

A General Approach to Phase Relationships: The Method of Representations

BY CARMELO GIACOVAZZO

Istituto di Mineralogia e Petrografia, Università degli Studi, Bari, Italy

(Received 16 December 1976; accepted 24 April 1977)

A theory is described which, for any universal structure invariant or structure seminvariant, Φ , is able to arrange the set of reflexions in shells, each shell a subset of the succeeding one, with the property that Φ may be estimated in terms of the magnitudes constituting any shell. The theory uses the idea that numberless representations exist of any universal structure invariant or seminvariant, each of them able to contribute to the estimate of Φ .

Introduction

The joint probability distribution of a set of n normalized structure factors was first introduced by Hauptman & Karle (1953). The importance of their work lies in the introduction of the idea that a certain phase or combination of phases may be calculated when other related structure factors have their observed values.

A crucial result for the probabilistic approach to the phase solution was the discovery of the properties of the invariant and seminvariant phases. Hauptman & Karle (1956, 1959) showed that in any space group linear combinations of phases exist whose cosines are in principle fixed by the $|E|$ magnitudes alone (universal cosine invariants) or by the $|E|$'s and the trigonometric form of the structure factor (cosine seminvariant). Thus the individual phases are uniquely determined by the values of the cosine invariants and seminvariants. This result greatly stimulated the calculation of the conditional distribution functions

$$P(\Phi | R_1, \dots, R_n),$$

where Φ is a seminvariant or invariant phase and R_j is the modulus of the j th normalized structure factor E_j . It has also become clear that the nature and the dependence of Φ on the fixed magnitudes R are strongly related to the mutual correlation of the magnitudes. Therefore, a properly chosen set of structure factors must be identified in order to obtain reliable estimates of $\cos \Phi$. Obviously this set is not unique. In general, several sets may be found, each of them giving rise to a different dependence of Φ on the selected magnitudes.

The formulation of the 'nested neighbourhoods principle' first fixed the idea of defining a sequence of sets of reflexions (sequence of nested neighbourhoods), each contained within the succeeding one and having the property that the cosine invariant or seminvariant may be estimated *via* the magnitudes constituting any neighbourhood.

Heuristic methods of finding sequences of nested neighbourhoods for certain universal structure invariants or structure seminvariants have recently been presented by Hauptman (1976). However, different

sequences for the same seminvariants or universal structure invariants exist. It seems then useful to have a theory which is able, for any universal structure invariant or structure seminvariant, to arrange in a general way the set of the reflexions in a sequence of subsets whose order is that of the expected effectiveness (in the statistical sense) for the estimation of Φ . The description of such a theory is the first aim of this paper. In particular, the theory introduces the idea of the upper representations of a universal structure invariant or structure seminvariant. It will lead us to organize the set of reflexions in a sequence of subsets, each contained in the succeeding one, which does not coincide in general with the corresponding nested neighbourhood sequence given by Hauptman. These subsets will be called phasing shells in order to stress this difference.

The theory of the representations has given the author new insights into direct methods of phase determination and has already proved to be a fruitful field of investigation (Giacovazzo, 1976*a*, 1977*a, b, c*). Future publications will deal with the practical applications of these ideas.

Abbreviations

u.s.i. = universal structure invariant, s.s. = structure seminvariant.

1. Preliminary definitions

Let $\{W\}$ be the set of the R magnitudes provided experimentally by diffraction or suitably chosen by the crystallographer. $\{W\}$ will be called the sample space. Define the n -fold Cartesian product

$$\{C\}_n = \{W\} \times \{W\} \times \dots \times \{W\}$$

to be the collection of all ordered n -ples

$$(R_{\mathbf{h}_1}, R_{\mathbf{h}_2}, \dots, R_{\mathbf{h}_n}) \quad (1)$$

where $\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_n$ are reciprocal vectors. Suppose next that the subset $\{Q\}_n$ of the points in $\{C\}_n$ consists of all the n -ples (1) for which

$$\Phi = \sum_{i=1}^n A_i \varphi_{\mathbf{h}_i} \quad [A_i \equiv 0 \pmod{1}], \quad (2)$$

is a s.s. for the actual space group. Of course

$$\sum_{i=1}^t A_i = n.$$

Thus the sets $\{Q\}_1, \{Q\}_2, \{Q\}_3, \dots$ may be formed.

For every set $\{Q\}_n$ there is associated the set $\{S\}_n$ which is the collection of the linear combinations of phases Φ defined by (2).

A subset of $\{Q\}_n$ is the $\{P\}_n$ set which consists of all the n -ples (1) for which

$$\sum A_i h_i = 0 \quad [i = 1, \dots, t; A_i \equiv 0 \pmod{1}].$$

For every set $\{P\}_n$ there is associated the set $\{U\}_n$, which is the collection of the linear combination of phases which are u.s.i.

For the sake of clarity we observe that $\{P\}_3$ coincides with the $\Sigma_1 + \Sigma_2$ listing, $\{P\}_4$ with the quartet listing, $\{Q\}_2$ with the listing of the two-phase seminvariants, etc.

2. A general phase-interrelationship principle

Let

$$\Phi_j = \sum_{i=1}^t A_i \varphi_{h_i}, \quad j = 1, \dots, n,$$

be n u.s.i. or s.s. for the actual space group which satisfy the condition

$$\sum_{j=1}^n B_j \sum_{i=1}^t A_i \varphi_{h_i} \equiv a \pmod{2\pi},$$

where a and the B 's are fixed values. The Φ 's may in general belong to different subsets $\{S\}_n$ and $\{U\}_n$.

Denote now by $\{R\}_j, j = 1, \dots, n, n$ subsets of the sample space $\{W\}$ such that any distribution of the system

$$P_1(\Phi_1 | \{R\}_1) \tag{3a}$$

$$P_2(\Phi_2 | \{R\}_2) \tag{3b}$$

$$\dots \dots \dots$$

$$P_n(\Phi_n | \{R\}_n) \tag{3n}$$

can in principle be different from $1/2\pi$. Furthermore, the intersections

$$\{R\}_i \cap \{R\}_j, \quad i, j = 1, \dots, n,$$

may or may not be the empty set. Denote next by $\{R\}_U$ the set union

$$\{R\}_U = \{R\}_1 \cup \{R\}_2 \cup \dots \cup \{R\}_n.$$

If $\{R\}_U$ does not coincide with $\{R\}_1$ then

$$P_1(\Phi_1 | \{R\}_U) \tag{4}$$

will not coincide in general with $P_1(\Phi_1 | \{R\}_1)$ and it is able to provide estimates of Φ_1 more accurate (in the statistical sense) than those given by (3a).

This principle may be justified by observing that, from the probability densities (3) the supplementary probability density

$$P(B_1 \Phi_1 + B_2 \Phi_2 + \dots + B_n \Phi_n | \{R\}_U)$$

arises, which is a two-valued function. It equals unity when

$$B_1 \Phi_1 + B_2 \Phi_2 + \dots + B_n \Phi_n \equiv a \pmod{2\pi},$$

and zero in all other cases. In other words, if $\Phi_2, \Phi_3, \dots, \Phi_n$ are contemporaneously estimated *via* the sets $\{R\}_2, \{R\}_3, \dots, \{R\}_n$, a consequent effect is the phase indication

$$B_1 \Phi_1 = a - B_2 \Phi_2 - \dots - B_n \Phi_n$$

which is supplementary to that given by (3a).

3. The first representation of a s.s.

Let us denote by $C_p \equiv (\mathbf{R}_p, \mathbf{T}_p), p = 1, \dots, m$, the m symmetry operators (\mathbf{R}_p rotation component, \mathbf{T}_p translation component) of the actual space group and suppose that

$$\Phi = A_1 \varphi_{h_1} + A_2 \varphi_{h_2} + \dots + A_n \varphi_{h_n} \tag{5}$$

be a s.s. Let us suppose that at least one phase φ_h and two symmetry operators C_p and C_q exist in principle (\mathbf{R}_h may be or may not be experimentally measured) such that

$$\begin{aligned} \Psi_1 &= \Phi' + \varphi_{h\mathbf{R}_p} - \varphi_{h\mathbf{R}_q} \\ &= A_1 \varphi_{h_1 \mathbf{R}_p} + A_2 \varphi_{h_2 \mathbf{R}_p} + \dots + A_n \varphi_{h_n \mathbf{R}_p} + \varphi_{h\mathbf{R}_p} - \varphi_{h\mathbf{R}_q} \end{aligned} \tag{6}$$

is a u.s.i.. Φ is then a s.s. of first rank. The collection of the s.s.'s of first rank will be denoted by $\{S_1\}_v$, where $v = A_1 + A_2 + \dots + A_n$; obviously $\{S_1\}_v \in \{S\}_v$. The set $\{S\}_v - \{S_1\}_v$ may or may not be the empty set. For example, in $P2$ all the s.s.'s are of first rank. In fact, in this space group Φ is a s.s. if

$$\sum A_i h_i \equiv 0 \pmod{(2, 0, 2)}.$$

Denoting

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

we have

$$\mathbf{h}(\mathbf{I} - \mathbf{R}_2) \equiv 0 \pmod{(2, 0, 2)},$$

from which (6) may be easily stated by choosing $\mathbf{R}_p = \mathbf{I}, \mathbf{R}_q = \mathbf{R}_2$.

As a further example, $\varphi_{-h_1 \mathbf{R}_s + h_2 \mathbf{R}_v} + \varphi_{h_1 - h_2}$ is a s.s. of first rank for all the space groups which have the symmetry operators C_s and C_v (Giacovazzo, 1977b, c). The vectors \mathbf{h} and the expressions of Φ' for which Ψ_1 is a u.s.i. are defined by

$$\begin{aligned} (a) \quad \Phi' &= \varphi_{-h_1 \mathbf{R}_s + h_2 \mathbf{R}_v} + \varphi_{(h_1 - h_2) \mathbf{R}_v}; \\ &\quad \mathbf{h} = \mathbf{h}_1 + \mathbf{k}; \quad \mathbf{R}_p = \mathbf{R}_v; \quad \mathbf{R}_q = \mathbf{R}_s; \end{aligned}$$

\mathbf{k} is a vector for which $\mathbf{k}(\mathbf{R}_v - \mathbf{R}_s) = 0$;

$$\begin{aligned} (b) \quad \Phi' &= \varphi_{-h_1 \mathbf{R}_s + h_2 \mathbf{R}_v} + \varphi_{(h_1 + h_2) \mathbf{R}_s}; \\ &\quad \mathbf{h} = \mathbf{h}_2 + \mathbf{k}; \quad \mathbf{R}_p = \mathbf{R}_s; \quad \mathbf{R}_q = \mathbf{R}_v; \end{aligned}$$

\mathbf{k} is the same vector as in (a).

If the conditions (a) and (b) hold we obtain the respective invariants

$$\Psi'_1 = \varphi_{-h_1R_s + h_2R_v} + \varphi_{(h_1-h_2)R_v} + \varphi_{(h_1+k)R_s} - \varphi_{(h_1+k)R_v}$$

$$\Psi''_1 = \varphi_{-h_1R_s + h_2R_v} + \varphi_{(h_1-h_2)R_s} + \varphi_{(h_2+k)R_s} - \varphi_{(h_2+k)R_v}$$

We note in (6) that Ψ_1 differs from Φ by a constant which arises because of the translational symmetry. In fact, as

$$\varphi_{hR} = \varphi_h - 2\pi hT, \tag{7}$$

then

$$\Psi_1 - \Phi = -2\pi[A_1h_1T_s + A_2h_2T_t + \dots + A_nh_nT_v + h(T_p - T_q)],$$

which is a constant if the trigonometric form of the structure factor has been fixed.

Suppose now that Φ is a s.s. for which (6) cannot be stated: two phases φ_h and φ_l and four symmetry operators C_p, C_q, C_i, C_j exist in principle (R_h and R_l may or may not be experimentally measured) such that

$$\Psi_1 = \Phi' + \varphi_{hR_p} - \varphi_{hR_q} + \varphi_{lR_i} - \varphi_{lR_j}$$

$$= A_1\varphi_{h_1R_s} + \dots + A_n\varphi_{h_nR_v} + \varphi_{hR_p} - \varphi_{hR_q} + \varphi_{lR_i} - \varphi_{lR_j} \tag{8}$$

is a u.s.i.. Φ is then a s.s. of second rank. The collection of the s.s.'s of second rank will be denoted by $\{S_{II}\}_v$, where $v = A_1 + A_2 + \dots + A_n$; obviously $\{S_{II}\}_v \in \{S\}_v$.

As an example, the reader will easily verify that in $P2_12_12_1$ the s.s.'s for which

$$\sum A_i h_i \equiv 0 \pmod{(0,2,2)}$$

$$\sum A_i h_i \equiv 0 \pmod{(2,0,2)}$$

or

$$\sum A_i h_i \equiv 0 \pmod{(2,2,0)}$$

are of first rank, whereas s.s.'s for which

$$\sum A_i h_i \equiv 0 \pmod{(2,2,2)}$$

are of second rank. A short survey of the rank of the s.s.'s in all the space groups is made in the Appendix.

We conclude by stating that each element of the set $\{S_I\}_v$ has its first representation (or representation of first order) in the set $\{U\}_{v+2}$. Each element of the set $\{S_{II}\}_v$ has its first representation in the set $\{U\}_{v+4}$. More precisely, the first representation of a s.s. of first rank $\Phi \in \{S\}_v$ is the collection of the u.s.i.'s $\Psi_1 \in \{U\}_{v+2}$ which are carried out by means of (6). The first representation of a s.s. of second rank $\Phi \in \{S_{II}\}_v$ is the collection of the u.s.i.'s $\Psi_1 \in \{U\}_{v+4}$ which are carried out by means of (8). We will denote the first representation of a s.s. Φ by $\{\Psi\}_1$, and the generic element of $\{\Psi\}_1$ by Ψ_1 .

In accordance with this notation, for example, the first representation in $P\bar{1}$ of the two-phase s.s. of first rank $\Phi = \varphi_{h+k} + \varphi_{h-k}$ coincides with the collection of the two quartet invariants

$$\Psi'_1 = \varphi_{h+k} + \varphi_{h-k} - \varphi_h - \varphi_h,$$

$$\Psi''_1 = \varphi_{h+k} - \varphi_{h-k} - \varphi_k - \varphi_k.$$

Again, the first representation in $P\bar{1}$ of the three-phase

s.s. $\Phi = \varphi_h + \varphi_k + \varphi_{h+k+2l}$ is the collection of the four quintet invariants

$$\Psi'_1 = \varphi_h + \varphi_k + 2\varphi_l - \varphi_{h+k+2l},$$

$$\Psi''_1 = -\varphi_h + \varphi_k + 2\varphi_{h+l} - \varphi_{h+k+2l},$$

$$\Psi'''_1 = \varphi_h - \varphi_k + 2\varphi_{k+l} - \varphi_{h+k+2l},$$

$$\Psi''''_1 = -\varphi_h - \varphi_k + 2\varphi_{h+k+l} - \varphi_{h+k+2l}.$$

In conclusion, the first representation in $P2_12_12_1$ of the one-phase s.s. of second rank $\Phi = \varphi_h [h \equiv 0 \pmod{(2,2,2)}]$ is the collection of quintet invariants

$$\Psi = \varphi_h + \varphi_{kR_i} - \varphi_{kR_j} + \varphi_{lR_p} - \varphi_{lR_q}$$

for which

- (a) $h + k(R_i - R_j) + l(R_p - R_q) = 0,$
- (b) the vectors $k(R_i - R_j)$ and $l(R_p - R_q)$ satisfy one of the following conditions:

- (1) $k(R_i - R_j) \equiv 0 \pmod{(2,2,0)},$
 $l(R_p - R_q) \equiv 0 \pmod{(0,2,2)},$
- (2) $k(R_i - R_j) \equiv 0 \pmod{(2,2,0)},$
 $l(R_p - R_q) \equiv 0 \pmod{(2,0,2)},$
- (3) $k(R_i - R_j) \equiv 0 \pmod{(0,2,2)},$
 $l(R_p - R_q) \equiv 0 \pmod{(2,0,2)}.$

In conclusion we note that, in terms of the phase-interrelationship principle described in §3 the realization of any Ψ_1 starting from a s.s. of first rank involves in (3)

$$\Phi_1 = \Phi = A_1\varphi_{h_1} + A_2\varphi_{h_2} + \dots + A_n\varphi_{h_n},$$

$$\Phi_2 = \Psi_1 = -A_1\varphi_{h_1R_s} - A_2\varphi_{h_2R_t} - \dots - A_n\varphi_{h_nR_v} + \varphi_{hR_p} - \varphi_{hR_q}.$$

A similar observation holds for the s.s.'s of second rank.

4. The multipoles

The results obtained in §3 suggest a simplification in the practical use of the phase-interrelationship principle. As each u.s.i. Ψ_1 differs from Φ by a constant angle, one can state, for each seminvariant Φ in (3), the equivalent

$$P(\Phi | \{R\}) = P(\Psi_1 | \{R\}).$$

Thus, without any loss of generality, one may consider that in (3) all the Φ 's belong to the sets $\{U\}$.

We state then that the set

$$\Phi_1, \Phi_2, \dots, \Phi_n$$

constitutes a multipole. Some multipoles have already had specific denominations: for example, the quadripoles (four triplet seminvariants), the hexapoles (six triplet invariants), the tripoles (three quartet invariants).

5. The first representation of a u.s.i. The basis magnitudes of a u.s.i. or s.s.

Let Φ , as given by (5), be a u.s.i. If the magnitudes $R_{h_i}, i = 1, \dots, n$ are known, one is always able to derive a

first estimation of Φ . $\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_n$ will be called the basis vectors and the $R_{\mathbf{h}_j}$ the basis magnitudes of Φ .

If the crystal symmetry is higher than triclinic a number of symmetry operators may be found in favourable cases such that one or more u.s.i.,

$$\begin{aligned}\Psi_1 &= A_1\varphi_{\mathbf{h}'_1} + A_2\varphi_{\mathbf{h}'_2} + \dots + A_n\varphi_{\mathbf{h}'_n} \\ &= A_1\varphi_{\mathbf{h}_1\mathbf{R}_s} + A_2\varphi_{\mathbf{h}_2\mathbf{R}_t} + \dots + A_n\varphi_{\mathbf{h}_n\mathbf{R}_v},\end{aligned}\quad (9)$$

arise in which at least one of the \mathbf{h}'_j vectors does not coincide with \mathbf{h}_j . Because of (7) $\Psi_1 - \Phi$ is a constant if the geometrical form of the structure factor has been fixed. The collection of the distinct u.s.i., Ψ_1 , obtained when $\mathbf{R}_s, \mathbf{R}_t, \dots, \mathbf{R}_v$ vary in the set of the m rotation matrices of the actual space group, is defined to be the first representation of Φ and will be denoted by $\{\Psi\}_1$. Any one of the sets

$$\mathbf{h}'_1, \mathbf{h}'_2, \dots, \mathbf{h}'_n$$

may be defined to be the set of the basis vectors of $\{\Psi\}_1$ and the corresponding magnitudes to be the basis magnitudes of $\{\Psi\}_1$. For example (Giacovazzo, 1976b), the first representation in $P2_1/c$ of the quartet

$$\Phi = \varphi_{234} + \varphi_{112} + \varphi_{2\bar{1}3} - \varphi_{539}$$

contains, in addition to Φ , also

$$\Psi_1 = \varphi_{234} + \varphi_{1\bar{1}2} + \varphi_{213} - \varphi_{539}.$$

If Φ is a s.s., it is not possible from only the knowledge of the $R_{\mathbf{h}_i}$'s, to obtain information about Φ . For example, the estimation of the phase $\varphi_{2\mathbf{h}}$ in $P\bar{1}$ requires, in addition to $R_{2\mathbf{h}}$, that the magnitude $R_{\mathbf{h}}$ is known. In the same way, one needs to know in $P\bar{1}$ at least $R_{\mathbf{h}_1}$ or $R_{\mathbf{h}_2}$, in addition to $R_{\mathbf{h}_1+\mathbf{h}_2}$ and $R_{\mathbf{h}_1-\mathbf{h}_2}$, in order to estimate $\varphi_{\mathbf{h}_1+\mathbf{h}_2} - \varphi_{\mathbf{h}_1-\mathbf{h}_2}$.

This observation may be generalized by observing that at least the basis magnitudes of one u.s.i. Ψ_1 must be known in order to derive a first estimate of Φ .

We state that the basis vectors of all the u.s.i. which constitute $\{\Psi\}_1$ are the basis vectors of the first representation of Φ and that the corresponding magnitudes are the basis magnitudes of the first representation of Φ . For example, if

$$\Phi = \varphi_{-\mathbf{h}_1\mathbf{R}_s + \mathbf{h}_2\mathbf{R}_v} + \varphi_{\mathbf{h}_1 - \mathbf{h}_2},$$

then $R_{-\mathbf{h}_1\mathbf{R}_s + \mathbf{h}_2\mathbf{R}_v}$, $R_{\mathbf{h}_1 - \mathbf{h}_2}$ and all the magnitudes $R_{\mathbf{h}_1+\mathbf{k}}$, $R_{\mathbf{h}_2+\mathbf{k}}$ for which $\mathbf{k}(\mathbf{R}_v - \mathbf{R}_s) = 0$ are basis magnitudes of the first representation of Φ .

6. The cross-magnitudes

As well known (Schenk, 1973; Hauptman, 1975; Giacovazzo, 1975a) the cross-vectors of the quartet invariant

$$\Phi = \varphi_{\mathbf{h}_1} + \varphi_{\mathbf{h}_2} + \varphi_{\mathbf{h}_3} + \varphi_{-\mathbf{h}_1 - \mathbf{h}_2 - \mathbf{h}_3}$$

may be defined in $P1$ or $P\bar{1}$ as the collection of the distinct vectors

$$\mathbf{h}_i + \mathbf{h}_j, \quad (i, j = 1, \dots, n) \quad (10)$$

where $n=4$. This definition is quite valid also for the quintet invariant ($n=5$)

$$\Phi = \varphi_{\mathbf{h}_1} + \varphi_{\mathbf{h}_2} + \varphi_{\mathbf{h}_3} + \varphi_{\mathbf{h}_4} + \varphi_{-\mathbf{h}_1 - \mathbf{h}_2 - \mathbf{h}_3 - \mathbf{h}_4},$$

(see Hauptman & Fortier, 1977).

For reasons which will become evident in § 14, the set of the cross-vectors of the u.s.i. (5) is defined to be the collection of the distinct vectors

$$m_1\mathbf{h}_1 + \dots + m_n\mathbf{h}_n \quad (m_p = 0, \dots, A_p), \quad (11)$$

which do not coincide with any basis vector of Φ or with 0. The corresponding magnitudes are the cross-magnitudes of Φ . (10) and (11) are not equivalent in general. For example

$$\Phi = 2\varphi_{\mathbf{h}_1} + 3\varphi_{\mathbf{h}_2} - \varphi_{2\mathbf{h}_1 + 3\mathbf{h}_2}$$

has the following cross-magnitudes

$$\begin{aligned}R_{2\mathbf{h}_1}, R_{\mathbf{h}_1 + \mathbf{h}_2}, R_{\mathbf{h}_1 + 3\mathbf{h}_2}, R_{2\mathbf{h}_2}, R_{2\mathbf{h}_1 + 2\mathbf{h}_2}, R_{3\mathbf{h}_2}, \\ R_{2\mathbf{h}_1 + 3\mathbf{h}_2}, R_{2\mathbf{h}_1 + \mathbf{h}_2}, R_{\mathbf{h}_1 + 2\mathbf{h}_2}\end{aligned}$$

of which only the first five are obtained according to (10). A further generalization is suitable for including crystal symmetries of higher order than triclinic. As stated in § 5, in fact, one or more combinations of symmetry operators may be found in favourable cases such that one or more u.s.i. (9) arise, in which at least one of the \mathbf{h}'_j vectors does not coincide with \mathbf{h}_j . Not all the cross-magnitudes of Ψ_1 as given by (9) will coincide then with that of Φ . However, as $\Psi_1 - \Phi$ is a constant which arises because of translational symmetry, the eventual estimation of any Ψ_1 in terms of its basis and cross-magnitudes defines also Φ . In other words, the cross-magnitudes of Ψ_1 can be considered in every practical respect to be cross-magnitudes of Φ . We state then that the collection of the distinct vectors, which are cross-vectors of at least one u.s.i. Ψ_1 and do not coincide with any basis vectors or with 0, constitute the set of the cross-vectors of Φ . Or, in accordance with § 5, they constitute the set of the cross-vectors of $\{\Psi\}_1$. The corresponding magnitudes are the cross-magnitudes of $\{\Psi\}_1$. It should be useful to note that these definitions are of more than passing interest if one thinks that all the cross-vectors of the first representation of Φ , together with its basis magnitudes, contribute to the estimate of Φ [see Giacovazzo (1976b) for the quartet invariants].

If Φ is a s.s. of first rank let us denote in (6)

$$\begin{aligned}\mathbf{h}'_1 &= \mathbf{h}_1\mathbf{R}_s, \mathbf{h}'_2 = \mathbf{h}_2\mathbf{R}_t, \dots, \\ \mathbf{h}'_n &= \mathbf{h}_n\mathbf{R}_v, \mathbf{h}'_{n+1} = \mathbf{h}\mathbf{R}_p, \mathbf{h}'_{n+2} = \mathbf{h}\mathbf{R}_q; \\ A_{n+1} &= 1, \quad A_{n+2} = -1.\end{aligned}$$

Of the set of distinct vectors

$$m_1\mathbf{h}'_1 + \dots + m_{n+2}\mathbf{h}'_{n+2}, \quad (m_p = 0, \dots, A_p)$$

which are obtained when Ψ_1 varies throughout $\{\Psi\}_1$, those which do not coincide with any basis vector of Φ (or with 0) are the cross-vectors of Φ . The corre-

sponding magnitudes are the cross-magnitudes of Φ . A similar definition may be easily stated for the s.s. of second rank.

7. Phasing magnitudes of a u.s.i. or of a s.s.

The basis and the cross-magnitudes of $\{\Psi\}_1$ are defined to be the phasing magnitudes of $\{\Psi\}_1$. The set of the phasing magnitudes of $\{\Psi\}_1$ will be denoted by $\{B\}_1$. For example, the phasing magnitudes of $\{\Psi\}_1$ for the s.s.

$$\Phi = \varphi_h + \varphi_k + \varphi_{h+k+2l}$$

in $P\bar{I}$ are the 17 magnitudes [compare with the 13 magnitudes given by Hauptman (1976) in the first two neighbourhoods]:

$$\begin{aligned} &R_h, R_k, R_{h+k+2l}, R_l, R_{h+l}, R_{k+l}, R_{h+k+l}, R_{h+k}, R_{h-k}, \\ &R_{h+2l}, R_{k+2l}, R_{h+2k+2l}, R_{2h+k+2l}, R_{2l}, R_{2h+2l}, \\ &R_{2k+2l}, R_{2h+2k+2l}. \end{aligned}$$

8. The order of the phase relationship associated with a s.s.

It is the practice of the crystallographer to state that a triplet relationship is a phase of order $1/\sqrt{N}$, a quartet is a phase relationship of order $1/N$, a quintet of order $1/N\sqrt{N}$, etc. This may be easily justified in a qualitative way from the point of view of the joint probability distribution methods. Let

$$C(\varrho_1, \dots, \varrho_n, \Psi_1, \dots, \Psi_n) \tag{12}$$

be the characteristic function of the distribution

$$P(R_1, \dots, R_n, \varphi_1, \dots, \varphi_n). \tag{13}$$

We may write (Giacovazzo, 1977c)

$$C = \exp \left\{ \left[-\frac{1}{2}(\varrho_1^2 + \dots + \varrho_n^2) \right] + \sum_{\substack{v \\ 3 \nmid v}}^{\infty} \frac{S_v}{N^{(v/2-1)}} \right\}, \tag{14}$$

where

$$S_v = \sum_{r+s+\dots+w=v} \frac{\lambda_{rs\dots w}}{r!s!\dots w!} (i\varrho_1 \cos \Psi_1)^r (i\varrho_2 \cos \Psi_2)^s \dots (i\varrho_n \sin \Psi_n)^w.$$

$\lambda_{rs\dots w}$ are the standardized cumulants of the distribution.

By a Gram-Charlier expansion of (14) one obtains

$$\begin{aligned} C = \exp \left[-\frac{1}{2}(\varrho_1^2 + \dots + \varrho_n^2) \right] &\left\{ 1 + \frac{S_3}{N^{1/2}} \right. \\ &+ \frac{1}{N} \left(S_4 + \frac{S_3^2}{2} \right) + \frac{1}{N\sqrt{N}} \left(S_5 + S_3 S_4 + \frac{S_3^3}{6} \right) \\ &\left. + \frac{1}{N^2} \left(S_6 + \frac{S_4^2}{2} + S_3 S_5 + \frac{S_3^2 S_4}{2} + \frac{S_3^4}{24} \right) + \dots \right\}, \tag{15} \end{aligned}$$

which is a series organized in a strictly asymptotic series of powers of $N^{-1/2}$. The Fourier transform of (15)

is also organized in an asymptotic series of powers of $1/\sqrt{N}$ and is a useful (for our purposes) approximation of (13). It may be written:

$$\begin{aligned} P(R_1, \dots, R_n, \varphi_1, \dots, \varphi_n) = &\frac{1}{(2\pi)^{2n}} \exp \{ -R_1^2 - \dots - R_n^2 \} \\ &\times \left\{ 1 + \frac{T_3}{\sqrt{N}} + \frac{1}{N} \left(T_4 + \frac{T_3^2}{2} \right) + \dots \right\}. \tag{16} \end{aligned}$$

From (16) one may easily select the terms of the distribution which chiefly contribute to the estimation of a given u.s.i. and recognize their order (in terms of powers of $1/\sqrt{N}$). These terms will be called the phasing terms of the distribution. The minimum order of the phasing terms may be considered the order of the phase relationship.

If a u.s.i. is estimated by a joint probability distribution which involves its phasing magnitudes one easily observes: (a) T_3 is the first term of the distribution which contributes to the estimation of a triplet; (b) T_4 and T_3^2 are the phasing terms for a quartet; (c) T_5 , $T_3 T_4$ and T_3^3 are the phasing terms for a quintet; etc.

In conclusion, one may shortly state that with a u.s.i. Φ is associated a phase relationship of order $1/\sqrt{N}$ if Φ is a triplet invariant, of order $1/N$ if Φ is a quartet invariant, etc. Clearly, the order of a s.s. Φ is that associated with any u.s.i. Ψ_1 . For example, in $P\bar{I}$ with the s.s. $\Phi = \varphi_{2h}$ is associated a phase relationship of order $1/\sqrt{N}$; with $\Phi = \varphi_{h_1+h_2} - \varphi_{h_1-h_2}$ is associated a phase relationship of order $1/N$; etc.

Further considerations need some special u.s.i.'s which play an important role in this paper. We limit ourselves to giving one example: the reader will surely be able to generalize it. Let

$$\Phi = \varphi_{h_1} + \varphi_{h_2} - \varphi_{h_1+h_2}$$

and

$$\Psi_2 = \varphi_{h_1} + \varphi_{h_2} - \varphi_{h_1+h_2} + \varphi_k - \varphi_k$$

be two u.s.i.'s. Ψ_2 is formally a quintet; however, its value is quite definite by the triplet invariant Φ . Of the two distribution functions

$$\begin{aligned} &P(E_{h_1}, E_{h_2}, E_{h_1+h_2}), \\ &P(E_{h_1}, E_{h_2}, E_{h_1+h_2}, E_k, \dots \text{ cross-magnitudes}), \end{aligned}$$

the first estimates Φ in terms of the three magnitudes $|E_{h_1}|, |E_{h_2}|, |E_{h_1+h_2}|$; the second in terms of $|E_{h_1}|, |E_{h_2}|, |E_{h_1+h_2}|, |E_k|$ and of the cross-magnitudes of Ψ_2 . In accordance with the phase-interrelationship principle, the estimate of Φ is more accurate (in a probabilistic sense) when derived from the second than from the first distribution. In particular, the second distribution will give rise to the same phasing terms of order $1/\sqrt{N}$ as from the first distribution, but it will be able to correct them (in the probabilistic sense) by means of magnitudes not included in the first distribution.

9. The upper representations of a u.s.i.

Let

$$\Phi = A_1\phi_{h_1} + A_2\phi_{h_2} + \dots + A_n\phi_{h_n}$$

be a u.s.i.. Then also

$$\Psi_2 = \Psi_1 + \phi_k - \phi_k \tag{17}$$

is a u.s.i.. In (17) $\Psi_1 \in \{\Psi\}_1$ and \mathbf{k} is a free vector. We denote by $\{\Psi\}_2$ the collection of the distinct u.s.i. obtained by (17) when \mathbf{k} varies throughout reciprocal space and Ψ_1 throughout $\{\Psi\}_1$. Ψ_2 will denote the generic u.s.i. belonging to $\{\Psi\}_2$. As any Ψ_2 differs from Φ by a constant arising because of the translational symmetry, $\{\Psi\}_2$ will be called the representation of the second order of Φ .

The set $\{\Psi\}_3$, in its turn, is defined to be the collection of all the distinct u.s.i.

$$\Psi_3 = \Psi_2 + \phi_l - \phi_l = \Psi_1 + \phi_k - \phi_k + \phi_l - \phi_l$$

obtained when Ψ_1 varies in $\{\Psi\}_1$ and \mathbf{k} and \mathbf{l} throughout reciprocal space (or, similarly, Ψ_2 throughout $\{\Psi\}_2$ and \mathbf{l} throughout reciprocal space). The upper representations $\{\Psi\}_4, \{\Psi\}_5, \dots$ are likewise obtained.

As we see, a u.s.i. defined in $\{U\}_v$ has its representation in $\{U\}_v, \{U\}_{v+2}, \{U\}_{v+4}, \{U\}_{v+6}, \dots$

10. The upper representation of a s.s.

In §3 we have already defined the first representation $\{\Psi\}_1$ of a s.s. Φ whatever its rank may be. As $\{\Psi\}_1$ is a collection of u.s.i.'s the procedure fixed in §9 enables us to define $\{\Psi\}_2, \{\Psi\}_3, \{\Psi\}_4, \dots$. In conclusion, one may state that a s.s. defined in $\{S_I\}_v$ has its first representations in $\{U\}_{v+2}$, its second representation in $\{U\}_{v+4}, \dots$. A s.s. defined in $\{S_{II}\}_v$ has its first representation in $\{U\}_{v+4}$, its second in $\{U\}_{v+6}, \dots$

11. Phasing magnitudes of the upper representation of Φ

The definitions given in the §§ 5, 6, 7 enable us to define easily the basis and the cross-reflexions of any

$$\Psi_n = \Psi_1 + \phi_k - \phi_k + \dots + \phi_l - \phi_l.$$

So the phasing magnitudes of the n th representation of Φ are defined to be all the magnitudes which are phasing magnitudes of at least one of the u.s.i.'s Ψ_n . The set of the phasing magnitudes of $\{\Psi\}_n$ will be denoted by $\{B\}_n$.

12. Phasing shells

With any representation $\{\Psi\}_v$ of Φ a real non-negative conditional probability function

$$P(\Phi | \{B\}_v) \tag{18}$$

may in principle be associated, which is able to estimate Φ if the phasing magnitudes of the v th representation are known. As

$$\{B_1\} \subset \{B\}_2 \subset \{B\}_3 \dots, \tag{19}$$

the phase-interrelationship principle assures that (18) is able to give a more accurate estimate (in a probabilistic sense) of Φ than

$$P(\Phi | \{B\}_r)$$

where $r < v$.

Furthermore, (19) suggests that the sample space $\{W\}$ may be arranged for a given Φ in a sequence of phasing shell, each shell a subset of the succeeding one, with the property that Φ may be estimated in terms of the magnitudes constituting any shell.

In accordance with this purpose, we state that $\{B\}_1$ constitutes the first phasing shell. The magnitudes in $\{B\}_2$ which are not in $\{B\}_1$ belong to the second phasing shell. The magnitudes in $\{B\}_3$ which do not belong to $\{B\}_2$ constitute the third phasing shell, etc.

A general representative scheme is shown in Fig. 1.

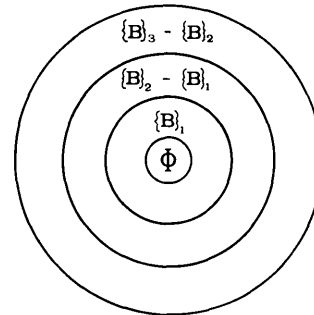


Fig. 1. A general scheme showing the general sequence of the phasing shells for a given u.s.i. or s.s. Φ .

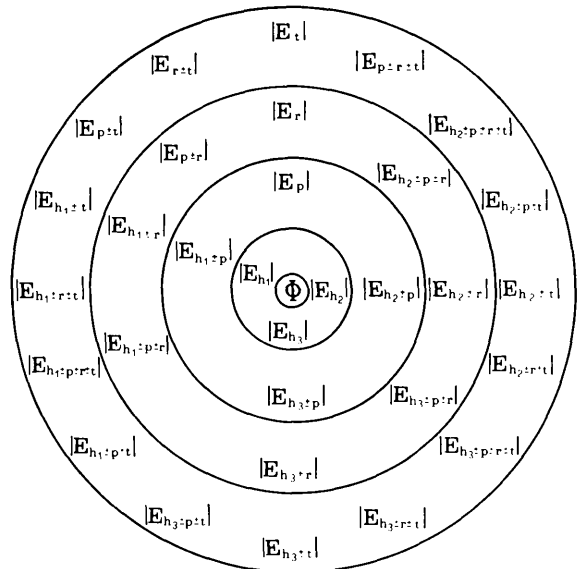


Fig. 2. The sequence of the first four phasing shells in P_I or $P_{\bar{I}}$ for the structure invariant $\Phi = \phi_{h_1} + \phi_{h_2} + \phi_{h_3}$. The reciprocal vectors $\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3$ satisfy $\mathbf{h}_1 + \mathbf{h}_2 + \mathbf{h}_3 = 0$.

In Figs. 2–12 the scheme is explained for some important u.s.i. and s.s.. The reader will surely be able to derive explicit sequences of phasing shells for other u.s.i. or s.s..

13. The estimation of Φ via its upper representations

Each Ψ_n may be estimated *via* its phasing magnitudes. As $\Psi_n - \Phi$ is a constant, the estimation of Ψ_n gives also an estimate of Φ . As the set of the phasing magnitudes of any Ψ_n includes that of Φ , the estimation of Φ in terms of the phasing magnitudes of Ψ_n is more accurate

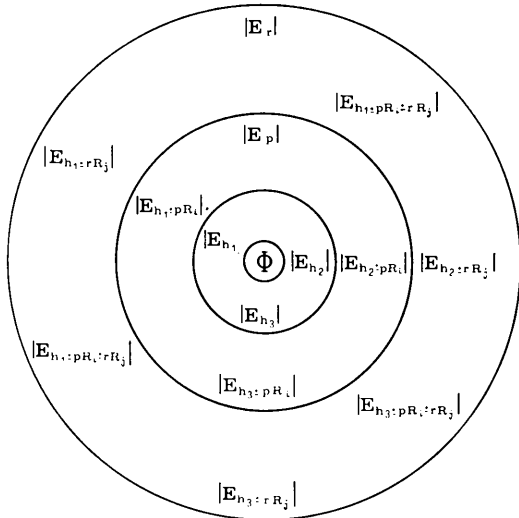


Fig. 3. The sequence of the first three phasing shells in any space group of order m for the structure invariant $\Phi = \varphi_{h_1} + \varphi_{h_2} + \varphi_{h_3}$. The reciprocal vectors h_1, h_2, h_3 satisfy $h_1 + h_2 + h_3 = 0$, p and r are arbitrary vectors. $E_{h_1 \pm p, R_1}, E_{h_1 \pm p, R_2}, \dots$ etc. represent the sets of magnitudes which arise when i, j, \dots vary from 1 to m .

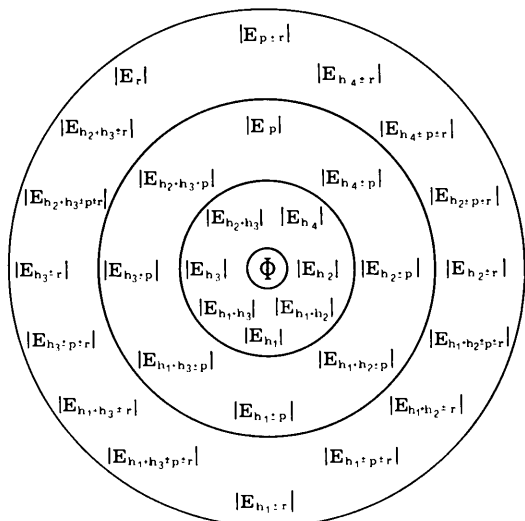


Fig. 4. The sequence of the first three phasing shells in $P1$ and $P\bar{1}$ for the structure invariant $\Phi = \varphi_{h_1} + \varphi_{h_2} + \varphi_{h_3} + \varphi_{h_4}$. The reciprocal vectors h_1, h_2, h_3, h_4 satisfy $h_1 + h_2 + h_3 + h_4 = 0$; the vectors p and r are arbitrary.

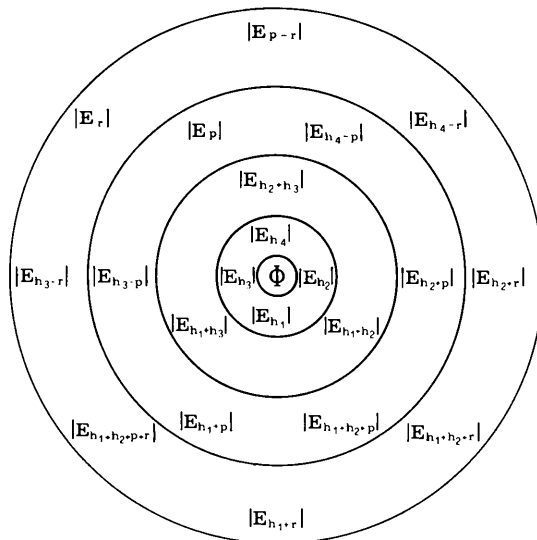


Fig. 5. Hauptman's sequence of nested neighbourhoods for the structure invariant $\Phi = \varphi_{h_1} + \varphi_{h_2} + \varphi_{h_3} + \varphi_{h_4}$. The reciprocal vectors h_1, h_2, h_3, h_4 satisfy $h_1 + h_2 + h_3 + h_4 = 0$; p and r are arbitrary.

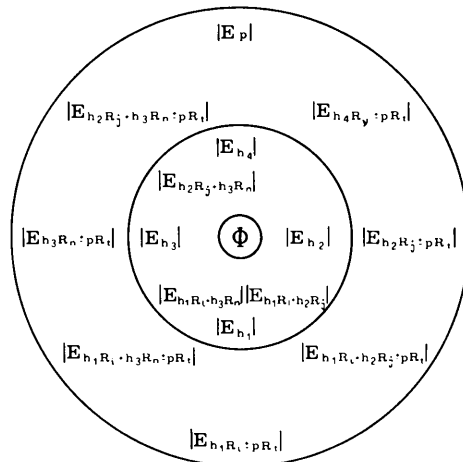


Fig. 6. The sequence of the first two phasing shells in any space group of order m for the structure invariant $\Phi = \varphi_{h_1} + \varphi_{h_2} + \varphi_{h_3} + \varphi_{h_4}$. The reciprocal vectors h_1, h_2, h_3, h_4 satisfy $h_1 + h_2 + h_3 + h_4 = 0$, the vector p is arbitrary. R_1, R_2, R_3, R_4 are rotation matrices of the space group for which $h_1 R_1 + h_2 R_2 + h_3 R_3 + h_4 R_4 = 0$, R_i is an arbitrary rotation matrix of the space group.

(in the probabilistic sense) than that expressible in terms of the sole phasing magnitudes of Φ . The chief advantage of the concept of the upper representations is that Φ may be estimated by collecting the contributions which arise from any element of $\{\Psi\}_n$ and from any $\{\Psi\}_n$.

For example, if one were to estimate in $P1$ $\Phi = \varphi_{h_1} + \varphi_{h_2} - \varphi_{h_1 + h_2}$ *via* its second representation, one could associate with any

$$\Psi_2 = \varphi_{h_1} + \varphi_{h_2} - \varphi_{h_1 + h_2} + \varphi_k - \varphi_k$$

the joint probability distribution function

$$P(E_{h_1}, E_{h_2}, E_{h_1+h_2}, E_k, E_{h_1 \pm k}, E_{h_2 \pm k}, E_{h_1+h_2 \pm k}). \quad (20)$$

Of the phasing magnitudes which appear in (20), the last six are the cross-magnitudes of Ψ_2 . As \mathbf{k} is a free vector, a figure of seven modes arises from (20) which, when \mathbf{k} varies, sweeps out all reciprocal space remaining parallel to itself. In general we conclude that each representation $\{\Psi\}_n$ with $n > 1$ may be described by

means of a geometrical figure which sweeps out the reciprocal space by means of one or more reciprocal vectors. The greater the order of the representation, the greater the number of vertices of the figure and the number of degrees of freedom of the sweeping.

This suggests a further advantage of the idea of the upper representations. In fact, when Φ is estimated by involving its upper representations, a large number of

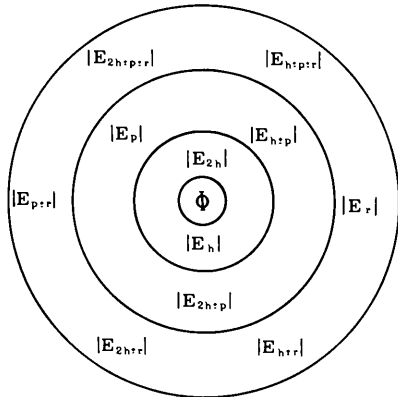


Fig. 7. The first three phasing shells in $P\bar{I}$ for the structure seminvariant φ_{2h} . The reciprocal vectors \mathbf{p} and \mathbf{r} are arbitrary.

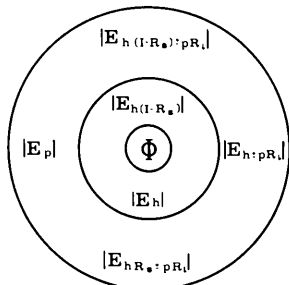


Fig. 8. The first two phasing shells for the structure seminvariant $\varphi_H = \varphi_{h(i-R_s)}$. The reciprocal vector \mathbf{p} is arbitrary. The reciprocal vector \mathbf{h} is arbitrary on condition that $\mathbf{h}(i-R_s) = \mathbf{H}$. $E_{h \pm p R_i}$; $E_{h R_s \pm p R_i}$; $E_{h(i-R_s) \pm p R_i}$ represent the sets of magnitudes which arise when i varies from 1 to m (m is the order of the space group).

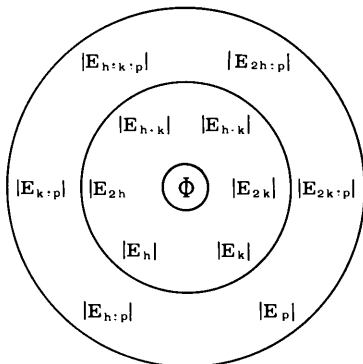


Fig. 9. The first two phasing shells for the structure seminvariant $\Phi_i = \varphi_{h+k} - \varphi_{h-k}$ in $P\bar{I}$. The reciprocal vector \mathbf{p} is arbitrary.

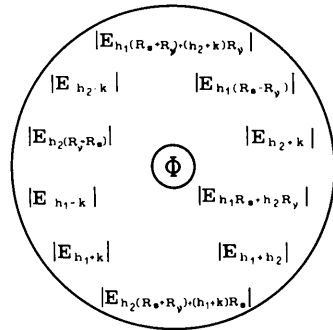


Fig. 10. The first phasing shell for the structure seminvariant $\Phi = \varphi_{h_1 R_s + h_2 R_v} - \varphi_{h_1 + h_2}$. The reciprocal vector \mathbf{k} is arbitrary on condition that $\mathbf{k}(R_v - R_s) = 0$.

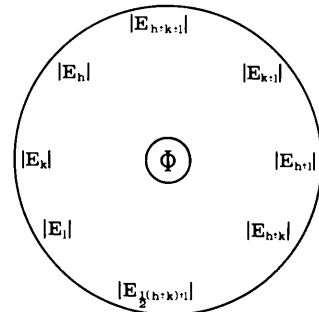


Fig. 11. The first phasing shell of the structure seminvariant $\Phi = \varphi_h + \varphi_k + \varphi_1$ in $P\bar{I}$.

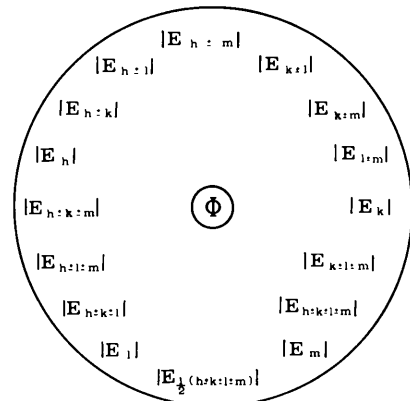


Fig. 12. The first phasing shell of the structure seminvariant $\Phi = \varphi_h + \varphi_k + \varphi_1 + \varphi_m$ in $P\bar{I}$.

magnitudes is checked so that the estimate of Φ is less sensitive to the troubles which arise when the conditions concerning the rational independence between the atomic coordinates are not fulfilled. Furthermore, the method may take noticeable advantage of the actual crystal symmetry. Referring for example to the case of the triplet invariant $\Phi = \varphi_{h_1} + \varphi_{h_2} + \varphi_{h_3}$ in a space group of order m , we observe that any Ψ_2 may be estimated *via* the phasing magnitudes

$$R_{h_1}, R_{h_2}, R_{h_3}, R_k, R_{h_1 \pm kR_p}, R_{h_2 \pm kR_p}, R_{h_3 \pm kR_p}, \quad p = 1, \dots, m.$$

Their number increases with m , which increases also the probability that any Ψ_2 is estimated with good accuracy. In conclusion, the estimation of Φ *via* its second representation may be carried out by restricting k to an asymmetrical region of reciprocal space, on condition that all the cross-magnitudes are taken into account. This conclusion is easily generalizable to other u.s.i.'s or s.s.'s and to representations of higher order.

However, the reader has surely become aware that the calculation of the probability function (17) when v is large may be a difficult task for the crystallographer. Thus it may be helpful to calculate probability functions of type

$$P(\Phi | \{R\}_v),$$

where $\{R\}_v$ is a suitable subset of $\{B\}_v$. From the point of view of the phase-interrelationship principle this procedure may be so justified. The phasing magnitudes of any representation of Φ enjoy the important quality that they rise to a number of multipoles each contributing to the estimate of Φ (see also §14). $\{R\}_v$ is an useful subset if $P(\Phi | \{R\}_v)$ is able to exploit at least one multipole contributing to the estimate of Φ . For example, if one wishes to estimate $\Phi = \varphi_{h_1} + \varphi_{h_2} - \varphi_{h_1+h_2}$ by $\{\Psi\}_2$, the distribution

$$P(E_{h_1}, E_{h_2}, E_{h_1+h_2}, E_k, E_{h_1+k}, E_{h_1+h_2+k}) \quad (21)$$

may be used (Giacovazzo, 1976a) instead of (20). In fact (21) is able to exploit the three (not independent) triplets

$$\begin{aligned} \Phi_1 = \Phi &= \varphi_{h_1} + \varphi_{h_2} - \varphi_{h_1+h_2} \\ \Phi_2 &= -\varphi_{h_1} - \varphi_{h_2} - \varphi_k + \varphi_{h_1+h_2+k} \\ \Phi_3 &= \varphi_{h_1+h_2} + \varphi_k - \varphi_{h_1+h_2+k}, \end{aligned} \quad (22)$$

$$\begin{aligned} \Phi_1 &= \Phi \\ \Phi_4 &= -\varphi_{h_1} + \varphi_{h_1+h_2} + \varphi_{h_1+k} - \varphi_{h_1+h_2+k} \\ \Phi_5 &= -\varphi_{h_2} - \varphi_{h_1+k} + \varphi_{h_1+h_2+k}, \end{aligned} \quad (23)$$

$$\begin{aligned} \Phi_1 &= \Phi \\ \Phi_6 &= -\varphi_{h_2} + \varphi_{h_1+h_2} + \varphi_k - \varphi_{h_1+k} \\ \Phi_7 &= -\varphi_{h_1} - \varphi_{k_1} + \varphi_{h_1+k}, \end{aligned} \quad (24)$$

each giving information on Φ . It may be noted that only two of the three cross-reflexions of Φ_2, Φ_4, Φ_6 are

contained in (21). Since a good estimation of the quartets can improve the estimation of Φ , more promising than (21) seems to be the study of a distribution which is able to exploit the knowledge of all three cross-magnitudes for at least one quartet. That is the case of the distribution

$$P(E_{h_1}, E_{h_2}, E_{h_1+h_2}, E_k, E_{h_1+k}, E_{h_1+h_2+k}, E_{h_2+k}) \quad (25)$$

which exploits, besides (22), (23), (24), the following (not independent) triplets

$$\begin{aligned} \Phi_1 &= \Phi \\ \Phi_8 &= -\varphi_{h_2} + \varphi_{h_1+h_2} + \varphi_{h_2+k} - \varphi_{h_1+h_2+k} \\ \Phi_9 &= -\varphi_{h_1} - \varphi_{h_2+k} + \varphi_{h_1+h_2+k}, \\ \Phi_{10} &= \Phi \\ \Phi_{10} &= -\varphi_{h_1} + \varphi_{h_1+h_2} + \varphi_k - \varphi_{h_2+k} \\ \Phi_{11} &= -\varphi_{h_2} - \varphi_k + \varphi_{h_2+k}. \end{aligned}$$

Two of the cross-magnitudes of $\Phi_4, \Phi_6, \Phi_8, \Phi_{10}$ and all three cross-magnitudes of Φ_2 are contained in (25). This procedure may be developed as far as to exploit all the triplets associated with (20).

14. The order of a multipole

The order of the multipole $\Phi_1, \Phi_2, \dots, \Phi_n$ with respect to Φ_1 is the product of the orders of the phase relationships associated with $\Phi_2, \Phi_3, \dots, \Phi_n$. For example, (22) is a multipole of order $1/N|N$ for Φ_1 .

The lower the order of the multipole for Φ_1 , the more effective (in a probabilistic sense) the magnitudes involved in the multipole will be in defining Φ_1 . The multipoles of the lowest order for Φ_1 (*i.e.* of the same order as the phase relationship associated with Φ_1) are those carried out by means of the cross-reflexions of Φ_1 . For example, the triplet

$$\begin{aligned} \Phi_1 &= \varphi_{h_1} + \varphi_{h_2} + \varphi_{h_3} - \varphi_{h_1+h_2+h_3} \\ \Phi_2 &= -\varphi_{h_2} - \varphi_{h_3} + \varphi_{h_2+h_3} \\ \Phi_3 &= -\varphi_{h_1} - \varphi_{h_2+h_3} + \varphi_{h_1+h_2+h_3} \end{aligned}$$

is of order $1/N$ for Φ_1 , just as the phase relationship associated with Φ_1 ($E_{h_2+h_3}$ is a cross-reflexion of Φ_1). This property may be useful in order to define the cross magnitudes of the u.s.i. Φ given by (5). We note in fact, that the triplet

$$\begin{aligned} \Phi_1 &= A_1\varphi_{h_1} + A_2\varphi_{h_2} + \dots + A_n\varphi_{h_n} \\ \Phi_2 &= -m_1\varphi_{h_1} - m_2\varphi_{h_2} - \dots - m_n\varphi_{h_n} + \varphi_{m_1h_1 + \dots + m_nh_n} \\ \Phi_3 &= -(A_1 - m)\varphi_{h_1} - \dots - (A_n - m_n)\varphi_{h_n} \\ &\quad - \varphi_{m_1h_1 + \dots + m_nh_n} \end{aligned}$$

with $m_i = 0, \dots, A_i$, is for Φ just of the same order as the phase relationship associated with Φ_1 . Thus any

$$E_{m_1h_1 + \dots + m_nh_n}$$

is a cross-reflexion of Φ (provided it does not coincide with 0 or with some basis vector), when the integers m_i vary between 0 and A_i .

Thus, $R_{h_1+h_2+h_3}$ is a cross-magnitude for the sextet invariant because it gives rise to the tripole

$$\Phi_1 = \varphi_{h_1} + \varphi_{h_2} + \varphi_{h_3} + \varphi_{h_4} + \varphi_{h_5} - \varphi_{h_1+h_2+h_3+h_4+h_5}$$

$$\Phi_2 = \varphi_{h_1+h_2+h_3} - \varphi_{h_1} - \varphi_{h_2} - \varphi_{h_3}$$

$$\Phi_3 = -\varphi_{h_1+h_2+h_3} - \varphi_{h_4} - \varphi_{h_5} + \varphi_{h_1+h_2+h_3+h_4+h_5}$$

which is of order $1/N^2$ for Φ_1 .

15. A comparison with the nested-neighbourhoods theory

The main purpose of the nested-neighbourhoods theory, as well as of our method of the upper representations, is that of arranging the sample space $\{W\}$ in a sequence of sets, each set being a subset of the succeeding one, with the property that Φ may be estimated in terms of the magnitudes constituting any set. The two methods have different backgrounds: in general they suggest different approaches to the solution of the phase problem. It therefore comes as no surprise that the two methods do not arrange $\{W\}$ in identical sequences of sets.

However, in spite of the different backgrounds, the two methods must have several points of connexion. Illuminating in this sense may be the comparison of the sequence in $P1$ of the phasing shells of the quartet invariant $\Phi = \varphi_{h_1} + \varphi_{h_2} + \varphi_{h_3} + \varphi_{h_4}$, described in Fig. 4 with the nested-neighbourhoods sequence given by Hauptman (1977a) and shown in our notation in Fig. 5. Besides the trivial observation that the magnitudes in the two first neighbourhoods coincide with those in the first phasing shell, we point out that any neighbourhood of higher order is an unsymmetrical (regarding h_1, h_2, h_3, h_4) subset of a corresponding phasing shell. Suitably, Hauptman noticed that Fig. 2 represents one of the possible sequences of nested neighbourhoods of Φ . Several other examples of this type may be derived if one compares the sequences of neighbourhoods given by Hauptman for some u.s.i. or s.s. with the corresponding sequences of phasing shells given in this paper. For instance, 17 and 48 phasing magnitudes are in the first phasing shells of the three-phase s.s. $\Phi = \varphi_h + \varphi_k + \varphi_l$ and of the four-phase s.s. $\Phi = \varphi_h + \varphi_k + \varphi_l + \varphi_m$ in $P\bar{1}$ respectively (see Figs. 11, 12). The first two neighbourhoods given by Hauptman (1976) for the same s.s. contain 13 and 22 magnitudes respectively.

16. A branch: the method of the complementary invariants

Until now we have shown that a u.s.i. or s.s. Φ may be estimated *via* one or more u.s.i. $\Psi_1 \in \{\Psi\}_1$, one or more $\Psi_2 \in \{\Psi\}_2, \dots$ etc., where

$$\Psi_1 = \Phi + \varphi_p - \varphi_p$$

$$\Psi_2 = \Psi_1 + \varphi_r - \varphi_r, \text{ etc.}$$

An obvious variant may be that of associating with Φ one or more u.s.i. or s.s. $\Phi_f, \Phi_g, \dots, \Phi_q$ such that

$$\Phi' = \Phi + \Phi_f + \Phi_g + \dots + \Phi_q \quad (26)$$

is a u.s.i.. Φ' is not an element of an upper representation of Φ because $\Phi' - \Phi$ is not a constant which arises because of translational symmetry. We state that Φ' is a complementary u.s.i. of Φ with respect to Φ_f, Φ_g, Φ_q . If $\Phi', \Phi_f, \Phi_g, \dots, \Phi_q$ are estimated *via* their representations, Φ is in consequence evaluated. From a mathematical point of view the estimation of Φ requires that from the probability distribution

$$P(\Phi, \Phi_f, \Phi_g, \dots, \Phi_q, \Phi' | \{R\})$$

one arrives at

$$P(\Phi | \Phi_f, \Phi_g, \dots, \Phi_q, \Phi', \{R\}).$$

The method seems rather promising since one or more reciprocal vectors may vary throughout reciprocal space, so giving rise, for a fixed Φ , to numerous collections $\Phi_f, \dots, \Phi_q, \Phi'$.

The first applications of this method are encouraging. Giacobazzo (1975b) associated in $P\bar{1}$ with the s.s. $\Phi = \varphi_{2h}$ the u.s.i. $\Phi_f = \varphi_h + \varphi_k - \varphi_{h+k}$ so that

$$\Phi' = \varphi_{2h} - \varphi_h + \varphi_k - \varphi_{h+k}.$$

As $\varphi_h = -\varphi_h$ in $P\bar{1}$, the estimate of any Φ' (k is a free vector) *via* the six magnitudes $R_{2h}, R_h, R_k, R_{h+k}, R_{2h+k}$ and of Φ_f *via* R_h, R_k, R_{h+k} allows one to evaluate Φ . More recently, Sheldrick (1976) extended the idea to any space group. Giacobazzo (1977a) and Burla, Polidori, Nunzi, Cascarano & Giacobazzo (1977) estimate one and two-phase seminvariants in $P\bar{1}$ by means of quintet complementary invariants.

It should be useful to emphasize that the methods of representations and complementary invariants are strictly related to one another: in particular, they work on the same sets of reflexions. For example, if one wishes to estimate φ_{2h} in $P\bar{1}$ by means of quartet complementary invariants, the following quartets should be evaluated:

$$\Phi'_1 = \varphi_{2h} - \varphi_h + \varphi_k - \varphi_{h+k},$$

$$\Phi'_2 = \varphi_{2h} - \varphi_h - \varphi_k - \varphi_{h-k}.$$

This may be performed just by means of the second phasing shell of φ_{2h} (see Fig. 7).

If φ_{2h} is estimated in $P\bar{1}$ by quintet complementary invariants, the following quintets should be evaluated

$$\Phi'_1 = \varphi_{2h} - \varphi_h + \varphi_k + \varphi_l - \varphi_{h+k+l},$$

$$\Phi'_2 = \varphi_{2h} - \varphi_h - \varphi_k - \varphi_l - \varphi_{h-k-l},$$

$$\Phi'_3 = \varphi_{2h} - \varphi_h + \varphi_k - \varphi_l - \varphi_{h+k-l},$$

$$\Phi'_4 = \varphi_{2h} - \varphi_h - \varphi_k + \varphi_l - \varphi_{h-k+l}.$$

This may be performed just by means of the third phasing shell of Φ_{2h} (see Fig. 7).

As a further example, the two-phase seminvariant $\varphi_{h+k} + \varphi_{h-k}$ may be estimated in $P\bar{I}$ by means of the following quintet complementary invariants

$$\begin{aligned} \Phi'_1 &= \varphi_{h+k} + \varphi_{h-k} - \varphi_h - \varphi_p - \varphi_{h-p}, \\ \Phi'_2 &= \varphi_{h+k} + \varphi_{h-k} - \varphi_h + \varphi_p - \varphi_{h+p}, \\ \Phi'_3 &= \varphi_{h+k} - \varphi_{h-k} - \varphi_k + \varphi_p - \varphi_{k+p}, \\ \Phi'_4 &= \varphi_{h+k} - \varphi_{h-k} - \varphi_k - \varphi_p - \varphi_{k-p}, \end{aligned}$$

which require that the magnitudes of the second phasing shell are known (see Fig. 9).

However, even if the methods of representations and complementary invariants work on the same sets of reflexions, they deal with them in different ways. In the complementary-invariants method each Φ' is considered independent of the others. This condition is not assumed in the method of representations. Furthermore, we can note that the method of complementary invariants exploits in $P\bar{I}$ the second phasing shell of φ_{2h} in order to use phase relationships of order $1/N$ (quartets), the third phasing shell in order to use phase relationships of order $1/N\sqrt{N}$ (quintets), etc. On the other hand the method of representations exploits the second phasing shell of φ_{2h} in order to calculate phase relationships of order $1/N\sqrt{N}$, the third phasing shell in order to calculate phase relationships of order $1/N^2\sqrt{N}$, etc. Similar considerations hold for other s.s.'s or u.s.i.'s.

17. Conclusions

The theory of representations of a u.s.i. or a s.s. Φ , allows the arrangement of the sample space $\{W\}$ in subsets (i.e. phasing shells), each a subset of the succeeding one, such that Φ can be estimated in terms of the magnitudes belonging to the phasing shell. The 'nested-neighbourhoods principle' formulated by Hauptman is also able to fix sequences of sets of magnitudes. We have shown that it is a consequence of the more general phase-interrelationships principle stated here which, joined with the theory of representations, is able to give general sequences of magnitudes.

This work was supported by the Consiglio Nazionale delle Ricerche (grant No. 75.1066.05.115.4593).

APPENDIX

Here are summarized the results of our analysis on the rank of the s.s.'s in all the primitive space groups. Since there is the set of rotation matrices which fixes the rank of a s.s., the results for non-primitive space groups are easily derivable. We have preferred to describe our results in terms of Hauptman-Karle groups (see Giacobozzo, 1975*b* for the notation) because of their large use in the direct procedures for phase solution. The following statements hold. (1) All the s.s.'s in the

centrosymmetric space groups are of first rank. (2) All the s.s.'s in the following Hauptman-Karle groups:

$$(h, k, l)P(2, 0, 2); (h, k, l)P(0, 2, 0); (h, k, l)P(2, 2, 0); (h+k, l)P(2, 0); (h-k, l)P(3, 0); (h+k+l)P(0); \quad (A.1)$$

are of first rank.

For each Hauptman-Karle group which is not in (A.1) a list of symmetry classes in which s.s.'s of second rank occur is given below. For each symmetry class in the list, the algebraic conditions characterizing the s.s.'s of first rank are given. Of course the conditions hold for all the space groups in the H-K group belonging to that symmetry class.

$$(h, k, l)P(2, 2, 2) \\ (h, k, l) \equiv 0 \pmod{(2, 2, 0) \text{ or } (2, 0, 2) \text{ or } (0, 2, 2)}$$

$$(h+k, l)P(2, 2) - \text{symmetry class 422} \\ (h, k, l) \equiv 0 \pmod{(2, 2, 0) \text{ or } (2, 0, 2) \text{ or } (0, 2, 2)} \\ (h \pm k, l) \equiv 0 \pmod{(0, 2) \text{ or } (2, 0)}$$

$$(h-k, l)P(3, 2) - \text{symmetry class 32} \\ (h-k, l) \equiv 0 \pmod{(3, 0)} \\ (h-k, l) \text{ or } (2h+k, l) \text{ or } (h+2k, l) \equiv 0 \pmod{(0, 2)}$$

$$IP(0) - \text{symmetry class } 3m \\ (h-k, l) \equiv 0 \pmod{(3, 0)} \\ (h+k, l) \equiv 0 \pmod{(0, 0)} \\ (h, k, l) \equiv 0 \pmod{(1, 0, 0) \text{ or } (0, 1, 0)}$$

$$IP(2) - \text{symmetry class 32} \\ (h-k, l) \equiv 0 \pmod{(3, 0)} \\ (h+k, l) \equiv 0 \pmod{(0, 2)} \\ (h, k, l) \equiv 0 \pmod{(1, 0, 2) \text{ or } (0, 1, 2)}$$

$$- \text{symmetry class } 622 \\ (h \pm k, l) \equiv 0 \pmod{(0, 2)} \\ (h, k, l) \equiv 0 \pmod{(1, 0, 2) \text{ or } (0, 1, 2)} \\ (2h+k, l) \text{ or } (h+2k, l) \equiv 0 \pmod{(0, 2)} \\ (h, k, l) \equiv 0 \pmod{(1, 1, 0)}$$

$$- \text{symmetry class } \bar{6}m2 \\ (h-k, l) \equiv 0 \pmod{(3, 2)} \\ (h+k, l) \equiv 0 \pmod{(0, 2)} \\ (h, k, l) \equiv 0 \pmod{(1, 0, 2) \text{ or } (0, 1, 2)}$$

$$(h+k+l)P(2) - \text{symmetry class 32 (rhombohedral lattice)} \\ (h-k, l) \text{ or } (h-l, k) \text{ or } (k-l, h) \equiv 0 \pmod{(0, 2)} \\ (h+k+l) \equiv 0 \pmod{(0)}$$

$$- \text{symmetry class 23} \\ (h, k, l) \equiv 0 \pmod{(0, 2, 2) \text{ or } (2, 0, 2) \text{ or } (2, 2, 0)} \\ (h \pm k \pm l) \equiv 0 \pmod{(0)}$$

$$- \text{symmetry class 432} \\ (h \pm k, l) \text{ or } (h \pm l, k) \text{ or } (k \pm l, h) \equiv 0 \pmod{(2, 0) \text{ or } (0, 2)} \\ (h \pm k \pm l) \equiv 0 \pmod{(0)}$$

– symmetry class $\bar{4}3m$

$$(h+k, l) \text{ or } (h+l, k) \text{ or } (k+l, h) \equiv 0 \pmod{(2, 2)}$$

$$(h, k, l) \equiv 0 \pmod{(0, 2, 2) \text{ or } (2, 0, 2) \text{ or } (2, 2, 0)}$$

$$(h \pm k \pm l) \equiv 0 \pmod{(0)}$$

References

- BURLA, M. C., POLIDORI, G., NUNZI, A., CASCARANO, G. & GIACOVAZZO, C. (1977). *Acta Cryst.* A **33**, 949–953.
 GIACOVAZZO, C. (1974). *Acta Cryst.* A **30**, 390–395.
 GIACOVAZZO, C. (1975a). *Acta Cryst.* A **31**, 252–259.
 GIACOVAZZO, C. (1975b). *Acta Cryst.* A **31**, 602–609.
 GIACOVAZZO, C. (1976a). *Acta Cryst.* A **32**, 967–976.
 GIACOVAZZO, C. (1976b). *Acta Cryst.* A **32**, 958–966.

- GIACOVAZZO, C. (1977a). *Acta Cryst.* A **33**, 527–531.
 GIACOVAZZO, C. (1977b). *Acta Cryst.* A **33**, 531–538.
 GIACOVAZZO, C. (1977c). *Acta Cryst.* A **33**, 539–547.
 HAUPTMAN, H. (1975). *Acta Cryst.* A **31**, 671–679.
 HAUPTMAN, H. (1976). *Acta Cryst.* A **32**, 934–940.
 HAUPTMAN, H. & FORTIER, S. (1977). *Acta Cryst.* A **33**, 575–580.
 HAUPTMAN, H. & KARLE, J. (1953). *Solution of the Phase Problem. I. The Centrosymmetric Crystal*. ACA Monograph No. 3. Pittsburg: Polycrystal Book Service.
 HAUPTMAN, H. & KARLE, J. (1956). *Acta Cryst.* **9**, 45–55.
 HAUPTMAN, H. & KARLE, J. (1959). *Acta Cryst.* **12**, 93–97.
 SCHENK, H. (1973). *Acta Cryst.* A **29**, 77–82.
 SHELDRICK, G. M. (1976). Buffalo Intercongress Symposium, August 3. Abstr. SE2.

Acta Cryst. (1977). A **33**, 944–948

Quintets in $P\bar{1}$ and Related Phase Relationships: a Probabilistic Approach

BY C. GIACOVAZZO

Istituto di Mineralogia, Università degli Studi, Piazza Umberto I 1, 70121 Bari, Italy

(Received 18 April 1977; accepted 11 June 1977)

A probabilistic theory is described which is able to estimate in $P\bar{1}$ the signs of the quintet invariants. An investigation is carried out on the use of special quintets in order to estimate one and two-phase seminvariants by means of the complementary-invariants method.

1.1. Introduction

Let $\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3, \mathbf{h}_4, \mathbf{h}_5$ be reciprocal vectors for which

$$\mathbf{h}_1 + \mathbf{h}_2 + \mathbf{h}_3 + \mathbf{h}_4 + \mathbf{h}_5 = 0.$$

Then the linear combination of phases

$$\varphi = \varphi_{\mathbf{h}_1} + \varphi_{\mathbf{h}_2} + \varphi_{\mathbf{h}_3} + \varphi_{\mathbf{h}_4} + \varphi_{\mathbf{h}_5} \quad (1)$$

is a structure invariant. The theory of representations (Giacovazzo, 1977) states that φ may be evaluated in $P1$ or $P\bar{1}$ via its first phasing shell by means of the 15 magnitudes

$$E_{m_1\mathbf{h}_1 + \dots + m_5\mathbf{h}_5} \quad (m_p = 0, 1). \quad (2)$$

Schenk (1975) spoke of quintets at the Tenth International Congress of Crystallography. The main result presented was a linear trend of φ versus the sum of the cross-magnitudes. At the Buffalo Symposium on Direct Methods more detailed analysis was presented by Schenk (1976) by a semi-empirical method and by Fortier & Hauptman (1976) with the theory of the joint probability distribution functions. More recently, Fortier & Hauptman (1977) described a probabilistic approach in $P\bar{1}$ which is able to predict the sign of a quintet by means of a formula which involves a summation over 1024 contributions. This paper describes a probabilistic approach to quintets in $P\bar{1}$ which leads to formulae more tractable than Fortier & Haupt-

man's. Special quintets are also studied which may allow good estimates of one and two-phase structure seminvariants.

1.2. The mathematical approach

The method to be described requires the derivation of a variety of conditional probability distributions. If we denote by $P(E_1, E_2, \dots, E_n)$ the joint probability function of n normalized structure factors, its characteristic function may be expanded in a Gram–Charlier series:

$$C(u_1, \dots, u_n) = \exp \left[-\frac{1}{2}(u_1^2 + \dots + u_n^2) \right] \\ \times \left[1 + S_3/t^{3/2} + (S_4/t^2 + S_3^2/2t^3) \right. \\ \left. + (S_5/t^{5/2} + S_3S_4/t^{7/2} + S_3^3/6t^{9/2}) + \dots \right], \quad (3)$$

where u_i , $i = 1, \dots, n$ are carrying variables associated with E_i , t is the number of independent atoms in the unit cell,

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

and

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

$K_{rs\dots w}$ are the cumulants of the distribution and m is the order of the space group. $P(E_1, E_2, \dots, E_n)$ is the Fourier transform of (3).