

vanishes. Fig. 2 shows a comparison of the first term of (55) with Wilson's formula. They agree well with each other, especially for $b > |h|$. Consequently, Wilson's formula is almost the same as the first term of the asymptotic expression of $I_h(b)$ in (55), which is not as correct as $I_h(b)$ of (13), and hence Wilson's formula gives a poor approximation as the crystal size decreases.

For a spherical crystallite, since $V_h(r) = V_r\{1 - 3r/2D + \frac{1}{2}(r/D)^3\}$, (19), and $V_{h\theta}(r) = V_{h00}(r) = 0$, (50) agrees with the exact form (20). As shown for a cubic crystallite in Fig. 1, the asymptotic expression (50) gives an excellent agreement with the exact

intensity. It can be concluded that the intensity profile can be exactly calculated by the asymptotic expansion including three single integrals, which can be calculated as easily as the Wilson's formula.

References

- INO, T. & MINAMI, N. (1979). *Acta Cryst.* **A35**, 163–170.
 LAUE, M. VON (1926). *Z. Kristallogr.* **64**, 115–142.
 LIGHTHILL, M. J. (1958). *An Introduction to Fourier Analysis and Generalized Function*, pp. 46–57. Cambridge University Press.
 PATTERSON, A. L. (1939). *Phys. Rev.* **56**, 978–982.
 WARREN, B. E. (1978). *J. Appl. Cryst.* **11**, 695–698.
 WILSON, A. J. C. (1962). *X-ray Optics*, 2nd ed. London: Methuen.

Acta Cryst. (1984). **A40**, 544–548

A Simple Interpretation of Integrated Direct Methods – Isomorphous Replacement Probability Distributions

BY SUZANNE FORTIER

Department of Chemistry, Queen's University, Kingston, Canada K7L 3N6

AND CHARLES M. WEEKS AND HERBERT HAUPTMAN

Medical Foundation of Buffalo, Inc., 73 High Street, Buffalo, NY 14203, USA

(Received 9 November 1983; accepted 12 April 1984)

Abstract

Integrated direct methods – isomorphous probability distributions [Hauptman (1982). *Acta Cryst.* **A38**, 289–294] are interpreted in terms of the familiar parameters of the isomorphous replacement method, the diffraction ratio and the differences in the diffraction intensities of a native protein and its heavy-atom derivative. The analysis shows that the reliability of the phase estimates is a function of the degree of heavy-atom substitution in the derivative. It clearly pinpoints the most favorable conditions for retrieving phase information from the intensity data of an isomorphous pair of structures. Finally, it provides a means to determine *a priori* the overall reliability of the phase estimates and to design the calculations accordingly.

1. Introduction

Crystallographic studies of molecular structures have been traditionally divided into two groups, those of macromolecules and those of small molecules. Although the central problem of any crystal structure

determination, the so-called phase problem, is shared by both groups, the techniques developed in the pursuit of a solution to this problem have followed significantly different courses. Bearing in mind the intrinsic differences between the two groups, it has nevertheless been felt, in the last few years, that the expertise acquired in both fields, when properly integrated, could strengthen the present methods of structure determination, whether applied to small or large molecules. The theoretical basis of integrated direct methods – isomorphous replacement techniques has been introduced recently (Hauptman, 1982). The distributions, although presented in a form suitable for computation, are rather complex and impermeable to straightforward interpretation in terms of the diffraction experiment performed. In the present paper we wish to show how a form of the distributions, in terms of the experimental parameters, can be obtained easily through simple mathematical manipulations. The gains from such an exercise are twofold. Not only can we acquire a better understanding of the behavior and scope of the distributions but we can also gain valuable information on how to improve the experiment.

2. Integrated direct methods – isomorphous replacement approach

2.1. The conditional probability distributions for a pair of isomorphous structures in $P1$

The data base at hand consists of two diffraction data sets. For each reciprocal-lattice vector \mathbf{H} , there exist two normalized structure factors $E_{\mathbf{H}}$ and $G_{\mathbf{H}}$. For a triplet of reciprocal-lattice vectors $\mathbf{H}, \mathbf{K}, \mathbf{L}$ satisfying $\mathbf{H} + \mathbf{K} + \mathbf{L} = 0$, there exist eight structure invariants

$$\begin{aligned}\omega_1 &= \varphi_{\mathbf{H}} + \varphi_{\mathbf{K}} + \varphi_{\mathbf{L}} \\ \omega_2 &= \varphi_{\mathbf{H}} + \varphi_{\mathbf{K}} + \psi_{\mathbf{L}} \\ \omega_3 &= \varphi_{\mathbf{H}} + \psi_{\mathbf{K}} + \varphi_{\mathbf{L}} \\ \omega_4 &= \psi_{\mathbf{H}} + \varphi_{\mathbf{K}} + \varphi_{\mathbf{L}} \\ \omega_5 &= \varphi_{\mathbf{H}} + \psi_{\mathbf{K}} + \psi_{\mathbf{L}} \\ \omega_6 &= \psi_{\mathbf{H}} + \varphi_{\mathbf{K}} + \psi_{\mathbf{L}} \\ \omega_7 &= \psi_{\mathbf{H}} + \psi_{\mathbf{K}} + \varphi_{\mathbf{L}} \\ \omega_8 &= \psi_{\mathbf{H}} + \psi_{\mathbf{K}} + \psi_{\mathbf{L}},\end{aligned}\quad (1)$$

where the φ 's and ψ 's are the phases associated with the isomorphous pair of structures.

The task at hand is to extract from the six magnitudes $|E_{\mathbf{H}}|, |E_{\mathbf{K}}|, |E_{\mathbf{L}}|, |G_{\mathbf{H}}|, |G_{\mathbf{K}}|, |G_{\mathbf{L}}|$ (the first neighborhood) estimates of the eight invariants (1). The first step is to derive the joint probability distribution of the six normalized structure factors $E_{\mathbf{H}}, E_{\mathbf{K}}, E_{\mathbf{L}}, G_{\mathbf{H}}, G_{\mathbf{K}}, G_{\mathbf{L}}$. From that distribution, the conditional probability distribution of each of the eight invariants, given the six magnitudes $|E_{\mathbf{H}}|, |E_{\mathbf{K}}|, |E_{\mathbf{L}}|, |G_{\mathbf{H}}|, |G_{\mathbf{K}}|, |G_{\mathbf{L}}|$, is obtained. A detailed account of this work has been reported (Hauptman, 1982).

Let

$$\begin{aligned}|E_{\mathbf{H}}| &= R_1, |E_{\mathbf{K}}| = R_2, |E_{\mathbf{L}}| = R_3; \\ |G_{\mathbf{H}}| &= S_1, |G_{\mathbf{K}}| = S_2, |G_{\mathbf{L}}| = S_3.\end{aligned}\quad (2)$$

The final formula is

$$P_i(\Omega_i | R_1, R_2, R_3, S_1, S_2, S_3) \approx 1/K_i \exp(A_i \cos \Omega_i), \quad i = 1, 2, \dots, 8, \quad (3)$$

where

$$K_i = 2\pi I_0(A_i) \quad (4)$$

and I_0 is the modified Bessel function.

The distribution (3) is analogous to that of the traditional three-phase invariant (Cochran, 1955; Hauptman, 1976). The A_i term, however, is significantly different.

$$\begin{aligned}A_i &= 2\{\beta_1 \tau_1 R_1 R_2 R_3 \\ &+ \beta_2 [\tau_{21} R_1 R_2 S_3 + \tau_{22} R_1 S_2 R_3 + \tau_{23} S_1 R_2 R_3] \\ &+ \beta_3 [\tau_{31} R_1 S_2 S_3 + \tau_{32} S_1 R_2 S_3 + \tau_{33} S_1 S_2 R_3] \\ &+ \beta_4 \tau_4 S_1 S_2 S_3\},\end{aligned}\quad (5)$$

where $\tau = C_1 C_2 C_3$ is obtained by comparing the i th structure factor associated with the coefficient of τ with the i th structure factor associated with the invariant. If they are of the same type, *i.e.* both R or both S , then $C_i = 1.0$, $i = 1, 2, 3$. If one is of type R and the other of type S , then

$$C_i = I_1(2\gamma R_i S_i) / I_0(2\gamma R_i S_i), \quad i = 1, 2, 3, \quad (6)$$

$$\gamma = \alpha_{20}^{1/2} \alpha_{02}^{1/2} a_{11} / (\alpha_{20} \alpha_{02} - \alpha_{11}^2), \quad (7)$$

I_1 and I_0 are the modified Bessel functions,

$$\begin{aligned}\beta_1 &= \frac{\alpha_{20}^{3/2}}{(\alpha_{20} \alpha_{02} - \alpha_{11}^2)^3} [\alpha_{30} \alpha_{02}^3 - 3\alpha_{21} \alpha_{02}^2 \alpha_{11} \\ &+ 3\alpha_{12} \alpha_{02} \alpha_{11}^2 - \alpha_{03} \alpha_{11}^3],\end{aligned}\quad (8)$$

$$\begin{aligned}\beta_2 &= \frac{\alpha_{20} \alpha_{02}^{1/2}}{(\alpha_{20} \alpha_{02} - \alpha_{11}^2)^3} [(\alpha_{21} \alpha_{20} - \alpha_{30} \alpha_{11}) \alpha_{02}^2 \\ &- 2(\alpha_{12} \alpha_{20} - \alpha_{21} \alpha_{11}) \alpha_{02} \alpha_{11} \\ &+ (\alpha_{03} \alpha_{20} - \alpha_{12} \alpha_{11}) \alpha_{11}^2],\end{aligned}\quad (9)$$

$$\begin{aligned}\beta_3 &= \frac{\alpha_{20}^{1/2} \alpha_{02}}{(\alpha_{20} \alpha_{02} - \alpha_{11}^2)^3} [(\alpha_{12} \alpha_{02} - \alpha_{03} \alpha_{11}) \alpha_{20}^2 \\ &- 2(\alpha_{21} \alpha_{02} - \alpha_{12} \alpha_{11}) \alpha_{20} \alpha_{11} \\ &+ (\alpha_{30} \alpha_{02} - \alpha_{21} \alpha_{11}) \alpha_{11}^2],\end{aligned}\quad (10)$$

$$\begin{aligned}\beta_4 &= \frac{\alpha_{02}^{3/2}}{(\alpha_{20} \alpha_{02} - \alpha_{11}^2)^3} [\alpha_{03} \alpha_{20}^3 - 3\alpha_{12} \alpha_{20}^2 \alpha_{11} \\ &+ 3\alpha_{21} \alpha_{20} \alpha_{11}^2 - \alpha_{30} \alpha_{11}^3],\end{aligned}\quad (11)$$

$$\alpha_{mn} = \sum_{j=1}^N f_j^m g_j^n,$$

where f_j and g_j denote atomic structure factors for a corresponding pair of isomorphous structures.

The formula (3) obtained by Hauptman (1982) is completely general and is valid for any isomorphous pair of structures in $P1$. It includes as special cases native protein and heavy-atom isomorphous derivatives as well as X-ray and neutron diffraction data. For the special case of a native protein and a heavy-atom isomorphous derivative, we will show how the parameters of the A term (5) are related to the diffraction ratio and the intensity differences, the familiar parameters of the isomorphous replacement method.

Let

$$\begin{aligned}S_1 &= R_1 + \Delta_1, \\ S_2 &= R_2 + \Delta_2, \\ S_3 &= R_3 + \Delta_3;\end{aligned}\quad (12)$$

then (5) can be rewritten as

$$\begin{aligned}
 A_i = & 2\{R_1 R_2 R_3[\beta_1 \tau_1 + \beta_2(\tau_{21} + \tau_{22} + \tau_{23}) \\
 & + \beta_3(\tau_{31} + \tau_{32} + \tau_{33}) + \beta_4 \tau_4] \\
 & + R_1 R_2 \Delta_3[\beta_2 \tau_{21} + \beta_3(\tau_{31} + \tau_{32}) + \beta_4 \tau_4] \\
 & + R_1 \Delta_2 R_3[\beta_2 \tau_{22} + \beta_3(\tau_{31} + \tau_{33}) + \beta_4 \tau_4] \\
 & + \Delta_1 R_2 R_3[\beta_2 \tau_{23} + \beta_3(\tau_{32} + \tau_{33}) + \beta_4 \tau_4] \\
 & + R_1 \Delta_2 \Delta_3[\beta_3 \tau_{31} + \beta_4 \tau_4] \\
 & + \Delta_1 R_2 \Delta_3[\beta_3 \tau_{32} + \beta_4 \tau_4] \\
 & + \Delta_1 \Delta_2 R_3[\beta_3 \tau_{33} + \beta_4 \tau_4] \\
 & + \Delta_1 \Delta_2 \Delta_3 \beta_4 \tau_4\}. \quad (13)
 \end{aligned}$$

In particular, when the τ 's approach 1, *i.e.* when the $2\gamma R_i S_i$'s are large, we have

$$\begin{aligned}
 A_i = & 2\{R_1 R_2 R_3[\beta_1 + 3\beta_2 + 3\beta_3 + \beta_4] \\
 & + [R_1 R_2 \Delta_3 + R_1 \Delta_2 R_3 + \Delta_1 R_2 R_3][\beta_2 + 2\beta_3 + \beta_4] \\
 & + [R_1 \Delta_2 \Delta_3 + \Delta_1 R_2 \Delta_3 + \Delta_1 \Delta_2 R_3][\beta_3 + \beta_4] \\
 & + \Delta_1 \Delta_2 \Delta_3 \beta_4\}. \quad (14)
 \end{aligned}$$

If we assume that the atomic content of the derivative equals the atomic content of the native protein (p) plus the heavy-atom content (H) then

$$\gamma = \alpha_{20}^{1/2} \alpha_{02}^{1/2} / (\alpha_{02} - \alpha_{20}), \quad (15)$$

$$\begin{aligned}
 \beta_1 = & \frac{-(\alpha_{03} - \alpha_{30})}{(\alpha_{02} - \alpha_{20})^3} \alpha_{20}^{3/2} + \frac{\alpha_{30}}{\alpha_{20}^{3/2}} \\
 \approx & \frac{-(\alpha_{03} - \alpha_{30}) \alpha_{20}^{3/2}}{(\alpha_{02} - \alpha_{20})^3}, \quad (16)
 \end{aligned}$$

$$\beta_2 = \frac{(\alpha_{03} - \alpha_{30}) \alpha_{20} \alpha_{02}^{1/2}}{(\alpha_{02} - \alpha_{20})^3}, \quad (17)$$

$$\beta_3 = \frac{-(\alpha_{03} - \alpha_{30}) \alpha_{20}^{1/2} \alpha_{02}}{(\alpha_{02} - \alpha_{20})^3}, \quad (18)$$

$$\beta_4 = \frac{(\alpha_{03} - \alpha_{30}) \alpha_{02}^{3/2}}{(\alpha_{02} - \alpha_{20})^3} \quad (19)$$

and

$$\begin{aligned}
 A_i = & 2 \frac{\alpha_{03} - \alpha_{30}}{(\alpha_{02} - \alpha_{20})^3} \{R_1 R_2 R_3 (\alpha_{02}^{1/2} - \alpha_{20}^{1/2})^3 \\
 & + [R_1 R_2 \Delta_3 + R_1 \Delta_2 R_3 + \Delta_1 R_2 R_3] \alpha_{02}^{1/2} (\alpha_{02}^{1/2} - \alpha_{20}^{1/2})^2 \\
 & + [R_1 \Delta_2 \Delta_3 + \Delta_1 R_2 \Delta_3 + \Delta_1 \Delta_2 R_3] \alpha_{02} (\alpha_{02}^{1/2} - \alpha_{20}^{1/2}) \\
 & + \Delta_1 \Delta_2 \Delta_3 \alpha_{02}^{3/2}\}. \quad (20)
 \end{aligned}$$

For abbreviation, define

$$\begin{aligned}
 \sum_{j \in p} Z_j^n &= \sum Z_p^n, \\
 \sum_{k \in H} Z_k^n &= \sum Z_H^n. \quad (21)
 \end{aligned}$$

The coefficient $(\alpha_{03} - \alpha_{30}) / (\alpha_{02} - \alpha_{20})^3$ can be expressed as $\sum Z_H^3 / (\sum Z_H^2)^3$. The distribution does not depend, as in the case of the traditional three-phase invariant, on the total number of atoms per unit cell but rather on the scattering difference between the native protein and the derivative – that is, on the scattering of the heavy atoms in the derivative. For the special case in which the heavy atoms in the derivative are of equal weight and equal occupancy, the distribution depends on the number of heavy atoms in the derivative. Since this number is usually small, it becomes clear that the distribution is capable of yielding extremely reliable estimates.

Finally, using the binomial series to evaluate the $\alpha_{20}^{1/2}$ and $\alpha_{02}^{1/2}$ terms, the A_i term is reduced to

$$\begin{aligned}
 A_i = & \sum Z_H^3 / 4 (\sum Z_p^2)^{3/2} \{R_1 R_2 R_3 \\
 & + 2\gamma (R_1 R_2 \Delta_3 + R_1 \Delta_2 R_3 + \Delta_1 R_2 R_3) \\
 & + 4\gamma^2 (R_1 \Delta_2 \Delta_3 + \Delta_1 R_2 \Delta_3 + \Delta_1 \Delta_2 R_3) \\
 & + 8\gamma^3 \Delta_1 \Delta_2 \Delta_3\}, \quad (22)
 \end{aligned}$$

where

$$2\gamma \approx (1 + 2\sum Z_p^2 / \sum Z_H^2). \quad (23)$$

The coefficient 2γ (23) is related to the diffraction ratio. This ratio, which is a measure of the average change in intensity due to the addition of heavy atoms, is estimated at low resolution as

$$(2\sum Z_H^2 / \sum Z_p^2)^{1/2},$$

(Crick & Magdoff, 1956). Hence,

$$2\gamma \approx [1 + 4/(\text{diffraction ratio})^2]. \quad (24)$$

The term 2γ is usually large compared to unity while the term $\sum Z_H^3 / 4 (\sum Z_p^2)^{3/2}$ is usually very small as can be seen from Table 1, illustrating the changes in these parameters as platinum atoms are added to proteins of different complexities.

2.2. The optimal case

The eight magnitude product terms of the distribution, $R_1 R_2 R_3$, $R_1 R_2 \Delta_3$, $R_1 \Delta_2 R_3$, $\Delta_1 R_2 R_3$, $R_1 \Delta_2 \Delta_3$, $\Delta_1 R_2 \Delta_3$, $\Delta_1 \Delta_2 R_3$, $\Delta_1 \Delta_2 \Delta_3$, do not contribute equally to the A_i term as can be seen from Table 2.

Comparison of these coefficients shows that the predominant terms of the distribution are the $\Delta\Delta\Delta$ term and, to a lesser extent, the $R\Delta\Delta$ terms.

The distribution is capable of yielding extremely reliable estimates, particularly in those cases where both the Δ 's and the 2γ coefficient are large. The distribution clearly pinpoints the most favorable experimental conditions for retrieving phase information from intensity data of isomorphous pairs of structures. In terms of the experiment, these conditions are satisfied when the diffraction ratio, or the degree of heavy-atom substitution, is relatively small and yet differences in the diffraction intensities between the

Table 1. Variation of 2γ and $\sum Z_H^3/4(\sum Z_P^2)^{3/2}$ with the diffraction ratio

Native protein, $M_R = 15\ 000$			
Number of Pt atoms	Diffraction ratio	2γ	$\sum Z_H^3/4(\sum Z_P^2)^{3/2}$
1	0.492	17.519	0.010
2	0.696	9.259	0.022
3	0.852	6.506	0.032
4	0.984	5.130	0.042
Native protein, $M_R = 25\ 000$			
Number of Pt atoms	Diffraction ratio	2γ	$\sum Z_H^3/4(\sum Z_P^2)^{3/2}$
1	0.381	28.531	0.004
2	0.539	14.766	0.010
3	0.660	10.177	0.014
4	0.762	7.883	0.020
5	0.852	6.506	0.024
6	0.934	5.589	0.030

Native protein, $M_R = 45\ 000$

Number of Pt atoms	Diffraction ratio	2γ	$\sum Z_H^3/4(\sum Z_P^2)^{3/2}$
1	0.284	50.556	0.002
2	0.402	25.778	0.004
3	0.492	17.519	0.006
4	0.568	13.389	0.008
5	0.635	10.911	0.010
6	0.696	9.259	0.012
7	0.752	8.079	0.014
8	0.804	7.195	0.016

Table 2. Variation of the coefficients of the magnitude product terms with the diffraction ratio

$M_R = 15\ 000$					
Number of Pt atoms	Diffraction ratio	RRR coefficient	RR Δ coefficient	R $\Delta\Delta$ coefficient	$\Delta\Delta\Delta$ coefficient
1	0.492	0.010	0.184	3.232	56.628
2	0.696	0.022	0.196	1.806	16.722
3	0.852	0.032	0.206	1.338	8.702
4	0.984	0.042	0.216	1.108	5.686

native protein and the derivative can be observed. There is clearly an optimal amount of heavy-atom substitution that leads to both sufficiently large differences in the intensities (and consequently in the normalized structure factors) and sufficiently large 2γ coefficient. Furthermore, provided that the diffraction ratio is known, it becomes possible to evaluate *a priori* the overall reliability of the invariant estimates and to design the calculations accordingly. For example, in optimal cases the estimates are extremely reliable and their calculations worthwhile. When the diffraction ratio becomes exceedingly large, however, the distribution converges to the traditional three-phase invariant and the benefits from the calculation of the integrated direct methods—*isomorphous replacement distributions* are marginal. In the optimal cases, since the predominant terms of the distribution are the $\Delta_1\Delta_2\Delta_3$ term, and to a lesser extent the $R_1\Delta_2\Delta_3$, $\Delta_1R_2\Delta_3$, $\Delta_1\Delta_2R_3$ terms, it follows that reliable esti-

Table 3. Contribution to the A value from the eight magnitude product terms in the distribution

$M_R = 15\ 000$ One Pt derivative			
$R_1 = R_2 = R_3 = 2.0$			
$\Delta_1 = \Delta_2 = \Delta_3 = 0.5$			
Contributors to the A value			
$R_1R_2R_3$	$R_1R_2\Delta_3$ $R_1\Delta_2R_3$ $\Delta_1R_2R_3$	$R_1\Delta_2\Delta_3$ $\Delta_1R_2\Delta_3$ $\Delta_1\Delta_2R_3$	$\Delta_1\Delta_2\Delta_3$
0.08	0.37 0.37 0.37	1.62 1.62 1.62	7.08
0.08	1.10	4.85	7.08

mates (*i.e.* large A values) can be obtained, even when the normalized structure factors themselves are small, provided that the differences between the normalized structure factors of the native protein and the derivative are large. Furthermore, since the Δ 's are signed values, both 0 and 180° estimates are obtainable.

2.3. Comparison with Karle's simple rule

A simple rule for estimating the values of triplet phase invariants (0 or 180°) in *isomorphous replacement procedures* was presented recently (Karle, 1983). When the simple rule, R_{iso} , is used with normalized structure factors, it corresponds to evaluating the $\Delta_1\Delta_2\Delta_3$ term in the present analysis. In many instances, the simple rule and the full distribution will yield identical estimates. The associated A values, however, can be significantly different as shown in Table 3. If all the terms are used an A value of 13.1 is obtained compared to one of 7.08 when only the $\Delta_1\Delta_2\Delta_3$ term is used. Since the A term is normally used to assign a weight to the invariant estimate, such differences could be important. Furthermore, as would be expected, the time involved in the calculation of the full distribution is negligible compared to the time involved in the generation of the invariants. Proper comparison of both methods, however, cannot be made until they are both tested against real diffraction data.

2.4. Evaluation of the diffraction ratio

As can be seen from the form of the A term (22) and from Table 1, detailed knowledge of the heavy-atom content in the derivative is not needed for the evaluation of invariant phase estimates. The parameters of the distribution can be calculated from the diffraction ratio alone. However, in many instances, because of uncertainty in the relative scaling of the native and derivative intensity data, the diffraction ratio cannot be evaluated with great accuracy. An alternative method to determine this ratio has been proposed recently (Hauptman, 1982).

$$r = \frac{\langle (|E_H|^2 - \overline{|E_H|^2}) (|G_H|^2 - \overline{|G_H|^2}) \rangle_H}{\langle (|E_H|^2 - \overline{|E_H|^2})^2 \rangle_H^{1/2} \langle (|G_H|^2 - \overline{|G_H|^2})^2 \rangle_H^{1/2}} \quad (25)$$

is shown to be equal to $\alpha_{11}^2 / \alpha_{20} \alpha_{02}$. Under the assumption that the atomic content of the derivative equals the atomic content of the native protein (p) plus the heavy-atom content (H),

$$r = 1 / (1 + \sum Z_H^2 / \sum Z_p^2) = 1 / [1 + (\text{diffraction ratio})^2 / 2]. \quad (26)$$

2.5. Phase accessibility

As is the case of the traditional three-phase invariant, although probability distributions can be calculated for the whole family of invariants (at a given resolution) only a subset of these invariants can be reliably estimated. In the present case, however, this subset consists largely of invariants whose phases are associated with normalized structure factors for which large differences between the native and derivative diffraction intensities are observed. Consequently only a subset of the structure factors can be reliably phased. Unfortunately, this subset, in some cases, may not coincide with that needed to calculate an interpretable density map. It is hoped that, in those instances, the phase set may be extended through the use of quartet invariants, standard phase extension techniques or density modification procedures.

3. Concluding remarks

In recent years, a formal mathematical integration of the techniques of direct methods and isomorphous replacement has been undertaken (Hauptman, 1982; Fortier, Weeks & Hauptman, 1984). The amount of

information contained in the probability distributions is extensive, although often hidden behind the mathematical complexity of the formulae. As we have shown, it is relatively easy to translate the distributions into the usual experimental parameters. Through such an exercise, a better understanding of the nature and scope of the distributions is attained. Conversely, the mathematical formulae yield a better understanding of the experiment, and indicate ways to improve and gauge the experiment. The exact role of direct methods in macromolecular structure determination cannot be predicted at this point. As has been the case with the traditional direct methods, several years of experience in the application of these methods will probably be needed before an accurate evaluation can be made. With the extensive theoretical base now at hand, and the extremely promising results obtained to date, we are now in a position to address many of the unanswered questions, principally pertaining to the application of the methods to real diffraction data.

This research was supported in part by the Natural Sciences and Engineering Research Council of Canada (SF), a Queen's University Research Award (SF); Grant No. CHE-8203930 from the National Science Foundation (CMW and HH) and a grant from the James H. Cummings Foundation (CMW and HH).

References

- COCHRAN, W. (1955). *Acta Cryst.* **8**, 473–478.
 CRICK, F. H. C. & MAGDOFF, B. (1956). *Acta Cryst.* **9**, 901–908.
 FORTIER, S., WEEKS, C. M. & HAUPTMAN, H. (1984). *Acta Cryst.* Submitted.
 HAUPTMAN, H. (1976). *Acta Cryst.* **A32**, 877–882.
 HAUPTMAN, H. (1982). *Acta Cryst.* **A38**, 289–294.
 KARLE, J. (1983). *Acta Cryst.* **A39**, 800–805.

Acta Cryst. (1984). **A40**, 548–558

Measurability of Bijvoet Differences in Triclinic, Monoclinic and Orthorhombic Crystals

BY D. VELMURUGAN AND S. PARTHASARATHY

Department of Crystallography and Biophysics, University of Madras, Guindy Campus, Madras-600 025, India*

(Received 7 September 1982; accepted 11 April 1984)

Abstract

Theoretical expressions for the measurability of Bijvoet differences have been derived for triclinic, monoclinic and orthorhombic crystals containing p

(= 1 or 2) anomalous scatterers and many normal scatterers per asymmetric unit. Results for the many-atom cases (*i.e.* $P = MN$ and MC cases) in space group $P1$ are also obtained. The theory takes into consideration the effect of data truncation due to unobserved reflections. The measurability values for the various cases are given in the form of compact

* Contribution No. 641.