Acta Cryst. (1978). A**34**, 562–574

The Estimation of the One-Phase Structure Seminvariants of First Rank by Means of Their First and Second Representations

By Carmelo Giacovazzo

Istituto di Mineralogia, Palazzo Ateneo, Università, 70121 *Bari, Italy*

(*Received* 30 *December* 1977; *accepted* 27 *February* 1978)

The estimation of the one-phase structure seminvariants of first rank is carried out in any centrosymmetric or non-centrosymmetric space group. Representations theory [Giacovazzo (1977), *Acta Cryst.* A33, 933–944] is suitably associated to the joint probability distribution method: the diffraction magnitudes belonging to the first and second phasing shells of any seminvariant are exploited in order to give probabilistic estimates of the seminvariant cosines. Probabilistic formulae are also derived *via* the method of complementary invariants.

1. Introduction

One-phase structure seminvariants have been studied by several authors (Hauptman & Karle, 1953; Cochran & Woolfson, 1954; Klug, 1958). Their probabilistic result in $P\bar{1}$ is

$$P_+(E_{2\mathbf{h}}) \simeq 0{\cdot}5 + 0{\cdot}5 \tanh \left[\frac{|E_{2\mathbf{h}}|}{2\sqrt{N}} (E_{\mathbf{h}}^2 - 1)\right], \quad (1)$$

which may be obtained *via* the joint probability distribution function $P(E_{2\mathbf{h}},E_{\mathbf{h}})$. Unfortunately the frequency of failures suggests that it should not be used in the early stages of the direct procedures. Joint probability distributions more complex than $P(E_{2\mathbf{h}},E_{\mathbf{h}})$ were therefore studied in order to obtain more accurate estimates of the sign of $E_{2\mathbf{h}}$. For example, $P(E_{2\mathbf{h}}, E_{\mathbf{k}}, E_{\mathbf{h}+\mathbf{k}})$ was studied by Hauptman & Karle (1953) and by Cochran & Woolfson (1955); $P(E_{2\mathbf{h}}, E_{\mathbf{h}}, E_{\mathbf{k}}, E_{\mathbf{h}+\mathbf{k}})$ by Cochran (1954) and by Hauptman & Karle (1957); $P(E_{2\mathbf{h}}, E_{\mathbf{h}}, E_{\mathbf{k}}, E_{\mathbf{h}+\mathbf{k}}, E_{2\mathbf{h}+K})$ and $P(E_{2\mathbf{h}}, E_{\mathbf{h}}, E_{\mathbf{k}}, E_{\mathbf{h}-\mathbf{k}}, E_{2\mathbf{h}-\mathbf{k}}, E_{\mathbf{h}+\mathbf{k}})$ by Giacovazzo (1976*b*) and Giacovazzo (1975) respectively. The theory of representations (Giacovazzo, 1977*a*) has given the author new insight into probabilistic methods of obtaining accurate estimates of the phase invariants or seminvariants. This theory is able, for any universal structure invariant or structure seminvariant Φ, to arrange in a general way the set of the reflexions in a sequence of subsets, each contained in the succeeding one, whose order is that of the expected effectiveness (in the statistical sense) for the estimation of Φ. From each subset $\{B\}_n$, which was called a phasing shell of nth order for Φ, one is able to estimate a collection of structure invariants (denoted in the quoted papers as $\{\psi\}_n$) whose values differ from Φ by constants which arise because of the translational symmetry.

In order to state the aim of this paper we recall some definitions (Giacovazzo, 1977*a*).

(*a*) Let $\mathbf{C}_p = (\mathbf{R}_p,\mathbf{T}_p)$, $p = 1,\ldots, m$ denote the m symmetry operators ($\mathbf{R}_p$ rotation component, $\mathbf{T}_p$ translation component) of the actual space group and let us suppose that $\Phi = \varphi_{\mathbf{H}}$ is a one-phase structure seminvariant. If at least one phase $\varphi_{\mathbf{h}}$ and two symmetry operators $\mathbf{C}_p$ and $\mathbf{C}_q$ exist in principle ($|E_{\mathbf{h}}|$ may or may not be in the measurements) such that

$$\psi_1 = \varphi_{\mathbf{H}} - \varphi_{\mathbf{h}\mathbf{R}_p} + \varphi_{\mathbf{h}\mathbf{R}_q} \quad (2)$$

is a universal structure invariant, then $\varphi_{\mathbf{H}}$ is a structure seminvariant of first rank. The collection of the invariants (2) obtained when $\mathbf{h}$ ranges over reciprocal space and $\mathbf{R}_p,\mathbf{R}_q$ over the set of the rotation matrices constitutes the first representation of $\varphi_{\mathbf{H}}$ and is denoted by $\{\psi\}_1$. The first phasing shell $\{B\}_1$ is defined to be the collection of the distinct $|E|$'s associated with any $\psi_1 \in \{\psi\}_1$.

(*b*) If $\varphi_{\mathbf{H}}$ is a structure seminvariant for which (2) cannot be stated, then two phases $\varphi_{\mathbf{h}}$ and $\varphi_{\mathbf{l}}$ and four symmetry operators exist in principle ($|E_{\mathbf{h}}|$ and $|E_{\mathbf{l}}|$ may or may not be in the measurements) such that

$$\psi_1 = \varphi_{\mathbf{H}} - \varphi_{\mathbf{h}\mathbf{R}_p} + \varphi_{\mathbf{h}\mathbf{R}_q} - \varphi_{\mathbf{l}\mathbf{R}_i} + \varphi_{\mathbf{l}\mathbf{R}_j} \quad (3)$$

is a universal structure invariant. $\varphi_{\mathbf{H}}$ is then a structure seminvariant of second rank. For example, φ_{eee} is a structure seminvariant of second rank in all the space groups belonging to the symmetry class 222.

The first aim of this paper is to estimate in any space group the one-phase seminvariants of first rank by means of their first and second representations. We recall in this connexion that the second representation of a one-phase structure seminvariant of first rank is the collection of the invariants

$$\psi_2 = \psi_1 + \varphi_{\mathbf{k}} - \varphi_{\mathbf{k}}$$

which arise when ψ_1 varies within $\{\psi\}_1$ and $\mathbf{k}$ over reciprocal space. Any ψ_2 is a special quintet. The collection of the distinct basis and cross-magnitudes associated with various ψ_2's constitutes the second phasing shell $\{B\}_2$ of $\varphi_{\mathbf{H}}$.

2. The mathematical approach

The mathematical device of joint probability distribution functions will be used. We assume that the reciprocal vectors are fixed and that the atomic coordinates are the primitive random variables. We suppose that a crystal structure consists of N identical atoms in the unit cell and that m is the order of the space group: $t = N/m$ is the number of atoms in the asymmetric unit. In the centrosymmetric space groups the characteristic function $C(u_1,u_2,\ldots,u_n)$ of the multivariate distribution $P(E_1,\ldots,E_n)$ is given by

$$C(u_1,\ldots,u_n) = \exp\left[\sum_{2\,v}^{\infty} \frac{S_v}{t^{v/2}}\right], \tag{4}$$

where

$$S_v = t \sum_{r+s+\ldots+w=v} \frac{\lambda_{rs\ldots w}}{r!s!\ldots w!} (iu_1)^r (iu_2)^2 \ldots (iu_n)^w.$$

$\lambda_{rs\ldots w}$ are the standardized cumulants of the distribution. In accordance with preceding papers (Giacovazzo, 1977*b*) the density function $P(E_1,\ldots,E_n)$ will be calculated *via* the Fourier transform of (4) or *via* its Gram–Charlier expansion (Klug, 1958)

$$\exp[-\tfrac{1}{2}(u_1^2+\ldots+u_n^2)] \left(1 + \frac{S_3}{t^{3/2}} + \frac{S_4}{t^2} + \frac{S_3^2}{2t^3} + \frac{S_3S_4}{t^{7/2}} + \frac{S_5}{t^{5/2}} + \frac{S_3^3}{t^{9/2}} + \ldots\right). \tag{5}$$

In non-centrosymmetric space groups we denote by $P(A_1,A_2,\ldots,A_n,B_1,B_2,\ldots,B_n)$ the joint probability distribution function of n normalized structure factors: A_j and B_j represent the real and imaginary parts respectively of the jth factor. The characteristic function of the distribution is

$$C(u_1,u_2,\ldots,u_n,v_1,\ldots,v_n) = \exp\left[\sum_{2\,v}^{\infty} \frac{S_v}{t^{v/2}}\right], \tag{6}$$

where $u_j,v_j,j = 1,\ldots,n$, are carrying variables associated with A_j and B_j respectively,

$$S_v = t \sum_{r+s+\ldots+w=v} \frac{1}{2^{v/2}} \frac{\lambda_{rs\ldots w}}{r!s!\ldots w!} (iu_1)^r (iu_2)^s \ldots (iv_n)^w,$$

$$\lambda_{rs\ldots w} = \frac{K_{rs\ldots w}}{m^{(r+s+\ldots+w)/2}}.$$

$K_{rs\ldots w}$ are cumulants of the distribution. After suitable change of variables

$$P(R_1,\ldots,R_n,\varphi_1,\ldots,\varphi_n) = \frac{1}{(2\pi)^{2n}} \times \int_0^\infty \ldots \int_0^\infty \int_0^{2\pi} \ldots \int_0^{2\pi} \exp\{-i[\rho_1 R_1 \cos(\psi_1 - \varphi_1) + \ldots + \rho_n R_n \cos(\psi_n - \varphi_n)]\} \times \exp\left\{[-\tfrac{1}{4}(\rho_1^2 + \ldots + \rho_n^2)] + \sum_{3\,v}^{\infty} \frac{S'_v}{t^{v/2}}\right\} \times R_1 R_2 \ldots R_n \rho_1 \rho_2 \ldots \rho_n \,\mathrm{d}\rho_1 \ldots \mathrm{d}\rho_n \,\mathrm{d}\psi_1 \ldots \mathrm{d}\psi_n, \tag{7}$$

where

$$S'_v = t \sum_{r+s+\ldots+w=v} \frac{\lambda_{rs\ldots w}}{r!s!\ldots w!} \times (i\rho_1 \cos \psi_1)^r (i\rho_2 \cos \psi_2)^s \ldots (i\rho_n \sin \psi_n)^w,$$

and R_j,φ_j are the modulus and the phase of the jth structure factor respectively. Apart from (7), $P(R_1,\ldots,\varphi_n)$ will be calculated *via* the Gram–Charlier expansion of (6) (see, in a different context, Giacovazzo, 1977*c*).

3. Algebraic properties of the one-phase structure seminvariants of first rank

Let $\varphi_{\mathbf{H}}$ be a structure seminvariant in $P\bar{1}$: then only one vector $\mathbf{h} = \mathbf{H}/2$ exists for which (2) holds. If the symmetry is higher than in $P1$ more vectors $\mathbf{h}$ may exist in principle (*i.e.* some of them may or may not be in the measurements) for which (2) holds. We will denote by $\{\mathbf{h}\}$ the set of vectors $\mathbf{h}$ which satisfy (2) and by $\{R_{\mathbf{h}}\}$ the corresponding set of observable magnitudes. We intend to exploit the algebraic properties of the system

$$\mathbf{h}(\mathbf{R}_p - \mathbf{R}_q) = \mathbf{H} \tag{8}$$

which arises from (2), in order to find $\{\mathbf{h}\}$. The result may be useful from a theoretical point of view (it may suggest the type of joint probability distribution which has to be studied) as well as from a practical point of view (it may allow a fast evaluation of $\varphi_{\mathbf{H}}$). We make use of some properties which explicitly are the following.

Property 1. $\varphi_{\mathbf{h}(\mathbf{R}_p-\mathbf{R}_q)}$ is a structure seminvariant of first rank for any space group which presents the rotation matrices $\mathbf{R}_p$ and $\mathbf{R}_q$, whatever $\mathbf{h}$ may be.

Proof. Since

$$\varphi_{\mathbf{h}(\mathbf{R}_p-\mathbf{R}_q)} - \varphi_{\mathbf{hR}_p} + \varphi_{\mathbf{hR}_q}$$

is a universal structure invariant, its value is a constant whatever the origin may be. As

$$\varphi_{\mathbf{hR}_p} - \varphi_{\mathbf{hR}_q} = 2\pi\mathbf{h}(\mathbf{T}_q - \mathbf{T}_p)$$

is a constant when the algebraic form of the structure factor has been fixed, $\varphi_{\mathbf{h}(\mathbf{R}_p-\mathbf{R}_q)}$ has the same property.

As $(\mathbf{R}_p - \mathbf{R}_q)$ is in general a singular matrix, some of the algebraic properties of (8) may be described by introducing the concept of the reflexive generalized inverse of a matrix.

Definition. If $\mathbf{A}$ is a $m \times n$ matrix, a $n \times m$ matrix $\mathbf{A}^*$ is said to be a reflexive generalized inverse of $\mathbf{A}$ provided $\mathbf{AA}^*\mathbf{A} = \mathbf{A}$ and $\mathbf{A}^*\mathbf{AA}^* = \mathbf{A}^*$.

Property 2. A system of linear equations

$$\mathbf{Ax} = \mathbf{b} \tag{9}$$

has a solution if and only if $\mathbf{AA}^*\mathbf{b} = \mathbf{b}$. Furthermore if it has a solution then

$$\mathbf{x} = \mathbf{A}^*\mathbf{b} + (\mathbf{I} - \mathbf{A}^*\mathbf{A})\mathbf{z} \tag{10}$$

where $\mathbf{z}$ is an arbitrary vector.

Property 3. In (9) $\mathbf{A} = (\mathbf{R}_p - \mathbf{R}_q)^+$ and $\mathbf{b} = \mathbf{H}$ are integral matrix and vectors respectively ($\mathbf{R}^+$ is the transpose of $\mathbf{R}$): furthermore, we are interested only in the integral solutions $\mathbf{h}$. We use then a theorem of Hurt & Waid (1970) for diophantine systems, according to which if $\mathbf{A}$ and $\mathbf{b}$ are integral, (9) has an integral solution if and only if

$$\mathbf{A}^*\mathbf{b} \equiv 0(\text{mod } 1). \tag{11}$$

In that case the general integral solution of (9) is given by (10), where $\mathbf{z}$ is an arbitrary integer vector.

The above-mentioned theorems allows us to derive, for a given structure seminvariant of first rank $\varphi_{\mathbf{H}}$, the matrices $\mathbf{A}_{pq} = (\mathbf{R}_p - \mathbf{R}_q)^+$ which satisfy (8) and, for every $\mathbf{A}_{pq}$, the general integral solution $\mathbf{h}$. We note that the number of matrices $\mathbf{A}_{pq}$ which, for fixed $\mathbf{H}$, make (8) consistent may be larger than unity. For example, let in $P2_12_12_1$

$$\mathbf{R}_1 = \mathbf{I};\quad \mathbf{R}_2 = \begin{vmatrix} 1 & 0 & 0 \\ 0 & \bar{1} & 0 \\ 0 & 0 & \bar{1} \end{vmatrix};$$

$$\mathbf{R}_3 = \begin{vmatrix} \bar{1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \bar{1} \end{vmatrix};\quad \mathbf{R}_4 = \begin{vmatrix} \bar{1} & 0 & 0 \\ 0 & \bar{1} & 0 \\ 0 & 0 & 1 \end{vmatrix}$$

and $\mathbf{H} = (800)$. Then all the matrices

$$\mathbf{A}_{13}, \mathbf{A}_{14}, \mathbf{A}_{23}, \mathbf{A}_{24}, \mathbf{A}_{31}, \mathbf{A}_{41}, \mathbf{A}_{32}, \mathbf{A}_{42}$$

make (8) consistent. The only independent solutions are

$$\mathbf{h} = (4k0) \quad \text{and} \quad \mathbf{h} = (40l)$$

obtained by means of $\mathbf{A}_{13}$ and $\mathbf{A}_{14}$ respectively. Obviously k and l are free integers.

Property 4. Let us state the following identities:

$$\begin{aligned} \mathbf{H} &= \mathbf{h}(\mathbf{R}_p - \mathbf{R}_q) = \mathbf{h}(\mathbf{I} - \mathbf{R}_q\mathbf{R}_p^{-1})\mathbf{R}_p \\ &= \mathbf{h}(\mathbf{I} - \mathbf{R}_n)\mathbf{R}_p = \mathbf{H}'\mathbf{R}_p, \end{aligned} \tag{12}$$

where

$$\mathbf{R}_n = \mathbf{R}_q\mathbf{R}_p^{-1} \quad \text{and} \quad \mathbf{H}' = \mathbf{H}\mathbf{R}_p^{-1}. \tag{13}$$

(12) tells us that if $\mathbf{h}$, for fixed $\mathbf{H}$, $\mathbf{R}_p$ and $\mathbf{R}_q$, is a solution of (8), then it is also a solution of

$$\mathbf{H}' = \mathbf{h}(\mathbf{I} - \mathbf{R}_n) \tag{8'}$$

and *vice versa*. From (13) $E_{\mathbf{H}'}$ is a reflexion symmetry-equivalent to $E_{\mathbf{H}}$. That allows us to simplify our notation: from now on, without any loss of generality, we will denote by $E_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)}$ the more general expression of a one-phase structure seminvariant of first rank.

Property 5. If $\mathbf{h}$ is a solution of (8′), $\mathbf{h} + \mathbf{k}$ is also provided $\mathbf{k}'(\mathbf{I} - \mathbf{R}_n) = 0$.

The proof is trivial. The theorem allows one to construct in a simple way $\{\mathbf{h}\}$ from any element $\mathbf{h}$.

The following property may be useful in space groups with symmetry higher than orthorhombic.

Property 6. If $\mathbf{h}$ satisfies $\mathbf{h}(\mathbf{I} - \mathbf{R}_n) = \mathbf{H}$ then $\mathbf{h}' = -\mathbf{h}\mathbf{R}_n$ satisfies $\mathbf{h}'(\mathbf{I} - \mathbf{R}_n^{-1}) = \mathbf{H}$ and *vice versa.*

Proof. $-\mathbf{h}\mathbf{R}_n(\mathbf{I} - \mathbf{R}_n^{-1}) = -\mathbf{h}(\mathbf{R}_n - \mathbf{I}) = \mathbf{h}(\mathbf{I} - \mathbf{R}_n) = \mathbf{H}$. This property tells us that the sets $\{|E_{\mathbf{h}}|\}$ and $\{|E_{\mathbf{h}'}|\}$ are symmetry-related (*i.e.* any element $E_{\mathbf{h}}$ is symmetry-equivalent to any $E_{\mathbf{h}'}$). Therefore (8′) needs to be exploited only for one between $\mathbf{R}_n$ and $\mathbf{R}_n^{-1}$.

In $P3$, for example, let

$$\mathbf{R}_1 = \mathbf{I};\quad \mathbf{R}_2 = \begin{vmatrix} 0 & \bar{1} & 0 \\ 1 & \bar{1} & 0 \\ 0 & 0 & 1 \end{vmatrix};\quad \mathbf{R}_3 = \begin{vmatrix} \bar{1} & 1 & 0 \\ \bar{1} & 0 & 0 \\ 0 & 0 & 1 \end{vmatrix};$$

where $\mathbf{R}_2 = \mathbf{R}_3^{-1}$.

All the seminvariants are of first rank and they satisfy the condition $(H - K, L) \equiv 0$ mod(3,0). Fixing $\mathbf{H} = (470)$ we obtain $\mathbf{h} = (51l)$ from $\mathbf{h}(\mathbf{I} - \mathbf{R}_2) = \mathbf{H}$, and $\mathbf{h} = (\bar{1}6l)$ from $\mathbf{h}(\mathbf{I} - \mathbf{R}_3) = \mathbf{H}$, in accordance with the fixed property.

4. The expected sign of $E_{\mathbf{H}} = E_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)}$ from its first representation in the centrosymmetric space groups

The idea of representations suggests the study of the joint probability distribution $P(E_{\mathbf{H}},\{R_{\mathbf{h}}\})$ from which the conditional probability $P(\varphi_{\mathbf{H}}|R_{\mathbf{H}},\{R_{\mathbf{h}}\})$ may in principle be derived. For example, in $P\bar{1}$ (8′) is satisfied only if

$$\mathbf{H} \equiv 0 \text{ mod}(2,2,2).$$

Then (10) gives $\{\mathbf{h}\} = \mathbf{h} = H/2$. The distribution $P(E_{2\mathbf{h}},E_{\mathbf{h}})$ is therefore suggested which, according to Cochran & Woolfson (1955), leads to (1).

In centrosymmetric space groups with crystal symmetry higher than $P\bar{1}$ one obtains (Naya, Nitta & Oda, 1964; Giacovazzo, 1974) *via* the Gram–Charlier expansion of (4),

$$P_+(E_{\mathbf{H}}) \simeq 0{\cdot}5 + 0{\cdot}5 \tanh\left[\frac{1}{2\sqrt{N}}\,|E_{\mathbf{H}}| \times \sum_n{}' \sum_j W_{\mathbf{H},\mathbf{h}_j}(E^2_{\mathbf{h}_j} - 1)(-1)^{2\mathbf{h}_j\mathbf{T}_n}\right], \tag{14}$$

where: (*a*) the first summation, in accordance with (11), goes within the set of matrices $\mathbf{R}_n$ for which $(\mathbf{I} - \mathbf{R}_n^+)^* \mathbf{H} \equiv 0(\mathrm{mod}\ 1)$. Only the matrices $\mathbf{R}_n$ must be considered which give independent $\mathbf{h}$ solutions. (*b*) For each $\mathbf{R}_n$ the second summation is over the general integral solution given by (10). (*c*) $W_{\mathbf{H},\mathbf{h}_j}$, whose value depends on the statistical nature of the reflexions, is a weight which is defined in Appendix *A*.

The Fourier transform of the exponential form of the characteristic function gives in $P\bar{1}$ (see Appendix *B*)

$$P_+(E_{\mathbf{H}}) \simeq \frac{P_+^0}{P_+^0 + P_-^0}, \tag{15}$$

where

$$P_\pm^0 = (f^\pm)^{-1/2} \exp(-R_{\mathbf{h}}^2/2f^\pm), \tag{16}$$

$$f^\pm = 1 \pm R_{\mathbf{H}}/\sqrt{N}.$$

(15) may usefully be compared with (1). From the power series $(1+x)^q \simeq 1 + qx + \ldots$, (16) becomes

$$(1 \mp R_{\mathbf{H}}/2\sqrt{N}) \exp[-R_{\mathbf{h}}^2(1 \mp R_{\mathbf{H}}/\sqrt{N})/2].$$

Putting $1 \pm x \simeq \exp(\pm x)$ (15) reduces to

$$P_+(E_h) \simeq \frac{\exp[R_{\mathbf{H}}(R_{\mathbf{h}}^2-1)/2\sqrt{N}]}{\exp[R_{\mathbf{H}}(R_{\mathbf{h}}^2-1)/2\sqrt{N}] + \exp[-R_{\mathbf{H}}(R_{\mathbf{h}}^2-1)/2\sqrt{N}]}$$

which coincides with (1). A numerical comparison between (1) and (15) is shown in Fig. 1 which shows that in many cases (1) is a useful approximation of (15). If the actual space group presents symmetry higher than $P\bar{1}$ more $\mathbf{h}$ exist in principle for a given $\mathbf{H}$. Denoting for every $\mathbf{h}_j$

$$P_{j,n}^\pm = (f_{j,n}^\pm)^{-1/2} \exp(-R_{\mathbf{h}}^2/2f_{j,n}^\pm),$$

$$f_{j,n}^\pm = 1 \pm \frac{W_{\mathbf{H},\mathbf{h}_j} R_{\mathbf{H}} (-1)^{2\mathbf{h}_j\mathbf{T}_n}}{\sqrt{N}},$$

the sign probability for $E_{\mathbf{H}}$ is given by

$$P_+(E_{\mathbf{H}}) = \frac{\Pi P_{j,n}^+}{\Pi P_{j,n}^+ + \Pi P_{j,n}^-}. \tag{17}$$

(17) may be compared with (14) in the same way as (15) with (1).

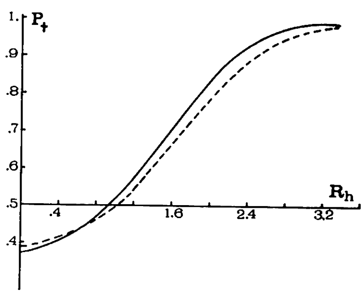

Fig. 1. The probability densities (1) (---) and (15) (——) as a function of $R_{\mathbf{h}}$ for fixed $R_{2\mathbf{h}} = 3{\cdot}0$ and $N = 50$.

5. The estimation of $\varphi_{\mathbf{H}} = \varphi_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)}$ from its first representation in non-centrosymmetric space groups

Using the Gram–Charlier expansion of (6) we obtain for $\mathbf{h} \in \{\mathbf{h}\}$,

$$P(\varphi_{\mathbf{H}}|R_{\mathbf{H}},R_{\mathbf{h}}) \simeq \frac{1}{L} \exp[G \cos(\varphi_{\mathbf{H}} - \Delta_{\mathbf{h},n})], \tag{18}$$

where

$$G = W_{\mathbf{H},\mathbf{h}} R_{\mathbf{H}} (R_{\mathbf{h}}^2 - 1)/\sqrt{N},$$

$$L = 2\pi I_0(G), \quad \Delta_{\mathbf{h},n} = 2\pi \mathbf{h}\mathbf{T}_n,$$

and $W_{\mathbf{H},\mathbf{h}}$ is a statistical weight justified in Appendix *C*. (18) holds when both $E_{\mathbf{H}}$ and $E_{\mathbf{h}}$ are non-centrosymmetric reflexions. For brevity its derivation is not described. (18) is a unimodal distribution which has its maximum at $\varphi_{\mathbf{H}} = \Delta_{\mathbf{h},n}$ if $G > 0$, at $\varphi_{\mathbf{H}} = \Delta_{\mathbf{h},n} + \pi$ if $G < 0$. If $G = 0$ (18) always equals $\frac{1}{2}\pi$. Furthermore,

$$\langle \cos \varphi_{\mathbf{H}} \rangle = \cos \Delta_{\mathbf{h},n} I_1(G)/I_0(G), \tag{19}$$

$$\mathrm{var}(\cos \varphi_{\mathbf{H}}) = \frac{1 + \cos 2\Delta_{\mathbf{h},n}}{2} \left[1 - \frac{I_1^2(G)}{I_0^2(G)}\right] - \frac{I_1(G)}{G I_0(G)} \cos 2\Delta_{\mathbf{h},n}, \tag{20}$$

$$\langle \sin \varphi_{\mathbf{H}} \rangle = \sin \Delta_{\mathbf{h},n} I_1(G)/I_0(G), \tag{21}$$

$$\mathrm{var}(\sin \varphi_{\mathbf{H}}) = \left(\frac{1 - \cos 2\Delta_{\mathbf{h},n}}{2}\right) \left[1 - \frac{I_1^2(G)}{I_0^2(G)}\right] + \frac{I_1(G)}{G I_0(G)} \cos 2\Delta_{\mathbf{h},n}. \tag{22}$$

The value of $\Delta_{\mathbf{h},n}$ may play a critical role in assigning the average and the variance values. In particular, we emphasize that the variance of $\cos \varphi_{\mathbf{H}}$ is not always smaller than that of $\sin \varphi_{\mathbf{H}}$ (*e.g.* if $\Delta_{\mathbf{h},n} = \pi/4$). Furthermore, the variance of $\varphi_{\mathbf{H}}$ depends on $|G|$:

$$\mathrm{var}(\varphi_{\mathbf{H}}) = \frac{\pi^2}{3} + [I_0(|G|)]^{-1} \sum_{1p}^{\infty} \left[\frac{I_{2p}(|G|)}{p^2}\right] - 4[I_0(|G|)]^{-1} \sum_{0p}^{\infty} \left[\frac{I_{2p+1}(|G|)}{(2p+1)^2}\right]. \tag{23}$$

If more R's belonging to the first phasing shell of $\varphi_{\mathbf{H}}$ are known then

$$P(\varphi_{\mathbf{H}}|R_{\mathbf{H}},\{R_{\mathbf{h}}\}) = \frac{\exp\left[\sum_n{}' \sum_j G_j \cos(\varphi_{\mathbf{H}} - \Delta_{j,n})\right]}{\int_{-\pi}^{\pi} \exp\left[\sum_n{}' \sum_j G_j \cos(\varphi_{\mathbf{H}} - \Delta_{j,n})\right] \mathrm{d}\varphi_{\mathbf{H}}} = \frac{1}{L'} \exp[A \cos(\varphi_{\mathbf{H}} - \theta)], \tag{24}$$

where

$$G_j = W_{\mathbf{H},\mathbf{h}_j} R_{\mathbf{H}} (R^2_{\mathbf{h}_j} - 1)/\sqrt{N},$$
$$\Delta_{j,n} = 2\pi \mathbf{h}_j \mathbf{T}_n,$$

$$A = \left[\left(\sum_n{}' \sum_j G_j \cos \Delta_{j,n}\right)^2 + \left(\sum_n \sum_j G_j \sin \Delta_{j,n}\right)^2\right]^{1/2},$$

$$\tan \theta \simeq \frac{\sum_n{}' \sum_j G_j \sin \Delta_{j,n}}{\sum_n{}' \sum_j G_j \cos \Delta_{j,n}}$$

$$L' = 2\pi I_0(A).$$

The same considerations described for (14) hold for the summations $\sum_n'$ and $\sum_j$.

(24) is a unimodal distribution which has its maximum at $\varphi_{\mathbf{H}} = \theta$. The variance of $\varphi_{\mathbf{H}}$ is given by (23) if $|G|$ is replaced by A. The expected cosine and sine values of $\varphi_{\mathbf{H}}$ and their variances according to (24) are again given by (19)–(22) if A replaces G and θ replaces $\Delta_{\mathbf{h},n}$. We note that (24) is formally different from (18) because always $A \geq 0$, whereas G may assume positive and negative values. However (25) and (19) give equivalent results when only one magnitude in $\{R_{\mathbf{h}}\}$ is known. In fact $\theta = \Delta_{j,n}$ if $G_j > 0$, $\theta = \Delta_{j,n} + \pi$ if $G < 0$. If the exponential form of (6) is directly used we obtain

$$P(\varphi_{\mathbf{H}}|R_{\mathbf{H}}, R_{\mathbf{h}}) \simeq \frac{\sqrt{d}}{2\pi I_0(q/d)} \exp\left(\frac{R^2_{\mathbf{h}}}{d}\right) \times [1 + a\cos(\varphi_{\mathbf{H}} - \Delta_{\mathbf{h},n})]^{-1} \times \exp\{-R^2_{\mathbf{h}}[1 + a\cos(\varphi_{\mathbf{H}} - \Delta_{\mathbf{h},n})]^{-1}\}, \quad (25)$$

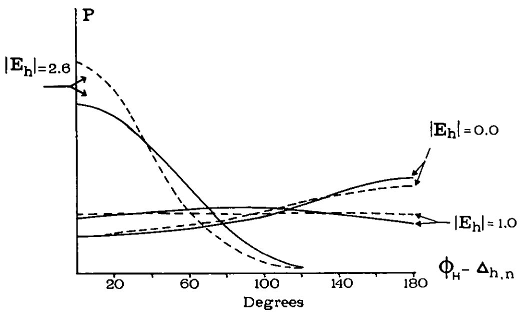

Fig. 2. The distributions (25) (——) and (18) (– – –) for the shown values of $R_{\mathbf{h}}$.

where

$$a = W_{\mathbf{H},\mathbf{h}} R_{\mathbf{H}}/\sqrt{N}, \quad q = R_{\mathbf{H}} R^2_{\mathbf{h}}/\sqrt{N},$$
$$d = 1 - R^2_{\mathbf{H}}/N, \quad \Delta_{\mathbf{h},n} = 2\pi \mathbf{h} \mathbf{T}_n.$$

The derivation of (25) is described in Appendix C.

When more $|E|$'s belonging to the first phasing shell of $\varphi_{\mathbf{H}}$ are known, the conditional probability of $\varphi_{\mathbf{H}}$ may be calculated by numerical techniques by means of

$$P(\varphi_{\mathbf{H}}|R_{\mathbf{H}}, \{R_{\mathbf{h}}\}) \simeq \frac{\prod_{j,n} Q_{j,n}}{\int_{-\pi}^{\pi} \left(\prod_{j,n} Q_{j,n}\right) \mathrm{d}\varphi_{\mathbf{H}}}, \quad (26)$$

where

$$Q_{j,n} = c^{-1}_{j,n} \exp(-R^2_j/c_{j,n}),$$
$$c_{j,n} = 1 + W_{\mathbf{H},\mathbf{h}_j} \frac{R_{\mathbf{H}}}{\sqrt{N}} \cos(\varphi_{\mathbf{H}} - \Delta_{j,n}).$$

The average values of $|\varphi_{\mathbf{H}}|$, the most probable value of $|\varphi_{\mathbf{H}}|$ ($|\varphi_{\mathbf{H}}|$mode), the average values of $\cos \varphi_{\mathbf{H}}$, $\sin \varphi_{\mathbf{H}}$ and their variances are also readily calculated from (26) by numerical techniques. We compare now (25) with (18): the comparison of (26) with (24) follows from there. In Fig. 2 the probability values according to (18) and (25) are shown for some values of $R_{\mathbf{h}}$ when $N = 50$ and $E_{\mathbf{H}} = 3{\cdot}0$. In Table 1 some expected cosines are calculated. We note: (a) the probability densities and the expected cosine values are significantly different only when $R_{\mathbf{H}}$ and $R_{\mathbf{h}}$ are large. Under these conditions (18) seems to overestimate the expected value of $\cos(\varphi_{\mathbf{H}} - \Delta_{\mathbf{h},n})$. ($b$) When $R_{\mathbf{h}} = 1$ (18) is a straight line whereas (25) is a rather flat curve with mode at $\varphi_{\mathbf{H}} = \Delta_{\mathbf{h},n} \pm 90°$.

We conclude that no large errors will arise in the direct procedures for phase determination if (18) is used instead of (25) [or (24) instead of (26)], thus saving computing time.

6. The expected signs of the centrosymmetric one-phase seminvariants in non-centrosymmetric space groups from the first representation

All one-phase seminvariants of first rank are centrosymmetric reflexions in up to orthorhombic space groups. Thus, it should be useful to calculate suitable sign probabilities for cases in which $E_{\mathbf{H}}$ is a centrosymmetric reflexion while $E_{\mathbf{h}}$ is not. If the Gram–Charlier expansion of (6) is used, (14) holds if $\cos \Delta_{j,n}$ replaces

Table 1. *Values of* $\langle\cos(\varphi_{\mathbf{H}} - \Delta_{\mathbf{h},n})\rangle$ *according to* (18) *and* (26) *for some values of* $R_{\mathbf{h}}$ *when* $R_{\mathbf{H}} = 3{\cdot}0$ *and* $N = 50$

$R_{\mathbf{h}}$	2·6	1·60	1·20	1·00	0·6	0·2	0·0
$\langle\cos(\varphi_{\mathbf{H}} - \Delta_{\mathbf{h},n})\rangle_{(18)}$	0·759	0·320	0·101	0·0	−0·144	−0·210	−0·217
$\langle\cos(\varphi_{\mathbf{H}} - \Delta_{\mathbf{h},n})\rangle_{(25)}$	0·700	0·337	0·23	0·022	−0·134	−0·213	−0·223

$(-1)^{2\mathbf{h}_j\mathbf{T}_n}$. If the exponential form of (6) is directly used when only a magnitude $R_{\mathbf{h}}$ is known we have for $E_{\mathbf{H}}$ the sign probability

$$P^+ = \frac{P^0_+}{P^0_+ + P^0_-}, \quad (27)$$

where

$$P^0_\pm \simeq \left(1 \pm \frac{R_{\mathbf{H}}}{2\sqrt{N}} \cos \Delta_{\mathbf{h},n}\right)^{-1} \times \exp\left[-R^2_{\mathbf{h}}\left(1 \pm \frac{R_{\mathbf{H}}}{2\sqrt{N}} \cos \Delta_{\mathbf{h},n}\right)^{-1}\right].$$

If more $R_{\mathbf{h}}$ belonging to the first phasing shell of $\varphi_{\mathbf{H}}$ are known then

$$P^+ = \frac{\Pi P^+_{j,n}}{\Pi P^+_{j,n} + \Pi P^-_{j,n}}, \quad (28)$$

where

$$P^\pm_{j,n} = \left(1 \pm W_{\mathbf{H},\mathbf{h}_j} \frac{R_{\mathbf{H}}}{2\sqrt{N}} \cos \Delta_{j,n}\right)^{-1} \times \exp\left[-R^2_{\mathbf{h}_j}\left(1 \pm W_{\mathbf{H},\mathbf{h}_j} \frac{R_{\mathbf{H}}}{2\sqrt{N}} \cos \Delta_{j,n}\right)^{-1}\right].$$

(27) and (28) may be usefully compared with (15) and (17).

7. The expected sign of $E_{2\mathbf{h}}$ in $P\bar{1}$ from its second representation

In accordance with § 1 the second representation of $E_{2\mathbf{h}}$ in $P\bar{1}$ is the collection of the special quintet invariants

$$\psi_2 = \varphi_{2\mathbf{h}} - 2\varphi_{\mathbf{h}} + \varphi_{\mathbf{k}} - \varphi_{\mathbf{k}}; \quad (29)$$

$\mathbf{k}$ is a free vector which varies throughout reciprocal space. The second phasing shell is then

$$\{B\}_2 = (R_{2\mathbf{h}}, R_{\mathbf{h}}, R_{\mathbf{k}}, R_{\mathbf{h}+\mathbf{k}}, R_{\mathbf{h}-\mathbf{k}}, R_{2\mathbf{h}+\mathbf{k}}, R_{2\mathbf{h}-\mathbf{k}}), \quad (30)$$

which suggests the study of the joint probability distribution

$$P(E_{2\mathbf{h}}, E_{\mathbf{h}}, E_{\mathbf{k}}, E_{\mathbf{h}+\mathbf{k}}, E_{\mathbf{h}-\mathbf{k}}, E_{2\mathbf{h}+\mathbf{k}}, E_{2\mathbf{h}-\mathbf{k}}). \quad (31)$$

In order to keep the notation of this paragraph similar to that in the following paragraphs where we deal with symmetries higher than $P\bar{1}$, we denote

$$E_1 = E_{2\mathbf{h}};\quad E_2 = E_{\mathbf{h}};\quad E_3 = E_{\mathbf{k}};\quad E_{4,1} = E_{\mathbf{h}+\mathbf{k}};$$
$$E_{4,2} = E_{\mathbf{h}-\mathbf{k}};\quad E_{5,1} = E_{2\mathbf{h}+\mathbf{k}};\quad E_{5,2} = E_{2\mathbf{h}-\mathbf{k}}.$$

From the Gram–Charlier expansion of the characteristic function of (31) we obtain

$$P_+(E_{2\mathbf{h}}) \simeq 0{\cdot}5 + 0{\cdot}5 \tanh \frac{|E_1|}{2\sqrt{N}}\left\{\varepsilon_2 + \frac{A_{\mathbf{k}}}{B_{\mathbf{k}}}\right\}, \quad (32)$$

where

$$A_{\mathbf{k}} = [2\varepsilon_2\varepsilon_3(\varepsilon_{4,1}\varepsilon_{5,1} + \varepsilon_{4,2}\varepsilon_{5,2} + \varepsilon_{4,1}\varepsilon_{4,2}) + \varepsilon_3(\varepsilon_{4,1}\varepsilon_{5,1} + \varepsilon_{4,2}\varepsilon_{5,2} + \varepsilon_{4,1}\varepsilon_{4,2}) - 0{\cdot}5\varepsilon_3(\varepsilon_{4,1} + \varepsilon_{4,2}) - 0{\cdot}5(\varepsilon_{4,1}\varepsilon_{5,1} + \varepsilon_{4,2}\varepsilon_{5,2})]/N,$$

$$B_{\mathbf{k}} \simeq 1 + Q_{\mathbf{k}}/2N - \tfrac{1}{8}\sum_{1\,i}^{7} H_4(E_i),$$

$$Q_{\mathbf{k}} \simeq \tfrac{1}{4}\varepsilon_1 H_4(E_2) + \varepsilon_1\varepsilon_3(\varepsilon_{5,1} + \varepsilon_{5,2}) + \varepsilon_2\varepsilon_3(\varepsilon_{4,1} + \varepsilon_{4,2}) + \varepsilon_2(\varepsilon_{4,1}\varepsilon_{5,1} + \varepsilon_{4,2}\varepsilon_{5,2}) + \varepsilon_1\varepsilon_{4,1}\varepsilon_{4,2},$$

$$\varepsilon_i = E_i^2 - 1,$$

$$H_4(E_i) = E_i^4 - 6E_i^2 + 3.$$

Some properties of (32) deserve to be stressed. The first two terms in $A_{\mathbf{k}}$ arise from the contribution of the terms denoted in (5) by $S_3^3/6t^{9/2}$ and $S_3S_4/t^{7/2}$. They give concordant information if $|E_{\mathbf{h}}| > 1$, disagreement if $|E_{\mathbf{h}}| < 1$. Their sum is $(2E_2^2 - 1)\varepsilon_3(\varepsilon_{4,1}\varepsilon_{5,1} + \varepsilon_{4,2}\varepsilon_{5,2} + \varepsilon_{4,1}\varepsilon_{4,2})$ which, unlike preceding formulae [*e.g.* Giacovazzo, 1976*b*, equation (14)], is able in principle to give information about the sign of $E_{2\mathbf{h}}$ even when $|E_{\mathbf{h}}| = 1$. Less favourable estimations are expected when $|E_{\mathbf{h}}| < 1$.

As $\mathbf{k}$ is a free vector in (29) or (30), in addition to (31) we have studied the joint probability distribution

$$P(E_{2\mathbf{h}}, E_{\mathbf{h}}, E_{\mathbf{k}_1}, E_{\mathbf{h}+\mathbf{k}_1}, \ldots, E_{2\mathbf{h}-\mathbf{k}_1}, E_{\mathbf{k}_2}, \ldots, E_{2\mathbf{h}-\mathbf{k}_2}, E_{\mathbf{k}_3}, \ldots).$$

The final sign probability for $E_{2\mathbf{h}}$ may still be described by means of (32) provided

$$A = \sum{}' A_{\mathbf{k}},\quad B = \sum{}' B_{\mathbf{k}} \quad (33)$$

replace $A_{\mathbf{k}}$ and $B_{\mathbf{k}}$. The prime to the summations in (33) warns the reader that precautions have to be taken in order to avoid duplication in the contributions.

8. The expected sign of $E_{\mathbf{H}} = E_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)}$ in any centrosymmetric space group from its second representation

In a centrosymmetric space group of order m for fixed $\mathbf{h} \in \{\mathbf{h}\}$ and $\mathbf{k}$ one may construct the set of special quintets

$$\psi_2 = \varphi_{\mathbf{H}} - \varphi_{\mathbf{h}} + \varphi_{\mathbf{hR}_n} - \varphi_{\mathbf{kR}_j} + \varphi_{\mathbf{kR}_j},\quad j = 1,\ldots,m/2. \quad (34)$$

In (34) $\mathbf{R}_j$ varies over the subset of matrices not related by the centre of symmetry. The second representation of $E_{\mathbf{h}}$ is then the collection of quintets (34) obtained when $\mathbf{h}$ varies over $\{\mathbf{h}\}$ and $\mathbf{k}$ over the asymmetric region of reciprocal space. The cross-magnitudes of any ψ_2 are

$$R_{\mathbf{H}\pm\mathbf{kR}_j}, R_{\mathbf{h}\pm\mathbf{kR}_j}, R_{\mathbf{hR}_n\pm\mathbf{kR}_j},\quad j = 1,\ldots,m/2.$$

As $R_{\mathbf{hR}_p \pm \mathbf{kR}_q}$ is symmetry-equivalent to $R_{\mathbf{h} \pm \mathbf{kR}_j}$, where $\mathbf{R}_j = \mathbf{R}_q\mathbf{R}_p^{-1}$, the second phasing shell of $E_{\mathbf{H}}$ reduces to

$$\{B\}_2 = (R_{\mathbf{H}}, R_{\mathbf{h}}, R_{\mathbf{k}}, R_{\mathbf{h}+\mathbf{kR}_j}, R_{\mathbf{H}+\mathbf{kR}_j}, j = 1,\ldots,m). \quad (35)$$

That suggests the study of the distribution

$$P(E_{\mathbf{H}}, E_{\mathbf{h}}, E_{\mathbf{k}}, E_{\mathbf{h}+\mathbf{kR}_1},\ldots, E_{\mathbf{h}+\mathbf{kR}_n}, E_{\mathbf{H}+\mathbf{kR}_1},\ldots, E_{\mathbf{H}+\mathbf{kR}_n}).$$

In order to describe in a simple way our results whatever the space group may be we denote

$$E_1 = E_{\mathbf{H}};\quad E_2 = E_{\mathbf{h}};\quad E_3 = E_{\mathbf{k}};$$
$$E_{4,j} = E_{\mathbf{h}+\mathbf{kR}_j};\quad E_{5,j} = E_{\mathbf{H}+\mathbf{kR}_j}.$$

We obtain

$$P_+(E_{\mathbf{H}}) \simeq 0{\cdot}5 + 0{\cdot}5 \tanh\left[\frac{|E_1|}{2\sqrt{N}}\left(\varepsilon_2 + \frac{A_{\mathbf{h,k}}}{B_{\mathbf{h,k}}}\right)(-1)^{2\mathbf{hT}_n}\right], \quad (36)$$

where

$$A_{\mathbf{h,k}} = \left[(2R_2^2-1)\varepsilon_3\left(\sum_{\substack{\mathbf{R}_i=\mathbf{R}_j\\ \mathbf{R}_j+\mathbf{R}_i\mathbf{R}_n=0}} \varepsilon_{4,i}\varepsilon_{5,j} + \sum_{\substack{\mathbf{R}_j=\mathbf{R}_i\mathbf{R}_n\\ \mathbf{R}_i=\mathbf{R}_i\mathbf{R}_n}} \varepsilon_{4,i}\varepsilon_{4,j}\right) - \tfrac{1}{2}\varepsilon_3\sum_{1j}^{m}\varepsilon_{4,j} - \tfrac{1}{2}\sum_{\substack{\mathbf{R}_j=\mathbf{R}_i\\ \mathbf{R}_j+\mathbf{R}_i\mathbf{R}_n=0}} \varepsilon_{4,i}\varepsilon_{5,j}\right]\Big/N,$$

$$B_{\mathbf{h,k}} = 1 - \frac{1}{8N}\sum_j H_4(E_j) + Q_{\mathbf{h,k}}/2N$$

$$Q_{\mathbf{h,k}} = \varepsilon_1\varepsilon_3\sum_{1j}^{m}\varepsilon_{5,j} + \varepsilon_1\sum_{\substack{\mathbf{R}_j=\mathbf{R}_i\mathbf{R}_n\\ \mathbf{R}_i=\mathbf{R}_j\mathbf{R}_n}}\varepsilon_{4,i}\varepsilon_{4,j} + \varepsilon_2\varepsilon_3\sum_{1j}^{m}\varepsilon_{4,j} + \varepsilon_2\sum_{\substack{\mathbf{R}_j=\mathbf{R}_i\\ \mathbf{R}_i+\mathbf{R}_i\mathbf{R}_n=0}}\varepsilon_{4,i}\varepsilon_{5,j} + \tfrac{1}{4}\varepsilon_1 H_4(E_2).$$

The derivation of (36) is not straightforward and requires an application of space-group algebra to the joint probability distribution method. In order to justify the constraints in the summations in (36) some elements of this algebra are described in Appendix *D*. In this connexion we note: (*a*) the condition $\mathbf{R}_i = \mathbf{R}_j\mathbf{R}_n$ coincides with $\mathbf{R}_j = \mathbf{R}_i\mathbf{R}_n$ if $\mathbf{R}_n$ corresponds to a symmetry operator of order two. In fact multiplying by $\mathbf{R}_n^{-1}$ both sides of $\mathbf{R}_i = \mathbf{R}_j\mathbf{R}_n$ we obtain $\mathbf{R}_i\mathbf{R}_n^{-1} = \mathbf{R}_i\mathbf{R}_n = \mathbf{R}_j$. (*b*) If $\mathbf{R}_n = -\mathbf{I}$ the condition $\mathbf{R}_j + \mathbf{R}_i\mathbf{R}_n = 0$ coincides with $\mathbf{R}_i = \mathbf{R}_j$.

Since $\mathbf{h}$ under certain conditions ($\mathbf{h} \in \{\mathbf{h}\}$) and $\mathbf{k}$ are free vectors in (35), the more complex joint probability distribution

$$P(E_{\mathbf{H}}, \{E_{\mathbf{h}}\}, \{E_{\mathbf{k}}\}, \{E_{\mathbf{h}+\mathbf{kR}_j}\}, \{E_{\mathbf{H}+\mathbf{kR}_j}\}, \ j = 1,\ldots,m) \quad (37)$$

has to be studied. In (37) $\{E_{\mathbf{h}}\}$ is the set of structure factors whose indices belong to $\{\mathbf{h}\}$, $\{E_{\mathbf{k}}\}$ is any chosen set in the asymmetric region of reciprocal space, $\{E_{\mathbf{h}+\mathbf{kR}_j}\}$ and $\{E_{\mathbf{H}+\mathbf{kR}_j}\}$ are sets obtainable from the specified conditions on $\mathbf{h}$ and $\mathbf{k}$. We have found that the final sign probability for $E_{\mathbf{H}}$ may be still described by means of (36) provided

$$\sum_{\mathbf{h,k}}{}' A_{\mathbf{h,k}} \quad\text{and}\quad \sum_{\mathbf{h,k}}{}' B_{\mathbf{h,k}} \quad (38)$$

replace $A_{\mathbf{h,k}}$ and $B_{\mathbf{h,k}}$. The prime to the summations in (38) warns the reader that precautions have to be taken in order to avoid duplication in the contributions.

9. The expected value of $\varphi_{\mathbf{H}} = \varphi_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)}$ in non-centrosymmetric space groups from its second representation

The second phasing shell of $E_{\mathbf{H}}$ is

$$\{B\}_2 = (R_{\mathbf{H}}, R_{\mathbf{h}}, R_{\mathbf{k}}, R_{\mathbf{h}+\mathbf{kR}_j}, R_{\mathbf{h}-\mathbf{kR}_j}, R_{\mathbf{H}+\mathbf{kR}_j}, j = 1,\ldots,m), \quad (39)$$

where $\mathbf{h}$ is a free vector under the condition $\mathbf{h} \in \{\mathbf{h}\}$ and $\mathbf{k}$ is any vector in the asymmetric region of reciprocal space. If we introduce the fictitious (not belonging to the space group) symmetry operators $\mathbf{C}_{m+j} = (-\mathbf{R}_j, -\mathbf{T}_j), j = 1,\ldots,m$, (39) may be written as

$$\{B\}_2 = (R_{\mathbf{H}}, R_{\mathbf{h}}, R_{\mathbf{k}}, R_{\mathbf{h}+\mathbf{kR}_j}, R_{\mathbf{H}+\mathbf{kR}_j}, j = 1,\ldots,m'),$$

where $m' = 2m$.

We may then use the compact notation of § 8 and obtain finally

$$\langle\cos\varphi_{\mathbf{H}}\rangle \simeq \frac{I_1(G)}{I_0(G)}, \quad (40)$$

where

$$G = \frac{|E_1|}{\sqrt{N}}\left(\varepsilon_2 + \frac{A_{\mathbf{h,k}}}{B_{\mathbf{h,k}}}\right)\cos 2\pi\mathbf{hT}_n.$$

$A_{\mathbf{h,k}}$ and $B_{\mathbf{h,k}}$ are the same as in § 8. The extension to any number of $\mathbf{h}$ and $\mathbf{k}$ vectors is straightforward.

10. The estimation of $\varphi_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)}$ *via* the method of complementary invariants

A branch of the representations method is the method of complementary invariants. To the structure seminvariant Φ which one wishes to estimate, one or more structure invariants or seminvariants $\Phi_f, \Phi_g, \ldots, \Phi_q$ can be associated such that

$$\Phi' = \Phi + \Phi_f + \Phi_g + \ldots + \Phi_q$$

is a structure invariant. If $\Phi', \Phi_f, \ldots, \Phi_q$ are estimated, Φ is in consequence evaluated. The method has been successfully applied in $P\bar{1}$ by Giacovazzo (1975) to quartet complementary invariants. Sheldrick (1976) extended in an empirical way the idea to all the space groups. We wish to give here general probabilistic formulae valid in any space group when the complementary invariants are given by

$$\Phi' = \varphi_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)} - \varphi_{\mathbf{h}} - \varphi_{\mathbf{k}} + \varphi_{\mathbf{h}\mathbf{R}_n+\mathbf{k}}. \tag{41}$$

$\varphi_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)}$ will be estimated from all the quartets which arise when $\mathbf{k}$ varies over reciprocal space on condition that triplets $\varphi_{\mathbf{h}} + \varphi_{\mathbf{k}} - \varphi_{\mathbf{h}\mathbf{R}_n+\mathbf{k}}$ are known. As one of the cross-vectors of (41) coincides with a basis vector, quartets such as Φ' cannot be estimated by formulae given by Hauptman (1975) for general quartets. Special formulae have already been given by Giacovazzo (1975) *via* the Gram–Charlier expansion of the characteristic function of the six-term distribution

$$P[E_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)}, E_{\mathbf{h}}, E_{\mathbf{k}}, E_{\mathbf{h}\mathbf{R}_n+\mathbf{k}}, E_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)-\mathbf{k}}, E_{\mathbf{h}+\mathbf{k}}]. \tag{42}$$

We give in Appendix E some conditional distributions derived from (42) *via* the exponential form of its characteristic function.

For brevity, formulae in Appendix E and in this paragraph will not be proved. We obtain in centrosymmetric space groups from (42)

$$P^{\pm}[E_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)}] \simeq a^{+} \exp\left[\pm \frac{1}{2\sqrt{N}} R_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)} R_{\mathbf{h}}^2 (-1)^{2\mathbf{h}\mathbf{T}_n} \mp \frac{2}{N} R_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)} E_{\mathbf{h}} E_{\mathbf{k}} E_{\mathbf{h}\mathbf{R}_n+\mathbf{k}}\right] \times \left\{\cosh\left[\frac{X^{\pm} R_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)-\mathbf{k}}}{\sqrt{N}}\right] \cosh\left(\frac{Y^{\pm} R_{\mathbf{h}+\mathbf{k}}}{\sqrt{N}}\right)\right\}, \tag{43}$$

where

$$X^{\pm} = [R_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)} E_{\mathbf{k}} \pm E_{\mathbf{h}} E_{\mathbf{h}\mathbf{R}_n+\mathbf{k}}],$$

$$Y^{\pm} = [R_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)} E_{\mathbf{h}\mathbf{R}_n+\mathbf{k}} \pm E_{\mathbf{h}} E_{\mathbf{k}}],$$

$$a^{\pm} = \left[1 \pm \frac{R_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)}}{\sqrt{N}} (-1)^{2\mathbf{h}\mathbf{T}_n}\right]^{-1/2}.$$

(43) gives the probability that the sign of $E_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)}$ is positive when the signs of the triplet $E_{\mathbf{h}\mathbf{R}_n} E_{\mathbf{k}} E_{\mathbf{h}\mathbf{R}_n+\mathbf{k}}$ and all the magnitudes in (42) are known.

If we use (E.5), (43) may be written

$$P^{\pm}[E_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)}] \simeq \exp\left[\pm \frac{1}{2\sqrt{N}} R_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)} (R_{\mathbf{h}}^2 - 1)(-1)^{2\mathbf{h}\mathbf{T}_n} \mp \frac{2}{N} R_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)} R_{\mathbf{h}} R_{\mathbf{k}} R_{\mathbf{h}\mathbf{R}_n+\mathbf{k}}\right] \times \left\{\cosh\left[\frac{X^{\pm} R_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)-\mathbf{k}}}{\sqrt{N}}\right] \cosh\left(\frac{Y^{\pm} R_{\mathbf{h}+\mathbf{k}}}{\sqrt{N}}\right)\right\}. \tag{44}$$

If $\mathbf{k}$ varies over the region of reciprocal space for which $E_{\mathbf{h}\mathbf{R}_n} E_{\mathbf{k}} E_{\mathbf{h}\mathbf{R}_n+\mathbf{k}}$ are large, then in (43) and (44) $E_{\mathbf{h}} E_{\mathbf{k}} E_{\mathbf{h}\mathbf{R}_n+\mathbf{k}}$ may be replaced by $(-1)^{2\mathbf{h}\mathbf{T}_n} R_{\mathbf{h}} R_{\mathbf{k}} R_{\mathbf{h}\mathbf{R}_n+\mathbf{k}}$ and $X^{\pm}$, $Y^{\pm}$ become

$$[R_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)} R_{\mathbf{k}} (-1)^{2\mathbf{h}\mathbf{T}_n} \pm R_{\mathbf{h}} R_{\mathbf{h}\mathbf{R}_n+\mathbf{k}}]$$

and

$$[R_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)} R_{\mathbf{h}\mathbf{R}_n+\mathbf{k}} \pm R_{\mathbf{h}} R_{\mathbf{k}} (-1)^{2\mathbf{h}\mathbf{T}_n}]$$

respectively.

Series expansion of (44) may be readily applied when more than one quartet is used in order to estimate the sign probability of $E_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)}$. We obtain

$$P_{+}[E_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)}] \simeq 0{\cdot}5 + 0{\cdot}5 \tanh\left[\frac{R_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)}}{2\sqrt{N}} \left(\varepsilon_{\mathbf{h}} + \frac{2}{\sqrt{N}} \times \frac{{}_{\mathbf{h},\mathbf{k}}}{1 + \sum_{\mathbf{h},\mathbf{k}} B_{\mathbf{h},\mathbf{k}}}\right)(-1)^{2\mathbf{h}\mathbf{T}_n}\right], \tag{45}$$

where

$$A_{\mathbf{h},\mathbf{k}} \simeq R_{\mathbf{h}} R_{\mathbf{k}} R_{\mathbf{h}\mathbf{R}_n+\mathbf{k}} [\varepsilon_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)-\mathbf{k}} + \varepsilon_{\mathbf{h}+\mathbf{k}}],$$

$$B_{\mathbf{h},\mathbf{k}} \simeq \{\varepsilon_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)} [\varepsilon_{\mathbf{k}} \varepsilon_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)-\mathbf{k}} + \varepsilon_{\mathbf{h}\mathbf{R}_n+\mathbf{k}} \varepsilon_{\mathbf{h}+\mathbf{k}}] + \varepsilon_{\mathbf{h}} [\varepsilon_{\mathbf{k}} \varepsilon_{\mathbf{h}+\mathbf{k}} + \varepsilon_{\mathbf{h}\mathbf{R}_n+\mathbf{k}} \varepsilon_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)-\mathbf{k}}]\}/2N.$$

If $\varphi_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)}$ is a non-centrosymmetric phase, $I_1(2G)/I_0(G)$ is the expected value of $\cos \varphi_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)}$, where G is the argument of the hyperbolic tangent in (45).

11. Conclusions

A theory has been described which is able to estimate the value of a one-phase structure seminvariant of first rank given the magnitudes in the first and second phasing shells. When only magnitudes in the first phasing shells are known probabilistic formulae are given which formally differ from the hyperbolic tangent formulation obtained by other authors and reduce to them if suitable approximations are introduced. New formulae are also given *via* the method of complementary invariants. The first applications of the theory proved satisfactory and are described by Burla, Polidori, Nunzi & Giacovazzo (1978).

This work was supported by the Consiglio Nazionale delle Ricerche (grant No. 77.00968.05.115.4593).

APPENDIX *A*

In this and the following appendices we denote by ξ the trigonometric form of the structure factor in the actual space group, by ψ and η the real and imaginary parts of ξ in non-centrosymmetric space groups. Since

$$\mathbf{E_h} = F_\mathbf{h}/\sqrt{p_\mathbf{h}} \left[\sum_{1j}^{N} f_j^2(\mathbf{h})\right]^{1/2}$$

we denote

$$\xi'(\mathbf{h}) = \xi(\mathbf{h})/\sqrt{p_\mathbf{h}}; \quad \psi'(\mathbf{h}) = \psi(\mathbf{h})/\sqrt{p_\mathbf{h}};$$
$$\eta'(\mathbf{h}) = \eta(\mathbf{h})/\sqrt{p_\mathbf{h}},$$

where $p_\mathbf{h}$ is the statistical weight of the reflexion $\mathbf{h}$.

Denoting $\mathbf{H} = \mathbf{h}(\mathbf{I} - \mathbf{R}_n)$, in accordance with preceding papers (Giacovazzo, 1974, 1976*a*) the weight $W_{\mathbf{H,h}}$ involved in (14) and (17) is given by

$$W_{\mathbf{H,h}} = \frac{\langle \xi(\mathbf{H})\xi^2(\mathbf{h})\rangle}{mp_\mathbf{h}\sqrt{p_\mathbf{H}}} = \frac{\langle \xi'(\mathbf{H})\xi'^2(\mathbf{h})\rangle}{m}$$
$$= \left\langle \sum_{1\,p,q}^{m} \xi[\mathbf{h}(\mathbf{I} - \mathbf{R}_n + \mathbf{R}_q - \mathbf{R}_p\mathbf{R}_q)] \times \exp 2\pi i\mathbf{h}(\mathbf{T}_q - \mathbf{R}_p\mathbf{T}_q - \mathbf{T}_p)\right\rangle \Big/ mp_\mathbf{h}\sqrt{p_\mathbf{H}}. \quad (A.1)$$

The value of (A.1) is different from zero for all $\mathbf{R}_p$, $\mathbf{R}_q$ operators for which

$$\mathbf{h}(\mathbf{I} - \mathbf{R}_n + \mathbf{R}_q - \mathbf{R}_p\mathbf{R}_q) = 0. \quad (A.2)$$

For example, whatever $\mathbf{h}$ and $\mathbf{H}$ may be, (A.2) is satisfied at least when

$$(a)\ \mathbf{R}_q = -\mathbf{I}, \quad \mathbf{R}_p = \mathbf{R}_n;$$
$$(b)\ \mathbf{R}_q = \mathbf{R}_n, \quad \mathbf{R}_p = \mathbf{R}_n^{-1}.$$

If $\mathbf{h}$ is a systematically absent reflexion, then $\xi(\mathbf{h}) = 0$ and $W_{\mathbf{H,h}} = 0$. Numerical values of $W_{\mathbf{H,h}}$ for different parity classes are shown in Table 2 for the space group *Pmmm*.

APPENDIX *B*

Denoting $E_\mathbf{H} = E_{2\mathbf{h}}$, we calculate in $P\bar{1}$ $P(E_\mathbf{H},E_\mathbf{h})$ directly *via* the exponential form of the characteristic function (4):

$$P(E_\mathbf{H},E_\mathbf{h}) \simeq \frac{1}{(2\pi)^2} \int_{-\infty}^{-\infty}\int_{-\infty}^{+\infty} \exp\left[-\tfrac{1}{2}(u_\mathbf{H}^2 + u_\mathbf{h}^2) - i\left(E_\mathbf{H}u_\mathbf{H} + E_\mathbf{h}u_\mathbf{h} + \frac{1}{2\sqrt{N}}u_\mathbf{H}u_\mathbf{h}^2\right)\right] \mathrm{d}u_\mathbf{H}\mathrm{d}u_\mathbf{h}. \quad (B.1)$$

The integration of (B.1) with respect to $u_\mathbf{H}$ is readily carried out:

$$P \simeq (2\pi)^{-3/2} \exp(-E_\mathbf{H}^2/2) \times \int_{-\infty}^{+\infty} \exp\left[-\left(\frac{1}{2} + \frac{E_\mathbf{H}}{2\sqrt{N}}\right)u_\mathbf{h}^2 - iE_\mathbf{h}u_\mathbf{h}\right] \mathrm{d}u_\mathbf{h}. \quad (B.2)$$

The integration of (B.2) with respect to $u_\mathbf{h}$ is carried out by means of the integral relation

$$\int_{-\infty}^{+\infty} \exp(-p^2x^2 \pm qx)\,\mathrm{d}x = \exp\left(\frac{q^2}{4p^2}\right)\frac{\sqrt{\pi}}{p}\ (p > 0).$$

Then

$$P \simeq \frac{1}{2\pi} f^{-1/2} \exp\left(-\frac{E_\mathbf{H}^2}{2} - \frac{E_\mathbf{h}^2}{2f}\right) \quad (B.3)$$

where

$$f = 1 + E_\mathbf{H}/\sqrt{N}.$$

From (B.3) (15) is easily found.

Let us now calculate in any centrosymmetric space group with crystal symmetry higher than $P\bar{1}$ the multivariate probability $P(E_\mathbf{H},\{E_\mathbf{h}\})$. We obtain

$$P(E_\mathbf{H},\{E_\mathbf{h}\}) \simeq \frac{1}{(2\pi)^{n+1}} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \exp\left\{-\tfrac{1}{2}\left(u_\mathbf{H}^2 + \sum_p{}' \sum_\mathbf{h} u_\mathbf{h}^2\right) - i\left[E_\mathbf{H}u_\mathbf{H} + \sum_p{}' \sum_\mathbf{h} E_\mathbf{h}u_\mathbf{h} + \sum_p{}' \sum_\mathbf{n} \frac{W_{\mathbf{H,h}}}{2\sqrt{N}} u_\mathbf{H}u_\mathbf{h}^2(-1)^{2\mathbf{hT}_p} + \ldots\right]\right\}, \quad (B.4)$$

where n is the number of the distinct $\mathbf{h}$ vectors belonging to the first phasing shell of $E_\mathbf{H}$. The same

Table 2. *Numerical values of $W_{\mathbf{H,h}}$ for different parity classes for space group Pmmm*

$\mathbf{H}$	$2h\,2k\,2l$	$2h\,0\,2l$	$2h\,0\,2l$	$2h\,0\,0$	$2h\,0\,0$	$2h\,0\,0$
$\mathbf{h}$	$h\,k\,l$	$h\,k\,l$	$h\,0\,l$	$h\,k\,l$	$h\,k\,0$	$h\,0\,0$
$\langle \xi(\mathbf{H})\xi^2(\mathbf{h})\rangle$	8	16	32	32	64	128
$\langle \xi'(\mathbf{H})\xi'^2(\mathbf{h})\rangle$	8	$16/\sqrt{2}$	$16/\sqrt{2}$	16	16	16
$W_{\mathbf{H,h}}$	1	$\sqrt{2}$	$\sqrt{2}$	2	2	2

considerations made for (14) hold for the symbols $\sum'_n$, $\sum_{\mathbf{h}}$ and $W_{\mathbf{H,h}}$. After integration we obtain

$$P \simeq \frac{1}{\sqrt{2\pi}} \exp(-E_{\mathbf{h}}^2/2) \prod_{p,\mathbf{h}} \frac{1}{(2\pi)^{1/2}} f_{\mathbf{h}}^{-1/2} \exp\left\{-\frac{E_{\mathbf{h}}^2}{2} f_{\mathbf{h}}^{-1}\right\}, \quad (B.5)$$

where

$$f_{\mathbf{h}} = 1 + \frac{W_{\mathbf{H,h}} E_{\mathbf{H}} (-1)^{2\mathbf{hT}_p}}{\sqrt{N}}.$$

(17) is easily derived from (B.5).

APPENDIX C

Denoting $\mathbf{H} = \mathbf{h}(\mathbf{I} - \mathbf{R}_n)$ and $A_{\mathbf{H}}$, $A_{\mathbf{h}}$, $B_{\mathbf{H}}$, $B_{\mathbf{h}}$ the real and imaginary parts of the normalized structure factors $E_{\mathbf{H}}$ and $E_{\mathbf{h}}$, in accordance with (6) their characteristic function is

$$C(u_{\mathbf{H}}, u_{\mathbf{h}}, v_{\mathbf{H}}, v_{\mathbf{h}}) = \exp\left[-\frac{K_{2000}}{2m} u_{\mathbf{H}}^2 - \ldots - \frac{K_{0002}}{2m} v_{\mathbf{h}}^2 - \frac{i}{\sqrt{N}}\left(\frac{K_{1200}}{2m} u_{\mathbf{H}} u_{\mathbf{h}}^2 + \frac{K_{1020}}{2m} u_{\mathbf{H}} v_{\mathbf{h}}^2 + \ldots + \frac{K_{0012}}{2m} v_{\mathbf{H}} u_{\mathbf{h}}^2\right)\right]. \quad (C.1)$$

This expression is able to take into account both the incidental centrosymmetric nature of $E_{\mathbf{H}}$ or $E_{\mathbf{h}}$ and their statistical nature. We remember in this connexion that

$$K_{2000} = m_{2000} = \langle \psi_{\mathbf{H}}'^2 \rangle, \ldots$$

$$K_{1200} = \langle \psi_{\mathbf{H}}' \psi_{\mathbf{h}}'^2 \rangle, \ldots$$

Furthermore, if $E_{\mathbf{h}}$ is a centrosymmetric reflexion,

$$\langle \psi'^2(\mathbf{h}) \rangle = m \quad \text{or} \quad \langle \eta'^2(\mathbf{h}) \rangle = m$$

whatever its statistical weight may be. For a non-centrosymmetric reflexion always

$$\langle \psi'^2(\mathbf{h}) \rangle = \langle \eta'^2(\mathbf{h}) \rangle = m/2.$$

By a change of variable

$$u_{\mathbf{H}} = \rho_{\mathbf{H}} \cos \psi_{\mathbf{H}}, \quad v_{\mathbf{H}} = \rho_{\mathbf{H}} \sin \psi_{\mathbf{H}}, \ldots$$

$$A_{\mathbf{H}} = R_{\mathbf{H}} \cos \varphi_{\mathbf{H}}, \quad B_{\mathbf{H}} = R_{\mathbf{H}} \sin \varphi_{\mathbf{H}}, \ldots$$

we obtain, in the case that both $E_{\mathbf{H}}$ and $E_{\mathbf{h}}$ are non-centrosymmetric reflexions,

$$P(\varphi_{\mathbf{H}}, \varphi_{\mathbf{h}}, R_{\mathbf{H}}, R_{\mathbf{h}}) \simeq \frac{1}{(2\pi)^4} R_{\mathbf{H}} R_{\mathbf{h}} \int_0^{2\pi}\int_0^{2\pi}\int_0^{\infty}\int_0^{\infty} \rho_{\mathbf{H}} \rho_{\mathbf{h}} \times \exp\left\{-\tfrac{1}{4}(\rho_{\mathbf{H}}^2 + \rho_{\mathbf{h}}^2) - i\left[R_{\mathbf{H}} \rho_{\mathbf{H}} \cos(\varphi_{\mathbf{H}} - \psi_{\mathbf{H}}) + R_{\mathbf{h}} \rho_{\mathbf{h}} \cos(\varphi_{\mathbf{h}} - \psi_{\mathbf{h}}) - \frac{W_{\mathbf{H,h}}}{8\sqrt{N}} \rho_{\mathbf{H}} \rho_{\mathbf{h}}^2 \cos(\psi_{\mathbf{H}} - \Delta)\right]\right\} d\rho_{\mathbf{H}} d\rho_{\mathbf{h}} d\psi_{\mathbf{H}} d\psi_{\mathbf{h}}, \quad (C.2)$$

where $\Delta = 2\pi \mathbf{hT}_n$ and $W_{\mathbf{H,h}}$ is a statistical weight analogous to that derived for centrosymmetric space groups in Appendix A. The integration of (C.2) may be performed by formulae such as

$$\frac{1}{2\pi} \int_0^{\infty} \int_0^{2\pi} \exp(-p^2 t^2 - iat \cos \varphi) \, t dt \, d\varphi = \frac{1}{2p^2} \exp\left(-\frac{a^2}{4p^2}\right).$$

We obtain

$$P(\varphi_{\mathbf{H}}, \varphi_{\mathbf{h}}, R_{\mathbf{H}}, R_{\mathbf{h}}) = \frac{R_{\mathbf{H}} R_{\mathbf{h}}}{\pi^2} \frac{1}{1 + \dfrac{W_{\mathbf{H,h}}}{\sqrt{N}} R_{\mathbf{H}} \cos(\varphi_{\mathbf{H}} - \Delta)} \times \exp\left[-R_{\mathbf{H}}^2 - R_{\mathbf{h}}^2 \Big/ \left(1 + \frac{W_{\mathbf{H,h}}}{\sqrt{N}} R_{\mathbf{H}} \cos(\varphi_{\mathbf{H}} - \Delta)\right)\right]. \quad (C.3)$$

In order to calculate $P(\varphi_{\mathbf{H}} | \varphi_{\mathbf{h}}, R_{\mathbf{H}}, R_{\mathbf{h}})$ we need to estimate

$$I = \int_{-\pi}^{\pi} \frac{1}{1 + a \cos(\varphi_{\mathbf{H}} - \Delta)} \times \exp\left[-\frac{b}{1 + a \cos(\varphi_{\mathbf{H}} - \Delta)}\right] d\varphi_{\mathbf{H}},$$

where $a = W_{\mathbf{H,h}} R_{\mathbf{H}} / \sqrt{N}$ and $b = R_{\mathbf{h}}^2$. By a change of variable $x = 1/[1 + a \cos(\varphi_{\mathbf{H}} - \Delta)]$ we obtain

$$I = 2 \int_c^d \frac{\exp(-bx)}{[(a^2 - 1)x^2 + 2x - 1]^{1/2}} dx,$$

where $c = 1/(1 + a)$ and $d = 1/(1 - a)$.

c and d are real roots of $(a^2 - 1)x^2 + 2x - 1 = 0$ in the same way as 0 and 2 are roots of $y(y - 2) = 0$ in the definite integral

$$\int_0^2 \frac{\exp(-py)}{[y(2 - y)]^{1/2}} dy = \pi \exp(-p) I_0(p) \quad (p > 0).$$

This suggests the further transformation

$$x = \tfrac{1}{2}[(d - c)y + 2c],$$

which leads us to

$$I = \frac{2}{(1 - a^2)^{1/2}} \exp[-b/(1 + a)] \times \int_0^2 \frac{\exp[-bay/(1 - a^2)]}{[y(2 - y)]^{1/2}} dy = \frac{2\pi}{(1 - a^2)^{1/2}} \exp[-b/(1 - a^2)] I_0[ba/(1 - a^2)]. \quad (C.4)$$

Since

$$P(\varphi_{\mathbf{H}}|\varphi_{\mathbf{h}},R_{\mathbf{H}},R_{\mathbf{h}}) = \frac{[1+a\cos(\varphi_{\mathbf{H}}-\Delta)]^{-1}\exp\{-b/[1+a\cos(\varphi_{\mathbf{H}}-\Delta)]\}}{I}$$

(25) is proved.

Let us now define the form of the distribution for cases in which $E_{\mathbf{H}}$ is a centrosymmetric reflexion while $E_{\mathbf{h}}$ is not. Then (C.2) becomes

$$P(E_{\mathbf{H}},\varphi_{\mathbf{h}},R_{\mathbf{h}}) \simeq \frac{1}{(2\pi)^3} R_{\mathbf{h}} \int_{-\infty}^{+\infty}\int_{0}^{\infty}\int_{0}^{2\pi} \rho_{\mathbf{h}} \times \exp\left\{-\tfrac{1}{2}u_{\mathbf{H}}^2 - \tfrac{1}{4}\rho_{\mathbf{h}}^2 - i\left[E_{\mathbf{H}}u_{\mathbf{H}} + R_{\mathbf{h}}\rho_{\mathbf{h}}\cos(\psi_{\mathbf{h}}-\varphi_{\mathbf{h}}) + \frac{W_{\mathbf{H},\mathbf{h}}}{8\sqrt{N}}u_{\mathbf{H}}\rho_{\mathbf{h}}^2\cos(\varphi_{\mathbf{H}}-\Delta)\right]\right\} \mathrm{d}u_{\mathbf{H}}\mathrm{d}\rho_{\mathbf{h}}\mathrm{d}\varphi_{\mathbf{h}}. \quad (C.5)$$

The calculation of (C.5) may be made by mathematical techniques similar to that previously described and leads to (27).

APPENDIX D

In § 8 several multivariate distributions are studied which require calculation of standardized cumulants of low order. The algebraic form of the conclusive probabilistic formulae depends on the type of non-vanishing cumulants which can be found. The estimation of all the cumulants used is too long to be described; however, we present an example. We deal here only with cumulants of order $1/\sqrt{N}$.

(a) $\langle\xi[\mathbf{h}(\mathbf{I}-\mathbf{R}_n)+\mathbf{k}\mathbf{R}_j]\xi[-\mathbf{h}(\mathbf{I}-\mathbf{R}_n)]\xi(-\mathbf{k})\rangle$

$$= \sum_{v,\psi=1}^{m} \langle\xi[\mathbf{H}(\mathbf{I}-\mathbf{R}_v)+\mathbf{k}(\mathbf{R}_j-\mathbf{R}_\psi)](-1)^{2(\mathbf{H}\mathbf{T}_v+\mathbf{k}\mathbf{T}_\psi)}\rangle. \quad (D.1)$$

(D.1) does not vanish for all R_v and R_ψ for which

$$\mathbf{H}(\mathbf{I}-\mathbf{R}_v)=0, \quad \mathbf{R}_j-\mathbf{R}_\psi=0. \quad (D.2)$$

Condition (D.2) is satisfied by more than one pair of matrices $\mathbf{R}_v$, $\mathbf{R}_\psi$ if $\mathbf{H}$ and $\mathbf{k}$ are special vectors (*i.e.* their statistical weight is larger than unity). Drawing of special-vector covariances in a direct procedure which uses the second representation of a one-phase seminvariant would be too expensive. Therefore, for the sake of brevity we limit ourselves to consider that (D.2) is satisfied only when $\mathbf{R}_v = \mathbf{I}$ and $\mathbf{R}_\psi = \mathbf{R}_j$. In conclusion (D.1) does not vanish whatever j may be and its value equals $m(-1)^{2\mathbf{k}\mathbf{T}_j}$.

(b) $\langle\xi[\mathbf{h}(\mathbf{I}-\mathbf{R}_n)]\xi[-(\mathbf{h}+\mathbf{k}\mathbf{R}_i)]\xi[\mathbf{h}+\mathbf{k}\mathbf{R}_j]\rangle$

$$= \sum_{v,\psi=1}^{m} \langle\xi[\mathbf{h}(\mathbf{I}-\mathbf{R}_n-\mathbf{R}_v+\mathbf{R}_\psi)+\mathbf{k}(-\mathbf{R}_i\mathbf{R}_v+\mathbf{R}_j\mathbf{R}_\psi)] \times (-1)^{2[-(\mathbf{h}+\mathbf{k}\mathbf{R}_i)\mathbf{T}_v+(\mathbf{h}+\mathbf{k}\mathbf{R}_j)\mathbf{T}_\psi]}\rangle. \quad (D.3)$$

(D.3) does not vanish when (1) $\mathbf{R}_v = \mathbf{I}$, $\mathbf{R}_\psi = \mathbf{R}_n$, $\mathbf{R}_i = \mathbf{R}_j\mathbf{R}_n$; (2) $\mathbf{R}_\psi = -\mathbf{I}$, $\mathbf{R}_v = -\mathbf{R}_n$, $\mathbf{R}_j = \mathbf{R}_i\mathbf{R}_n$. (1) and (2) coincide if $\mathbf{R}_n$ corresponds to a symmetry operator of order two. Since $\mathbf{R}_j\mathbf{T}_n = \mathbf{T}_i - \mathbf{T}_j$ when $\mathbf{R}_j\mathbf{R}_n = \mathbf{R}_i$ in both the cases the value of (D.3) equals $m(-1)^{2[\mathbf{h}+\mathbf{k}(\mathbf{T}_i-\mathbf{T}_j)]}$.

(c) $\langle\xi(-\mathbf{h})\xi(-\mathbf{k})\xi(\mathbf{h}+\mathbf{k}\mathbf{R}_j)\rangle$

$$= \sum_{v,\psi=1}^{m} \langle\xi[\mathbf{h}(\mathbf{I}-\mathbf{R}_v)+\mathbf{k}(\mathbf{R}_j-\mathbf{R}_\psi)] \times \exp 2\pi i(-\mathbf{h}\mathbf{T}_v-\mathbf{k}\mathbf{T}_\psi)\rangle. \quad (D.4)$$

(D.4) does not vanish when $\mathbf{R}_v = \mathbf{I}$ and $\mathbf{R}_\psi = \mathbf{R}_j$. In other words (D.2) does not vanish whatever j may be and its value equals $m(-1)^{2\mathbf{k}\mathbf{T}_j}$.

(d) $\langle\xi[\mathbf{h}(\mathbf{I}-\mathbf{R}_n)+\mathbf{k}\mathbf{R}_j]\xi(\mathbf{h}+\mathbf{k}\mathbf{R}_i)\xi(-\mathbf{h})\rangle$

$$= \sum_{v,\psi=1}^{m} \langle\xi[\mathbf{h}(\mathbf{I}-\mathbf{R}_n+\mathbf{R}_v-\mathbf{R}_\psi)+\mathbf{k}(\mathbf{R}_j+\mathbf{R}_i\mathbf{R}_v)] \times \exp 2\pi i[-\mathbf{h}\mathbf{T}_\psi+(\mathbf{h}+\mathbf{k}\mathbf{R}_i)\mathbf{T}_v]\rangle. \quad (D.5)$$

(D.5) does not vanish when (1) $\mathbf{R}_\psi = \mathbf{I}$, $\mathbf{R}_v = \mathbf{R}_n$, $\mathbf{R}_j + \mathbf{R}_i\mathbf{R}_n = 0$; (2) $\mathbf{R}_v = -\mathbf{I}$, $\mathbf{R}_\psi = -\mathbf{R}_n$, $\mathbf{R}_i = \mathbf{R}_j$. (1) and (2) coincide if $\mathbf{R}_n = -\mathbf{I}$. In conclusion (D.5) does not vanish when $\mathbf{R}_i = \mathbf{R}_j$ or $\mathbf{R}_j + \mathbf{R}_i\mathbf{R}_n = 0$ and its value equals $m(-1)^{2\mathbf{h}\mathbf{T}_n}$.

APPENDIX E

Retaining terms up to $1/N$ order we obtain in centrosymmetric space groups

$$P(E_1,E_2,E_3,E_4,E_5,E_6) \simeq \frac{1}{(2\pi)^3} a \exp\left\{-\tfrac{1}{2}\sum_{j=1}^{6} E_j^2 + \frac{1}{\sqrt{N}}[\tfrac{1}{2}E_1E_2^2(-1)^{2\mathbf{h}\mathbf{T}_n} + E_1E_3E_5 + E_1E_4E_6 + E_2E_3E_4(-1)^{2\mathbf{h}\mathbf{T}_n} + E_2E_3E_6 + E_2E_4E_5] - \frac{1}{2N}[E_2^2E_3E_5 + E_2^2E_4E_6 + E_3E_4^2E_5 + E_3^2E_4E_6](-1)^{2\mathbf{h}\mathbf{T}_n} - \frac{1}{N}[E_1E_2E_4E_5(-1)^{2\mathbf{h}\mathbf{T}_n} + E_1E_2E_3E_6(-1)^{2\mathbf{h}\mathbf{T}_n} + E_1E_2E_5E_6 + E_3E_4E_5E_6 + 2E_1E_2E_3E_4]\right\}, \quad [E.1]$$

where

$$E_1 = E_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)};\quad E_2 = E_{\mathbf{h}};\quad E_3 = E_{\mathbf{k}};\quad E_4 = E_{\mathbf{hR}_n+\mathbf{k}};$$
$$E_5 = E_{\mathbf{h}(\mathbf{I}-\mathbf{R}_n)-\mathbf{k}};\quad E_6 = E_{\mathbf{h}+\mathbf{k}},\qquad (E.2)$$

$$a = \left[1 + \frac{E_1}{\sqrt{N}}(-1)^{2\mathbf{hT}_n}\right]^{-1/2}.$$

From (E.1) the probability that the sign of $E_1E_2E_3E_4$ is positive given six magnitudes is

$$P_Q^+ \simeq \frac{P^+}{P^+ + P^-},\qquad (E.3)$$

where

$$P^{\pm} = \exp(\mp 2B)\cosh\left(\frac{R_5 Z_5^{\pm}}{\sqrt{N}}\right)\cosh\left(\frac{R_6 Z_6^{\pm}}{\sqrt{N}}\right)\alpha^{\pm},\qquad (E.4)$$

$$B = R_1R_2R_3R_4/N,\quad Z_5^{\pm} = (R_1R_3 \pm R_2R_4),$$
$$Z_6^{\pm} = (R_2R_3 \pm R_1R_4)$$

$$\alpha^+ = \beta^+ \exp\left[\left(\frac{1}{2\sqrt{N}}R_1R_2^2 + \frac{R_2R_3R_4}{\sqrt{N}}\right)(-1)^{2\mathbf{hT}_n}\right] + \beta^- \exp\left[\left(-\frac{1}{2\sqrt{N}}R_1R_2^2 - \frac{R_2R_3R_4}{\sqrt{N}}\right)(-1)^{2\mathbf{hT}_n}\right],$$

$$\alpha^- = \beta^+ \exp\left[\left(\frac{1}{2\sqrt{N}}R_1R_2^2 - \frac{R_2R_3R_4}{\sqrt{N}}\right)(-1)^{2\mathbf{hT}_n}\right] + \beta^- \exp\left[\left(-\frac{1}{2\sqrt{N}}R_1R_2^2 + \frac{R_2R_3R_4}{\sqrt{N}}\right)(-1)^{2\mathbf{hT}_n}\right]$$

$$\beta^{\pm} = \left[1 \pm \frac{R_1}{\sqrt{N}}(-1)^{2\mathbf{hT}_n}\right]^{-1/2}.$$

From the approximation

$$\left[1 \pm \frac{R_1}{\sqrt{N}}(-1)^{2\mathbf{hT}_n}\right]^{-1/2} \simeq 1 \mp \frac{R_1}{2\sqrt{N}}(-1)^{2\mathbf{hT}_n} \simeq \exp\left[\mp\frac{R_1}{2\sqrt{N}}(-1)^{2\mathbf{hT}_n}\right]\qquad (E.5)$$

(E.4) becomes

$$P^{\pm} \simeq \exp(\mp 2B)\cosh\left(\frac{R_5 R_5^{\pm}}{\sqrt{N}}\right)\cosh\left(\frac{R_6 Z_6^{\pm}}{\sqrt{N}}\right)\cosh(\gamma^{\pm}),\qquad (E.6)$$

where

$$\gamma^{\pm} = \frac{1}{\sqrt{N}}\left[\frac{R_5(R_2^2-1)}{2} \pm R_2R_3R_4\right].\qquad (E.7)$$

In order to understand what (E.6) means we expand it in series and obtain

$$P_Q^+ \simeq 0{\cdot}5 + 0{\cdot}5\tanh\left(\frac{R_1R_2R_3R_4}{N}\,\frac{A}{1+B}\right),\qquad (E.8)$$

where $A = \varepsilon_5 + \varepsilon_6 + \varepsilon_2/2$,

$$B = 1 + [\varepsilon_1(\varepsilon_3\varepsilon_5 + \varepsilon_4\varepsilon_6) + \varepsilon_2(\varepsilon_3\varepsilon_4 + \varepsilon_3\varepsilon_6 + \varepsilon_4\varepsilon_5)]/2N.$$

(E.8) tells us that if ε_2 is large enough the quartet is expected to be positive no matter how small the cross-magnitudes R_5 and R_6 are.

Retaining terms up to $1/N$ order we obtain in non-centrosymmetric space groups

$$P(\Phi'|R_1,\ldots,R_6) \simeq \frac{1}{L}\exp(-4B\cos\Phi) \times I_0\left(\frac{2X_2}{\sqrt{N}}\right) I_0\left(\frac{2R_5X_5}{\sqrt{N}}\right) I_0\left(\frac{2R_6X_6}{\sqrt{N}}\right),\qquad (E.9)$$

where

$$\Phi' = \varphi_1 - \varphi_2 - \varphi_3 + \varphi_4,\qquad (E.10)$$
$$X_5 = [R_1^2R_3^2 + R_2^2R_4^2 + 2R_1R_2R_3R_4\cos\Phi']^{1/2},$$
$$X_6 = [R_1^2R_4^2 + R_2^2R_3^2 + 2R_1R_2R_3R_4\cos\Phi']^{1/2},$$
$$X_2 = \left[\frac{R_1^2}{4}(R_2^2-1)^2 + R_2^2R_3^2R_4^2 + R_1(R_2^2-1)R_2R_3R_4\cos\Phi'\right]^{1/2},$$

L is a normalizing parameter which does not depend on Φ', I_0 is the modified Bessel function of order zero and $R_1\ldots,R_6$ denote diffraction magnitudes according to (E.2). The algebraic form of X_2 and $\gamma^{\pm}$ suggests: (a) special quartets such as (E.10) are on average less gathered around zero than general quartets; (b) their estimation must be carried out by means of special formulae such as (E.4), (E.6), (E.8) or (E.9) rather than by general formulae as given by Hauptman (1975) and Hauptman & Green (1976).

References

Burla, M. C., Polidori, G., Nunzi, A. & Giacovazzo, C. (1978). To be published.

Cochran, W. (1954). *Acta Cryst.* **7**, 581–583.

Cochran, W. & Woolfson, M. M. (1954). *Acta Cryst.* **7**, 450–451.

Cochran, W. & Woolfson, M. M. (1955). *Acta Cryst.* **8**, 1–12.

Giacovazzo, C. (1974). *Acta Cryst.* A**30**, 626–630.

Giacovazzo, C. (1975). *Acta Cryst.* A**31**, 602–609.

Giacovazzo, C. (1976*a*). *Acta Cryst.* A**32**, 958–966.

Giacovazzo, C. (1976*b*). *Acta Cryst.* A**32**, 967–976.

Giacovazzo, C. (1977*a*). *Acta Cryst.* A**33**, 933–944.

Giacovazzo, C. (1977*b*). *Acta Cryst.* A**33**, 531–538.

Giacovazzo, C. (1977*c*). *Acta Cryst.* A**33**, 539–547.

Hauptman, H. (1975). *Acta Cryst.* A**31**, 680–687.

HAUPTMAN, H. & GREEN, E. A. (1976). *Acta Cryst.* A**32**, 45–49.
HAUPTMAN, H. & KARLE, J. (1953). *Solution of the Phase Problem. I. The Centrosymmetric Crystal.* ACA Monograph No. 3. Pittsburgh: Polycrystal Book Service.
HAUPTMAN, H. & KARLE, J. (1957). *Acta Cryst.* **10**, 267–270.
HURT, M. F. & WAID, C. (1970). *SIAM* (*Soc. Ind. Appl. Math.*) *J. Appl. Math.* **19**, 547–550.
KLUG, A. (1958). *Acta Cryst.* **11**, 513–543.
NAYA, S., NITTA, I. & ODA, T. (1964). *Acta Cryst.* **17**, 421–433.
SHELDRICK, G. M. (1976). Intercongr. Symp. Direct Methods Crystallogr., Buffalo, Abstr. SE2.

Acta Cryst. (1978). A**34**, 574–578

Electronic Polarizabilities of Ions in Doubly Refracting Crystals

BY DIETER POHL

Mineralogisch-Petrographisches Institut der Universität Hamburg, Grindelallee 48, *D*2000 *Hamburg* 13, *Federal Republic of Germany*

(*Received* 20 *September* 1977; *accepted* 7 *March* 1978)

A method is described for determining electronic polarizabilities of ions in doubly refracting ionic crystals solely from crystal data. An expression for computing polarizabilities by a least-squares fit can be derived. Such a method is used to obtain values for the polarizability of cations for the sodium *D* line in Å^3: Na^+ 1·7, K^+ 11·6, Rb^+ 19·1, Tl^+ 48·7, Ca^{2+} 5·2, Sr^{2+} 11·1, Ba^{2+} 23·2, Pb^{2+} 38·3. Evidence is given of decreasing polarizability of the O^{2-} ion in aragonite-type carbonates with decreasing cation sizes.

Introduction

Tessman, Kahn & Shockley (1953) (TK & S) reported a method for evaluating electronic polarizability values of ions in ionic crystals from both optical and structural data. By applying their method to a great number of crystals, they obtained a list of polarizability values of ions. TK & S polarizability values differ considerably from polarizability values that have been otherwise determined (Pauling, 1927; Born & Heisenberg, 1924; Mayer & Goeppert-Mayer, 1933; Fajans & Joos, 1924; Langhoff, 1965; Cohen, 1965, 1966; Lahiri & Mukherji, 1967). The cation polarizabilities are generally higher, and the anion polarizabilities lower, than those obtained by other methods for gaseous ions. Efforts have been made to account for these differences (Ruffa, 1963; Jain, Shanker & Khandelwal, 1975), as well as to develop the TK & S method so as to reduce the discrepancies (Pirenne & Kartheuser, 1964). As pointed out by Batsanov (1966), the main deficiency of the TK & S method stems from neglect of the fact that small cations tend to reduce the polarizability of the anions. The phenomenon of *interacting individual ions*, as suggested by Pirenne & Kartheuser (1964), is not sufficiently explained though it improves the correspondence between the polarizabilities of ions in ionic crystals and the polarizabilities of gaseous ions.

Since TK & S and Pirenne & Kartheuser (1964) restricted themselves to isotropic crystals, they could only determine the sum of the polarizabilities from the refractive index. They lacked the additional data required to evaluate the individual polarizabilities. The difficulty of having more adjustable parameters than experimental measurements may be handled, at least in principle, by extending the TK & S method to doubly refracting crystals. The optical properties of a diatomic crystal, for example, can be described sufficiently by two polarizabilities while there are one, two or three indices of refraction measured, depending on symmetry. It has been suggested (Batsanov, 1966) that the use of salts with complex oxygen-containing anions could be advantageous in determining cation polarizabilities, since the polarizabilities of the complex anions alter but little with the cation sizes. With these arguments in mind, I decided to examine some sulfates and carbonates to obtain values for electronic polarizabilities of ions in doubly refracting crystals. In this paper a method is presented for determining the polarizabilities of ions in ionic crystals. The resulting polarizabilities do not depend as they do in the work of TK & S and of Pirenne & Kartheuser (1964) on an arbitrary selection of the electronic polarizability of one of the ions. Moreover, we shall see that the trend of decreasing polarizability of the anions with decreasing sizes of the cations is confirmed. The calculations remove the discrepancies between the electronic polarizabilities of ions in crystals and the values for gaseous ions.