Acta Cryst. (1982). A**38**, 201–207

A General Theory of X-ray Intensity Statistics for Twins by Merohedry

BY D. C. REES

Gibbs Chemical Laboratory, Harvard University, Cambridge, MA 02138, *USA*

(*Received* 24 *March* 1980; *accepted* 17 *September* 1981)

Abstract

A general theory of X-ray intensity statistics for twins by merohedry is developed based on mathematical analogy of this problem to a one-dimensional random walk. Probability distribution functions are derived (*a*) for intensities from a single twinned specimen, and (*b*) for the differences in intensities from two separate twinned specimens. Theoretical values of a discrepancy index between observed intensities from a twin, and calculated intensities from both correct and incorrect structural models are evaluated. A new method for determining volume fractions of crystals in a twin is also proposed, based on the distribution of differences in intensities between twin-related reflections. These results have implications for the treatment of isomorphous replacement data collected from twins by hemihedry.

Twinning by merohedry is often difficult to detect in practice since the reciprocal lattices of crystals in a twin superimpose exactly. As a result, there is no obvious indication in the diffraction pattern of the composite nature of the reciprocal lattice. Unsuccessful attempts at a structure determination, or unsatisfactory refinement of a trial structure, may be the first sign that the 'crystals' used for data collection were actually twinned. If twinning is suspected, it is first necessary to determine the twin law (how many crystals are present in the twin, and by which twin-symmetry operation they are related), and whether one has twinning by true merohedry (twin-lattice symmetry) or by pseudo-merohedry (twin-lattice quasi-symmetry) (Donnay & Donnay, 1974). Next, one must determine the volume fractions of the crystals in the twin. This may be accomplished using methods based on comparison of intensities of twin-related reflections (Britton, 1972; Murray-Rust, 1973; Fisher & Sweet, 1980) or from the distribution of diffracted intensities (Rees, 1980). With this information, it is possible to correct the observed intensities for twinning. The accuracy of this correction is sensitive to both the magnitudes of the volume fractions, and to uncertainties in the measurements of the intensities and volume fractions (Grainger, 1969). If the degree of twinning is small, it may be possible to correct the data for twinning, and then proceed as in a normal structure determination. When the degree of twinning is sufficiently large, however, it is no longer advisable to correct the intensities for twinning. If less twinned specimens of a sample are not available, one is then faced with the prospect of solving a structure with twin data.

In order to take full advantage of intensity data from a twin, it is necessary to understand how twinning influences the statistical properties and distribution of intensities. This is particularly relevant when direct methods are employed, since the selection of reflections for the phase relationships depends upon the magnitude of the structure factors, which may be greatly perturbed by twinning. The probability distribution functions for intensities from twins by hemihedry have been derived as a function of the volume fraction (Rees, 1980), assuming Wilson's (1949) intensity distribution functions for intensities from untwinned crystals. In this communication, a general theory of X-ray intensity statistics for twins by merohedry is presented. Of practical significance, a new method for estimating volume fractions of crystals in a twin is described, and theoretical values for a discrepancy index between observed intensities from a twin, and calculated intensities from correct and incorrect models for the structure are evaluated. Mathematical details of the theory are developed in the Appendix. In the text, application of these methods to the estimation of volume fractions and the treatment of isomorphous replacement data is discussed in some detail for the case of twinning by hemihedry.

The intensity, p, of a reflection measured from a twin by hemihedry is a linear combination of the crystal intensities, j_1 and j_2, of two reflections related by the twin law:

$$p = (1 - \alpha) j_1 + \alpha j_2, \qquad (1a)$$

where α is the volume fraction (twin fraction) of the smaller crystal in the twin. If the intensity of the same reflection is measured from a second specimen of twin fraction β, the observed value of the intensity, q, is given by

$$q = (1 - \beta) j_1 + \beta j_2. \qquad (1b)$$

0567-7394/82/020201-07$01.00

Since twinning by merohedry superimposes symmetry-equivalent reflections for certain zones, we will restrict the following discussion to the general case that j_1 and j_2 are independent.

A discrepancy index between intensities p and q may be defined:

$$R_{\alpha\beta} = \frac{\langle |p-q| \rangle}{\langle p \rangle}. \tag{2}$$

From equation (1) and the equality $\langle j_1 \rangle = \langle j_2 \rangle$, equation (2) reduces to:

$$\begin{aligned} R_{\alpha\beta} &= \frac{\langle |j_2 - j_1| \rangle}{\langle j_1 \rangle} |\beta - \alpha| \\ &= \langle R \rangle |\beta - \alpha|, \end{aligned} \tag{3}$$

where $\langle R \rangle$ is the average value of the discrepancy index $\langle |j_2 - j_1| \rangle / \langle j_1 \rangle$ for random crystal intensities. Assuming Wilson's (1949) statistics to be valid for the intensities from untwinned crystals, $\langle R \rangle = 1{\cdot}0$ for non-centrosymmetric reflections, and $4/\pi$ for centrosymmetric reflections (Srinivasan & Parthasarathy, 1976). This result is also derived using a more general technique in the Appendix. In a similar fashion, the mean-square difference between the normalized intensities p and q reduces to

$$\begin{aligned} r^2_{\alpha\beta} &= (p-q)^2 \\ &= (\alpha - \beta)^2 \langle r^2 \rangle, \end{aligned} \tag{4}$$

where $\langle r^2 \rangle$ equals 2 or 4 for non-centrosymmetric and centrosymmetric reflections, respectively (Srinivasan & Parthasarathy, 1976).

The correlation of intensities of twin-related reflections from the same specimen is formally equivalent to the correlation of intensities from different specimens of twinning fractions α and $\beta = 1 - \alpha$. In this case,

$$\begin{aligned} R_\alpha &= (1 - 2\alpha) \langle R \rangle \\ r^2_\alpha &= (1 - 2\alpha)^2 \langle r^2 \rangle. \end{aligned} \tag{5}$$

With perfectly measured data, the twinning fraction may be determined from R_α or r^2_α. As an example, Murray-Rust (1973) found for data collected from a twinned centrosymmetric specimen of $[Co(NH_3)_6]_4$-Cu_5Cl_{17} that $R_\alpha = 0{\cdot}36$. With $\langle R \rangle = 4/\pi$, (5) predicts $\alpha = 0{\cdot}36$, in reasonable agreement with the value $\alpha = 0{\cdot}32$ which Murray-Rust found by minimizing the difference between observed structure factors corrected for twinning and calculated structure factors.

The presence of random errors in the intensity measurements contributes an additional term to the experimental values of R_α and r^2_α, which will decrease the apparent twinning fraction of crystals calculated with (5). However, the effect of these errors may be compensated for as follows. If Δ^2 is the mean-square difference between intensities from twin-related reflections, and δ^2 is the average variance in the intensity measurements, then

$$\Delta^2 = 2\delta^2 + r^2_\alpha \tag{6}$$

so that the twinning fraction may be estimated by rearranging equations (5) and (6) to give

$$\alpha = \frac{1}{2}\left[1 - \left(\frac{\Delta^2 - 2\delta^2}{\langle r^2 \rangle}\right)^{1/2}\right], \tag{7}$$

where all quantities on the right-hand side can be determined experimentally.

If a structure determination by the method of isomorphous replacement is attempted using crystals twinned by merohedry, the intensity changes between the native and derivative data sets will include a contribution from the difference in twinning fractions of the two data sets. In an extreme case, an unsubstituted native crystal may appear to be a potential derivative

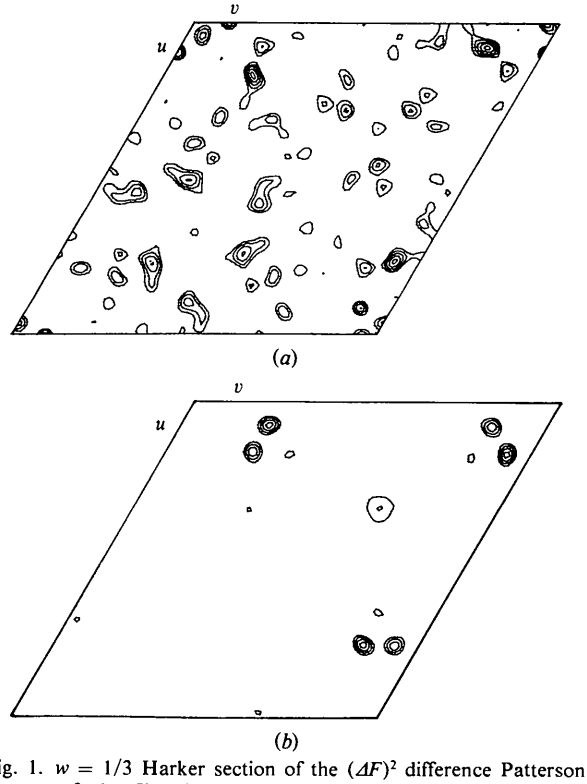

Fig. 1. $w = 1/3$ Harker section of the $(\Delta F)^2$ difference Patterson map of the $K_2Pt(NO_2)_4$ derivative of the protein complex described in the text. The space group of this complex is $P3_2$. (*a*) Difference coefficients calculated from uncorrected data sets (with twinning fractions of 0·10 and 0·42 for the native and $K_2Pt(NO_2)_4$ derivative data sets, respectively). (*b*) Difference coefficients calculated using perfectly twinned native data and the uncorrected $K_2Pt(NO_2)_4$ data set. Contours are at equal and arbitrary levels, with the first level above zero omitted. Further analysis of the derivative failed to identify any heavy-atom sites consistent with the extra peaks in Fig. 1(*a*) (Rees & Lipscomb, 1980).

based entirely on intensity changes due to twinning. It is possible, however, to determine the root-mean-square contributions of both the substituent and twinning to the intensity changes. Making the assumptions that the intensity changes due to the substituent are small, and that all sources of variance are statistically independent, then the mean-square intensity difference between native and derivative data sets, ε^2, will consist of several components:

$$\varepsilon^2 = \delta^2_{\rm nat} + \delta^2_{\rm der} + r^2_{\alpha\beta} + z^2_{\rm iso}, \qquad (8)$$

where $\delta^2_{\rm nat}$ and $\delta^2_{\rm der}$ are the variances of the intensity measurements for the native and derivative data sets, respectively, and $z^2_{\rm iso}$ is the mean-square intensity change due to the substituent. As every term except $z^2_{\rm iso}$ in (8) may be experimentally measured, the true contribution of the substituent to the intensity change can be simply determined by solving (8) for $z^2_{\rm iso}$.

A serious problem may arise in the calculation of difference Fourier and Patterson maps when there are differences in the twinning fractions of the two data sets used in the calculation. Under these circumstances, (8) indicates that it is possible to minimize errors in the difference coefficients by using data sets which have similar twinning fractions. For example, the native and $K_2Pt(NO_2)_4$ derivative data sets of the protein complex between carboxypeptidase-A and the potato carboxypeptidase inhibitor have twinning fractions of 0·10 and 0·42, respectively (Rees & Lipscomb, 1980). Since errors associated with twin correcting the $K_2Pt(NO_2)_4$ data set would be quite large (Grainger, 1969), it is not feasible to calculate a difference Patterson map with structure factors corrected for twinning. The difference Patterson map was considerably improved, however, by perfectly twinning the native data set used in the difference coefficient calculation (see Fig. 1). This procedure increased the ratio of the maximum peak height to root-mean-square value of the Harker section in the Patterson map from 3·6 to 5·4. Paradoxically, even though this method actually increases the difference between intensities and their true, untwinned values, it produces difference coefficients which are closer approximations to the correct terms.

The interest of Drs R. LeSar, M. Lewis, J. Hogle, and Professors W. N. Lipscomb and J. D. H. Donnay is gratefully acknowledged. This work was supported by NIH grant GM 06920.

APPENDIX
Probability distribution functions for X-ray intensities from a twin by merohedry

In this Appendix, we derive the probability distributions for X-ray intensities from a twin by merohedry. The first section treats the case of diffracted intensities from a single twinned specimen, while the second section examines differences in observed intensities from two separate twinned specimens. From these results, a new method of determining the volume fractions of crystals in a twin is proposed. Theoretical values for a discrepancy index between observed intensities from a twin, and calculated intensities from both correct and incorrect models are also evaluated.

It is assumed throughout the discussion that Wilson's (1949) statistics provide an adequate description of the probability distribution for intensities from untwinned specimens. No assumptions are made concerning the number of crystals, N, in a twinned specimen: twinning by merohedry properly includes twinning by hemihedry ($N = 2$), tetartohedry ($N = 4$), and ogdohedry ($N = 8$) (Catti & Ferraris, 1976). A range of values for N are permissible for twinning by pseudo-merohedry, depending on the crystal symmetry and fortuitous relationships between the cell constants. Implicit in the following discussion is the assumption that reflections from the individual twin domains are independent. This restriction eliminates the case in which reflections related by the twinning operation are symmetry related. Normalized intensities are used throughout the Appendix.

I. Probability distribution functions for X-ray intensities from a single twinned specimen

The general expression for an observed intensity, p, from a twinned specimen with N independent twin domains is

$$\begin{aligned} p &= \sum_{k=1}^{N} \alpha_k j_k \\ &\equiv \sum_{k=1}^{N} u_k, \end{aligned} \qquad (A1)$$

where α_k and j_k are the volume fraction and diffracted intensity, respectively, from the kth twin domain. The probability distribution of p may be derived by noting the similarity of this problem to a random-walk problem. In essence, we want to find the probability distribution for the length of an N-step one-dimensional random walk, given the probability distribution of the kth step. The general solution to a random-walk problem has been discussed by many authors (see, for example, Chandrasekhar, 1943), and may be broken into two parts:

(i) Calculation of λ_k, the characteristic function for the kth step, from the Fourier transform of the probability distribution function for the kth step, $P(u_k)$:

$$\lambda_k(\varphi) = \int_{-\infty}^{\infty} \exp(i\varphi u_k)\, P(u_k)\, {\rm d}u_k, \qquad (A2)$$

where $i^2 = -1$.

(ii) The probability distribution function of p, $P_N(p)$, is then given by the inverse Fourier transform of the repeated product of all N characteristic functions:

$$P_N(p) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-i\varphi p) \left\{ \prod_{k=1}^{N} \lambda_k(\varphi) \right\} d\varphi. \qquad (A3)$$

The relevant probability distribution and characteristic functions are given in Tables 1 and 2 for the non-centrosymmetric and centrosymmetric cases, respectively. The expressions for the $P(j_k)$ are given by Wilson's (1949) statistics, and the $P(u_k)$ are calculated from $P(j_k)$ by a change of variable from j_k to u_k (Srinivasan & Parthasarathy, 1976, p. 204). The probability distribution functions $P_N(p)$ were evaluated as follows:

(*a*) for non-centrosymmetric reflections, the integrand of (*A*3) has poles at $\varphi = -i/\alpha_k$, so that the integral may be evaluated using Cauchy's integral theorem and Jordan's lemma (Whittaker & Watson, 1927). Expressions for $P_N(p)$ were explicitly evaluated for the two limiting cases:
(i) all the α_k are different (Table 1, entry 4);
(ii) all the α_k are equal to $1/N$ (the case of perfect twinning; Table 1, entry 5).

Table 1. *Probability distribution and characteristic functions used in deriving the distribution of non-centrosymmetric X-ray intensities from a twin by merohedry*

Functions and parameters are defined in the text.

1.	$P(j_k)$	$\exp(-j_k)$
2.	$P(u_k)$	$[\exp - (u_k/\alpha_k)]/\alpha_k$
3.	$\lambda_k(\varphi)$	$[(i/\alpha_k) + \varphi]^{-1}(i/\alpha_k)$
4.	$P_N(p)$	$\sum_{k=1}^{N} \{[\exp - (p/\alpha_k)]/\alpha_k\} \left[\prod_{\substack{m=1 \\ m \neq k}}^{N} \frac{\alpha_k}{\alpha_k - \alpha_m} \right]$
5.	$P_N(p)$	$\frac{N^N p^{N-1}}{(N-1)!} \exp(-Np)$

Table 2. *Probability distribution and characteristic functions used in deriving the distribution of centrosymmetric X-ray intensities from a twin by merohedry*

$\Gamma(x)$ and ${}_1F_1(p;q;r)$ are the gamma function and the confluent hypergeometric function (Erdelyi, 1954) respectively. All other functions and parameters are defined in the text.

1.	$P(j_k)$	$[\exp - (j_k/2)]/(2\pi j_k)^{1/2}$
2.	$P(u_k)$	$[\exp - (u_k/2\alpha_k)]/(2\pi u_k \alpha_k)^{1/2}$
3.	$\lambda_k(\varphi)$	$[\varphi + (i/2\alpha_k)]^{-1/2}(i/2\alpha_k)^{1/2}$
4.	$P_N(p)$	$[(2\alpha)^{-m/2}(2\beta)^{-(N-m)/2}] \exp[-(p/2\alpha)]\, p^{(N-2)/2}[\Gamma(N/2)]^{-1} \times {}_1F_1\{(N-m)/2; N/2; [(\beta-\alpha)/2\alpha\beta]p\}$
5.	$P_N(p)$	$(N/2)^{N/2}[\Gamma(N/2)]^{-1} \exp - (Np/2)\, p^{(N-2)/2}$

(*b*) In the general case for centrosymmetric reflections, the integrand (*A*3) has branch points at $\varphi = -i/2\alpha_k$, which greatly complicate evaluation of the integral. Consequently, $P_N(p)$ was evaluated for only two cases, using expressions tabulated by Erdelyi (1954):
(i) m twin domains have volume fraction α, while the remaining $(N-m)$ twin domains have volume fraction $\beta = (1 - m\alpha)/(N-m)$ (Table 2, entry 4);
(ii) all the α_k are equal to $1/N$ (the case of perfect twinning; Table 2, entry 5).

The expressions for $P_N(p)$ in the case of perfect twinning ($\alpha_k = 1/N$) for both the centrosymmetric and non-centrosymmetric cases were first derived by Stanley (1972). Stanley's approach was only suitable for the special case of perfect twinning, however.

Although expressions for $P_N(p)$ have been explicitly evaluated for only a few specific examples, it should be stressed that for other situations $P_N(p)$ may be calculated by substitution of the appropriate values of $\lambda_k(\varphi)$ into (*A*3). While it may be necessary to evaluate numerically the Fourier transform in (*A*3) for certain centrosymmetric problems, it should be possible to obtain analytical solutions for all non-centrosymmetric cases.

II. Distribution of differences in intensities from two twinned specimens

The random-walk approach used to determine the probability distributions for intensities from a single twinned specimen may also be used to calculate the distribution of differences in intensities from two twinned specimens. This aspect of the theory will be illustrated through two specific examples: (i) the dependence of the distribution of intensity differences on the twinning fractions of two specimens, each one a twin by hemihedry; (ii) the evaluation of a discrepancy index for the comparison of observed intensities from perfect twins with calculated intensities from both correct and incorrect model structures.

II.i. *Twinning fraction determination*

Consider the observed intensities, p and q, from two specimens twinned by hemihedry, with twinning fractions α and β, respectively. From equations (1*a*) and (1*b*) in the main text, the difference in intensities, $p - q$, is

$$\begin{aligned} y &\equiv (p - q) \\ &= (\beta - \alpha)(j_1 - j_2). \end{aligned} \qquad (A4)$$

Comparing expressions (*A*1) and (*A*4), we see that α_1 and α_2 in (*A*1) correspond to $(\beta - \alpha)$ and $(\alpha - \beta)$, respectively, in (*A*4). Using the methods described in § I of the Appendix, the probability distribution

function $P_s(|y|)$ for the intensity differences [where $P_s(|y|) = P_2(y > 0) + P_2(y < 0)$] may be determined. For non-centrosymmetric reflections

$$P_s(|y|) = [\exp -(|y|/|\beta - \alpha|)]/|\beta - \alpha|, \quad (A5)$$

while for centrosymmetric reflections

$$P_s(|y|) = [K_0(|y|/2|\beta - \alpha|)]/(2\pi|\beta - \alpha|), \quad (A6)$$

where K_0 is a modified Bessel function of the second kind, of order zero.

If p and q are the intensities of twin-related reflections from the same specimen, then $\beta = 1 - \alpha$, and the terms $(\beta - \alpha)$ in (A5) and (A6) may be replaced by $(1 - 2\alpha)$. The sensitivity of $P_s(|y|)$ to α may be demonstrated by considering a function related to $P_s(|y|)$, the cumulative function $S(\Delta)$:

$$S(\Delta) = \int_0^{\Delta} P_s(|y|)\, dy. \quad (A7)$$

For non-centrosymmetric reflections, $S(\Delta)$ may be evaluated analytically:

$$S(\Delta) = 1 - \exp -[\Delta/(1 - 2\alpha)], \quad (A8)$$

while for centrosymmetric reflections the integral was evaluated using Gauss–Legendre quadratures (Stroud & Secrest, 1966). The dependence of $S(\Delta)$ on α is shown in Figs. 2 and 3 for non-centrosymmetric and centrosymmetric reflections, respectively. Even though these curves neglect errors in the intensity measurements, the sensitivity of $S(\Delta)$ to α indicates that the value of α may be estimated from $S(\Delta)$ *versus* Δ plots.

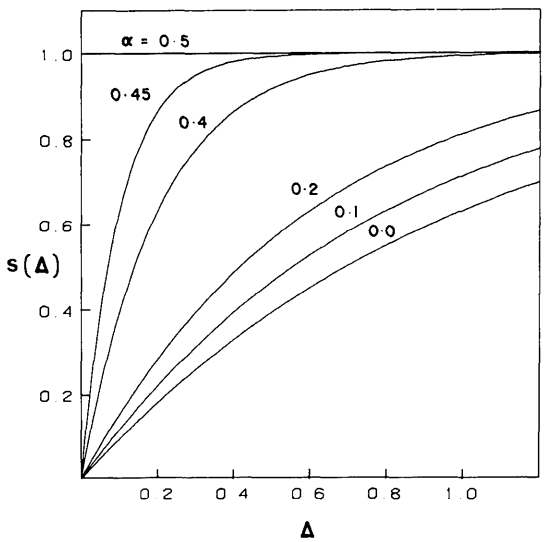

Fig. 2. The cumulative distribution function $S(\Delta)$ for non-centrosymmetric reflections from specimens twinned by hemihedry, for various values of the twinning fraction α.

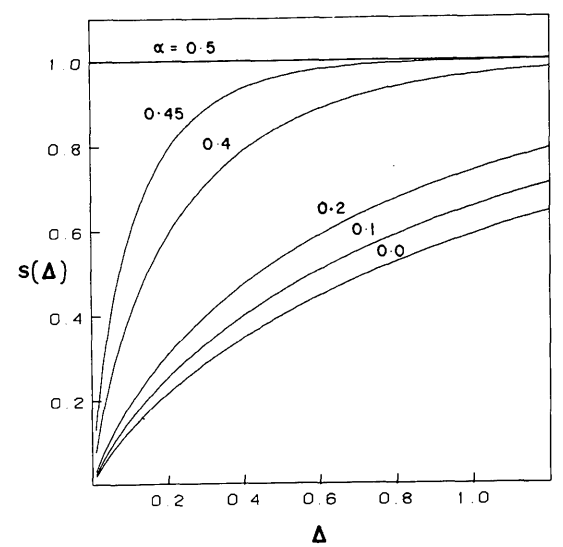

Fig. 3. The cumulative distribution function $S(\Delta)$ for centrosymmetric reflections from specimens twinned by hemihedry, for various values of the twinning fraction α.

II.ii. *Effect of twinning on discrepancy indices*

If a structure determination using intensity data from a twin is attempted, it is important to have theoretical values for a discrepancy index between observed and calculated intensities for both correct and incorrect structure models. This is especially true when the specimens used for data collection are highly twinned, so there is no prospect of correcting the observed intensities for twinning.

A discrepancy index, R, between the observed intensities, I_o, and intensities calculated from a structural model, I_c, may be defined

$$\begin{aligned} R &= \frac{\langle |I_o - I_c| \rangle}{\langle I_o \rangle} \\ &= \langle |I_o - I_c| \rangle \\ &\equiv \langle |\delta| \rangle, \end{aligned} \quad (A9)$$

where the second relationship is valid since the intensities are normalized. With equation (A1), I_o, I_c and δ may be expressed as

$$\begin{aligned} I_o &= \sum_{k=1}^{N} \alpha_k j_k \\ I_c &= \sum_{k=1}^{M} \beta_k t_k \\ \delta &= \sum_{k=1}^{N} \alpha_k j_k - \sum_{k=1}^{M} \beta_k t_k. \end{aligned} \quad (A10)$$

The probability distribution of $|\delta|$, $P_s(|\delta|)$, may be evaluated using analogous methods to those described

in previous sections. R is then calculated from $P_s(|\delta|)$ by the integral

$$R = \int_0^\infty \delta P_s(|\delta|)\, d\delta. \qquad (A11)$$

Theoretical values of R have been evaluated for three cases. In every case, I_o is assumed to be measured from a perfectly twinned specimen, while the I_c are defined as follows.

Case (i). The I_c are from a correct, untwinned model, so that $M = 1$, $\beta_1 = 1$, and $t_1 = j_1$.

Case (ii). The I_c are from an incorrect, untwinned model, so that $M = 1$, $\beta_1 = 1$, but the value of t_1 is independent of any j_k.

Case (iii). The I_c are from an incorrect, but perfectly twinned, model, so that $M = N$, $\beta_k = 1/N$, and the values for the t_k are independent of any j_k.

Integrals required for evaluating $P_s(|\delta|)$ and R were calculated using Cauchy's integral theorem and Jordan's lemma for the non-centrosymmetric case; and were taken from Erdelyi (1954) for the centrosymmetric case.

Tables 3 and 4 list expressions for R as a function of N for the non-centrosymmetric and centrosymmetric cases, respectively. Values of R as a function of N for the three cases are illustrated in Figs. 4 and 5 for the non-centrosymmetric and centrosymmetric cases, respectively. Asymptotic forms of the various R factors are seen to approach the expected values: since as N tends to infinity, the values of all I_o will approach 1, the asymptotic value for R in cases (i) and (ii) should approach $\langle |I_c - 1| \rangle$, which equals 0·736 and 0·968 for the non-centrosymmetric and centrosymmetric cases, respectively. In case (iii), I_c will also approach 1 as N approaches infinity, so that the value of R will tend to 0.

Table 3. *Theoretical values for the discrepancy index R defined by equation (A9) between observed and calculated non-centrosymmetric intensities*

The details of the three cases are described in the text. N and $\Gamma(x)$ are the number of crystals in the twin, and the gamma function (Erdelyi, 1954), respectively.

Case	R
(i)	$2[(N-1)/N]^N$
(ii)	$2[N/(N+1)]^N$
(iii)	$\dfrac{2\Gamma[(2N+1)/2]}{N\sqrt{\pi}\,\Gamma(N)}$

Table 4. *Theoretical values for the discrepancy index R defined by equation (A9) between observed and calculated centrosymmetric intensities*

The details of the three cases are described in the text. N, $\Gamma(x)$, and ${}_2F_1(p,q;r;s)$ are the number of crystals in the twin, the gamma function and the Gauss hypergeometric function (Erdelyi, 1954), respectively.

Case	R
(i)	$\dfrac{16(N-1)^{(N-3)/2}\Gamma[(N+2)/2]}{3\sqrt{\pi}N^{N/2}\Gamma[(N-1)/2]}\,{}_2F_1[2, (3-N)/2; 5/2; 1/(1-N)]$
(ii)	$\dfrac{16N^{(N-4)/2}\Gamma[(N+3)/2]}{3\sqrt{\pi}(N+1)^{(N-1)/2}\Gamma(N/2)}\,{}_2F_1[2, (2-N)/2; 5/2; -1/N]$
(iii)	$\dfrac{4\Gamma[(N+1)/2]}{N\sqrt{\pi}\Gamma(N/2)}$

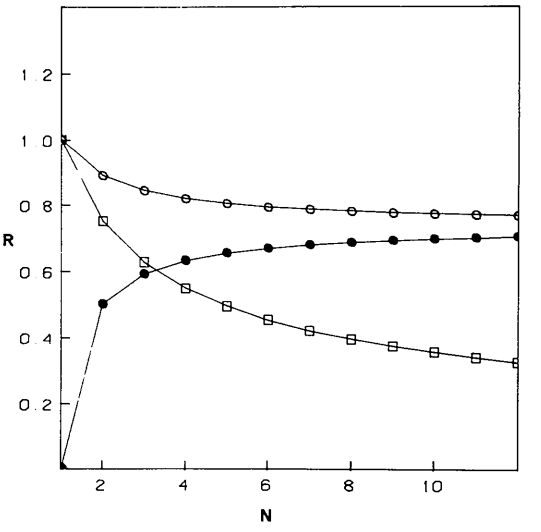

Fig. 4. Dependence of the discrepancy index R, defined in equation (A9), for non-centrosymmetric reflections, on the number of crystals in the twin, N, for cases (i) (●), (ii) (○), and (iii) (□). The details of the three cases are described in the text.

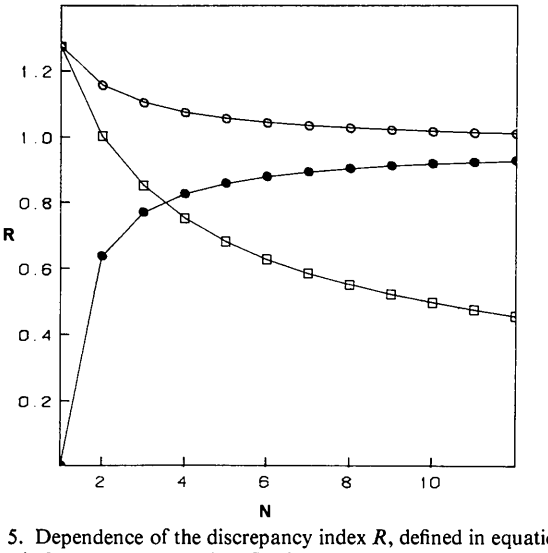

Fig. 5. Dependence of the discrepancy index R, defined in equation (A9), for centrosymmetric reflections, on the number of crystals in the twin, N, for cases (i) (●), (ii) (○), and (iii) (□). The details of the three cases are described in the text.

By replacing $P_s(|\delta|)$ in (A11) with $P_s(|y|)$ from (A5) and (A6), one may verify the expressions for $R_{\alpha\beta}$ derived in equation (3) of the main text for the case of twinning by hemihedry.

An especially striking feature of these results is the deleterious effect of perfect twinning on the R factor for a correct structure model even at low values of N: for $N = 2$, the theoretical minimum values for R are 0·50 and 0·63 for non-centrosymmetric and centrosymmetric data, respectively.

References

BRITTON, D. (1972). *Acta Cryst.* A**28**, 296–297.

CATTI, M. & FERRARIS, G. (1976). *Acta Cryst.* A**32**, 163–165.

CHANDRASEKHAR, S. (1943). *Rev. Mod. Phys.* **15**, 1–89.

DONNAY, G. & DONNAY, J. D. H. (1974). *Can. Mineral.* **12**, 422–425.

ERDELYI, A. L. (1954). *Tables of Integral Transforms*, Vol. I. New York: McGraw-Hill.

FISHER, R. G. & SWEET, R. M. (1980). *Acta Cryst.* A**36**, 755–760.

GRAINGER, C. T. (1969). *Acta Cryst.* A**25**, 427–434.

MURRAY-RUST, P. (1973). *Acta Cryst.* B**29**, 2559–2566.

REES, D. C. (1980). *Acta Cryst.* A**36**, 578–581.

REES, D. C. & LIPSCOMB, W. N. (1980). *Proc. Natl Acad. Sci. USA*, **77**, 4633–4637.

SRINIVASAN, R. & PARTHASARATHY, S. (1976). *Some Statistical Applications in X-ray Crystallography*. Oxford: Pergamon.

STANLEY, E. (1972). *J. Appl. Cryst.* **5**, 191–194.

STROUD, A. H. & SECREST, D. (1966). *Gaussian Quadrature Formulas*. Englewood Cliffs: Prentice-Hall.

WHITTAKER, E. T. & WATSON, G. N. (1927). *Modern Analysis*, 4th ed. Cambridge Univ. Press.

WILSON, A. J. C. (1949). *Acta Cryst.* **2**, 318–321.

Acta Cryst. (1982). A**38**, 207–211

The Effect of Crystal Bending on Direct Phasing of Electron Diffraction Data from Cytosine

BY BARBARA MOSS AND DOUGLAS L. DORSET

Molecular Biophysics Department, Medical Foundation of Buffalo, Inc., 73 High Street, Buffalo, New York 14203, *USA*

(*Received* 29 *June* 1981; *accepted* 21 *September* 1981)

Abstract

Limiting conditions for the crystal structure analysis of organics using electron diffraction intensity data from elastically bent microcrystals are shown for a representative aromatic structure, cytosine, $C_4H_5N_3O$. In a projection down the longest unit-cell axis, the normalized structure-factor magnitudes are greatly changed by slight bends, making the diffraction data useless for crystal structure analysis. This alteration of intensity is less severe for a projection down the shortest cell axis and allows a correct structure analysis for bends comparable to those measured experimentally. The correct crystal orientation, moreover, is only achieved by epitaxial growth and not solution growth.

Introduction

As effects due to crystal bending are routinely observed in experimental electron microscopy, their consequence for the crystal structure analysis from electron diffraction data should be considered. Direct phase determination has been successfully applied to electron diffraction intensity data from an orthorhombic paraffin and a cephalin (Dorset & Hauptman, 1976) to elucidate the aliphatic chain packing. The effect of n-beam dynamical scattering on the success of this phasing procedure has been recently reported (Dorset, Jap, Ho & Glaeser, 1979). Two organic structures, anhydrous cytosine and disodium 4-oxo-pyrimidine-2-sulphinate hexahydrate, were used to investigate the dependence of correct phasing on crystal thickness and electron beam energy. Dynamical structure factors from crystals up to about 75 Å thick, at 100 kV, gave correct kinematical phases and yielded correct crystal structures.

As shown by Cowley (1961), elastic bends also can appreciably affect diffraction patterns, altering both intensities and the apparent symmetry of the zone. This has been demonstrated with experimental data from several long-chain paraffinic materials (Dorset, 1979; Dorset, 1980). Since solvent-grown crystals of these materials have a long unit-cell edge parallel to the

0567-7394/82/020207-05$01.00