Acta Cryst. (1987). A**43**, 30–36

An Isomorphous Replacement Method for Phasing Twinned Structures

BY T. O. YEATES AND D. C. REES

Department of Chemistry and Biochemistry and Molecular Biology Institute, University of California, Los Angeles, CA 90024, USA

(*Received* 5 *August* 1985; *accepted* 12 *May* 1986)

Abstract

A linear least-squares formulation of the method of isomorphous replacement is presented. With data from untwinned crystals, this approach is shown to be equivalent to the phasing representation developed by Hendrickson & Lattman [*Acta Cryst.* (1970). B**26**, 136–143]. A general method for calculating the most probable phase is described and applied to the higher-dimensional problem of phase determination for twinned structures. A method for calculating the best phase with intensity data from twinned crystals is also presented. The dependences of these phasing procedures on the number of derivatives and accuracy of the data sets are evaluated in test calculations.

Introduction

In a structure determination confronted with twinning by merohedry, two options are generally available: either (i) the observed intensities must be corrected for twinning, or (ii) crystallization conditions yielding untwinned crystals must be found. The former is possible provided that the twinning fraction is sufficiently low (Grainger, 1969), while the latter may or may not be possible. If neither choice can be satisfied, the project is effectively stopped. The fundamental problem with intensity data measured from twins by merohedry is that the square root of the observed intensities is not simply related to the structure-factor amplitudes of individual reflections. Without direct knowledge of $|F|$, conventional approaches for determining phases by the method of isomorphous replacement (Blow & Crick, 1959) are not feasible.

Inability to solve the phase problem is the basic stumbling block to the determination of twinned structures. As described below, however, the phasing problem may be recast in such a way as to permit direct calculation of the complex protein structure factors from isomorphous replacement data obtained from a partially or perfectly twinned structure. [A similar problem of phasing cylindrically averaged diffraction patterns has been treated by Stubbs & Diamond (1975).] With this information, electron density maps of the untwinned structure may be calculated, thus permitting the structure determination to proceed.

Theoretical background

The basic approach will be illustrated explicitly for the method of isomorphous replacement in the absence of twinning, but the extension to include twinning is straightforward. In an isomorphous replacement experiment, the intensities $|F_P|^2$ and $|F_{H_j}|^2$ corresponding to the native and jth heavy-atom derivative are measured. If the heavy atoms have been located, the scattering factor of these groups, $f_j = a_j + ib_j$, may be calculated. In the absence of errors, the following relationship holds:

$$F_{H_j}^2 = F_P^2 + f_j^2 + 2a_jA + 2b_jB \qquad (1)$$

where A and B are the real and imaginary parts of the native structure factor. Since all terms except A and B may be calculated or experimentally determined, (1) describes a line in (A, B) space. If a second isomorphous derivative is available, a different, but analogous, equation corresponding to (1) will be obtained. These two lines should intersect at a point corresponding to the correct values of A and B for the native structure factor. Furthermore, these values for A and B must be on the circle defined by the constraint $A^2 + B^2 = F_P^2$. These geometric relationships are illustrated in Fig. 1.

In the absence of measurement errors, the value for (A, B) may be obtained by direct solution of the linear equations (1) (Ramachandran & Srinivasan, 1970). Introduction of experimental errors leads to a more complicated situation, however, since the lines and circle need not intersect at a single point (Fig. 2). This problem may be solved as follows. The heavy-atom equations (1) may be generally represented by

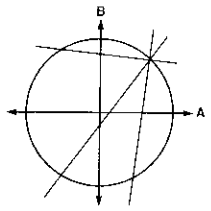

Fig. 1. Linear relationships representing the method of isomorphous replacement. The phasing lines [(1)] corresponding to three independent derivatives are illustrated for the case of perfect data. The circle describes the condition $A^2 + B^2 = F_P^2$. The common point of intersection indicates the native structure factor.

0108-7673/87/010030-07$01.50

the system of equations:

$$\begin{aligned} a_1A + b_1B &= (F_{H_1}^2 - F_P^2 - f_1^2)/2 \\ &= c_1 \\ a_2A + b_2B &= c_2 \\ \vdots \quad &\quad \vdots \end{aligned} \tag{2}$$

These equations may be written in matrix form:

$$\mathbf{Gx} = \mathbf{c} \tag{3}$$

where $\mathbf{x}$ is the vector with components (A, B) in the present problem, and $\mathbf{G}$ is the matrix containing the coefficients of A, B in (2).

A least-squares solution to these equations may be easily obtained, but in general it will not satisfy the constraint

$$|\mathbf{x}|^2 = A^2 + B^2 = F_P^2. \tag{4}$$

Instead, we wish to find the point on the phase circle which best satisfies a least-squares error function determined by the linear equations (1).

For every point in the A, B plane, the error function

$$\varepsilon^2(A, B) = \sum_j [|f_j| d_j(A, B)]^2 \tag{5}$$

assigns a value determined by the sum of the squared weighted distances from the lines [(1)] to the point in the plane. The distance to each line, d_j, is weighted by the magnitude of the heavy-atom contribution, $|f_j|$, for that derivative. When reduced to a function of phase angle alone, (5) is equivalent to the error function described by the lack of closure based upon intensities. This equivalence can be demonstrated by showing that identical expressions for the Hendrickson & Lattman (1970) coefficients may be obtained for the two error functions (Appendix I). Equation (5) is therefore an appropriate error function.

Determination of the most probable phase

The most probable phase is identified by the point on the phase circle for which the error function (5) is a global minimum. A geometrical interpretation of the problem may be developed as follows. The contours of a linear least-squares function in the plane form a family of ellipses. For points on the phase circle at which the error function is a local extremum (along the circle), the contour of the error function must be tangent to the circle. A maximum of four such points may be present, as shown in Fig. 3.

An analytical determination of these points is possible using Lagrange multipliers. Therefore, we require the gradient of the error function to be proportional to the gradient of the constraint condition:

$$\mathbf{G}^T\mathbf{Gx} - \mathbf{G}^T\mathbf{c} = \lambda\mathbf{x} \tag{6}$$

where 2λ is the Lagrange multiplier, and the gradient of the error function (5) is $2\mathbf{G}^T(\mathbf{Gx} - \mathbf{c})$. Forcing the solution to satisfy the constraint (4) is equivalent to assigning all errors to the derivative data. Values of $\mathbf{x}$ satisfying (4) and (6) will represent extrema to (5) along the constraining circle.

To solve for $\mathbf{x}$, the following coordinate transformation step is employed. Let $\mathbf{E}$ be the matrix composed of the column eigenvectors of $\mathbf{G}^T\mathbf{G}$. Since the eigenvectors are orthonormal, multiplication by $\mathbf{E}$ or $\mathbf{E}^T$ performs a simple coordinate rotation. If $\mathbf{x}' = \mathbf{E}^T\mathbf{x}$, then $|\mathbf{x}'| = |\mathbf{x}|$. Introduction of the identity factor $\mathbf{EE}^T$ and premultiplication by $\mathbf{E}^T$ in (6) gives:

$$\mathbf{E}^T(\mathbf{G}^T\mathbf{G})(\mathbf{EE}^T)\mathbf{x} - \mathbf{E}^T\mathbf{G}^T\mathbf{c} = \lambda\mathbf{E}^T\mathbf{x} \tag{7a}$$

$$\mathbf{E}^T(\mathbf{G}^T\mathbf{G})\mathbf{Ex}' - \mathbf{B} = \lambda\mathbf{x}' \tag{7b}$$

$$\mathbf{Dx}' - \mathbf{B} = \lambda\mathbf{x}' \tag{7c}$$

where $\mathbf{B} = \mathbf{E}^T\mathbf{G}^T\mathbf{c}$. Since $\mathbf{E}^T$ and $\mathbf{E}$ diagonalize $\mathbf{G}^T\mathbf{G}$, $\mathbf{D} = \mathbf{E}^T\mathbf{G}^T\mathbf{GE}$ is the diagonal matrix composed of the eigenvalues of $\mathbf{G}^T\mathbf{G}$. Equation (7c) may be solved for x_i':

$$x_i' = B_i/(D_{ii} - \lambda). \tag{8}$$

Recall $|x'|^2 = |x|^2 = F_P^2$. Imposing this condition upon (8) leads to the relationship:

$$\sum_{i=1}^{N}\left[B_i^2 \prod_{j\neq i}^{N}(D_{jj} - \lambda)^2\right] = F_P^2 \prod_{k=1}^{N}(D_{kk} - \lambda)^2 \tag{9}$$

where N is the dimension of the vector $\mathbf{x}$. Equation (9) describes a polynomial in λ of degree $2N$, which

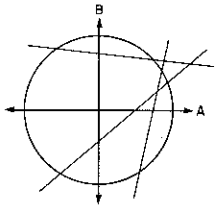

Fig. 2. The same relationships as Fig. 1, but for imperfect data. In this case, the phasing lines will not intersect at a single point on the phase circle.

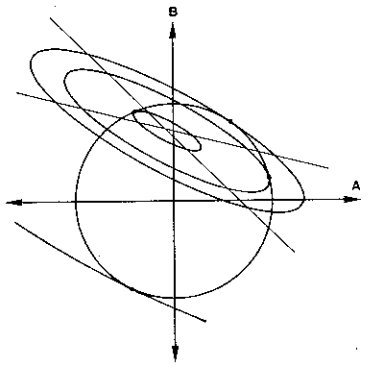

Fig. 3. Error ellipses corresponding to the error function (4). For points on the phase circle at which the error function is a local extremum, the elliptic contours of the error function will be tangent to the circle.

may be solved by standard numerical methods. Substitution of each value into (8) yields x_i', and the minima may be identified from (5).

If the $\mathbf{G}^T\mathbf{G}$ matrix is degenerate (as in the single isomorphous replacement case), then some of the eigenvalues D_{ii} and the corresponding B_i will vanish, and the value $\lambda = 0$ will be a degenerate solution. Under these conditions, not all of the x_i are independent, and multiple minima solutions occur. In the *SIR* case, two equally probable solutions exist, corresponding to the non-tangential intersection of a line with a circle. In this situation, either phase solution is possible, and an average of the two minima may be selected.

In addition to isomorphous replacement data, linear equations of the form (1) may be obtained from anomalous-dispersion measurements and molecular-replacement data.

Application to twinning

In the case of twinning by merohedry, the twinning operation exactly superimposes non-equivalent reflections from the twin domains. In the simplest case of twinning by hemihedry, two crystals are present in the twin. The following analysis will describe a phasing procedure which is valid for intensity data measured from hemihedral twins, although it may be generalized to other types of twinning in a straightforward manner. For the case of twinning by hemihedry, the observed native intensity is the weighted sum of intensities from the two twin-related reflections I_1 and I_2:

$$I_N = \alpha I_1 + (1-\alpha) I_2 \tag{10}$$

where α is the volume fraction of twin 1. The observed intensity for the jth derivative, I_{H_j}, is given by

$$I_{H_j} = \beta(I_1 + f_{j1}^2 + 2a_{j1}A_1 + 2b_{j1}B_1) + (1-\beta)(I_2 + f_{j2}^2 + 2a_{j2}A_2 + 2b_{j2}B_2) \tag{11}$$

where β is the volume fraction of twin 1 of derivative j. α and β may be estimated through a variety of methods (Britton, 1972; Murray-Rust, 1973; Fisher & Sweet, 1980; Rees, 1982). If α or $\beta = \frac{1}{2}$ (perfect twinning), then a single equation, linear in terms of the native structure-factor components (A_1, B_1, A_2, B_2), is obtained:

$$I_{H_j} = I_N + f_{j1}^2/2 + f_{j2}^2/2 + a_{j1}A_1 + b_{j1}B_1 + a_{j2}A_2 + b_{j2}B_2. \tag{12}$$

The magnitude constraint for the case of perfect hemihedral twinning is

$$(A_1^2 + B_1^2 + A_2^2 + B_2^2)/2 = I_N. \tag{13}$$

When neither α nor β is equal to $\frac{1}{2}$, two independent equations, similar to (12), are obtained from each derivative (Appendix II).

The phasing problem in the presence of twinning by hemihedry is four-dimensional. The corresponding geometric interpretation is to describe the intersection of hyperplanes with a hypersphere in four-space. This generates an eighth-order equation in λ analogous to (9). Identification of the appropriate minimum (or minima) then generates the A's and B's corresponding to the most probable phase. Four or more independent phasing equations are required to specify a unique point on the hypersphere. If the native and derivative crystals are perfectly twinned, four independent isomorphous derivatives are needed. For the partial twinning case, two derivatives are sufficient to determine uniquely the most probable phase. However, calculations on test data (described below) show that phases of moderate quality can be obtained from fewer than four phasing equations. (This is analogous to *SIR* phasing in the absence of twinning.)

Determination of the best phase for perfectly twinned structures

Since the pioneering studies of Blow & Crick (1959), Fourier syntheses have generally been calculated using 'best' coefficients. These coefficients are defined by the centroid of the structure-factor probability distribution.

For the case of perfect twinning ($\alpha = \frac{1}{2}$), the points (A_1, B_1, A_2, B_2) which satisfy the constraint of (13) describe a hypersphere. Any point on this hypersphere may be described by three angles (γ, δ, τ) using the following definitions.

$$\gamma = \text{phase of twin } 1 = \tan^{-1}(B_1/A_1) \qquad 0 \le \gamma < 2\pi \tag{14a}$$

$$\delta = \text{phase of twin } 2 = \tan^{-1}(B_2/A_2) \qquad 0 \le \delta < 2\pi \tag{14b}$$

$$\tau = \tan^{-1}(|F_1|/|F_2|) \qquad 0 \le \tau \le \pi/2. \tag{14c}$$

Conversely,

$$A_1 = (2I_N)^{1/2} \sin\tau \cos\gamma \tag{14d}$$

$$B_1 = (2I_N)^{1/2} \sin\tau \sin\gamma \tag{14e}$$

$$A_2 = (2I_N)^{1/2} \cos\tau \cos\delta \tag{14f}$$

$$B_2 = (2I_N)^{1/2} \cos\tau \sin\delta. \tag{14g}$$

We now define a lack of closure in four-space, based upon intensities, as:

$$\varepsilon_j(\gamma, \delta, \tau) = I_{H_j} - (A_1 + a_{j1})^2/2 - (B_1 + b_{j1})^2/2 - (A_2 + a_{j2})^2/2 - (B_2 + b_{j2})^2/2 \tag{15}$$

where A_1, B_1, A_2 and B_2 can all be expressed as functions of γ, δ and τ. The lack-of-closure error function $\varepsilon_j^2(\gamma, \delta, \tau)$ can be shown to take the form of a truncated Fourier series (Appendix III). The

Table 1. *Accuracy of the most probable phase in the degenerate case*

% error represents the random error introduced into the derivative intensity data. $\Delta\alpha$ is the average phase error in degrees.

Number of derivatives	1% error $\Delta\alpha$	4% error $\Delta\alpha$
1	61·9	63·0
2	44·5	47·9
3	27·4	35·6

arguments of the sine and cosine terms are linear combinations of γ, δ and τ. For each pair of twin-related reflections, 18 real Fourier coefficients are required to describe completely the squared closure error over the hypersphere. When data from four or more derivatives are available, storing the phase probability distributions in this form requires less space than storing derivative intensities and heavy-atom scattering factors separately. When computer storage space is not a concern, it is unnecessary to expand $\varepsilon_j^2(\gamma, \delta, \tau)$ as a Fourier series, since $\varepsilon_j(\gamma, \delta, \tau)$ may be computed directly from (14) and (15).

The probability distribution, based upon a single derivative, may be given by

$$P_j(\gamma, \delta, \tau) = N_j \exp\left[-\varepsilon_j^2(\gamma, \delta, \tau)/2E_j\right] \quad (16)$$

where N_j is a normalization constant and E_j is the e.s.d. in errors for derivative j. The total probability, as a function of γ, δ and τ, resulting from multiple derivatives, is given by the product of the P_j's.

$$P(\gamma, \delta, \tau) = N \exp\left\{-\sum_j \left[\varepsilon_j^2(\gamma, \delta, \tau)/2E_j\right]\right\}. \quad (17)$$

Therefore, the corresponding coefficients from all available derivatives may be summed to give a complete description of the total probability distribution. Following Hendrickson & Lattman (1970), calculation of the best structure factor requires evaluation of the following five integrals:

$$\begin{pmatrix} Q_1 \\ Q_2 \\ Q_3 \\ Q_4 \\ Q_5 \end{pmatrix} = \int\int\int \begin{pmatrix} \sin\tau\cos\gamma \\ \sin\tau\sin\gamma \\ \cos\tau\cos\delta \\ \cos\tau\sin\delta \\ 1 \end{pmatrix} \times P(\gamma, \delta, \tau) \sin\tau\cos\tau\, d\gamma\, d\delta\, d\tau, \quad (18)$$

where $\sin\tau\cos\tau\, d\gamma\, d\delta\, d\tau$ is the integration element for a hyperspherical surface in our orthogonal curvilinear coordinate system. These integrals may be approximated numerically. The final structure factor would be given by

$$\begin{pmatrix} A_1 \\ B_1 \\ A_2 \\ B_2 \end{pmatrix} = \left[(2I_N)^{1/2}/Q_5\right] \begin{pmatrix} Q_1 \\ Q_2 \\ Q_3 \\ Q_4 \end{pmatrix}. \quad (19)$$

[The problem of determining the best phase for the case of partial twinning may be treated by generalizing (15) to be the sum of the squared errors in equations (*B*3*a*) and (*B*3*b*) in Appendix II.]

Test calculations for the case of perfect hemihedral twinning

Atomic coordinates for the complex between carboxypeptidase-A and a protease inhibitor from potatoes (CPA-PCI) (Rees & Lipscomb, 1982) were used to calculate structure factors between 20·0 and 4·5 Å. In addition, five different derivative data sets were generated by introducing a single mercury atom at arbitrary positions in the asymmetric unit [(0·23, 0·33, 0·10), (0·42, 0·61, 0·26), (0·12, 0·30, 0·03), (0·68, 0·66, 0·16), (0·26, 0·21, 0·28)]. The average R factor on $|F|$ between native and derivative data sets was 8·0%. The space group for the CPA-PCI complex is $P3_2$. The crystals grow as twins, with a $[11\bar{2}0]$ twinning axis. To simulate data from a perfectly twinned crystal, intensities of twin-related reflections were averaged so that $I_{hkl} = I_{kh\bar{l}}$. Approximately Gaussian errors were introduced into the derivative intensities. Fortran programs were written to calculate the most probable phase and the best phase from perfectly twinned data. The programs used subroutines from the IBM Scientific Subroutine Package and were implemented on a VAX 11/780.

For fewer than four derivatives, the most probable phase solution is degenerate. Calculations were performed with one, two and three derivatives. For each reflection, the degenerate solutions were identified and averaged (as in the conventional *SIR* case). The average absolute phase error ($\Delta\alpha$) was calculated between the known and most probable phase sets (Table 1). With three derivatives, the phase error averages approximately 30° (depending on the size of the error introduced into the derivative intensities). Even with one derivative, the calculated phases are still systematically biased towards the correct phases.

For more than three independent derivatives, the most probable phase solution is unique. The average phase error ($\Delta\alpha$) and R factor on $|F|$ (between the known structure-factor amplitudes and the calculated amplitudes) were calculated for four and five derivatives (Table 2). The best phase was also determined for four and five derivatives. The average phase error and R factor for the best phase method are given in Table 3. The integrals of (18) were evaluated by sampling the hyperspherical surface at 10° intervals for the three angles. The lack-of-closure error was computed directly from (14) and (15). Calculations of the most probable phase were based upon 1136 pairs of twin-related reflections. Calculations of the best phase were based upon 100 pairs of twin-related reflections. For both methods, the accuracies of the phase and amplitude estimates improve as the number

Table 2. *Accuracy of the most probable phase in the non-degenerate case*

% error represents the random error introduced into the derivative intensity data. $\Delta\alpha$ is the average phase error in degrees. $R(\%)$ is the R factor on $|F|$ between the known and calculated structure-factor amplitudes.

Number of derivatives	1% error		4% error	
	$\Delta\alpha$	R (%)	$\Delta\alpha$	R (%)
4	9·2	8·5	24·9	19·9
5	5·1	5·6	18·4	16·9

Table 3. *Accuracy of the best phase in the non-degenerate case*

% error represents the random error introduced into the derivative intensity data. $\Delta\alpha$ is the average phase error in degrees. $R(\%)$ is the R factor on $|F|$ between the known and calculated structure-factor amplitudes.

Number of derivatives	1% error		4% error	
	$\Delta\alpha$	R (%)	$\Delta\alpha$	R (%)
4	10·8	9·2	24·7	19·4
5	4·8	5·6	17·2	16·4

of derivatives increases and the error in the derivative data decreases. With data from four derivatives (perfectly twinned) and 4% intensity errors, phase errors of less than 25° and an R factor below 0·2 may be achieved. Although the best-phase algorithm provides slightly better results, this advantage is offset by the increased computer time (approximately a factor of 50) required to calculate the best phase, relative to the most probable phase.

Concluding remarks

The least-squares formulation of the method of isomorphous replacement provides geometrical insight into the phase problem, and can be directly extended to the higher-dimensional problem of phasing twinned structures. The most probable phase may be determined rapidly by the method of Lagrange multipliers. Calculation of the best phase for twinned structures is accomplished by expressing a four-dimensional lack-of-closure function in terms of three hyperspherical angles. With these methods, the technique of isomorphous replacement may be used to determine phases for twinned structures, even in the case of perfect twinning.

This work was supported in part by USPHS grant GM31875 and USPHS NRS award GM07185 to TOY.

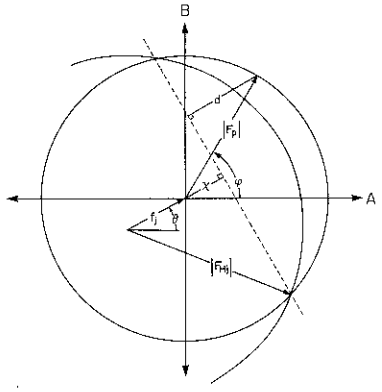

Fig. 4. Phase circle diagram illustrating the relationships between structure-factor amplitudes and phase angles described in Appendix I.

APPENDIX I

Equivalence of a linear least-squares error function to the lack-of-closure error

Let θ = phase of the heavy-atom contribution (f_j), φ = possible choice for phase of the native structure factor (F_P), χ = shortest distance from the origin to the line defined by the points of intersection of the two circles (Fig. 4).

From Fig. 4,

$$d^2 = [F_P \cos(\varphi - \theta) - \chi]^2$$
$$= F_P^2 \cos(2\varphi - 2\theta)/2 + F_P^2/2$$
$$- 2F_P \cos(\varphi - \theta)\chi + \chi^2. \qquad (A1)$$

Since

$$F_P^2 - \chi^2 = F_H^2 - (|f_j| + \chi)^2, \qquad (A2a)$$
$$\chi = (F_H^2 - F_P^2 - f_j^2)/(2|f_j|). \qquad (A2b)$$

Equations ($A1$) and ($A2b$) give

$$d^2 = F_P^2 \{\cos(2\varphi)[\cos^2\theta - \sin^2\theta]$$
$$+ \sin(2\varphi) 2\cos\theta \sin\theta\}/2$$
$$+ F_P(F_P^2 + f_j^2 - F_H^2)$$
$$\times(\cos\varphi\cos\theta + \sin\varphi\sin\theta)/|f_j|$$
$$+ (F_H^2 - F_P^2 - f_j^2)^2/(4f_j^2) + F_P^2/2. \qquad (A3)$$

If we let $\varepsilon_j^2(\varphi) = [|f_j| d(\varphi)]^2$ [(5)], then

$$\varepsilon_j^2(\varphi) = F_P^2(a_j^2 - b_j^2)\cos(2\varphi)/2 + F_P^2 a_j b_j \sin(2\varphi)$$
$$+ F_P(F_P^2 + f_j^2 - F_H^2) a_j \cos\varphi$$
$$+ F_P(F_P^2 + f_j^2 - F_H^2) b_j \sin\varphi$$
$$+ [2F_P^2 f_j^2 + (F_P^2 + f_j^2 - F_H^2)^2]/4. \qquad (A4)$$

These coefficients are proportional, by a factor of $\frac{1}{4}$, to those obtained by Hendrickson & Lattman (1970) for the lack-of-closure error, based upon intensities. In short, the lack-of-closure error,

$$\sum_j [F_H^2 - |(F_P + f_j)|^2]^2,$$

is equal to

$$4\sum_j [|f_j| d_j(\varphi)]^2.$$

Since the constant factor does not alter the positions of the extrema or the probability centroid it will not affect either the most probable phase or the best phase.

APPENDIX II

Linear isomorphous replacement equations for the case of hemihedral twinning with arbitrary twinning fractions

In the case of hemihedral twinning, the two crystal intensities (I_1 and I_2) related by the twinning operation contribute to two different observed intensities, designated I_{N-} and I_{N+}. (These two observed intensities are equal in the case of perfect twinning.)

$$I_{N+} = \alpha I_1 + (1-\alpha) I_2 \qquad (B1a)$$

$$I_{N-} = (1-\alpha) I_1 + \alpha I_2 \qquad (B1b)$$

where α is the twinning fraction of the native specimen. For a derivative with twinning fraction β,

$$I_{H+} = \beta(I_1 + f_1^2 + 2a_1A_1 + 2b_1B_1) + (1-\beta)(I_2 + f_2^2 + 2a_2A_2 + 2b_2B_2) \qquad (B2a)$$

$$I_{H-} = (1-\beta)(I_1 + f_1^2 + 2a_1A_1 + 2b_1B_1) + \beta(I_2 + f_2^2 + 2a_2A_2 + 2b_2B_2). \qquad (B2b)$$

From these equations, we can obtain two independent linear equations in A_1, B_1, A_2 and B_2:

$$I_{H+} = I_{N+}(\alpha+\beta-1)/(2\alpha-1) + I_{N-}(\alpha-\beta)/(2\alpha-1) + \beta f_1^2 + (1-\beta) f_2^2 + 2\beta a_1A_1 + 2\beta b_1B_1 + 2(1-\beta)a_2A_2 + 2(1-\beta)b_2B_2 \qquad (B3a)$$

and

$$I_{H-} = I_{N+}(\alpha-\beta)/(2\alpha-1) + I_{N-}(\alpha+\beta-1)/(2\alpha-1) + (1-\beta)f_1^2 + \beta f_2^2 + 2(1-\beta)a_1A_1 + 2(1-\beta)b_1B_1 + 2\beta a_2A_2 + 2\beta b_2B_2. \qquad (B3b)$$

These equations may be used to determine the most probable phase (as described in the text). When $\beta = \frac{1}{2}$, $(B3a)$ and $(B3b)$ are identical and reduce to (12). When $\alpha = \frac{1}{2}$, these equations cannot be evaluated. Equation (12) still holds, however, with $I_H = (I_{H+} + I_{H-})/2$. Therefore, two phasing equations are obtained per derivative when neither α nor β are equal to $\frac{1}{2}$. Otherwise, only one phasing equation is obtained per derivative.

When α is near $\frac{1}{2}$, measurement errors are amplified by solving for I_1 and I_2 directly from $(B1a)$ and $(B1b)$ above (Grainger, 1969). This problem is minimized in the present treatment, however, since the most probable structure factor is not required to satisfy $(B1a)$ and $(B1b)$ exactly. For the partial twinning case, the most probable structure factor must only satisfy the total magnitude constraint

$$A_1^2 + B_1^2 + A_2^2 + B_2^2 = I_1 + I_2 = I_{N+} + I_{N-} \qquad (B4)$$

as well as (6) in the text.

APPENDIX III

Fourier coefficients for the lack-of-closure error in the case of perfect twinning by hemihedry

From $(14d)$-$(14g)$ and (15) in the text, it can be shown that the lack-of-closure error function, $\varepsilon_j^2(\gamma, \delta, \tau)$, takes the form of a finite Fourier series in the hyperspherical angle variables γ, δ and τ:

$$\begin{aligned}
2\varepsilon_j^2(\gamma,\delta,\tau) &= I_N(a_2^2 + b_2^2 - a_1^2 - b_1^2)\cos(2\tau) \\
&+ I_N(a_1^2 - b_1^2)\cos(2\gamma) + 2I_N a_1 b_1 \sin(2\gamma) \\
&+ I_N(a_2^2 - b_2^2)\cos(2\delta) + 2I_N a_2 b_2 \sin(2\delta) \\
&- Xb_1(2I_N)^{1/2}\cos(\tau+\gamma) + Xa_1(2I_N)^{1/2}\sin(\tau+\gamma) \\
&+ Xb_1(2I_N)^{1/2}\cos(\tau-\gamma) + Xa_1(2I_N)^{1/2}\sin(\tau-\gamma) \\
&+ Xa_2(2I_N)^{1/2}\cos(\tau+\delta) + Xb_2(2I_N)^{1/2}\sin(\tau+\delta) \\
&+ Xa_2(2I_N)^{1/2}\cos(\tau-\delta) - Xb_2(2I_N)^{1/2}\sin(\tau-\delta) \\
&- [I_N(a_1^2 - b_1^2)/2]\cos(2\tau+2\gamma) \\
&- I_N a_1 b_1 \sin(2\tau+2\gamma) \\
&- [I_N(a_1^2 - b_1^2)/2]\cos(2\tau-2\gamma) \\
&+ I_N a_1 b_1 \sin(2\tau-2\gamma) \\
&+ [I_N(a_2^2 - b_2^2)/2]\cos(2\tau+2\delta) \\
&+ I_N a_2 b_2 \sin(2\tau+2\delta) \\
&+ [I_N(a_2^2 - b_2^2)/2]\cos(2\tau-2\delta) \\
&- I_N a_2 b_2 \sin(2\tau-2\delta) \\
&- I_N(b_1a_2 + a_1b_2)\cos(2\tau+\gamma+\delta) \\
&+ I_N(a_1a_2 - b_1b_2)\sin(2\tau+\gamma+\delta) \\
&- I_N(b_1a_2 - a_1b_2)\cos(2\tau+\gamma-\delta) \\
&+ I_N(a_1a_2 + b_1b_2)\sin(2\tau+\gamma-\delta) \\
&+ I_N(b_1a_2 - a_1b_2)\cos(2\tau-\gamma+\delta) \\
&+ I_N(a_1a_2 + b_1b_2)\sin(2\tau-\gamma+\delta) \\
&+ I_N(b_1a_2 + a_1b_2)\cos(2\tau-\gamma-\delta) \\
&+ I_N(a_1a_2 - b_1b_2)\sin(2\tau-\gamma-\delta) \\
&+ I_N(a_1^2 + b_1^2 + a_2^2 + b_2^2) + X^2/2 \qquad (C1)
\end{aligned}$$

where X is defined as

$$2I_N + a_1^2 + b_1^2 + a_2^2 + b_2^2 - 2I_H.$$

In the expression above, only 18 of the 30 terms have unique coefficients.

References

BLOW, D. M. & CRICK, F. H. C. (1959). *Acta Cryst.* **12**, 794–802.
BRITTON, D. (1972). *Acta Cryst.* A**28**, 296–297.
FISHER, R. G. & SWEET, R. M. (1980). *Acta Cryst.* A**36**, 755–760.
GRAINGER, C. T. (1969). *Acta Cryst.* A**25**, 427–434.
HENDRICKSON, W. A. & LATTMAN, E. E. (1970). *Acta Cryst.* B**26**, 136–143.
MURRAY-RUST, P. (1973). *Acta Cryst.* B**29**, 2559–2566.
RAMACHANDRAN, G. N. & SRINIVASAN, R. (1970). *Fourier Methods in Crystallography.* New York: Wiley.
REES, D. C. (1982). *Acta Cryst.* A**38**, 201–207.
REES, D. C. & LIPSCOMB, W. N. (1982). *J. Mol. Biol.* **160**, 475–498.
STUBBS, G. J. & DIAMOND, R. (1975). *Acta Cryst.* A**31**, 709–718.

Acta Cryst. (1987). A**43**, 36–42

(3+1)-Dimensional Patterson and Fourier Methods for the Determination of One-Dimensionally Modulated Structures

BY W. STEURER

Institut für Kristallographie und Mineralogie der Universität, Theresienstrasse 41, *D*-8000 *München* 2, *Federal Republic of Germany*

(*Received* 15 *July* 1985; *accepted* 12 *May* 1986)

Abstract

Patterson and Fourier methods applied to one-dimensionally modulated structures in the (3+1)-dimensional space R_{3+1} can be very helpful tools for the calculation of starting parameters of the atomic modulation functions. The characteristics of the (3+1)-dimensional Patterson function [(3+1)-PF] are discussed for some typical modulation waves from a geometrical point of view as well as with the aid of known modulated phases. The influence of series termination errors, resulting from incomplete data sets, is demonstrated. The (3+1)-PF, in any case, yields sufficient basic information even if first-order satellites only are accessible. Of course, it is necessary to include higher orders if one wants to learn something about the shape of the modulation wave. A comparison is made with the Patterson methods used for the solution of modulated structures until now, and it is shown that they are special cases of the (3+1)-PF. Some applications are given for Fourier methods in R_{3+1}, for example, to detect fluctuations of the phase or the amplitude of the modulation wave.

1. Introduction

In recent years an increasing number of commensurately and incommensurately modulated structures has been determined. In general the solution of the average structure presents no difficulties, but it can be problematical to find workable starting parameters for the refinement of the atomic modulation functions. Direct methods are available for superstructures (*cf. e.g.* Böhme, 1982) but there is no incommensurately modulated structure known to the author which has been solved using them, and the usual Patterson methods are hardly interpretable for more complicated cases. Consequently, most of the modulated structures have been solved based on model considerations or in a rather straightforward way (*cf. e.g.* Horst, Tagai, Korekawa & Jagodzinski, 1981; Yamamoto, Nakazawa, Kitamura & Morimoto, 1984; Steurer & Adlhart, 1983*a*, *b*). However, there have been many attempts to apply Patterson techniques to get information independent of models. The most frequently used way is to describe the incommensurate modulation in a commensurate supercell, approximately, and to calculate the 'partial' or 'difference' Patterson synthesis using the superstructure reflections alone (*cf. e.g.* Frueh, 1953; Takéuchi, 1972; Böhm, 1978; Tomeoka & Ohmasa, 1982). As a result the Patterson map of the 'complementary' structure, the difference between the real modulated structure and the average structure, is obtained.

Another method has been derived by Toman & Frueh (1973*a*, *b*) by calculating the Patterson synthesis in the subcell using 'one set' of satellite reflections. The 'plus and minus' difference Patterson function (McConnell & Heine, 1984) is a similar approach and has been used primarily to obtain symmetry information. A detailed discussion of all these methods will be given in § 5 of this paper. The purpose of this study is to discuss the properties of the (3+1)-PF for some fundamental modulation functions and to give an aid to the application of this method in practical structure determination. For the sake of a clear representation plane modulation waves with equal amplitudes are investigated in the first place, but the (3+1)-PF is interpretable in more general cases as well. Naturally, in the case of asymmetric functions of very different shape the definition of a phase relation between two modulation waves will

0108-7673/87/010036-07$01.50