

On Integrating Direct Methods and Isomorphous Replacement Techniques. I. A Distribution Function for Quartet Invariants

CARMELO GIACOVAZZO^{a*} AND DRITAN SILIQI^b

^a*Dipartimento Geomineralogico, Università di Bari, Campus Universitario, Via Orabona 4, 70125 Bari, Italy, and*

^b*Laboratory of X-ray Diffraction, Department of Inorganic Chemistry, Faculty of Natural Sciences, Tirana University, Tirana, Albania. E-mail: giacovazzo@area.ba.cnr.it*

(Received 6 April 1995; accepted 30 August 1995)

Abstract

For two isomorphous structures, the joint probability distribution function of seven pairs of structure factors has been derived. The vectorial indices of the reflexions are the basis and the cross vectors of a quartet invariant. The atomic positions are assumed to be the primitive random variables. The characteristic function of the distribution is expanded in a Gram–Charlier series: the distribution of the structure factors is first obtained by a Fourier transform operation and then modified into the exponential form.

1. Notation

The notation is basically the same as that used in the paper by Giacovazzo, Cascarano & Zheng (1988) (GCZ from now on). Some notation changes and new symbols are, however, necessary: they are defined in the text.

2. Introduction

The integration of direct methods with isomorphous replacement techniques was initiated by Hauptman (1982). The joint probability distribution of the triplet invariant was derived when one pair of isomorphous diffraction data is available. In that probabilistic approach, the primitive random variable is the ordered triple $(\mathbf{h}, \mathbf{k}, \mathbf{l})$ of reciprocal vectors, which is assumed to be uniformly distributed over the subset of vectors satisfying the condition

$$\mathbf{h} + \mathbf{k} + \mathbf{l} = 0. \quad (1)$$

Then, the structure factors of the native and the isomorphous derivative are functions of the primitive random variables $\mathbf{h}, \mathbf{k}, \mathbf{l}$, so that they are themselves random variables.

In Hauptman notation, the normalized structure factors of the native and of the derivatives are

$$E_{\mathbf{h}} = |E_{\mathbf{h}}| \exp(i\varphi_{\mathbf{h}}) = (1/\alpha_{20}^{1/2}) \sum_{j=1}^N f_j \exp 2\pi i \mathbf{h} \mathbf{r}_j$$

and

$$G_{\mathbf{h}} = |G_{\mathbf{h}}| \exp(i\psi_{\mathbf{h}}) = (1/\alpha_{02}^{1/2}) \sum_{j=1}^N g_j \exp 2\pi i \mathbf{h} \mathbf{r}_j,$$

respectively, where

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

The main points of the Hauptman paper are:

(a) the joint probability distribution

$$P(\phi_{\mathbf{h}}, \psi_{\mathbf{h}}, |E_{\mathbf{h}}|, |G_{\mathbf{h}}|) \quad (2)$$

was found, from which the conditional distribution

$$P(\phi_{\mathbf{h}} - \psi_{\mathbf{h}} | |E_{\mathbf{h}}|, |G_{\mathbf{h}}|)$$

was derived;

(b) the joint probability distribution function

$$P(\phi_1, \phi_2, \phi_3, \psi_1, \psi_2, \psi_3, R_1, R_2, R_3, S_1, S_2, S_3) \quad (3)$$

was derived, where $\phi_1, \phi_2, \phi_3, \psi_1, \psi_2, \psi_3$ stand for $\phi_{\mathbf{h}}, \phi_{\mathbf{k}}, \phi_{\mathbf{l}}, \psi_{\mathbf{h}}, \psi_{\mathbf{k}}, \psi_{\mathbf{l}}$, respectively, and $R_1, R_2, R_3, S_1, S_2, S_3$ represent $|E_{\mathbf{h}}|, |E_{\mathbf{k}}|, |E_{\mathbf{l}}|, |G_{\mathbf{h}}|, |G_{\mathbf{k}}|, |G_{\mathbf{l}}|$.

The first application of the method to error-free data was successful (Hauptman, Potter & Weeks, 1982). The advantage of the method was clearly outlined by Fortier, Weeks & Hauptman (1984): the accuracy of the distribution depends on the scattering difference between the native protein and the derivative.

The Hauptman approach has been revisited by Giacovazzo, Cascarano & Zheng (1988). The main points of this paper may be described as:

(a) It was shown that the joint probability distributions (2) and (3) can also be obtained by considering the reciprocal vectors as fixed and the atomic positions as the primitive random variables;

(b) Some inadequacies of the Hauptman approach were corrected. In particular, the parameter

$$\alpha = \alpha_{11}/(\alpha_{20}\alpha_{02})^{1/2}$$

was considered a resolution-independent parameter; in Hauptman's paper, its value was calculated *via zero-*

Table 1. *Differences in notation between GCZ and this paper*

GCZ paper	This paper
γ_{123}	γ_{123}
$\gamma_{12\bar{3}}$	$\gamma_{12\bar{3}}$
γ_{135}	$\gamma_{13\bar{2}} \equiv \gamma_{1\bar{2}3}$
γ_{156}	$\gamma_{1\bar{2}\bar{3}}$
γ_{234}	$\gamma_{23\bar{1}} \equiv \gamma_{123}$
γ_{345}	$\gamma_{31\bar{2}} \equiv \gamma_{1\bar{2}3}$
γ_{246}	$\gamma_{21\bar{3}} \equiv \gamma_{123}$
γ_{456}	$\gamma_{1\bar{2}\bar{3}}$
β_0	β_{123}
β_{11}	$\beta_{1\bar{2}3}$
β_{12}	$\beta_{1\bar{2}3}$
β_{13}	$\beta_{1\bar{2}3}$
β_{21}	$\beta_{1\bar{2}3}$
β_{22}	$\beta_{1\bar{2}3}$
β_{23}	$\beta_{1\bar{2}3}$
β_{33}	$\beta_{1\bar{2}3}$

angle scattering factors of f and g . Thus, in (3), no difference is made between α_h , α_k and α_l in spite of the fact that h , k and l can have quite different resolutions. The final formula was

$$\begin{aligned}
 & P(\phi_1, \phi_2, \phi_3, \psi_1, \psi_2, \psi_3, R_1, R_2, R_3, S_1, S_2, S_3) \\
 &= \prod_{i=1}^3 (1/\pi^2) [R_i S_i / (1 - \alpha_i^2)] \exp - \{ [1 / (1 - \alpha_i^2)] \\
 &\quad \times [R_i^2 + S_i^2 - 2\beta_{0i} R_i S_i \cos(\psi_i - \phi_i)] \} \\
 &\quad \times \exp [2\beta_{123} R_1 R_2 R_3 \cos(\phi_1 + \phi_2 + \phi_3) \\
 &\quad + 2\beta_{1\bar{2}3} S_1 R_2 R_3 \cos(\psi_1 + \phi_2 + \phi_3) \\
 &\quad + 2\beta_{1\bar{2}\bar{3}} R_1 S_2 R_3 \cos(\phi_1 + \psi_2 + \phi_3) \\
 &\quad + 2\beta_{12\bar{3}} R_1 R_2 S_3 \cos(\phi_1 + \phi_2 + \psi_3) \\
 &\quad + 2\beta_{1\bar{2}\bar{3}} R_1 S_2 S_3 \cos(\phi_1 + \psi_2 + \psi_3) \\
 &\quad + 2\beta_{123} S_1 R_2 S_3 \cos(\psi_1 + \phi_2 + \psi_3) \\
 &\quad + 2\beta_{1\bar{2}3} S_1 S_2 R_3 \cos(\psi_1 + \psi_2 + \phi_3) \\
 &\quad + 2\beta_{1\bar{2}\bar{3}} S_1 S_2 S_3 \cos(\psi_1 + \psi_2 + \psi_3)]. \quad (4)
 \end{aligned}$$

The reader will notice that we have used a notation slightly different from that employed in GCZ. The correspondence is shown in Table 1.

(c) A formula was derived that holds when derivatives are obtained by addition of heavy atoms. It was shown that the triplet phase

$$\Phi = \phi_h + \phi_k + \phi_l$$

is distributed according to the von Mises function

$$P(\Phi | R_i, S_i, i = 1, 2, 3) = [2\pi I_0(A)]^{-1} \exp(A \cos \Phi),$$

where

$$A = 2[\sigma_3/\sigma_2^{3/2}]_P R_1 R_2 R_3 + 2[\sigma_3/\sigma_2^{3/2}]_H \Delta_1 \Delta_2 \Delta_3,$$

where

$$\Delta = (E'_d - E'_p),$$

$$E'_p = F_p / \Sigma_H^{1/2}, \quad E'_d = F_d / \Sigma_H^{1/2}.$$

E'_p and E'_d are the structure factors of the protein and the derivative, respectively, normalized with respect to the heavy-atom structure.

The Hauptman mathematical approach is based on the preliminary calculation of the characteristic function of the distribution (4). Such a function was assumed to have exponential form. The question now is: if the characteristic function is expanded in a Gram-Charlier series, is some information lost during the subsequent calculations? In particular, can Hauptman's or GCZ's distribution be derived *via* a characteristic function expanded in a Gram-Charlier series?

There are several reasons that make this problem relevant:

(a) The Gram-Charlier expansion of the characteristic function was used (Giacovazzo, 1975, 1976) for estimating quartet invariants. The conclusive formula is different from Hauptman's formula but has equivalent accuracy and proved very useful in a large variety of applications (Altomare, Burla, Cascarano, Giacovazzo & Guagliardi, 1993).

(b) The use of the Gram-Charlier expansion simplified the calculations necessary for the derivation of the quartet formula. Thus, its use could also simplify the calculations necessary to evaluate quartet invariants when isomorphous derivative data are available. The probabilistic theory of quartet invariants when isomorphous data are available has not been settled so far. In order to derive the joint probability distribution of isomorphous data sets, Kiriakidis, Peschar & Schenk (1993) have applied a technique that relies on the use of the so-called single difference between two isomorphous structure factors as a variable, instead of using the individual structure factors separately as done by Hauptman and by GCZ. While this technique succeeded in reproducing triplet distributions, it did not provide accurate probability distributions for quartet invariants (Kiriakidis, 1993). In this paper, we aim at deriving the joint probability distribution of isomorphous structure factors defining the quartet phases. In order to do that, we will first derive the joint probability distribution function

$$P \equiv P(\phi_1, \phi_2, \phi_3, \psi_1, \psi_2, \psi_3, R_1, R_2, R_3, S_1, S_2, S_3)$$

via Gram-Charlier expansion of the characteristic function and we will show that both Hauptman and GCZ distributions may be obtained in this way. Since the method allows remarkable simplifications in the calculations, we will study by the same technique the more complex distribution

$$P(\phi_h, \phi_k, \phi_l, \phi_m, \phi_{h+k}, \phi_{h+l}, \phi_{k+l}, R_h, R_k, R_l, \dots, S_{h+l}, S_{k+l})$$

in order to derive a formula for the estimation of the quartet invariants when derivative data are available. In spite of the simplifications involved by the method, the calculations are rather extensive. For reader usefulness, the key formulas are quoted in Appendices A, B and C. In

the text, we will only note the strictly necessary intermediate results.

3. Triplet estimation via the Gram–Charlier expansion of the characteristic function

Let us denote by

$$C(\nu_1, \nu_2, \nu_3, \mu_1, \mu_2, \mu_3, \rho_1, \rho_2, \rho_3, \gamma_1, \gamma_2, \gamma_3)$$

the characteristic function of the distribution

$$P(\phi_1, \phi_2, \phi_3, \psi_1, \psi_2, \psi_3, R_1, R_2, R_3, S_1, S_2, S_3).$$

$\nu_i, \mu_i, \rho_i, \gamma_i$, for $i = 1, 2, 3$, are the carrying variables associated with ϕ_i, ψ_i, R_i, S_i , $i = 1, 2, 3$, respectively.

If the Gram–Charlier expansion of the characteristic function is used (Giacovazzo, 1980) up to and including triplet relationships, the joint probability distribution P may be written in the form

$$\begin{aligned} P &\simeq (2\pi)^{-12} 2^6 R_1 R_2 R_3 S_1 S_2 S_3 \\ &\times \int_0^\infty \dots \int_0^{2\pi} \int_0^{2\pi} \dots \int_0^{2\pi} \rho_1 \rho_2 \rho_3 \gamma_1 \gamma_2 \gamma_3 \\ &\times \exp\{-i2^{1/2}[\rho_1 R_1 \cos(\phi_1 - \nu_1) \\ &+ \rho_2 R_2 \cos(\phi_2 - \nu_2) + \rho_3 R_3 \cos(\phi_3 - \nu_3) \\ &+ \gamma_1 S_1 \cos(\psi_1 - \mu_1) + \gamma_2 S_2 \cos(\psi_2 - \mu_2) \\ &+ \gamma_3 S_3 \cos(\psi_3 - \mu_3)] - \frac{1}{2}(\rho_1^2 + \rho_2^2 + \rho_3^2 \\ &+ \gamma_1^2 + \gamma_2^2 + \gamma_3^2) - \alpha_1 \rho_1 \gamma_1 \cos(\nu_1 - \mu_1) \\ &- \alpha_2 \rho_2 \gamma_2 \cos(\nu_2 - \mu_2) - \alpha_3 \rho_3 \gamma_3 \cos(\nu_3 - \mu_3)] \\ &\times \{1 - 2^{-1/2} i [\gamma_{123} \rho_1 \rho_2 \rho_3 \cos(\nu_1 + \nu_2 + \nu_3) \\ &+ \gamma_{123} \rho_1 \rho_2 \gamma_3 \cos(\nu_1 + \nu_2 + \mu_3) \\ &+ \gamma_{1\bar{2}3} \rho_1 \gamma_2 \rho_3 \cos(\nu_1 + \mu_2 + \nu_3) \\ &+ \gamma_{1\bar{2}\bar{3}} \rho_1 \gamma_2 \gamma_3 \cos(\nu_1 + \mu_2 + \mu_3) \\ &+ \gamma_{123} \gamma_1 \rho_2 \rho_3 \cos(\mu_1 + \nu_2 + \nu_3) \\ &+ \gamma_{1\bar{2}3} \gamma_1 \gamma_2 \rho_3 \cos(\mu_1 + \nu_2 + \nu_3) \\ &+ \gamma_{1\bar{2}\bar{3}} \gamma_1 \gamma_2 \rho_3 \cos(\mu_1 + \nu_2 + \mu_3) \\ &+ \gamma_{1\bar{2}\bar{3}} \gamma_1 \gamma_2 \gamma_3 \cos(\mu_1 + \mu_2 + \mu_3)]\} \\ &\times d\rho_1 \dots d\gamma_1 \dots d\nu_1 \dots d\mu_3. \end{aligned} \quad (5)$$

We have used in (5) a notation slightly different from that used in the GCZ paper. Differences are shown in Table 1.

We first integrate the zero-order term in (5). The result is (see Appendices A and B)

$$\begin{aligned} &\prod_{i=1}^3 (1/\pi^2) [R_i S_i / (1 - \alpha_i^2)] \exp\{-[1/(1 - \alpha_i^2)] \\ &\times [R_i^2 + S_i^2 - 2\alpha_i R_i S_i \cos(\psi_i - \phi_i)]\}, \end{aligned}$$

which coincides with (4) when the triplet contributions are excluded.

Let us now calculate the contribution to (5) arising from the first triplet term in (5), that is from

$$-2^{-1/2} \gamma_{123} \rho_1 \rho_2 \rho_3 \cos(\nu_1 + \nu_2 + \nu_3).$$

In accordance with Appendix C, we obtain the contribution

$$\begin{aligned} &\prod_{i=1}^3 (1/\pi^2) [R_i S_i / (1 - \alpha_i^2)] \exp\{-[1/(1 - \alpha_i^2)] \\ &\times [R_i^2 + S_i^2 - 2\alpha_i R_i S_i \cos(\psi_i - \phi_i)]\} \\ &\times 2\gamma_{123} \left[\prod_{i=1}^3 (1 - \alpha_i^2) \right]^{-1} [R_1 R_2 R_3 \cos(\phi_1 + \phi_2 + \phi_3) \\ &- \alpha_1 G_1 R_2 R_3 \cos(\psi_1 + \phi_2 + \phi_3) \\ &- \alpha_2 R_1 G_2 R_3 \cos(\phi_1 + \psi_2 + \phi_3) \\ &- \alpha_3 R_1 R_2 G_3 \cos(\phi_1 + \phi_2 + \psi_3) \\ &+ \alpha_1 \alpha_2 G_1 G_2 R_3 \cos(\psi_1 + \psi_2 + \phi_3) \\ &+ \alpha_1 \alpha_3 G_1 R_2 G_3 \cos(\psi_1 + \phi_2 + \psi_3) \\ &+ \alpha_2 \alpha_3 R_1 G_2 G_3 \cos(\phi_1 + \psi_2 + \psi_3) \\ &- \alpha_1 \alpha_2 \alpha_3 G_1 G_2 G_3 \cos(\psi_1 + \psi_2 + \psi_3)]. \end{aligned} \quad (6)$$

If the calculation is repeated for all the triplet terms in (5), one obtains

$$\begin{aligned} P &\simeq \prod_{i=1}^3 (1/\pi^2) [R_i S_i / (1 - \alpha_i^2)] \exp\{-[1/(1 - \alpha_i^2)] \\ &\times [R_i^2 + S_i^2 - 2\alpha_i R_i S_i \cos(\psi_i - \phi_i)]\} \\ &\times [1 + 2\beta_{123} R_1 R_2 R_3 \cos(\phi_1 + \phi_2 + \phi_3) \\ &+ 2\beta_{1\bar{2}3} S_1 R_2 R_3 \cos(\psi_1 + \phi_2 + \phi_3) \\ &+ 2\beta_{1\bar{2}\bar{3}} R_1 S_2 R_3 \cos(\phi_1 + \psi_2 + \phi_3) \\ &+ 2\beta_{1\bar{2}\bar{3}} R_1 R_2 S_3 \cos(\phi_1 + \phi_2 + \psi_3) \\ &+ 2\beta_{1\bar{2}\bar{3}} R_1 S_2 S_3 \cos(\phi_1 + \psi_2 + \psi_3) \\ &+ 2\beta_{\bar{1}23} S_1 R_2 S_3 \cos(\psi_1 + \phi_2 + \psi_3) \\ &+ 2\beta_{\bar{1}2\bar{3}} S_1 S_2 R_3 \cos(\psi_1 + \psi_2 + \phi_3) \\ &+ 2\beta_{\bar{1}\bar{2}\bar{3}} S_1 S_2 S_3 \cos(\psi_1 + \psi_2 + \psi_3)]. \end{aligned} \quad (7)$$

The transformation

$$(1 + x) \simeq \exp x$$

makes (7) equal to (4).

It is therefore verified that the use of the Gram–Charlier expansion of the characteristic function provides, for triplet invariants, the same results derived via the exponential form of the characteristic function. The key to the method, the truncated series expansion $\exp x \rightarrow (1 + x)$ of the characteristic function and the back-transformation $(1 + x) \rightarrow \exp x$ for the distribution function, seems to disturb the identification of the probability distribution minimally. We can therefore expect that the method may be usefully applied to quartet invariant distributions.

4. The characteristic function of the joint probability distribution function of seven isomorphous pairs of structure factors

Let

$$C(v_1, \dots, v_7, \mu_1, \dots, \mu_7, \rho_1, \dots, \rho_7, \gamma_1, \dots, \gamma_7)$$

be the characteristic function of the distribution

$$P_7 \equiv P(\phi_1, \dots, \phi_7, \psi_1, \dots, \psi_7, R_1, \dots, R_7, S_1, \dots, S_7),$$

where

$$\begin{aligned} E_1 &= R_1 \exp(i\phi_1) = R_h \exp(i\phi_h), \\ E_2 &= R_2 \exp(i\phi_2) = R_k \exp(i\phi_k), \\ E_3 &= R_3 \exp(i\phi_3) = R_l \exp(i\phi_l), \\ E_4 &= R_4 \exp(i\phi_4) = R_m \exp(i\phi_m), \\ E_5 &= R_5 \exp(i\phi_5) = R_{h+k} \exp(i\phi_{h+k}), \\ E_6 &= R_6 \exp(i\phi_6) = R_{h+l} \exp(i\phi_{h+l}), \\ E_7 &= R_7 \exp(i\phi_7) = R_{k+l} \exp(i\phi_{k+l}), \\ G_1 &= S_1 \exp(i\psi_1) = S_h \exp(i\psi_h), \\ G_2 &= S_2 \exp(i\psi_2) = S_k \exp(i\psi_k), \\ G_3 &= S_3 \exp(i\psi_3) = S_l \exp(i\psi_l), \\ G_4 &= S_4 \exp(i\psi_4) = S_m \exp(i\psi_m), \\ G_5 &= S_5 \exp(i\psi_5) = S_{h+k} \exp(i\psi_{h+k}), \\ G_6 &= S_6 \exp(i\psi_6) = S_{h+l} \exp(i\psi_{h+l}), \\ G_7 &= S_7 \exp(i\psi_7) = S_{k+l} \exp(i\psi_{k+l}), \end{aligned}$$

$v_i, \mu_i, \rho_i, \gamma_i$, for $i = 1, \dots, 7$, are the carrying variables associated with ϕ_i, ψ_i, R_i, S_i for $i = 1, \dots, 7$, respectively, and

$$h + k + l + m = 0.$$

The characteristic function, expanded in a Gram-Charlier series, may be written as

$$\begin{aligned} P_7 &= \int_0^\infty \dots \int_0^\infty \int_0^{2\pi} \dots \int_0^{2\pi} \prod_{i=1}^{14} \{ (1/2\pi^2) R_i S_i \rho_i \gamma_i \\ &\times \exp[-\frac{1}{2}(\rho_i^2 + \gamma_i^2) - i2^{1/2} \rho_i R_i \cos(\phi_i - v_i) \\ &- i2^{1/2} \gamma_i S_i \cos(\psi_i - \mu_i) - \alpha_i \rho_i \gamma_i \cos(v_i - \mu_i)] \} \\ &\times \{ 1 - 2^{-1/2} i [\gamma_{125} \rho_1 \rho_2 \rho_5 \cos(v_1 + v_2 - v_5) + \circlearrowleft \\ &+ \gamma_{345} \rho_3 \rho_4 \rho_5 \cos(v_3 + v_4 + v_5) + \circlearrowleft \\ &+ \gamma_{136} \rho_1 \rho_3 \rho_6 \cos(v_1 + v_3 - v_6) + \circlearrowleft \\ &+ \gamma_{246} \rho_2 \rho_4 \rho_6 \cos(v_2 + v_4 + v_6) + \circlearrowleft \\ &+ \gamma_{147} \rho_1 \rho_4 \rho_7 \cos(v_1 + v_4 + v_7) + \circlearrowleft \\ &+ \gamma_{237} \rho_2 \rho_3 \rho_7 \cos(v_2 + v_3 - v_7) + \circlearrowleft \\ &- \frac{1}{4} \gamma_{125}^2 \rho_1^2 \rho_2^2 \rho_5^2 \cos^2(v_1 + v_2 - v_5) + \circlearrowleft \\ &- \frac{1}{4} \gamma_{345}^2 \rho_3^2 \rho_4^2 \rho_5^2 \cos^2(v_3 + v_4 + v_5) + \circlearrowleft \\ &- \frac{1}{4} \gamma_{136}^2 \rho_1^2 \rho_3^2 \rho_6^2 \cos^2(v_1 + v_3 - v_6) + \circlearrowleft \} \end{aligned}$$

$$\begin{aligned} &- \frac{1}{4} \gamma_{246}^2 \rho_2^2 \rho_4^2 \rho_6^2 \cos^2(v_2 + v_4 + v_6) + \circlearrowleft \\ &- \frac{1}{4} \gamma_{147}^2 \rho_1^2 \rho_4^2 \rho_7^2 \cos^2(v_1 + v_4 + v_7) + \circlearrowleft \\ &- \frac{1}{4} \gamma_{237}^2 \rho_2^2 \rho_3^2 \rho_7^2 \cos^2(v_2 + v_3 - v_7) + \circlearrowleft \\ &+ \gamma_{1234} \rho_1 \rho_2 \rho_3 \rho_4 \cos(v_1 + v_2 + v_3 + v_4) + \circlearrowleft \\ &- \frac{1}{2} \gamma_{125} \gamma_{345} \rho_1 \rho_2 \rho_3 \rho_4 \rho_5^2 \cos(v_1 + v_2 + v_3 + v_4) + \circlearrowleft \\ &- \frac{1}{2} \gamma_{136} \gamma_{246} \rho_1 \rho_2 \rho_3 \rho_4 \rho_6^2 \cos(v_1 + v_2 + v_3 + v_4) + \circlearrowleft \\ &- \frac{1}{2} \gamma_{147} \gamma_{237} \rho_1 \rho_2 \rho_3 \rho_4 \rho_7^2 \cos(v_1 + v_2 + v_3 + v_4) + \circlearrowleft \\ &+ \dots \}. \end{aligned} \tag{8}$$

The number of terms in the distribution (8) is extremely large. We have quoted only those that significantly contribute to the estimation of

$$\Phi = \phi_h + \phi_k + \phi_l + \phi_m.$$

We note the following.

(i) We have used a curved arrow to represent the 'cyclic terms' of a prototype term [only the prototypes are quoted in (8)]. For example:

(a) The complete set of cyclic terms for $\gamma_{125} \rho_1 \rho_2 \rho_5 \cos(v_1 + v_2 - v_5)$ (the prototype included) contains [see the distribution (5)] the eight terms

$$\begin{aligned} &\gamma_{125} \rho_1 \rho_2 \rho_5 \cos(v_1 + v_2 - v_5), \quad \gamma_{12\bar{5}} \rho_1 \rho_2 \rho_5 \cos(v_1 + v_2 - \mu_5), \\ &\gamma_{1\bar{2}5} \rho_1 \gamma_2 \rho_5 \cos(v_1 + \mu_2 - v_5), \quad \gamma_{1\bar{2}\bar{5}} \rho_1 \gamma_2 \rho_5 \cos(v_1 + \mu_2 - \mu_5), \\ &\gamma_{125} \gamma_1 \rho_2 \rho_5 \cos(\mu_1 + v_2 - v_5), \quad \gamma_{1\bar{2}5} \gamma_1 \rho_2 \rho_5 \cos(\mu_1 + \mu_2 - v_5), \\ &\gamma_{1\bar{2}\bar{5}} \gamma_1 \rho_2 \rho_5 \cos(\mu_1 + v_2 - \mu_5), \quad \gamma_{1\bar{2}\bar{5}} \gamma_1 \gamma_2 \rho_5 \cos(\mu_1 + \mu_2 - \mu_5). \end{aligned}$$

In all, the distribution involves 48 different triplet terms and we quote in (8) only the six prototypes.

(b) The cyclic terms of $\gamma_{1234} \rho_1 \rho_2 \rho_3 \rho_4 \cos(v_1 + v_2 + v_3 + v_4)$ (the prototype included) are the 16 terms

$$\begin{aligned} &\gamma_{1234} \rho_1 \rho_2 \rho_3 \rho_4 \cos(v_1 + v_2 + v_3 + v_4), \\ &\gamma_{123\bar{4}} \rho_1 \rho_2 \rho_3 \rho_4 \cos(v_1 + v_2 + v_3 + \mu_4), \\ &\gamma_{1\bar{2}34} \rho_1 \rho_2 \rho_3 \rho_4 \cos(v_1 + v_2 + \mu_3 + \mu_4), \dots, \\ &\gamma_{1\bar{2}\bar{3}\bar{4}} \rho_1 \rho_2 \rho_3 \rho_4 \cos(\mu_1 + \mu_2 + \mu_3 + \mu_4). \end{aligned} \tag{9}$$

We have not quoted in (8) quartets like

$$\begin{aligned} &\gamma_{1267} \rho_1 \rho_2 \rho_6 \rho_7 \cos(v_1 - v_2 - v_6 + v_7) + \circlearrowleft \\ &\gamma_{1357} \rho_1 \rho_3 \rho_5 \rho_7 \cos(v_1 - v_3 - v_5 + v_7) + \circlearrowleft \\ &\gamma_{2356} \rho_2 \rho_3 \rho_5 \rho_6 \cos(v_2 - v_3 - v_5 + v_6) + \circlearrowleft \\ &\gamma_{1456} \rho_1 \rho_4 \rho_5 \rho_6 \cos(v_1 - v_4 - v_5 - v_6) + \circlearrowleft \\ &\gamma_{2457} \rho_2 \rho_4 \rho_5 \rho_7 \cos(v_2 - v_4 - v_5 - v_7) + \circlearrowleft \\ &\gamma_{3467} \rho_3 \rho_4 \rho_6 \rho_7 \cos(v_3 - v_4 - v_6 - v_7) + \circlearrowleft. \end{aligned}$$

It was shown by Giacovazzo (1975, 1976) that their contribution for the estimation of

$$\Phi = \phi_h + \phi_k + \phi_l + \phi_m$$

is of order higher than that arising from the quartets (9).

(c) There are (16×4) cyclic terms of the prototype $\gamma_{125}\gamma_{345}\rho_1\rho_2\rho_3\rho_4\rho_5^2 \cos(\nu_1 + \nu_2 + \nu_3 + \nu_4)$, which may be obtained by permuting on $\rho_1\rho_2\rho_3\rho_4$ as in (9) [16 cases] and associating to each of them the permutations on ρ_5^2 . For example, if we select the first of the 16 permutations in (9), we must consider the following terms:

$$\begin{aligned} &\gamma_{125}\gamma_{345}\rho_1\rho_2\rho_3\rho_4\rho_5^2 \cos(\nu_1 + \nu_2 + \nu_3 + \nu_4), \\ &\gamma_{125}\gamma_{345}\rho_1\rho_2\rho_3\rho_4\gamma_5^2 \cos(\nu_1 + \nu_2 + \nu_3 + \nu_4), \quad (10a) \\ &\gamma_{125}\gamma_{345}\rho_1\rho_2\rho_3\rho_4\rho_5\gamma_5 \cos(\nu_1 + \nu_2 + \nu_3 + \nu_4 + \nu_5 - \mu_5), \\ &\gamma_{125}\gamma_{345}\rho_1\rho_2\rho_3\rho_4\rho_5\gamma_5 \cos(\nu_1 + \nu_2 + \nu_3 + \nu_4 - \nu_5 + \mu_5). \end{aligned}$$

In the same way, for the permutation $\rho_1\rho_2\rho_3\rho_4$ in (9), we must consider the terms

$$\begin{aligned} &\gamma_{125}\gamma_{345}\rho_1\rho_2\rho_3\rho_4\rho_5^2 \cos(\nu_1 + \nu_2 + \nu_3 + \mu_4), \\ &\gamma_{125}\gamma_{345}\rho_1\rho_2\rho_3\rho_4\gamma_5^2 \cos(\nu_1 + \nu_2 + \nu_3 + \mu_4), \\ &\gamma_{125}\gamma_{345}\rho_1\rho_2\rho_3\rho_4\rho_5\gamma_5 \cos(\nu_1 + \nu_2 + \nu_3 + \mu_4 + \nu_5 - \mu_5), \\ &\gamma_{125}\gamma_{345}\rho_1\rho_2\rho_3\rho_4\rho_5\gamma_5 \cos(\nu_1 + \nu_2 + \nu_3 + \mu_4 - \nu_5 + \mu_5). \end{aligned}$$

We have not quoted in (8) (see again Giacovazzo, 1975, 1976) terms like

$$\begin{aligned} &\gamma_{136}\gamma_{237}\rho_1\rho_2\rho_6\rho_7\rho_3^2 \cos(\nu_1 - \nu_2 - \nu_6 + \nu_7) + \odot \\ &\gamma_{125}\gamma_{237}\rho_1\rho_2\rho_5\rho_7\rho_2^2 \cos(\nu_1 - \nu_3 - \nu_5 + \nu_7) + \odot \quad (10b) \end{aligned}$$

since they provide a contribution of a higher order (for the estimation of Φ) than that arising from the quartets (9). In conclusion, the total number of quartet terms in (8) that will be involved in the next calculations is $16 + 3(16 \times 4) = 208$.

(ii) γ_{ijl} are defined in the GCZ paper; γ_{1234} and cyclic parameters are defined as follows:

$$\begin{aligned} \gamma_{1234} &= \left(\sum_1 \sum_2 \sum_3 \sum_4 \right)^{-1/2} \sum_{j=1}^N f_j(\mathbf{h})f_j(\mathbf{k})f_j(\mathbf{l})f_j(\mathbf{m}) \\ \gamma_{1234} &= \left(\sum_1 \sum_2 \sum_3 \sum_4 \right)^{-1/2} \sum_{j=1}^N f_j(\mathbf{h})f_j(\mathbf{k})f_j(\mathbf{l})g_j(\mathbf{m}) \\ &\vdots \\ &\vdots \\ \gamma_{1234} &= \left(\sum_1 \sum_2 \sum_3 \sum_4 \right)^{-1/2} \sum_{j=1}^N g_j(\mathbf{h})g_j(\mathbf{k})g_j(\mathbf{l})g_j(\mathbf{m}). \end{aligned}$$

5. The contribution to P_7 of some prototype terms

There are too many terms in (8) to register the contribution of each single term. We therefore decided to calculate the contribution of the prototype terms and then to derive by a symmetry rule all the other useful contributions.

We first integrate the term of order zero in the Gram–Charlier expansion (8). We have (see Appendices A and B)

$$\begin{aligned} P_0 &\simeq \prod_{i=1}^7 (1/\pi^2)[R_i S_i / (1 - \alpha_i^2)] \exp\{-[1/(1 - \alpha_i^2)] \\ &\quad \times [R_i^2 + S_i^2 - 2\alpha_i R_i S_i \cos(\psi_i - \phi_i)]\}. \quad (11) \end{aligned}$$

The contribution to P_7 of the prototype quartet term $\gamma_{1234}\rho_1\rho_2\rho_3\rho_4 \cos(\nu_1 + \nu_2 + \nu_3 + \nu_4)$ is

$$P_0 \prod_{i=1}^4 [1 - \alpha_i^2]^{-1} \gamma_{1234} Z_{1234}, \quad (12)$$

where

$$\begin{aligned} Z_{1234} &= 16[R_1 R_2 R_3 R_4 \cos(\phi_1 + \phi_2 + \phi_3 + \phi_4) \\ &\quad - \alpha_1 S_1 R_2 R_3 R_4 \cos(\psi_1 + \phi_2 + \phi_3 + \phi_4) \\ &\quad - \alpha_2 R_1 S_2 R_3 R_4 \cos(\phi_1 + \psi_2 + \phi_3 + \phi_4) \\ &\quad - \alpha_3 R_1 R_2 S_3 R_4 \cos(\phi_1 + \phi_2 + \psi_3 + \phi_4) \\ &\quad - \alpha_4 R_1 R_2 R_3 S_4 \cos(\phi_1 + \phi_2 + \phi_3 + \psi_4) \\ &\quad + \alpha_1 \alpha_2 S_1 S_2 R_3 R_4 \cos(\psi_1 + \psi_2 + \phi_3 + \phi_4) \\ &\quad + \alpha_1 \alpha_3 S_1 R_2 S_3 R_4 \cos(\psi_1 + \phi_2 + \psi_3 + \phi_4) \\ &\quad + \alpha_1 \alpha_4 S_1 R_2 R_3 S_4 \cos(\psi_1 + \phi_2 + \phi_3 + \psi_4) \\ &\quad + \alpha_2 \alpha_3 R_1 S_2 S_3 R_4 \cos(\phi_1 + \psi_2 + \psi_3 + \phi_4) \\ &\quad + \alpha_2 \alpha_4 R_1 S_2 S_3 R_4 \cos(\phi_1 + \psi_2 + \psi_3 + \phi_4) \\ &\quad + \alpha_3 \alpha_4 R_1 R_2 S_3 S_4 \cos(\phi_1 + \phi_2 + \psi_3 + \psi_4) \\ &\quad - \alpha_1 \alpha_2 \alpha_3 S_1 S_2 S_3 R_4 \cos(\psi_1 + \psi_2 + \psi_3 + \phi_4) \\ &\quad - \alpha_1 \alpha_2 \alpha_4 S_1 S_2 R_3 S_4 \cos(\psi_1 + \psi_2 + \phi_3 + \psi_4) \\ &\quad - \alpha_1 \alpha_3 \alpha_4 S_1 R_2 S_3 S_4 \cos(\psi_1 + \phi_2 + \psi_3 + \psi_4) \\ &\quad - \alpha_2 \alpha_3 \alpha_4 R_1 S_2 S_3 S_4 \cos(\phi_1 + \psi_2 + \psi_3 + \psi_4) \\ &\quad + \alpha_1 \alpha_2 \alpha_3 \alpha_4 S_1 S_2 S_3 S_4 \cos(\psi_1 + \psi_2 + \psi_3 + \psi_4)]. \end{aligned}$$

The contribution to P_7 of the prototype quartet term

$$-\left(\frac{1}{2}\right)\gamma_{125}\gamma_{345}\rho_1\rho_2\rho_3\rho_4\rho_5^2 \cos(\nu_1 + \nu_2 + \nu_3 + \nu_4)$$

is

$$\begin{aligned} P_0 &\left[\prod_{i=1}^5 (1 - \alpha_i^2) \right]^{-1} \gamma_{125}\gamma_{345} Z_{1234} \{ [1/(1 - \alpha_5^2)] \\ &\quad \times [R_5^2 + \alpha_5^2 S_5^2 - 2\alpha_5 R_5 S_5 \cos(\phi_5 - \psi_5)] - 1 \}. \quad (13) \end{aligned}$$

The contribution to P_7 of the prototype quartet term

$$-\left(\frac{1}{2}\right)\gamma_{125}\gamma_{345}\rho_1\rho_2\rho_3\rho_4\gamma_5^2 \cos(\nu_1 + \nu_2 + \nu_3 + \nu_4)$$

is

$$\begin{aligned} P_0 &\left[\prod_{i=1}^5 (1 - \alpha_i^2) \right]^{-1} \gamma_{125}\gamma_{345} Z_{1234} \{ [1/(1 - \alpha_5^2)] \\ &\quad \times [S_5^2 + \alpha_5^2 R_5^2 - 2\alpha_5 R_5 S_5 \cos(\phi_5 - \psi_5)] - 1 \}. \quad (14) \end{aligned}$$

The contribution to P_7 of the prototype quartet term

$$-\left(\frac{1}{2}\right)\gamma_{125}\gamma_{345}\rho_1\rho_2\rho_3\rho_4\rho_5\gamma_5 \cos(\nu_1+\nu_2+\nu_3+\nu_4+\nu_5-\mu_5)$$

is

$$P_0 \left[\prod_{i=1}^5 (1 - \alpha_i^2) \right]^{-1} \gamma_{125}\gamma_{345} (R_5 S_5 Z_{1234}^{\text{mod}(1)} - \alpha_5 S_5^2 Z_{1234} - \alpha_5 \{ [1/(1 - \alpha_5^2)] [R_5^2 + \alpha_5^2 S_5^2 - 2\alpha_5 R_5 S_5 \cos(\phi_5 - \psi_5)] - 1 \} Z_{1234}), \quad (15)$$

where

$$Z_{1234}^{\text{mod}(1)} = 16[R_1 R_2 R_3 R_4 \cos(\phi_1 + \phi_2 + \phi_3 + \phi_4 + \psi_5 - \phi_5) - \alpha_1 S_1 R_2 R_3 R_4 \cos(\psi_1 + \phi_2 + \phi_3 + \phi_4 + \psi_5 - \phi_5) - \alpha_2 R_1 S_2 R_3 R_4 \cos(\phi_1 + \psi_2 + \phi_3 + \phi_4 + \psi_5 - \phi_5) + \dots + \alpha_1 \alpha_2 \alpha_3 \alpha_4 S_1 S_2 S_3 S_4 \times \cos(\psi_1 + \psi_2 + \psi_3 + \psi_4 + \psi_5 - \phi_5)].$$

The contribution to P_7 of the prototype quartet term

$$-\left(\frac{1}{2}\right)\gamma_{125}\gamma_{345}\rho_1\rho_2\rho_3\rho_4\rho_5\gamma_5 \cos(\nu_1+\nu_2+\nu_3+\nu_4-\nu_5+\mu_5)$$

is

$$P_0 \left[\prod_{i=1}^5 (1 - \alpha_i^2) \right]^{-1} \gamma_{125}\gamma_{345} (R_5 S_5 Z_{1234}^{\text{mod}(2)} - \alpha_5 S_5^2 Z_{1234} - \alpha_5 \{ [1/(1 - \alpha_5^2)] [R_5^2 + \alpha_5^2 S_5^2 - 2\alpha_5 R_5 S_5 \cos(\phi_5 - \psi_5)] - 1 \} Z_{1234}), \quad (16)$$

where

$$Z_{1234}^{\text{mod}(2)} = 16[R_1 R_2 R_3 R_4 \cos(\phi_1 + \phi_2 + \phi_3 + \phi_4 + \phi_5 - \psi_5) - \alpha_1 S_1 R_2 R_3 R_4 \cos(\psi_1 + \phi_2 + \phi_3 + \phi_4 + \phi_5 - \psi_5) - \alpha_2 R_1 S_2 R_3 R_4 \cos(\phi_1 + \psi_2 + \phi_3 + \phi_4 + \phi_5 - \psi_5) + \dots + \alpha_1 \alpha_2 \alpha_3 \alpha_4 S_1 S_2 S_3 S_4 \times \cos(\psi_1 + \psi_2 + \psi_3 + \psi_4 + \phi_5 - \psi_5)].$$

Let us now collect together the contributions arising from the following terms:

$$\gamma_{1234}\rho_1\rho_2\rho_3\rho_4 \cos(\nu_1+\nu_2+\nu_3+\nu_4)$$

and, for $i = 5, 6, 7$,

$$\begin{aligned} &-\frac{1}{2}\gamma_{12i}\gamma_{34i}\rho_1\rho_2\rho_3\rho_4\rho_i^2 \cos(\nu_1+\nu_2+\nu_3+\nu_4), \\ &-\frac{1}{2}\gamma_{12i}\gamma_{34i}\rho_1\rho_2\rho_3\rho_4\gamma_i^2 \cos(\nu_1+\nu_2+\nu_3+\nu_4), \\ &-\frac{1}{2}\gamma_{12i}\gamma_{34i}\rho_1\rho_2\rho_3\rho_4\rho_i\gamma_i \cos(\nu_1+\nu_2+\nu_3+\nu_4+\nu_i-\mu_i), \\ &-\frac{1}{2}\gamma_{12i}\gamma_{34i}\rho_1\rho_2\rho_3\rho_4\rho_i\gamma_i \cos(\nu_1+\nu_2+\nu_3+\nu_4+\mu_i-\nu_i). \end{aligned}$$

From (12)–(16), we find that the total contribution of the above terms is

$$P_0 \left[\prod_{i=1}^4 (1 - \alpha_i^2) \right]^{-1} \left\{ Z_{1234}\gamma_{1234} + \sum_{i=5}^7 Z_{1234}(1 - \alpha_i^2)^{-1} \times [\gamma_{12i}\gamma_{34i}L_i + \gamma_{12i}\gamma_{34i}L_i + (\gamma_{12i}\gamma_{34i} + \gamma_{12i}\gamma_{34i})\alpha_i L_i - (\gamma_{12i}\gamma_{34i} + \gamma_{12i}\gamma_{34i})\alpha_i S_i^2] + \sum_{i=5}^7 (1 - \alpha_i^2)^{-1} \times [\gamma_{12i}\gamma_{34i}Z_{1234}^{\text{mod}(2)} + \gamma_{12i}\gamma_{34i}Z_{1234}^{\text{mod}(1)}] R_i S_i \right\}, \quad (17)$$

where

$$L_i = (1 - \alpha_i^2)^{-1} [R_i^2 + \alpha_i^2 S_i^2 - 2\alpha_i R_i S_i \cos(\phi_i - \psi_i)] - 1$$

$$L_i = (1 - \alpha_i^2)^{-1} [S_i^2 + \alpha_i^2 R_i^2 - 2\alpha_i R_i S_i \cos(\phi_i - \psi_i)] - 1.$$

6. The joint probability distribution function of seven pairs of structure factors

In order to write the complete expression for the distribution P_7 , let us consider how the Fourier transform works when applied to (8). We have seen in the preceding section that the term

$$\rho_1\rho_2\rho_3\rho_4 \cos(\nu_1+\nu_2+\nu_3+\nu_4)$$

in the characteristic function C gives rise to the 16-term function Z_{1234} . By analogy, each of the 16 cyclic terms of $\gamma_{1234}\rho_1\rho_2\rho_3\rho_4 \cos(\nu_1+\nu_2+\nu_3+\nu_4)$ [see (9)] will produce a suitable Z_{ijpd} function. For example, the term $\rho_1\rho_2\rho_3\gamma_4 \cos(\nu_1+\nu_2+\nu_3+\mu_4)$ will give rise to $Z_{123\bar{4}}$, the term $\rho_1\rho_2\gamma_3\gamma_4 \cos(\nu_1+\nu_2+\nu_3+\mu_4)$ will give rise to $Z_{12\bar{3}\bar{4}}$ etc. For reader usefulness, we show here only the expression for $Z_{123\bar{4}}$; the reader can derive the other Z functions in a cyclic way.

We have

$$\begin{aligned} Z_{123\bar{4}} = &16[R_1 R_2 R_3 S_4 \cos(\phi_1 + \phi_2 + \phi_3 + \psi_4) \\ &- \alpha_1 S_1 R_2 R_3 S_4 \cos(\psi_1 + \phi_2 + \phi_3 + \psi_4) \\ &- \alpha_2 R_1 S_2 R_3 S_4 \cos(\phi_1 + \psi_2 + \phi_3 + \psi_4) \\ &- \alpha_3 R_1 R_2 S_3 S_4 \cos(\phi_1 + \phi_2 + \psi_3 + \psi_4) \\ &- \alpha_4 R_1 R_2 R_3 R_4 \cos(\phi_1 + \phi_2 + \phi_3 + \phi_4) \\ &+ \alpha_1 \alpha_2 S_1 S_2 R_3 S_4 \cos(\psi_1 + \psi_2 + \phi_3 + \psi_4) \\ &+ \alpha_1 \alpha_3 S_1 R_2 S_3 S_4 \cos(\psi_1 + \phi_2 + \psi_3 + \psi_4) \\ &+ \alpha_1 \alpha_4 S_1 R_2 R_3 R_4 \cos(\psi_1 + \phi_2 + \phi_3 + \phi_4) \\ &+ \alpha_2 \alpha_3 R_1 S_2 S_3 S_4 \cos(\phi_1 + \psi_2 + \psi_3 + \psi_4) \\ &+ \alpha_2 \alpha_4 R_1 S_2 R_3 R_4 \cos(\phi_1 + \psi_2 + \phi_3 + \phi_4) \\ &+ \alpha_3 \alpha_4 R_1 R_2 S_3 R_4 \cos(\phi_1 + \phi_2 + \psi_3 + \phi_4) \\ &+ \alpha_1 \alpha_2 \alpha_3 S_1 S_2 S_3 S_4 \cos(\psi_1 + \psi_2 + \psi_3 + \psi_4) \\ &+ \alpha_1 \alpha_2 \alpha_4 S_1 S_2 R_3 R_4 \cos(\psi_1 + \psi_2 + \phi_3 + \phi_4) \\ &+ \alpha_1 \alpha_3 \alpha_4 S_1 R_2 S_3 R_4 \cos(\psi_1 + \phi_2 + \psi_3 + \phi_4) \\ &+ \alpha_2 \alpha_3 \alpha_4 R_1 S_2 S_3 R_4 \cos(\phi_1 + \psi_2 + \psi_3 + \phi_4) \\ &+ \alpha_1 \alpha_2 \alpha_3 S_1 S_2 S_3 R_4 \cos(\psi_1 + \psi_2 + \psi_3 + \phi_4)]. \end{aligned}$$

If we regroup all the terms in the various Z_{ijpd} functions into: (a) the subset relative to $\cos(\phi_1 + \phi_2 + \phi_3 + \phi_4)$; (b) the subset relative to $\cos(\psi_1 + \phi_2 + \phi_3 + \phi_4)$; (c) etc., we will obtain the following 16 cumulative terms:

$$\begin{aligned} &2\beta_{1234}R_1R_2R_3R_4 \cos(\phi_1 + \phi_2 + \phi_3 + \phi_4) \\ &+ 2\beta_{123\bar{4}}S_1R_2R_3R_4 \cos(\psi_1 + \phi_2 + \phi_3 + \phi_4) \\ &+ 2\beta_{1\bar{2}34}R_1S_2R_3R_4 \cos(\phi_1 + \psi_2 + \phi_3 + \phi_4) \\ &+ 2\beta_{12\bar{3}\bar{4}}R_1R_2S_3R_4 \cos(\phi_1 + \phi_2 + \psi_3 + \phi_4) \\ &+ \dots + 2\beta_{1\bar{2}\bar{3}\bar{4}}S_1S_2S_3S_4 \cos(\psi_1 + \psi_2 + \psi_3 + \psi_4), \end{aligned}$$

(c) the B_{ijl} are defined as

$$\begin{aligned} B_{ijl} &= \gamma_{ijl}^2 [(1 - \alpha_i^2)(1 - \alpha_j^2)(1 - \alpha_l^2)]^{-1} L_i L_j L_l \\ B_{i\bar{j}\bar{l}} &= \gamma_{i\bar{j}\bar{l}}^2 [(1 - \alpha_i^2)(1 - \alpha_j^2)(1 - \alpha_l^2)]^{-1} L_i L_j L_l \\ &\vdots \\ B_{i\bar{j}l} &= \gamma_{i\bar{j}l}^2 [(1 - \alpha_i^2)(1 - \alpha_j^2)(1 - \alpha_l^2)]^{-1} L_i L_j L_l \end{aligned}$$

(d) the various B_{ijlp} in the distribution (18) and their modified forms are defined as

$$\begin{aligned} B_{1234i} &= \left[\prod_{i=1}^4 (1 - \alpha_i^2)^{-1} \right] (\gamma_{12i}\gamma_{34i} - \gamma_{12i}\gamma_{34i}\alpha_4 \\ &\quad - \gamma_{12i}\gamma_{34i}\alpha_3 - \dots + \gamma_{1\bar{2}i}\gamma_{3\bar{4}i}\alpha_1\alpha_2\alpha_3\alpha_4) \\ B_{123\bar{4}i} &= \left[\prod_{i=1}^4 (1 - \alpha_i^2)^{-1} \right] (\gamma_{12i}\gamma_{3\bar{4}i} - \gamma_{12i}\gamma_{34i}\alpha_4 \\ &\quad - \gamma_{12i}\gamma_{3\bar{4}i}\alpha_3 - \dots + \gamma_{1\bar{2}i}\gamma_{3\bar{4}i}\alpha_1\alpha_2\alpha_3\alpha_4) \\ &\vdots \\ B_{1\bar{2}\bar{3}4i} &= \left[\prod_{i=1}^4 (1 - \alpha_i^2)^{-1} \right] (\gamma_{1\bar{2}i}\gamma_{3\bar{4}i} - \gamma_{1\bar{2}i}\gamma_{34i}\alpha_4 \\ &\quad - \gamma_{1\bar{2}i}\gamma_{3\bar{4}i}\alpha_3 - \dots + \gamma_{12i}\gamma_{34i}\alpha_1\alpha_2\alpha_3\alpha_4) \\ B_{1234\bar{i}} &= \left[\prod_{i=1}^4 (1 - \alpha_i^2)^{-1} \right] (\gamma_{12\bar{i}}\gamma_{34\bar{i}} - \gamma_{12\bar{i}}\gamma_{3\bar{4}\bar{i}}\alpha_4 \\ &\quad - \dots + \gamma_{1\bar{2}\bar{i}}\gamma_{3\bar{4}\bar{i}}\alpha_1\alpha_2\alpha_3\alpha_4) \\ B_{123\bar{4}\bar{i}} &= \left[\prod_{i=1}^4 (1 - \alpha_i^2)^{-1} \right] (\gamma_{12\bar{i}}\gamma_{3\bar{4}\bar{i}} - \gamma_{12\bar{i}}\gamma_{34\bar{i}}\alpha_4 \\ &\quad - \dots + \gamma_{1\bar{2}\bar{i}}\gamma_{3\bar{4}\bar{i}}\alpha_1\alpha_2\alpha_3\alpha_4) \\ &\vdots \\ B_{1\bar{2}\bar{3}4\bar{i}} &= \left[\prod_{i=1}^4 (1 - \alpha_i^2)^{-1} \right] (\gamma_{1\bar{2}\bar{i}}\gamma_{3\bar{4}\bar{i}} - \gamma_{1\bar{2}\bar{i}}\gamma_{34\bar{i}}\alpha_4 \\ &\quad - \dots + \gamma_{12\bar{i}}\gamma_{34\bar{i}}\alpha_1\alpha_2\alpha_3\alpha_4) \\ B_{1234i}^{\text{mod}(2)} &= \left[\prod_{i=1}^4 (1 - \alpha_i^2)^{-1} \right] (\gamma_{12i}\gamma_{34i} - \gamma_{12i}\gamma_{3\bar{4}i}\alpha_4 \\ &\quad - \dots + \gamma_{1\bar{2}i}\gamma_{3\bar{4}i}\alpha_1\alpha_2\alpha_3\alpha_4) \\ B_{1234i}^{\text{mod}(2)} &= \left[\prod_{i=1}^4 (1 - \alpha_i^2)^{-1} \right] (\gamma_{12i}\gamma_{34i} - \gamma_{12i}\gamma_{34i}\alpha_4 \\ &\quad - \dots + \gamma_{1\bar{2}i}\gamma_{34i}\alpha_1\alpha_2\alpha_3\alpha_4) \\ &\vdots \\ B_{1\bar{2}\bar{3}4i}^{\text{mod}(2)} &= \left[\prod_{i=1}^4 (1 - \alpha_i^2)^{-1} \right] (\gamma_{1\bar{2}i}\gamma_{3\bar{4}i} - \gamma_{1\bar{2}i}\gamma_{3\bar{4}i}\alpha_4 \\ &\quad - \dots + \gamma_{12i}\gamma_{34i}\alpha_1\alpha_2\alpha_3\alpha_4) \end{aligned}$$

$$\begin{aligned} B_{1234i}^{\text{mod}(1)} &= \left[\prod_{i=1}^4 (1 - \alpha_i^2)^{-1} \right] (\gamma_{12i}\gamma_{34i} - \gamma_{12i}\gamma_{3\bar{4}i}\alpha_4 \\ &\quad - \dots + \gamma_{1\bar{2}i}\gamma_{3\bar{4}i}\alpha_1\alpha_2\alpha_3\alpha_4) \\ B_{1234i}^{\text{mod}(1)} &= \left[\prod_{i=1}^4 (1 - \alpha_i^2)^{-1} \right] (\gamma_{12\bar{i}}\gamma_{34\bar{i}} - \gamma_{12\bar{i}}\gamma_{3\bar{4}\bar{i}}\alpha_4 \\ &\quad - \dots + \gamma_{1\bar{2}\bar{i}}\gamma_{3\bar{4}\bar{i}}\alpha_1\alpha_2\alpha_3\alpha_4) \\ &\vdots \\ B_{1\bar{2}\bar{3}4i}^{\text{mod}(1)} &= \left[\prod_{i=1}^4 (1 - \alpha_i^2)^{-1} \right] (\gamma_{1\bar{2}i}\gamma_{3\bar{4}i} - \gamma_{1\bar{2}i}\gamma_{3\bar{4}i}\alpha_4 \\ &\quad - \dots + \gamma_{12i}\gamma_{34i}\alpha_1\alpha_2\alpha_3\alpha_4). \end{aligned}$$

7. Conclusions

The distribution (18) is the main goal of this paper. It is the joint probability distribution of seven pairs of structure factors, whose indices are chosen so as to allow the estimate of the quartet phase invariants when the diffraction data from two isomorphous structures are available. Expression (18) is very complicated and does not immediately reveal its basic features. In particular, it is cumbersome for the protein case, to which this paper is particularly addressed, where the derivatives are obtained by addition of heavy atoms. In such a situation, it may be expected that (18) may be substantially simplified [see GCZ for simplification of the distribution (4)]. In the following paper (Giacovazzo & Siliqi, 1996), we will derive such a simpler distribution function and will apply it to real protein data. It is anticipated here that the results will be satisfactory and that the theoretical implications of the quartet theory contribute to a sounder interpretation of the Hauptman and GCZ formulas for triplet invariants.

APPENDIX A

The basic formulas used for the derivation of the joint probability distribution function (18) are here collected.

(a)

$$\begin{aligned} &\int_0^{2\pi} \cos(\tau - s) \exp\{-it[a \cos(q - \tau) - b \cos(r - \tau)]\} d\tau \\ &= (-2\pi i)[J_1(tQ)/Q][a \cos(q - s) - b \cos(r - s)], \end{aligned} \quad (19)$$

where

$$Q^2 = a^2 + b^2 - 2ab \cos(q - r)$$

and J_1 is the Bessel function of the first kind of order 1.

(b)

$$\int_0^\infty t^2 J_1(Qt) \exp(-p^2 t^2 / 2) dt = (Q/p^4) \exp(-Q^2 / 2p^2). \quad (20)$$

$$\begin{aligned}
 (c) \quad & \int_0^\infty t \exp(-\frac{1}{2}p^2 t^2) dt \int_0^{2\pi} \exp\{-it[a \cos(q - \tau) \\
 & - b \cos(r - \tau)]\} d\tau \\
 & = (2\pi/p^2) \exp(-Q^2/2p^2). \quad (21)
 \end{aligned}$$

$$\begin{aligned}
 (d) \quad & \int_0^\infty t^2 \exp(-\frac{1}{2}p^2 t^2) dt \int_0^{2\pi} \cos(s - \tau) \exp\{-it[a \cos(q - \tau) \\
 & - b \cos(r - \tau)]\} d\tau \\
 & = (-2\pi i/p^4)[a \cos(q - s) - b \cos(r - s)] \\
 & \quad \times \exp(-Q^2/2p^2). \quad (22)
 \end{aligned}$$

$$\begin{aligned}
 (e) \quad & \int_0^\infty t^3 \exp(-\frac{1}{2}p^2 t^2) dt \int_0^{2\pi} \exp\{-it[a \cos(q - \tau) \\
 & - b \cos(r - \tau)]\} d\tau \\
 & = (4\pi/p^4) \exp(-Q^2/2p^2)(1 - Q^2/2p^2). \quad (23)
 \end{aligned}$$

APPENDIX B

The integral of the first-order term in (5) may be written as

$$\begin{aligned}
 (2\pi)^{-12} 2^6 & \left[\prod_{i=1}^3 (R_i S_i) \right] \int_0^\infty \rho_1 \exp(-\rho_1^2/2) \int_0^\infty \rho_2 \exp(-\rho_2^2/2) \\
 & \times \int_0^\infty \rho_3 \exp(-\rho_3^2/2) d\rho_1 d\rho_2 d\rho_3 \int_0^{2\pi} \dots \int_0^{2\pi} \exp\{-i2^{1/2} \\
 & \times [\rho_1 R_1 \cos(\phi_1 - \nu_1) + \rho_2 R_2 \cos(\phi_2 - \nu_2) \\
 & + \rho_3 R_3 \cos(\phi_3 - \nu_3)]\} d\nu_1 d\nu_2 d\nu_3 \prod_{i=1}^3 \int_0^\infty \gamma_i \exp(-\gamma_i^2/2) \\
 & \times \int_0^{2\pi} \exp\{-i\gamma_i [2^{1/2} S_i \cos(\psi_i - \mu_i) \\
 & - i\alpha_i \rho_i \cos(\nu_i - \mu_i)]\} d\gamma_i d\mu_i. \quad (24)
 \end{aligned}$$

The last limit may be evaluated by means of (21) and gives

$$(2\pi)^3 \exp\left\{-\sum_{i=1}^3 [S_i^2 - \alpha_i^2 \rho_i^2/2 - i2^{1/2} \alpha_i \rho_i S_i \cos(\psi_i - \nu_i)]\right\}. \quad (25)$$

Then, (24) reduces to

$$\begin{aligned}
 (2\pi)^{-9} 2^6 & \prod_{i=1}^3 [R_i S_i \exp(-S_i^2)] \prod_{i=1}^3 \left\{ \int_0^\infty \rho_i \exp[-\frac{1}{2}(1 - \alpha_i^2) \rho_i^2] d\rho_i \right. \\
 & \times \int_0^{2\pi} \exp\{-i\rho_i [2^{1/2} R_i \cos(\phi_i - \nu_i) \\
 & \left. - 2^{1/2} \alpha_i S_i \cos(\phi_i - \nu_i)]\} d\nu_i \right\},
 \end{aligned}$$

which, because of (21), reduces to

$$\begin{aligned}
 (2\pi)^{-9} 2^6 & \prod_{i=1}^3 [R_i S_i \exp(-S_i^2)] \prod_{i=1}^3 \left\{ [2\pi/(1 - \alpha_i^2)] \right. \\
 & \times \exp\{-[1/(1 - \alpha_i^2)][R_i^2 + \alpha_i^2 S_i^2 \\
 & \left. - 2\alpha_i R_i S_i \cos(\psi_i - \phi_i)]\} \right\} \\
 & = \pi^{-6} \prod_{i=1}^3 [R_i S_i / (1 - \alpha_i^2)] \exp\{-[1/(1 - \alpha_i^2)] \\
 & \times [R_i^2 + S_i^2 - 2\alpha_i R_i S_i \cos(\psi_i - \phi_i)]\}.
 \end{aligned}$$

APPENDIX C

The integral of the term in (5) containing γ_{123} may be written as

$$\begin{aligned}
 (2\pi)^{-12} 2^6 & \prod_{i=1}^3 (R_i S_i) \left\{ -i2^{-1/2} \gamma_{123} \int_0^\infty \rho_1^2 \exp(-\rho_1^2/2) \right. \\
 & \times \int_0^\infty \rho_2^2 \exp(-\rho_2^2/2) \int_0^\infty \rho_3^2 \exp(-\rho_3^2/2) d\rho_1 d\rho_2 d\rho_3 \\
 & \times \int_0^{2\pi} \dots \int_0^{2\pi} \cos(\nu_1 + \nu_2 + \nu_3) \exp\{-i2^{1/2} \\
 & \times [\rho_1 R_1 \cos(\phi_1 - \nu_1) + \rho_2 R_2 \cos(\phi_2 - \nu_2) \\
 & + \rho_3 R_3 \cos(\phi_3 - \nu_3)]\} d\nu_1 d\nu_2 d\nu_3 \\
 & \times \prod_{i=1}^3 \int_0^\infty \gamma_i \exp(-\frac{1}{2} \gamma_i^2) \int_0^{2\pi} \exp\{-i\gamma_i [2^{1/2} S_i \cos(\psi_i - \mu_i) \\
 & \left. - i\alpha_i \rho_i \cos(\nu_i - \mu_i)]\} d\gamma_i d\mu_i \right\}. \quad (26)
 \end{aligned}$$

The last line has been evaluated in Appendix B [see (25)]. Accordingly, (26) reduces to

$$\begin{aligned}
 (2\pi)^{-9} 2^6 & \prod_{i=1}^3 [R_i S_i \exp(-S_i^2)] \left\{ -i2^{-1/2} \gamma_{123} \right. \\
 & \times \int_0^\infty \rho_2^2 \exp[-\frac{1}{2}(1 - \alpha_2^2) \rho_2^2] d\rho_2 \\
 & \times \int_0^\infty \rho_3^2 \exp[-\frac{1}{2}(1 - \alpha_3^2) \rho_3^2] d\rho_3 \\
 & \times \int_0^{2\pi} \exp\{-i\rho_2 [2^{1/2} R_2 \cos(\phi_2 - \nu_2) \\
 & - 2^{1/2} \alpha_2 S_2 \cos(\psi_2 - \nu_2)]\} d\nu_2 \\
 & \times \int_0^{2\pi} \exp\{-i\rho_3 [2^{1/2} R_3 \cos(\phi_3 - \nu_3) \\
 & - 2^{1/2} \alpha_3 S_3 \cos(\psi_3 - \nu_3)]\} d\nu_3 \\
 & \times \int_0^{2\pi} \rho_1^2 \exp[-\frac{1}{2}(1 - \alpha_1^2) \rho_1^2] d\rho_1 \int_0^{2\pi} \cos(\nu_1 + \nu_2 + \nu_3) \\
 & \times \exp\{-i\rho_1 [2^{1/2} R_1 \cos(\phi_1 - \nu_1) \\
 & \left. - 2^{1/2} \alpha_1 S_1 \cos(\psi_1 - \nu_1)]\} d\nu_1 \right\}.
 \end{aligned}$$

The last integral can be evaluated with (22) and is

$$\begin{aligned}
 & - [2\pi i / (1 - \alpha_1^2)^2] [2^{1/2} R_1 \cos(\phi_1 + \nu_2 + \nu_3) \\
 & - 2^{1/2} \alpha_1 S_1 \cos(\psi_1 + \nu_2 + \nu_3)] \exp\{-[1/(1 - \alpha_1^2)] \\
 & \times [R_1^2 + \alpha_1^2 S_1^2 - 2\alpha_1 R_1 S_1 \cos(\phi_1 - \psi_1)]\}.
 \end{aligned}$$

Integration with respect to the pairs (ρ_2, ν_2) and (ρ_3, ν_3) may be accomplished *via* the same technique.

References

- Altomare, A., Burla, M. C., Cascarano, G., Giacovazzo, C. & Guagliardi, A. (1993). *Acta Cryst.* **A49**, 342–346.
- Fortier, S., Weeks, C. M. & Hauptman, H. (1984). *Acta Cryst.* **A40**, 544–548.
- Giacovazzo, C. (1975). *Acta Cryst.* **A31**, 252–259.
- Giacovazzo, C. (1976). *Acta Cryst.* **A32**, 91–99.
- Giacovazzo, C. (1980). *Direct Methods in Crystallography*. London: Academic Press.
- Giacovazzo, C., Cascarano, G. & Zheng, C. (1988). *Acta Cryst.* **A44**, 45–51.
- Giacovazzo, C. & Siliqi, D. (1996). *Acta Cryst.* **A52**, 143–151.
- Hauptmann, H. (1982). *Acta Cryst.* **A38**, 289–294.
- Hauptman, H., Potter, S. & Weeks, C. M. (1982). *Acta Cryst.* **A38**, 294–300.
- Kiriakidis, C. E. (1993). Thesis, University of Amsterdam, The Netherlands.
- Kiriakidis, C. E., Peschar, R. & Schenk, H. (1993). *Acta Cryst.* **A49**, 557–569.