17].
- Schwarzenberger, D.R., 1959, *Phil. Mag.* 47, 1242. [Ch. 14].
- Seeman, H., 1919, *Phys. Zeit.* 20, 169. [Ch. 14].
- Seeman, H., 1926, *Naturwiss.* 23, 735. [Ch. 14].
- Self, P., M.A. O'Keefe, P. Buseck and A.E.C. Spargo, 1983, *Ultramicros.* 11, 35. [Ch. 11].
- Self, P.G. and M.A. O'Keefe, 1988, *Calculation of Diffraction Patterns and Images for Fast Electrons*, in *High Resolution Electron Microscopy and Associated Techniques*, Eds. P.R. Buseck, J.M. Cowley and L. Eyring (Oxford Univ. Press, Oxford). [Ch. 11,13].
- Sellar, J.R. and J.M. Cowley, 1973, in *Scanning electron microscopy, 1973*, Ed. Om Johari (I.I.T. Research Institute, Chicago) p. 243. [Ch. 13].
- Shinohara, K., 1932, *Sci. Pap. Inst. Phys. Chem. Res. Tokyo* 21, 21. [Ch. 14].
- Shirley, C.G. and S. Wilkins, 1972, *Phys. Rev.* B6, 1252. [Ch. 17].
- Shmueli, U., Ed., 1993, *International Tables for Crystallography, Vol. B; Reciprocal Space*, (Kluwer Academic Publ., Dordrecht). [Ch. 2,4].
- Shull, C., 1968, *Phys. Rev. Letters* 21, 1585. [Ch. 9].

- Shuman, H., 1977, *Ultramicros.* 2, 361. [Ch. 19].
- Smith, D.J., J.M. Cowley, 1971, *J. Appl Cryst.* 4, 482. [Ch. 14].
- Smith, D.J., 1989, *Advances Optical Electron Micros.* 11, 1. [Ch. 13].
- Smith, D.J., Z.G. Li, Ping Lu, M.R. McCartney and S.-C.Y. Tsen, 1991, *Ultramicros.* 37, 169. [Ch. 19].
- Somorjai, G.A., 1981, in *Fifty Years of Electron Diffraction*, Ed. P. Goodman, (Reidel Publ. Co, Dordrecht). [Ch. 19].
- Spence, J.C.H. and A.E.C. Spargo, 1970, *Phys. Letters* 33A, 116. [Ch. 12].
- Spence, J.C.H. and J.M. Cowley, 1978, *Optik* 50, 129. [Ch. 13].
- Spence, J.C.H., 1983, *Ultramicros.* 11, 117. [Ch. 19].
- Spence, J.C.H., 1988, *Experimental High-Resolution Electron Microscopy*, 2nd Edition (Oxford Univ. Press, Oxford). [Ch. 13,18].
- Spence, J.C.H., 1992, *Electron Diffraction Techniques, Vol. 1*, Ed. J.M. Cowley, (Oxford Univ. Press, Oxford). [Ch. 9,12,14,15].
- Spence, J.C.H. and J.M. Zuo, 1992, *Electron Microdiffraction*, (Plenum Press, New York). [Ch. 15].
- Stroke, W.G., M. Halioua, F. Thon and D. Willasch, 1974, *Optik* 41, 319. [Ch. 13].
- Sturkey, L., 1957, *Acta Cryst.* 10, 858. [Ch. 8].
- Sturkey, L., 1962, *Proc. Phys. Soc.* 80, 321. [Ch. 8,10,11].
- Sumida, N., Y. Uchida, F. Fujimoto and H. Fujita, 1977, in: *High voltage electron microscopy, 1977*, Eds. T. Imura and H. Hashimoto (Japanese Soc. Electron Microscopy, Tokyo) p. 281. [Ch. 14].
- Tachiki, M. and K. Teramoto, 1966, *J. Phys. Chem. Sol.* 27, 335. [Ch. 17].
- Tafto, J. and J.C.H. Spence, 1982, *Science* 218, 49, and *Ultramicros.* 9, 243. [Ch. 14].
- Takagi, S., 1958, *J. Phys. Soc. Japan* 13, 278, 287. [Ch. 12].
- Takagi, S., 1962, *Acta Cryst.* 15, 1311. [Ch. 18].
- Takayanagi, K., 1981, *Surface Sci.* 104, 527. [Ch. 19].
- Takayanagi, K., 1984, *J. Microscopy* 136, 287. [Ch. 19].
- Takayanagi, K., Y. Tanishiro, S. Takahashi and M. Takahashi, 1985, *Surface Sci.* 164, 367. [Ch. 19].
- Talbot, F., 1836, *Phil. Mag.* 9, 401. [Ch. 1].
- Tanaka, M. and M. Terauchi, 1985, *Convergent Beam Electron Diffraction*, (JEOL, Tokyo). [Ch. 15].
- Tanaka, M., M. Terauchi and T. Kaneyama, 1988, *Convergent Beam Electron Diffraction II*, (JEOL, Tokyo). [Ch. 15].
- Tanaka, M., M. Terauchi and K. Tsuda, 1994, *Convergent Beam Electron Diffraction III*, (JEOL, Tokyo). [Ch. 15].
- Tanaka, N. and J.M. Cowley, 1987, *Acta Cryst.* A43, 337. [Ch. 17].
- Taylor, C.A. and H. Lipson, 1964, *Optical transforms* (G. Bell and Sons, London). [Ch. 3,6].
- Tewordt, L., 1958, *Phys. Rev.* 109, 61. [Ch. 12].
- Thomas, G. and M.J. Goringe, 1979, *Transmission Electron Microscopy of Materials*, (Wiley, New York). [Ch. 13].

- Tibballs, J.E., 1975, *J. Appl. Cryst.* 8, 111. [Ch. 17].
- Tixier, R. and C. Wache, 1970, *J. Appl. Cryst.* 3, 466. [Ch. 14]
- Tong, S.Y., H. Li & H. Huang, 1994, *Surface Rev. Letters* 1, 303. [Ch. 19].
- Tong, S.Y., T.C. Zhao and H.C. Poon, 1988, in *Reflection High Energy Electron Diffraction and Reflection Electron Imaging of Surfaces*, Eds. P.K. Larsen and P.J. Dobson, (Plenum Press, New York). [Ch. 19].
- Tomomura, A., 1992, *Advances Physics* 41, 59. [Ch. 13].
- Tournarie, M., 1960, *Bull. Soc. Franc. Miner. Crist.* 83, 179. [Ch. 8,11].
- Tournarie, M., 1961, *C.R. Acad. Sci.* 252, 2862. [Ch. 8,11].
- Treacy, M.M.J., J.M. Newsam and M.W. Deem, 1991, *Proc. Roy. Soc. London*, A433, 499. [Ch. 18].
- Treacy, M.M.J., J.M. Newsam and M.W. Deem, 1993, *Ultramicros.* 52, 512. [Ch. 18].
- Turner, P.S. and J.M. Cowley, 1969, *Acta Cryst.* A25, 475 [Ch. 16].
- Unwin, P.N.T., 1971, *Phil. Trans. Roy. Soc. London* B261, 95. [Ch. 13].
- Unwin, P.N.T. and R. Henderson, 1975, *J. Molec. Biol.* 94, 425. [Ch. 13].
- Uyeda, N., T. Kobayashi, E. Suito, Y. Harada and M. Watanabe, 1970, in: *Microscopie électronique, 1970*, Proc. 7th Int. Conf. on Electron Microscopy, Grenoble, Vol. 1, p. 23; 1972, *J. Appl. Phys.* 43, 5181. [Ch. 13].
- Uyeda, R. and M. Nonoyama, 1965, *Jap. Journal Appl. Phys.* 4, 498. [Ch. 9].
- Uyeda, R., 1968, *Acta Cryst.* A24, 175. [Ch. 10,15].
- Vainshtein, B.K., 1964, *Structure analysis by electron diffraction*, Translated from the Russian (Pergamon Press, Oxford). [Ch. 6,16].
- Vainshtein, B.K., B.B. Zvyagin and A.S. Avilov, 1992, in *Electron Diffraction Techniques, Vol. 1*, Ed. J.M. Cowley, (Oxford Univ. Press, Oxford). [Ch. 6,16].
- Vainshtein, B.K., 1993, in *International Tables for Crystallography, Vol. B*, Ed. U. Shmueli, (Kluwer Academic Publ., Dordrecht) Section 2.5.5. [Ch. 13].
- Van Dyck, D., 1978, in: *Electron microscopy 1978*, Vol. 1, Ed., J.M. Sturgess (Microscopical Soc., Toronto, Canada) p. 196. [Ch. 13].
- Van Dyck, D., 1983, *J. Microscopy* 132, 31. [Ch. 11].
- Van Dyck, D., 1985, *Image calculation in electron microscopy: problems, progress and prospects*. in *Advances Electronics Electron Phys.*, Ed., P.W. Hawkes (Academic Press, New York), Vol. 65, p. 255. [Ch. 11].
- Van Dyck, D., J. Danckaert, W. Coene, E. Selderslaghs, D. Broddin, J. Van Landuyt and S. Amelinckx, 1989, in *Computer Simulation of Electron Microscope Diffraction and Images*, Eds. W. Krakow and M.A. O'Keefe, (Minerals, Metals and Materials Soc., Warrendale, PA). [Ch. 11,17].

- Van Dyck, D. and Op de Beeck, 1990, in *Proc. XIIth Internat. Conf. Electron Microscopy*, Eds., L.D. Peachey, D.B. Williams, (San Francisco Press, San Francisco) Vol. 1, p. 26. [Ch. 13].
- Van Hove, L., 1954, *Phys. Rev.* 93, 268; 95, 249 and 1374. [Ch. 5,12].
- Van Hove, M.A. and S.Y. Tong, 1979, *Surface Crystallography by LEED*, (Springer-Verlag, Berlin). [Ch. 19].
- Van Hove, M.A., 1981, in *Fifty Years of Electron Diffraction*, Ed., P. Goodman, (Reidel Publ. Co, Dordrecht). [Ch. 19].
- Van Hove, M.A. and S.Y. Tong, Eds., 1985, *The Structure of Surfaces*, (Springer-Verlag, Berlin). [Ch. 19].
- Vine, W.J., R. Vincent, P. Spellward and J.W. Steeds, 1992, *Ultramicros.* 41, 423. [Ch. 15].
- Voges, H., 1936, *Ann. Phys.* 27, 694. [Ch. 14].
- Wang, S.-Y. and J.M. Cowley, 1995, *Micros. Res. Tech.* 30, 181. [Ch. 3].
- Wang, Z.L., P. Lu and J.M. Cowley, 1987, *Ultramicros.* 23, 205. [Ch. 19].
- Wang, Z.L. and J.M. Cowley, 1988, *Surface Sci.* 193, 501. [Ch. 19].
- Wang, Z.L. and J.M. Cowley, 1990, *Ultramicros.* 32, 275. [Ch. 13].
- Wang, Z.L. and P. Lu, 1994, *Ultramicros.* 26, 217. [Ch. 19].
- Wang, Z.L., 1995, *Elastic and Inelastic Scattering in Electron Diffraction and Imaging*. (Plenum Press, New York). [Ch. 8,12,13].
- Warren, B.E. and B.L. Averbach, 1950, *J. Appl. Phys.* 21, 595. [Ch. 16].
- Warren, B.E., B.L. Averbach and B.W. Roberts, 1951, *J. Appl. Cryst.* 22, 1493. [Ch. 17].
- Warren, B.E. and B.L. Averbach, 1952, *J. Appl. Phys.* 23, 497. [Ch. 16].
- Warren, B.E., 1969, *X-ray diffraction* (Addison-Wesley, Reading, Mass.). [Ch. 5,8,12,16,18].
- Watanabe, D. and P.M.J. Fisher, 1965, *J. Phys. Soc. Japan* 20, 2170. [Ch. 17].
- Watanabe, D., R. Uyeda and M. Kogiso, 1968, *Acta Cryst.* A24, 249. [Ch. 15].
- Watanabe, D., R. Uyeda and A. Fukuhara, 1969, *Acta Cryst.* A25, 138. [Ch. 15].
- Watanabe, H., 1966, in *Electron microscopy, 1966*, 6th Int. Conf. on Electron Microscopy, Kyoto, Vol. 1 (Maruzen Co., Tokyo) p. 63. [Ch. 12].
- Wei, C.M., 1994, *Surface Rev. Letters* 1, 335. [Ch. 19].
- Weisel, H., 1910, *Ann. Phys. Lpz.* 33, 995. [Ch. 1].
- Whelan, M.J., 1965, *J. Appl. Phys.* 36, 2099. [Ch. 12].
- Willis, B.T.M. and A.W. Pryor, 1975, *Thermal vibrations in crystallography* (Cambridge Univ. Press, Oxford). [Ch. 12].
- Willis, B.T.M. and A. Albinati, 1992, in *International Tables for Crystallography*, Vol. C, Ed., A.J.C. Wilson (Kluwer Academic Publ., Dordrecht). [Ch. 16].
- Wilson, A.J.C., 1942, *Proc. Roy. Soc.* A180, 277. [Ch. 18].

- Wilson, A.J.C., 1952, *Acta Cryst.* 5, 318. [Ch. 18].
- Wilson, A.J.C., Ed., 1992, *International Tables for Crystallography, Vol. C: Mathematical, Physical and Chemical Tables.* (Kluwer Academic Publ., Dordrecht) [Ch. 4,5].
- Winick, H., Ed. 1980, *Synchrotron Radiation Research*, (Plenum Press, New York). [Ch. 4].
- Wolfke, M., 1913, *Ann. Phys. Lpz.* 40, 194. [Ch. 1].
- Wu, T. and T. Ohmura, 1962, *Quantum theory of scattering* (Prentice Hall, New York). [Ch. 1,4].
- Yagi, K., 1987, *J. Appl. Crystallogr.* 20, 147. [Ch. 19].
- Yagi, K., 1993, in *Electron Diffraction Techniques, Vol. 2*, Ed. J.M. Cowley, (Oxford Univ. Press, Oxford). [Ch. 19].
- Yoshioka, H., 1957, *J. Phys. Soc. Japan* 12, 618. [Ch. 12].
- You, H., 1992, in *Surface X-ray and Neutron Scattering*, Eds. H. Zabel and I.K. Robinson. (Springer-Verlag, Berlin). [Ch. 19].
- Zabel, H. and I.K. Robinson, Eds., 1992, *Surface X-ray and Neutron Scattering.* (Springer-Verlag, Berlin). [Ch. 19].
- Zachariasen, W.H., 1945, *Theory of X-ray diffractions in crystals* (Wiley, New York). [Ch. 8,14].
- Zachariasen, W.H., 1968, *Acta Cryst.* A24, 212, 324, 421, 425. [Ch. 16].
- Zachariasen, W.H., 1969, *Acta Cryst.* A25, 102. [Ch. 16].
- Zhu, J., H.Q. Ye and J.M. Cowley, 1985, *Ultramicros.* 18, 111. [Ch. 15].
- Zu, P., E.J. Kirkland, J. Silcox and R. Keyse, 1990, *Ultramicros.* 32, 93. [Ch. 13].
- Zuo, J.M., J.C.H. Spence and M. O'Keefe, 1988, *Phys. Rev. Letters* 61, 353. [Ch. 15].

INDEX

- AgTlSe₂, 383
- Abbe theory of imaging, 55, 292
- Absorption, 86, 197, 232, 279
 - electrons, 281-285, 294, 299
 - from thermal vibrations, 282
 - X-rays, 87, 280
- Absorption coefficients, 198, 283
- Aluminum, 336, 352
- Anomalous transmission, 205
- Anpassung, 174
- Anti-ferromagnetism, 140
- Astigmatism, 288
- Auger electrons, 292

- BiOCl, 383
- Babinet's principle, 47
- Beta-brass, 386
- Bethe equations, 175, 229
- Bethe potentials, 182
- Bethe second approximation, 183, 361, 382
- Blackman curve, 381
- Bloch waves, 214
- Bloch's theorem, 171
- Born approximations, 15
 - first, 15, 29, 85
 - higher, 15, 85
- Born series, 15, 210
- Borrmann effect, 205, 335
- Boundary conditions, 173
- Bragg case, 185
- Bragg-Williams long range order
 - parameters, 389, 394
- Bragg's law, 127
- Buerger precession camera, 119
- Burgers vector, 424

- Cu-hexadecachlorophthalocyanine, 301
- Cadmium sulfide, 353
- Carbon black, 161
- Center of symmetry, 133, 363
- Channelling, 190, 343-346
 - high-voltage electrons, 345, 430
 - low-energy electrons, 347
 - protons, 343
 - secondary radiations, 346
 - surface, 447
- Chromatic aberration, 293

- Clay minerals, 157
- Cobalt, 411
- Coherence, 9
 - chromatic, 71
 - degree of, 70
 - illumination, 73
 - lateral, 71
- Column approximation, 227, 430
- Compton scattering, 92, 112
- Compton wavelength, 84
- Computing methods, 248
 - consistency test, 247
 - diffuse scattering, 274-279
 - dislocation images, 229, 252, 429
 - images of crystals, 246, 252, 305
 - matrix method, 225
 - multi-slice method, 243
 - non-periodic object, 249, 318, 430
 - real-space method, 253
- Conduction electrons, 177
- Convergent-beam electron diffraction:
 - accurate structure amplitudes, 358
 - coherent, 367
 - production, 194
 - symmetry determination, 363-367
- Convolutions, 26-29
- Copper-gold alloys, 386
 - CuAu I, 386
 - CuAu II, 405-407
 - CuAu₃, 404
 - Cu₃Au, 224, 227, 386, 393, 394, 410
- Copper-zinc, 386
- Cornu spiral, 24
- Correlation functions, 387, 396, 403
- Critical temperature, 394
- Critical voltage, 359
- Crystal structure analysis, 131-138
 - electron diffraction, 141, 326
 - neutron diffraction, 138
 - X-ray diffraction, 131-138

- Debye-Waller factor, 140, 155, 362
- Defects in crystals, 146, 368
 - atom displacements, 154
 - clustered vacancies, 152
 - diffraction by, 147-163
 - point defects, 149
 - random vacancies, 150,
 - split interstitials, 163

- Delta functions, 25
- Detectors, 115
 - incoherent, 116
- Diamond, 368
- Diffraction grating, 43-45
- Diffusion of atoms, 107
- Direct methods, 136
- Dislocations,
 - diffraction effects, 424
 - edge, 424
 - electron microscope images, 426
 - X-ray topographs, 427
 - screw, 424
- Disordered alloys, 385-407
- Dispersion correction, 81
- Dispersion surfaces, 174, 176
- Displacements of atoms, 263
- Dynamical diffraction:
 - absorption effects, 202, 279
 - Bragg case - reflection, 185, 449
 - beam reduction, 222
 - Bethe theory, 170-175, 211, 449
 - crystal imperfections, 422
 - Darwin treatment, 169, 220
 - differential equations, 170, 220
 - diffuse scattering, 274, 402-405
 - disordered crystals, 278, 402
 - dispersion equations, 171
 - electron intensities, 142
 - Howie-Whelan treatment, 169, 220
 - incoherent, 276
 - integral equation methods, 169
 - intensities, 190, 380
 - LEED, 452
 - matrix equation, 170, 172, 211
 - multi-slice approaches, 231-253, 449
 - n -beam, 209-224, 351-354, 383
 - neutrons, 204
 - physical-optics method, 169, 210, 231-250, 449
 - perfect crystal, 167
 - polarization, 170
 - real-space formulations, 251
 - scattering matrix, 216, 226
 - small-angle approximation, 181, 213
 - three-beam, 211
 - two-beam, 175, 187, 218, 243, 351-354
 - X-rays, 178, 199
- Dynamical scattering factor, 279
- Electron diffraction:
 - bent crystal, 130, 192, 371
 - convergent beam, 193, 207, 358, 367
 - crystal wedges, 195
 - emitted electrons, 453
 - intensities, 377
 - microdiffraction, 291
 - reflection, 186, 207, 440, 444-447
 - relativistic effects, 209
 - selected area, 141, 288
 - structure analysis, 129, 141, 376
- Electron excitations, 270
 - plasmons, 271
 - single electrons, 272
- Electron microscopes:
 - optics, 287
 - scanning transmission, 141, 289
- Electron microscopy:
 - biological objects, 299, 327
 - column approximation, 313, 429
 - combination with electron diffraction, 326
 - crystal defects, 313, 405, 422-430
 - crystal profiles, 440
 - crystal structures, 301
 - dark-field, 300
 - incoherent factors, 293, 306
 - inelastic scattering, 307, 328
 - lattice fringes, 309
 - reflection, 447-450
 - resolution, 298
 - scanning transmission, 289, 313-318
 - stacking faults, 422
 - transfer function, 293
 - transmission, 287
 - weak-beam, 429
- Electrons:
 - absorption effects, 86
 - atom scattering amplitudes, 82, 85
 - sources, 82
- Ewald sphere, 114, 130, 236
- Excitation errors, 236
- Extinction, 372-374
 - correction factor, 373, 383
 - primary, 372
 - secondary, 372
- Extinction contours, 191, 192
- Extinction distance, 226
- Fermi surface, 401, 407
- Ferromagnetism, 140
- Flinn occupation parameters, 264, 387
- Forbidden reflections, 363
- Fourier images, 19-21, 24
- Fourier transforms, 30-48
 - convolution theorem, 33
 - definitions, 30
 - examples, 35-48
 - list of, 32
 - multiplication theorem, 33

- Fractional coordinates, 125
 Fraunhofer diffraction, 21-24
 Fresnel diffraction, 17, 24
 Fresnel integrals, 18
 Friedel's Law, 97, 363
 Fundamental length standards, 351
- Gas:
 distribution function, 109
 ideal monatomic, 106
 radial distribution function, 111
 real monatomic, 109
 scattering, 86
- Gaussian function, 45
 Gold, 305
 Graphite, 367
 Green's function, 13
- Heavy-atom method, 135
 Higher-order Laue zones, 364
 Holographic diffraction, 454
 Holography, 63-67
 electron, 318-326
 Gabor concept, 64
 in-line, 64, 320
 magnetic and electric fields, 323
 STEM, 325
 off-axis, 66, 321
- Howie-Whelan equations, 169, 220
 Huang scattering, 266
 Huyghens's Principle, 11, 29
 Hydrogen atom:
 elastic scattering, 112
- Inequality relations, 136
 Image-seeking, 100
 Imaging theory:
 Abbe theory, 55
 geometric optics, 51
 incoherent, 54
 multi-component systems, 67, 233
 small-angle approximation, 56
 wave-optics, 51-73
- Imperfect crystals:
 disordered orientations, 370
 disordered stacking, 157-160
- Incident beam convergence, 306
 Incommensurate structures, 409
 Inelastic scattering:
 electrons, 270-274, 450
 X-rays, 269
- Integrated intensities, 370
 Interaction constant, 84
 Intersecting Kikuchi lines, 362
 Ising model, 385
 Isomorphous replacement, 135
- K-line patterns, 329-343, 445
 lattice parameters, 341
 parabolas, 447
 surface atoms, 453
- Kikuchi bands, 336, 445
 Kikuchi lines, 335, 445
 intersecting, 362
 Kinematical approximation, 93, 370
 Kirchhoff formula, 11, 29
 Kossel patterns, 329-335
- Lattice relaxation, 154
 Laue conditions, 126
 Laue monatomic scattering, 268
 Laue transmission case, 178
 Least resolvable distance, 298
 LEED (low-energy electron diffraction),
 82, 187, 433, 451
 Line profile analysis, 378
 Liquids:
 distribution function, 108
 ideal, monatomic, 106
 real, monatomic, 109
- Long-range order, 362, 389, 394
 Long-range oscillatory potential,
 407, 409
- Lorentz factor, 118, 121
- Magnesium fluoro-germanate, 414
 Magnesium oxide, 194, 207, 328, 356,
 435
 Manganese fluoride, 143
 Many-atom correlation parameters, 388,
 396
- Mica crystals, 193
 Modulated structures, 408
 Moliere high-energy approximation, 84
 Molybdenum oxide, 131
 Mosaic crystals, 370-374
 Mott formula, 83, 92
- Multiple scattering:
 coherent, 94, 167, 238
 incoherent, 94
- Neutrons:
 atomic scattering, 88, 138
 Breit-Wigner formula, 88
 incoherent scattering, 89
 isotopic disorder, 89
 magnetic scattering, 91, 139
 spin scattering, 89
 thermal vibrations, 91
- Nickel oxide, 304
 Null-matrix alloys, 399
- Oblique-texture patterns, 376
 Obliquity factor, 12

- Optical diffractometer, 143
 Optical transfer function, 54
 Order parameters, 386-389
 Cu₃Au, 393
 Measurement, 394
 Ordering energies, 400
 Out-of-focus images, 60, 295-300
 Out-of-phase domains, 368, 408
- Parseval's theorem, 48
 Patterson function, 98, 120, 132, 134, 162, 259, 390, 412
 crystal defects, 148
 crystal disorder, 157, 161, 390
 four-dimensional, 104, 260, 270
 generalized, 48, 98, 412
 no average structure, 149, 157
 periodic, 102, 126, 134
 spatial correlation, 148, 390
 surface structure, 455
 vacancies, 150, 152
 Pendellösung or pendulum solution, 190
 Phase contrast, 59-63, 294-301
 aperture limitation, 62
 out-of-focus, 60
 surface structure, 435
 Zernike, 62, 295
 Phase-object approximation, 84, 241, 294, 382
 Phase problem, 97, 132
 Phonons, 258, 277
 correlation length, 258, 277
 dispersion curves, 261
 Photographic recording, 116
 Planar faults, 157, 411-424
 close-packed structures, 420
 dynamical diffraction, 368, 422
 general treatment, 414
 probability matrix, 418
 stacking faults, 420
 Plasmons, 270
 surface, 271
 Poisson's equation, 82
 Polarization effects, 190, 204
 Polycrystalline materials, 375-384
 Positron diffraction, 232
 Powder patterns, 375-384
 line-profile analysis, 378
 Rietveld analysis, 379
 Poynting vector, 202
 Probability tree, 411
 Projected charge-density approximation, 295
 Projections:
 one-dimensional, 10, 119-121
 Propagation function, 119-121
- Protein molecules, 326
 Pseudo-kinematical theory, 86
 Pseudo-Kossel patterns, 339, 341
- R-factor, 134
 Radio-interferometer, 47, 49
 Reciprocal lattice, 96
 Reciprocal space, 95, 113
 Reciprocity, 16
 for electron microscopy, 16
 for inelastic scattering, 16
 for Kossel lines, 16, 330
 theorem, 16
 Reitveld analysis, 379
 Reflection electron microscopy, 447-450
 Reflection high-energy electron diffraction (RHEED), 444-447
 Refractive index:
 electrons, 84
 X-rays, 81, 443
 Rocking curves, 191, 357
- Sayre's equation, 137
 Scattering matrix, 216, 226
 Scattering power distribution, 146
 Scattering theory, 13
 potential field, 14
 Schrödinger equation, 94
 Secondary elastic scattering, 94, 374
 Secondary radiation, 206
 Sections:
 one-dimensional, 119-121
 two-dimensional, 119-121
 Selected-area electron diffraction, 141, 288
 Shape-convolution, 101
 Shape function, 101
 Shape-transform, 127
 needle-like crystal, 129
 Short-range order, 387-390
 Short-range order diffuse scattering, 392-399
 size effects, 390, 395-399
 Silicon, 356, 439
 Single-electron excitation, 273
 Size-effects, 390-392, 395-399
 Snell's Law, 173
 Space-groups, 133, 365
 Spherical aberration, 61, 293
 Spread functions, 28, 293
 Stacking faults, 157, 412-424
 Standing waves, 443
 electrons, 443
 X-rays, 444
 Structure amplitudes, 132
 accurate measurement, 354-362

- Surface structure, 434-455
 - close-packed structures, 436
 - emitted electrons, 453-455
 - imaging, 435, 448
 - LEED, 451-453
 - superlattices, 439
 - X-ray diffraction, 441
- Symmetries, 133
 - n*-beam diffraction, 365
- Systematic absences, 133, 365
- Systematic interactions, 183
- Systematic sets, 183

- Thermal diffuse scattering, 156, 257-262
- Thermal vibrations:
 - Einstein model, 156
 - phonons, 258, 283
- Thickness fringes, 355
- Thomson scattering theory, 78
- Three-phase invariants, 137, 362
- Titanium-niobium oxides, 301
- Topographs,
 - projection, 200
 - section, 202
- Tungsten bronze, 163
- Twins, 411
- Two-beam approximation, 175, 310

- Uncertainty Principle, 9

- Warren short-range order parameter, 388
- Wave equation:
 - electromagnetic waves, 6
 - electrons, 7
- Wave functions, 5
 - plain wave, 8
 - spherical wave, 8
- Wavelength, electron, 82
 - relativistic, 82
- Wave-length spread, 117
- Weak-phase-object approximation,
 - 296, 435

- X-ray interferometry, 349
- X-ray reflectivity, 442
- X-ray topography, 200
- X-rays:
 - atomic scattering factor, 80
 - characteristic radiation, 77
 - coherence length, 77
 - synchrotron radiation, 71
 - white-radiation, 71

- Young's fringe experiment, 49

- Zernike-Van Cittert theorem, 71
- Zernike phase contrast, 62, 295
- Zero-order Laue zone, 364, 445
- Zinc, 142
- Zinc oxide, 129

This Page Intentionally Left Blank