INTERNATIONAL TABLES FOR CRYSTALLOGRAPHY

Volume B

RECIPROCAL SPACE

Edited by

U. SHMUELI

Contributing authors

E. Arnold: CABM & Rutgers University, 679 Hoes Lane, Piscataway, New Jersey 08854-5638, USA. [2.3]

M. I. Aroyo: Faculty of Physics, University of Sofia, bulv. J. Boucher 5, 1164 Sofia, Bulgaria. [1.5]

A. Authier: Laboratoire de Minéralogie-Cristallographie, Université P. et M. Curie, 4 Place Jussieu, F-75252 Paris CEDEX 05, France. [5.1]

G. Bricogne: MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, England, and LURE, Bâtiment 209D, Université Paris-Sud, 91405 Orsay, France. [1.3]

P. Coppens: Department of Chemistry, Natural Sciences & Mathematics Complex, State University of New York at Buffalo, Buffalo, New York 14260-3000, USA. [1.2]

J. M. Cowley: Arizona State University, Box 871504, Department of Physics and Astronomy, Tempe, AZ 85287-1504, USA. [2.5.1, 2.5.2, 4.3, 5.2]

R. Diamond: MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, England. [3.3]

D. L. Dorset: ExxonMobil Research and Engineering Co., 1545 Route 22 East, Clinton Township, Annandale, New Jersey 08801, USA. [2.5.7, 4.5.1, 4.5.3]

F. Frey: Institut für Kristallographie und Mineralogie, Universität, Theresienstrasse 41, D-8000 München 2, Germany. [4.2]

C. Giacovazzo: Dipartimento Geomineralogico, Campus Universitario, I-70125 Bari, Italy. [2.2]

J. K. Gjønnes: Institute of Physics, University of Oslo, PO Box 1048, N-0316 Oslo 3, Norway. [4.3]

P. Goodman† [2.5.3, 5.2]

R. W. Grosse-Kunstleve: Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Mailstop 4-230, Berkeley, CA 94720, USA. [1.4]

J.-P. Guigay: European Synchrotron Radiation Facility, BP 220, F-38043 Grenoble, France. [5.3]

T. Haibach: Laboratory of Crystallography, Swiss Federal Institute of Technology, CH-8092 Zurich, Switzerland. [4.6]

S. R. Hall: Crystallography Centre, University of Western Australia, Nedlands 6907, WA, Australia. [1.4]

H. Jagodzinski: Institut für Kristallographie und Mineralogie, Universität, Theresienstrasse 41, D-8000 München 2, Germany. [4.2]

R. E. Marsh: The Beckman Institute–139–74, California Institute of Technology, 1201 East California Blvd, Pasadena, California 91125, USA. [3.2]

R. P. Millane: Whistler Center for Carbohydrate Research, and Computational Science and Engineering Program, Purdue University, West Lafayette, Indiana 47907-1160, USA. [4.5.1, 4.5.2]

A. F. Moodie: Department of Applied Physics, Royal Melbourne Institute of Technology, 124 La Trobe Street, Melbourne, Victoria 3000, Australia. [5.2]

P. S. Pershan: Division of Engineering and Applied Science and The Physics Department, Harvard University, Cambridge, MA 02138, USA. [4.4]

S. Ramaseshan: Raman Research Institute, Bangalore 560 080, India. [2.4]

M. G. Rossmann: Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907, USA. [2.3]

D. E. Sands: Department of Chemistry, University of Kentucky, Chemistry–Physics Building, Lexington, Kentucky 40506-0055, USA. [3.1]

M. Schlenker: Laboratoire Louis Néel du CNRS, BP 166, F-38042 Grenoble CEDEX 9, France. [5.3]

V. Schomaker† [3.2]

U. Shmueli: School of Chemistry, Tel Aviv University, 69 978 Tel Aviv, Israel. [1.1, 1.4, 2.1]

W. Steurer: Laboratory of Crystallography, Swiss Federal Institute of Technology, CH-8092 Zurich, Switzerland. [4.6]

B. K. Vainshtein† [2.5.4, 2.5.5, 2.5.6]

M. Vijayan: Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India. [2.4]

D. E. Williams: Department of Chemistry, University of Louisville, Louisville, Kentucky 40292, USA. [3.4]

B. T. M. Willis: Chemical Crystallography Laboratory, University of Oxford, 9 Parks Road, Oxford OX1 3PD, England. [4.1]

A. J. C. Wilson† [2.1]

H. Wondratschek: Institut für Kristallographie, Universität, D-76128 Karlsruhe, Germany. [1.5]

B. B. Zvyagin: Institute of Ore Mineralogy (IGEM), Academy of Sciences of Russia, Staromonetny 35, 109017 Moscow, Russia. [2.5.4]

† Deceased.

† Deceased.

Contents

Preface

By Uri Shmueli

The purpose of Volume B of *International Tables for Crystallography* is to provide the user or reader with accounts of some well established topics, of importance to the science of crystallography, which are related in one way or another to the concepts of reciprocal lattice and, more generally, reciprocal space. Efforts have been made to extend the treatment of the various topics to include X-ray, electron, and neutron diffraction techniques, and thereby do some justice to the inclusion of the present Volume in the new series of *International Tables for Crystallography*.

An important crystallographic aspect of symmetry in reciprocal space, space-group-dependent expressions of trigonometric structure factors, already appears in Volume I of *International Tables for X-ray Crystallography*, and preliminary plans for incorporating this and other crystallographic aspects of reciprocal space in the new edition of *International Tables* date back to 1972. However, work on a volume of *International Tables for Crystallography*, largely dedicated to the subject of reciprocal space, began over ten years later. The present structure of Volume B, as determined in the years preceding the 1984 Hamburg congress of the International Union of Crystallography (IUCr), is due to (i) computer-controlled production of concise structure-factor tables, (ii) the ability to introduce many more aspects of reciprocal space – as a result of reducing the effort of producing the above tables, as well as their volume, and (iii) suggestions by the National Committees and individual crystallographers of some additional interesting topics. It should be pointed out that the initial plans for the present Volume and Volume C (*Mathematical, Physical and Chemical Tables*, edited by Professor A. J. C. Wilson), were formulated and approved during the same period.

The obviously delayed publication of Volume B is due to several reasons. Some minor delays were caused by a requirement that potential contributors should be approved by the Executive Committee prior to issuing relevant invitations. Much more serious delays were caused by authors who failed to deliver their contributions. In fact, some invited contributions had to be excluded from this first edition of Volume B. Some of the topics here treated are greatly extended, considerably updated or modern versions of similar topics previously treated in the old Volumes I, II, and IV. Most of the subjects treated in Volume B are new to *International Tables*.

I gratefully thank Professor A. J. C. Wilson, for suggesting that I edit this Volume and for sharing with me his rich editorial experience. I am indebted to those authors of Volume B who took my requests and deadlines seriously, and to the Computing Center of Tel Aviv University for computing facilities and time. Special thanks are due to Mrs Z. Stein (Tel Aviv University) for skilful assistance in numeric and symbolic programming, involved in my contributions to this Volume.

I am most grateful to many colleagues–crystallographers for encouragement, advice, and suggestions. In particular, thanks are due to Professors J. M. Cowley, P. Goodman and C. J. Humphreys, who served as Chairmen of the Commission on Electron Diffraction during the preparation of this Volume, for prompt and expert help at all stages of the editing. The kind assistance of Dr J. N. King, the Executive Secretary of the IUCr, is also gratefully acknowledged. Last, but certainly not least, I wish to thank Mr M. H. Dacombe, the Technical Editor of the IUCr, and his staff for the skilful and competent treatment of the variety of drafts and proofs out of which this Volume arose.

Preface to the second edition

By Uri Shmueli

The first edition of Volume B appeared in 1993, and was followed by a corrected reprint in 1996. Although practically all the material for the second edition was available in early 1997, its publication was delayed by the decision to translate all of Volume B, and indeed all the other volumes of *International Tables for Crystallography*, to Standard Generalized Markup Language (SGML) and thus make them available also in an electronic form suitable for modern publishing procedures.

During the preparation of the second edition, most chapters that appeared in the first edition have been corrected and/or revised, some were rather extensively updated, and five new chapters were added. The overall structure of the second edition is outlined below.

After an introductory chapter, Part 1 presents the reader with an account of structure-factor formalisms, an extensive treatment of the theory, algorithms and crystallographic applications of Fourier methods, and treatments of symmetry in reciprocal space. These are here enriched with more advanced aspects of representations of space groups in reciprocal space.

In Part 2, these general accounts are followed by detailed expositions of crystallographic statistics, the theory of direct methods, Patterson techniques, isomorphous replacement and anomalous scattering, and treatments of the role of electron microscopy and diffraction in crystal structure determination. The latter topic is here enhanced by applications of direct methods to electron crystallography.

Part 3, *Dual Bases in Crystallographic Computing*, deals with applications of reciprocal space to molecular geometry and 'best'-plane calculations, and contains a treatment of the principles of molecular graphics and modelling and their applications; it concludes with the presentation of a convergence-acceleration method, of importance in the computation of approximate lattice sums.

Part 4 contains treatments of various diffuse-scattering phenomena arising from crystal dynamics, disorder and low dimensionality (liquid crystals), and an exposition of the underlying theories and/or experimental evidence. The new additions to this part are treatments of polymer crystallography and of reciprocal-space images of aperiodic crystals.

Part 5 contains introductory treatments of the theory of the interaction of radiation with matter, the so-called dynamical theory, as applied to X-ray, electron and neutron diffraction techniques. The chapter on the dynamical theory of neutron diffraction is new.

I am deeply grateful to the authors of the new contributions for making their expertise available to Volume B and for their

excellent collaboration. I also take special pleasure in thanking those authors of the first edition who revised and updated their contributions in view of recent developments. Last but not least, I wish to thank all the authors for their contributions and their patience, and am grateful to those authors who took my requests seriously. I hope that the updating and revision of future editions will be much easier and more expedient, mainly because of the new format of *International Tables*.

Four friends and greatly respected colleagues who contributed to the second edition of Volume B are no longer with us. These are Professors Arthur J. C. Wilson, Peter Goodman, Verner Schomaker and Boris K. Vainshtein. I asked Professors Michiyoshi Tanaka, John Cowley and Douglas Dorset if they were prepared to answer queries related to the contributions of the late Peter Goodman and Boris K. Vainshtein to Chapter 2.5. I am most grateful for their prompt agreement.

This editorial work was carried out at the School of Chemistry and the Computing Center of Tel Aviv University. The facilities they put at my disposal are gratefully acknowledged on my behalf and on behalf of the IUCr. I wish to thank many colleagues for interesting conversations and advice, and in particular Professor Theo Hahn with whom I discussed at length problems regarding Volume B and *International Tables* in general.

Given all these expert contributions, the publication of this volume would not have been possible without the expertise and devotion of the Technical Editors of the IUCr. My thanks go to Mrs Sue King, for her cooperation during the early stages of the work on the second edition of Volume B, while the material was being collected, and to Dr Nicola Ashcroft, for her collaboration during the final stages of the production of the volume, for her most careful and competent treatment of the proofs, and last but not least for her tactful and friendly attitude.

SAMPLE PAGES

1.1. Reciprocal space in crystallography

BY U. SHMUELI

1.1.1. Introduction

The purpose of this chapter is to provide an introduction to several aspects of reciprocal space, which are of general importance in crystallography and which appear in the various chapters and sections to follow. We first summarize the basic definitions and briefly inspect some fundamental aspects of crystallography, while recalling that they can be usefully and simply discussed in terms of the concept of the reciprocal lattice. This introductory section is followed by a summary of the basic relationships between the direct and associated reciprocal lattices. We then introduce the elements of tensor-algebraic formulation of such dual relationships, with emphasis on those that are important in many applications of reciprocal space to crystallographic algorithms. We proceed with a section that demonstrates the role of mutually reciprocal bases in transformations of coordinates and conclude with a brief outline of some important analytical aspects of reciprocal space, most of which are further developed in other parts of this volume.

1.1.2. Reciprocal lattice in crystallography

The notion of mutually reciprocal triads of vectors dates back to the introduction of vector calculus by J. Willard Gibbs in the 1880s (*e.g.* Wilson, 1901). This concept appeared to be useful in the early interpretations of diffraction from single crystals (Ewald, 1913; Laue, 1914) and its first detailed exposition and the recognition of its importance in crystallography can be found in Ewald's (1921) article. The following free translation of Ewald's (1921) introduction, presented in a somewhat different notation, may serve the purpose of this section:

> To the set of $\mathbf{a}_i$, there corresponds in the vector calculus a set of 'reciprocal vectors' $\mathbf{b}_i$, which are defined (by Gibbs) by the following properties:
>
> $$\mathbf{a}_i \cdot \mathbf{b}_k = 0 \quad (\text{for } i \neq k) \tag{1.1.2.1}$$
>
> and
>
> $$\mathbf{a}_i \cdot \mathbf{b}_i = 1, \tag{1.1.2.2}$$
>
> where i and k may each equal 1, 2 or 3. The first equation, (1.1.2.1), says that each vector $\mathbf{b}_k$ is perpendicular to two vectors $\mathbf{a}_i$, as follows from the vanishing scalar products. Equation (1.1.2.2) provides the norm of the vector $\mathbf{b}_i$: the length of this vector must be chosen such that the projection of $\mathbf{b}_i$ on the direction of $\mathbf{a}_i$ has the length $1/a_i$, where a_i is the magnitude of the vector $\mathbf{a}_i$. . ..

The consequences of equations (1.1.2.1) and (1.1.2.2) were elaborated by Ewald (1921) and are very well documented in the subsequent literature, crystallographic as well as other.

As is well known, the reciprocal lattice occupies a rather prominent position in crystallography and there are nearly as many accounts of its importance as there are crystallographic texts. It is not intended to review its applications, in any detail, in the present section; this is done in the remaining chapters and sections of the present volume. It seems desirable, however, to mention by way of an introduction some fundamental geometrical, physical and mathematical aspects of crystallography, and try to give a unified demonstration of the usefulness of mutually reciprocal bases as an interpretive tool.

Consider the equation of a lattice plane in the direct lattice. It is shown in standard textbooks (*e.g.* Buerger, 1941) that this equation is given by

$$hx + ky + lz = n, \tag{1.1.2.3}$$

where h, k and l are relatively prime integers (*i.e.* not having a common factor other than $+1$ or -1), known as Miller indices of the lattice plane, x, y and z are the coordinates of any point lying in the plane and are expressed as fractions of the magnitudes of the basis vectors **a**, **b** and **c** of the direct lattice, respectively, and n is an integer denoting the serial number of the lattice plane within the family of parallel and equidistant (hkl) planes, the interplanar spacing being denoted by d_{hkl}; the value $n = 0$ corresponds to the (hkl) plane passing through the origin.

Let $\mathbf{r} = x\mathbf{a} + y\mathbf{b} + z\mathbf{c}$ and $\mathbf{r}_L = u\mathbf{a} + v\mathbf{b} + w\mathbf{c}$, where u, v, w are any integers, denote the position vectors of the point xyz and a lattice point uvw lying in the plane (1.1.2.3), respectively, and assume that $\mathbf{r}$ and $\mathbf{r}_L$ are different vectors. If the plane normal is denoted by **N**, where **N** is proportional to the vector product of two in-plane lattice vectors, the vector form of the equation of the lattice plane becomes

$$\mathbf{N} \cdot (\mathbf{r} - \mathbf{r}_L) = 0 \quad \text{or} \quad \mathbf{N} \cdot \mathbf{r} = \mathbf{N} \cdot \mathbf{r}_L. \tag{1.1.2.4}$$

For equations (1.1.2.3) and (1.1.2.4) to be identical, the plane normal **N** must satisfy the requirement that $\mathbf{N} \cdot \mathbf{r}_L = n$, where n is an (unrestricted) integer.

Let us now consider the basic diffraction relations (*e.g.* Lipson & Cochran, 1966). Suppose a parallel beam of monochromatic radiation, of wavelength λ, falls on a lattice of identical point scatterers. If it is assumed that the scattering is elastic, *i.e.* there is no change of the wavelength during this process, the wavevectors of the incident and scattered radiation have the same magnitude, which can conveniently be taken as $1/\lambda$. A consideration of path and phase differences between the waves outgoing from two point scatterers separated by the lattice vector $\mathbf{r}_L$ (defined as above) shows that the condition for their maximum constructive interference is given by

$$(\mathbf{s} - \mathbf{s}_0) \cdot \mathbf{r}_L = n, \tag{1.1.2.5}$$

where $\mathbf{s}_0$ and **s** are the wavevectors of the incident and scattered beams, respectively, and n is an arbitrary integer.

Since $\mathbf{r}_L = u\mathbf{a} + v\mathbf{b} + w\mathbf{c}$, where u, v and w are unrestricted integers, equation (1.1.2.5) is equivalent to the equations of Laue:

$$\mathbf{h} \cdot \mathbf{a} = h, \quad \mathbf{h} \cdot \mathbf{b} = k, \quad \mathbf{h} \cdot \mathbf{c} = l, \tag{1.1.2.6}$$

where $\mathbf{h} = \mathbf{s} - \mathbf{s}_0$ is the diffraction vector, and h, k and l are integers corresponding to orders of diffraction from the three-dimensional lattice (Lipson & Cochran, 1966). The diffraction vector thus has to satisfy a condition that is analogous to that imposed on the normal to a lattice plane.

The next relevant aspect to be commented on is the Fourier expansion of a function having the periodicity of the crystal lattice. Such functions are *e.g.* the electron density, the density of nuclear matter and the electrostatic potential in the crystal, which are the operative definitions of crystal structure in X-ray, neutron and electron-diffraction methods of crystal structure determination. A Fourier expansion of such a periodic function may be thought of as a superposition of waves (*e.g.* Buerger, 1959), with wavevectors related to the interplanar spacings d_{hkl}, in the crystal lattice. Denoting the wavevector of a Fourier wave by **g** (a function of hkl), the phase of the Fourier wave at the point **r** in the crystal is given by $2\pi\mathbf{g} \cdot \mathbf{r}$, and the triple Fourier series corresponding to the expansion of the periodic function, say $G(\mathbf{r})$, can be written as

$$G(\mathbf{r}) = \sum_{\mathbf{g}} C(\mathbf{g}) \exp(-2\pi i \mathbf{g} \cdot \mathbf{r}), \tag{1.1.2.7}$$

where $C(\mathbf{g})$ are the amplitudes of the Fourier waves, or Fourier

1.2. The structure factor

BY P. COPPENS

1.2.1. Introduction

The *structure factor* is the central concept in structure analysis by diffraction methods. Its modulus is called the *structure amplitude*. The structure amplitude is a function of the indices of the set of scattering planes h, k and l, and is defined as the amplitude of scattering by the contents of the crystallographic unit cell, expressed in units of scattering. For X-ray scattering, that unit is the scattering by a single electron (2.82×10^{-15} m), while for neutron scattering by atomic nuclei, the unit of scattering length of 10^{-14} m is commonly used. The complex form of the structure factor means that the phase of the scattered wave is not simply related to that of the incident wave. However, the observable, which is the scattered intensity, must be real. It is proportional to the square of the scattering amplitude (see, *e.g.*, Lipson & Cochran, 1966).

The structure factor is directly related to the distribution of scattering matter in the unit cell which, in the X-ray case, is the electron distribution, time-averaged over the vibrational modes of the solid.

In this chapter we will discuss structure-factor expressions for X-ray and neutron scattering, and, in particular, the modelling that is required to obtain an analytical description in terms of the features of the electron distribution and the vibrational displacement parameters of individual atoms. We concentrate on the most basic developments; for further details the reader is referred to the cited literature.

1.2.2. General scattering expression for X-rays

The total scattering of X-rays contains both elastic and inelastic components. Within the first-order Born approximation (Born, 1926) it has been treated by several authors (*e.g.* Waller & Hartree, 1929; Feil, 1977) and is given by the expression

$$I_{\rm total}(\mathbf{S}) = I_{\rm classical}\sum_n \left|\int \psi_n^* \exp(2\pi i \mathbf{S}\cdot\mathbf{r}_j)\psi_0 \, {\rm d}\mathbf{r}\right|^2, \qquad (1.2.2.1)$$

where $I_{\rm classical}$ is the classical Thomson scattering of an X-ray beam by a free electron, which is equal to $(e^2/mc^2)^2(1+\cos^2 2\theta)/2$ for an unpolarized beam of unit intensity, ψ is the n-electron space-wavefunction expressed in the $3n$ coordinates of the electrons located at $\mathbf{r}_j$ and the integration is over the coordinates of all electrons. $\mathbf{S}$ is the scattering vector of length $2\sin\theta/\lambda$.

The coherent elastic component of the scattering, in units of the scattering of a free electron, is given by

$$I_{\rm coherent,\ elastic}(\mathbf{S}) = \left|\int \psi_0^*\right|\sum_j \exp(2\pi i \mathbf{S}\cdot\mathbf{r}_j)|\psi_0 \, {\rm d}\mathbf{r}|^2. \qquad (1.2.2.2)$$

If integration is performed over all coordinates but those of the jth electron, one obtains after summation over all electrons

$$I_{\rm coherent,\ elastic}(\mathbf{S}) = |\int \rho(\mathbf{r}) \exp(2\pi i \mathbf{S}\cdot\mathbf{r}) \, {\rm d}\mathbf{r}|^2, \qquad (1.2.2.3)$$

where $\rho(\mathbf{r})$ is the electron distribution. The scattering amplitude $A(\mathbf{S})$ is then given by

$$A(\mathbf{S}) = \int \rho(\mathbf{r}) \exp(2\pi i \mathbf{S}\cdot\mathbf{r}) \, {\rm d}\mathbf{r} \qquad (1.2.2.4a)$$

or

$$A(\mathbf{S}) = \hat{F}\{\rho(\mathbf{r})\}, \qquad (1.2.2.4b)$$

where $\hat{F}$ is the Fourier transform operator.

1.2.3. Scattering by a crystal: definition of a structure factor

In a crystal of infinite size, $\rho(\mathbf{r})$ is a three-dimensional periodic function, as expressed by the convolution

$$\rho_{\rm crystal}(\mathbf{r}) = \sum_n\sum_m\sum_p \rho_{\rm unit\ cell}(\mathbf{r}) * \delta(\mathbf{r} - n\mathbf{a} - m\mathbf{b} - p\mathbf{c}), \quad (1.2.3.1)$$

where n, m and p are integers, and δ is the Dirac delta function.

Thus, according to the Fourier convolution theorem,

$$\begin{aligned} A(\mathbf{S}) &= \hat{F}\{\rho(\mathbf{r})\} \\ &= \sum_n\sum_m\sum_p \hat{F}\{\rho_{\rm unit\ cell}(\mathbf{r})\}\hat{F}\{\delta(\mathbf{r} - n\mathbf{a} - m\mathbf{b} - p\mathbf{c}\}, \end{aligned} \quad (1.2.3.2)$$

which gives

$$A(\mathbf{S}) = \hat{F}\{\rho_{\rm unit\ cell}(\mathbf{r})\}\sum_h\sum_k\sum_l \delta(\mathbf{S} - h\mathbf{a}^* - k\mathbf{b}^* - l\mathbf{c}^*). \quad (1.2.3.3)$$

Expression (1.2.3.3) is valid for a crystal with a very large number of unit cells, in which particle-size broadening is negligible. Furthermore, it does not account for multiple scattering of the beam within the crystal. Because of the appearance of the delta function, (1.2.3.3) implies that $\mathbf{S} = \mathbf{H}$ with $\mathbf{H} = h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^*$.

The first factor in (1.2.3.3), the scattering amplitude of one unit cell, is defined as the structure factor F:

$$F(\mathbf{H}) = \hat{F}\{\rho_{\rm unit\ cell}(\mathbf{r})\} = \int_{\rm unit\ cell}\rho(\mathbf{r}) \exp(2\pi i \mathbf{H}\cdot\mathbf{r}) \, {\rm d}\mathbf{r}. \quad (1.2.3.4)$$

1.2.4. The isolated-atom approximation in X-ray diffraction

To a reasonable approximation, the unit-cell density can be described as a superposition of isolated, spherical atoms located at $\mathbf{r}_j$.

$$\rho_{\rm unit\ cell}(\mathbf{r}) = \sum_j \rho_{{\rm atom},\,j}(\mathbf{r}) * \delta(\mathbf{r} - \mathbf{r}_j). \qquad (1.2.4.1)$$

Substitution in (1.2.3.4) gives

$$F(\mathbf{H}) = \sum_j \hat{F}\{\rho_{{\rm atom},\,j}\}\hat{F}\{\delta(\mathbf{r} - \mathbf{r}_j)\} = \sum_j f_j \exp(2\pi i \mathbf{H}\cdot\mathbf{r}_j) \qquad (1.2.4.2a)$$

or

$$\begin{aligned} F(h,k,l) &= \sum_j f_j \exp 2\pi i(hx_j + ky_j + lz_j) \\ &= \sum_j f_j\{\cos 2\pi(hx_j + ky_j + lz_j) \\ &\quad + i\sin 2\pi(hx_j + ky_j + lz_j)\}. \end{aligned} \qquad (1.2.4.2b)$$

$f_j(S)$, the spherical atomic scattering factor, or form factor, is the Fourier transform of the spherically averaged atomic density $\rho_j(r)$, in which the polar coordinate r is relative to the nuclear position. $f_j(S)$ can be written as (James, 1982)

Table 1.2.7.4. *Closed-form expressions for Fourier transform of Slater-type functions (Avery & Watson, 1977; Su & Coppens, 1990)*

$\langle j_k \rangle \equiv \int_0^\infty r^N \exp(-Zr) j_k(Kr)\, \mathrm{d}r, K = 4\pi \sin\theta/\lambda.$

	N							
k	1	2	3	4	5	6	7	8
0	$\frac{1}{K^2+Z^2}$	$\frac{2Z}{(K^2+Z^2)^2}$	$\frac{2(3Z^2-K^2)}{(K^2+Z^2)^3}$	$\frac{24Z(Z^2-K^2)}{(K^2+Z^2)^4}$	$\frac{24(5Z^2-10K^2Z^2+K^4)}{(K^2+Z^2)^5}$	$\frac{240Z(K^2-3Z^2)(3K^2-Z^2)}{(K^2+Z^2)^6}$	$\frac{720(7Z^6-35K^2Z^4+21K^4Z^2-K^6)}{(K^2+Z^2)^7}$	$\frac{40320(Z^7-7K^2Z^5+7K^4Z^3-K^6Z)}{(K^2+Z^2)^8}$
1		$\frac{2K}{(K^2+Z^2)^2}$	$\frac{8KZ}{(K^2+Z^2)^3}$	$\frac{8K(5Z^2-K^2)}{(K^2+Z^2)^4}$	$\frac{48KZ(5Z^2-3K^2)}{(K^2+Z^2)^5}$	$\frac{48K(35Z^4-42K^2Z^2+3K^4)}{(K^2+Z^2)^6}$	$\frac{1920KZ(7Z^4-14K^2Z^2+3K^4)}{(K^2+Z^2)^7}$	$\frac{5760K(21Z^6-63K^2Z^4+27K^4Z^2-K^6)}{(K^2+Z^2)^8}$
2			$\frac{8K^2}{(K^2+Z^2)^3}$	$\frac{48K^2Z}{(K^2+Z^2)^4}$	$\frac{48K^2(7Z^2-K^2)}{(K^2+Z^2)^5}$	$\frac{384K^2Z(7Z^2-3K^2)}{(K^2+Z^2)^6}$	$\frac{1152K^2(21Z^4-18K^2Z^2+K^4)}{(K^2+Z^2)^7}$	$\frac{11520K^2Z(21Z^4-30K^2Z^2+5K^4)}{(K^2+Z^2)^8}$
3				$\frac{48K^3}{(K^2+Z^2)^4}$	$\frac{384K^3Z}{(K^2+Z^2)^5}$	$\frac{384K^3(9Z^2-K^2)}{(K^2+Z^2)^6}$	$\frac{11520K^3Z(3Z^2-K^2)}{(K^2+Z^2)^7}$	$\frac{11520K^3(33Z^4-22K^2Z^2+K^4)}{(K^2+Z^2)^8}$
4					$\frac{384K^4}{(K^2+Z^2)^5}$	$\frac{3840K^4Z}{(K^2+Z^2)^6}$	$\frac{3840K^4(11Z^2-K^2)}{(K^2+Z^2)^7}$	$\frac{46080K^4Z(11Z^2-3K^2)}{(K^2+Z^2)^8}$
5						$\frac{3840K^5}{(K^2+Z^2)^6}$	$\frac{46080K^5Z}{(K^2+Z^2)^7}$	$\frac{40680K^5(13Z^2-K^2)}{(K^2+Z^2)^8}$
6							$\frac{46080K^6}{(K^2+Z^2)^7}$	$\frac{645120K^6Z}{(K^2+Z^2)^8}$
7								$\frac{645120K^7}{(K^2+Z^2)^8}$

$$P(\mathbf{u}) = \frac{|\boldsymbol{\sigma}^{-1}|^{1/2}}{(2\pi)^{3/2}} \exp\{-\tfrac{1}{2}\boldsymbol{\sigma}^{-1}_{jk}(u^j u^k)\}. \qquad (1.2.10.2a)$$

Here σ is the variance–covariance matrix, with covariant components, and $|\boldsymbol{\sigma}^{-1}|$ is the determinant of the inverse of σ. Summation over repeated indices has been assumed. The corresponding equation in matrix notation is

$$P(\mathbf{u}) = \frac{|\boldsymbol{\sigma}^{-1}|^{1/2}}{(2\pi)^{3/2}} \exp\{-\tfrac{1}{2}(\mathbf{u})^T \boldsymbol{\sigma}^{-1}(\mathbf{u})\}, \qquad (1.2.10.2b)$$

where the superscript T indicates the transpose.

The characteristic function, or Fourier transform, of $P(\mathbf{u})$ is

$$T(\mathbf{H}) = \exp\{-2\pi^2 \sigma^{jk} h_j h_k\} \qquad (1.2.10.3a)$$

or

$$T(\mathbf{H}) = \exp\{-2\pi^2 \mathbf{H}^T \boldsymbol{\sigma} \mathbf{H}\}. \qquad (1.2.10.3b)$$

With the change of variable $b^{jk} = 2\pi^2 \sigma^{jk}$, (1.2.10.3*a*) becomes

$$T(\mathbf{H}) = \exp\{-b^{jk} h_j h_k\}.$$

1.2.11. Rigid-body analysis

The treatment of rigid-body motion of molecules or molecular fragments was developed by Cruickshank (1956) and expanded into a general theory by Schomaker & Trueblood (1968). The theory has been described by Johnson (1970*b*) and by Dunitz (1979). The latter reference forms the basis for the following treatment.

The most general motions of a rigid body consist of rotations about three axes, coupled with translations parallel to each of the axes. Such motions correspond to screw rotations. A libration around a vector $\boldsymbol{\lambda}$ (λ_1, λ_2, λ_3), with length corresponding to the magnitude of the rotation, results in a displacement $\delta\mathbf{r}$, such that

$$\delta\mathbf{r} = (\boldsymbol{\lambda} \times \mathbf{r}) = \mathbf{D}\mathbf{r} \qquad (1.2.11.1)$$

with

$$\mathbf{D} = \begin{bmatrix} 0 & -\lambda_3 & \lambda_2 \\ \lambda_3 & 0 & -\lambda_1 \\ -\lambda_2 & \lambda_1 & 0 \end{bmatrix}, \qquad (1.2.11.2)$$

or in tensor notation, assuming summation over repeated indices,

$$\delta r_i = D_{ij} r_j = -\varepsilon_{ijk} \lambda_k r_j \qquad (1.2.11.3)$$

where the permutation operator ε_{ijk} equals +1 for i, j, k a cyclic permutation of the indices 1, 2, 3, or -1 for a non-cyclic permutation, and zero if two or more indices are equal. For $i = 1$, for example, only the ε_{123} and ε_{132} terms occur. Addition of a translational displacement gives

$$\delta r_i = D_{ij} r_j + t_i. \qquad (1.2.11.4)$$

When a rigid body undergoes vibrations the displacements vary with time, so suitable averages must be taken to derive the mean-square displacements. If the librational and translational motions are independent, the cross products between the two terms in (1.2.11.4) average to zero and the elements of the mean-square displacement tensor of atom n, U^n_{ij}, are given by

$$\begin{aligned}
U^n_{11} &= +L_{22} r_3^2 + L_{33} r_2^2 - 2L_{23} r_2 r_3 + T_{11} \\
U^n_{22} &= +L_{33} r_1^2 + L_{11} r_3^2 - 2L_{13} r_1 r_3 + T_{22} \\
U^n_{33} &= +L_{11} r_2^2 + L_{22} r_1^2 - 2L_{12} r_1 r_2 + T_{33} \\
U^n_{12} &= -L_{33} r_1 r_2 - L_{12} r_3^2 + L_{13} r_2 r_3 + L_{23} r_1 r_3 + T_{12} \\
U^n_{13} &= -L_{22} r_1 r_3 + L_{12} r_2 r_3 - L_{13} r_2^2 + L_{23} r_1 r_2 + T_{13} \\
U^n_{23} &= -L_{11} r_2 r_3 + L_{12} r_1 r_3 - L_{13} r_1 r_2 - L_{23} r_1^2 + T_{23},
\end{aligned} \qquad (1.2.11.5)$$

where the coefficients $L_{ij} = \langle \lambda_i \lambda_j \rangle$ and $T_{ij} = \langle t_i t_j \rangle$ are the elements of the 3×3 libration tensor $\mathbf{L}$ and the 3×3 translation tensor $\mathbf{T}$,

1.3.2.7.4. *Matrix representation of the discrete Fourier transform (DFT)*

By virtue of definitions (i) and (ii),

$$\mathscr{F}(\mathbf{N})\mathbf{v}_{k^*} = \frac{1}{|\det \mathbf{N}|}\sum_{k} \exp[-2\pi i k^* \cdot (\mathbf{N}^{-1}k)]\mathbf{u}_k$$
$$\bar{\mathscr{F}}(\mathbf{N})\mathbf{u}_k = \sum_{k^*} \exp[+2\pi i k^* \cdot (\mathbf{N}^{-1}k)]\mathbf{v}_{k^*}$$

so that $\mathscr{F}(\mathbf{N})$ and $\bar{\mathscr{F}}(\mathbf{N})$ may be represented, in the canonical bases of $W_{\mathbf{N}}$ and $W^*_{\mathbf{N}}$, by the following matrices:

$$[\mathscr{F}(\mathbf{N})]_{kk^*} = \frac{1}{|\det \mathbf{N}|}\exp[-2\pi i k^* \cdot (\mathbf{N}^{-1}k)]$$
$$[\bar{\mathscr{F}}(\mathbf{N})]_{k^*k} = \exp[+2\pi i k^* \cdot (\mathbf{N}^{-1}k)].$$

When $\mathbf{N}$ is symmetric, $\mathbb{Z}^n/\mathbf{N}\mathbb{Z}^n$ and $\mathbb{Z}^n/\mathbf{N}^T\mathbb{Z}^n$ may be identified in a natural manner, and the above matrices are symmetric.

When $\mathbf{N}$ is diagonal, say $\mathbf{N} = \mathrm{diag}(\nu_1, \nu_2, \ldots, \nu_n)$, then the tensor product structure of the full multidimensional Fourier transform (Section 1.3.2.4.2.4)

$$\mathscr{F}_{\mathbf{x}} = \mathscr{F}_{x_1} \otimes \mathscr{F}_{x_2} \otimes \ldots \otimes \mathscr{F}_{x_n}$$

gives rise to a tensor product structure for the DFT matrices. The tensor product of matrices is defined as follows:

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & \ldots & a_{1n}\mathbf{B} \\ \vdots & & \vdots \\ a_{n1}\mathbf{B} & \ldots & a_{nn}\mathbf{B} \end{pmatrix}.$$

Let the index vectors k and k^* be ordered in the same way as the elements in a Fortran array, *e.g.* for k with k_1 increasing fastest, k_2 next fastest, $\ldots, k_n$ slowest; then

$$\mathscr{F}(\mathbf{N}) = \mathscr{F}(\nu_1) \otimes \mathscr{F}(\nu_2) \otimes \ldots \otimes \mathscr{F}(\nu_n),$$

where

$$[\mathscr{F}(\nu_j)]_{k_j, k_j^*} = \frac{1}{\nu_j}\exp\left(-2\pi i \frac{k_j^* k_j}{\nu_j}\right),$$

and

$$\bar{\mathscr{F}}(\mathbf{N}) = \bar{\mathscr{F}}(\nu_1) \otimes \bar{\mathscr{F}}(\nu_2) \otimes \ldots \otimes \bar{\mathscr{F}}(\nu_n),$$

where

$$[\bar{\mathscr{F}}_{\nu_j}]_{k_j^*, k_j} = \exp\left(+2\pi i \frac{k_j^* k_j}{\nu_j}\right).$$

1.3.2.7.5. *Properties of the discrete Fourier transform*

The DFT inherits most of the properties of the Fourier transforms, but with certain numerical factors ('Jacobians') due to the transition from continuous to discrete measure.

(1) *Linearity* is obvious.

(2) *Shift property*. If $(\tau_a \psi)(k) = \psi(k - a)$ and $(\tau_{a^*}\Psi)(k^*) = \Psi(k^* - a^*)$, where subtraction takes place by modular vector arithmetic in $\mathbb{Z}^n/\mathbf{N}\mathbb{Z}^n$ and $\mathbb{Z}^n/\mathbf{N}^T\mathbb{Z}^n$, respectively, then the following identities hold:

$$\bar{\mathscr{F}}(\mathbf{N})[\tau_k \psi](k^*) = \exp[+2\pi i k^* \cdot (\mathbf{N}^{-1}k)]\bar{\mathscr{F}}(\mathbf{N})[\psi](k^*)$$
$$\mathscr{F}(\mathbf{N})[\tau_{k^*}\Psi](k) = \exp[-2\pi i k^* \cdot (\mathbf{N}^{-1}k)]\mathscr{F}(\mathbf{N})[\Psi](k).$$

(3) *Differentiation identities*. Let vectors $\boldsymbol{\psi}$ and $\boldsymbol{\Psi}$ be constructed from $\varphi^0 \in \mathscr{E}(\mathbb{R}^n)$ as in Section 1.3.2.7.3, hence be related by the DFT. If $D^{\mathbf{p}}\boldsymbol{\psi}$ designates the vector of sample values of $D^{\mathbf{p}}_{\mathbf{x}}\varphi^0$ at the points of $\Lambda_{\mathbf{B}}/\Lambda_{\mathbf{A}}$, and $D^{\mathbf{p}}\boldsymbol{\Psi}$ the vector of values of $D^{\mathbf{p}}_{\boldsymbol{\xi}}\Phi^0$ at points of $\Lambda^*_{\mathbf{A}}/\Lambda^*_{\mathbf{B}}$, then for all multi-indices $\mathbf{p} = (p_1, p_2, \ldots, p_n)$

$$(D^{\mathbf{p}}\boldsymbol{\psi})(k) = \bar{\mathscr{F}}(\mathbf{N})[(+2\pi i k^*)^{\mathbf{p}}\boldsymbol{\Psi}](k)$$
$$(D^{\mathbf{p}}\boldsymbol{\Psi})(k^*) = \mathscr{F}(\mathbf{N})[(-2\pi i k)^{\mathbf{p}}\boldsymbol{\psi}](k^*)$$

or equivalently

$$\mathscr{F}(\mathbf{N})[D^{\mathbf{p}}\boldsymbol{\psi}](k^*) = (+2\pi i k^*)^{\mathbf{p}}\boldsymbol{\Psi}(k^*)$$
$$\bar{\mathscr{F}}(\mathbf{N})[D^{\mathbf{p}}\boldsymbol{\Psi}](k) = (-2\pi i k)^{\mathbf{p}}\boldsymbol{\psi}(k).$$

(4) *Convolution property*. Let $\boldsymbol{\varphi} \in W_{\mathbf{N}}$ and $\boldsymbol{\Phi} \in W^*_{\mathbf{N}}$ (respectively $\boldsymbol{\psi}$ and $\boldsymbol{\Psi}$) be related by the DFT, and define

$$(\boldsymbol{\varphi} * \boldsymbol{\psi})(k) = \sum_{k' \in \mathbb{Z}^n/\mathbf{N}\mathbb{Z}^n} \boldsymbol{\varphi}(k')\boldsymbol{\psi}(k - k')$$
$$(\boldsymbol{\Phi} * \boldsymbol{\Psi})(k^*) = \sum_{k^{*\prime} \in \mathbb{Z}^n/\mathbf{N}^T\mathbb{Z}^n} \boldsymbol{\Phi}(k^{*\prime})\boldsymbol{\Psi}(k^* - k^{*\prime}).$$

Then

$$\bar{\mathscr{F}}(\mathbf{N})[\boldsymbol{\Phi} * \boldsymbol{\Psi}](k) = |\det \mathbf{N}|\boldsymbol{\varphi}(k)\boldsymbol{\psi}(k)$$
$$\mathscr{F}(\mathbf{N})[\boldsymbol{\varphi} * \boldsymbol{\psi}](k^*) = \boldsymbol{\Phi}(k^*)\boldsymbol{\Psi}(k^*)$$

and

$$\bar{\mathscr{F}}(\mathbf{N})[\boldsymbol{\varphi} \times \boldsymbol{\psi}](k^*) = \frac{1}{|\det \mathbf{N}|}(\boldsymbol{\Phi} * \boldsymbol{\Psi})(k^*)$$
$$\mathscr{F}(\mathbf{N})[\boldsymbol{\Phi} \times \boldsymbol{\Psi}](k) = (\boldsymbol{\varphi} * \boldsymbol{\psi})(k).$$

Since addition on $\mathbb{Z}^n/\mathbf{N}\mathbb{Z}^n$ and $\mathbb{Z}^n/\mathbf{N}^T\mathbb{Z}^n$ is modular, this type of convolution is called *cyclic* convolution.

(5) *Parseval/Plancherel property*. If $\boldsymbol{\varphi}$, $\boldsymbol{\psi}$, $\boldsymbol{\Phi}$, $\boldsymbol{\Psi}$ are as above, then

$$(\mathscr{F}(\mathbf{N})[\boldsymbol{\Phi}], \mathscr{F}(\mathbf{N})[\boldsymbol{\Psi}])_W = \frac{1}{|\det \mathbf{N}|}(\boldsymbol{\Phi}, \boldsymbol{\Psi})_{W^*}$$
$$(\bar{\mathscr{F}}(\mathbf{N})[\boldsymbol{\varphi}], \bar{\mathscr{F}}(\mathbf{N})[\boldsymbol{\psi}])_W = \frac{1}{|\det \mathbf{N}|}(\boldsymbol{\varphi}, \boldsymbol{\psi})_W.$$

(6) *Period* 4. When $\mathbf{N}$ is symmetric, so that the ranges of indices k and k^* can be identified, it makes sense to speak of powers of $\mathscr{F}(\mathbf{N})$ and $\bar{\mathscr{F}}(\mathbf{N})$. Then the 'standardized' matrices $(1/|\det \mathbf{N}|^{1/2})\mathscr{F}(\mathbf{N})$ and $(1/|\det \mathbf{N}|^{1/2})\bar{\mathscr{F}}(\mathbf{N})$ are *unitary* matrices whose fourth power is the identity matrix (Section 1.3.2.4.3.4); their eigenvalues are therefore ± 1 and $\pm i$.

1.3.3. Numerical computation of the discrete Fourier transform

1.3.3.1. *Introduction*

The Fourier transformation's most remarkable property is undoubtedly that of turning convolution into multiplication. As distribution theory has shown, other valuable properties – such as the shift property, the conversion of differentiation into multiplication by monomials, and the duality between periodicity and sampling – are special instances of the convolution theorem.

This property is exploited in many areas of applied mathematics and engineering (Campbell & Foster, 1948; Sneddon, 1951; Champeney, 1973; Bracewell, 1986). For example, the passing of a signal through a linear filter, which results in its being convolved with the response of the filter to a δ-function 'impulse', may be modelled as a multiplication of the signal's transform by the transform of the impulse response (also called transfer function). Similarly, the solution of systems of partial differential equations may be turned by Fourier transformation into a division problem for distributions. In both cases, the formulations obtained after Fourier transformation are considerably simpler than the initial ones, and lend themselves to constructive solution techniques.

B is a diagonal matrix of multiplications,

C is a matrix with entries 0, ± 1, $\pm i$, defining the 'post-additions'.

The elements on the diagonal of **B** can be shown to be either real or pure imaginary, by the same argument as in Section 1.3.3.2.3.1. Matrices **A** and **C** may be rectangular rather than square, so that intermediate results may require extra storage space.

1.3.3.3. *Multidimensional algorithms*

From an algorithmic point of view, the distinction between one-dimensional (1D) and multidimensional DFTs is somewhat blurred by the fact that some factoring techniques turn a 1D transform into a multidimensional one. The distinction made here, however, is a practical one and is based on the dimensionality of the indexing sets for data and results. This section will therefore be concerned with the problem of factoring the DFT when the *indexing sets* for the input data and output results are multidimensional.

1.3.3.3.1. *The method of successive one-dimensional transforms*

The DFT was defined in Section 1.3.2.7.4 in an n-dimensional setting and it was shown that when the decimation matrix **N** is diagonal, say $\mathbf{N} = \mathrm{diag}(N^{(1)}, N^{(2)}, \ldots, N^{(n)})$, then $\bar{F}(N)$ has a tensor product structure:

$$\bar{F}(\mathbf{N}) = \bar{F}(N^{(1)}) \otimes \bar{F}(N^{(2)}) \otimes \ldots \otimes \bar{F}(N^{(n)}).$$

This may be rewritten as follows:

$$\begin{aligned}\bar{F}(\mathbf{N}) = {} & [\bar{F}(N^{(1)}) \otimes I_{N^{(2)}} \otimes \ldots \otimes I_{N^{(n)}}] \\ & \times [I_{N^{(1)}} \otimes \bar{F}(N^{(2)}) \otimes \ldots \otimes I_{N^{(n)}}] \\ & \times \ldots \\ & \times [I_{N^{(1)}} \otimes I_{N^{(2)}} \otimes \ldots \otimes \bar{F}(N^{(n)}],\end{aligned}$$

where the I's are identity matrices and $\times$ denotes ordinary matrix multiplication. The matrix within each bracket represents a one-dimensional DFT along one of the n dimensions, the other dimensions being left untransformed. As these matrices commute, the order in which the successive 1D DFTs are performed is immaterial.

This is the most straightforward method for building an n-dimensional algorithm from existing 1D algorithms. It is known in crystallography under the name of 'Beevers–Lipson factorization' (Section 1.3.4.3.1), and in signal processing as the 'row–column method'.

1.3.3.3.2. *Multidimensional factorization*

Substantial reductions in the arithmetic cost, as well as gains in flexibility, can be obtained if the factoring of the DFT is carried out in several dimensions simultaneously. The presentation given here is a generalization of that of Mersereau & Speake (1981), using the abstract setting established independently by Auslander, Tolimieri & Winograd (1982).

Let us return to the general n-dimensional setting of Section 1.3.2.7.4, where the DFT was defined for an arbitrary decimation matrix **N** by the formulae (where $|\mathbf{N}|$ denotes $|\det \mathbf{N}|$):

$$F(\mathbf{N}): \quad X(\mathbf{k}) = \frac{1}{|\mathbf{N}|} \sum_{\mathbf{k}^*} X^*(\mathbf{k}^*) e[-\mathbf{k}^* \cdot (\mathbf{N}^{-1}\mathbf{k})]$$

$$\bar{F}(\mathbf{N}): \quad X^*(\mathbf{k}^*) = \sum_{\mathbf{k}} X(\mathbf{k}) e[\mathbf{k}^* \cdot (\mathbf{N}^{-1}\mathbf{k})]$$

with

$$\mathbf{k} \in \mathbb{Z}^n/\mathbf{N}\mathbb{Z}^n, \quad \mathbf{k}^* \in \mathbb{Z}^n/\mathbf{N}^T\mathbb{Z}^n.$$

1.3.3.3.2.1. *Multidimensional Cooley–Tukey factorization*

Let us now assume that this decimation can be factored into d successive decimations, *i.e.* that

$$\mathbf{N} = \mathbf{N}_1\mathbf{N}_2 \ldots \mathbf{N}_{d-1}\mathbf{N}_d$$

and hence

$$\mathbf{N}^T = \mathbf{N}_d^T\mathbf{N}_{d-1}^T \ldots \mathbf{N}_2^T\mathbf{N}_1^T.$$

Then the coset decomposition formulae corresponding to these successive decimations (Section 1.3.2.7.1) can be combined as follows:

$$\begin{aligned}\mathbb{Z}^n &= \bigcup_{\mathbf{k}_1} (\mathbf{k}_1 + \mathbf{N}_1\mathbb{Z}^n) \\ &= \bigcup_{\mathbf{k}_1} \left\{ \mathbf{k}_1 + \mathbf{N}_1 \left[\bigcup_{\mathbf{k}_2} (\mathbf{k}_2 + \mathbf{N}_2\mathbb{Z}^n) \right] \right\} \\ &= \ldots \\ &= \bigcup_{\mathbf{k}_1} \ldots \bigcup_{\mathbf{k}_d} (\mathbf{k}_1 + \mathbf{N}_1\mathbf{k}_2 + \ldots + \mathbf{N}_1\mathbf{N}_2 \times \ldots \times \mathbf{N}_{d-1}\mathbf{k}_d + \mathbf{N}\mathbb{Z}^n)\end{aligned}$$

with $\mathbf{k}_j \in \mathbb{Z}^n/\mathbf{N}_j\mathbb{Z}^n$. Therefore, any $\mathbf{k} \in \mathbb{Z}/\mathbf{N}\mathbb{Z}^n$ may be written uniquely as

$$\mathbf{k} = \mathbf{k}_1 + \mathbf{N}_1\mathbf{k}_2 + \ldots + \mathbf{N}_1\mathbf{N}_2 \times \ldots \times \mathbf{N}_{d-1}\mathbf{k}_d.$$

Similarly:

$$\begin{aligned}\mathbb{Z}^n &= \bigcup_{\mathbf{k}_d^*} (\mathbf{k}_d^* + \mathbf{N}_d^T\mathbb{Z}^n) \\ &= \ldots \\ &= \bigcup_{\mathbf{k}_d^*} \ldots \bigcup_{\mathbf{k}_1^*} (\mathbf{k}_d^* + \mathbf{N}_d^T\mathbf{k}_{d-1}^* + \ldots + \mathbf{N}_d^T \times \ldots \times \mathbf{N}_2^T\mathbf{k}_1^* \\ &\quad + \mathbf{N}^T\mathbb{Z}^n)\end{aligned}$$

so that any $\mathbf{k}^* \in \mathbb{Z}^n/\mathbf{N}^T\mathbb{Z}^n$ may be written uniquely as

$$\mathbf{k}^* = \mathbf{k}_d^* + \mathbf{N}_d^T\mathbf{k}_{d-1}^* + \ldots + \mathbf{N}_d^T \times \ldots \times \mathbf{N}_2^T\mathbf{k}_1^*$$

with $\mathbf{k}_j^* \in \mathbb{Z}^n/\mathbf{N}_j^T\mathbb{Z}^n$. These decompositions are the vector analogues of the multi-radix number representation systems used in the Cooley–Tukey factorization.

We may then write the definition of $\bar{F}(\mathbf{N})$ with $d = 2$ factors as

$$\begin{aligned}X^*(\mathbf{k}_2^* + \mathbf{N}_2^T\mathbf{k}_1^*) = {} & \sum_{\mathbf{k}_1} \sum_{\mathbf{k}_2} X(\mathbf{k}_1 + \mathbf{N}_1\mathbf{k}_2) \\ & \times e[(\mathbf{k}_2^{*T} + \mathbf{k}_1^{*T}\mathbf{N}_2)\mathbf{N}_2^{-1}\mathbf{N}_1^{-1}(\mathbf{k}_1 + \mathbf{N}_1\mathbf{k}_2)].\end{aligned}$$

The argument of $e(-)$ may be expanded as

$$\mathbf{k}_2^* \cdot (\mathbf{N}^{-1}\mathbf{k}_1) + \mathbf{k}_1^* \cdot (\mathbf{N}_1^{-1}\mathbf{k}_1) + \mathbf{k}_2^* \cdot (\mathbf{N}_2^{-1}\mathbf{k}_2) + \mathbf{k}_1^* \cdot \mathbf{k}_2.$$

The first summand may be recognized as a twiddle factor, the second and third as the kernels of $\bar{F}(\mathbf{N}_1)$ and $\bar{F}(\mathbf{N}_2)$, respectively, while the fourth is an integer which may be dropped. We are thus led to a 'vector-radix' version of the Cooley–Tukey algorithm, in which the successive decimations may be introduced in all n dimensions simultaneously by general integer matrices. The computation may be decomposed into five stages analogous to those of the one-dimensional algorithm of Section 1.3.3.2.1:

(i) form the $|\mathbf{N}_1|$ vectors $\mathbf{Y}_{\mathbf{k}_1}$ of shape $\mathbf{N}_2$ by

$$Y_{\mathbf{k}_1}(\mathbf{k}_2) = X(\mathbf{k}_1 + \mathbf{N}_1\mathbf{k}_2), \quad \mathbf{k}_1 \in \mathbb{Z}^n/\mathbf{N}_1\mathbb{Z}^n, \quad \mathbf{k}_2 \in \mathbb{Z}^n/\mathbf{N}_2\mathbb{Z}^n;$$

Table A1.4.3.3. *Monoclinic space groups*

Each expression for A or B in the monoclinic system and for the space-group settings chosen in *IT* A is represented in terms of one of the following symbols:

$$\begin{aligned}
&\mathrm{c}(hl)\mathrm{c}(ky) = \cos[2\pi(hx+lz)]\cos(2\pi ky), \quad && \mathrm{c}(hk)\mathrm{c}(lz) = \cos[2\pi(hx+ky)]\cos(2\pi lz),\\
&\mathrm{c}(hl)\mathrm{s}(ky) = \cos[2\pi(hx+lz)]\sin(2\pi ky), \quad && \mathrm{c}(hk)\mathrm{s}(lz) = \cos[2\pi(hx+ky)]\sin(2\pi lz),\\
&\mathrm{s}(hl)\mathrm{c}(ky) = \sin[2\pi(hx+lz)]\cos(2\pi ky), \quad && \mathrm{s}(hk)\mathrm{c}(lz) = \sin[2\pi(hx+ky)]\cos(2\pi lz),\\
&\mathrm{s}(hl)\mathrm{s}(ky) = \sin[2\pi(hx+lz)]\sin(2\pi ky), \quad && \mathrm{s}(hk)\mathrm{s}(lz) = \sin[2\pi(hx+ky)]\sin(2\pi lz),
\end{aligned} \tag{A1.4.3.1}$$

where the left-hand column of expressions corresponds to space-group representations in the second setting, with b taken as the unique axis, and the right-hand column corresponds to representations in the first setting, with c taken as the unique axis.

The lattice types in this table are P, A, B, C and I, and are all explicit in the full space-group symbol only (see below). Note that s(hl), s(hk), s(ky) and s(lz) are zero for $h = l = 0$, $h = k = 0$, $k = 0$ and $l = 0$, respectively.

No.	Group symbol		Parity	A	B	Unique axis
	Short	Full				
3	$P2$	$P121$		2c(hl)c(ky)	2c(hl)s(ky)	b
3	$P2$	$P112$		2c(hk)c(lz)	2c(hk)s(lz)	c
4	$P2_1$	$P12_11$	$k = 2n$	2c(hl)c(ky)	2c(hl)s(ky)	b
			$k = 2n+1$	−2s(hl)s(ky)	2s(hl)c(ky)	
4	$P2_1$	$P112_1$	$l = 2n$	2c(hk)c(lz)	2c(hk)s(lz)	c
			$l = 2n+1$	−2s(hk)s(lz)	2s(hk)c(lz)	
5	$C2$	$C121$		4c(hl)c(ky)	4c(hl)s(ky)	b
5	$C2$	$A121$		4c(hl)c(ky)	4c(hl)s(ky)	b
5	$C2$	$I121$		4c(hl)c(ky)	4c(hl)s(ky)	b
5	$C2$	$A112$		4c(hk)c(lz)	4c(hk)s(lz)	c
5	$C2$	$B112$		4c(hk)c(lz)	4c(hk)s(lz)	c
5	$C2$	$I112$		4c(hk)c(lz)	4c(hk)s(lz)	c
6	Pm	$P1m1$		2c(hl)c(ky)	2s(hl)c(ky)	b
6	Pm	$P11m$		2c(hk)c(lz)	2s(hk)c(lz)	c
7	Pc	$P1c1$	$l = 2n$	2c(hl)c(ky)	2s(hl)c(ky)	b
			$l = 2n+1$	−2s(hl)s(ky)	2c(hl)s(ky)	
7	Pc	$P1n1$	$h+l = 2n$	2c(hl)c(ky)	2s(hl)c(ky)	b
			$h+l = 2n+1$	−2s(hl)s(ky)	2c(hl)s(ky)	
7	Pc	$P1a1$	$h = 2n$	2c(hl)c(ky)	2s(hl)c(ky)	b
			$h = 2n+1$	−2s(hl)s(ky)	2c(hl)s(ky)	
7	Pc	$P11a$	$h = 2n$	2c(hk)c(lz)	2s(hk)c(lz)	c
			$h = 2n+1$	−2s(hk)s(lz)	2c(hk)s(lz)	
7	Pc	$P11n$	$h+k = 2n$	2c(hk)c(lz)	2s(hk)c(lz)	c
			$h+k = 2n+1$	−2s(hk)s(lz)	2c(hk)s(lz)	
7	Pc	$P11b$	$k = 2n$	2c(hk)c(lz)	2s(hk)c(lz)	c
			$k = 2n+1$	−2s(hk)s(lz)	2c(hk)s(lz)	
8	Cm	$C1m1$		4c(hl)c(ky)	4s(hl)c(ky)	b
8	Cm	$A1m1$		4c(hl)c(ky)	4s(hl)c(ky)	b
8	Cm	$I1m1$		4c(hl)c(ky)	4s(hl)c(ky)	b
8	Cm	$A11m$		4c(hk)c(lz)	4s(hk)c(lz)	c
8	Cm	$B11m$		4c(hk)c(lz)	4s(hk)c(lz)	c
8	Cm	$I11m$		4c(hk)c(lz)	4s(hk)c(lz)	c
9	Cc	$C1c1$	$l = 2n$	4c(hl)c(ky)	4s(hl)c(ky)	b
			$l = 2n+1$	−4s(hl)s(ky)	4c(hl)s(ky)	
9	Cc	$A1n1$	$h+l = 2n$	4c(hl)c(ky)	4s(hl)c(ky)	b
			$h+l = 2n+1$	−4s(hl)s(ky)	4c(hl)s(ky)	
9	Cc	$I1a1$	$h = 2n$	4c(hl)c(ky)	4s(hl)c(ky)	b
			$h = 2n+1$	−4s(hl)s(ky)	4c(hl)s(ky)	
9	Cc	$A11a$	$h = 2n$	4c(hk)c(lz)	4s(hk)c(lz)	c
			$h = 2n+1$	−4s(hk)s(lz)	4c(hk)s(lz)	
9	Cc	$B11n$	$h+k = 2n$	4c(hk)c(lz)	4s(hk)c(lz)	c
			$h+k = 2n+1$	−4s(hk)s(lz)	4c(hk)s(lz)	
9	Cc	$I11b$	$k = 2n$	4c(hk)c(lz)	4s(hk)c(lz)	c
			$k = 2n+1$	−4s(hk)s(lz)	4c(hk)s(lz)	
10	$P2/m$	$P12/m1$		4c(hl)c(ky)	0	b

Table A1.4.4.1. *Crystallographic space groups in reciprocal space* (*cont.*)

$I4_1md$ No. 109 (162)

(1) hkl:	(2) $\bar{h}\bar{k}l$: $-111/2$	(3) $k\bar{h}l$: $-021/4$	(4) $\bar{k}hl$: $-203/4$
(5) $h\bar{k}l$:	(6) $\bar{h}kl$: $-111/2$	(7) $\bar{k}\bar{h}l$: $-203/4$	(8) khl: $-021/4$

$I4_1cd$ No. 110 (163)

(1) hkl:	(2) $\bar{h}\bar{k}l$: $-111/2$	(3) $k\bar{h}l$: $-021/4$	(4) $\bar{k}hl$: $-203/4$
(5) $h\bar{k}l$: $-001/2$	(6) $\bar{h}kl$: $-110/2$	(7) $\bar{k}\bar{h}l$: $-201/4$	(8) khl: $-023/4$

Point group: $\bar{4}2m$ Tetragonal Laue group: $4/mmm$

$P\bar{4}2m$ No. 111 (164)

(1) hkl:	(2) $\bar{h}\bar{k}l$:	(3) $k\bar{h}\bar{l}$:	(4) $k\bar{h}l$:
(5) $\bar{h}k\bar{l}$:	(6) $h\bar{k}l$:	(7) $\bar{k}hl$:	(8) khl:

$P\bar{4}2c$ No. 112 (165)

(1) hkl:	(2) $\bar{h}\bar{k}l$:	(3) $k\bar{h}\bar{l}$:	(4) $k\bar{h}l$:
(5) $\bar{h}k\bar{l}$: $-001/2$	(6) $h\bar{k}l$: $-001/2$	(7) $\bar{k}hl$: $-001/2$	(8) khl: $-001/2$

$P\bar{4}2_1m$ No. 113 (166)

(1) hkl:	(2) $\bar{h}\bar{k}l$:	(3) $k\bar{h}\bar{l}$:	(4) $k\bar{h}l$:
(5) $\bar{h}k\bar{l}$: $-110/2$	(6) $h\bar{k}l$: $-110/2$	(7) $\bar{k}hl$: $-110/2$	(8) khl: $-110/2$

$P\bar{4}2_1c$ No. 114 (167)

(1) hkl:	(2) $\bar{h}\bar{k}l$:	(3) $k\bar{h}\bar{l}$:	(4) $k\bar{h}l$:
(5) $\bar{h}k\bar{l}$: $-111/2$	(6) $h\bar{k}l$: $-111/2$	(7) $\bar{k}hl$: $-111/2$	(8) khl: $-111/2$

$P\bar{4}m2$ No. 115 (168)

(1) hkl:	(2) $\bar{h}\bar{k}l$:	(3) $\bar{k}h\bar{l}$:	(4) $k\bar{h}l$:
(5) $h\bar{k}l$:	(6) $\bar{h}kl$:	(7) $kh\bar{l}$:	(8) $\bar{k}hl$:

$P\bar{4}c2$ No. 116 (169)

(1) hkl:	(2) $\bar{h}\bar{k}l$:	(3) $\bar{k}h\bar{l}$:	(4) $k\bar{h}l$:
(5) $h\bar{k}l$: $-001/2$	(6) $\bar{h}kl$: $-001/2$	(7) $kh\bar{l}$: $-001/2$	(8) $\bar{k}hl$: $-001/2$

$P\bar{4}b2$ No. 117 (170)

(1) hkl:	(2) $\bar{h}\bar{k}l$:	(3) $\bar{k}h\bar{l}$:	(4) $k\bar{h}l$:
(5) $h\bar{k}l$: $-110/2$	(6) $\bar{h}kl$: $-110/2$	(7) $kh\bar{l}$: $-110/2$	(8) $\bar{k}hl$: $-110/2$

$P\bar{4}n2$ No. 118 (171)

(1) hkl:	(2) $\bar{h}\bar{k}l$:	(3) $\bar{k}h\bar{l}$:	(4) $k\bar{h}l$:
(5) $h\bar{k}l$: $-111/2$	(6) $\bar{h}kl$: $-111/2$	(7) $kh\bar{l}$: $-111/2$	(8) $\bar{k}hl$: $-111/2$

$I\bar{4}m2$ No. 119 (172)

(1) hkl:	(2) $\bar{h}\bar{k}l$:	(3) $\bar{k}h\bar{l}$:	(4) $k\bar{h}l$:
(5) $h\bar{k}l$:	(6) $\bar{h}kl$:	(7) $kh\bar{l}$:	(8) $\bar{k}hl$:

$I\bar{4}c2$ No. 120 (173)

(1) hkl:	(2) $\bar{h}\bar{k}l$:	(3) $\bar{k}h\bar{l}$:	(4) $k\bar{h}l$:
(5) $h\bar{k}l$: $-001/2$	(6) $\bar{h}kl$: $-001/2$	(7) $kh\bar{l}$: $-001/2$	(8) $\bar{k}hl$: $-001/2$

$I\bar{4}2m$ No. 121 (174)

(1) hkl:	(2) $\bar{h}\bar{k}l$:	(3) $\bar{k}h\bar{l}$:	(4) $k\bar{h}l$:
(5) $\bar{h}k\bar{l}$:	(6) $h\bar{k}l$:	(7) $\bar{k}hl$:	(8) khl:

$I\bar{4}2d$ No. 122 (175)

(1) hkl:	(2) $\bar{h}\bar{k}l$:	(3) $\bar{k}h\bar{l}$:	(4) $k\bar{h}l$:
(5) $\bar{h}k\bar{l}$: $-203/4$	(6) $h\bar{k}l$: $-203/4$	(7) $\bar{k}hl$: $-021/4$	(8) khl: $-021/4$

Point group: $4/mmm$ Tetragonal Laue group: $4/mmm$

$P4/mmm$ No. 123 (176)

(1) hkl:	(2) $\bar{h}\bar{k}l$:	(3) $k\bar{h}\bar{l}$:	(4) $\bar{k}hl$:
(5) $\bar{h}k\bar{l}$:	(6) $h\bar{k}\bar{l}$:	(7) $kh\bar{l}$:	(8) $\bar{k}\bar{h}\bar{l}$:

$P4/mcc$ No. 124 (177)

(1) hkl:	(2) $\bar{h}\bar{k}l$:	(3) $k\bar{h}\bar{l}$:	(4) $\bar{k}hl$:
(5) $\bar{h}k\bar{l}$: $-001/2$	(6) $h\bar{k}\bar{l}$: $-001/2$	(7) $kh\bar{l}$: $-001/2$	(8) $\bar{k}\bar{h}\bar{l}$: $-001/2$

$P4/nbm$ Origin 1 No. 125 (178)

(1) hkl:	(2) $\bar{h}\bar{k}l$:	(3) $k\bar{h}\bar{l}$:	(4) $\bar{k}hl$:
(5) $\bar{h}k\bar{l}$:	(6) $h\bar{k}\bar{l}$:	(7) $kh\bar{l}$:	(8) $\bar{k}\bar{h}\bar{l}$:
(9) $\bar{h}\bar{k}l$: $-110/2$	(10) $hk\bar{l}$: $-110/2$	(11) $\bar{k}h\bar{l}$: $-110/2$	(12) $k\bar{h}\bar{l}$: $-110/2$
(13) $h\bar{k}l$: $-110/2$	(14) $\bar{h}kl$: $-110/2$	(15) $\bar{k}\bar{h}l$: $-110/2$	(16) khl: $-110/2$

$P4/nbm$ Origin 2 No. 125 (179)

(1) hkl:	(2) $\bar{h}\bar{k}l$: $-110/2$	(3) $k\bar{h}\bar{l}$: $-100/2$	(4) $\bar{k}hl$: $-010/2$
(5) $\bar{h}k\bar{l}$: $-100/2$	(6) $h\bar{k}\bar{l}$: $-010/2$	(7) $kh\bar{l}$:	(8) $\bar{k}\bar{h}\bar{l}$: $-110/2$

$P4/nnc$ Origin 1 No. 126 (180)

(1) hkl:	(2) $\bar{h}\bar{k}l$:	(3) $k\bar{h}\bar{l}$:	(4) $\bar{k}hl$:
(5) $\bar{h}k\bar{l}$:	(6) $h\bar{k}\bar{l}$:	(7) $kh\bar{l}$:	(8) $\bar{k}\bar{h}\bar{l}$:
(9) $\bar{h}\bar{k}l$: $-111/2$	(10) $hk\bar{l}$: $-111/2$	(11) $\bar{k}h\bar{l}$: $-111/2$	(12) $k\bar{h}\bar{l}$: $-111/2$
(13) $h\bar{k}l$: $-111/2$	(14) $\bar{h}kl$: $-111/2$	(15) $\bar{k}\bar{h}l$: $-111/2$	(16) khl: $-111/2$

$P4/nnc$ Origin 2 No. 126 (181)

(1) hkl:	(2) $\bar{h}\bar{k}l$: $-110/2$	(3) $k\bar{h}\bar{l}$: $-100/2$	(4) $\bar{k}hl$: $-010/2$
(5) $\bar{h}k\bar{l}$: $-101/2$	(6) $h\bar{k}\bar{l}$: $-011/2$	(7) $kh\bar{l}$: $-001/2$	(8) $\bar{k}\bar{h}\bar{l}$: $-111/2$

$P4/mbm$ No. 127 (182)

(1) hkl:	(2) $\bar{h}\bar{k}l$:	(3) $k\bar{h}\bar{l}$:	(4) $\bar{k}hl$:
(5) $\bar{h}k\bar{l}$: $-110/2$	(6) $h\bar{k}\bar{l}$: $-110/2$	(7) $kh\bar{l}$: $-110/2$	(8) $\bar{k}\bar{h}\bar{l}$: $-110/2$

$P4/mnc$ No. 128 (183)

(1) hkl:	(2) $\bar{h}\bar{k}l$:	(3) $k\bar{h}\bar{l}$:	(4) $\bar{k}hl$:
(5) $\bar{h}k\bar{l}$: $-111/2$	(6) $h\bar{k}\bar{l}$: $-111/2$	(7) $kh\bar{l}$: $-111/2$	(8) $\bar{k}\bar{h}\bar{l}$: $-111/2$

$P4/nmm$ Origin 1 No. 129 (184)

(1) hkl:	(2) $\bar{h}\bar{k}l$:	(3) $k\bar{h}\bar{l}$: $-110/2$	(4) $\bar{k}hl$: $-110/2$
(5) $\bar{h}k\bar{l}$: $-110/2$	(6) $h\bar{k}\bar{l}$: $-110/2$	(7) $kh\bar{l}$:	(8) $\bar{k}\bar{h}\bar{l}$:
(9) $\bar{h}\bar{k}l$: $-110/2$	(10) $hk\bar{l}$: $-110/2$	(11) $\bar{k}h\bar{l}$:	(12) $k\bar{h}\bar{l}$:
(13) $h\bar{k}l$:	(14) $\bar{h}kl$:	(15) $\bar{k}\bar{h}l$: $-110/2$	(16) khl: $-110/2$

$P4/nmm$ Origin 2 No. 129 (185)

(1) hkl:	(2) $\bar{h}\bar{k}l$: $-110/2$	(3) $k\bar{h}\bar{l}$: $-100/2$	(4) $\bar{k}hl$: $-010/2$
(5) $\bar{h}k\bar{l}$: $-010/2$	(6) $h\bar{k}\bar{l}$: $-100/2$	(7) $kh\bar{l}$: $-110/2$	(8) $\bar{k}\bar{h}\bar{l}$:

$P4/ncc$ Origin 1 No. 130 (186)

(1) hkl:	(2) $\bar{h}\bar{k}l$:	(3) $k\bar{h}\bar{l}$: $-110/2$	(4) $\bar{k}hl$: $-110/2$
(5) $\bar{h}k\bar{l}$: $-111/2$	(6) $h\bar{k}\bar{l}$: $-111/2$	(7) $kh\bar{l}$: $-001/2$	(8) $\bar{k}\bar{h}\bar{l}$: $-001/2$
(9) $\bar{h}\bar{k}l$: $-110/2$	(10) $hk\bar{l}$: $-110/2$	(11) $\bar{k}h\bar{l}$:	(12) $k\bar{h}\bar{l}$:
(13) $h\bar{k}l$: $-001/2$	(14) $\bar{h}kl$: $-001/2$	(15) $\bar{k}\bar{h}l$: $-111/2$	(16) khl: $-111/2$

$P4/ncc$ Origin 2 No. 130 (187)

(1) hkl:	(2) $\bar{h}\bar{k}l$: $-110/2$	(3) $k\bar{h}\bar{l}$: $-100/2$	(4) $\bar{k}hl$: $-010/2$
(5) $\bar{h}k\bar{l}$: $-011/2$	(6) $h\bar{k}\bar{l}$: $-101/2$	(7) $kh\bar{l}$: $-111/2$	(8) $\bar{k}\bar{h}\bar{l}$: $-001/2$

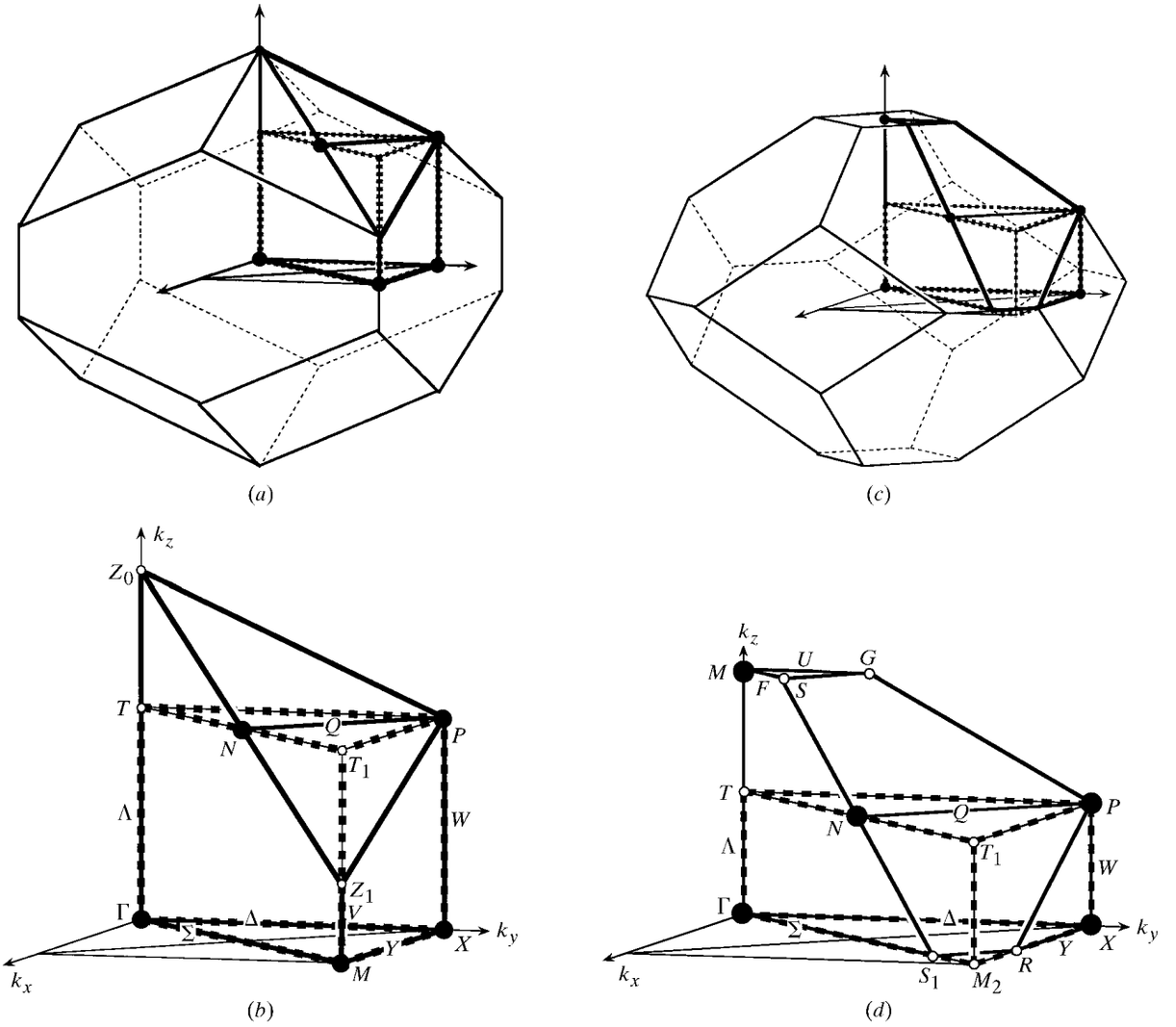

Fig. 1.5.5.3. (*a*), (*b*). Symmorphic space group $I4/mmm$ (isomorphic to the reciprocal-space group $\mathcal{G}^*$ of $4/mmmI$). Diagrams for $a > c$, *i.e.* $c^* > a^*$. In the figures $a = 1.25c$, *i.e.* $c^* = 1.25a^*$. (*a*) Representation domain (thick lines) and asymmetric unit (thick dashed lines, partly protruding) imbedded in the Brillouin zone, which is a tetragonal elongated rhombdodecahedron. (*b*) Representation domain ΓMXZ_1PZ_0 and asymmetric unit $\Gamma MXTT_1P$ of $I4/mmm$, *IT* A, p. 468. The part $\Gamma MXTNZ_1P$ is common to both bodies; the part $TNPZ_0$ is equivalent to the part NZ_1PT_1 by a twofold rotation around the axis $Q = NP$. Coordinates of the *points*: $\Gamma = 0,0,0$; $X = 0, \frac{1}{2}, 0$; $M = \frac{1}{2}, \frac{1}{2}, 0$; $P = 0, \frac{1}{2}, \frac{1}{4}$; $N = \frac{1}{4}, \frac{1}{4}, \frac{1}{4}$; $T = 0,0, \frac{1}{4} \sim T_1 = \frac{1}{2}, \frac{1}{2}, \frac{1}{4}$; $Z_0 = 0,0,z_0 \sim Z_1 = \frac{1}{2}, \frac{1}{2}, z_1$ with $z_0 = [1 + (c/a)^2]/4$; $z_1 = \frac{1}{2} - z_0$; the sign $\sim$ means symmetrically equivalent. *Lines*: $\Lambda = \Gamma Z_0 = 0,0,z$; $V = Z_1M = \frac{1}{2}, \frac{1}{2}, z$; $W = XP = 0, \frac{1}{2}, z$; $\Sigma = \Gamma M = x,x,0$; $\Delta = \Gamma X = 0,y,0$; $Y = XM = x, \frac{1}{2}, 0$; $Q = PN = x, \frac{1}{2} - x, \frac{1}{4}$. The lines Z_0Z_1, Z_1P and PZ_0 have no special symmetry but belong to special planes. *Planes*: $C = \Gamma MX = x,y,0$; $B = \Gamma Z_0Z_1M = x,x,z$; $A = \Gamma XPZ_0 = 0,y,z$; $E = MXPZ_1 = x, \frac{1}{2}, z$. The plane Z_0Z_1P belongs to the general position *GP*. Large black circles: special points belonging to the representation domain; small open circles: $T \sim T_1$ and $Z_0 \sim Z_1$ belonging to special lines; thick lines: edges of the representation domain and special line $Q = NP$; dashed lines: edges of the asymmetric unit. For the parameter ranges see Table 1.5.5.3.

(*c*), (*d*). Symmorphic space group $I4/mmm$ (isomorphic to the reciprocal-space group $\mathcal{G}^*$ of $4/mmmI$). Diagrams for $c > a$, *i.e.* $a^* > c^*$. In the figures $c = 1.25a$, *i.e.* $a^* = 1.25c^*$. (*c*) Representation domain (thick lines) and asymmetric unit (dashed lines, partly protruding) imbedded in the Brillouin zone, which is a tetragonal cuboctahedron. (*d*) Representation domain $\Gamma S_1RXPMSG$ and asymmetric unit ΓM_2XTT_1P of $I4/mmm$, *IT* A, p. 468. The part ΓS_1RXTNP is common to both bodies; the part $TNPMSG$ is equivalent to the part $T_1NPM_2S_1R$ by a twofold rotation around the axis $Q = NP$. Coordinates of the *points*: $\Gamma = 0,0,0$; $X = 0, \frac{1}{2}, 0$; $N = \frac{1}{4}, \frac{1}{4}, \frac{1}{4}$; $M = 0,0, \frac{1}{2} \sim M_2 = \frac{1}{2}, \frac{1}{2}, 0$; $T = 0,0, \frac{1}{4} \sim T_1 = \frac{1}{2}, \frac{1}{2}, \frac{1}{4}$; $P = 0, \frac{1}{2}, \frac{1}{4}$; $S = s,s, \frac{1}{2} \sim S_1 = s_1,s_1,0$ with $s = [1 - (a/c)^2]/4$; $s_1 = \frac{1}{2} - s$; $R = r, \frac{1}{2}, 0 \sim G = 0,g, \frac{1}{2}$ with $r = (a/c)^2/2$; $g = \frac{1}{2} - r$; the sign $\sim$ means symmetrically equivalent. *Lines*: $\Lambda = \Gamma M = 0,0,z$; $W = XP = 0, \frac{1}{2}, z$; $\Sigma = \Gamma S_1 = x,x,0$; $F = MS = x,x, \frac{1}{2}$; $\Delta = \Gamma X = 0,y,0$; $Y = XR = x, \frac{1}{2}, 0$; $U = MG = 0,y, \frac{1}{2}$; $Q = PN = x, \frac{1}{2} - x, \frac{1}{4}$. The lines $GS \sim S_1R$, $SN \sim NS_1$ and $GP \sim PR$ have no special symmetry but belong to special planes. *Planes*: $C = \Gamma S_1RX = x,y,0$; $D = MSG = x,y, \frac{1}{2}$; $B = \Gamma S_1SM = x,x,z$; $A = \Gamma XPGM = 0,y,z$; $E = RXP = x, \frac{1}{2}, z$. The plane S_1RPGS belongs to the general position *GP*. Large black circles: special points belonging to the representation domain; small open circles: $M_2 \sim M$; the points $T \sim T_1$, $S \sim S_1$ and $G \sim R$ belong to special lines; thick lines: edges of the representation domain and special line $Q = NP$; dashed lines: edges of the asymmetric unit. For the parameter ranges see Table 1.5.5.3.

notorious for its rather indeterminate mean and 'infinite' variance, resulting from the 'tail' of the denominator distributions extending through zero to negative values. The leading terms of the ratio distribution are given by Kendall & Stuart (1977, p. 288).

2.1.7. Non-ideal distributions: the correction-factor approach

2.1.7.1. *Introduction*

The probability density functions (p.d.f.'s) of the magnitude of the structure factor, presented in Section 2.1.5, are based on the central-limit theorem discussed above. In particular, the centric and acentric p.d.f.'s given by equations (2.1.5.11) and (2.1.5.8), respectively, are expected to account for the statistical properties of diffraction patterns obtained from crystals consisting of nearly equal atoms, which obey the fundamental assumptions of uniformity and independence of the atomic contributions and are not affected by noncrystallographic symmetry and dispersion. It is also assumed there that the number of atoms in the asymmetric unit is large. Distributions of structure-factor magnitudes which are based on the central-limit theorem, and thus obey the above assumptions, have been termed 'ideal', and the subjects of the following sections are those distributions for which some of the above assumptions/restrictions are not fulfilled; the latter distributions will be called 'non-ideal'.

We recall that the assumption of uniformity consists of the requirement that the fractional part of the scalar product $hx + ky + lz$ be uniformly distributed over the [0, 1] interval, which holds well if x, y, z are rationally independent (Hauptman & Karle, 1953), and permits one to regard the atomic contribution to the structure factor as a random variable. This is of course a necessary requirement for any statistical treatment. If, however, the atomic composition of the asymmetric unit is widely heterogeneous, the structure factor is then a sum of unequally distributed random variables and the Lindeberg–Lévy version of the central-limit theorem (*cf.* Section 2.1.4.4) cannot be expected to apply. Other versions of this theorem might still predict a normal p.d.f. of the sum, but at the expense of a correspondingly large number of terms/atoms. It is well known that atomic heterogeneity gives rise to severe deviations from ideal behaviour (*e.g.* Howells *et al.*, 1950) and one of the aims of crystallographic statistics has been the introduction of a correct dependence on the atomic composition into the non-ideal p.d.f.'s [for a review of the early work on non-ideal distributions see Srinivasan & Parthasarathy (1976)]. A somewhat less well known fact is that the dependence of the p.d.f.'s of $|E|$ on space-group symmetry becomes more conspicuous as the composition becomes more heterogeneous (*e.g.* Shmueli, 1979; Shmueli & Wilson, 1981). Hence both the composition and the symmetry dependence of the intensity statistics are of interest. Other problems, which likewise give rise to non-ideal p.d.f.'s, are the presence of heavy atoms in (variable) special positions, heterogeneous structures with complete or partial noncrystallographic symmetry, and the presence of outstandingly heavy dispersive scatterers.

The need for theoretical representations of non-ideal p.d.f.'s is exemplified in Fig. 2.1.7.1(*a*), which shows the ideal centric and acentric p.d.f.'s together with a frequency histogram of $|E|$ values, recalculated for a centrosymmetric structure containing a platinum atom in the asymmetric unit of $P\bar{1}$ (Faggiani *et al.*, 1980). Clearly, the deviation from the Gaussian p.d.f., predicted by the central-limit theorem, is here very large and a comparison with the possible ideal distributions can (in this case) lead to wrong conclusions.

Two general approaches have so far been employed in derivations of non-ideal p.d.f.'s which account for the above-mentioned problems: the correction-factor approach, to be dealt with in the following sections, and the more recently introduced Fourier method, to which Section 2.1.8 is dedicated. In what follows, we introduce briefly the mathematical background of the correction-factor approach, apply this formalism to centric and acentric non-ideal p.d.f.'s, and present the numerical values of the moments of the trigonometric structure factor which permit an approximate evaluation of such p.d.f.'s for all the three-dimensional space groups.

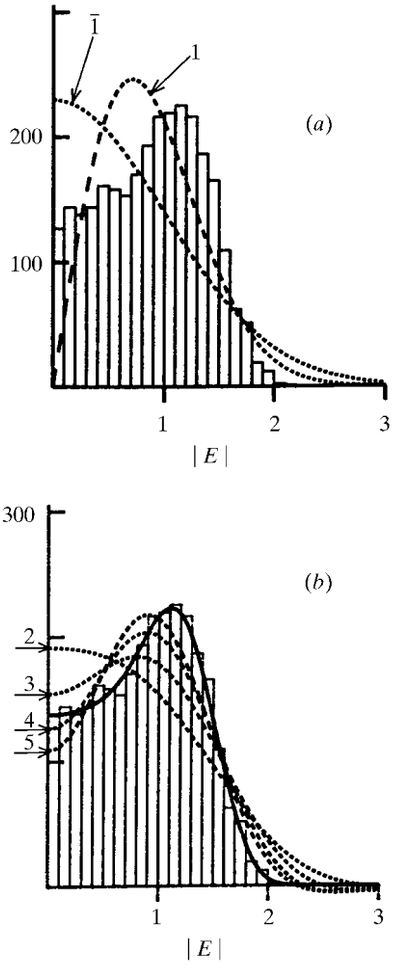

Fig. 2.1.7.1. Atomic heterogeneity and intensity statistics. The histogram appearing in (*a*) and (*b*) was constructed from $|E|$ values which were recalculated from atomic parameters published for the centrosymmetric structure of $C_6H_{18}Cl_2N_4O_4Pt$ (Faggiani *et al.*, 1980). The space group of the crystal is $P\bar{1}$, $Z = 2$, *i.e.* all the atoms are located in general positions. (*a*) A comparison of the recalculated distribution of $|E|$ with the ideal centric [equation (2.1.5.11)] and acentric [equation (2.1.5.8)] p.d.f.'s, denoted by $\bar{1}$ and 1, respectively. (*b*) The same recalculated histogram along with the centric correction-factor p.d.f. [equation (2.1.7.5)], truncated after two, three, four and five terms (dashed lines), and with that accurately computed for the correct space-group Fourier p.d.f. [equations (2.1.8.5) and (2.1.8.22)] (solid line).

2.1.7.2. *Mathematical background*

Suppose that $p(x)$ is a p.d.f. which accurately describes the experimental distribution of the random variable x, where x is related to a sum of random variables and can be assumed to obey (to some approximation) an ideal p.d.f., say $p^{(0)}(x)$, based on the central-limit theorem. In the correction-factor approach we seek to represent $p(x)$ as

2.2. Direct methods

BY C. GIACOVAZZO

2.2.1. List of symbols and abbreviations

f_j atomic scattering factor of jth atom
Z_j atomic number of jth atom
N number of atoms in the unit cell
m order of the point group

$$[\sigma_r]_p, [\sigma_r]_q, [\sigma_r]_N, \ldots = \sum_{j=1}^{p} Z_j^r, \sum_{j=1}^{q} Z_j^r, \sum_{j=1}^{N} Z_j^r, \ldots$$

$[\sigma_r]_N$ is always abbreviated to σ_r when N is the number of atoms in the cell

$$\sum_p, \sum_q, \sum_N, \ldots = \sum_{j=1}^{p} f_j^2, \sum_{j=1}^{q} f_j^2, \sum_{j=1}^{N} f_j^2, \ldots$$

s.f. structure factor
n.s.f. normalized structure factor
cs. centrosymmetric
ncs. noncentrosymmetric
s.i. structure invariant
s.s. structure seminvariant
$\mathbf{C} = (\mathbf{R}, \mathbf{T})$ symmetry operator; $\mathbf{R}$ is the rotational part, $\mathbf{T}$ the translational part
$\varphi_{\mathbf{h}}$ phase of the structure factor $F_{\mathbf{h}} = |F_{\mathbf{h}}| \exp(i\varphi_{\mathbf{h}})$

2.2.2. Introduction

Direct methods are today the most widely used tool for solving small crystal structures. They work well both for equal-atom molecules and when a few heavy atoms exist in the structure. In recent years the theoretical background of direct methods has been improved to take into account a large variety of prior information (the form of the molecule, its orientation, a partial structure, the presence of pseudosymmetry or of a superstructure, the availability of isomorphous data or of data affected by anomalous-dispersion effects, ...). Owing to this progress and to the increasing availability of powerful computers, a number of effective, highly automated packages for the practical solution of the phase problem are today available to the scientific community.

The *ab initio* crystal structure solution of macromolecules seems not to exceed the potential of direct methods. Many efforts will certainly be devoted to this task in the near future: a report of the first achievements is given in Section 2.2.10.

This chapter describes both the traditional direct methods tools and the most recent and revolutionary techniques suitable for macromolecules.

The theoretical background and tables useful for origin specification are given in Section 2.2.3; in Section 2.2.4 the procedures for normalizing structure factors are summarized. Phase-determining formulae (inequalities, probabilistic formulae for triplet, quartet and quintet invariants, and for one- and two-phase s.s.'s, determinantal formulae) are given in Section 2.2.5. In Section 2.2.6 the connection between direct methods and related techniques in real space is discussed. Practical procedures for solving crystal structures are described in Sections 2.2.7 and 2.2.8, and references to the most extensively used packages are given in Section 2.2.9. The techniques suitable for the *ab initio* crystal structure solution of macromolecules are described in Section 2.2.10.2. The integration of direct methods with isomorphous-replacement and anomalous-dispersion techniques is briefly described in Sections 2.2.10.3 and 2.2.10.4.

The reader will find full coverage of the most important aspects of direct methods in the recent books by Giacovazzo (1998) and Woolfson & Fan (1995).

2.2.3. Origin specification

(*a*) Once the origin has been chosen, the symmetry operators $\mathbf{C}_s \equiv (\mathbf{R}_s, \mathbf{T}_s)$ and, through them, the algebraic form of the s.f. remain fixed.

A shift of the origin through a vector with coordinates $\mathbf{X}_0$ transforms $\varphi_{\mathbf{h}}$ into

$$\varphi'_{\mathbf{h}} = \varphi_{\mathbf{h}} - 2\pi \mathbf{h} \cdot \mathbf{X}_0 \qquad (2.2.3.1)$$

and the symmetry operators $\mathbf{C}_s$ into $\mathbf{C}'_s = (\mathbf{R}'_s, \mathbf{T}'_s)$, where

$$\mathbf{R}'_s = \mathbf{R}_s; \mathbf{T}'_s = \mathbf{T}_s + (\mathbf{R}_s - \mathbf{I})\mathbf{X}_0 \quad s = 1, 2, \ldots, m. \qquad (2.2.3.2)$$

(*b*) *Allowed* or *permissible origins* (Hauptman & Karle, 1953, 1959) for a given algebraic form of the s.f. are all those points in direct space which, when taken as origin, maintain the same symmetry operators $\mathbf{C}_s$. The allowed origins will therefore correspond to those points having the same symmetry environment in the sense that they are related to the symmetry elements in the same way. For instance, if $\mathbf{T}_s = 0$ for $s = 1, \ldots, 8$, then the allowed origins in *Pmmm* are the eight inversion centres.

To each functional form of the s.f. a set of permissible origins will correspond.

(*c*) A translation between permissible origins will be called a *permissible* or *allowed translation.* Trivial allowed translations correspond to the lattice periods or to their multiples. A change of origin by an allowed translation does not change the algebraic form of the s.f. Thus, according to (2.2.3.2), all origins allowed by a fixed functional form of the s.f. will be connected by translational vectors $\mathbf{X}_p$ such that

$$(\mathbf{R}_s - \mathbf{I})\mathbf{X}_p = \mathbf{V}, \quad s = 1, 2, \ldots, m, \qquad (2.2.3.3)$$

where $\mathbf{V}$ is a vector with zero or integer components.

In centred space groups, an origin translation corresponding to a centring vector $\mathbf{B}_v$ does not change the functional form of the s.f. Therefore all vectors $\mathbf{B}_v$ represent permissible translations. $\mathbf{X}_p$ will then be an allowed translation (Giacovazzo, 1974) not only when, as imposed by (2.2.3.3), the difference $\mathbf{T}'_s - \mathbf{T}_s$ is equal to one or more lattice units, but also when, for any s, the condition

$$(\mathbf{R}_s - \mathbf{I})\mathbf{X}_p = \mathbf{V} + \alpha \mathbf{B}_v, \quad s = 1, 2, \ldots, m; \quad \alpha = 0, 1 \qquad (2.2.3.4)$$

is satisfied.

We will call any set of cs. or ncs. space groups having the same allowed origin translations a Hauptman–Karle group (H–K group). The 94 ncs. primitive space groups, the 62 primitive cs. groups, the 44 ncs. centred space groups and the 30 cs. centred space groups can be collected into 13, 4, 14 and 5 H–K groups, respectively (Hauptman & Karle, 1953, 1956; Karle & Hauptman, 1961; Lessinger & Wondratschek, 1975). In Tables 2.2.3.1–2.2.3.4 the H–K groups are given together with the allowed origin translations.

(*d*) Let us consider a product of structure factors

$$F_{\mathbf{h}_1}^{A_1} \times F_{\mathbf{h}_2}^{A_2} \times \ldots \times F_{\mathbf{h}_n}^{A_n} = \prod_{j=1}^{n} F_{\mathbf{h}_j}^{A_j} = \exp\left(i \sum_{j=1}^{n} A_j \varphi_{\mathbf{h}_j}\right) \prod_{j=1}^{n} |F_{\mathbf{h}_j}|^{A_j}, \qquad (2.2.3.5)$$

A_j being integer numbers.

The factor $\sum_{j=1}^{n} A_j \varphi_{\mathbf{h}_j}$ is the phase of the product (2.2.3.5). A *structure invariant* (s.i.) is a product (2.2.3.5) such that

Table 2.2.3.1. *Allowed origin translations, seminvariant moduli and phases for centrosymmetric primitive space groups*

	H–K group					
	$(h,k,l)\underline{P}(2,2,2)$		$(h+k,l)\underline{P}(2,2)$		$(l)\underline{P}(2)$	$(h+k+l)\underline{P}(2)$
Space group	$P\bar{1}$	$Pmna$	$P\frac{4}{m}$	$P\frac{4}{n}mm$	$P\bar{3}$	$R\bar{3}$
	$P\frac{2}{m}$	$Pcca$	$P\frac{4_2}{m}$	$P\frac{4}{n}cc$	$P\bar{3}1m$	$R\bar{3}m$
	$P\frac{2_1}{m}$	$Pbam$	$P\frac{4}{n}$	$P\frac{4_2}{m}mc$	$P\bar{3}1c$	$R\bar{3}c$
	$P\frac{2}{c}$	$Pccn$	$P\frac{4_2}{n}$	$P\frac{4_2}{m}cm$	$P\bar{3}m1$	$Pm\bar{3}$
	$P\frac{2_1}{c}$	$Pbcm$	$P\frac{4}{m}mm$	$P\frac{4_2}{n}bc$	$P\bar{3}c1$	$Pn\bar{3}$
	$Pmmm$	$Pnnm$	$P\frac{4}{m}cc$	$P\frac{4_2}{n}nm$	$P\frac{6}{m}$	$Pa\bar{3}$
	$Pnnn$	$Pmmn$	$P\frac{4}{n}bm$	$P\frac{4_2}{m}bc$	$P\frac{6_3}{m}$	$Pm\bar{3}m$
	$Pccm$	$Pbcn$	$P\frac{4}{n}nc$	$P\frac{4_2}{m}nm$	$P\frac{6}{m}mm$	$Pn\bar{3}n$
	$Pban$	$Pbca$	$P\frac{4}{m}bm$	$P\frac{4_2}{n}mc$	$P\frac{6}{m}cc$	$Pm\bar{3}n$
	$Pmma$	$Pnma$	$P\frac{4}{m}nc$	$P\frac{4_2}{n}cm$	$P\frac{6_3}{m}cm$	$Pn\bar{3}m$
	$Pnna$				$P\frac{6_3}{m}mc$	
Allowed origin translations	(0, 0, 0);	$(0,\frac{1}{2},\frac{1}{2})$	(0, 0, 0)		(0, 0, 0)	(0, 0, 0)
	$(\frac{1}{2},0,0)$;	$(\frac{1}{2},0,\frac{1}{2})$	$(0,0,\frac{1}{2})$		$(0,0,\frac{1}{2})$	$(\frac{1}{2},\frac{1}{2},\frac{1}{2})$
	$(0,\frac{1}{2},0)$;	$(\frac{1}{2},\frac{1}{2},0)$	$(\frac{1}{2},\frac{1}{2},0)$			
	$(0,0,\frac{1}{2})$;	$(\frac{1}{2},\frac{1}{2},\frac{1}{2})$	$(\frac{1}{2},\frac{1}{2},\frac{1}{2})$			
Vector $\mathbf{h}_s$ seminvariantly associated with $\mathbf{h}=(h,k,l)$	(h,k,l)		$(h+k,l)$		(l)	$(h+k+l)$
Seminvariant modulus $\boldsymbol{\omega}_s$	(2, 2, 2)		(2, 2)		(2)	(2)
Seminvariant phases	φ_{eee}		φ_{eee}; φ_{ooe}		φ_{eee}; φ_{eoe}	φ_{eee}; φ_{ooe}
					φ_{oee}; φ_{ooe}	φ_{oeo}; φ_{eoo}
Number of semindependent phases to be specified	3		2		1	1

$$\sum_{j=1}^{n} A_j\mathbf{h}_j = 0. \qquad (2.2.3.6)$$

Since $|F_{\mathbf{h}_j}|$ are usually known from experiment, it is often said that s.i.'s are combinations of phases

$$\sum_{j=1}^{n} A_j\varphi_{\mathbf{h}_j}, \qquad (2.2.3.7)$$

for which (2.2.3.6) holds.

F_0, $F_{\mathbf{h}}F_{-\mathbf{h}}$, $F_{\mathbf{h}}F_{\mathbf{k}}F_{\overline{\mathbf{h}+\mathbf{k}}}$, $F_{\mathbf{h}}F_{\mathbf{k}}F_{\mathbf{l}}F_{\overline{\mathbf{h}+\mathbf{k}+\mathbf{l}}}$, $F_{\mathbf{h}}F_{\mathbf{k}}F_{\mathbf{l}}F_{\mathbf{p}}F_{\overline{\mathbf{h}+\mathbf{k}+\mathbf{l}+\mathbf{p}}}$ are examples of s.i.'s for $n=1,2,3,4,5$.

The value of any s.i. does not change with an arbitrary shift of the space-group origin and thus it will depend on the crystal structure only.

(*e*) A *structure seminvariant* (s.s.) is a product of structure factors [or a combination of phases (2.2.3.7)] whose value is unchanged when the origin is moved by an allowed translation.

Let $\mathbf{X}_p$'s be the permissible origin translations of the space group. Then the product (2.2.3.5) [or the sum (2.2.3.7)] is an s.s., if, in accordance with (2.2.3.1),

$$\sum_{j=1}^{n} A_j(\mathbf{h}_j\cdot\mathbf{X}_p) = r, \quad p=1,2,\ldots \qquad (2.2.3.8)$$

where r is a positive integer, null or a negative integer.

Conditions (2.2.3.8) can be written in the following more useful form (Hauptman & Karle, 1953):

$$\sum_{j=1}^{n} A_j\mathbf{h}_{s_j} \equiv 0 \pmod{\boldsymbol{\omega}_s}, \qquad (2.2.3.9)$$

where $\mathbf{h}_{s_j}$ is the vector seminvariantly associated with the vector $\mathbf{h}_j$ and $\boldsymbol{\omega}_s$ is the seminvariant modulus. In Tables 2.2.3.1–2.2.3.4, the reflection $\mathbf{h}_s$ seminvariantly associated with $\mathbf{h}=(h,k,l)$, the seminvariant modulus $\boldsymbol{\omega}_s$ and seminvariant phases are given for every H–K group.

The symbol of any group (*cf.* Giacovazzo, 1974) has the structure $\mathbf{h}_sL\boldsymbol{\omega}_s$, where L stands for the lattice symbol. This symbol is underlined if the space group is cs.

By definition, if the class of permissible origin has been chosen, that is to say, if the algebraic form of the symmetry operators has been fixed, then the value of an s.s. does not depend on the origin but on the crystal structure only.

2.3. Patterson and molecular-replacement techniques

BY M. G. ROSSMANN AND E. ARNOLD

2.3.1. Introduction

2.3.1.1. *Background*

Historically, the Patterson has been used in a variety of ways to effect the solutions of crystal structures. While some simple structures (Ketelaar & de Vries, 1939; Hughes, 1940; Speakman, 1949; Shoemaker *et al*., 1950) were solved by direct analysis of Patterson syntheses, alternative methods have largely superseded this procedure. An early innovation was the heavy-atom method which depends on the location of a small number of relatively strong scatterers (Harker, 1936). Image-seeking methods and Patterson superposition techniques were first contemplated in the late 1930s (Wrinch, 1939) and applied sometime later (Beevers & Robertson, 1950; Clastre & Gay, 1950; Garrido, 1950*a*; Buerger, 1959). This experience provided the encouragement for computerized vector-search methods to locate individual atoms automatically (Mighell & Jacobson, 1963; Kraut, 1961; Hamilton, 1965; Simpson *et al*., 1965) or to position known molecular fragments in unknown crystal structures (Nordman & Nakatsu, 1963; Huber, 1965). The Patterson function has been used extensively in conjunction with the isomorphous replacement method (Rossmann, 1960; Blow, 1958) or anomalous dispersion (Rossmann, 1961*a*) to determine the position of heavy-atom substitution. Pattersons have been used to detect the presence and relative orientation of multiple copies of a given chemical motif in the crystallographic asymmetric unit in the same or different crystals (Rossmann & Blow, 1962). Finally, the orientation and placement of known molecular structures ('molecular replacement') into unknown crystal structures can be accomplished *via* Patterson techniques.

The function, introduced by Patterson in 1934 (Patterson, 1934*a*,*b*), is a convolution of electron density with itself and may be defined as

$$P(\mathbf{u}) = \int_V \rho(\mathbf{x}) \cdot \rho(\mathbf{u}+\mathbf{x})\, \mathrm{d}\mathbf{x}, \qquad (2.3.1.1)$$

where $P(\mathbf{u})$ is the 'Patterson' function at $\mathbf{u}$, $\rho(\mathbf{x})$ is the crystal's periodic electron density and V is the volume of the unit cell. The Patterson function, or F^2 series, can be calculated directly from the experimentally derived X-ray intensities as

$$P(\mathbf{u}) = \frac{2}{V^2} \sum_{\mathbf{h}}^{\text{hemisphere}} |\mathbf{F}_\mathbf{h}|^2 \cos 2\pi\mathbf{h}\cdot\mathbf{u}. \qquad (2.3.1.2)$$

The derivation of (2.3.1.2) from (2.3.1.1) can be found in this volume (see Section 1.3.4.2.1.6) along with a discussion of the physical significance and symmetry of the Patterson function, although the principal properties will be restated here.

The Patterson can be considered to be a vector map of all the pairwise interactions between the atoms in a unit cell. The vectors in a Patterson correspond to vectors in the real (direct) crystal cell but translated to the Patterson origin. Their weights are proportional to the product of densities at the tips of the vectors in the real cell. The Patterson unit cell has the same size as the real crystal cell. The symmetry of the Patterson comprises the Laue point group of the crystal cell plus any additional lattice symmetry due to Bravais centring. The reduction of the real space group to the Laue symmetry is produced by the translation of all vectors to the Patterson origin and the introduction of a centre of symmetry. The latter is a consequence of the relationship between the vectors **AB** and **BA**. The Patterson symmetries for all 230 space groups are tabulated in *IT* A (1983).

An analysis of Patterson peaks can be obtained by considering N atoms with form factors f_i in the unit cell. Then

$$\mathbf{F}_\mathbf{h} = \sum_{i=1}^{N} f_i \exp(2\pi i\mathbf{h}\cdot\mathbf{x}_i).$$

Using Friedel's law,

$$|\mathbf{F}_\mathbf{h}|^2 = \mathbf{F}_\mathbf{h}\cdot\mathbf{F}_\mathbf{h}^* = \left[\sum_{i=1}^{N} f_i \exp(2\pi i\mathbf{h}\cdot\mathbf{x}_i)\right]\left[\sum_{j=1}^{N} f_j \exp(-2\pi i\mathbf{h}\cdot\mathbf{x}_j)\right],$$

which can be decomposed to

$$|\mathbf{F}_\mathbf{h}|^2 = \sum_{i=1}^{N} f_i^2 + \sum_{i\neq j}^{N}\sum^{N} f_i f_j \exp[2\pi i\mathbf{h}\cdot(\mathbf{x}_i-\mathbf{x}_j)]. \qquad (2.3.1.3)$$

On substituting (2.3.1.3) in (2.3.1.2), we see that the Patterson consists of the sum of N^2 total interactions of which N are of weight f_i^2 at the origin and $N(N-1)$ are of weight $f_i f_j$ at $\mathbf{x}_i-\mathbf{x}_j$.

The weight of a peak in a real cell is given by

$$w_i = \int_U \rho_i(\mathbf{x})\, \mathrm{d}\mathbf{x} = Z_i \text{ (the atomic number)},$$

where U is the volume of the atom i. By analogy, the weight of a peak in a Patterson (form factor $f_i f_j$) will be given by

$$w_{ij} = \int_U P_{ij}(\mathbf{u})\, \mathrm{d}\mathbf{u} = Z_i Z_j.$$

Although the maximum height of a peak will depend on the spread of the peak, it is reasonable to assume that heights of peaks in a Patterson are proportional to the products of the atomic numbers of the interacting atoms.

There are a total of N^2 interactions in a Patterson due to N atoms in the crystal cell. These can be represented as an $N\times N$ square matrix whose elements $\mathbf{u}_{ij}$, w_{ij} indicate the position and weight of the peak produced between atoms i and j (Table 2.3.1.1). The N vectors corresponding to the diagonal of this matrix are located at the Patterson origin and arise from the convolution of each atom with itself. This leaves $N(N-1)$ vectors whose locations depend on the relative positions of all of the atoms in the crystal cell and whose weights depend on the atom types related by the vector. Complete specification of the unique non-origin Patterson vectors requires description of only the $N(N-1)/2$ elements in either the upper or the lower triangle of this matrix, since the two sets of vectors represented by the two triangles are related by a centre of symmetry $[\mathbf{u}_{ij}\equiv\mathbf{x}_i-\mathbf{x}_j=-\mathbf{u}_{ij}\equiv-(\mathbf{x}_j-\mathbf{x}_i)]$. Patterson vector positions are usually represented as $\langle uvw\rangle$, where u, v and w are expressed as fractions of the Patterson cell axes.

2.3.1.2. *Limits to the number of resolved vectors*

If we assume a constant number of atoms per unit volume, the number of atoms N in a unit cell increases in direct proportion with the volume of the unit cell. Since the number of non-origin peaks in the Patterson function is $N(N-1)$ and the Patterson cell is the same size as the real cell, the problem of overlapping peaks in the Patterson function becomes severe as N increases. To make matters worse, the breadth of a Patterson peak is roughly equal to the sum of the breadth of the original atoms. The effective width of a Patterson peak will also increase with increasing thermal motion, although this effect can be artificially reduced by sharpening techniques. Naturally, a loss of attainable resolution at high scattering angles

2.4. Isomorphous replacement and anomalous scattering

BY M. VIJAYAN AND S. RAMASESHAN

2.4.1. Introduction

Isomorphous replacement is among the earliest methods to be employed for crystal structure determination (Cork, 1927). The power of this method was amply demonstrated in the classical X-ray work of J. M. Robertson on phthalocyanine in the 1930s using centric data (Robertson, 1936; Robertson & Woodward, 1937). The structure determination of strychnine sulfate pentahydrate by Bijvoet and others provides an early example of the application of this method to acentric reflections (Bokhoven *et al.*, 1951). The usefulness of isomorphous replacement in the analysis of complex protein structures was demonstrated by Perutz and colleagues (Green *et al.*, 1954). This was closely followed by developments in the methodology for the application of isomorphous replacement to protein work (Harker, 1956; Blow & Crick, 1959) and rapidly led to the first ever structure solution of two related protein crystals, namely, those of myoglobin and haemoglobin (Kendrew *et al.*, 1960; Cullis *et al.*, 1961*b*). Since then isomorphous replacement has been the method of choice in macromolecular crystallography and most of the subsequent developments in and applications of this method have been concerned with biological macromolecules, mainly proteins (Blundell & Johnson, 1976; McPherson, 1982).

The application of anomalous-scattering effects has often developed in parallel with that of isomorphous replacement. Indeed, the two methods are complementary to a substantial extent and they are often treated together, as in this article. Although the most important effect of anomalous scattering, namely, the violation of Friedel's law, was experimentally observed as early as 1930 (Coster *et al.*, 1930), two decades elapsed before this effect was made use of for the first time by Bijvoet and his associates for the determination of the absolute configuration of asymmetric molecules as well as for phase evaluation (Bijvoet, 1949, 1954; Bijvoet *et al.*, 1951). Since then there has been a phenomenal spurt in the application of anomalous-scattering effects (Srinivasan, 1972; Ramaseshan & Abrahams, 1975; Vijayan, 1987). A quantitative formulation for the determination of phase angles using intensity differences between Friedel equivalents was derived by Ramachandran & Raman (1956), while Okaya & Pepinsky (1956) successfully developed a Patterson approach involving anomalous effects. The anomalous-scattering method of phase determination has since been used in the structure analysis of several structures, including those of a complex derivative of vitamin B_{12} (Dale *et al.*, 1963) and a small protein (Hendrickson & Teeter, 1981). In the meantime, the effect of changes in the real component of the dispersion correction as a function of the wavelength of the radiation used, first demonstrated by Mark & Szillard (1925), also received considerable attention. This effect, which is formally equivalent to that of isomorphous replacement, was demonstrated to be useful in structure determination (Ramaseshan *et al.*, 1957; Ramaseshan, 1963). Protein crystallographers have been quick to exploit anomalous-scattering effects (Rossmann, 1961; Kartha & Parthasarathy, 1965; North, 1965; Matthews, 1966; Hendrickson, 1979) and, as in the case of the isomorphous replacement method, the most useful applications of anomalous scattering during the last two decades have been perhaps in the field of macromolecular crystallography (Kartha, 1975; Watenpaugh *et al.*, 1975; Vijayan, 1981). In addition to anomalous scattering of X-rays, that of neutrons was also found to have interesting applications (Koetzle & Hamilton, 1975; Sikka & Rajagopal, 1975). More recently there has been a further revival in the development of anomalous-scattering methods with the advent of synchrotron radiation, particularly in view of the possibility of choosing any desired wavelength from a synchrotron-radiation source (Helliwell, 1984).

It is clear from the foregoing that the isomorphous replacement and the anomalous-scattering methods have a long and distinguished history. It is therefore impossible to do full justice to them in a comparatively short presentation like the present one. Several procedures for the application of these methods have been developed at different times. Many, although of considerable historical importance, are not extensively used at present for a variety of reasons. No attempt has been made to discuss them in detail here; the emphasis is primarily on the state of the art as it exists now. The available literature on isomorphous replacement and anomalous scattering is extensive. The reference list given at the end of this part is representative rather than exhaustive.

During the past few years, rapid developments have taken place in the isomorphous replacement and anomalous-scattering methods, particularly in the latter, as applied to macromolecular crystallography. These developments will be described in detail in *International Tables for Crystallography*, Volume F (2001). Therefore, they have not been dealt with in this chapter. Significant developments in applications of direct methods to macromolecular crystallography have also occurred in recent years. A summary of these developments as well as the traditional direct methods on which the recent progress is based are presented in Chapter 2.2.

2.4.2. Isomorphous replacement method

2.4.2.1. *Isomorphous replacement and isomorphous addition*

Two crystals are said to be isomorphous if (*a*) both have the same space group and unit-cell dimensions and (*b*) the types and the positions of atoms in both are the same except for a replacement of one or more atoms in one structure with different types of atoms in the other (isomorphous replacement) or the presence of one or more additional atoms in one of them (isomorphous addition). Consider two crystal structures with identical space groups and unit-cell dimensions, one containing N atoms and the other M atoms. The N atoms in the first structure contain subsets P and Q whereas the M atoms in the second structure contain subsets P, Q' and R. The subset P is common to both structures in terms of atomic positions and atom types. The atomic positions are identical in subsets Q and Q', but at any given atomic position the atom type is different in Q and Q'. The subset R exists only in the second structure. If $\mathbf{F}_N$ and $\mathbf{F}_M$ denote the structure factors of the two structures for a given reflection,

$$\mathbf{F}_N = \mathbf{F}_P + \mathbf{F}_Q \qquad (2.4.2.1)$$

and

$$\mathbf{F}_M = \mathbf{F}_P + \mathbf{F}_{Q'} + \mathbf{F}_R, \qquad (2.4.2.2)$$

where the quantities on the right-hand side represent contributions from different subsets. From (2.4.2.1) and (2.4.2.2) we have

$$\mathbf{F}_M - \mathbf{F}_N = \mathbf{F}_H = \mathbf{F}_{Q'} - \mathbf{F}_Q + \mathbf{F}_R. \qquad (2.4.2.3)$$

The above equations are illustrated in the Argand diagram shown in Fig. 2.4.2.1. $\mathbf{F}_Q$ and $\mathbf{F}_{Q'}$ would be collinear if all the atoms in Q were of the same type and those in Q' of another single type, as in the replacement of chlorine atoms in a structure by bromine atoms.

We have a case of 'isomorphous replacement' if $\mathbf{F}_R = 0$ ($\mathbf{F}_H = \mathbf{F}_{Q'} - \mathbf{F}_Q$) and a case of 'isomorphous addition' if $\mathbf{F}_Q = \mathbf{F}_{Q'} = 0$ ($\mathbf{F}_H = \mathbf{F}_R$). Once $\mathbf{F}_H$ is known, in addition to the magnitudes of $\mathbf{F}_N$ and $\mathbf{F}_M$, which can be obtained experimentally, the two cases can be treated in an equivalent manner in reciprocal space. In deference to common practice, the term 'isomorphous

2.5. Electron diffraction and electron microscopy in structure determination

BY J. M. COWLEY, P. GOODMAN, B. K. VAINSHTEIN, B. B. ZVYAGIN AND D. L. DORSET

2.5.1. Foreword (J. M. COWLEY)

Given that electrons have wave properties and the wavelengths lie in a suitable range, the diffraction of electrons by matter is completely analogous to the diffraction of X-rays. While for X-rays the scattering function is the electron-density distribution, for electrons it is the potential distribution which is similarly peaked at the atomic sites. Hence, in principle, electron diffraction may be used as the basis for crystal structure determination. In practice it is used much less widely than X-ray diffraction for the determination of crystal structures but is receiving increasing attention as a means for obtaining structural information not readily accessible with X-ray- or neutron-diffraction techniques.

Electrons having wavelengths comparable with those of the X-rays commonly used in diffraction experiments have energies of the order of 100 eV. For such electrons, the interactions with matter are so strong that they can penetrate only a few layers of atoms on the surfaces of solids. They are used extensively for the study of surface structures by low-energy electron diffraction (LEED) and associated techniques. These techniques are not covered in this series of volumes, which include the principles and practice of only those diffraction and imaging techniques making use of high-energy electrons, having energies in the range of 20 keV to 1 MeV or more, in transmission through thin specimens.

For the most commonly used energy ranges of high-energy electrons, 100 to 400 keV, the wavelengths are about 50 times smaller than for X-rays. Hence the scattering angles are much smaller, of the order of 10^{-2} rad, the recording geometry is relatively simple and the diffraction pattern represents, to a useful first approximation, a planar section of reciprocal space.

The elastic scattering of electrons by atoms is several orders of magnitude greater than for X-rays. This fact has profound consequences, which in some cases are highly favourable and in other cases are serious hindrances to structure analysis work. On the one hand it implies that electron-diffraction patterns can be obtained from very small single-crystal regions having thicknesses equal to only a few layers of atoms and, with recently developed techniques, having diameters equivalent to only a few interatomic distances. Hence single-crystal patterns can be obtained from microcrystalline phases.

However, the strong scattering of electrons implies that the simple kinematical single-scattering approximation, on which most X-ray diffraction structure analysis is based, fails for electrons except for very thin crystals composed of light-atom materials. Strong dynamical diffraction effects occur for crystals which may be 100 Å thick, or less for heavy-atom materials. As a consequence, the theory of dynamical diffraction for electrons has been well developed, particularly for the particular special diffracting conditions relevant to the transmission of fast electrons (see Chapter 5.2), and observations of dynamical diffraction effects are commonly made and quantitatively interpreted. The possibility has thus arisen of using the observation of dynamical diffraction effects as the basis for obtaining crystal structure information. The fact that dynamical diffraction is dependent on the relative phases of the diffracted waves then implies that relative phase information can be deduced from the diffraction intensities and the limitations of kinematical diffraction, such as Friedel's law, do not apply. The most immediately practicable method for making use of this possibility is convergent-beam electron diffraction (CBED) as described in Section 2.5.3.

A further important factor, determining the methods for observing electron diffraction, is that, being charged particles, electrons can be focused by electromagnetic lenses. The irreducible aberrations of cylindrical magnetic lenses have, to date, limited the resolution of electron microscopes to the extent that the least resolvable distances (or 'resolutions') are about 100 times the electron wavelength. However, with microscopes having a resolution of better than 2 Å it is possible to distinguish the individual rows of atoms, parallel to the incident electron beam, in the principal orientations of many crystalline phases. Thus 'structure images' can be obtained, sometimes showing direct representation of projections of crystal structures [see *IT* C (1999), Section 4.3.8]. However, the complications of dynamical scattering and of the coherent imaging processes are such that the image intensities vary strongly with crystal thickness and tilt, and with the defocus or other parameters of the imaging system, making the interpretation of images difficult except in special circumstances. Fortunately, computer programs are readily available whereby image intensities can be calculated for model structures [see *IT* C (1999), Section 4.3.6] Hence the means exist for deriving the projection of the structure if only by a process of trial and error and not, as would be desirable, from a direct interpretation of the observations.

The accuracy with which the projection of a structure can be deduced from an image, or series of images, improves as the resolution of the microscope improves but is not at all comparable with the accuracy attainable with X-ray diffraction methods. A particular virtue of high-resolution electron microscopy as a structural tool is that it may give information on individual small regions of the sample. Structures can be determined of 'phases' existing over distances of only a few unit cells and the defects and local disorders can be examined, one by one.

The observation of electron-diffraction patterns forms an essential part of the technique of structure imaging in high-resolution electron microscopy, because the diffraction patterns are used to align the crystals to appropriate axial orientations. More generally, for all electron microscopy of crystalline materials the image interpretation depends on knowledge of the diffraction conditions. Fortunately, the diffraction pattern and image of any specimen region can be obtained in rapid succession by a simple switching of lens currents. The ready comparison of the image and diffraction data has become an essential component of the electron microscopy of crystalline materials but has also been of fundamental importance for the development of electron-diffraction theory and techniques.

The individual specimen regions giving single-crystal electron-diffraction patterns are, with few exceptions, so small that they can be seen only by use of an electron microscope. Hence, historically, it was only after electron microscopes were commonly available that the direct correlations of diffraction intensities with crystal size and shape could be made, and a proper basis was available for the development of the adequate dynamical diffraction theory.

For the complete description of a diffraction pattern or image intensities obtained with electrons, it is necessary to include the effects of inelastic scattering as well as elastic scattering. In contrast to the X-ray diffraction case, the inelastic scattering does not produce just a broad and generally negligible background. The average energy loss for an inelastic scattering event is about 20 eV, which is small compared with the energy of about 100 keV for the incident electrons. The inelastically scattered electrons have a narrow angular distribution and are diffracted in much the same way as the incident or elastically scattered electrons in a crystal. They therefore produce a highly modulated contribution to the diffraction pattern, strongly peaked about the Bragg spot positions (see Chapter 4.3). Also, as a result of the inelastic scattering processes, including thermal diffuse scattering, an effective

3.1. Distances, angles, and their standard uncertainties

BY D. E. SANDS

3.1.1. Introduction

A crystal structure analysis provides information from which it is possible to compute distances between atoms, angles between interatomic vectors, and the uncertainties in these quantities. In Cartesian coordinate systems, these geometric computations require the Pythagorean theorem and elementary trigonometry. The natural coordinate systems of crystals, though, are determined by symmetry, and only in special cases are the basis vectors (or coordinate axes) of these systems constrained to be of equal lengths or mutually perpendicular.

It is possible, of course, to transform the positional parameters of the atoms to a Cartesian system and perform the subsequent calculations with the transformed coordinates. Along with the coordinates, the transformations must be applied to anisotropic thermal factors, variance–covariance matrices and other important quantities. Moreover, leaving the natural coordinate system of the crystal sacrifices the simplified relationships imposed by translational and point symmetry; for example, if an atom has fractional coordinates x^1, x^2, x^3, an equivalent atom will be at $1 + x^1$, x^2, x^3, *etc.*

Fortunately, formulation of the calculations in generalized rectilinear coordinate systems is straightforward, and readily adapted to computer languages (Section 3.1.12 illustrates the use of Fortran for such calculations). The techniques for these computations are those of tensor analysis, which provides a compact and elegant notation. While an effort will be made to be self-sufficient in this chapter, some proficiency in vector algebra is assumed, and the reader not familiar with the basics of tensor analysis should refer to Chapter 1.1 and Sands (1982*a*).

3.1.2. Scalar product

The scalar product of vectors **u** and **v** is defined as

$$\mathbf{u} \cdot \mathbf{v} = uv \cos\varphi, \tag{3.1.2.1}$$

where u and v are the lengths of the vectors and φ is the angle between them. In terms of components,

$$\mathbf{u} \cdot \mathbf{v} = (u^i \mathbf{a}_i) \cdot (v^j \mathbf{a}_j) \tag{3.1.2.2}$$

$$\mathbf{u} \cdot \mathbf{v} = u^i v^j \mathbf{a}_i \cdot \mathbf{a}_j \tag{3.1.2.3}$$

$$\mathbf{u} \cdot \mathbf{v} = u^i v^j g_{ij}. \tag{3.1.2.4}$$

In all equations in this chapter, the convention is followed that summation is implied over an index that is repeated once as a subscript and once as a superscript in an expression; thus, the right-hand side of (3.1.2.4) implies the sum of nine terms

$$u^1 v^1 g_{11} + u^1 v^2 g_{12} + \ldots + u^3 v^3 g_{33}.$$

The g_{ij} in (3.1.2.4) are the components of the metric tensor [see Chapter 1.1 and Sands (1982*a*)]

$$g_{ij} = \mathbf{a}_i \cdot \mathbf{a}_j. \tag{3.1.2.5}$$

Subscripts are used for quantities that transform the same way as the basis vectors $\mathbf{a}_i$; such quantities are said to transform covariantly. Superscripts denote quantities that transform the same way as coordinates x^i; these quantities are said to transform contravariantly (Sands, 1982*a*).

Equation (3.1.2.4) is in a form convenient for computer evaluation, with indices i and j taking successively all values from 1 to 3. The matrix form of (3.1.2.4) is useful both for symbolic manipulation and for computation,

$$\mathbf{u} \cdot \mathbf{v} = \boldsymbol{u}^T \boldsymbol{g} \boldsymbol{v}, \tag{3.1.2.6}$$

where the superscript italic T following a matrix symbol indicates a transpose. Written out in full, (3.1.2.6) is

$$\mathbf{u} \cdot \mathbf{v} = (u^1 u^2 u^3) \begin{pmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{pmatrix} \begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix}. \tag{3.1.2.7}$$

If $\boldsymbol{u}$ is the column vector with components u^1, u^2, u^3, $\boldsymbol{u}^T$ is the corresponding row vector shown in (3.1.2.7).

3.1.3. Length of a vector

By (3.1.2.1), the scalar product of a vector with itself is

$$\mathbf{v} \cdot \mathbf{v} = (v)^2. \tag{3.1.3.1}$$

The length of **v** is, therefore, given by

$$v = (v^i v^j g_{ij})^{1/2}. \tag{3.1.3.2}$$

Computation of lengths in a generalized rectilinear coordinate system is thus simply a matter of evaluating the double summation $v^i v^j g_{ij}$ and taking the square root.

3.1.4. Angle between two vectors

By (3.1.2.1) and (3.1.2.4), the angle φ between vectors **u** and **v** is given by

$$\varphi = \cos^{-1}[u^i v^j g_{ij}/(uv)]. \tag{3.1.4.1}$$

An even more concise expression of equations such as (3.1.4.1) is possible by making use of the ability of the metric tensor $\boldsymbol{g}$ to convert components from contravariant to covariant (Sands, 1982*a*). Thus,

$$v_i = g_{ij} v^j, \quad u_j = g_{ij} u^i, \tag{3.1.4.2}$$

and (3.1.2.4) may be written succinctly as

$$\mathbf{u} \cdot \mathbf{v} = u^i v_i \tag{3.1.4.3}$$

or

$$\mathbf{u} \cdot \mathbf{v} = u_i v^i. \tag{3.1.4.4}$$

With this notation, the angle calculation of (3.1.4.1) becomes

$$\varphi = \cos^{-1}[u^i v_i/(uv)] = \cos^{-1}[u_i v^i/(uv)]. \tag{3.1.4.5}$$

The summations in (3.1.4.3), (3.1.4.4) and (3.1.4.5) include only three terms, and are thus equivalent in numerical effort to the computation in a Cartesian system, in which the metric tensor is represented by the unit matrix and there is no numerical distinction between covariant components and contravariant components.

Appreciation of the elegance of tensor formulations may be enhanced by noting that corresponding to the metric tensor **g** with components g_{ij} there is a contravariant metric tensor $\mathbf{g}^*$ with components

$$g^{ij} = \mathbf{a}^i \cdot \mathbf{a}^j. \tag{3.1.4.6}$$

The $\mathbf{a}^i$ are contravariant basis vectors, known to crystallographers as reciprocal axes. Expressions parallel to (3.1.4.2) may be written, in which $\mathbf{g}^*$ plays the role of converting covariant components to contravariant components. These tensors thus express mathematically the crystallographic notions of crystal space and reciprocal space [see Chapter 1.1 and Sands (1982*a*)].

3.2. The least-squares plane

BY R. E. MARSH AND V. SCHOMAKER

3.2.1. Introduction

By way of introduction, we remark that in earlier days of crystal structure analysis, before the advent of high-speed computers and routine three-dimensional analyses, molecular planarity was often assumed so that atom coordinates along the direction of projection could be estimated from two-dimensional data [see, *e.g.*, Robertson (1948)]. Today, the usual aim in deriving the coefficients of a plane is to investigate the degree of planarity of a group of atoms as found in a full, three-dimensional structure determination. We further note that, for such purposes, a crystallographer will often be served just as well by establishing the plane in an almost arbitrary fashion as by resorting to the most elaborate, nit-picking and pretentious least-squares treatment. The approximate plane and the associated perpendicular distances of the atoms from it will be all he needs as scaffolding for his geometrical and structural imagination; reasonable common sense will take the place of explicit attention to error estimates.

Nevertheless, we think it appropriate to lay out in some detail the derivation of the 'best' plane, in a least-squares sense, through a group of atoms and of the standard uncertainties associated with this plane. We see two cases: (1) The weights of the atoms in question are considered to be isotropic and uncorrelated (*i.e.* the weight matrix for the positions of all the atoms is diagonal, when written in terms of Cartesian axes, and for each atom the three diagonal elements are equal). In such cases the weights may have little or nothing to do with estimates of random error in the atom positions (they may have been assigned merely for convenience or convention), and, therefore, no one should feel that the treatment is proper in respect to the theory of errors. Nevertheless, it may be desired to incorporate the error estimates (variances) of the atom positions into the *results* of such calculations, whereupon these variances (which may be anisotropic, with correlation between atoms) need to be propagated. In this case the distinction between *weights* (or their inverses) and *variances* must be kept very clear. (2) The weights are anisotropic and are presumably derived from a variance–covariance matrix, which may include correlation terms between different atoms; the objective is to achieve a truly proper Gaussian least-squares result.

3.2.2. Least-squares plane based on uncorrelated, isotropic weights

This is surely the most common situation; it is not often that one will wish to take the trouble, or be presumptive enough, to assign anisotropic or correlated weights to the various atoms. And one will sometimes, perhaps even often, not be genuinely interested in the hypothesis that the atoms actually are rigorously coplanar; for instance, one might be interested in examining the best plane through such a patently non-planar molecule as cyclohexane. Moreover, the calculation is simple enough, given the availability of computers and programs, as to be a practical realization of the off-the-cuff treatment suggested in our opening paragraph. The problem of deriving the plane's coefficients is intrinsically nonlinear in the way first discussed by Schomaker *et al.* (1959; SWMB). Any formulation other than as an eigenvalue–eigenvector problem (SWMB), as far as we can tell, will sometimes go astray. As to the propagation of errors, numerous treatments have been given, but none that we have seen is altogether satisfactory.

We refer all vectors and matrices to Cartesian axes, because that is the most convenient in calculation. However, a more elegant formulation can be written in terms of general axes [*e.g.*, as in Shmueli (1981)].

The notation is troublesome. Indices are needed for atom number and Cartesian direction, and the exponent 2 is needed as well, which is difficult if there are superscript indices. The best way seems to be to write all the indices as subscripts and distinguish among them by context – *i*, *j*, 1, 2, 3 for directions; *k*, *l*, *p* (and sometimes *K*, . . .) for atoms. In any case, *atom* first then *direction* if there are two subscripts; *direction*, if only one index for a vector component, but *atom* (in this section at least) if for a weight or a vector. And σ_{d_1}, *e.g.*, for the standard uncertainty of the distance of atom 1 from a plane. For simplicity in practice, we use Cartesian coordinates throughout.

The first task is to find the plane, which we write as

$$0 = \mathbf{m}\cdot\mathbf{r} - d \equiv \mathsf{m}^T\mathsf{r} - d,$$

where $\mathbf{r}$ is here the vector from the origin to any point on the plane (but usually represents the measured position of an atom), $\mathbf{m}$ is a unit vector parallel to the normal from the origin to the plane, d is the length of the normal, and m and r are the column representations of $\mathbf{m}$ and $\mathbf{r}$. The least-squares condition is to find the stationary values of $S \equiv [w_k(\mathsf{m}^T\mathsf{r}_k - d)^2]$ subject to $\mathsf{m}^T\mathsf{m} = 1$, with r_k, $k = 1, \ldots, n$, the vector from the origin to atom k and with weights, w_k, isotropic and without interatomic correlations for the n atoms of the plane. We also write S as $S \equiv [w(\mathsf{m}^T\mathsf{r} - d)^2]$, the subscript for atom number being implicit in the Gaussian summations $([\ldots])$ over all atoms, as it is also in the angle-bracket notation for the weighted average over all atoms, for example in $\langle\mathsf{r}\rangle$ – the weighted centroid of the groups of atoms – just below.

First solve for d, the origin-to-plane distance.

$$0 = -\frac{1}{2}\frac{\partial S}{\partial d} = [w(\mathsf{m}^T\mathsf{r} - d)] = 0,$$
$$d = [w\mathsf{m}^T\mathsf{r}]/[w] \equiv \mathsf{m}^T\langle\mathsf{r}\rangle.$$

Then

$$S \equiv [w(\mathsf{m}^T\mathsf{r} - d)^2] = [w\{\mathsf{m}^T(\mathsf{r} - \langle\mathsf{r}\rangle)\}^2]$$
$$\equiv [w(\mathsf{m}^T\mathsf{s})^2] \equiv \mathsf{m}^T[w\mathsf{s}\mathsf{s}^T]\mathsf{m} \equiv \mathsf{m}^T\mathbf{A}\mathsf{m}.$$

Here $\mathbf{s}_k \equiv \mathsf{r}_k - \langle\mathsf{r}\rangle$ is the vector from the centroid to atom k. Then solve for $\mathbf{m}$. This is the eigenvalue problem – to diagonalize $\mathbf{A}$ (bear in mind that $\mathbf{A}_{ij}$ is just $[ws_is_j]$) by rotating the coordinate axes, *i.e.*, to find the 3×3 arrays $\mathbf{M}$ and $\mathbf{L}$, $\mathbf{L}$ diagonal, to satisfy

$$\mathbf{M}^T\mathbf{A}\mathbf{M} = \mathbf{L}, \qquad \mathbf{M}^T\mathbf{M} = \mathbf{I}.$$

$\mathbf{A}$ and $\mathbf{M}$ are symmetric; the columns m of $\mathbf{M}$ are the direction cosines of, and the diagonal elements of $\mathbf{L}$ are the sums of weighted squares of residuals from, the best, worst and intermediate planes, as discussed by SWMB.

3.2.2.1. *Error propagation*

Waser *et al.* (1973; WMC) carefully discussed how the random errors of measurement of the atom positions propagate into the derived quantities in the foregoing determination of a least-squares plane. This section presents an extension of their discussion. To begin, however, we first show how standard first-order perturbation theory conveniently describes the propagation of error into $\mathbf{M}$ and $\mathbf{L}$ when the positions r_k of the atoms are incremented by the amounts $\delta\mathsf{r}_k \equiv \xi_k$ and the corresponding quantities $\mathbf{s}_k \equiv \mathsf{r}_k - \langle\mathsf{r}\rangle$ (the vectors from the centroid to the atoms) by the amounts η_k, $(\mathbf{s}\rightarrow\mathbf{s}+\eta)$, $\eta_k \equiv \xi_k - \langle\xi\rangle$. (The need to account for the variation in position of the centroid, *i.e.* to distinguish between η_k and ξ_k, was overlooked by WMC.) The consequent increments in $\mathbf{A}, \mathbf{M}$ and $\mathbf{L}$ are

3.3. Molecular modelling and graphics

BY R. DIAMOND

3.3.1. Graphics

3.3.1.1. *Coordinate systems, notation and standards*

3.3.1.1.1. *Cartesian and crystallographic coordinates*

It is usual, for purposes of molecular modelling and of computer graphics, to adopt a Cartesian coordinate system using mutually perpendicular axes in a right-handed system using the ångström unit or the nanometre as the unit of distance along such axes, and largely to ignore the existence of crystallographic coordinates expressed as fractions of unit-cell edges. Transformations between the two are thus associated, usually, with the input and output stages of any software concerned with modelling and graphics, and it will be assumed after this section that all coordinates are Cartesian using the chosen unit of distance as the unit of coordinates. For a discussion of coordinate transformations and rotations without making this assumption see Chapter 1.1 in which formulations using co- and contravariant forms are presented.

The relationship between these systems may be written

$$\mathbf{X} = \boldsymbol{M}\mathbf{x} \quad \mathbf{x} = \boldsymbol{M}^{-1}\mathbf{X}$$

in which **X** and **x** are position vectors in direct space, written as column vectors, with **x** expressed in crystallographic fractional coordinates (dimensionless) and **X** in Cartesian coordinates (dimension of length).

There are two forms of $\boldsymbol{M}$ in common use. The first of these sets the first component of **X** parallel to $\mathbf{a}^*$ and the third parallel to **c** and is

$$\boldsymbol{M} = \begin{pmatrix} a\varphi/\sin\alpha & 0 & 0 \\ a(\cos\gamma - \cos\alpha\cos\beta)/\sin\alpha & b\sin\alpha & 0 \\ a\cos\beta & b\cos\alpha & c \end{pmatrix}$$

$$\boldsymbol{M}^{-1} = \begin{pmatrix} \sin\alpha/a\varphi & 0 & 0 \\ (\cos\alpha\cos\beta - \cos\gamma)/b\varphi\sin\alpha & 1/b\sin\alpha & 0 \\ (\cos\alpha\cos\gamma - \cos\beta)/c\varphi\sin\alpha & -1/c\tan\alpha & 1/c \end{pmatrix}$$

in which

$$\varphi = \sqrt{1 - \cos^2\alpha - \cos^2\beta - \cos^2\gamma + 2\cos\alpha\cos\beta\cos\gamma} = \sin\alpha\sin\beta\sin\gamma^*.$$

φ is equal to the volume of the unit cell divided by *abc*, and is unchanged by cyclic permutation of α, β and γ and of α^*, β^* and γ^*. The Cartesian and crystallographic axes have the same chirality if the positive square root is taken.

The second form sets the first component of **X** parallel to **a** and the third component of **X** parallel to $\mathbf{c}^*$ and is

$$\boldsymbol{M} = \begin{pmatrix} a & b\cos\gamma & c\cos\beta \\ 0 & b\sin\gamma & c(\cos\alpha - \cos\beta\cos\gamma)/\sin\gamma \\ 0 & 0 & c\varphi/\sin\gamma \end{pmatrix}$$

$$\boldsymbol{M}^{-1} = \begin{pmatrix} 1/a & -1/a\tan\gamma & (\cos\alpha\cos\gamma - \cos\beta)/a\varphi\sin\gamma \\ 0 & 1/b\sin\gamma & (\cos\beta\cos\gamma - \cos\alpha)/b\varphi\sin\gamma \\ 0 & 0 & \sin\gamma/c\varphi \end{pmatrix}.$$

A third form, suitable only for rhombohedral cells, is

$$\boldsymbol{M} = \frac{a}{3}\begin{pmatrix} p+2q & p-q & p-q \\ p-q & p+2q & p-q \\ p-q & p-q & p+2q \end{pmatrix}$$

$$\boldsymbol{M}^{-1} = \frac{1}{3a}\begin{pmatrix} \frac{1}{p}+\frac{2}{q} & \frac{1}{p}-\frac{1}{q} & \frac{1}{p}-\frac{1}{q} \\ \frac{1}{p}-\frac{1}{q} & \frac{1}{p}+\frac{2}{q} & \frac{1}{p}-\frac{1}{q} \\ \frac{1}{p}-\frac{1}{q} & \frac{1}{p}-\frac{1}{q} & \frac{1}{p}+\frac{2}{q} \end{pmatrix}$$

in which

$$p = \pm\sqrt{1 + 2\cos\alpha} \quad q = \pm\sqrt{1 - \cos\alpha},$$

which preserves the equivalence of axes. Here the chiralities of the Cartesian and crystallographic axes are the same if p is chosen positive, and different otherwise, and the two sets of axes coincide in projection along the triad if q is chosen positive and are π out of phase otherwise.

3.3.1.1.2. *Homogeneous coordinates*

Homogeneous coordinates have found wide application in computer graphics. For some equipment their use is essential, and they are of value analytically even if the available hardware does not require their use.

Homogeneous coordinates employ four quantities, *X*, *Y*, *Z* and *W*, to define the position of a point, rather than three. The fourth coordinate has a scaling function so that it is the quantity X/W (as delivered to the display hardware) which controls the left–right positioning of the point within the picture. A point with $|X/W| < 1$ is in the picture, normally, and those with $|X/W| > 1$ are outside it, but see Section 3.3.1.3.5.

There are many reasons why homogeneous coordinates may be adopted, among them the following:

(i) *X*, *Y*, *Z* and *W* may be held as integers, thus enabling fast arithmetic whilst offering much of the flexibility of floating-point working. A single *W* value may be common to a whole array of *X*, *Y*, *Z* values.

(ii) Perspective transformations can be implemented without the need for any division. Only high-speed matrix multiplication using integer arithmetic is necessary, provided only that the drawing hardware can provide displacements proportional to the ratio of two signals, *X* and *W* or *Y* and *W*. Rotation, translation, scaling and the application of perspective are all affected by operations of the same form, namely multiplication of a four-vector by a 4×4 matrix. The hardware may thus be kept relatively simple since only one type of operation needs to be provided for.

(iii) Since *kX*, *kY*, *kZ*, *kW* represents the same point as *X*, *Y*, *Z*, *W*, the hardware may be arranged to maximize resolution without risk of integer overflow.

For analytical purposes it is convenient to regard homogeneous transformations in terms of partitioned matrices

$$\begin{pmatrix} \boldsymbol{M} & \mathbf{V} \\ \mathbf{U} & N \end{pmatrix}\begin{pmatrix} \mathbf{X} \\ W \end{pmatrix},$$

where $\boldsymbol{M}$ is a 3×3 matrix, **V** and **X** are three-element column vectors, **U** is a three-element row vector and *N* and *W* are scalars.

Matrices and vectors which are equivalent under the considerations of (iii) above will be related by the sign $\simeq$ in what follows.

Table 3.4.2.2. *Untreated lattice-sum results for the dispersion energy ($n = 6$) of crystalline benzene (kJ mol^{-1}, Å)*

Truncation limit	Number of molecules	Number of terms	Calculated energy
6.0	26	524	−69.227
8.0	51	1313	−76.007
10.0	77	2631	−78.179
12.0	126	4718	−79.241
14.0	177	7531	−79.726
16.0	265	11274	−80.013
18.0	344	15904	−80.178
20.0	439	22049	−80.295
Converged value			−80.589

to Parseval's theorem (described below) this integral is equal to an integral of the product of the two Fourier transforms of the functions. Finally, the integral over the Fourier transforms of the functions is converted to a sum in reciprocal (or Fourier-transform) space. The choice of the convergence function $W(R)$ is not unique; an obvious requirement is that the relevant Fourier transforms must exist and have correct limiting behaviour. Nijboer and DeWette suggested using the incomplete gamma function for $W(R)$. More recently, Fortuin (1977) showed that this choice of convergence function leads to optimal convergence of the sums in both direct and reciprocal space:

$$W(R) = \Gamma(n/2, \pi w^2 R^2)/\Gamma(n/2),$$

where $\Gamma(n/2)$ and $\Gamma(n/2, \pi w^2 R^2)$ are the gamma function and the incomplete gamma function, respectively:

$$\Gamma(n/2, \pi w^2 R^2) = \int_{\pi w^2 R^2}^{\infty} t^{(n/2)-1} \exp(-t)\, \mathrm{d}t$$

and

$$\Gamma(n/2) = \Gamma(n/2, 0).$$

The complement of the incomplete gamma function is

$$\gamma(n/2, \pi w^2 R^2) = \Gamma(n/2) - \Gamma(n/2, \pi w^2 R^2).$$

3.4.4. Preliminary derivation to obtain a formula which accelerates the convergence of an R^{-n} sum over lattice points X(d)

The three-dimensional direct-space crystal lattice is specified by the origin vectors $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_3$. A general vector in direct space is defined as

$$\mathbf{X}(\mathbf{x}) = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3,$$

where x_1, x_2, x_3 are the fractional cell coordinates of $\mathbf{X}$. A lattice vector in direct space is defined as

$$\mathbf{X}(\mathbf{d}) = d_1\mathbf{a}_1 + d_2\mathbf{a}_2 + d_3\mathbf{a}_3,$$

where d_1, d_2, d_3 are integers (specifying particular values of x_1, x_2, x_3) designating a lattice point. V_d is the direct-cell volume which is equal to $\mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3$. A general point in the direct lattice is $\mathbf{X}(\mathbf{x})$; the contents of the lattice are by definition identical as the components of $\mathbf{x}$ are increased or decreased by integer amounts.

The reciprocal-lattice vectors are defined by the relations

$$\begin{aligned} \mathbf{a}_j \cdot \mathbf{b}_k &= 1 \qquad j = k \\ &= 0 \qquad j \neq k. \end{aligned}$$

A general vector in reciprocal space $\mathbf{H}(\mathbf{r})$ is defined as

$$\mathbf{H}(\mathbf{r}) = r_1\mathbf{b}_1 + r_2\mathbf{b}_2 + r_3\mathbf{b}_3.$$

A reciprocal-lattice vector $\mathbf{H}(\mathbf{h})$ is defined by the integer triplet h_1, h_2, h_3 (specifying particular values of r_1, r_2, r_3) so that

$$\mathbf{H}(\mathbf{h}) = h_1\mathbf{b}_1 + h_2\mathbf{b}_2 + h_3\mathbf{b}_3.$$

In other sections of this volume a shortened notation $\mathbf{h}$ is used for the reciprocal-lattice vector. In this section the symbol $\mathbf{H}(\mathbf{h})$ is used to indicate that it is a particular value of $\mathbf{H}(\mathbf{r})$.

The three-dimensional Fourier transform $g(\mathbf{t})$ of a function $f(\mathbf{x})$ is defined by

$$g(\mathbf{t}) = FT_3[f(\mathbf{x})] = \int f(\mathbf{x}) \exp(2\pi i \mathbf{x} \cdot \mathbf{t})\, \mathrm{d}\mathbf{x}.$$

The Fourier transform of the set of points defining the direct lattice is the set of points defining the reciprocal lattice, scaled by the direct-cell volume. It is useful for our purpose to express the lattice transform in terms of the Dirac delta function $\delta(x - x_o)$ which is defined so that for any function $f(\mathbf{x})$

$$f(\mathbf{x}_o) = \int \delta(\mathbf{x} - \mathbf{x}_o) f(\mathbf{x})\, \mathrm{d}\mathbf{x}.$$

We then write

$$FT_3\{\sum_{\mathbf{d}} \delta[\mathbf{X}(\mathbf{x}) - \mathbf{X}(\mathbf{d})]\} = V_d^{-1} \sum_{\mathbf{h}} \delta[\mathbf{H}(\mathbf{r}) - \mathbf{H}(\mathbf{h})].$$

First consider the lattice sum over the direct-lattice points $\mathbf{X}(\mathbf{d})$, relative to a particular point $\mathbf{X}(\mathbf{x}) = \mathbf{R}$, with omission of the origin lattice point.

$$S'(n, \mathbf{R}) = \sum_{\mathbf{d} \neq 0} |\mathbf{X}(\mathbf{d}) - \mathbf{R}|^{-n}.$$

The special case with $\mathbf{R} = 0$ will also be needed:

$$S'(n, 0) = \sum_{\mathbf{d} \neq 0} |\mathbf{X}(\mathbf{d})|^{-n}.$$

Now define a sum of Dirac delta functions

$$f'[\mathbf{X}(\mathbf{d})] = \sum_{d \neq 0} \delta[\mathbf{X}(\mathbf{x}) - \mathbf{X}(\mathbf{d})].$$

Then S' can be represented as an integral

$$S'(n, \mathbf{R}) = \int f'[\mathbf{X}(\mathbf{d})] |\mathbf{X} - \mathbf{R}|^{-n}\, \mathrm{d}\mathbf{X},$$

in which a term is contributed to S' whenever the direct-space vector $\mathbf{X}$ coincides with the lattice vector $\mathbf{X}(\mathbf{d})$, except for $\mathbf{d} = 0$. Now apply the convergence function to S':

$$\begin{aligned} S'(n, \mathbf{R}) = {} & [\Gamma(n/2)]^{-1} \int f'[\mathbf{X}(\mathbf{d})] |\mathbf{X} - \mathbf{R}|^{-n} \\ & \times \Gamma(n/2, \pi w^2 |\mathbf{X} - \mathbf{R}|^2)\, \mathrm{d}\mathbf{X} \\ & + [\Gamma(n/2)]^{-1} \int f'[\mathbf{X}(\mathbf{d})] |\mathbf{X} - \mathbf{R}|^{-n} \\ & \times \gamma(n/2, \pi w^2 |\mathbf{X} - \mathbf{R}|^2)\, \mathrm{d}\mathbf{X}. \end{aligned}$$

The first integral is shown here only for the purpose of giving a consistent representation of S'; in fact, the first integral will be reconverted back into a sum and evaluated in direct space. The second integral will be transformed to reciprocal space using Parseval's theorem [see, for example, Arfken (1970)], which states that

$$\int f(\mathbf{X}) g^*(\mathbf{X})\, \mathrm{d}\mathbf{X} = \int FT_3[f(\mathbf{X})] FT_3[g^*(\mathbf{X})]\, \mathrm{d}\mathbf{H}.$$

4.1. Thermal diffuse scattering of X-rays and neutrons

BY B. T. M. WILLIS

4.1.1. Introduction

Thermal motion of the atoms in a crystal gives rise to a reduction in the intensities of the Bragg reflections and to a diffuse distribution of non-Bragg scattering in the rest of reciprocal space. This distribution is known as thermal diffuse scattering (TDS). Measurement and analysis of TDS gives information about the lattice dynamics of the crystal, *i.e.* about the small oscillatory displacements of the atoms from their equilibrium positions which arise from thermal excitations. Lattice-dynamical models form the basis for interpreting many physical properties – for example, specific heat and thermal conductivity – which cannot be explained by a static model of the crystal.

Reference to a lattice-dynamical model is found in Newton's *Principia*, which contains a discussion of the vibrations of a linear chain of equidistant mass points connected by springs. The model was used to estimate the speed of sound in air. The vibrational properties of a one-dimensional crystal treated as a linear chain of atoms provide the starting point for several modern treatises on the lattice dynamics of crystals.

The classical theory of the dynamics of three-dimensional crystals is based on the treatment of Born & von Kármán (1912, 1913). In this theory, the restoring force on an atom is determined not by the displacement of the atom from its equilibrium position, but by its displacement relative to its neighbours. The atomic motion is then considered in terms of travelling waves, or 'lattice vibrations', extending throughout the whole crystal. These waves are the normal modes of vibration, in which each mode is characterized by a wavevector $\mathbf{q}$, an angular frequency $\omega(\mathbf{q})$ and certain polarization properties.

For twenty years after its publication the Born–von Kármán treatment was eclipsed by the theory of Debye (1912). In the Debye theory the crystal is treated as a continuous medium instead of a discrete array of atoms. The theory gives a reasonable fit to the integral vibrational properties (for example, the specific heat or the atomic temperature factor) of simple monatomic crystals. It fails to account for the form of the frequency distribution function which relates the number of modes and their frequency.

An even simpler model than Debye's is due to Einstein (1907), who considered the atoms in the crystal to be vibrating independently of each other and with the same frequency ω_E. By quantizing the energy of each atom in units of $\hbar\omega_E$, Einstein showed that the specific heat falls to zero at $T = 0$ K and rises asymptotically to the Dulong and Petit value for T much larger than $\hbar\omega_E/k_B$. ($\hbar$ is Planck's constant divided by 2π and k_B is Boltzmann's constant.) His theory accounts satisfactorily for the breakdown of equipartition of energy at low temperatures, but it predicts a more rapid fall-off of specific heat with decreasing temperature than is observed.

Deficiencies in the Debye theory were noted by Blackman (1937), who showed that they are overcome satisfactorily using the more rigorous Born–von Kármán theory. Extensive X-ray studies of Laval (1939) on simple structures such as sylvine, aluminium and diamond showed that the detailed features of the TDS could only be explained in terms of the Born–von Kármán theory. The X-ray work on aluminium was developed further by Olmer (1948) and by Walker (1956) to derive the phonon dispersion relations (see Section 4.1.5) along various symmetry directions in the crystal.

It is possible to measure the vibrational frequencies directly with X-rays, but such measurements are very difficult as lattice vibrational energies are many orders of magnitude less than X-ray energies. The situation is much more favourable with thermal neutrons because their wavelength is comparable with interatomic spacings and their energy is comparable with a quantum of vibrational energy (or phonon). The neutron beam is scattered inelastically by the lattice vibrations, exchanging energy with the phonons. By measuring the energy change for different directions of the scattered beam, the dispersion relations $\omega(\mathbf{q})$ can be determined. Brockhouse & Stewart (1958) reported the first dispersion curves to be derived in this way; since then the neutron technique has become the principal experimental method for obtaining detailed information about lattice vibrations.

In this chapter we shall describe briefly the standard treatment of the lattice dynamics of crystals. There follows a section on the theory of the scattering of X-rays by lattice vibrations, and a similar section on the scattering of thermal neutrons. We then refer briefly to experimental work with X-rays and neutrons. The final section is concerned with the measurement of elastic constants: these constants are required in calculating the TDS correction to measured Bragg intensities (see Section 7.4.2 of *IT* C, 1999).

4.1.2. Dynamics of three-dimensional crystals

For modes of vibration of very long wavelength, the crystal can be treated as a homogeneous elastic continuum without referring to its crystal or molecular structure. The theory of the propagation of these elastic waves is based on Hooke's law of force and on Newton's equations of motion. As the wavelength of the vibrations becomes shorter and shorter and approaches the separation of adjacent atoms, the calculation of the vibrational properties requires a knowledge of the crystal structure and of the nature of the forces between adjacent atoms. The three-dimensional treatment is based on the formulation of Born and von Kármán, which is discussed in detail in the book by Born & Huang (1954) and in more elementary terms in the books by Cochran (1973) and by Willis & Pryor (1975).

Before setting up the equations of motion, it is necessary to introduce three approximations:

(i) *The harmonic approximation*. When an atom is displaced from its equilibrium position, the restoring force is assumed to be proportional to the displacement, measured relative to the neighbouring atoms. The approximation implies no thermal expansion and other properties not possessed by real crystals; it is a reasonable assumption in the lattice-dynamical theory provided the displacements are not too large.

(ii) *The adiabatic approximation*. We wish to set up a potential function for the crystal describing the binding between the atoms. However, the binding involves electronic motions whereas the dynamics involve nuclear motions. The adiabatic approximation, known as the Born–Oppenheimer approximation in the context of molecular vibrations, provides the justification for adopting the same potential function to describe both the binding and the dynamics. Its essence is that the electronic and nuclear motions may be considered separately. This is possible if the nuclei move very slowly compared with the electrons: the electrons can then instantaneously take up a configuration appropriate to that of the displaced nuclei without changing their quantum state. The approximation holds well for insulators, where electronic transition energies are high owing to the large energy gap between filled and unfilled electron states. Surprisingly, it even works for metals, because (on account of the Pauli principle) only a few electrons near the Fermi level can make transitions.

(iii) *Periodic boundary conditions*. These are introduced to avoid problems associated with the free surface. The system is treated as an infinite crystal made up of contiguous, repeating blocks of the actual crystal. The periodic (or cyclic) boundary conditions require that the displacements of corresponding atoms in different blocks are identical. The validity of the conditions was challenged by

4.2. Disorder diffuse scattering of X-rays and neutrons

BY H. JAGODZINSKI AND F. FREY

4.2.1. Scope of this chapter

Diffuse scattering of X-rays, neutrons and other particles is an accompanying effect in all diffraction experiments aimed at structure analysis with the aid of so-called elastic scattering. In this case the momentum exchange of the scattered photon (or particle) includes the crystal as a whole; the energy transfer involved becomes negligibly small and need not be considered in diffraction theory. Inelastic scattering processes, however, are due to excitation processes, such as ionization, phonon scattering *etc.* Distortions as a consequence of structural changes cause typical elastic or inelastic diffuse scattering. All these processes contribute to scattering, and a general theory has to include all of them. Hence, the exact treatment of diffuse scattering becomes very complex. Fortunately, approximations treating the phenomena independently are possible in most cases, but it should be kept in mind that difficulties may occasionally arise.

A separation of elastic from inelastic diffuse scattering may be made if detectors sensitive to the energy of radiation are used. Difficulties may sometimes result from small energy exchanges, which cannot be resolved for experimental reasons. The latter is true for scattering of X-rays by phonons which have energies of the order of 10^{-2}–10^{-3} eV, a value which is considerably smaller than 10 keV, a typical value for X-ray quanta. Another equivalent explanation, frequently forwarded in the literature, is the high speed of X-ray photons, such that the rather slow motion of atoms cannot be 'observed' by them during diffraction. Hence, all movements appear as static displacement waves of atoms, and temperature diffuse scattering is pseudo-elastic for X-rays. This is not true in the case of thermal neutrons, which have energies comparable to those of phonons. Since thermal diffuse scattering is discussed in Chapter 4.1, this chapter is mainly concerned with the elastic (or pseudo-elastic other than thermal) part of diffuse scattering.

The full treatment of the complicated theoretical background for all other kinds of diffuse scattering lies beyond the scope of this article. It is also impossible to refer to all papers in this wide and complicated field. Different theoretical treatments of one and the same subject are often developed, but only some are given here, in most cases those which may be understood most easily – at least to the authors' feeling. As shown in this chapter, electron-density fluctuations and distribution functions of defects play an important role for the complete interpretation of diffraction patterns. Both quantities may best be studied in the low-angle scattering range, which occasionally represents the only Bragg peak dealing with the full information of the distribution function of the defects. Hence, many problems cannot be solved without a detailed interpretation of low-angle diffraction.

Disorder phenomena in magnetic structures are not specifically discussed here. Magnetic diffuse neutron scattering and special experimental techniques themselves constitute a large subject. Many aspects, however, may be analysed along similar lines as given here. For this particular topic the reader is referred to textbooks of neutron scattering, where the theory of diffraction by magnetic materials is generally included (see, *e.g.*, Lovesey, 1984).

Glasses, liquids or liquid crystals show typical diffuse diffraction phenomena. Particle-size effects and strains have an important influence on the diffuse scattering. The same is true for dislocations and point defects such as interstitials or vacancies. These defects are mainly described by their strain field which influences the intensities of sharp reflections like an artificial temperature factor: the Bragg peaks diminish in intensity, while the diffuse scattering increases predominantly close to them. These phenomena are less important from a structural point of view, at least in the case of metals or other simple structures. This statement is true as long as the structure of the 'kernel' of defects may be neglected when compared with the influence of the strain field. Whether dislocations in more complicated structures meet this condition is not yet known.

Radiation damage in crystals represents another field of diffuse scattering which cannot be treated here explicitly. As long as point defects only are generated, the strain field around these defects is the most important factor governing diffuse scattering. Particles with high energy, such as fast neutrons, protons and others, generate complicated defect structures which have to be treated with the aid of the cluster method described below, but no special reference is given here because of the complexity of these phenomena.

Diffuse scattering related to phase transitions, in particular the critical diffuse scattering observed at or close to the transition temperature, cannot be discussed here. In simple cases a satisfactory description may be given with the aid of a 'soft phonon', which freezes at the critical temperature, thus generating typical temperature-dependent diffuse scattering. If the geometry of the lattice is maintained during the transformation (no breakdown into crystallites of different cell geometry), the diffuse scattering is very similar to diffraction phenomena described in this article. Sometimes, however, very complicated interim stages (ordered or disordered) are observed demanding a complicated theory for their full explanation (see, *e.g.*, Dorner & Comes, 1977).

Commensurate and incommensurate modulated structures as well as quasicrystals are frequently accompanied by a typical diffuse scattering, demanding an extensive experimental and theoretical study in order to arrive at a satisfactory explanation. A reliable structure determination becomes very difficult in cases where the interpretation of diffuse scattering has not been incorporated. Many erroneous structural conclusions have been published in the past. The solution of problems of this kind needs careful thermodynamical consideration as to whether a plausible explanation of the structural data can be given.

Obviously, there is a close relationship between thermodynamics and diffuse scattering in disordered systems, representing a stable or metastable thermal equilibrium. From the thermodynamical point of view the system is then characterized by its grand partition function, which is intimately related to the correlation functions used in the interpretation of diffuse scattering. The latter is nothing other than a kind of 'partial partition function' where two atoms, or two cell occupations, are fixed such that the sum of all partial partition functions represents the grand partition function. This fact yields the useful correlation between thermodynamics and diffuse scattering mentioned above, which may well be used for a determination of thermodynamical properties of the crystal. This subject could not be included here for the following reason: real three-dimensional crystals generally exhibit diffuse scattering by defects and/or disordering effects which are not in thermal equilibrium. They are created during crystal growth, or are frozen-in defects formed at higher temperatures. Hence, a thermodynamical interpretation of diffraction data needs a careful study of diffuse scattering as a function of temperature or some other thermodynamical parameters. This can be done in very rare cases only, so the omission of this subject seems justified.

For all of the reasons mentioned above, this article cannot be complete. It is hoped, however, that it will provide a useful guide for those who need the information for the full understanding of the crystal chemistry of a given structure.

There is no comprehensive treatment of all aspects of diffuse scattering. Essential parts are treated in the textbooks of James (1954), Wilson (1962), Wooster (1962) and Schwartz & Cohen (1977); handbook articles are written by Jagodzinski (1963,

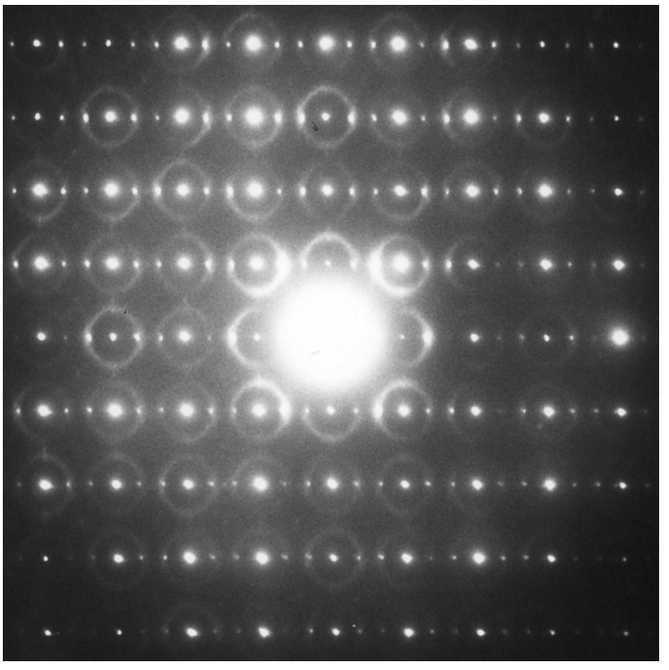

Fig. 4.3.1.2. Electron-diffraction pattern from a disordered crystal of $17Nb_2O_5.48WO_3$ close to the [001] orientation of the tetragonal tungsten-bronze-type structure (Iijima & Cowley, 1977).

These experimental and theoretical aspects of electron diffraction have influenced the ways in which it has been applied in studies of diffuse scattering.

In general, we may distinguish three different approaches to the interpretation of diffuse scattering:

(*a*) The crystallographic way, in which the Patterson- or correlation-function representation of the local order is emphasized, *e.g.* by use of short-range-order parameters.

(*b*) The physical model in terms of excitations. These are usually described in reciprocal (momentum) space: phonons, plasmons *etc.*

(*c*) Structure models in direct space. These must be derived by trial or by chemical considerations of bonds, coordinates *etc.*

Owing to the difficulties of separating the different components in the diffuse scattering, most work on diffuse scattering of electrons has followed one or both of the two last approaches, although Patterson-type interpretation, based upon kinematical scattering including some dynamical corrections, has also been tried.

4.3.2. Inelastic scattering

In the kinematical approximation, a general expression which includes inelastic scattering can be written in the form quoted by Van Hove (1954)

$$I(\mathbf{u},\nu) = \frac{m^3}{2^2h^6}\frac{k}{k_o} \times W(\mathbf{u}) \sum P_{n_o} \sum_{\eta} \sum_{j=1}^{Z} |\langle n_o| \exp\{2\pi i \mathbf{u}\cdot\mathbf{R}_j\}|n\rangle|^2 \times \delta\left(\nu + \frac{E_n - E_{n_o}}{h}\right) \quad (4.3.2.1)$$

for the intensity of scattering as function of energy transfer and momentum transfer from a system of Z identical particles, $\mathbf{R}_j$. Here m and h have their usual meanings; k_o and k, E_{n_o} and E_n are wavevectors and energies before and after the scattering between object states n_o and n; P_{n_o} are weights of the initial states; $W(\mathbf{u})$ is a form factor (squared) for the individual particle.

In equation (4.3.2.1), $\mathbf{u}$ is essentially momentum transfer. When the energy transfer is small ($\Delta E/E \ll \theta$), we can still write $|\mathbf{u}| = 2\sin\theta/\lambda$, then the sum over final states n is readily performed and an expression of the Waller–Hartree type is obtained for the total inelastic scattering as a function of angle:

$$I_{\text{inel}}(\mathbf{u}) \propto \frac{S}{u^4},$$

where

$$S(u) = Z - \sum_{j=1}^{Z} |f_{jj}(u)|^2 - \sum_{j}^{Z}\sum_{\neq k}^{Z} |f_{jk}(u)|^2, \quad (4.3.2.2)$$

and where the one-electron f's for Hartree–Fock orbitals, $f_{jk}(\mathbf{u}) = \langle j|\exp(2\pi i\mathbf{u}\cdot\mathbf{r})|k\rangle$, have been calculated by Freeman (1959, 1960) for atoms up to $Z = 30$. The last sum is over electrons with the same spin only.

The Waller–Hartree formula may be a very good approximation for Compton scattering of X-rays, where most of the scattering occurs at high angles and multiple scattering is no problem. With electrons, it has several deficiencies. It does not take into account the electronic structure of the solid, which is most important at low values of u. It does not include the energy distribution of the scattering. It does not give a finite cross section at zero angle, if u is interpreted as an angle. In order to remedy this, we should go back to equation (4.3.1.2) and decompose $\mathbf{u}$ into two components, one tangential part which is associated with angle in the usual way and one normal component along the beam direction, u_n, which may be related to the excitation energy $\Delta E = E_n - E_{n_o}$ by the expression $u_n = \Delta E_k/2E$. This will introduce a factor $1/(u^2 + u_n^2)$ in the intensity at small angles, often written as $1/(\theta^2 + \theta_E^2)$, with ΔE estimated from ionization energies *etc.* (Strictly speaking, ΔE is not a constant, not even for scattering from one shell. It is a weighted average which will vary with u.)

Calculations beyond this simple adjustment of the Waller–Hartree-type expression are few. Plasmon scattering has been treated on the basis of a nearly free electron model by Ferrel (1957):

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}(\Delta E)\,\mathrm{d}\Omega} = (1/\pi^2 a_H m v^2 N)(-\mathrm{Im}\{1/\varepsilon\})/(\theta^2 + \theta_E^2), \quad (4.3.2.3)$$

where m, v are relativistic mass and velocity of the incident electron, N is the density of the valence electrons and $\varepsilon(\Delta E, \theta)$ their dielectric constant. Upon integration over ΔE:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{E_p}{2\pi a_H m v N}[1/(\theta^2 + \theta_E^2)G(\theta,\theta_c)], \quad (4.3.2.4)$$

where $G(\theta,\theta_c)$ takes account of the cut-off angle Θ_c. Inner-shell excitations have been studied because of their importance to spectroscopy. The most realistic calculations may be those of Leapman *et al.* (1980) where one-electron wavefunctions are determined for the excited states in order to obtain 'generalized oscillator strengths' which may then be used to modify equation (4.3.1.2).

At high energies and high momentum transfer, the scattering will approach that of free electrons, *i.e.* a maximum at the so-called Bethe ridge, $E = h^2u^2/2m$.

A complete and detailed picture of inelastic scattering of electrons as a function of energy and angle (or scattering variable) is lacking, and may possibly be the least known area of diffraction by solids. It is further complicated by the dynamical scattering, which involves the incident and diffracted electrons and also the ejected atomic electron (see *e.g.* Maslen & Rossouw, 1984).

4.4. Scattering from mesomorphic structures

BY P. S. PERSHAN

4.4.1. Introduction

The term mesomorphic is derived from the prefix 'meso-', which is defined in the dictionary as 'a word element meaning middle', and the term '-morphic', which is defined as 'an adjective termination corresponding to morph or form'. Thus, mesomorphic order implies some 'form', or order, that is 'in the middle', or intermediate between that of liquids and crystals. The name liquid crystalline was coined by researchers who found it to be more descriptive, and the two are used synonymously. It follows that a mesomorphic, or liquid-crystalline, phase must have more symmetry than any one of the 230 space groups that characterize crystals.

A major source of confusion in the early liquid-crystal literature was concerned with the fact that many of the molecules that form liquid crystals also form true three-dimensional crystals with diffraction patterns that are only subtly different from those of other liquid-crystalline phases. Since most of the original mesomorphic phase identifications were performed using a 'miscibility' procedure, which depends on optically observed changes in textures accompanying variation in the sample's chemical composition, it is not surprising that some three-dimensional crystalline phases were mistakenly identified as mesomorphic. Phases were identified as being either the same as, or different from, phases that were previously observed (Liebert, 1978; Gray & Goodby, 1984), and although many of the workers were very clever in deducing the microscopic structure responsible for the microscopic textures, the phases were labelled in the order of discovery as smectic-A, smectic-B *etc.* without any attempt to develop a systematic nomenclature that would reflect the underlying order. Although different groups did not always assign the same letters to the same phases, the problem is now resolved and the assignments used in this article are commonly accepted (Gray & Goodby, 1984).

Fig. 4.4.1.1 illustrates the way in which increasing order can be assigned to the series of mesomorphic phases in three dimensions listed in Table 4.4.1.1. Although the phases in this series are the most thoroughly documented mesomorphic phases, there are others not included in the table which we will discuss below.

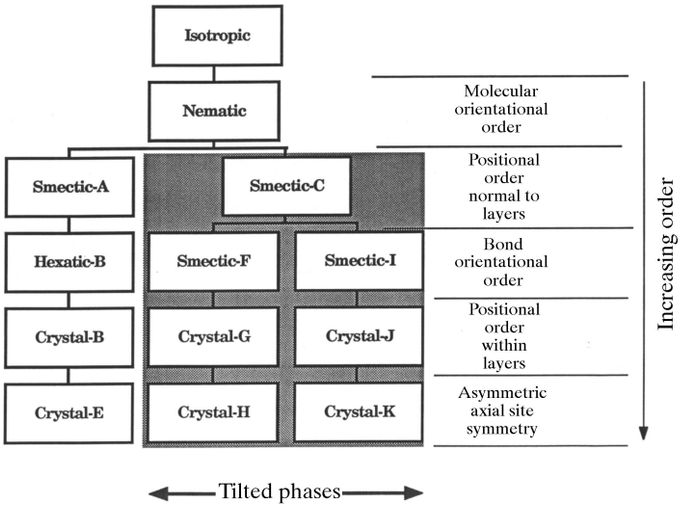

Fig. 4.4.1.1. Illustration of the progression of order throughout the sequence of mesomorphic phases that are based on 'rod-like' molecules. The shaded section indicates phases in which the molecules are tilted with respect to the smectic layers.

The progression from the completely symmetric isotropic liquid through the mesomorphic phases into the crystalline phases can be described in terms of three separate types of order. The first, or the molecular orientational order, describes the fact that the molecules have some preferential orientation analogous to the spin orientational order of ferromagnetic materials. In the present case, the molecular quantity that is oriented is a symmetric second-rank tensor, like the moment of inertia or the electric polarizability, rather than a magnetic moment. This is the only type of long-range order in the nematic phase and as a consequence its physical properties are those of an anisotropic fluid; this is the origin of the name liquid crystal. Fig. 4.4.1.2(*a*) is a schematic illustration of the nematic order if it is assumed that the molecules can be represented by oblong ellipses. The average orientation of the ellipses is aligned; however, there is no long-range order in the relative positions of the ellipses. Nematic phases are also observed for disc-shaped molecules and for clusters of molecules that form micelles. These all share the common properties of being optically anisotropic and fluid-like, without any long-range positional order.

The second type of order is referred to as bond orientational order. Consider, for example, the fact that for dense packing of spheres on a flat surface most of the spheres will have six neighbouring spheres distributed approximately hexagonally around it. If a perfect two-dimensional triangular lattice of indefinite size were constructed of these spheres, each hexagon on the lattice would be oriented in the same way. Within the last few years, we have come to recognize that this type of order, in which the hexagons are everywhere parallel to one another, is possible even when there is no lattice. This type of order is referred to as bond orientational order, and bond orientational order in the absence of a lattice is the essential property defining the hexatic phases (Halperin & Nelson, 1978; Nelson & Halperin, 1979; Young, 1979; Birgeneau & Litster, 1978).

Table 4.4.1.1. *Some of the symmetry properties of the series of three-dimensional phases described in Fig. 4.4.1.1*

The terms LRO and SRO imply long-range or short-range order, respectively, and QLRO refers to 'quasi-long-range order' as explained in the text.

Phase	Molecular orientation order within layer	Bond orientation order	Positional order	
			Normal to layer	Within layer
Smectic-A (SmA)	SRO	SRO	SRO	SRO
Smectic-C (SmC)	LRO	LRO*	SRO	SRO
Hexatic-B	LRO*	LRO	QLRO	SRO
Smectic-F (SmF)	LRO	LRO	QLRO	SRO
Smectic-I (SmI)	LRO	LRO	QLRO	SRO
Crystalline-B (CrB)	LRO	LRO	LRO	LRO
Crystalline-G (CrG)	LRO	LRO	LRO	LRO
Crystalline-J (CrJ)	LRO	LRO	LRO	LRO
Crystalline-E (CrE)	LRO	LRO	LRO	LRO
Crystalline-H (CrH)	LRO	LRO	LRO	LRO
Crystalline-K (CrK)	LRO	LRO	LRO	LRO

* Theoretically, the existence of LRO in the molecular orientation, or tilt, implies that there must be some LRO in the bond orientation and *vice versa*.

its structure is described by too many variable parameters for the parameter space to be explored *a priori*. It is then necessary to phase the fibre diffraction data and construct an electron-density map into which the molecular structure can be fitted and then refined. The second example above would belong to this class. The second class of methods therefore mimics conventional protein crystallography quite closely. The third class of problems applies when the structure is large, but there are too few diffraction data to attempt phasing and the usual determination of atomic coordinates. The solution to such problems varies from case to case and usually involves modelling and optimization of some kind.

An important parameter in structure determination by fibre diffraction is the degree of overlap (that results from the cylindrical averaging) in the data. This parameter is equal to the number of significant terms in equation (4.5.2.17) or the number of independent terms in equation (4.5.2.24), and depends on the position in reciprocal space and, for a polycrystalline fibre, the space-group symmetry. The number of degrees of freedom in a particular datum is equal to twice this number (since each structure factor generally has real and imaginary parts), and is denoted in this section by m. Determination of the $G_{nl}(R)$ from the cylindrically averaged data $I_l(R)$ therefore involves separating the $m/2$ amplitudes $|G_{nl}(R)|$ and assigning phases to each. The electron density can be calculated from the $G_{nl}(R)$ using equations (4.5.2.7) and (4.5.2.11).

4.5.2.6.2. *Helix symmetry, cell constants and space-group symmetry*

The first step in analysis of any fibre diffraction pattern is determination of the molecular helix symmetry u_v. Only the zero-order Bessel term contributes diffracted intensity on the meridian, and referring to equation (4.5.2.6) shows that the zero-order term occurs only on layer lines for which l is a multiple of u. Therefore, inspection of the distribution of diffraction along the meridian allows the value of u to be inferred. This procedure is usually effective, but can be difficult if u is large, because the first meridional maximum may be on a layer line that is difficult to measure. This difficulty was overcome in one case by Franklin & Holmes (1958) by noting that the second Bessel term on the equator is $n = u$, estimating $G_{00}(R)$ using data from a heavy-atom derivative (see Section 4.5.2.6.6), subtracting this from $I_0(R)$, and using the behaviour of the remaining intensity for small R to infer the order of the next Bessel term [using equation (4.5.2.14)] and thence u.

Referring to equations (4.5.2.6) and (4.5.2.14) shows that the distribution of $R_{\min}$ for $0 < l < u$ depends on the value of v. Therefore, inspection of the intensity distribution close to the meridian often allows v to be inferred. Note, however, that the distribution of $R_{\min}$ does not distinguish between the helix symmetries u_v and u_{u-v}. Any remaining ambiguities in the helix symmetry need to be resolved by steric considerations, or by detailed testing of models with the different symmetries against the available data.

For a polycrystalline system, the cell constants are determined from the (R, Z) coordinates of the spots on the diffraction pattern as described in Section 4.5.2.6.4. Space-group assignment is based on analysis of systematic absences, as in conventional crystallography. However, in some cases, because of possible overlap of systematic absences with other reflections, there may be some ambiguity in space-group assignment. However, the space group can always be limited to one of a few possibilities, and ambiguities can usually be resolved during structure determination (Section 4.5.2.6.4).

4.5.2.6.3. *Patterson functions*

In fibre diffraction, the conventional Patterson function cannot be calculated since the individual structure-factor intensities are not available. However, MacGillavry & Bruins (1948) showed that the *cylindrically averaged Patterson function* can be calculated from fibre diffraction data. Consider the function $\hat{Q}(r, z)$ defined by

$$\hat{Q}(r,z) = \sum_{l=0}^{\infty} \int_0^{\infty} \varepsilon_l I_l(R) J_0(2\pi R r) \cos(2\pi l z/c) 2\pi R \, \mathrm{d}R, \quad (4.5.2.58)$$

where $\varepsilon_l = 1$ for $l = 0$ and 2 for $l > 0$, which can be calculated from the intensity distribution on a continuous fibre diffraction pattern. Using equations (4.5.2.7), (4.5.2.10), (4.5.2.17) and (4.5.2.58) shows that $\hat{Q}(r, z)$ is the cylindrical average of the Patterson function, $\hat{P}(r, \varphi, z)$, of one molecule, *i.e.*

$$\hat{Q}(r,z) = (1/2\pi) \int_0^{2\pi} \hat{P}(r,\varphi,z) \, \mathrm{d}\varphi. \quad (4.5.2.59)$$

The $\hat{}$ symbols on $\hat{P}(r, \varphi, z)$ and $\hat{Q}(r, z)$ indicate that these are Patterson functions of a single molecule, as distinct from the usual Patterson function of a crystal, which contains intermolecular interatomic vectors and is periodic with the same periodicity as the crystal. $\hat{P}(r, \varphi, z)$ is periodic only along z and is therefore, strictly, a Patterson function along z and an autocorrelation function along x and y (Millane, 1990*b*). The cylindrically averaged Patterson contains information on interatomic separations along the axial direction and in the lateral plane, but no information on orientations of the vectors in the lateral plane.

For a polycrystalline system; consider the function $Q(r, z)$ given by

$$Q(r,z) = \sum_l \sum_{h,k} R_{hk} I_l(R_{hk}) J_0(2\pi R_{hk} r) \cos(2\pi l z/c), \quad (4.5.2.60)$$

where the sums are over all the overlapped reflections $I_l(R_{hk})$ on the diffraction pattern, given by equation (4.5.2.24). It is easily shown that $Q(r, z)$ is related to the Patterson function $P(r, \varphi, z)$ by

$$Q(r,z) = (1/2\pi) \int_0^{2\pi} P(r,\varphi,z) \, \mathrm{d}\varphi, \quad (4.5.2.61)$$

where, in this case, $P(r, \varphi, z)$ is the usual Patterson function (expressed in cylindrical polar coordinates), *i.e.* it contains all intermolecular (both intra- and inter-unit cell) interatomic vectors and has the same translational symmetry as the unit cell. The cylindrically averaged Patterson function for polycrystalline fibres therefore contains the same information as it does for noncrystalline fibres (*i.e.* no angular information in the lateral plane), except that it also contains information on intermolecular separations.

Low resolution and cylindrical averaging, in addition to the usual difficulties with interpretation of Patterson functions, has resulted in the cylindrically averaged Patterson function not playing a major role in structure determination by fibre diffraction. However, information provided by the cylindrically averaged Patterson function has, in a number of instances, been a useful component in fibre diffraction analyses. A good review of the application of Patterson functions in fibre diffraction is given by Stubbs (1987). Removing data from the low-resolution part (or all) of the equator when calculating the cylindrically averaged Patterson function removes the strong vectors related to axially invariant (or cylindrically symmetric) parts of the map, and can aid interpretation (Namba *et al.*, 1980; Stubbs, 1987). It is also important when calculating cylindrically averaged Patterson functions to use data only at a resolution that is appropriate to the size and spacings of features one is looking for (Stubbs, 1987).

Cylindrically averaged Patterson functions were used in early applications of fibre diffraction analysis (Franklin & Gosling, 1953; Franklin & Klug, 1955). The intermolecular peaks that usually dominate in a cylindrically averaged Patterson function can help to define the locations of multiple molecules in the unit cell.

4.6. Reciprocal-space images of aperiodic crystals

BY W. STEURER AND T. HAIBACH

4.6.1. Introduction

The discovery of materials with icosahedral diffraction symmetry (Shechtman *et al.*, 1984) was the main reason for the reassessment of the definition of *crystallinity* and for the introduction of the concept of *aperiodic crystals*. The first aperiodic crystal, *i.e.* a material with Bragg reflections not located only at reciprocal-lattice nodes, was identified long before (Dehlinger, 1927). In the following decades a wealth of incommensurately modulated phases and composite crystals were discovered. Nevertheless, only a few attempts have been made to develop a crystallography of aperiodic crystals; the most powerful of these was the higher-dimensional approach (see de Wolff, 1974, 1977; Janner & Janssen, 1979, 1980*a*,*b*; de Wolff *et al.*, 1981). In fact, incommensurate structures can be easily described using the higher-dimensional approach and also, fully equivalently, in a dual way: as a three-dimensional (3D) combination of one or more periodic basic structures and one or several modulation waves (de Wolff, 1984). However, with the discovery of quasicrystals and their noncrystalline symmetries, the latter approach failed and geometrical crystallography including the higher-dimensional approach received new attention. For more recent reviews of the crystallography of all three types of aperiodic crystals see van Smaalen (1995), of incommensurately modulated structures see Cummins (1990), of quasicrystals see Steurer (1990, 1996), of quasicrystals and their crystalline approximants see Goldman & Kelton (1993) and Kelton (1995). Textbooks on quasicrystals have been written by Janot (1994) and Senechal (1995).

According to the traditional crystallographic definition, an *ideal crystal* corresponds to an infinite 3D periodic arrangement of identical structure motifs. Its symmetry can be described by one of the 230 3D space groups. Mathematically, a periodic structure can be generated by the convolution of a function representing the structure motif with a lattice function. The structure motif can be given, for instance, by the electron-density distribution $\rho(\mathbf{r})$ of one primitive unit cell of the structure. The lattice function $g(\mathbf{r})$ is represented by a set of δ functions at the nodes $\mathbf{r} = \sum_{i=1}^{3} k_i \mathbf{a}_i$ of a 3D lattice Λ with basis $\mathbf{a}_i$, $i = 1, \ldots, 3$, and $k_i \in \mathbb{Z}$ ($\mathbb{Z}$ is the set of integer numbers). In reciprocal space, this convolution corresponds to the product of the Fourier transform $G(\mathbf{H})$ of the lattice function $g(\mathbf{r})$ and the Fourier transform $F(\mathbf{H}) = \int_v \rho(\mathbf{r}) \exp(2\pi i \mathbf{H}\cdot\mathbf{r})\, \mathrm{d}\mathbf{r}$ of the structure motif $\rho(\mathbf{r})$. $G(\mathbf{H})$ is represented by the reciprocal lattice Λ^* decorated with δ functions on the reciprocal-lattice nodes $\mathbf{H} = \sum_{i=1}^{3} h_i \mathbf{a}_i^*$, with the reciprocal-basis vectors $\mathbf{a}_i^*$, $i = 1, \ldots, 3$, defined by $\mathbf{a}_i \cdot \mathbf{a}_j^* = \delta_{ij}$ and $h_i \in \mathbb{Z}$. The product $G(\mathbf{H}) \times F(\mathbf{H})$ is called the weighted reciprocal lattice; the weights are given by the structure factors $F(\mathbf{H})$. Thus, the characteristic feature of an ideal crystal in direct and reciprocal space is the existence of a lattice. In direct space, this lattice is decorated with identical structure motifs preserving translational and point symmetry in the framework of *space-group symmetry*. In reciprocal space, only the point symmetry between structure factors is maintained. The *Fourier spectrum* (or *Fourier image*, *i.e.* the Fourier transform) of the electron-density distribution of an ideal crystal consists of a countably infinite set of discrete Bragg peaks with a strictly defined minimum distance.

This crystal definition can be generalized to $n > 3$ dimensions. A d-dimensional (dD) *ideal aperiodic crystal* can be defined as a dD irrational section of an n-dimensional (nD, $n > d$) crystal with nD lattice symmetry. The intersection of the nD *hypercrystal* with the dD physical space is equivalent to a projection of the weighted nD reciprocal lattice $\Sigma^* = \left\{\mathbf{H} = \sum_{i=1}^{n} h_i \mathbf{d}_i^* | h_i \in \mathbb{Z}\right\}$ onto the dD physical space. The resulting set (Fourier module) $M^* = \left\{\mathbf{H}^{\parallel} = \sum_{i=1}^{n} h_i \mathbf{a}_i^* | h_i \in \mathbb{Z}\right\}$ is countably dense. *Countably dense* means that the dense set of Bragg peaks can be mapped one-to-one onto the set of natural numbers. Hence, the Bragg reflections can be indexed with integer indices on an appropriate basis. The Fourier module of the projected reciprocal-lattice vectors $\mathbf{H}^{\parallel}$ has the structure of a $\mathbb{Z}$ module of rank n. A $\mathbb{Z}$ module is a free Abelian group, its rank n is given by the number of free generators (rationally independent vectors). The dimension of a $\mathbb{Z}$ module is that of the vector space spanned by it. The vectors $\mathbf{a}_i^*$ are the images of the vectors $\mathbf{d}_i^*$ projected onto the physical space $\mathbf{V}^{\parallel}$. Thus, by definition, the 3D reciprocal space of an ideal *aperiodic crystal* consists of a countably dense set of Bragg reflections only. Contrary to an ideal *crystal*, a minimum distance between Bragg reflections does not exist in an aperiodic one. In summary, it may be stressed that the terms *aperiodic* and *periodic* refer to properties of crystal structures in dD space. In nD space, as considered here, lattice symmetry is always present and, therefore, the term *crystal* is used.

Besides the *aperiodic crystals* mentioned above, other classes of *aperiodic structures* with strictly defined construction rules exist (see Axel & Gratias, 1995). Contrary to the kind of aperiodic crystals dealt with in this chapter, the Fourier spectra of aperiodic structures considered in the latter reference are continuous and contain only in a few cases additional sharp Bragg reflections (δ peaks).

Experimentally, the borderline between aperiodic crystals and their periodic approximations (*crystalline approximants*) is not sharply defined. Finite crystal size, static and dynamic disorder, chemical impurities and defects broaden Bragg peaks and cause diffuse diffraction phenomena. Furthermore, the resolution function of the diffraction equipment is limited.

However, the concept of describing an aperiodic structure as a dD physical-space section of an nD crystal (see Section 4.6.2) is only useful if it significantly simplifies the description of its structural order. Thus, depending on the shape of the *atomic surfaces*, which gives information on the atomic ordering, incommensurately modulated structures (IMSs, Sections 4.6.2.2 and 4.6.3.1), composite structures (CSs, Sections 4.6.2.3 and 4.6.3.2), or quasiperiodic structures (QSs, Sections 4.6.2.4 and 4.6.3.3) can be obtained from irrational cuts. The atomic surfaces are continuous $(n-d)$-dimensional objects for IMSs and CSs, and discrete $(n-d)$-dimensional objects for QSs. A class of aperiodic crystals with discrete fractal atomic surfaces also exists (Section 4.6.2.5). In this case the Hausdorff dimension (Hausdorff, 1919) of the atomic surface is not an integer number and smaller than $n-d$. The most outstanding characteristic feature of a fractal is its scale invariance: the object appears similar to itself 'from near as from far, that is, whatever the scale' (Gouyet, 1996).

To overcome the problems connected with experimental resolution, the translational symmetry of periodic crystals is used as a hard constraint in the course of the determination of their structures. Hence, space-group symmetry is taken for granted and only the local atomic configuration in a unit cell (actually, asymmetric unit) remains to be determined. In reciprocal space, this assumption corresponds to a condensation of Bragg reflections with finite full width at half maximum (FWHM) to δ peaks accurately located at the reciprocal-lattice nodes. Diffuse diffraction phenomena are mostly neglected. This extrapolation to the existence of an ideal crystal is generally out of the question even if samples of very poor quality (high mosaicity, microdomain structure, defects, . . .) are investigated.

The same practice is convenient for the determination of real aperiodic structures once the type of idealized aperiodic ordering is 'known'. Again, the global ordering principle is taken as a hard

Table 4.6.2.1. *Expansion of the Fibonacci sequence* $B_n = \sigma^n(\mathrm{L})$ *by repeated action of the substitution rule* σ: $\mathrm{S} \to \mathrm{L}$, $\mathrm{L} \to \mathrm{LS}$

ν_L, ν_S are the frequencies of the letters L and S in word B_n.

B_n	ν_L	ν_S	n
L	1	0	0
LS	1	1	1
LSL	2	1	2
LSLLS	3	2	3
LSLLSLSL	5	3	4
LSLLSLSLLSLLS	8	5	5
LSLLSLSLLSLLSLSLLSLSL	13	8	6
	⋮	⋮	⋮
	F_{n+1}	F_n	n

$$\tau = 1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \ldots}}},$$

contains only the number 1. This means that τ is the 'most irrational' number, *i.e.* the irrational number with the worst truncated continued fraction approximation to it. This might be one of the reasons for the stability of quasiperiodic systems, where τ plays a role. The strong irrationality may impede the lock-in into commensurate systems (*rational approximants*).

By associating intervals (*e.g.* atomic distances) with length ratio τ to 1 to the letters L and S, a quasiperiodic structure $s(\mathbf{r})$ (*Fibonacci chain*) can be obtained. The invariance of the ratio of lengths $\mathrm{L/S} = (\mathrm{L} + \mathrm{S})/\mathrm{L} = \tau$ is responsible for the invariance of the Fibonacci chain under scaling by a factor $\tau^n, n \in \mathbb{Z}$. Owing to a minimum atomic distance S in real crystal structures, the full set of inverse symmetry operators τ^{-n} does not exist. Consequently, the set of scaling operators $s = \{\tau^0 = 1, \tau^1, \ldots\}$ forms only a semigroup, *i.e.* an associative groupoid. Groupoids are the most general algebraic sets satisfying only one of the group axioms: the associative law. The scaling properties of the Fibonacci sequence can be derived from the eigenvalues of the scaling matrix S. For this purpose the equation

$$\det |S - \lambda I| = 0$$

with eigenvalue λ and unit matrix I has to be solved. The evaluation of the determinant yields the characteristic polynomial

$$\lambda^2 - \lambda - 1 = 0,$$

yielding in turn the eigenvalues $\lambda_1 = [1 + (5)^{1/2}]/2 = \tau$, $\lambda_2 = [1 - (5)^{1/2}]/2 = -1/\tau$ and the eigenvectors $\mathbf{w}_1 = \begin{pmatrix} 1 \\ \tau \end{pmatrix}$, $\mathbf{w}_2 = \begin{pmatrix} 1 \\ -1/\tau \end{pmatrix}$. Rewriting the eigenvalue equation gives for the first (*i.e.* the largest) eigenvalue

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \tau \end{pmatrix} = \begin{pmatrix} \tau \\ 1 + \tau \end{pmatrix} = \begin{pmatrix} \tau \\ \tau^2 \end{pmatrix} = \tau \begin{pmatrix} 1 \\ \tau \end{pmatrix}.$$

Identifying the eigenvector $\begin{pmatrix} 1 \\ \tau \end{pmatrix}$ with $\begin{pmatrix} \mathrm{S} \\ \mathrm{L} \end{pmatrix}$ shows that an infinite Fibonacci sequence $s(\mathbf{r})$ remains invariant under scaling by a factor τ. This scaling operation maps each new lattice vector $\tau\mathbf{r}$ upon a vector $\mathbf{r}$ of the original lattice:

$$s(\tau\mathbf{r}) = s(\mathbf{r}).$$

Considering periodic lattices, these eigenvalues are integer numbers. For quasiperiodic 'lattices' (*quasilattices*) they always correspond to *algebraic numbers* (*Pisot numbers*). A Pisot number is the solution of a polynomial equation with integer coefficients. It is larger than one, whereas the modulus of its conjugate is smaller than unity: $\lambda_1 > 1$ and $|\lambda_2| < 1$ (Luck *et al.*, 1993). The total lengths l_n^A and l_n^B of the words A_n, B_n can be determined from the equations $l_n^A = \lambda_1^n l^A$ and $l_n^B = \lambda_1^n l^B$ with the eigenvalue λ_1. The left Perron–Frobenius eigenvector $\mathbf{w}_1$ of S, *i.e.* the left eigenvector associated with λ_1, determines the ratio S:L to 1:τ. The right Perron–Frobenius eigenvector $\mathbf{w}_1$ of S associated with λ_1 gives the relative frequencies, 1 and τ, for the letters S and L (for a definition of the Perron–Frobenius theorem see Luck *et al.*, 1993, and references therein).

The general case of an alphabet $A = \{\mathrm{L}_1 \ldots \mathrm{L}_k\}$ with k letters (intervals) L_i, of which at least two are on incommensurate length scales and which transform with the substitution matrix S,

$$\mathrm{L}_i' \to \sum_{j=1}^{k} S_{ij} \mathrm{L}_j,$$

can be treated analogously. S is a $k \times k$ matrix with non-negative integer coefficients. Its eigenvalues are solutions of a polynomial equation of rank k with integer coefficients (algebraic or Pisot numbers). The dimension n of the embedding space is generically equal to the number of letters (intervals) k involved by the substitution rule. From substitution rules, infinitely many different 1D quasiperiodic sequences can be generated. However, their atomic surfaces in the nD description are generically of fractal shape (see Section 4.6.2.5).

The quasiperiodic 1D density distribution $\rho(\mathbf{r})$ of the Fibonacci chain can be represented by the Fourier series

$$\rho(\mathbf{r}) = (1/V) \sum_{\mathbf{H}^\parallel} F(\mathbf{H}^\parallel) \exp(-2\pi i \mathbf{H}^\parallel \cdot \mathbf{r}),$$

with $\mathbf{H}^\parallel \in \mathbb{R}$ (the set of real numbers). The Fourier coefficients $F(\mathbf{H}^\parallel)$ form a Fourier module $M^* = \{\mathbf{H}^\parallel = \sum_{i=1}^{2} h_i \mathbf{a}_i^* | h_i \in \mathbb{Z}\}$ equivalent to a $\mathbb{Z}$ module of rank 2. Thus a periodic function in 2D space can be defined by

$$\rho(\mathbf{r}^\parallel, \mathbf{r}^\perp) = (1/V) \sum_{\mathbf{H}} F(\mathbf{H}) \exp[-2\pi i (\mathbf{H}^\parallel \cdot \mathbf{r}^\parallel + \mathbf{H}^\perp \cdot \mathbf{r}^\perp)],$$

where $\mathbf{r} = (\mathbf{r}^\parallel, \mathbf{r}^\perp) \in \Sigma$ and $\mathbf{H} = (\mathbf{H}^\parallel, \mathbf{H}^\perp) \in \Sigma^*$ are, by construction, direct and reciprocal lattice vectors (Figs. 4.6.2.8 and 4.6.2.9):

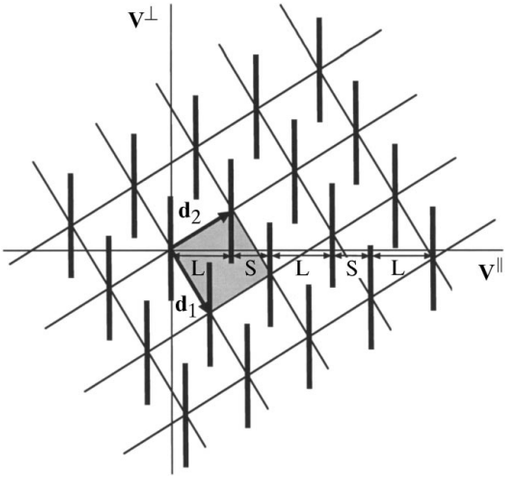

Fig. 4.6.2.8. 2D embedding of the Fibonacci chain. The short and long distances S and L, generated by the intersection of the atomic surfaces with the physical space $\mathbf{V}^\parallel$, are indicated. The atomic surfaces are represented by bars parallel to $\mathbf{V}^\perp$. Their lengths correspond to the projection of one unit cell (shaded) upon $\mathbf{V}^\perp$.

5.1. Dynamical theory of X-ray diffraction

BY A. AUTHIER

5.1.1. Introduction

The first experiment on X-ray diffraction by a crystal was performed by W. Friedrich, P. Knipping and M. von Laue in 1912 and Bragg's law was derived in 1913 (Bragg, 1913). Geometrical and dynamical theories for the intensities of the diffracted X-rays were developed by Darwin (1914*a*,*b*). His dynamical theory took into account the interaction of X-rays with matter by solving recurrence equations that describe the balance of partially transmitted and partially reflected amplitudes at each lattice plane. This is the first form of the dynamical theory of X-ray diffraction. It gives correct expressions for the reflected intensities and was extended to the absorbing-crystal case by Prins (1930). A second form of dynamical theory was introduced by Ewald (1917) as a continuation of his previous work on the diffraction of optical waves by crystals. He took into account the interaction of X-rays with matter by considering the crystal to be a periodic distribution of dipoles which were excited by the incident wave. This theory also gives the correct expressions for the reflected and transmitted intensities, and it introduces the fundamental notion of a wavefield, which is necessary to understand the propagation of X-rays in perfect or deformed crystals. Ewald's theory was later modified by von Laue (1931), who showed that the interaction could be described by solving Maxwell's equations in a medium with a continuous, triply periodic distribution of dielectric susceptibility. It is this form which is most widely used today and which will be presented in this chapter.

The geometrical (or kinematical) theory, on the other hand, considers that each photon is scattered only once and that the interaction of X-rays with matter is so small it can be neglected. It can therefore be assumed that the amplitude incident on every diffraction centre inside the crystal is the same. The total diffracted amplitude is then simply obtained by adding the individual amplitudes diffracted by each diffracting centre, taking into account only the geometrical phase differences between them and neglecting the interaction of the radiation with matter. The result is that the distribution of diffracted amplitudes in reciprocal space is the Fourier transform of the distribution of diffracting centres in physical space. Following von Laue (1960), the expression *geometrical theory* will be used throughout this chapter when referring to these geometrical phase differences.

The first experimentally measured reflected intensities were not in agreement with the theoretical values obtained with the more rigorous dynamical theory, but rather with the simpler geometrical theory. The integrated reflected intensities calculated using geometrical theory are proportional to the square of the structure factor, while the corresponding expressions calculated using dynamical theory for an infinite perfect crystal are proportional to the modulus of the structure factor. The integrated intensity calculated by geometrical theory is also proportional to the volume of the crystal bathed in the incident beam. This is due to the fact that one neglects the decrease of the incident amplitude as it progresses through the crystal and a fraction of it is scattered away. According to geometrical theory, the diffracted intensity would therefore increase to infinity if the volume of the crystal was increased to infinity, which is of course absurd. The theory only works because the strength of the interaction is very weak and if it is applied to very small crystals. How small will be shown quantitatively in Sections 5.1.6.5 and 5.1.7.2. Darwin (1922) showed that it can also be applied to large imperfect crystals. This is done using the model of mosaic crystals (Bragg *et al.*, 1926). For perfect or nearly perfect crystals, dynamical theory should be used. Geometrical theory presents another drawback: it gives no indication as to the phase of the reflected wave. This is due to the fact that it is based on the Fourier transform of the electron density limited by the external shape of the crystal. This is not important when one is only interested in measuring the reflected intensities. For any problem where the phase is important, as is the case for multiple reflections, interference between coherent blocks, standing waves *etc.*, dynamical theory should be used, even for thin or imperfect crystals.

Until the 1940s, the applications of dynamical theory were essentially intensity measurements. From the 1950s to the 1970s, applications were related to the properties (absorption, interference, propagation) of wavefields in perfect or nearly perfect crystals: anomalous transmission, diffraction of spherical waves, interpretation of images on X-ray topographs, accurate measurement of form factors, lattice-parameter mapping. In recent years, they have been concerned mainly with crystal optics, focusing and the design of monochromators for synchrotron radiation [see, for instance, Batterman & Bilderback (1991)], the location of atoms at crystal surfaces and interfaces using the standing-waves method [see, for instance, the reviews by Authier (1989) and Zegenhagen (1993)], attempts to determine phases using multiple reflections [see, for instance, Chang (1987) and Hümmer & Weckert (1995)], characterization of the crystal perfection of epilayers and superlattices by high-resolution diffractometry [see, for instance, Tanner (1990) and Fewster (1993)], *etc.*

For reviews of dynamical theory, see Zachariasen (1945), von Laue (1960), James (1963), Batterman & Cole (1964), Authier (1970), Kato (1974), Brümmer & Stephanik (1976), Pinsker (1978), Authier & Malgrange (1998), and Authier (2001). Topography is described in Chapter 2.7 of *IT* C (1999), in Tanner (1976) and in Tanner & Bowen (1992). For the use of Bragg-angle measurements for accurate lattice-parameter mapping, see Hart (1981).

A reminder of some basic concepts in electrodynamics is given in Section A5.1.1.1 of the Appendix.

5.1.2. Fundamentals of plane-wave dynamical theory

5.1.2.1. *Propagation equation*

The wavefunction Ψ associated with an electron or a neutron beam is *scalar* while an electromagnetic wave is a *vector* wave. When propagating in a medium, these waves are solutions of a *propagation equation*. For electrons and neutrons, this is Schrödinger's equation, which can be rewritten as

$$\Delta\Psi + 4\pi^2 k^2(1+\chi)\Psi = 0, \qquad (5.1.2.1)$$

where $k = 1/\lambda$ is the wavenumber in a vacuum, $\chi = \varphi/W$ (φ is the potential in the crystal and W is the accelerating voltage) in the case of electron diffraction and $\chi = -2mV(\mathbf{r})/h^2k^2$ [$V(\mathbf{r})$ is the Fermi pseudo-potential and h is Planck's constant] in the case of neutron diffraction. The dynamical theory of electron diffraction is treated in Chapter 5.2 [note that a different convention is used in Chapter 5.2 for the scalar wavenumber: $k = 2\pi/\lambda$; compare, for example, equation (5.2.2.1) and its equivalent, equation (5.1.2.1)] and the dynamical theory of neutron diffraction is treated in Chapter 5.3.

In the case of X-rays, the propagation equation is deduced from Maxwell's equations after neglecting the interaction with protons. Following von Laue (1931, 1960), it is assumed that the positive charge of the nuclei is distributed in such a way that the medium is everywhere locally neutral and that there is no current. As a first approximation, magnetic interaction, which is very weak, is not taken into account in this review. The propagation equation is

5.2. Dynamical theory of electron diffraction

BY A. F. MOODIE, J. M. COWLEY AND P. GOODMAN

5.2.1. Introduction

Since electrons are charged, they interact strongly with matter, so that the single scattering approximation has a validity restricted to thin crystals composed of atoms of low atomic number. Further, at energies of above a few tens of keV, the wavelength of the electron is so short that the geometry of two-beam diffraction can be approximated in only small unit cells.

It is therefore necessary to develop a scattering theory specific to electrons and, preferably, applicable to imaging as well as to diffraction. The development, started by Born (1926) and Bethe (1928), and continuing into the present time, is the subject of an extensive literature, which includes reviews [for instance: Howie (1978), Humphreys (1979)] and historical accounts (Goodman, 1981), and is incorporated in Chapter 5.1. Here, an attempt will be made to present only that outline of the main formulations which, it is hoped, will help the nonspecialist in the use of the tables. No attempt will be made to follow the historical development, which has been tortuous and not always logical, but rather to seek the simplest and most transparent approach that is consistent with brevity. Only key points in proofs will be sketched in an attempt to display the nature, rather than the rigorous foundations of the arguments.

5.2.2. The defining equations

No many-body effects have yet been detected in the diffraction of fast electrons, but the velocities lie well within the relativistic region. The one-body Dirac equation would therefore appear to be the appropriate starting point. Fujiwara (1962), using the scattering matrix, carried through the analysis for forward scattering, and found that, to a very good approximation, the effects of spin are negligible, and that the solution is the same as that obtained from the Schrödinger equation provided that the relativistic values for wavelength and mass are used. In effect a Klein–Gordon equation (Messiah, 1965) can be used in electron diffraction (Buxton, 1978) in the form

$$\nabla^2\psi_b + \frac{8\pi^2 m|e|\varphi}{h^2}\psi_b + \frac{8\pi^2 m_0|e|W}{h^2}\left(1+\frac{|e|W}{2m_0c^2}\right)\psi_b = 0.$$

Here, W is the accelerating voltage and φ, the potential in the crystal, is defined as being positive. The relativistic values for mass and wavelength are given by $m = m_0(1-v^2/c^2)^{-1/2}$, and taking '$e$' now to represent the modulus of the electronic charge, $|e|$,

$$\lambda = h[2m_0eW(1+eW/2m_0c^2)]^{-1/2},$$

and the wavefunction is labelled with the subscript b in order to indicate that it still includes back scattering, of central importance to LEED (low-energy electron diffraction).

In more compact notation,

$$[\nabla^2 + k^2(1+\varphi/W)]\psi_b = (\nabla^2 + k^2 + 2k\sigma\varphi)\psi_b = 0. \quad (5.2.2.1)$$

Here $k = |\mathbf{k}|$ is the scalar wavenumber of magnitude $2\pi/\lambda$, and the interaction constant $\sigma = 2\pi me\lambda/h^2$. This constant is approximately 10^{-3} for 100 kV electrons.

For fast electrons, φ/W is a slowly varying function on a scale of wavelength, and is small compared with unity. The scattering will therefore be peaked about the direction defined by the incident beam, and further simplification is possible, leading to a forward-scattering solution appropriate to HEED (high-energy electron diffraction).

5.2.3. Forward scattering

A great deal of geometric detail can arise at this point and, further, there is no generally accepted method for approximation, the various procedures leading to numerically negligible differences and to expressions of precisely the same form. Detailed descriptions of the geometry are given in the references.

The entrance surface of the specimen, in the form of a plate, is chosen as the x, y plane, and the direction of the incident beam is taken to be close to the z axis. Components of the wavevector are labelled with suffixes in the conventional way; $\mathbf{K}_0 = \mathbf{k}_x + \mathbf{k}_y$ is the transverse wavevector, which will be very small compared to $\mathbf{k}_z$. In this notation, the excitation error for the reflection is given by

$$\zeta_{\mathbf{h}} = \frac{K_0^2 - |\mathbf{K}_0 + 2\pi\mathbf{h}|^2}{4\pi|\mathbf{k}_z|}.$$

An intuitive method argues that, since $\varphi/W \ll 1$, then the component of the motion along z is little changed by scattering. Hence, making the substitution $\psi_b = \psi\exp\{ik_z z\}$ and neglecting $\partial^2\psi/\partial z^2$, equation (5.2.2.1) becomes

$$\frac{\partial\psi}{\partial z} = i\left[\frac{1}{2k_z}(\nabla^2_{x,y} + K_0^2) + \sigma\varphi\right]\psi, \quad (5.2.3.1)$$

where

$$\nabla^2_{x,y} \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2},$$

and $\psi(x,y,0) = \exp\{i(k_x x + k_y y)\}$.

Equation (5.2.3.1) is of the form of a two-dimensional time-dependent Schrödinger equation, with the z coordinate replacing time. This form has been extensively discussed. For instance, Howie (1966) derived what is essentially this equation using an expansion in Bloch waves, Berry (1971) used a Green function in a detailed and rigorous derivation, and Goodman & Moodie (1974), using methods due to Feynman, derived the equation as the limit of the multislice recurrence relation. A method due to Corones *et al.* (1982) brings out the relationship between the HEED and LEED equations. Equation (5.2.2.1) is cast in the form of a first-order system,

$$\frac{\partial}{\partial z}\begin{pmatrix}\psi_b \\ \partial\psi_b/\partial z\end{pmatrix} = \begin{pmatrix}0 & 1 \\ -(\nabla^2_{x,y} + k^2 + 2k\sigma\varphi) & 0\end{pmatrix}\begin{pmatrix}\psi_b \\ \partial\psi_b/\partial z\end{pmatrix}.$$

A splitting matrix is introduced to separate the wavefunction into the forward and backward components, $\psi_b^{\pm}$, and the fast part of the phase is factored out, so that $\psi_b^{\pm} = \psi^{\pm}\exp\{\pm ik_z z\}$. In the resulting matrix differential equation, the off-diagonal terms are seen to be small for fast electrons, and equation (5.2.2.1) reduces to the pair of equations

$$\frac{\partial\psi^{\pm}}{\partial z} = \pm i\left[\frac{1}{2k_z}(\nabla^2_{x,y} + K_0^2) + \sigma\varphi\right]\psi^{\pm}. \quad (5.2.3.2)$$

The equation for ψ^{+} is the Lontovich & Fock (1946) parabolic equation.

5.2.4. Evolution operator

Equation (5.2.3.1) is a standard and much studied form, so that many techniques are available for the construction of solutions. One of the most direct utilizes the causal evolution operator. A recent account is given by Gratias & Portier (1983).

5.3. Dynamical theory of neutron diffraction

BY M. SCHLENKER AND J.-P. GUIGAY

5.3.1. Introduction

Neutron and X-ray scattering are quite similar both in the geometry of scattering and in the orders of magnitude of the basic quantities. When the neutron spin is neglected, *i.e.* when dealing with scattering by perfect non-magnetic crystals, the formalism and the results of the dynamical theory of X-ray scattering can be very simply transferred to the case of neutrons (Section 5.3.2). Additional features of the neutron case are related to the neutron spin and appear in diffraction by magnetic crystals (Section 5.3.3). The low intensities available, coupled with the low absorption of neutrons by most materials, make it both necessary and possible to use large samples in standard diffraction work. The effect of extinction in crystals that are neither small nor bad enough to be amenable to the kinematical approximation is therefore very important in the neutron case, and will be discussed in Section 5.3.4 together with the effect of crystal distortion. Additional possibilities arise in the neutron case because the neutrons can be manipulated from outside through applied fields (Section 5.3.5). Reasonably extensive tests of the predictions of the dynamical theory of neutron diffraction have been performed, with the handicap of the very low intensities of neutron beams as compared with X-rays: these are described in Section 5.3.6. Finally, the applications of the dynamical theory in the neutron case, and in particular neutron interferometry, are reviewed in Section 5.3.7.

5.3.2. Comparison between X-rays and neutrons with spin neglected

5.3.2.1. *The neutron and its interactions*

An excellent introductory presentation of the production, properties and scattering properties of neutrons is available (Scherm & Fåk, 1993, and other papers in the same book). A stimulating review on neutron optics, including diffraction by perfect crystals, has been written by Klein & Werner (1983). X-rays and neutrons are compared in terms of the basic quantities in Table 4.1.3.1 of *IT* C (1999), where Chapter 4.4 is devoted to neutron techniques.

The neutron is a massive particle for which the values relevant to diffraction are: no electric charge, rest mass $m = 1.675 \times 10^{-27}$ kg, angular momentum eigenvalues along a given direction $\pm\hbar/2$ (spin $\frac{1}{2}$) and a magnetic moment of -1.913 nuclear magneton, meaning that its component along a quantization direction z can take eigenvalues $\mu_z = \mp 0.996 \times 10^{-26}$ A m^2. The de Broglie wavelength is $\lambda = h/p$ where h is Planck's constant ($h = 2\pi\hbar = 6.625 \times 10^{-34}$ J s) and p is the linear momentum; $p = mv$ in the non-relativistic approximation, which always applies in the context of this chapter, v being the neutron's velocity. The neutron's wavelength, λ, and kinetic energy, E_c, are thus related by $\lambda = h/(2mE_c)^{1/2}$, or, in practical units, λ [Å] $= 9.05/(E_c$ [meV]$)^{1/2}$. Thus, to be of interest for diffraction by materials, neutrons should have kinetic energies in the range 10^0 to 10^2 meV. In terms of the velocity, λ [Å] $= 3.956/(v$ [km s^{-1}]).

Neutron beams are produced by nuclear reactors or by spallation sources, usually pulsed. In either case they initially have an energy in the MeV range, and have to lose most of it before they can be used. The moderation process involves inelastic interactions with materials. It results in statistical distributions of energy, hence of velocity, close to the Maxwell distribution characteristic of the temperature T of the moderator. Frequently used moderators are liquid deuterium (D_2, *i.e.* $^2H_2$) at 25 K, heavy water (D_2O) at room temperature and graphite allowed to heat up to 2400 K; the corresponding neutron distributions are termed cold, thermal and hot, respectively.

The interaction of a neutron with an atom is usually described in terms of scattering lengths or of scattering cross sections. The main contribution corresponding to the nuclear interaction is related to the strong force. The interaction with the magnetic field created by atoms with electronic magnetic moments is comparable in magnitude to the nuclear term.

5.3.2.2. *Scattering lengths and refractive index*

The elastic scattering amplitude for scattering vector $\mathbf{s}$, $f(\mathbf{s})$, is defined by the wave scattered by an object placed at the origin when the incident plane wave is $\Psi_i = A\exp[i(\mathbf{k}_0 \cdot \mathbf{r} - \omega t)]$, written as $\Psi_s = A[f(\mathbf{s})/r]\exp[i(kr - \omega t)]$ with $k = |\mathbf{k}_0| = |\mathbf{k}_0 + \mathbf{s}| = 2\pi/\lambda$. In the case of the strong-force interaction with nuclei, the latter can be considered as point scatterers because the interaction range is very small, hence the scattering amplitude is isotropic (independent of the direction of $\mathbf{s}$). It is also independent of λ except in the vicinity of resonances. It is conventionally written as $-b$ so that most values of b, called the scattering length, are positive. A table of experimentally measured values of the scattering lengths b is given in *IT* C for the elements in their natural form as well as for many individual isotopes. It is apparent that the typical order of magnitude is the fm (femtometer, *i.e.* 10^{-15} m, or fermi), that there is no systematic variation with atomic number and that different isotopes have very different scattering lengths, including different signs. The first remark implies that scattering amplitudes of X-rays and of neutrons have comparable magnitudes, because the characteristic length for X-ray scattering (the scattering amplitude for forward scattering by one free electron) is $R = 2.8$ fm, the classical electron radius. The second and third points explain the importance of neutrons in structural crystallography, in diffuse scattering and in small-angle scattering. Scattering of neutrons by condensed matter implies the use of the bound scattering lengths, as tabulated in *IT* C. The 'free' scattering length, used in some presentations, is obtained by multiplying the bound scattering lengths by $A/(A+1)$, where A is the mass of the nucleus in atomic units.

A description in terms of an interaction potential is possible using the Fermi pseudo-potential, which in the case of the nuclear interaction with a nucleus at $\mathbf{r}_0$ can be written as $V(\mathbf{r}) = (h^2/2\pi m)b\delta(\mathbf{r} - \mathbf{r}_0)$, where δ denotes the three-dimensional Dirac distribution.

Refraction of neutrons at an interface can be conveniently described by assigning a refractive index to the material, such that the wavenumber in the material, k, is related to that in a vacuum, k_0, by $k = nk_0$. Here

$$n = \left(1 - \frac{\lambda^2}{\pi V}\sum_i b_i\right)^{1/2},$$

where the sum is over the nuclei contained in volume V. With typical values, n is very close to 1 and $1 - n = (\lambda^2/2\pi V)\sum_i b_i$ is typically of the order of 10^{-5}. This small value, in the same range as for X-rays, gives a feeling for the order of magnitude of key quantities of the dynamical theory, in particular the Darwin width 2δ as discussed in Chapter 5.1. It also makes total external reflection possible on materials for which $\sum_i b_i > 0$: this is the basis for the neutron guide tubes now installed in most research reactors, as well as for reflectometry.

The notations prevailing in X-ray and in neutron crystallography are slightly different, and the correspondence is very simple: X-ray atomic scattering factors and structure factors are numbers. When