

The authors thank Dr M. J. Cooper for provision of the sample.

References

- BUSING, W. R., MARTIN, K. O. & LEVY, H. A. (1962). *ORFLS*. Report ORNL-TM-305. Oak Ridge National Laboratory, Tennessee, USA.
- COOPER, M. J. & ROUSE, K. D. (1971). *Acta Cryst.* **A27**, 622–628.
- COOPER, M. J., ROUSE, K. D. & WILLIS, B. T. M. (1968). *Acta Cryst.* **A24**, 484–493.
- FORSYTH, J. B., LAWRENCE, R. T. & WILSON, C. C. (1988). *Nucl. Instrum. Methods A*. In the press.
- FORSYTH, J. B. & WILSON, C. C. (1988). Rutherford Appleton Laboratory Rep. In preparation.
- GUREVICH, I. I. & TARASOV, L. V. (1968). *Low-Energy Neutron Physics*. Amsterdam: North-Holland.
- MAIR, S. L. & BARNEA, Z. (1971). *Phys. Lett.* **35A**, 286–287.
- MAIR, S. L., BARNEA, Z., COOPER, M. J. & ROUSE, K. D. (1974). *Acta Cryst.* **A30**, 806–813.
- MATTHEWMAN, J. C., THOMPSON, P. & BROWN, P. J. (1982). *J. Appl. Cryst.* **15**, 167–173.
- TAYLOR, A. D. (1984). Report RAL-84-120. Rutherford Appleton Laboratory, Didcot, Oxon, England.
- WILLIS, B. T. M. & PRYOR, A. W. (1975). *Thermal Vibrations in Crystallography*. Cambridge Univ. Press.

Acta Cryst. (1989). **A45**, 247–254

On the Probabilistic Theory of Isomorphous Data Sets: General Joint Distributions for the SIR, SAS and Partial/Complete Structure Cases

BY SUZANNE FORTIER* AND GUR DAYAL NIGAM†

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

(Received 28 April 1988; accepted 5 October 1988)

Abstract

By characterizing isomorphism in reciprocal space [*i.e.* diffraction data sets are isomorphous if they have the same geometry (the same reciprocal-lattice unit cell) and the same symmetry] it is shown that the diffraction data of a native protein and of its heavy-atom derivatives, the calculated data of a partial structure and the observed data of its associated complete structure, and the Friedel-pair data of an anomalously scattering crystal structure all belong to the more general class of isomorphous data sets. Their joint probability distributions for two- and three-phase structure invariants are shown to be isomorphous: they have the same functional form and differ only in individual atomic scattering factors. General joint probability distributions, which can be used for any isomorphous data pairs, are presented.

1. Introduction

Isomorphism is usually defined in direct space as, for example, the 'similarity of crystal shape, unit-cell dimensions, and structure between substances of similar chemical composition' (Glusker & Trueblood, 1985). For the present work, it is convenient to characterize isomorphism in reciprocal space; *i.e.* diffraction

data sets are isomorphous if they have the same geometry (the same reciprocal-lattice unit cell) and the same symmetry. Differences between isomorphous data are thus to be found in the intensities of individual reflections, and therefore in the scattering power of a subset of atoms. This allows us to consider a variety of diffraction data sets as representing cases of isomorphous data sets: for example, the diffraction data of a native protein and of its heavy-atom derivatives, the calculated data of a partial structure and the observed data of its associated complete structure, X-ray and neutron data measured on the same substance, and, finally, the Friedel-pair data of an anomalously scattering crystal structure.

Through the use of the method of joint probability distributions, formulae have been obtained to estimate the value of two- and three-phase structure invariants for the cases of isomorphous replacement (Hauptman, 1982*a*; Giacovazzo, Cascarano & Zheng Chao-de, 1988; Srinivasan & Parthasarathy, 1976), anomalous dispersion (Hauptman, 1982*b*; Giacovazzo, 1983*a*) and partial/complete structure (Beurskens, Prick, Doesburg & Gould, 1979; Srinivasan & Parthasarathy, 1976; Giacovazzo, 1983*b*; Sim, 1959). In the present paper, it is shown that the joint probability distributions are isomorphous: that is, they have the same functional form and differ only in individual atomic scattering factors. General joint probability distributions, which can be used for any isomorphous data pairs, are presented.

* Author to whom all correspondence should be addressed.

† Present address: Department of Physics, Indian Institute of Technology, Kharagpur 721 302, India.

2. The probabilistic theory of the two-phase structure invariant $\varphi_{\mathbf{H}} + \psi_{\bar{\mathbf{H}}}$

Let $f_{j\mathbf{H}}$ and $g_{j\mathbf{H}}$ represent atomic scattering factors for a corresponding pair of isomorphous data sets in space group $P1$ and let us define them in a general way, so as to include the case where they are complex numbers, *i.e.*

$$f_{j\mathbf{H}} = |f_{j\mathbf{H}}| \exp(i\delta_{j\mathbf{H}}) \quad (2.1)$$

$$g_{j\mathbf{H}} = |g_{j\mathbf{H}}| \exp(i\eta_{j\mathbf{H}}). \quad (2.2)$$

The normalized structure factors, $E_{\mathbf{H}}$ and $G_{\bar{\mathbf{H}}}$, are defined by

$$\begin{aligned} E_{\mathbf{H}} &= |E_{\mathbf{H}}| \exp i\varphi_{\mathbf{H}} \\ &= \alpha_{\mathbf{H}}^{-1/2} \sum_{j=1}^N f_{j\mathbf{H}} \exp(2\pi i\mathbf{H} \cdot \mathbf{r}_j) \end{aligned} \quad (2.3)$$

$$\begin{aligned} G_{\bar{\mathbf{H}}} &= |G_{\bar{\mathbf{H}}}| \exp i\psi_{\bar{\mathbf{H}}} \\ &= \beta_{\bar{\mathbf{H}}}^{-1/2} \sum_{j=1}^N g_{j\mathbf{H}} \exp(-2\pi i\mathbf{H} \cdot \mathbf{r}_j) \end{aligned} \quad (2.4)$$

where

$$N = \text{number of atoms in the unit cell}, \quad (2.5)$$

$$\alpha_{\mathbf{H}} = \sum_{j=1}^N f_{j\mathbf{H}} f_{j\mathbf{H}}^* = \sum_{j=1}^N |f_{j\mathbf{H}}|^2, \quad (2.6)$$

$$\beta_{\bar{\mathbf{H}}} = \sum_{j=1}^N g_{j\mathbf{H}} g_{j\mathbf{H}}^* = \sum_{j=1}^N |g_{j\mathbf{H}}|^2. \quad (2.7)$$

Denote by $P(R, \bar{S}; \Phi, \bar{\Psi})$ the joint probability distribution of the magnitudes $|E_{\mathbf{H}}|$, $|G_{\bar{\mathbf{H}}}|$ and the phases $\varphi_{\mathbf{H}}$, $\psi_{\bar{\mathbf{H}}}$ of the complex normalized structure factors $E_{\mathbf{H}}$, $G_{\bar{\mathbf{H}}}$. Then, following methods previously described (Karle & Hauptman, 1958; Hauptman, 1982*a, b*) we can write

$$\begin{aligned} P(R, \bar{S}; \Phi, \bar{\Psi}) &= [R\bar{S}/\pi^2(1 - X_{\mathbf{H}}^2)] \exp\{[-R^2 - \bar{S}^2 \\ &\quad + 2XR\bar{S} \cos(\Phi + \bar{\Psi} + \xi_{\mathbf{H}})] \\ &\quad \times (1 - X_{\mathbf{H}}^2)^{-1}\} \end{aligned} \quad (2.8)$$

where

$$X_{\mathbf{H}} \cos \xi_{\mathbf{H}} = C_{\mathbf{H}}, \quad X_{\mathbf{H}} \sin \xi_{\mathbf{H}} = -S_{\mathbf{H}} \quad (2.9)$$

$$X_{\mathbf{H}} = (C_{\mathbf{H}}^2 + S_{\mathbf{H}}^2)^{1/2}, \quad \tan \xi_{\mathbf{H}} = -S_{\mathbf{H}}/C_{\mathbf{H}} \quad (2.10)$$

$$C_{\mathbf{H}} = \alpha_{\mathbf{H}}^{-1/2} \beta_{\bar{\mathbf{H}}}^{-1/2} \sum_{j=1}^N |f_{j\mathbf{H}}| |g_{j\mathbf{H}}| \cos(\delta_{j\mathbf{H}} + \eta_{j\mathbf{H}}) \quad (2.11)$$

$$S_{\mathbf{H}} = \alpha_{\mathbf{H}}^{-1/2} \beta_{\bar{\mathbf{H}}}^{-1/2} \sum_{j=1}^N |f_{j\mathbf{H}}| |g_{j\mathbf{H}}| \sin(\delta_{j\mathbf{H}} + \eta_{j\mathbf{H}}). \quad (2.12)$$

The joint distribution thus obtained can be used for any pair of isomorphous data sets by simply substituting, in each case, the appropriate scattering factors as shown below.

Case 1. Native protein/heavy-atom-derivative data sets (neglecting the effect of anomalous scattering)

When effects of anomalous scattering are not included, we have

$$S = \bar{S} \quad (2.13)$$

and

$$\Psi = -\bar{\Psi}. \quad (2.14)$$

Furthermore, if it is assumed that the atomic scattering factors, as functions of $(\sin \theta)/\lambda$, have the same shape for different atoms then $f_{\mathbf{H}}$ and $g_{\mathbf{H}}$ are real and equal to the atomic numbers Z_j . Using the notation of Hauptman (1982*a*) we obtain

$$C_{\mathbf{H}} = X_{\mathbf{H}} = \alpha_{11}/\alpha_{20}^{1/2} \alpha_{02}^{1/2} = \alpha \quad (2.15)$$

$$S_{\mathbf{H}} = 0 \quad (2.16)$$

$$\xi_{\mathbf{H}} = 0 \quad (2.17)$$

where

$$\alpha_{mn} = \sum_{j=1}^N |f_j|^m |g_j|^n \quad (2.18)$$

and

$$\begin{aligned} P(R, S; \Phi, \Psi) &= [RS/\pi^2(1 - \alpha^2)] \\ &\quad \times \exp\{-[R^2 + S^2 - 2\alpha RS \\ &\quad \times \cos(\Phi - \Psi)]/(1 - \alpha^2)\} \end{aligned} \quad (2.19)$$

which corresponds to formula (2.4) of Hauptman (1982*a*).

Case 2. Friedel-pair data of an anomalously scattering crystal structure

In this case we have, using the notation of Hauptman (1982*b*),

$$\bar{R} = \bar{S} \quad (2.20)$$

$$\bar{\Phi} = \bar{\Psi} \quad (2.21)$$

$$f_{j\mathbf{H}} = g_{j\mathbf{H}} \quad (2.22)$$

and therefore

$$C_{\mathbf{H}} = \alpha_{\mathbf{H}}^{-1} \sum_{j=1}^N |f_{j\mathbf{H}}|^2 \cos(2\delta_{j\mathbf{H}}) \quad (2.23)$$

$$S_{\mathbf{H}} = \alpha_{\mathbf{H}}^{-1} \sum_{j=1}^N |f_{j\mathbf{H}}|^2 \sin(2\delta_{j\mathbf{H}}). \quad (2.24)$$

$X_{\mathbf{H}}$ and $\xi_{\mathbf{H}}$ are defined by (2.9)–(2.10), and

$$\begin{aligned} P(R, \bar{R}; \Phi, \bar{\Phi}) &= [R\bar{R}/\pi^2(1 - X_{\mathbf{H}}^2)] \\ &\quad \times \exp\{[-R^2 - \bar{R}^2 + 2XR\bar{R} \\ &\quad \times \cos(\Phi + \bar{\Phi} + \xi_{\mathbf{H}})]/(1 - X_{\mathbf{H}}^2)\} \end{aligned} \quad (2.25)$$

which corresponds to formula (2.10) of Hauptman (1982*b*). We also note that Giacovazzo (1983*a*) has

obtained a distribution which is identical except for notation.

Case 3. Calculated data of a partial structure and observed data of its associated complete structure (neglecting the effect of anomalous scattering)

In the absence of anomalous scattering, we have

$$S = \bar{S} \quad (2.26)$$

$$\Psi = -\bar{\Psi}. \quad (2.27)$$

Let P = number of atoms in the partial structure unit cell and R = number of atoms ($N - P$) in the remainder of the structure. Then

$$C_H = X_H = \left(\sum_{j=1}^P |f_j|^2 \right)^{1/2} / \left(\sum_{j=1}^N |f_j|^2 \right)^{1/2} \quad (2.28)$$

$$S_H = 0 \quad (2.29)$$

$$\xi_H = 0. \quad (2.30)$$

Using the notation of Beurskens, Prick, Doesburg & Gould (1979) we have

$$p = \left(\sum_{j=1}^P |f_j|^2 \right)^{1/2} / \left(\sum_{j=1}^N |f_j|^2 \right)^{1/2} \quad (2.31)$$

$$r = \left(\sum_{j=1}^R |f_j|^2 \right)^{1/2} / \left(\sum_{j=1}^N |f_j|^2 \right)^{1/2} \quad (2.32)$$

and

$$\begin{aligned} P(R, S; \Phi, \Psi) &= [RS / \pi^2(r^2)] \\ &\times \exp \{ [-R^2 - S^2 + 2pRS \\ &\times \cos(\Phi - \Psi)] / r^2 \}. \end{aligned} \quad (2.33)$$

From (2.33), the conditional distribution of the structure invariant, $\Omega = \Phi - \Psi$, given the magnitudes R and S , is easily obtained. The result

$$\begin{aligned} P(\Omega | RS) &= [2\pi I_0(2pRS/r^2)]^{-1} \\ &\times \exp [(2pRS \cos \Omega) / r^2] \end{aligned} \quad (2.34)$$

can then be compared with equation (17) of Beurskens, Prick, Doesburg & Gould (1979). Related distributions were obtained by Srinivasan & Parthasarathy (1976) and were also derived for structure factors (Sim, 1959; Srinivasan & Parthasarathy, 1976) and pseudo-normalized (partial structure)/normalized (complete structure) structure factors (Giacovazzo, 1983*b*). In each case, it can be shown that the distributions can be obtained from the general distribution (2.8) by proper selection of the scattering-factor expressions.

3. The probabilistic theory of the three-phase structure invariant

For a triplet of reciprocal-lattice vectors, \mathbf{H} , \mathbf{K} , \mathbf{L} , satisfying

$$\mathbf{H} + \mathbf{K} + \mathbf{L} = \mathbf{0} \quad (3.1)$$

we consider the joint probability distribution $P(R_1, R_2, R_3, S_1, S_2, S_3; \Phi_1, \Phi_2, \Phi_3, \Psi_1, \Psi_2, \Psi_3)$ of the magnitudes $|E_H|$, $|E_K|$, $|E_L|$, $|G_H|$, $|G_K|$, $|G_L|$ and the phases φ_H , φ_K , φ_L , ψ_H , ψ_K , ψ_L of the complex normalized structure factors E_H , E_K , E_L , G_H , G_K , G_L . The normalized structure factors follow the definitions given by (2.3) and (2.4) and, in particular, include complex atomic scattering factors. Then by following methods previously described (Hauptman, 1982*b*), we obtain

$$\begin{aligned} P &\approx \frac{R_1 R_2 R_3 S_1 S_2 S_3}{\pi^6 (1 - X_H^2)(1 - X_K^2)(1 - X_L^2)} \\ &\times \exp \left(-\frac{R_1^2 + S_1^2}{1 - X_H^2} - \frac{R_2^2 + S_2^2}{1 - X_K^2} - \frac{R_3^2 + S_3^2}{1 - X_L^2} \right) \\ &\times \exp \left[\frac{2R_1 S_1 X_H}{1 - X_H^2} \cos(\Phi_1 + \Psi_1 + \xi_H) \right. \\ &+ \frac{2R_2 S_2 X_K}{1 - X_K^2} \cos(\Phi_2 + \Psi_2 + \xi_K) \\ &+ \left. \frac{2R_3 S_3 X_L}{1 - X_L^2} \cos(\Phi_3 + \Psi_3 + \xi_L) \right] \\ &\times \exp \left\{ \frac{2}{(1 - X_H^2)(1 - X_K^2)(1 - X_L^2)} \right. \\ &\times [Z_0 R_1 R_2 R_3 \cos(\Phi_1 + \Phi_2 + \Phi_3 - \zeta_0) \\ &+ Z'_0 S_1 S_2 S_3 \cos(\Psi_1 + \Psi_2 + \Psi_3 - \zeta'_0) \\ &+ Z_1 S_1 R_2 R_3 \cos(-\Psi_1 + \Phi_2 + \Phi_3 - \zeta_1) \\ &+ Z'_1 R_1 S_2 S_3 \cos(-\Phi_1 + \Psi_2 + \Psi_3 - \zeta'_1) \\ &+ Z_2 R_1 S_2 R_3 \cos(\Phi_1 - \Psi_2 + \Phi_3 - \zeta_2) \\ &+ Z'_2 S_1 R_2 S_3 \cos(\Psi_1 - \Phi_2 + \Psi_3 - \zeta'_2) \\ &+ Z_3 R_1 R_2 S_3 \cos(\Phi_1 + \Phi_2 - \Psi_3 - \zeta_3) \\ &+ \left. Z'_3 S_1 S_2 R_3 \cos(\Psi_1 + \Psi_2 - \Phi_3 - \zeta'_3) \right] \}. \end{aligned} \quad (3.2)$$

The Z_j , ζ_j and Z'_j , ζ'_j , $j = 0, 1, 2, 3$ are uniquely defined by the equations

$$\begin{aligned} Z_0 \exp(i\zeta_0) &= D_{\mathbf{H}\mathbf{K}\mathbf{L}} - D_{\mathbf{H}'\mathbf{K}'\mathbf{L}'} D_{\mathbf{H}} D_{\mathbf{K}} D_{\mathbf{L}} \\ &- D_{\mathbf{H}\mathbf{K}\mathbf{L}'} D_{\mathbf{L}} + D_{\mathbf{H}'\mathbf{K}'\mathbf{L}} D_{\mathbf{H}} D_{\mathbf{K}} \\ &- D_{\mathbf{H}\mathbf{K}\mathbf{L}'} D_{\mathbf{K}} + D_{\mathbf{H}'\mathbf{K}'\mathbf{L}} D_{\mathbf{H}} D_{\mathbf{L}} \\ &- D_{\mathbf{H}'\mathbf{K}\mathbf{L}} D_{\mathbf{H}} + D_{\mathbf{H}'\mathbf{K}'\mathbf{L}} D_{\mathbf{K}} D_{\mathbf{L}}, \end{aligned} \quad (3.3)$$

$$\begin{aligned} Z'_0 \exp(i\zeta'_0) &= D_{\mathbf{H}'\mathbf{K}'\mathbf{L}'} - D_{\mathbf{H}\mathbf{K}\mathbf{L}} D_{\mathbf{H}} D_{\mathbf{K}} D_{\mathbf{L}} \\ &- D_{\mathbf{H}'\mathbf{K}'\mathbf{L}} D_{\mathbf{L}} + D_{\mathbf{H}\mathbf{K}\mathbf{L}'} D_{\mathbf{H}} D_{\mathbf{K}} \\ &- D_{\mathbf{H}'\mathbf{K}'\mathbf{L}} D_{\mathbf{K}} + D_{\mathbf{H}\mathbf{K}\mathbf{L}'} D_{\mathbf{H}} D_{\mathbf{L}} \\ &- D_{\mathbf{H}'\mathbf{K}\mathbf{L}'} D_{\mathbf{H}} + D_{\mathbf{H}'\mathbf{K}'\mathbf{L}} D_{\mathbf{K}} D_{\mathbf{L}}, \end{aligned} \quad (3.4)$$

$$\begin{aligned}
Z_1 \exp(i\zeta_1) = & -D_{\text{HKL}}D_{\text{H}} + D_{\text{H}'\text{K}'\text{L}}D_{\text{K}}D_{\text{L}} \\
& + D_{\text{HKL}}D_{\text{H}}D_{\text{L}} - D_{\text{H}'\text{K}'\text{L}}D_{\text{K}} \\
& + D_{\text{H}'\text{K}'\text{L}}D_{\text{H}}D_{\text{K}} - D_{\text{H}'\text{K}'\text{L}}D_{\text{L}} \\
& + D_{\text{H}'\text{K}'\text{L}} - D_{\text{H}'\text{K}'\text{L}}D_{\text{H}}D_{\text{K}}D_{\text{L}}, \quad (3.5)
\end{aligned}$$

$$\begin{aligned}
Z'_1 \exp(i\zeta'_1) = & -D_{\text{H}'\text{K}'\text{L}}D_{\text{H}} + D_{\text{HKL}}D_{\text{K}}D_{\text{L}} \\
& + D_{\text{H}'\text{K}'\text{L}}D_{\text{H}}D_{\text{L}} - D_{\text{HKL}}D_{\text{K}} \\
& + D_{\text{H}'\text{K}'\text{L}}D_{\text{H}}D_{\text{K}} - D_{\text{H}'\text{K}'\text{L}}D_{\text{L}} \\
& + D_{\text{H}'\text{K}'\text{L}} - D_{\text{H}'\text{K}'\text{L}}D_{\text{H}}D_{\text{K}}D_{\text{L}}, \quad (3.6)
\end{aligned}$$

$$\begin{aligned}
Z_2 \exp(i\zeta_2) = & -D_{\text{HKL}}D_{\text{K}} + D_{\text{H}'\text{K}'\text{L}}D_{\text{H}}D_{\text{L}} \\
& + D_{\text{HKL}}D_{\text{K}}D_{\text{L}} - D_{\text{H}'\text{K}'\text{L}}D_{\text{H}} \\
& + D_{\text{H}'\text{K}'\text{L}} - D_{\text{H}'\text{K}'\text{L}}D_{\text{H}}D_{\text{K}}D_{\text{L}} \\
& + D_{\text{H}'\text{K}'\text{L}}D_{\text{H}}D_{\text{K}} - D_{\text{H}'\text{K}'\text{L}}D_{\text{L}}, \quad (3.7)
\end{aligned}$$

$$\begin{aligned}
Z'_2 \exp(i\zeta'_2) = & -D_{\text{H}'\text{K}'\text{L}}D_{\text{K}} + D_{\text{HKL}}D_{\text{H}}D_{\text{L}} \\
& + D_{\text{H}'\text{K}'\text{L}}D_{\text{K}}D_{\text{L}} - D_{\text{HKL}}D_{\text{H}} \\
& + D_{\text{H}'\text{K}'\text{L}} - D_{\text{H}'\text{K}'\text{L}}D_{\text{H}}D_{\text{K}}D_{\text{L}} \\
& + D_{\text{H}'\text{K}'\text{L}}D_{\text{H}}D_{\text{K}} - D_{\text{H}'\text{K}'\text{L}}D_{\text{L}}, \quad (3.8)
\end{aligned}$$

$$\begin{aligned}
Z_3 \exp(i\zeta_3) = & -D_{\text{HKL}}D_{\text{L}} + D_{\text{H}'\text{K}'\text{L}}D_{\text{H}}D_{\text{K}} \\
& + D_{\text{H}'\text{K}'\text{L}} - D_{\text{H}'\text{K}'\text{L}}D_{\text{H}}D_{\text{K}}D_{\text{L}} \\
& + D_{\text{H}'\text{K}'\text{L}}D_{\text{K}}D_{\text{L}} - D_{\text{H}'\text{K}'\text{L}}D_{\text{H}} \\
& + D_{\text{H}'\text{K}'\text{L}}D_{\text{H}}D_{\text{L}} - D_{\text{H}'\text{K}'\text{L}}D_{\text{K}},
\end{aligned}$$

$$\begin{aligned}
Z'_3 \exp(i\zeta'_3) = & -D_{\text{H}'\text{K}'\text{L}}D_{\text{L}} + D_{\text{HKL}}D_{\text{H}}D_{\text{K}} \\
& + D_{\text{H}'\text{K}'\text{L}} - D_{\text{H}'\text{K}'\text{L}}D_{\text{H}}D_{\text{K}}D_{\text{L}} \\
& + D_{\text{H}'\text{K}'\text{L}}D_{\text{K}}D_{\text{L}} - D_{\text{H}'\text{K}'\text{L}}D_{\text{H}} \\
& + D_{\text{H}'\text{K}'\text{L}}D_{\text{H}}D_{\text{L}} - D_{\text{H}'\text{K}'\text{L}}D_{\text{K}}, \quad (3.9)
\end{aligned}$$

where

$$\begin{aligned}
D_{\text{HKL}} = & C_{\text{HKL}} + iS_{\text{HKL}} \\
= & (\alpha_{\text{H}}\alpha_{\text{K}}\alpha_{\text{L}})^{-1/2} \sum_{j=1}^N |f_{\text{H}}f_{\text{K}}f_{\text{L}}| \\
& \times \exp[i(\delta_{\text{H}} + \delta_{\text{K}} + \delta_{\text{L}})], \quad (3.10)
\end{aligned}$$

$$\begin{aligned}
D_{\text{H}'\text{K}'\text{L}} = & C_{\text{H}'\text{K}'\text{L}} + iS_{\text{H}'\text{K}'\text{L}} \\
= & (\beta_{\text{H}}\beta_{\text{K}}\beta_{\text{L}})^{-1/2} \sum_{j=1}^N |g_{\text{H}}g_{\text{K}}g_{\text{L}}| \\
& \times \exp[i(\eta_{\text{H}} + \eta_{\text{K}} + \eta_{\text{L}})], \quad (3.11)
\end{aligned}$$

$$\begin{aligned}
D_{\text{H}'\text{KL}} = & C_{\text{H}'\text{KL}} + iS_{\text{H}'\text{KL}} \\
= & (\beta_{\text{H}}\alpha_{\text{K}}\alpha_{\text{L}})^{-1/2} \sum_{j=1}^N |g_{\text{H}}f_{\text{K}}f_{\text{L}}| \\
& \times \exp[i(-\eta_{\text{H}} + \delta_{\text{K}} + \delta_{\text{L}})], \quad (3.12)
\end{aligned}$$

$$\begin{aligned}
D_{\text{HK'L}} = & C_{\text{HK'L}} + iS_{\text{HK'L}} \\
= & (\alpha_{\text{H}}\beta_{\text{K}}\beta_{\text{L}})^{-1/2} \sum_{j=1}^N |f_{\text{H}}g_{\text{K}}g_{\text{L}}| \\
& \times \exp[i(-\delta_{\text{H}} + \eta_{\text{K}} + \eta_{\text{L}})], \quad (3.13)
\end{aligned}$$

$$\begin{aligned}
D_{\text{HK'L}} = & C_{\text{HK'L}} + iS_{\text{HK'L}} \\
= & (\alpha_{\text{H}}\beta_{\text{K}}\alpha_{\text{L}})^{-1/2} \sum_{j=1}^N |f_{\text{H}}g_{\text{K}}f_{\text{L}}| \\
& \times \exp[i(\delta_{\text{H}} - \eta_{\text{K}} + \delta_{\text{L}})], \quad (3.14)
\end{aligned}$$

$$\begin{aligned}
D_{\text{H'KL}} = & C_{\text{H'KL}} + iS_{\text{H'KL}} \\
= & (\beta_{\text{H}}\alpha_{\text{K}}\beta_{\text{L}})^{-1/2} \sum_{j=1}^N |g_{\text{H}}f_{\text{K}}g_{\text{L}}| \\
& \times \exp[i(\eta_{\text{H}} - \delta_{\text{K}} + \eta_{\text{L}})], \quad (3.15)
\end{aligned}$$

$$\begin{aligned}
D_{\text{HKL}} = & C_{\text{HKL}} + iS_{\text{HKL}} \\
= & (\alpha_{\text{H}}\alpha_{\text{K}}\beta_{\text{L}})^{-1/2} \sum_{j=1}^N |f_{\text{H}}f_{\text{K}}g_{\text{L}}| \\
& \times \exp[i(\delta_{\text{H}} + \delta_{\text{K}} - \eta_{\text{L}})], \quad (3.16)
\end{aligned}$$

$$\begin{aligned}
D_{\text{H'K'L}} = & C_{\text{H'K'L}} + iS_{\text{H'K'L}} \\
= & (\beta_{\text{H}}\beta_{\text{K}}\alpha_{\text{L}})^{-1/2} \sum_{j=1}^N |g_{\text{H}}g_{\text{K}}f_{\text{L}}| \\
& \times \exp[i(\eta_{\text{H}} + \eta_{\text{K}} - \delta_{\text{L}})], \quad (3.17)
\end{aligned}$$

$$\begin{aligned}
D_{\text{H}} = & C_{\text{H}} + iS_{\text{H}} \\
= & (\alpha_{\text{H}}\beta_{\text{H}})^{-1/2} \sum_{j=1}^N |f_{\text{H}}g_{\text{H}}| \\
& \times \exp[i(\delta_{\text{H}} + \eta_{\text{H}})]. \quad (3.18)
\end{aligned}$$

D_{K} and D_{L} are defined in a similar way.

X_{H} and ξ_{H} have already been defined [(2.9) and (2.10)]. X_{K} , ξ_{K} , X_{L} and ξ_{L} are defined in a similar way.

Case 1. Native protein/heavy-atom-derivative data sets (neglecting the effect of anomalous scattering)

When effects of anomalous scattering are not included, we have

$$S_1 = S_{\bar{1}}, \quad S_2 = S_{\bar{2}}, \quad S_3 = S_{\bar{3}} \quad (3.19)$$

$$\Psi_1 = -\Psi_{\bar{1}}, \quad \Psi_2 = -\Psi_{\bar{2}}, \quad \Psi_3 = -\Psi_{\bar{3}}. \quad (3.20)$$

The f_j and g_j are real and equal to the atomic numbers Z_j . In the notation of Hauptman (1982a), the D terms, equations (3.10)–(3.20), become

$$D_{\text{HKL}} = \alpha_{30}/\alpha_{20}^{3/2} \quad (3.21)$$

$$D_{\text{H}'\text{K}'\text{L}} = \alpha_{03}/\alpha_{02}^{3/2} \quad (3.22)$$

$$D_{\text{H}'\text{KL}} = D_{\text{HK'L}} = D_{\text{HKL}} = \alpha_{21}/(\alpha_{20}\alpha_{02}^{1/2}) \quad (3.23)$$

$$D_{\text{HK'L}} = D_{\text{H}'\text{K}'\text{L}} = D_{\text{H}'\text{KL}} = \alpha_{12}/(\alpha_{20}^{1/2}\alpha_{02}) \quad (3.24)$$

$$D_{\text{H}} = D_{\text{K}} = D_{\text{L}} = \alpha_{11}/(\alpha_{20}^{1/2}\alpha_{02}^{1/2}) = \alpha \quad (3.25)$$

and the Z , ζ , Z' , ζ' equations (3.3)-(3.9) become

$$Z_0 = \alpha_{20}^{-3/2} \alpha_{02}^{-3} [\alpha_{30} \alpha_{02}^3 - 3\alpha_{21} \alpha_{02}^2 \alpha_{11} + 3\alpha_{12} \alpha_{02} \alpha_{11}^2 - \alpha_{03} \alpha_{11}^3] \quad (3.26)$$

$$Z'_0 = \alpha_{20}^{-3} \alpha_{02}^{-3/2} [\alpha_{03} \alpha_{20}^3 - 3\alpha_{12} \alpha_{20}^2 \alpha_{11} + 3\alpha_{21} \alpha_{20} \alpha_{11}^2 - \alpha_{30} \alpha_{11}^3] \quad (3.27)$$

$$Z_1 = Z_2 = Z_3 = \alpha_{20}^{-2} \alpha_{02}^{-5/2} [(\alpha_{21} \alpha_{20} - \alpha_{30} \alpha_{11}) \alpha_{02}^2 - 2(\alpha_{12} \alpha_{20} - \alpha_{21} \alpha_{11}) \alpha_{02} \alpha_{11} + (\alpha_{03} \alpha_{20} - \alpha_{12} \alpha_{11}) \alpha_{11}^2] \quad (3.28)$$

$$Z'_1 = Z'_2 = Z'_3 = \alpha_{20}^{-5/2} \alpha_{02}^{-2} [(\alpha_{12} \alpha_{02} - \alpha_{03} \alpha_{11}) \alpha_{20}^2 - 2(\alpha_{21} \alpha_{02} - \alpha_{12} \alpha_{11}) \alpha_{20} \alpha_{11} + (\alpha_{30} \alpha_{02} - \alpha_{21} \alpha_{11}) \alpha_{11}^2] \quad (3.29)$$

$$\zeta_0 = \zeta'_0 = \zeta_1 = \zeta'_1 = \zeta_2 = \zeta'_2 = \zeta_3 = \zeta'_3 = 0. \quad (3.30)$$

Comparison of equations (3.34)-(3.37) with equations (3.6)-(3.9) of Hauptman (1982a) shows that

$$Z_0 = \beta_0 (1 - \alpha^2)^3 \quad (3.31)$$

$$Z'_0 = \beta_3 (1 - \alpha^2)^3 \quad (3.32)$$

$$Z_1 = Z_2 = Z_3 = \beta_1 (1 - \alpha^2)^3 \quad (3.33)$$

$$Z'_1 = Z'_2 = Z'_3 = \beta_2 (1 - \alpha^2)^3 \quad (3.34)$$

and, finally, substituting in (3.2) we have

$$P = \frac{R_1 R_2 R_3 S_1 S_2 S_3}{\pi^6 (1 - \alpha^2)^3} \times \exp \{ [-1/(1 - \alpha^2)] \times (R_1^2 + R_2^2 + R_3^2 + S_1^2 + S_2^2 + S_3^2) + [2\alpha/(1 - \alpha^2)] [R_1 S_1 \cos(\Phi_1 - \Psi_1) + R_2 S_2 \cos(\Phi_2 - \Psi_2) + R_3 S_3 \cos(\Phi_3 - \Psi_3)] + 2\beta_0 R_1 R_2 R_3 \cos(\Phi_1 + \Phi_2 + \Phi_3) + 2\beta_1 [R_1 R_2 S_3 \cos(\Phi_1 + \Phi_2 + \Psi_3) + R_1 S_2 R_3 \cos(\Phi_1 + \Psi_2 + \Phi_3) + S_1 R_2 R_3 \cos(\Psi_1 + \Phi_2 + \Phi_3)] + 2\beta_2 [R_1 S_2 S_3 \cos(\Phi_1 + \Psi_2 + \Psi_3) + S_1 R_2 S_3 \cos(\Psi_1 + \Phi_2 + \Psi_3) + S_1 S_2 R_3 \cos(\Psi_1 + \Psi_2 + \Phi_3)] + 2\beta_3 S_1 S_2 S_3 \cos(\Psi_1 + \Psi_2 + \Psi_3) \} \quad (3.35)$$

which is equation (3.4) of Hauptman (1982a).

The distribution presented by Giacovazzo, Cascarano & Zheng Chao-de (1988) [their equation (3)] is also readily obtained from (3.2). Indeed, the only differences between the two distributions are in the

atomic scattering factors, which are real in the former case and complex in the latter.

Case 2. Friedel-pair data of an anomalously scattering crystal structure

In this case we have, using the notation of Hauptman (1982b),

$$R_{\bar{1}} = S_{\bar{1}}, \quad R_{\bar{2}} = S_{\bar{2}}, \quad R_{\bar{3}} = S_{\bar{3}}, \quad (3.36)$$

$$\Phi_{\bar{1}} = \Psi_{\bar{1}}, \quad \Phi_{\bar{2}} = \Psi_{\bar{2}}, \quad \Phi_{\bar{3}} = \Psi_{\bar{3}}, \quad (3.37)$$

$$f_{j\bar{H}} = g_{j\bar{H}}, \quad f_{j\bar{K}} = g_{j\bar{K}}, \quad f_{j\bar{L}} = g_{j\bar{L}}. \quad (3.38)$$

The D terms, equations (3.10)-(3.20), become

$$D_{\text{HKL}} = D_{\text{H}'\text{K}'\text{L}'} = (\alpha_{\text{H}} \alpha_{\text{K}} \alpha_{\text{L}})^{-1/2} \sum_{j=1}^N |f_{j\text{H}} f_{j\text{K}} f_{j\text{L}}| \times \exp [i(\delta_{j\text{H}} + \delta_{j\text{K}} + \delta_{j\text{L}})], \quad (3.39)$$

$$D_{\text{H}'\text{K}'\text{L}'} = D_{\text{HKL}} = (\alpha_{\text{H}} \alpha_{\text{K}} \alpha_{\text{L}})^{-1/2} \sum_{j=1}^N |f_{j\text{H}} f_{j\text{K}} f_{j\text{L}}| \times \exp [i(-\delta_{j\text{H}} + \delta_{j\text{K}} + \delta_{j\text{L}})], \quad (3.40)$$

$$D_{\text{HKL}} = D_{\text{H}'\text{K}'\text{L}'} = (\alpha_{\text{H}} \alpha_{\text{K}} \alpha_{\text{L}})^{-1/2} \sum_{j=1}^N |f_{j\text{H}} f_{j\text{K}} f_{j\text{L}}| \times \exp [i(\delta_{j\text{H}} - \delta_{j\text{K}} + \delta_{j\text{L}})], \quad (3.41)$$

$$D_{\text{HKL}} = D_{\text{H}'\text{K}'\text{L}'} = (\alpha_{\text{H}} \alpha_{\text{K}} \alpha_{\text{L}})^{-1/2} \sum_{j=1}^N |f_{j\text{H}} f_{j\text{K}} f_{j\text{L}}| \times \exp [i(\delta_{j\text{H}} + \delta_{j\text{K}} - \delta_{j\text{L}})], \quad (3.42)$$

$$D_{\text{H}} = \alpha_{\text{H}}^{-1} \sum_{j=1}^N |f_{j\text{H}}|^2 \exp [i(2\delta_{j\text{H}})]. \quad (3.43)$$

D_{K} and D_{L} are defined in a similar way.

The Z , ζ , Z' , ζ' equations (3.3)-(3.9) become

$$Z_0 \exp(i\zeta_0) = Z'_0 \exp(i\zeta'_0) = D_{\text{HKL}} (1 - D_{\text{H}} D_{\text{K}} D_{\text{L}}) + D_{\text{HKL}} (-D_{\text{L}} + D_{\text{H}} D_{\text{K}}) + D_{\text{H}'\text{K}'\text{L}'} (-D_{\text{K}} + D_{\text{H}} D_{\text{L}}) + D_{\text{H}'\text{K}'\text{L}'} (-D_{\text{H}} + D_{\text{K}} D_{\text{L}}) \quad (3.44)$$

$$Z_1 \exp(i\zeta_1) = Z'_1 \exp(i\zeta'_1) = D_{\text{HKL}} (-D_{\text{H}} + D_{\text{K}} D_{\text{L}}) + D_{\text{HKL}} (-D_{\text{K}} + D_{\text{H}} D_{\text{L}}) + D_{\text{H}'\text{K}'\text{L}'} (-D_{\text{L}} + D_{\text{H}} D_{\text{K}}) + D_{\text{H}'\text{K}'\text{L}'} (1 - D_{\text{H}} D_{\text{K}} D_{\text{L}}), \quad (3.45)$$

$$\begin{aligned}
Z_2 \exp(i\zeta_2) &= Z'_2 \exp(i\zeta'_2) \\
&= D_{\text{HKL}}(-D_{\text{K}} + D_{\text{H}}D_{\text{L}}) \\
&\quad + D_{\text{HK'L}}(-D_{\text{H}} + D_{\text{K}}D_{\text{L}}) \\
&\quad + D_{\text{HK'L}}(1 - D_{\text{H}}D_{\text{K}}D_{\text{L}}) \\
&\quad + D_{\text{H'KL}}(D_{\text{L}} - D_{\text{H}}D_{\text{K}}), \quad (3.46)
\end{aligned}$$

$$\begin{aligned}
Z_3 \exp(i\zeta_3) &= Z'_3 \exp(i\zeta'_3) \\
&= D_{\text{HKL}}(-D_{\text{L}} + D_{\text{H}}D_{\text{K}}) \\
&\quad + D_{\text{HK'L}}(1 - D_{\text{H}}D_{\text{K}}D_{\text{L}}) \\
&\quad + D_{\text{HK'L}}(D_{\text{H}} - D_{\text{K}}D_{\text{L}}) \\
&\quad + D_{\text{H'KL}}(D_{\text{K}} - D_{\text{H}}D_{\text{L}}). \quad (3.47)
\end{aligned}$$

Finally, substituting in the general equation (3.2) we obtain

$$\begin{aligned}
P &= \frac{R_1 R_2 R_3 R_{\bar{1}} R_{\bar{2}} R_{\bar{3}}}{\pi^6 (1 - X_{\text{H}}^2)(1 - X_{\text{K}}^2)(1 - X_{\text{L}}^2)} \\
&\quad \times \exp\left(-\frac{R_1^2 + R_{\bar{1}}^2}{1 - X_{\text{H}}^2} - \frac{R_2^2 + R_{\bar{2}}^2}{1 - X_{\text{K}}^2} - \frac{R_3^2 + R_{\bar{3}}^2}{1 - X_{\text{L}}^2}\right) \\
&\quad \times \exp\{[2R_1 R_{\bar{1}} X_{\text{H}}/(1 - X_{\text{H}}^2)] \cos(\Phi_1 + \Phi_{\bar{1}} + \xi_{\text{H}}) \\
&\quad + [2R_2 R_{\bar{2}} X_{\text{K}}/(1 - X_{\text{K}}^2)] \cos(\Phi_2 + \Phi_{\bar{2}} + \xi_{\text{K}}) \\
&\quad + [2R_3 R_{\bar{3}} X_{\text{L}}/(1 - X_{\text{L}}^2)] \cos(\Phi_3 + \Phi_{\bar{3}} + \xi_{\text{L}})\} \\
&\quad \times \exp\{2[(1 - X_{\text{H}}^2)(1 - X_{\text{K}}^2)(1 - X_{\text{L}}^2)]^{-1} \\
&\quad \times Z_0[R_1 R_2 R_3 \cos(\Phi_1 + \Phi_2 + \Phi_3 - \zeta_0) \\
&\quad + R_{\bar{1}} R_{\bar{2}} R_{\bar{3}} \cos(\Phi_{\bar{1}} + \Phi_{\bar{2}} + \Phi_{\bar{3}} - \zeta_0)] \\
&\quad + Z_1[R_{\bar{1}} R_2 R_3 \cos(-\Phi_{\bar{1}} + \Phi_2 + \Phi_3 - \zeta_1) \\
&\quad + R_1 R_{\bar{2}} R_{\bar{3}} \cos(-\Phi_1 + \Phi_{\bar{2}} + \Phi_3 - \zeta_1)] \\
&\quad + Z_2[R_1 R_{\bar{2}} R_3 \cos(\Phi_1 - \Phi_{\bar{2}} + \Phi_3 - \zeta_2) \\
&\quad + R_{\bar{1}} R_2 R_{\bar{3}} \cos(\Phi_{\bar{1}} - \Phi_2 + \Phi_3 - \zeta_2)] \\
&\quad + Z_3[R_1 R_2 R_{\bar{3}} \cos(\Phi_1 + \Phi_2 - \Phi_{\bar{3}} - \zeta_3) \\
&\quad + R_{\bar{1}} R_{\bar{2}} R_3 \cos(\Phi_{\bar{1}} + \Phi_{\bar{2}} - \Phi_3 - \zeta_3)]\}. \quad (3.48)
\end{aligned}$$

The distribution corresponds to equation (I.4) of Hauptman (1982b). Giacovazzo (1983a) has obtained a distribution which is identical except for notation.

Case 3. Calculated data of a partial structure and observed data of its associated complete structure (neglecting the effect of anomalous scattering)

The three-phase invariant joint probability distribution was presented by Giacovazzo (1983b). The formula was derived using normalized structure factors for the complete structure, and pseudo-normalized structure factors for the partial structure. The

pseudo-normalized structure factors are defined by

$$E_{p,\text{h}} = F_{p,\text{h}}[|F_{p,\text{h}}|^2 + \sum_{\mathbf{q}} (\mathbf{h})]^{-1/2} \quad (3.49)$$

$$E_{q,\text{h}} = F_{q,\text{h}}[|F_{q,\text{h}}|^2 + \sum_{\mathbf{q}} (\mathbf{h})]^{-1/2} \quad (3.50)$$

where p = number of atoms in the partial structure unit cell and

$$q = N - p, \quad (3.51)$$

$$\sum_{\mathbf{q}} (\mathbf{h}) = \sum_{j=p+1}^N |f_j|^2(\mathbf{h}). \quad (3.52)$$

When effects of anomalous scattering are not included, then the D terms, equations (3.10)–(3.18), become

$$D_{\text{HKL}} = \sum_{j=1}^N |f_{\text{H}} f_{\text{K}} f_{\text{L}}| / (\alpha_{\text{H}} \alpha_{\text{K}} \alpha_{\text{L}})^{1/2} \quad (3.53)$$

$$D_{\text{H'KL}} = \sum_{j=1}^p |f_{\text{H}} f_{\text{K}} f_{\text{L}}| / (\beta_{\text{H}} \beta_{\text{K}} \beta_{\text{L}})^{1/2} \quad (3.54)$$

$$D_{\text{H'KL}} = \sum_{j=1}^p |f_{\text{H}} f_{\text{K}} f_{\text{L}}| / (\beta_{\text{H}} \alpha_{\text{K}} \alpha_{\text{L}})^{1/2} \quad (3.55)$$

$$D_{\text{HKL}} = \sum_{j=1}^p |f_{\text{H}} f_{\text{K}} f_{\text{L}}| / (\alpha_{\text{H}} \beta_{\text{K}} \beta_{\text{L}})^{1/2} \quad (3.56)$$

$$D_{\text{HKL}} = \sum_{j=1}^p |f_{\text{H}} f_{\text{K}} f_{\text{L}}| / (\alpha_{\text{H}} \beta_{\text{K}} \alpha_{\text{L}})^{1/2} \quad (3.57)$$

$$D_{\text{H'KL}} = \sum_{j=1}^p |f_{\text{H}} f_{\text{K}} f_{\text{L}}| / (\beta_{\text{H}} \alpha_{\text{K}} \beta_{\text{L}})^{1/2} \quad (3.58)$$

$$D_{\text{HKL}} = \sum_{j=1}^p |f_{\text{H}} f_{\text{K}} f_{\text{L}}| / (\alpha_{\text{H}} \alpha_{\text{K}} \beta_{\text{L}})^{1/2} \quad (3.59)$$

$$D_{\text{H'KL}} = \sum_{j=1}^p |f_{\text{H}} f_{\text{K}} f_{\text{L}}| / (\beta_{\text{H}} \beta_{\text{K}} \alpha_{\text{L}})^{1/2} \quad (3.60)$$

$$D_{\text{H}} = X_{\text{H}} = \beta_{\text{H}}^{1/2} / \alpha_{\text{H}}^{1/2}. \quad (3.61)$$

Similar expressions are found for D_{K} and D_{L} .

The α 's and β 's are defined by (2.6)–(2.7). Note that in this case β_{H} , for example, is simply

$$\sum_{j=1}^p |f_{\text{H}}|^2. \quad (3.62)$$

The Z , ζ , Z' , ζ' equations (3.3)–(3.9) become

$$Z_0 = \frac{\sum_{j=1}^N |f_{\text{H}} f_{\text{K}} f_{\text{L}}| - \sum_{j=1}^p |f_{\text{H}} f_{\text{K}} f_{\text{L}}|}{(\alpha_{\text{H}} \alpha_{\text{K}} \alpha_{\text{L}})^{1/2}}, \quad (3.63)$$

$$\begin{aligned}
Z'_0 &\approx -\frac{\sum_{j=1}^N |f_{\text{H}} f_{\text{K}} f_{\text{L}}| - \sum_{j=1}^p |f_{\text{H}} f_{\text{K}} f_{\text{L}}|}{(\alpha_{\text{H}} \alpha_{\text{K}} \alpha_{\text{L}})^{1/2}} \\
&\quad \times (\beta_{\text{H}} \beta_{\text{K}} \beta_{\text{L}})^{1/2} / (\alpha_{\text{H}} \alpha_{\text{K}} \alpha_{\text{L}})^{1/2}, \quad (3.64)
\end{aligned}$$

$$Z_1 = -\frac{\sum_{j=1}^N |f_{jH}f_{jK}f_{jL}| - \sum_{j=1}^p |f_{jH}f_{jK}f_{jL}|}{(\alpha_H\alpha_K\alpha_L)^{1/2}} \beta_H^{1/2}/\alpha_H^{1/2}, \quad (3.65)$$

$$Z'_1 = \frac{\sum_{j=1}^N |f_{jH}f_{jK}f_{jL}| - \sum_{j=1}^p |f_{jH}f_{jK}f_{jL}|}{(\alpha_H\alpha_K\alpha_L)^{1/2}} \times (\beta_K\beta_L)^{1/2}/(\alpha_K\alpha_L)^{1/2}, \quad (3.66)$$

$$Z_2 = \frac{\sum_{j=1}^N |f_{jH}f_{jK}f_{jL}| - \sum_{j=1}^p |f_{jH}f_{jK}f_{jL}|}{(\alpha_H\alpha_K\alpha_L)^{1/2}} \beta_K^{1/2}/\alpha_K^{1/2}, \quad (3.67)$$

$$Z'_2 = \frac{\sum_{j=1}^N |f_{jH}f_{jK}f_{jL}| - \sum_{j=1}^p |f_{jH}f_{jK}f_{jL}|}{(\alpha_H\alpha_K\alpha_L)^{1/2}} \times (\beta_H\beta_L)^{1/2}/(\alpha_H\alpha_L)^{1/2}, \quad (3.68)$$

$$Z_3 = -\frac{\sum_{j=1}^N |f_{jH}f_{jK}f_{jL}| - \sum_{j=1}^p |f_{jH}f_{jK}f_{jL}|}{(\alpha_H\alpha_K\alpha_L)^{1/2}} \beta_L^{1/2}/\alpha_L^{1/2}, \quad (3.69)$$

$$Z'_3 = \frac{\sum_{j=1}^N |f_{jH}f_{jK}f_{jL}| - \sum_{j=1}^p |f_{jH}f_{jK}f_{jL}|}{(\alpha_H\alpha_K\alpha_L)^{1/2}} \times (\beta_H\beta_K)^{1/2}/(\alpha_H\alpha_K)^{1/2}, \quad (3.70)$$

$$\zeta_0 = \zeta'_0 = \zeta_1 = \zeta'_1 = \zeta_2 = \zeta'_2 = \zeta_3 = \zeta'_3 = 0. \quad (3.71)$$

It should be noted that in the case of Z'_0 the expression on the right-hand side is an approximation only. Its validity [see Fortier, Weeks & Hauptman (1984a) and Giacovazzo, Cascarano & Zheng Chao-de (1988)] depends upon: (a) the total number of atoms in the complete structure unit cell; (b) the difference between the number of atoms in the partial and complete structures; and (c) the dependence of the scattering factors on the Bragg angle. The approximation is used here for the sake of comparison with the distribution presented by Giacovazzo (1983b). In order to compare (3.2) with the results of Giacovazzo (1983b) we also introduce pseudo-normalized structure factors for the partial structure,

$$\frac{R_{p1}}{\beta_H^{1/2}/\alpha_H^{1/2}} = S_1, \quad \frac{R_{p2}}{\beta_K^{1/2}/\alpha_K^{1/2}} = S_2, \quad \frac{R_{p3}}{\beta_L^{1/2}/\alpha_L^{1/2}} = S_3 \quad (3.72)$$

and

$$\Phi_{p1} = -\bar{\Psi}_1, \quad \Phi_{p2} = -\bar{\Psi}_2, \quad \Phi_{p3} = -\bar{\Psi}_3; \quad (3.73)$$

and finally, as in Giacovazzo's approach, we use $|F_{p,H}|^2$ rather than the average value $\sum_{j=1}^p f_{jH}^2$ and, similarly, $[|F_{p,H}|^2 + \sum_q (\mathbf{h})]$ rather than $\sum_{j=1}^N f_{jH}^2$.

Equation (3.2) then becomes

$$P \approx \frac{R_1 R_2 R_3 R_{1p} R_{2p} R_{3p}}{\pi^6 b_1 b_2 b_3} \times \exp \left(-\frac{R_1^2 + R_{1p}^2}{b_1} - \frac{R_2^2 + R_{2p}^2}{b_2} - \frac{R_3^2 + R_{3p}^2}{b_3} \right) \times \exp [(2R_1 R_{1p}/b_1) \cos(\Phi_1 - \Phi_{1p}) + (2R_2 R_{2p}/b_2) \cos(\Phi_2 - \Phi_{2p}) + (2R_3 R_{3p}/b_3) \cos(\Phi_3 - \Phi_{3p})] \times \exp \{ (2c/b_1 b_2 b_3) [R_1 R_2 R_3 \cos(\Phi_1 + \Phi_2 + \Phi_3) - R_{1p} R_{2p} R_{3p} \cos(\Phi_{1p} + \Phi_{2p} + \Phi_{3p}) - R_{1p} R_2 R_3 \cos(\Phi_{1p} + \Phi_2 + \Phi_3) + R_1 R_{2p} R_{3p} \cos(\Phi_1 + \Phi_{2p} + \Phi_{3p}) - R_1 R_{2p} R_3 \cos(\Phi_1 + \Phi_{2p} + \Phi_3) + R_{1p} R_2 R_{3p} \cos(\Phi_{1p} + \Phi_2 + \Phi_{3p}) - R_1 R_2 R_{3p} \cos(\Phi_1 + \Phi_2 + \Phi_{3p}) + R_{1p} R_{2p} R_3 \cos(\Phi_{1p} + \Phi_{2p} + \Phi_3)] \}. \quad (3.74)$$

Equation (19) of Giacovazzo (1983b) is readily obtained from this joint distribution. The definition of the parameters c and b_i , $i = 1, 2, 3$, can be found in the *Symbols and Abbreviations* section of Giacovazzo (1983b).

4. Concluding remarks

By characterizing isomorphism in reciprocal space, it becomes clear that SIR (single isomorphous replacement), SAS (single-wavelength anomalous scattering) and partial/complete structure data sets all belong to a more general class, that of isomorphous data sets. We can thus obtain general formulae that are applicable to any of these cases and, indeed, general joint probability distributions are presented here for two and three-phase structure invariants. The similarities among the three cases have been recognized before. For example, Srinivasan & Parthasarathy (1976) clearly stated that the distributions they obtained for a 'pair of crystal structures' could be used for either SIR or partial/complete structure data. Karle (1984, 1985), recognizing the similarity in the mathematical and physical properties of isomorphous-replacement and anomalous-scattering diffraction data, presented formulae that can be used with several sets of isomorphous data, several sets of anomalous-scattering data or a combination of both. Fan Hai-fu, Han Fu-son, Qian Jin-zi & Yao Jia-xing (1984) also presented formulae to treat in a unified manner the problem of phase ambiguities in the SIR and SAS techniques.

Clearly, there are important differences between the SIR, SAS and partial/complete structure cases.

Because of these differences, the conditional probability distributions which will be of most use may vary from one case to the other. Similarly, the phasing procedures which will be most effective may vary from one case to the other.

In this paper we have focused on the similarities among these cases, for the following reasons. In a direct-methods probabilistic approach, the derivation of a joint probability distribution is often a lengthy initial task. As shown here, this analysis need only be done once, if it is formulated in a general way. It is also an easy task to translate a distribution derived for a specific case into more general terms. Consequently, much of the already available theoretical foundation for either SIR, SAS or partial/complete structure data may be reformulated so that it can be used in any example of isomorphous data sets. For example, the joint probability distribution of a triplet of isomorphous data sets (Fortier, Weeks & Hauptman, 1984*b*) can be translated easily into any case of interest, such as the case of a native protein and a single heavy-atom derivative for which Friedel-pair data are available. Thus, as in the algebraic approach presented by Karle (1984, 1985), general formulae can be used on a large variety of combinations of the various cases. Finally, while there is still little experience in the use of direct methods in macromolecular structure determination, much valuable experience, both practical and theoretical, has been gained in the use of direct methods for *ab initio* phasing of isomorphous data sets in small molecules. In particular, much can be learned from the vast

amount of expertise that has been gained in the applications of direct methods to the problem of partial structure expansion in the *DIRDIF* system (Beurskens *et al.*, 1981).

Financial assistance from the Natural Sciences and Engineering Research Council of Canada is gratefully acknowledged.

References

- BEURSKENS, P. T., BOSMAN, W. P., DOESBURG, H. M., GOULD, R. O., VAN DEN HARK, TH. E. M., PRICK, P. A. J., NOORDIK, J. H., BEURSKENS, G. & PARTHASARATHY, V. (1981). Tech. Rep. 1981/2. Crystallography Laboratory, Toernooiveld, 6525ED Nijmegen, The Netherlands.
- BEURSKENS, P. T., PRICK, P. A. J., DOESBURG, H. M. & GOULD, R. O. (1979). *Acta Cryst.* **A35**, 765-772.
- FAN HAI-FU, HAN FU-SON, QIAN JIN-ZI & YAO JIA-XING (1984). *Acta Cryst.* **A40**, 489-495.
- FORTIER, S., WEEKS, C. M. & HAUPTMAN, H. (1984*a*). *Acta Cryst.* **A40**, 544-548.
- FORTIER, S., WEEKS, C. M. & HAUPTMAN, H. (1984*b*). *Acta Cryst.* **A40**, 646-651.
- GIACOVAZZO, C. (1983*a*). *Acta Cryst.* **A39**, 585-592.
- GIACOVAZZO, C. (1983*b*). *Acta Cryst.* **A39**, 685-692.
- GIACOVAZZO, C., CASCARANO, G. & ZHENG CHAO-DE (1988). *Acta Cryst.* **A44**, 45-51.
- GLUSKER, J. P. & TRUEBLOOD, K. N. (1985). *Crystal Structure Analysis: a Primer*. Oxford Univ. Press.
- HAUPTMAN, H. (1982*a*). *Acta Cryst.* **A38**, 289-294.
- HAUPTMAN, H. (1982*b*). *Acta Cryst.* **A38**, 632-641.
- KARLE, J. (1984). *Acta Cryst.* **A40**, 374-379.
- KARLE, J. (1985). *Acta Cryst.* **A41**, 182-189.
- KARLE, J. & HAUPTMAN, H. (1958). *Acta Cryst.* **11**, 264-269.
- SIM, G. A. (1959). *Acta Cryst.* **12**, 813-815.
- SRINIVASAN, R. & PARTHASARATHY, S. (1976). *Some Statistical Applications in X-ray Crystallography*. New York: Pergamon.

Acta Cryst. (1989). **A45**, 254-258

The Probability Distributions of X-ray Intensities in Fiber Diffraction: Largest Likely Values for Fiber Diffraction *R* Factors

BY GERALD STUBBS

Department of Molecular Biology, Vanderbilt University, Nashville, TN 37235, USA

(Received 11 March 1988; accepted 14 September 1988)

Abstract

R factors in fiber diffraction are generally lower than in conventional crystallography, because of the cylindrical averaging of fiber diffraction data. The probability distributions for fiber diffraction intensities, analogous to Wilson's distributions for crystal diffraction intensities, are derived, and from these the largest likely values of *R* are estimated. These values depend on the size and symmetry of the diffracting particle and on the resolution of the analysis, and range from 0.586 for systems for very high symmetry (as in crystal

diffraction) to much lower values for systems of low symmetry.

Introduction

The *R* factor, $R = \sum ||F_{\text{obs}}| - |F_{\text{calc}}|| / \sum |F_{\text{obs}}|$, has been used for many years as an index of the quality of crystallographic structure determinations. It is also widely quoted in descriptions of structures determined by refinement of models against fiber diffraction data, although in fiber diffraction $|F|$ must be replaced by $I^{1/2}$. (In fiber diffraction, $|F|$ is not gen-