

Integration of Patterson Information into Direct Methods. I. The Theoretical Background

BY CARMELO GIACOVAZZO

*Istituto di Ricerca per lo Sviluppo di Metodologie Cristallografiche, CNR,
c/o Dipartimento Geomineralogico, Campus Universitario, 70124 Bari, Italy*

(Received 30 March 1990; accepted 12 December 1990)

Abstract

A theoretical approach is described aiming at exploiting, by the method of joint probability distribution functions of structure factors, the information provided by a Patterson map. Among the various possible types of information the minimal one, *i.e.* position and intensity of a Patterson peak, has been taken into account. It is shown that Cochran's distribution of triplet phases is substantially modified when such prior information is considered.

1. Symbols and abbreviations

Throughout this paper a number of symbols will find frequent application. For the sake of simplicity they are here listed together.

$f(\mathbf{h})$: atomic scattering factor. The thermal factor is included; anomalous dispersion is not considered.

N : number of atoms in the cell.

V : volume of the cell.

t : number of symmetry-independent atoms in the cell.

m : order of the space group (it equals the number of symmetry operators).

$\rho(\mathbf{r})$: electron density function.

$F_{\mathbf{h}}$: structure factor with vectorial index \mathbf{h} .

$E'_{\mathbf{h}} = R'_{\mathbf{h}} \exp(i\varphi_{\mathbf{h}})$: normalized structure factor when the only available prior information is the atomicity. $R'_{\mathbf{h}}$ is the modulus of $E'_{\mathbf{h}}$, $\varphi_{\mathbf{h}}$ is the phase.

$E_{\mathbf{h}} = R_{\mathbf{h}} \exp(i\varphi_{\mathbf{h}})$: normalized structure factor when, besides atomicity, information on some interatomic vectors is also available.

\mathbf{r}_j : positional vector of the j th atom.

\mathbf{u} : interatomic vector.

$\mathbf{C}_s \equiv (\mathbf{R}_s, \mathbf{T}_s)$: s th symmetry operator. \mathbf{R}_s is the rotational part, \mathbf{T}_s the translational part.

$\mathbf{C}_s \mathbf{r}_j = \mathbf{R}_s \mathbf{r}_j + \mathbf{T}_s$: positional vector of the atom symmetrically equivalent to \mathbf{r}_j .

$$\sum_N(\mathbf{h}) = \sum_{j=1}^N f_j^2(\mathbf{h}).$$

$$\sum_3(\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3) = \sum_{j=1}^N f_j(\mathbf{h}_1) f_j(\mathbf{h}_2) f_j(\mathbf{h}_3).$$

$\Phi = \varphi_{\mathbf{h}_1} + \varphi_{\mathbf{h}_2} + \varphi_{\mathbf{h}_3}$ with $\mathbf{h}_1 + \mathbf{h}_2 + \mathbf{h}_3 = 0$.

$\sigma_n = \sum_{j=1}^N Z_j^n$ where Z_j is the atomic number of the j th atom.

$I_n(x)$: modified Bessel function of order n .

2. Introduction

Both Patterson and direct methods determine the atomic positions starting from diffraction intensities: direct methods are used in reciprocal space, Patterson methods in direct space. The connection between the two methods has been considered by many authors (for a review, see Giacovazzo, 1980).

Vaughan (1958) studied the function

$$P(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n) = \int_V \rho(\mathbf{r}) \rho(\mathbf{r} + \mathbf{u}_1) \dots \rho(\mathbf{r} + \mathbf{u}_n) dV$$

to derive phase information on triplet invariants (for $n=1$ the function P coincides with the usual Patterson function). Hauptman & Karle (1962) developed their algebraic 'vector interaction formula' which exploits pairs of interatomic vectors satisfying the relation $\mathbf{u}_{ij} + \mathbf{u}_{jk} = \mathbf{u}_{ik}$ where $\mathbf{u}_{ij} = \mathbf{r}_i - \mathbf{r}_j$. Hoppe (1963), Main & Woolfson (1963), Allegra (1979) and Rius & Miravittles (1989) described algebraic methods for phase estimation which make use of the zero points in the Patterson function. Further contributions have been made by Kroon & Krabbendam (1970) who showed how the molecular orientation can be related to some features of the double Patterson and how that can be used to determine the signs of triplet invariants. Krabbendam & Kroon (1971) emphasized the importance of the Hoppe (1957) sections of the double Patterson for phase determination. von Eller (1973) applied his polynomial method to a Patterson map in order to derive Cochran's relation (Cochran, 1955) for triplet estimation. Heinerman, Krabbendam & Kroon [1975; see also Heinerman, Krabbendam, Kroon & Spek (1978) and Pontenagel, Krabbendam & Kroon (1987)] studied in $P\bar{1}$ the joint probability distribution $P(E_{\mathbf{k}}, E_{\mathbf{n}+\mathbf{k}}, M)$ with \mathbf{h} fixed, where

$$M = \cos [2\pi(\mathbf{h}/2 + \mathbf{k}) \cdot \mathbf{u}]$$

and \mathbf{u} is the Patterson vector of the symmetry-equivalent atoms ($\mathbf{u} = 2\mathbf{r}$). New probabilities based

on the Patterson function were also derived by Brosius (1985) in $P\bar{1}$, with special interest in the sign of the seminvariant E_{2h} .

In the present paper (the first of a series), we study how the information arising from a Patterson map, used as prior information, may modify Cochran's distribution for triplet invariants. No interpretation of the Patterson map is needed for the application of the present theory: in particular, pairs of peaks of type $\mathbf{u}_{ij} = \mathbf{r}_i - \mathbf{r}_j$ and $\mathbf{u}_{ik} = \mathbf{r}_i - \mathbf{r}_k$ do not need to be sought. In practice, we intend to exploit in this paper the minimum level of information provided by a Patterson function, *i.e.* coordinates and intensities of one or more peaks. The use of more specific or elaborate information contained in a Patterson map [*i.e.* Harker peaks or pairs of vectors of type $(\mathbf{u}_{ij}, \mathbf{u}_{jk})$ *etc.*] will be discussed in the following papers of this series.

In the present approach, the set of symmetry-independent atomic positional vectors are chosen to be the random variables, while $\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3$ are a fixed triple of reciprocal vectors satisfying

$$\mathbf{h}_1 + \mathbf{h}_2 + \mathbf{h}_3 = 0. \quad (1)$$

In the absence of Patterson information, all the t positional vectors \mathbf{r}_j can be assumed to be statistically independent of one another and uniformly distributed in the asymmetric unit. Let us now suppose that, by inspection of a Patterson map, an interatomic vector, say $\mathbf{u} = \mathbf{r}_{j_1} - \mathbf{r}_{j_2}$, is *a priori* known. Also, let the scattering factors of the atoms at \mathbf{r}_{j_1} and \mathbf{r}_{j_2} be known (we will see later that such an assumption is not too critical in most cases). Then \mathbf{r}_{j_2} is no longer a random variable since it is completely determined in terms of \mathbf{r}_{j_1} and \mathbf{u} . Consequently, the number of primitive random vectors reduces from t to $t-1$.

The above considerations suggest that standard probabilistic methods for triplet invariant estimates have to be modified when prior information on an interatomic vector is available. This is the main aim of the present paper. It will also be shown that, under some general hypotheses, prior information from a Patterson peak arising from the overlapping of more interatomic vectors may also be treated by the theory here developed.

3. The role of the Patterson information in direct methods

It was shown by Cochran (1952) that, owing to the atomicity of the electron density, a crystal structure should satisfy

$$\int_V \rho^3(\mathbf{r}) \, d\mathbf{r} = \max. \quad (2)$$

Stanley (1979) used (2) as a specific criterion for direct determination of centrosymmetrical electron density distributions: the correct sign for a structure factor was assumed to be that giving rise to the largest

value of (2). This relationship may be considered the real-space counterpart of the tangent formula (Karle & Hauptman, 1956): their application in symmorphic space groups mostly produces the so-called Patterson-like function, where all the F 's have positive signs or have signs systematically defined by the parity of the indices.

Positivity and peakiness of the Patterson function may be used in combination with (2) to improve the efficiency of the phasing process. For the sake of simplicity, but without losing generality, let us assume that the space group is $P1$. If the position of a Patterson peak is known we can expect that in a suitable domain

$$S = \int_V \rho^3(\mathbf{r}) \, d\mathbf{r} + a \int_V \rho(\mathbf{r} + \mathbf{u})\rho^2(\mathbf{r}) \, d\mathbf{r} = \max \quad (3)$$

should be satisfied, where a is a suitable parameter. Relation (3) can be written as

$$S = P(0, 0) + aP(\mathbf{u}, 0) = \max,$$

where $P(\mathbf{u}, \mathbf{v})$ is the double Patterson function. Expressing ρ in terms of structure factors gives

$$S = \sum_{\mathbf{h}_1 + \mathbf{h}_2 + \mathbf{h}_3 = 0} F_{\mathbf{h}_1} F_{\mathbf{h}_2} F_{\mathbf{h}_3} [1 + a \exp(-2\pi i \mathbf{h}_1 \mathbf{u})] = \max.$$

The same triplet $F_{\mathbf{h}_1} F_{\mathbf{h}_2} F_{\mathbf{h}_3}$ will appear in S 12 times: its total contribution to $\int_V \rho(\mathbf{r} + \mathbf{u})\rho^2(\mathbf{r}) \, d\mathbf{r}$ is

$$\begin{aligned} & F_{\mathbf{h}_1} F_{\mathbf{h}_2} F_{\mathbf{h}_3} \exp(-2\pi i \mathbf{h}_1 \mathbf{u}) \\ & + F_{\mathbf{h}_1} F_{\mathbf{h}_3} F_{\mathbf{h}_2} \exp(-2\pi i \mathbf{h}_1 \mathbf{u}) \\ & + F_{\mathbf{h}_2} F_{\mathbf{h}_1} F_{\mathbf{h}_3} \exp(-2\pi i \mathbf{h}_2 \mathbf{u}) \\ & + F_{\mathbf{h}_2} F_{\mathbf{h}_3} F_{\mathbf{h}_1} \exp(-2\pi i \mathbf{h}_2 \mathbf{u}) \\ & + F_{\mathbf{h}_3} F_{\mathbf{h}_1} F_{\mathbf{h}_2} \exp(-2\pi i \mathbf{h}_3 \mathbf{u}) \\ & + F_{\mathbf{h}_3} F_{\mathbf{h}_2} F_{\mathbf{h}_1} \exp(-2\pi i \mathbf{h}_3 \mathbf{u}) \\ & + F_{-\mathbf{h}_1} F_{-\mathbf{h}_2} F_{-\mathbf{h}_3} \exp(2\pi i \mathbf{h}_1 \mathbf{u}) \\ & + F_{-\mathbf{h}_1} F_{-\mathbf{h}_3} F_{-\mathbf{h}_2} \exp(2\pi i \mathbf{h}_1 \mathbf{u}) + \dots \\ & = 2|F_{\mathbf{h}_1} F_{\mathbf{h}_2} F_{\mathbf{h}_3}| \exp(i\Phi) [\exp(-2\pi i \mathbf{h}_1 \mathbf{u}) \\ & \quad + \exp(-2\pi i \mathbf{h}_2 \mathbf{u}) + \exp(-2\pi i \mathbf{h}_3 \mathbf{u})] \\ & \quad + 2|F_{\mathbf{h}_1} F_{\mathbf{h}_2} F_{\mathbf{h}_3}| \exp(-i\Phi) [\exp(2\pi i \mathbf{h}_1 \mathbf{u}) \\ & \quad + \exp(2\pi i \mathbf{h}_2 \mathbf{u}) + \exp(2\pi i \mathbf{h}_3 \mathbf{u})] \\ & = 4|F_{\mathbf{h}_1} F_{\mathbf{h}_2} F_{\mathbf{h}_3}| \{ \cos(\Phi - 2\pi \mathbf{h}_1 \mathbf{u}) \\ & \quad + \cos(\Phi - 2\pi \mathbf{h}_2 \mathbf{u}) + \cos(\Phi - 2\pi \mathbf{h}_3 \mathbf{u}) \}. \end{aligned}$$

Accordingly, the total contribution in (3) of the triplet $F_{\mathbf{h}_1} F_{\mathbf{h}_2} F_{\mathbf{h}_3}$ to $\int_V \rho^3(\mathbf{r}) \, d\mathbf{r}$ is $12|F_{\mathbf{h}_1} F_{\mathbf{h}_2} F_{\mathbf{h}_3}| \cos \Phi$. Then

$$S = 4 \sum' |F_{\mathbf{h}_1} F_{\mathbf{h}_2} F_{\mathbf{h}_3}| \{ 3 \cos \Phi + a [\cos(\Phi - 2\pi \mathbf{h}_1 \mathbf{u}) + \cos(\Phi - 2\pi \mathbf{h}_2 \mathbf{u}) + \cos(\Phi - 2\pi \mathbf{h}_3 \mathbf{u})] \},$$

where the prime to the summation warns the reader that Friedel-related triplets do not enter into the sum.

We can also maximize

$$S = \int_V \rho^3(\mathbf{r}) \, d\mathbf{r} + a \left[\int_V \rho(\mathbf{r}+\mathbf{u})\rho^2(\mathbf{r}) \, d\mathbf{r} + \int_V \rho(\mathbf{r}-\mathbf{u})\rho^2(\mathbf{r}) \, d\mathbf{r} \right]$$

so obtaining

$$S = 4 \sum'_{\mathbf{h}_1+\mathbf{h}_2+\mathbf{h}_3=0} |F_{\mathbf{h}_1}F_{\mathbf{h}_2}F_{\mathbf{h}_3}| \cos \Phi \\ \times \{3 + a[2 \cos 2\pi\mathbf{h}_1\mathbf{u} + 2 \cos 2\pi\mathbf{h}_2\mathbf{u} + 2 \cos 2\pi\mathbf{h}_3\mathbf{u}]\} \\ = \max.$$

We are now able to exploit the space-group symmetry. The function to maximize is now

$$S = \int_V \rho^3(\mathbf{r}) \, d\mathbf{r} + a \left[\sum_{s=1}^m \int_V \rho(\mathbf{r}+\mathbf{R}_s\mathbf{u})\rho^2(\mathbf{r}) \, d\mathbf{r} + \sum_{s=1}^m \int_V \rho(\mathbf{r}-\mathbf{R}_s\mathbf{u})\rho^2(\mathbf{r}) \, d\mathbf{r} \right]$$

which yields

$$S = \sum'_{\mathbf{h}_1+\mathbf{h}_2+\mathbf{h}_3=0} |F_{\mathbf{h}_1}F_{\mathbf{h}_2}F_{\mathbf{h}_3}| \cos \Phi \\ \times \left\{ 3 + 2a \sum_{s=1}^m (\cos 2\pi\mathbf{h}_1\mathbf{R}_s\mathbf{u} + \cos 2\pi\mathbf{h}_2\mathbf{R}_s\mathbf{u} + \cos 2\pi\mathbf{h}_3\mathbf{R}_s\mathbf{u}) \right\} \\ = \max.$$

When S is maximized, due account has to be taken of the multiplicity (according to symmetry) of the reflection \mathbf{h}_1 , \mathbf{h}_2 , \mathbf{h}_3 and/or of the interatomic vectors \mathbf{u} .

With the assumption that all $\varphi_{\mathbf{h}_2}$ and $\varphi_{\mathbf{h}_3}$ are known an estimate of $\varphi_{\mathbf{h}_1}$ may be obtained by applying the condition $\delta S / \delta \varphi_{\mathbf{h}_1} \cong 0$. Then

$$\tan \varphi_{\mathbf{h}_1} = - \frac{\sum_j \beta_j \sin(\varphi_{\mathbf{h}_2} + \varphi_{\mathbf{h}_3})}{\sum_j \beta_j \cos(\varphi_{\mathbf{h}_2} + \varphi_{\mathbf{h}_3})}, \quad (4)$$

where j varies over the set of triplets involving $\varphi_{\mathbf{h}_1}$ and

$$\beta = |F_{\mathbf{h}_1}F_{\mathbf{h}_2}F_{\mathbf{h}_3}| \left\{ 3 + 2a \sum_{s=1}^m (\cos 2\pi\mathbf{h}_1\mathbf{R}_s\mathbf{u} + \cos 2\pi\mathbf{h}_2\mathbf{R}_s\mathbf{u} + \cos 2\pi\mathbf{h}_3\mathbf{R}_s\mathbf{u}) \right\}.$$

In conclusion, (3) seems able to exploit a larger amount of prior information than the Cochran relation

$$\sum'_{\mathbf{h}_1+\mathbf{h}_2+\mathbf{h}_3=0} F_{\mathbf{h}_1}F_{\mathbf{h}_2}F_{\mathbf{h}_3} \cos \Phi = \max$$

and may be immediately generalized to the case in which more symmetry-independent Patterson peaks are *a priori* known. On the other hand, we have no clear way of fixing the 'best' value of the parameter a : furthermore, no probabilistic guess may be made about the reliability of (4) owing to its asymptotical nature. In order to obtain probabilistic phase relationships, we will apply to our problem the method of joint probability distribution functions modified in such a way that prior information on a Patterson peak is taken into account.

4. The average value of $|F_{\mathbf{h}}|^2$ when one interatomic vector is *a priori* known

If no prior information is available except atomicity then (Wilson, 1949)

$$\langle |F_{\mathbf{h}}|^2 \rangle = \sum_{j=1}^N f_j^2(\mathbf{h}). \quad (5)$$

Suppose now that the interatomic vector $\mathbf{u} = \mathbf{r}_p - \mathbf{r}_q$ has been identified from a Patterson map and that f_p and f_q are the scattering factors of the atoms in p and q . For the sake of simplicity, we will assume in our mathematical treatment that both \mathbf{r}_p and \mathbf{r}_q denote general atomic positions. Also, the average value of $|F_{\mathbf{h}}|^2$ when the prior information on \mathbf{u} is available will be denoted by $\langle |F_{\mathbf{h}}|^2 | \mathbf{u} \rangle$.

(a) Non-centrosymmetrical space groups

According to our hypotheses

$$F_{\mathbf{h}} = \sum_{\substack{j=1 \\ j \neq p,q}}^i f_j(\mathbf{h}) \sum_{s=1}^m \exp 2\pi i \mathbf{h} \mathbf{C}_s \mathbf{r}_j \\ + f_p(\mathbf{h}) \sum_{s=1}^m \exp 2\pi i \mathbf{h} \mathbf{C}_s \mathbf{r}_p \\ + f_q(\mathbf{h}) \sum_{s=1}^m \exp 2\pi i \mathbf{h} \mathbf{C}_s (\mathbf{r}_p + \mathbf{u}).$$

Then

$$\langle |F_{\mathbf{h}}|^2 | \mathbf{u} \rangle = \varepsilon_{\mathbf{h}} \left\{ m \sum_{\substack{j=1 \\ j \neq p,q}}^i f_j^2(\mathbf{h}) + m f_p^2(\mathbf{h}) + m f_q^2(\mathbf{h}) + 2 f_p(\mathbf{h}) f_q(\mathbf{h}) \sum_{s=1}^m \cos 2\pi \mathbf{h} \mathbf{R}_s \mathbf{u} \right\} \\ = \varepsilon_{\mathbf{h}} \left\{ \sum_N(\mathbf{h}) + 2 f_p(\mathbf{h}) f_q(\mathbf{h}) \sum_{s=1}^m \cos 2\pi \mathbf{h} \mathbf{R}_s \mathbf{u} \right\} \quad (6)$$

where $\varepsilon_{\mathbf{h}}$ is the Wilson coefficient for the \mathbf{h} reflection.

The coefficient 2 to the second term at the right side of (6) takes into account the fact that two Patterson peaks exist, at \mathbf{u} and $-\mathbf{u}$, respectively.

(b) Centrosymmetrical space groups

$$F_{\mathbf{h}} = 2 \sum_{\substack{j=1 \\ j \neq p,q}}^l f_j(\mathbf{h}) \sum_{s=1}^{m/2} \cos 2\pi \mathbf{h} \mathbf{C}_s \mathbf{r}_j \\ + 2f_p(\mathbf{h}) \sum_{s=1}^{m/2} \cos 2\pi \mathbf{h} \mathbf{C}_s \mathbf{r}_j \\ + 2f_q(\mathbf{h}) \sum_{s=1}^{m/2} \cos 2\pi \mathbf{h} \mathbf{C}_s (\mathbf{r}_j + \mathbf{u}),$$

where $m/2$ is the number of symmetry operations not related by the inversion centre. Accordingly,

$$\langle |F_{\mathbf{h}}|^2 | \mathbf{u} \rangle = \varepsilon_{\mathbf{h}} \left\{ m \sum_{\substack{j=1 \\ j \neq p,q}}^l f_j^2(\mathbf{h}) + mf_p^2(\mathbf{h}) \right. \\ \left. + mf_q^2(\mathbf{h}) + 4f_p(\mathbf{h})f_q(\mathbf{h}) \sum_{s=1}^{m/2} \cos 2\pi \mathbf{h} \mathbf{R}_s \mathbf{u} \right\} \\ = \varepsilon_{\mathbf{h}} \left\{ \sum_N(\mathbf{h}) + 4f_p(\mathbf{h})f_q(\mathbf{h}) \sum_{s=1}^{m/2} \cos 2\pi \mathbf{h} \mathbf{R}_s \mathbf{u} \right\}. \tag{7}$$

The coefficient 4 in the second term on the right side of (7) takes into account the fact that, in centrosymmetrical space groups, besides the pair $(\mathbf{r}_p, \mathbf{r}_q)$ the pair $(-\mathbf{r}_p, -\mathbf{r}_q)$ also exists.

Comparison between (5) and (6)-(7) suggests that prior information on an interatomic vector may significantly change the expected values of the squared moduli, especially when the atoms in \mathbf{r}_p and \mathbf{r}_q represent a non-negligible percentage of the electrons in the unit cell. It may also be noted that (6) may be used for centrosymmetrical space groups too: this time m includes symmetry operators related by the inversion centre.

For subsequent use we define here two types of normalized (according to the prior information) structure factors:

$$E'_b = F_b / \langle |F_b|^2 \rangle^{1/2} \quad \text{and} \quad E_b = F_b / \langle |F_b|^2 | \mathbf{u} \rangle^{1/2}.$$

R'_b and R_b will denote the moduli of E'_b and E_b , respectively.

5. The conditional distribution $P(\Phi | R_{b_1}, R_{b_2}, R_{b_3}, \mathbf{u})$ in non-centrosymmetric space groups

If no prior information is available besides atomicity then

$$\langle F_{b_1} F_{b_2} F_{b_3} \rangle = \sum_3(\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3). \tag{8}$$

The conditional probability of Φ given $R'_{b_1}, R'_{b_2}, R'_{b_3}$ is (see also Cochran, 1955)

$$P(\Phi | R'_{b_1}, R'_{b_2}, R'_{b_3}) \cong [2\pi I_0(G')]^{-1} \exp(G' \cos \Phi),$$

where

$$G' = \frac{2\langle F_{b_1} F_{b_2} F_{b_3} \rangle}{[\langle |F_{b_1}|^2 \rangle \langle |F_{b_2}|^2 \rangle \langle |F_{b_3}|^2 \rangle]^{1/2}} R'_{b_1} R'_{b_2} R'_{b_3} \tag{9} \\ \cong 2\sigma_3 \sigma_2^{-3/2} R'_{b_1} R'_{b_2} R'_{b_3}.$$

For equal atoms $G' = 2R'_{b_1} R'_{b_2} R'_{b_3} / N^{1/2}$.

If the interatomic vector \mathbf{u} is *a priori* known then

$$\langle F_{b_1} F_{b_2} F_{b_3} | \mathbf{u} \rangle \\ = \left\langle \left[\sum_{\substack{j=1 \\ j \neq p,q}}^l f_j(\mathbf{h}_1) \sum_{s=1}^m \exp 2\pi i \mathbf{h}_1 \mathbf{C}_s \mathbf{r}_j \right. \right. \\ \left. \left. + f_p(\mathbf{h}_1) \sum_{s=1}^m \exp 2\pi i \mathbf{h}_1 \mathbf{C}_s \mathbf{r}_p \right. \right. \\ \left. \left. + f_q(\mathbf{h}_1) \exp 2\pi i \mathbf{h}_1 \mathbf{C}_s (\mathbf{r}_p + \mathbf{u}) \right] \right. \\ \times [\text{equivalent expression for } \mathbf{h}_2] \\ \left. \times [\text{equivalent expression for } \mathbf{h}_3] \right\rangle \\ = \sum_3(\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3) \\ + f_q(\mathbf{h}_1) f_p(\mathbf{h}_2) f_p(\mathbf{h}_3) \sum_{s=1}^m \exp 2\pi i \mathbf{h}_1 \mathbf{R}_s \mathbf{u} \\ + f_p(\mathbf{h}_1) f_q(\mathbf{h}_2) f_q(\mathbf{h}_3) \sum_{s=1}^m \exp -2\pi i \mathbf{h}_1 \mathbf{R}_s \mathbf{u} \\ + f_p(\mathbf{h}_1) f_q(\mathbf{h}_2) f_p(\mathbf{h}_3) \sum_{s=1}^m \exp 2\pi i \mathbf{h}_2 \mathbf{R}_s \mathbf{u} \\ + f_q(\mathbf{h}_1) f_p(\mathbf{h}_2) f_q(\mathbf{h}_3) \sum_{s=1}^m \exp -2\pi i \mathbf{h}_2 \mathbf{R}_s \mathbf{u} \\ + f_q(\mathbf{h}_1) f_q(\mathbf{h}_2) f_p(\mathbf{h}_3) \sum_{s=1}^m \exp -2\pi i \mathbf{h}_3 \mathbf{R}_s \mathbf{u} \\ + f_p(\mathbf{h}_1) f_p(\mathbf{h}_2) f_q(\mathbf{h}_3) \sum_{s=1}^m \exp +2\pi i \mathbf{h}_3 \mathbf{R}_s \mathbf{u} \\ = \left[\sum_3(\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3) + Q_1 \sum_{s=1}^m \cos 2\pi \mathbf{h}_1 \mathbf{R}_s \mathbf{u} \right. \\ \left. + Q_2 \sum_{s=1}^m \cos 2\pi \mathbf{h}_2 \mathbf{R}_s \mathbf{u} + Q_3 \sum_{s=1}^m \cos 2\pi \mathbf{h}_3 \mathbf{R}_s \mathbf{u} \right] \\ + i \left[T_1 \sum_{s=1}^m \sin 2\pi \mathbf{h}_1 \mathbf{R}_s \mathbf{u} + T_2 \sum_{s=1}^m \sin 2\pi \mathbf{h}_2 \mathbf{R}_s \mathbf{u} \right. \\ \left. + T_3 \sum_{s=1}^m \sin 2\pi \mathbf{h}_3 \mathbf{R}_s \mathbf{u} \right] \\ = M_R + iM_I = M \exp(i\theta),$$

where

$$M = (M_R^2 + M_I^2)^{1/2}, \\ \tan \varphi = M_I / M_R$$

and

$$\begin{aligned} Q_1 &= f_q(\mathbf{h}_1)f_p(\mathbf{h}_2)f_p(\mathbf{h}_3) + f_p(\mathbf{h}_1)f_q(\mathbf{h}_2)f_q(\mathbf{h}_3) \\ Q_2 &= f_p(\mathbf{h}_1)f_q(\mathbf{h}_2)f_p(\mathbf{h}_3) + f_q(\mathbf{h}_1)f_p(\mathbf{h}_2)f_q(\mathbf{h}_3) \\ Q_3 &= f_q(\mathbf{h}_1)f_q(\mathbf{h}_2)f_p(\mathbf{h}_3) + f_p(\mathbf{h}_1)f_p(\mathbf{h}_2)f_q(\mathbf{h}_3) \\ T_1 &= f_q(\mathbf{h}_1)f_p(\mathbf{h}_2)f_p(\mathbf{h}_3) - f_p(\mathbf{h}_1)f_q(\mathbf{h}_2)f_q(\mathbf{h}_3) \\ T_2 &= f_p(\mathbf{h}_1)f_q(\mathbf{h}_2)f_p(\mathbf{h}_3) - f_q(\mathbf{h}_1)f_p(\mathbf{h}_2)f_q(\mathbf{h}_3) \\ T_3 &= f_p(\mathbf{h}_1)f_p(\mathbf{h}_2)f_q(\mathbf{h}_3) - f_q(\mathbf{h}_1)f_q(\mathbf{h}_2)f_p(\mathbf{h}_3). \end{aligned}$$

$\langle F_{\mathbf{h}_1}F_{\mathbf{h}_2}F_{\mathbf{h}_3}|\mathbf{u} \rangle$ is in general a complex number the phase θ of which may lie anywhere between 0 and 2π . The conditional distribution $P(\Phi|R_{\mathbf{h}_1}, R_{\mathbf{h}_2}, R_{\mathbf{h}_3}, \mathbf{u})$ may then be written as

$$P(\Phi|R_{\mathbf{h}_1}, R_{\mathbf{h}_2}, R_{\mathbf{h}_3}, \mathbf{u}) \cong [2\pi I_0(G)]^{-1} \exp[G \cos(\Phi - \theta)] \quad (10)$$

where

$$G = \{M/[\langle |F_{\mathbf{h}_1}|^2|\mathbf{u} \rangle \langle |F_{\mathbf{h}_2}|^2|\mathbf{u} \rangle \langle |F_{\mathbf{h}_3}|^2|\mathbf{u} \rangle]^{1/2}\} \times 2R_{\mathbf{h}_1}R_{\mathbf{h}_2}R_{\mathbf{h}_3} \quad (11)$$

is a positive term and $\langle |F_{\mathbf{h}_i}|^2|\mathbf{u} \rangle$ is given by (6).

Owing to the prior information, the triplet phase is no longer expected to be zero. Such a result is not a surprise: indeed, at the same time at which we assume that atoms with scattering factors f_p and f_q lie at \mathbf{r}_p and \mathbf{r}_q , respectively, then we also fix the enantiomorph (associating f_p with the position \mathbf{r}_q and f_q with \mathbf{r}_p should fix the other enantiomorph). If $f_p \equiv f_q$ then $T_1 = T_2 = T_3 = 0$, $M_i = 0$ and the enantiomorph is no longer fixed by prior information. Also, for X-ray diffraction, if $Z_p \equiv Z_q$ or $|Z_p - Z_q| \ll F_{000}$, then the enantiomorph is not defined for practical applications.

It is worthwhile mentioning that in $P1$ the prior information about the interatomic vector \mathbf{u} may be exploited in a more effective way than by the present theory. Indeed, in $P1$ the origin is free in any direction: thus, the first of the two atoms connected by \mathbf{u} may be arbitrarily located, and the second may be placed in position \mathbf{u} or $-\mathbf{u}$ (if $Z_p \neq Z_q$ the enantiomorph is also fixed by any of the two choices). Such a two-atom fragment may be considered as a well located molecular fragment: recovery of the complete crystal structure may then be accomplished by techniques such as those described by Beurskens, Prick, Doesburg & Gould (1979) or by Burla, Cascarano, Fares, Giacovazzo, Polidori & Spagna (1989).

6. The conditional distribution $P(\Phi|R_{\mathbf{h}_1}, R_{\mathbf{h}_2}, R_{\mathbf{h}_3}, \mathbf{u})$ in centrosymmetric space groups

Since Φ may assume only values 0 or π , we state the probability density for $\Phi = 0$. We obtain

$$P(\Phi = 0) = \frac{1}{2} + \frac{1}{2} \tanh G_c \quad (12)$$

where

$$\begin{aligned} G_c &= \{M/[\langle |F_{\mathbf{h}_1}|^2|\mathbf{u} \rangle \langle |F_{\mathbf{h}_2}|^2|\mathbf{u} \rangle \langle |F_{\mathbf{h}_3}|^2|\mathbf{u} \rangle]^{1/2}\} \\ &\quad \times R_{\mathbf{h}_1}R_{\mathbf{h}_2}R_{\mathbf{h}_3}, \\ M &= \sum_3 (\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3) + 2Q_1 \sum_{s=1}^{m/2} \cos 2\pi \mathbf{h}_1 \mathbf{R}_s \mathbf{u} \\ &\quad + 2Q_2 \sum_{s=1}^{m/2} \cos 2\pi \mathbf{h}_2 \mathbf{R}_s \mathbf{u} \\ &\quad + 2Q_3 \sum_{s=1}^{m/2} \cos 2\pi \mathbf{h}_3 \mathbf{R}_s \mathbf{u} \end{aligned} \quad (13)$$

and $\langle |F_{\mathbf{h}_i}|^2|\mathbf{u} \rangle$ is given by (7).

7. The average value of $|F_{\mathbf{h}}|^2$ when more interatomic vectors overlap on \mathbf{u}

Let us suppose that the Patterson peak \mathbf{u} is generated by overlapping of more symmetry-independent interatomic vectors

$$\mathbf{r}_{j_1} - \mathbf{r}_{j_2} = \mathbf{r}_{j_3} - \mathbf{r}_{j_4} = \dots = \mathbf{u}$$

and that the scattering factor of the atom in each position \mathbf{r}_j is *a priori* known. We will also assume that

$$j_1 \neq j_2 \neq j_3 \neq j_4 \neq \dots \quad (14)$$

Such conditions are not too restrictive and are frequently satisfied, at least approximately, in usual crystal structures. In Fig. 1(a) a geometrical molecular model is shown: each of the two types of Patterson peaks (say 1-2 and 13-14) is the sum of six overlapping interatomic vectors. It is easily seen that condition (14) is satisfied for each of the two peaks. Thus it seems worthwhile exploring how the prior information on such overlapped interatomic vectors modifies the Wilson statistics and the standard probabilistic formulae for the estimation of triplet invariants.

We stress the point that results obtained in the present paper cannot be applied to a Patterson peak for which condition (14) is not satisfied. For example, in Fig. 1(b) the condition is violated for the peak $\mathbf{u} = \mathbf{r}_1 - \mathbf{r}_2 = \mathbf{r}_4 - \mathbf{r}_1$ and similar other peaks. If such cases are recognized a larger amount of prior informa-

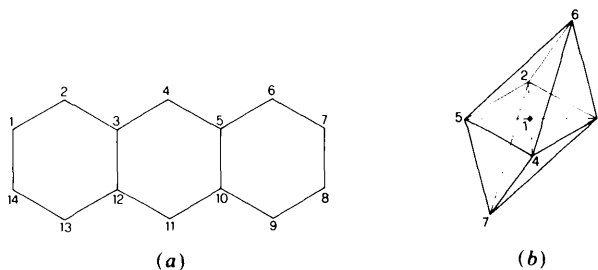


Fig. 1. Two geometrical patterns generating overlapping Patterson vectors.

tion is available in practice: for example, since $\mathbf{u} = \mathbf{r}_1 - \mathbf{r}_2 = \mathbf{r}_4 - \mathbf{r}_1$, the additional interatomic vector $\mathbf{r}_4 - \mathbf{r}_2$ remains completely defined. Special treatments will be devised for such cases in the following paper of this series.

The average of $|F_{\mathbf{h}}|^2$ when more interatomic vectors overlap in \mathbf{u} may be obtained by a generalization of the procedure described in § 3. The conclusive formulas are:

non-centrosymmetrical space groups

$$\langle |F_{\mathbf{h}}|^2 | \mathbf{u} \rangle = \varepsilon_{\mathbf{h}} \left\{ \sum_{\mathbf{N}} (\mathbf{h}) + 2 \left[\sum_{s=1}^m \cos 2\pi \mathbf{h} \mathbf{R}_s \mathbf{u} \right] \times \left[\sum'_{(p,q)} f_p(\mathbf{h}) f_q(\mathbf{h}) \right] \right\}; \quad (15)$$

centrosymmetrical space groups

$$\langle |F_{\mathbf{h}}|^2 | \mathbf{u} \rangle = \varepsilon_{\mathbf{h}} \left\{ \sum_{\mathbf{N}} (\mathbf{h}) + 4 \left[\sum_{s=1}^{m/2} \cos 2\pi \mathbf{h} \mathbf{R}_s \mathbf{u} \right] \times \left[\sum'_{(p,q)} f_p(\mathbf{h}) f_q(\mathbf{h}) \right] \right\}. \quad (16)$$

In (15) and (16) the primed summation goes over the pairs of symmetry-independent atoms referred to \mathbf{u} .

Because of the geometrical conditions at the basis of our statistical model, it will never occur that

$$\left| \left[\sum'_{(p,q)} f_p(\mathbf{h}) f_q(\mathbf{h}) \right] \left[\sum_{s=1}^m \cos 2\pi \mathbf{h} \mathbf{R}_s \mathbf{u} \right] \right| > \frac{1}{2} \sum_{\mathbf{N}} (\mathbf{h}),$$

otherwise negative values of $\langle |F_{\mathbf{h}}|^2 | \mathbf{u} \rangle$ could be obtained when $\mathbf{h} \mathbf{R}_s \mathbf{u}$ is a semi-integer value for any s .

8. The conditional distribution $P(\Phi | \mathbf{R}_{\mathbf{h}_1}, \mathbf{R}_{\mathbf{h}_2}, \mathbf{R}_{\mathbf{h}_3}, \mathbf{u})$ when more interatomic vectors overlap on \mathbf{u}

If some prior information is available which permits the association of specific scattering factors to each atomic position \mathbf{r}_j involved in the formation of the interatomic vector \mathbf{u} then the distribution (10) still holds, provided

$$\begin{aligned} Q_1 &= \sum'_{(p,q)} [f_q(\mathbf{h}_1) f_p(\mathbf{h}_2) f_p(\mathbf{h}_3) + f_p(\mathbf{h}_1) f_q(\mathbf{h}_2) f_q(\mathbf{h}_3)] \\ Q_2 &= \sum'_{(p,q)} [f_p(\mathbf{h}_1) f_q(\mathbf{h}_2) f_p(\mathbf{h}_3) + f_q(\mathbf{h}_1) f_p(\mathbf{h}_2) f_q(\mathbf{h}_3)] \\ Q_3 &= \sum'_{(p,q)} [f_q(\mathbf{h}_1) f_q(\mathbf{h}_2) f_p(\mathbf{h}_3) + f_p(\mathbf{h}_1) f_p(\mathbf{h}_2) f_q(\mathbf{h}_3)] \\ T_1 &= \sum'_{(p,q)} [f_q(\mathbf{h}_1) f_p(\mathbf{h}_2) f_p(\mathbf{h}_3) - f_p(\mathbf{h}_1) f_q(\mathbf{h}_2) f_q(\mathbf{h}_3)] \\ T_2 &= \sum'_{(p,q)} [f_p(\mathbf{h}_1) f_q(\mathbf{h}_2) f_p(\mathbf{h}_3) - f_q(\mathbf{h}_1) f_p(\mathbf{h}_2) f_q(\mathbf{h}_3)] \\ T_3 &= \sum'_{(p,q)} [f_p(\mathbf{h}_1) f_p(\mathbf{h}_2) f_q(\mathbf{h}_3) - f_q(\mathbf{h}_1) f_q(\mathbf{h}_2) f_p(\mathbf{h}_3)]. \end{aligned} \quad (17)$$

Unfortunately, inspection of a Patterson map usually provides estimates of the terms Q_j but not of the terms T_j . In this case, the most sensitive choice is to assume $T_j = 0$ for $j = 1, 2, 3$ and to renounce the enantiomorph

definition: then

$$P(\Phi | \mathbf{R}_{\mathbf{h}_1}, \mathbf{R}_{\mathbf{h}_2}, \mathbf{R}_{\mathbf{h}_3}, \mathbf{u}) \equiv [2\pi I_0(G)]^{-1} \exp(G \cos \Phi), \quad (18)$$

where

$$G = \{M / [\langle |F_{\mathbf{h}_1}|^2 | \mathbf{u} \rangle \langle |F_{\mathbf{h}_2}|^2 | \mathbf{u} \rangle \langle |F_{\mathbf{h}_3}|^2 | \mathbf{u} \rangle]^{1/2}\} 2 \mathbf{R}_{\mathbf{h}_1} \mathbf{R}_{\mathbf{h}_2} \mathbf{R}_{\mathbf{h}_3}, \quad (19)$$

and

$$\begin{aligned} M &= \sum_3 (\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3) + Q_1 \sum_{s=1}^m \cos 2\pi \mathbf{h}_1 \mathbf{R}_s \mathbf{u} \\ &+ Q_2 \sum_{s=1}^m \cos 2\pi \mathbf{h}_2 \mathbf{R}_s \mathbf{u} \\ &+ Q_3 \sum_{s=1}^m \cos 2\pi \mathbf{h}_3 \mathbf{R}_s \mathbf{u}. \end{aligned} \quad (20)$$

Q_1, Q_2, Q_3 are defined according to (17).

By analogy, when more interatomic vectors overlap on \mathbf{u} in centrosymmetrical space groups, distribution (12) will exactly hold provided (16) for $\langle |F_{\mathbf{h}_1}|^2 | \mathbf{u} \rangle$ and (17) for $Q_i, i = 1, 2, 3$, are used.

9. The case in which $2\mathbf{u}$ is a lattice vector

Let us suppose that: (a) atoms related by vector \mathbf{u} are all equal; (b) $2\mathbf{u}$ is a lattice vector [i.e. $\mathbf{u} = \mathbf{a}/2$ or $(\mathbf{a} + \mathbf{b})/2$ or ...]. Then an ideal pseudotranslational symmetry of order $n = 2$ exists: in this case the present theory overlaps with part of that described by Cascarano, Giacobozzo & Luić (1985, 1987; from now on papers I and II, respectively). In spite of the quite different notations the conclusive formulas for $n = 2$ should coincide. In order to check that, we first examine (I.5):

$$\langle |F_{\mathbf{h}}|^2 \rangle = \varepsilon_{\mathbf{h}} [\delta_{\mathbf{h}} \sum_{t_p} (\mathbf{h}) + \sum_q (\mathbf{h})] \quad (I.5)$$

where $\sum_{t_p} (\mathbf{h}) = \sum_{j=1}^{t_p} f_j^2$ and $\sum_q (\mathbf{h}) = \sum_{j=1}^q f_j^2$. t_p is the number of independent atoms that generate all the p atoms related by pseudotranslational symmetry when the pseudotranslation vector \mathbf{u} and the symmetry operators $\mathbf{C}_s, s = 1, \dots, m$, are applied. q is the number of atoms (symmetry equivalent included) whose positions are not related by any pseudotranslation vector.

When $n = 2, \delta_{\mathbf{h}}$ is defined by

$$\delta_{\mathbf{h}} = \sum_{s=1}^m \sin^2 2\pi \mathbf{h} \mathbf{R}_s \mathbf{u} / \sin^2 \pi \mathbf{h} \mathbf{R}_s \mathbf{u}. \quad (21)$$

On introducing in (21) the trigonometric formulae

$$\sin^2 x = 2 \sin x \cos x \quad \text{and} \quad 2 \cos^2 x = 1 + \cos 2x,$$

(I.5) reduces to

$$\langle |F_{\mathbf{h}}|^2 \rangle = \varepsilon_{\mathbf{h}} \left[2m \sum_t (\mathbf{h}) + 2 \sum_t (\mathbf{h}) \left(\sum_{s=1}^m \cos 2\pi \mathbf{h} \mathbf{R}_s \mathbf{u} \right) + \sum_q (\mathbf{h}) \right]. \quad (22)$$

Since $2m \sum_{i_p} = \sum_p$ and $\sum_p + \sum_q = \sum_N$, (22) may be written as

$$\langle |F_{\mathbf{h}}|^2 \rangle = \varepsilon_{\mathbf{h}} \left[\sum_N(\mathbf{h}) + 2 \sum_t(\mathbf{h}) \left(\sum_{s=1}^m \cos 2\pi \mathbf{h} \mathbf{R}_s \mathbf{u} \right) \right],$$

which coincides with our equation (15) [by hypothesis we must assume in (15) $f_q(\mathbf{h}) \equiv f_p(\mathbf{h})$].

When only a pseudotranslational vector of order two exists, the reliability parameter for a triplet phase in non-centrosymmetric space groups was given by (II.3):

$$G_{ps} = 2R_{\mathbf{h}_1} R_{\mathbf{h}_2} R_{\mathbf{h}_3} [\langle |F_{\mathbf{h}_1}|^2 \rangle \langle |F_{\mathbf{h}_2}|^2 \rangle \langle |F_{\mathbf{h}_3}|^2 \rangle]^{-1} M_{ps},$$

where

$$M_{ps} = \frac{1}{2m} \left[\sum_3(\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3) \right]_p \left(\sum_{s=1}^m \tau_s \right) + \left[\sum_3(\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3) \right]_q$$

and

$$\tau_s = \frac{\sin(2\pi \mathbf{h}_1 \mathbf{R}_s \mathbf{u}) \sin(2\pi \mathbf{h}_2 \mathbf{R}_s \mathbf{u}) \sin(2\pi \mathbf{h}_3 \mathbf{R}_s \mathbf{u})}{\sin(\pi \mathbf{h}_1 \mathbf{R}_s \mathbf{u}) \sin(\pi \mathbf{h}_2 \mathbf{R}_s \mathbf{u}) \sin(\pi \mathbf{h}_3 \mathbf{R}_s \mathbf{u})}.$$

The indices p and q to $\sum_3(\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3)$ warns the reader that the summation is extended only to p and q atoms, respectively.

Because of the result just obtained, G_{ps} will coincide with our G as defined by (19) if $M_{ps} \equiv M$. By application of trivial trigonometric formulae we find

$$\begin{aligned} \tau_s &= 8 \cos(\pi \mathbf{h}_1 \mathbf{R}_s \mathbf{u}) \cos(\pi \mathbf{h}_2 \mathbf{R}_s \mathbf{u}) \cos(\pi \mathbf{h}_3 \mathbf{R}_s \mathbf{u}) \\ &= 2(1 + \cos 2\pi \mathbf{h}_1 \mathbf{R}_s \mathbf{u} \\ &\quad + \cos 2\pi \mathbf{h}_2 \mathbf{R}_s \mathbf{u} + \cos 2\pi \mathbf{h}_3 \mathbf{R}_s \mathbf{u}) \end{aligned}$$

so that

$$\begin{aligned} M_{ps} &= \left[\sum_3(\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3) \right]_p \left[1 + \frac{1}{m} \left(\sum_{s=1}^m \cos 2\pi \mathbf{h}_1 \mathbf{R}_s \mathbf{u} \right. \right. \\ &\quad \left. \left. + \sum_{s=1}^m \cos 2\pi \mathbf{h}_2 \mathbf{R}_s \mathbf{u} \right. \right. \\ &\quad \left. \left. + \sum_{s=1}^m \cos 2\pi \mathbf{h}_3 \mathbf{R}_s \mathbf{u} \right) \right] + \left[\sum_3(\mathbf{h}_1 + \mathbf{h}_2 + \mathbf{h}_3) \right]_q. \end{aligned}$$

We observe now that

$$\begin{aligned} (a) \quad & \left[\sum_3(\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3) \right]_p + \left[\sum_3(\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3) \right]_q \\ &= \sum_3(\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3) \\ (b) \quad & \left[\sum_3(\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3) \right]_p / m = 2 \left[\sum_3(\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3) \right]_{t_p} \\ & \equiv 2Q_1 = 2Q_2 = 2Q_3. \end{aligned}$$

Finally, $M_{ps} \equiv M$ and $G_{ps} \equiv G$.

10. Triplets and Harker sections

If \mathbf{u} is a Harker peak the present mathematical model cannot be applied. It is, however, anticipated that Harker sections can provide additional information for triplet estimation. In particular, Harker sections will prove highly efficient for the estimation of one-phase seminvariants of first rank; this is in accordance with some recent algebraic results (Ardito, Cascarano, Giacobazzo & Luić, 1985; Cascarano, Giacobazzo, Luić, Pifferi & Spagna, 1987).

11. Concluding remarks

A Patterson map may be analyzed in order to obtain different types of information: the simplest consists of the knowledge of one or more interatomic peaks. More elaborated information is the identification of Harker peaks (false and true Harker peaks cannot always be distinguished). Additional work is required when the Patterson map is used in order to define the orientation of a molecular fragment and/or its location.

In this paper, the theoretical bases have been given in order to integrate the simplest type of Patterson information, the knowledge of a non-Harker peak, in the method of joint probability distributions of structure factors. It has been shown that the reliability of a triplet phase estimate may be strongly affected if the location and the intensity of a Patterson peak are used as prior information in the probabilistic approach. When $2\mathbf{u}$ is a lattice vector, the probabilistic formulae coincide with those obtained by Cascarano, Giacobazzo & Luić (1985, 1987) for the case in which only one pseudotranslational vector of order 2 exists.

Readers may think that the usefulness of our conclusive formula (20) is limited by the facts that not only may the interatomic peak involve unequal atoms but also that it seems necessary to know which atoms are involved in producing the Patterson peak. Actually that is not true. It will be shown in a paper by Altomare, Cascarano & Giacobazzo (1991) that estimates of the factors Q_1, Q_2, Q_3 in (20) may be expressed in terms of Patterson peak intensity.

It is also anticipated that a light-atom structure resistant to any attempt has been solved by application of such a theory. The Patterson map showed a very large peak (with intensity larger than half the origin peak) due to the superposition of several light-atom-light-atom interactions. Such a situation reduces the efficiency of traditional direct methods (indeed atoms are not uniformly distributed) but it is an ideal test for the present theory which can exploit prior information on the distribution of the atoms.

References

- ALLEGRA, G. (1979). *Acta Cryst.* **A35**, 213-220.
 ALTOMARE, A., CASCARANO, G. & GIACOVAZZO, C. (1991). In preparation.

- ARDITO, G., CASCARANO, G., GIACOVAZZO, C. & LUIĆ, M. (1985). *Z. Kristallogr.* **172**, 25-34.
- BEURSKENS, P. T., PRICK, A. J., DOESBURG, H. M. & GOULD, R. O. (1979). *Acta Cryst.* **A35**, 765-772.
- BROSIUS, J. (1985). *Acta Cryst.* **A41**, 613-617.
- BURLA, M. C., CASCARANO, G., FARES, V., GIACOVAZZO, C., POLIDORI, G. & SPAGNA, R. (1989). *Acta Cryst.* **A45**, 781-786.
- CASCARANO, G., GIACOVAZZO, C. & LUIĆ, M. (1985). *Acta Cryst.* **A41**, 544-551.
- CASCARANO, G., GIACOVAZZO, C. & LUIĆ, M. (1987). *Acta Cryst.* **A43**, 14-22.
- CASCARANO, G., GIACOVAZZO, C., LUIĆ, M., PIFFERI, A. & SPAGNA, R. (1987). *Z. Kristallogr.* **179**, 113-125.
- COCHRAN, W. (1952). *Acta Cryst.* **5**, 65-67.
- COCHRAN, W. (1955). *Acta Cryst.* **8**, 473-478.
- ELLER, G. VON (1973). *Acta Cryst.* **A29**, 63-67.
- GIACOVAZZO, C. (1980). *Direct Methods in Crystallography*. London: Academic Press.
- HAUPTMAN, H. & KARLE, J. (1962). *Acta Cryst.* **15**, 547-550.
- HEINERMAN, J. J. L., KRABBENDAM, H. & KROON, J. (1975). *Acta Cryst.* **A31**, 731-736.
- HEINERMAN, J. J. L., KRABBENDAM, H., KROON, J. & SPEK, A. L. (1978). *Acta Cryst.* **A34**, 447-450.
- HOPPE, W. (1957). *Acta Cryst.* **10**, 751.
- HOPPE, W. (1963). *Acta Cryst.* **16**, 1056-1057.
- KARLE, J. & HAUPTMAN, H. (1956). *Acta Cryst.* **9**, 635-651.
- KRABBENDAM, H. & KROON, J. (1971). *Acta Cryst.* **A27**, 362-367.
- KROON, J. & KRABBENDAM, H. (1970). *Acta Cryst.* **B26**, 312-314.
- MAIN, P. & WOOLFSON, M. M. (1963). *Acta Cryst.* **16**, 1046-1051.
- PONTENAGEL, W. M. G. F., KRABBENDAM, H. & KROON, J. (1987). *Acta Cryst.* **A43**, 55-60.
- RIUS, J. & MIRAVITLLES (1989). *Acta Cryst.* **A45**, 490-494.
- STANLEY, E. (1979). *Acta Cryst.* **A35**, 966-970.
- VAUGHAN, P. A. (1958). *Acta Cryst.* **11**, 111-115.
- WILSON, A. J. C. (1949). *Acta Cryst.* **2**, 318-321.

Acta Cryst. (1991). **A47**, 263-267

Polytypism in PbI₂-Doped Dendritic Crystals of Cadmium Iodide

BY BINAY KUMAR AND G. C. TRIGUNAYAT

Department of Physics and Astrophysics, University of Delhi, Delhi 110007, India

(Received 29 May 1990; accepted 12 December 1990)

Abstract

Cadmium iodide and lead iodide have been highly purified by a horizontal zone refining technique. Dendritic single crystals of PbI₂-doped cadmium iodide have been grown from vapour phase in vacuum. The effect of doping on polytype formation and other structural characteristics has been studied by X-ray diffraction. The crystals have shown the formation of rhombohedral polytype 12R in about 50% of cases, along with the common polytype 4H of cadmium iodide. This is in sharp contrast to earlier findings for undoped dendritic crystals which exclusively contained the polytype 4H. Further, unlike the undoped crystals, some of the doped crystals also show streaking and arcing on their X-ray photographs. The observed structural changes may be governed by both thermodynamical considerations and the kinetics of crystal growth and the observed streaking and arcing result from internal stresses built up around the relatively large Pb²⁺ ions in the host CdI₂ structure.

Introduction

A vast amount of experimental and theoretical work has been carried out on the polytypism of crystals, particularly in the last three decades. The crystals studied have had several different habits, *viz* plates, needles, prisms, pyramids *etc.*; the crystals of strongly polytypic materials like CdI₂, SiC and ZnS are usually

plate shaped. No work has been done on the polytypism of dendritic crystals. In fact, no attempts have previously been made to grow dendritic crystals of a polytypic material. However, it is now realized that the formation of polytypes is also related to crystal habit. For instance, GaSe crystals have been grown in many habits like plates, needles, thin ribbons, prisms, truncated pyramids *etc.*, but polytypism has been observed only in the needle-shaped crystals grown by sublimation (Terhell, 1983). Similarly, only needle-shaped AgI crystals show polytypism, although these crystals have been grown in other shapes, *viz* hexagonal plates, pyramids, prisms *etc.* (Prager, 1983; Cochrane, 1967).

We have successfully grown large dendritic single crystals of the richly polytypic material CdI₂ from the vapour phase, measuring up to nearly 10 mm long and nearly 5 mm wide. A preliminary report has been published (Kumar & Trigunayat, 1990). Besides the crystal habit, the polytypism of crystals is known to be affected by several other factors, of which an important one is the presence of impurities (Trigunayat, 1989). A study of the effect of the introduction of some selected impurities in highly pure melt-grown crystals of cadmium iodide has been made recently (Tyagi, 1988; Tyagi & Trigunayat, 1988, 1989). It was proposed to carry out a similar study on the role played by impurities in the formation of polytypes in the dendritic crystals grown by us. Accordingly, it was planned to dope CdI₂ crystals