References

BUSETTA, B. (1973). Thèse de Doctorat ès Sciences. Université de Bordeaux I.

BUSETTA, B. (1976). *Acta Cryst.* A**32**, 359–362.

COLENS, A., DECLERCQ, J. P., GERMAIN, G., PUTZEYS, J. P. & VAN MEERSCHE, M. (1974). *Cryst. Struct. Commun.* **3**, 119–122.

GIACOVAZZO, C. (1976*a*). *Acta Cryst.* A**32**, 967–976.

GIACOVAZZO, C. (1976*b*). *Acta Cryst.* A**32**, 958–967.

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

GIACOVAZZO, C. (1979). In preparation.

HAUPTMAN, H. (1972). *Crystal Structure Determination, The Role of the Cosine Seminvariants.* New York: Plenum.

HAUPTMAN, H. (1974). *Acta Cryst.* A**30**, 822–829.

KARLE, I. L. (1975). *J. Am. Chem. Soc.* **97**, 4379–4382.

MAIN, P., WOOLFSON, M. M. & GERMAIN, G. (1971). *MULTAN. A Computer Program for the Automatic Solution of Crystal Structures.* Univ. of York, England.

SCHENK, H. (1974). *Acta Cryst.* A**30**, 477–481.

Acta Cryst. (1980). A**36**, 74–82

Triplet and Quartet Relations: Their Use in Direct Procedures

BY CARMELO GIACOVAZZO

Istituto di Mineralogia e Petrografia, Palazzo Ateneo, Università di Bari, Italy

(*Received* 21 *June* 1979; *accepted* 30 *July* 1979)

Abstract

A phase relationship involving triplet and quartet contributions is given. It is able to take account of the correlation between triplet and quartet relations. The information exploited by the formula is discussed and compared with that exploited by phase relationships arising from the properties of the Fourier transform of periodic positive functions. In particular, the information contained in a Karle–Hauptman determinant of low order is briefly considered.

1. Introduction

The properties and the use of quartet relations in direct procedures can be approached from two basic points of view: (*a*) the properties of the Fourier transform of periodic positive functions (*i.e.* the electron density function); (*b*) the application of joint probability distribution methods. We show in this paper that phase relations based on (*a*) hold only if some restrictive conditions are satisfied. On the other hand, phase relations based on (*b*) can be extensively applied in the usual procedures for phase solution. In practice, the probabilistic approach can provide phase relations more useful than those provided by (*a*).

Triplet and quartet relations can be used together in phase-determination processes. Since a quartet is the sum of two triplets, a strong correlation may exist between the sets of estimated triplets and quartets. Unfortunately, triplets and quartets have been used so far as independent relationships. In practice this assumption can lead to undesirable effects, *e.g.* to the 'squaring' effect of Σ_2 relationships the 'cubing' effect of quartet relations is added. In this paper a formula is given, involving triplet and quartet relations, which is able to take account of the correlation between the two kinds of phase relationships.

The value of Karle–Hauptman determinants of low order is mostly determined by triplet and quartet contributions. We discuss the information contained in such determinants and briefly compare it with the information exploited in our probabilistic approach.

2. Phase relations arising from the Fourier transform of $\rho^n(\mathbf{r})$

For a structure containing atoms which are fully resolved from one another the operation of raising $\rho(\mathbf{r})$ to the nth power retains the conditions of resolved atoms but changes the shape of each atom. In practice it is possible to substitute, with an accuracy quite sufficient for the purpose of structure analysis, $\rho(\mathbf{r})$ by a sum of N spherically-symmetrical atomic functions:

$$\rho(\mathbf{r}) = \sum_{j=1}^{N} \rho_j(\mathbf{r} - \mathbf{r}_j).$$

$\rho_j(\mathbf{r})$ is an atomic function and $\mathbf{r}_j$ is the coordinate of the center of the atom. When a function is periodic, its nth power is periodic with the same unit cell but with a

0567-7394/80/010074-09$01.00

different set of Fourier coefficients. Since the atoms do not overlap,

$$\rho^n(\mathbf{r}) = \left[\sum_{j=1}^{N} \rho_j(\mathbf{r}-\mathbf{r}_j)\right]^n \simeq \sum_{j=1}^{N} \rho_j^n(\mathbf{r}-\mathbf{r}_j)$$

and its Fourier transform gives

$$\begin{aligned} {}_nF_{\mathbf{h}} &= \int_V \rho^n(\mathbf{r}) \exp(2\pi i\mathbf{hr})\,\mathrm{d}V \\ &= \sum_{j=1}^{N} \int_V \rho_j^n(\mathbf{r}-\mathbf{r}_j) \exp(2\pi i\mathbf{hr})\,\mathrm{d}V \\ &= \sum_{j=1}^{N} {}_nf_j \exp 2\pi i\mathbf{hr}_j. \end{aligned}$$

${}_nf_j$ is the scattering factor for the jth peak of $\rho^n(\mathbf{r})$,

$${}_nf_j = \int_V \rho_j^n(\mathbf{r}) \exp(2\pi i\mathbf{hr})\,\mathrm{d}\mathbf{r}.$$

If all atoms are equal so that $f_j \equiv f$ and ${}_nf_j \equiv {}_nf$ for any j, then

$$F_{\mathbf{h}} = (f/{}_nf)\,{}_nF_{\mathbf{h}} = \theta_n\,{}_nF_{\mathbf{h}}, \tag{1}$$

where θ_n is a function which corrects for the difference in shape of atoms with electron distributions $\rho(\mathbf{r})$ and $\rho^n(\mathbf{r})$. Since

$$\rho^n(\mathbf{r}) = \frac{1}{V^n} \sum_{\mathbf{h}_1,\ldots,\mathbf{h}_n = -\infty}^{+\infty} F_{\mathbf{h}_1} \ldots F_{\mathbf{h}_n} \exp[-2\pi i(\mathbf{h}_1 + \cdots + \mathbf{h}_n)\mathbf{r}],$$

the Fourier transform of both sides gives

$${}_nF_{\mathbf{h}} = \frac{1}{V^{n-1}} \sum_{\mathbf{h}_1,\ldots,\mathbf{h}_{n-1} = -\infty}^{+\infty} F_{\mathbf{h}_1} F_{\mathbf{h}_2} \cdots F_{\mathbf{h}-\mathbf{h}_1-\mathbf{h}_2-\ldots-\mathbf{h}_{n-1}},$$

from which the following relation arises:

$$F_{\mathbf{h}} = \theta_n \frac{1}{V^{n-1}} \sum_{\mathbf{h}_1,\ldots,\mathbf{h}_{n-1} = -\infty}^{+\infty} F_{\mathbf{h}_1} F_{\mathbf{h}_2} \cdots F_{\mathbf{h}-\mathbf{h}_1-\mathbf{h}_2-\cdots-\mathbf{h}_{n-1}}. \tag{2}$$

For $n = 2$, (2) reduces to Sayre's (1952) relation

$$F_{\mathbf{h}} = \theta_2 \frac{1}{V} \sum_{\mathbf{k}} F_{\mathbf{k}} F_{\mathbf{h}-\mathbf{k}}. \tag{3}$$

As assumed above, relations such as (2) are valid only when atoms are all of the same type and are fully resolved; their application requires that $\rho^n(\mathbf{r})$ is known. In reciprocal space this condition requires that all phases (of measured and unmeasured reflections) are known. This restrictive condition cannot be satisfied in practice. However, if the set of known phases is sufficiently large, then it may be considered as a unbiased sample of the entire population of phases. In this case (2) approximately holds.

3. Phase relations arising from the Fourier transform of linear combinations of powers of $\rho(\mathbf{r})$

If the structure contains resolved isotropic atoms of two types, P and Q, it is impossible to find a factor θ_2 such that the relation $F_{\mathbf{h}} = \theta_2\,{}_2F_{\mathbf{h}}$ holds, since this would imply values of θ_2 such that $({}_2f)_P = \theta_2(f)_P$ and $({}_2f)_Q = \theta_2(f)_Q$ simultaneously. Then (3) does not hold exactly. The Sayre relation, however, can be suitably modified for this case (Fan, 1965; Krabbendam & Kroon, 1971). A different approach was introduced by Woolfson (1958) who intrduced quartet invariants to correct the Sayre relation,

$$F_{\mathbf{h}} \simeq \frac{A_s}{V} \sum_{\mathbf{k}} F_{\mathbf{k}} F_{\mathbf{h}-\mathbf{k}} - \frac{B_s}{V^2} \sum_{\mathbf{k},\mathbf{l}} F_{\mathbf{k}} F_{\mathbf{l}} F_{\mathbf{h}-\mathbf{k}-\mathbf{l}}, \tag{4}$$

where A_s and B_s are adjustable positive parameters of $s = \sin\theta/\lambda$. Equation (4) can be generalized to structures containing resolved atoms of more than two types:

$$F_{\mathbf{h}} \simeq \frac{A_s}{V} \sum_{\mathbf{k}} F_{\mathbf{k}} F_{\mathbf{h}-\mathbf{k}} + \frac{B_s}{V^2} \sum_{\mathbf{k},\mathbf{l}} F_{\mathbf{k}} F_{\mathbf{l}} F_{\mathbf{h}-\mathbf{k}-\mathbf{l}} + \frac{C_s}{V^3} \sum_{\mathbf{k},\mathbf{l},\mathbf{p}} F_{\mathbf{k}} F_{\mathbf{l}} F_{\mathbf{p}} F_{\mathbf{h}-\mathbf{k}-\mathbf{l}-\mathbf{p}} + \cdots$$

The procedure so far described can be adapted to structures consisting of point atoms. For example, von Eller (1973) showed, for a point-atom structure with atoms of two types, that

$$E_{\mathbf{h}} \simeq \frac{Z_1 + Z_2}{Z_1 Z_2} \sigma_2^{1/2} \langle E_{\mathbf{k}} E_{\mathbf{h}-\mathbf{k}} \rangle_{\mathbf{k}} - \frac{\sigma_2}{Z_1 Z_2} \langle E_{\mathbf{k}} E_{\mathbf{l}} E_{\mathbf{h}-\mathbf{k}-\mathbf{l}} \rangle_{\mathbf{k},\mathbf{l}}, \tag{5}$$

where Z_1 and Z_2 are the atomic numbers of the first and second atomic species and $\sigma_2 = \sum_{j=1}^{N} Z_j^2$.

It may be concluded that (4) and (5) are valid only when two types of atoms are in the structure and the atomic electron densities are fully resolved. Their application requires that $\rho^2(\mathbf{r})$ and $\rho^3(\mathbf{r})$ are known. In reciprocal space this condition requires again that all phases are known. The Fourier transform of (4) and (5) gives

$$\rho \simeq a_1 \rho^2 - a_2 \rho^3, \tag{6}$$

where a_1 and a_2 are positive parameters which are chosen so that $\rho - (a_1 \rho^2 - a_2 \rho^3)$ is a minimum. If all the atoms are of the same type, $a_2 = 0$ and quartets do not appear in (4) and (5). A condition related to but not identical with (6) was proposed by Allegra & Colombo

(1974), according to which the difference between the real and the squared structure must approach zero. Therefore the integral

$$\int_V [\rho^2(\mathbf{r})/I_2 - \rho(\mathbf{r})/I_1]^2 \, dV, \qquad (7)$$

where $I_n = \int_V \rho^n(\mathbf{r}) \, dV$, must be as small as possible. From (7), we can obtain

$$\tan \varphi_{\mathbf{h}} \simeq \left\{ \sum_{\mathbf{k}} |E_{\mathbf{k}} E_{\mathbf{h}-\mathbf{k}}| \sin(\varphi_{\mathbf{k}} + \varphi_{\mathbf{h}-\mathbf{k}}) - A \sum_{\mathbf{k},\mathbf{l}} |E_{\mathbf{k}} E_{\mathbf{l}} E_{\mathbf{h}-\mathbf{k}-\mathbf{l}}| \sin(\varphi_{\mathbf{k}} + \varphi_{\mathbf{l}} + \varphi_{\mathbf{h}-\mathbf{k}-\mathbf{l}}) \right\} \times \left\{ \sum_{\mathbf{k}} |E_{\mathbf{k}} E_{\mathbf{h}-\mathbf{k}}| \cos(\varphi_{\mathbf{k}} + \varphi_{\mathbf{h}-\mathbf{k}}) - A \sum_{\mathbf{k},\mathbf{l}} |E_{\mathbf{k}} E_{\mathbf{l}} E_{\mathbf{h}-\mathbf{k}-\mathbf{l}}| \cos(\varphi_{\mathbf{k}} + \varphi_{\mathbf{l}} + \varphi_{\mathbf{h}-\mathbf{k}-\mathbf{l}}) \right\}, \qquad (8)$$

where $A = E_0/2 \sum_{\mathbf{k}} E_{\mathbf{k}}^2$.

As well as (4) and (5), (8) strictly holds when all the phases are known. However, unlike (4) and (5), quartets appear in (8) even if the atoms are of the same type.

4. Phase-refinement techniques in direct space by modification of the electron-density distribution

Several criteria can be found which modify the electron-density distribution: Simonov (1976) proposed modification of $\rho(\mathbf{r})$ by the following relationship

$$\rho^{\text{mod}}(\mathbf{r}) = \begin{cases} \rho(\mathbf{r}) \text{ if } \rho(\mathbf{r}) \geq q_1 \\ q_2 \text{ if } \rho(\mathbf{r}) < q_1 \end{cases}, \qquad (9)$$

where q_i are suitable parameters; Hoppe, Gassmann & Zechmeister (1970) proposed the substitution of $\rho(\mathbf{r})$ by

$$\rho^{\text{mod}}(\mathbf{r}) \simeq a\rho(\mathbf{r}) + b\rho^2(\mathbf{r}) + c\rho^3(\mathbf{r}), \qquad (10)$$

where a, b, c are chosen so that the weak peaks in the ρ map decrease further, peaks of average height remain at the previous level and strong peaks somewhat decrease. The equivalent of (9) in reciprocal space is

$$F_{\mathbf{h}}^{\text{mod}} \simeq aF_{\mathbf{h}} + b\langle F_{\mathbf{k}} F_{\mathbf{h}-\mathbf{k}} \rangle_{\mathbf{k}} + c\langle F_{\mathbf{k}} F_{\mathbf{l}} F_{\mathbf{h}-\mathbf{k}-\mathbf{l}} \rangle_{\mathbf{k},\mathbf{l}},$$

which is very similar to (4) and (5); Collins (1975) proposed the following modification:

$$\begin{cases} \rho^{\text{mod}}(\mathbf{r}) = 3\rho^2(\mathbf{r}) - 2\rho^3(\mathbf{r}) \text{ if } \rho(\mathbf{r}) > 0 \\ \rho^{\text{mod}}(\mathbf{r}) = 0 \text{ if } \rho(\mathbf{r}) \leq 0. \end{cases} \qquad (11)$$

Equation (11) should increase resolution and sharpness at low-density levels and leave the largest peaks unchanged. The similarity between (10), (11) and (6), (7) suggests that refinement techniques such as those described in this paragraph can be applied only if conditions nearly identical to those described in §§ 2 and 3 are satisfied.

5. A fundamental property of the probabilistic approaches

Procedures described in §§ 2, 3, 4 assign phases both by expressing $\rho(\mathbf{r})$ *via* a linear combination of suitable powers of $\rho(\mathbf{r})$ and by exploiting the condition $\rho(\mathbf{r}) \geq 0$. A proper choice of the parameters in the linear combinations of the powers of $\rho(\mathbf{r})$ can in principle allow the introduction of some additional chemical information about the structure. Although such relationships proved useful in several cases, they have two undesirable weak points:

(*a*) they are strictly applicable only when both the number and the quality of known phases are sufficiently high to provide, after a Fourier transform, a good approximation of $\rho(\mathbf{r})$. That may be a critical condition in some cases. For example, in protein crystallography both the atomicity and the non-negativity conditions of the electron density function can be violated if data do not extend to atomic resolution ($d > 2$ Å) or phases are not sufficiently accurate;

(*b*) they are unable to exploit information about diffraction magnitudes when corresponding phases are unknown. Representations theory (Giacovazzo, 1977*a*) showed in fact that, for any structure seminvariant, it is possible to arrange the set of diffraction magnitudes in a sequence of expected effectiveness (in the statistical sense) for the estimation of Φ. In particular, any structure seminvariant depends, in its first representation, on the set of magnitudes contained in the first phasing shell. This set contains both the basis and the cross magnitudes of the seminvariant.

A triplet invariant is a particular case in the family of invariants because its cross magnitudes coincide with its basis magnitudes. On the contrary, a quartet invariant depends in its first representation, besides the four basis magnitudes, on at least (Giacovazzo, 1976*b*) three cross magnitudes; a quintet invariant on at least ten magnitudes, *etc.* If upper representations are introduced into calculations then any invariant or seminvariant can depend on many magnitudes. These results suggest that a phase may be accurately evaluated provided a suitable set of phases and a sufficiently large number of magnitudes are known. As a simple example, let us suppose that we want to estimate $\varphi_{\mathbf{h}}$ when the values of $\varphi_{\mathbf{k}}$, $\varphi_{\mathbf{l}}$, $\varphi_{\mathbf{h}-\mathbf{k}-\mathbf{l}}$ and the corresponding $|E|$'s are known. Then we can write

$$\varphi_{\mathbf{h}} \simeq \varphi_{\mathbf{k}} + \varphi_{\mathbf{l}} + \varphi_{\mathbf{h}-\mathbf{k}-\mathbf{l}}$$

with a reliability proportional to $|E_{\mathbf{h}} E_{\mathbf{k}} E_{\mathbf{l}} E_{\mathbf{h}-\mathbf{k}-\mathbf{l}}|/N$. The evaluation appears inadequate (Hauptman, 1975;

Giacovazzo, 1975) if the cross magnitudes of the quartet $\varphi_{\mathbf{h}} - \varphi_{\mathbf{k}} - \varphi_{\mathbf{l}} - \varphi_{\mathbf{h-k-l}}$ are known too. The situation does not change if more quartets are used in order to estimate $\varphi_{\mathbf{h}}$.

The example shows one of the fundamental properties of the probabilistic approach; unlike relationships which arise from the properties of the electron density, probabilistic procedures are able to exploit information about magnitudes of structure factors even when the corresponding phases are unknown. This property has not been extensively used so far. We describe in the next paragraphs how information contained in the first representation of triplets and quartets can be exploited in order to assign and refine phases.

6. The role and the use of quartets in direct procedures

Different probabilistic procedures have been suggested which make use of quartets only. All are based on the atomicity and positivity properties of the electron density function. Among them:

(i) Giacovazzo (1976*a*) proposed the tangent formula

$$\tan \varphi_{\mathbf{h}} = \frac{\sum_j G_j \sin(\varphi_{\mathbf{k}_j} + \varphi_{\mathbf{l}_j} + \varphi_{\mathbf{h}-\mathbf{k}_j-\mathbf{l}_j})}{\sum_j G_j \cos(\varphi_{\mathbf{k}_j} + \varphi_{\mathbf{l}_j} + \varphi_{\mathbf{h}-\mathbf{k}_j-\mathbf{l}_j})}, \qquad (12)$$

where

$$G_j = 2C_j(1 + \varepsilon_{\mathbf{h}-\mathbf{k}_j} + \varepsilon_{\mathbf{h}-\mathbf{l}_j} + \varepsilon_{\mathbf{k}_j+\mathbf{l}_j})/(1 + Q_j), \qquad (13)$$

$$C_j = |E_{\mathbf{h}} E_{\mathbf{k}_j} E_{\mathbf{l}_j} E_{\mathbf{h}-\mathbf{k}_j-\mathbf{l}_j}|/N,$$

$$Q_j = [(\varepsilon_{\mathbf{h}} \varepsilon_{\mathbf{k}_j} + \varepsilon_{\mathbf{l}_j} \varepsilon_{\mathbf{h}-\mathbf{k}_j-\mathbf{l}_j})\varepsilon_{\mathbf{h}-\mathbf{k}_j} + (\varepsilon_{\mathbf{h}} \varepsilon_{\mathbf{l}_j} + \varepsilon_{\mathbf{k}_j} \varepsilon_{\mathbf{h}-\mathbf{k}_j-\mathbf{l}_j})\varepsilon_{\mathbf{h}-\mathbf{l}_j} + (\varepsilon_{\mathbf{h}} \varepsilon_{\mathbf{h}-\mathbf{k}_j-\mathbf{l}_j} + \varepsilon_{\mathbf{k}_j} \varepsilon_{\mathbf{l}_j})\varepsilon_{\mathbf{k}_j+\mathbf{l}_j}]/2N,$$

$$\varepsilon_i = (E_i^2 - 1).$$

If $Q_j \le 0$ then it was proposed to set $Q_j = 0$. If some cross reflection is not measured then the corresponding ε is replaced by zero. If the cross magnitudes are unknown, then $G_j = 2C_j$, and (12) reduces to the relation proposed by Simerska (1956), where no use is made of the cross magnitudes.

(ii) Gilmore (1977) suggested a least-squares procedure according to which the phase $\varphi_{\mathbf{h}}$ is determined by the minimum in the function

$$\sum_j t_j[\cos \varphi_j - b_j]/\sum_j t_j.$$

t_j is the inverse of the estimated variance for Φ_j, $\Phi_j = \varphi_{\mathbf{h}} - \varphi_{\mathbf{k}_j} - \varphi_{\mathbf{l}_j} - \varphi_{\mathbf{h}-\mathbf{k}_j-\mathbf{l}_j}$, and b_j is the estimated cosine for Φ_j.

(iii) Schenk & van der Putten (1978) proposed the modified tangent formula

$$\tan \varphi_{\mathbf{h}} \simeq \frac{\sum_j C_j \sin(\varphi_{\mathbf{k}_j} + \varphi_{\mathbf{l}_j} + \varphi_{\mathbf{h}-\mathbf{k}_j-\mathbf{l}_j} - S_j|p_j|)}{\sum_j C_j \cos(\varphi_{\mathbf{k}_j} + \varphi_{\mathbf{l}_j} + \varphi_{\mathbf{h}-\mathbf{k}_j-\mathbf{l}_j} - S_j|p_j|)},$$

where p_j is the estimated value of the quartet invariant and $S_j = \pm 1$ is chosen such that for each term the expression $-\varphi_{\mathbf{h}} + \varphi_{\mathbf{k}_j} + \varphi_{\mathbf{l}_j} + \varphi_{\mathbf{h}-\mathbf{k}_j-\mathbf{l}_j} - S|p_j|$ is closest to zero.

Procedures (ii) and (iii) are more suitable than (i) for enantiomorph-specific refinement processes. Although these procedures seem very attractive (they are able to exploit distributions with maxima other than at 2π; furthermore, the number of available quartet relationships for each **h** is larger than that of triplets), they suffer some severe limitations. In fact, quartets are phase relationships of order N, so the number of reliable quartets may be small in spite of the large total number of estimated quartets, and it decreases with the structural complexity more rapidly than the number of reliable triplets. Furthermore, N can assume high values even in structures with a small number of independent atoms but high point symmetry. This effect can be minimized if maximum use of the symmetry for estimating quartets is made (Giacovazzo, 1976*b*). In fact, even if quartets with more than three cross magnitudes in the first phasing shell are a small percentage of the total number of quartets, they can constitute a large percentage of the most reliable quartets (Busetta *et al.*, 1980).

These considerations suggest that quartets should not be used alone in phasing procedures, but always associated with triplets.

Giacovazzo (1976*a*) obtained by a probabilistic approach the following expression:

$$\tan \varphi_{\mathbf{h}} = Q/Q', \qquad (14)$$

where

$$Q = \sum_{\mathbf{k}} A_{\mathbf{k}} \sin(\varphi_{\mathbf{k}} + \varphi_{\mathbf{h-k}}) + \sum_{\mathbf{k,l}} G_{\mathbf{k,l}} \sin(\varphi_{\mathbf{k}} + \varphi_{\mathbf{l}} + \varphi_{\mathbf{h-k-l}}),$$

$$Q' = \sum_{\mathbf{k}} A_{\mathbf{k}} \cos(\varphi_{\mathbf{k}} + \varphi_{\mathbf{h-k}}) + \sum_{\mathbf{k,l}} G_{\mathbf{k,l}} \cos(\varphi_{\mathbf{k}} + \varphi_{\mathbf{l}} + \varphi_{\mathbf{h-k-l}}),$$

$$A_{\mathbf{k}} = 2|E_{\mathbf{h}} E_{\mathbf{k}} E_{\mathbf{h-k}}|/\sqrt{N}$$

and G is defined by (13). Equation (14) has some advantages over previous formulae. In fact, the reliability of $\varphi_{\mathbf{h}}$ in (14) may be estimated by

$$\alpha_{\mathbf{h}} = (Q^2 + Q'^2)^{1/2},$$

which suggests that $\varphi_{\mathbf{h}}$ may be determined by (14) with higher accuracy than by triplets or quartets separately. Furthermore, the number of links which are weak in a convergence procedure using only triplets may be reduced if (14) is used. However, (14) deals with triplets and quartets as if they were independent relationships, and this is not the case. In order to assess this dependence in different practical conditions we recall some basic ideas presented by Giacovazzo (1977*b*).

The quartet $\Phi = \varphi_{\mathbf{h}} + \varphi_{\mathbf{k}} + \varphi_{\mathbf{l}} + \varphi_{\mathbf{h+k+l}}$ exploits in its first representation three (not independent) tripoles:

$$\begin{aligned}
&\Phi_1 = \Phi,\\
&t_1 = -\varphi_{\mathbf{h}} - \varphi_{\mathbf{k}} + \varphi_{\mathbf{h+k}},\\
&t_2 = -\varphi_{\mathbf{l}} - \varphi_{\mathbf{h+k}} + \varphi_{\mathbf{h+k+l}};\\
&\Phi_1 = \Phi,\\
&t_3 = -\varphi_{\mathbf{h}} - \varphi_{\mathbf{l}} + \varphi_{\mathbf{h+l}},\\
&t_4 = -\varphi_{\mathbf{k}} - \varphi_{\mathbf{h+l}} + \varphi_{\mathbf{h+k+l}};\\
&\Phi_1 = \Phi,\\
&t_5 = -\varphi_{\mathbf{k}} - \varphi_{\mathbf{l}} + \varphi_{\mathbf{k+l}},\\
&t_6 = -\varphi_{\mathbf{h}} - \varphi_{\mathbf{k+l}} + \varphi_{\mathbf{h+k+l}}.
\end{aligned}$$

One obtains $-\Phi = t_1 + t_2 = t_3 + t_4 = t_5 + t_6$. Thus the expected value of a quartet estimates three sums of two triplets. Three typical cases can be described:

(*a*) *Strong positive quartets*. Let us suppose that all the reflections involved in the tripoles have magnitudes larger than E_t, where E_t is the minimum value of $|E|$ chosen for carrying out phase determination by means of triplets. If E_t is large enough and the probabilistic theories of triplets and quartets hold, then

$$\Phi \simeq t_1 \simeq t_2 \simeq t_6 \simeq 0.$$

In this case some authors claim that quartets contain the same information as triplets. Though this is not quite true because the expected value of a quartet estimates sums of two triplets, the correlation between triplet and quartet information is very high. Therefore a phase refinement which uses triplets and quartet relationships as independent could emphasize the inadequacies of the standard tangent formula instead of improving it. From the point of view of direct space, a 'cubing' effect should be added to the 'squaring' effect of the triplet relationships, so that the procedure should tend to strengthen the dominant features of the structure.

(*b*) *Enantiomorph-sensitive quartets*. Let us suppose all $|E|$'s are larger than E_t, except $|E_{\mathbf{h+l}}|$ and $|E_{\mathbf{k+l}}|$ which are smaller. In this case t_1 and t_2 are the only triplets which are estimated in the phasing procedure. In particular, if $\varphi_{\mathbf{h}} \simeq \varphi_{\mathbf{k}} \simeq \varphi_{\mathbf{l}} \simeq 0$, triplet theory gives $\varphi_{\mathbf{h+k+l}} \simeq 0$; on the other hand probabilistic quartet theories can give $\Phi \simeq \pm\pi/2$ so that $\varphi_{\mathbf{h+k+l}} = \pm\pi/2$. The dramatic change in the phase estimation occurs because quartet theory is able to exploit information not used by triplet theory (*i.e.* the knowledge that $|E_{\mathbf{h}_1+\mathbf{h}_3}|$ and $|E_{\mathbf{h}_2+\mathbf{h}_3}|$ are small).

(*c*) *Negative quartets*. If all the $|E|$'s are larger than E_t, except $|E_{\mathbf{h}_1+\mathbf{h}_2}|$, $|E_{\mathbf{h}_1+\mathbf{h}_3}|$, $|E_{\mathbf{h}_2+\mathbf{h}_3}|$, which are near zero, no triplet appearing in the tripoles is estimated in the direct procedure. However, quartet theory enables us to estimate $q \simeq \pi$. Information contained in categories (*a*), (*b*), (*c*) is differently correlated with information contained in triplets. This correlation decreases when we pass from category (*a*) to (*c*), so these categories can play different roles in direct procedures.

No convincing use has been made so far of triplets and quartets in category (*a*), whereas some specific use of categories (*b*) and (*c*) proved successful.

Category (*b*). If N is not too large, some quartets are reliably estimated to be near $\pm\pi/2$. So they can be used to define a starting set which is completely enantiomorph specific. Enantiomorph-sensitive quartets can also be used as a figure of merit in non-centrosymmetric structures for recognizing the correct solution in multisolution approaches (van der Putten, 1979; Gilmore, 1979). This figure of merit is defined as

$$\mathrm{ENQ} = \sum w_i(\Phi_{t,i} + S_i|\Phi_{e,i}|),$$

where $\Phi_{t,i}$ and $\Phi_{e,i}$ are the assigned and the estimated values of the ith quartet. S_i is a sign chosen such that each $(\Phi_{t,i} + S_i|\Phi_{e,i}|)$ is closest to zero, w_i is proportional to the inverse of the variance for Φ_i. The correct solution should be among the smallest values of ENQ.

Category (*c*). Although the number of reliable negative quartets is small compared with the number of triplets and positive quartets, they can be used successfully for a figure of merit for selecting the correct solution in the multisolution procedures (Schenk, 1974). The figure is expected to be more effective in symmorphic space groups because it exploits a kind of relation (negative invariants) which is not accessible to triplets (at least from their first representation). The most simple criterion is (Giacovazzo, 1976*a*)

$$\mathrm{NQ} = -\sum C(1 + \varepsilon_{\mathbf{h+k}} + \varepsilon_{\mathbf{h+l}} + \varepsilon_{\mathbf{k+l}}) \cos \Phi,$$

where the summation goes over the quartets with small values of the cross magnitudes. The solution with the most negative value of NQ is expected to be the correct solution. An analogous figure of merit has been proposed by De Titta, Edmonds, Langs & Hauptman (1975); the negative quartets are selected by sharp cutoff parameters for C and cross magnitudes.

At times negative quartets can be used in the first stages of a multisolution procedure (Silverton, 1978). Since the correct solution should have the most negative NQ value, it seems reasonable to use in the starting set some planes frequently occurring as basis reflections in the set of estimated negative quartets. From the above considerations it can be seen that, from the probabilistic point of view, triplet and quartet relationships can be strongly correlated but they never exploit the same information. Their simultaneous use may be of great help in practice provided a suitable strategy is followed.

In the next section we show how information available at any stage of the phasing process may be exploited *via* triplet and quartet relations.

7. A general formula using triplet and quartet relationships

The seven magnitudes in the first phasing shell of

$$\Phi = \varphi_{\mathbf{h}} + \varphi_{\mathbf{k}} + \varphi_{\mathbf{l}} - \varphi_{\mathbf{h+k+l}}$$

are denoted by

$$R_1 = |E_1| = |E_{\mathbf{h}}|; \ldots; R_4 = |E_4| = |E_{\mathbf{h+k+l}}|;$$

$$R_5 = |E_{\mathbf{h+k}}|; R_6 = |E_{\mathbf{h+l}}|; R_7 = |E_{\mathbf{k+l}}|;$$

and $\varepsilon_i = (R_i^2 - 1)$.

Let φ_5, φ_6, φ_7 be unknown and R_5, R_6, R_7 be known. Developments of the Hauptman (1975) quartet theory lead to the marginal probability density

$$P(\varphi_1 | R_1, R_2, \ldots, R_7, \varphi_2, \varphi_3, \varphi_4)$$

$$\simeq \frac{1}{L} \exp[-4C \cos(\varphi_1 + \varphi_2 + \varphi_3 - \varphi_4)]\, I_0(R_5 X_5) \times I_0(R_6 X_6)\, I_0(R_7 X_7), \qquad (15)$$

where L is a suitable normalizing parameter, and

$$C = R_1 R_2 R_3 R_4/N,$$

$$X_5 = \frac{2}{\sqrt{N}} [R_1^2 R_2^2 + R_3^2 R_4^2 + 2NC \cos(\varphi_1 + \varphi_2 + \varphi_3 - \varphi_4)]^{1/2},$$

$$X_6 = \frac{2}{\sqrt{N}} [R_1^2 R_3^2 + R_2^2 R_4^2 + 2NC \cos(\varphi_1 + \varphi_2 + \varphi_3 - \varphi_4)]^{1/2},$$

$$X_7 = \frac{2}{\sqrt{N}} [R_2^2 R_3^2 + R_1^2 R_4^2 + 2NC \cos(\varphi_1 + \varphi_2 + \varphi_3 - \varphi_4)]^{1/2}.$$

Under the same conditions, developments of the Giacovazzo (1976*a*) quartet theory lead to

$$P(\varphi_1 | R_1, \ldots, R_7, \varphi_2, \varphi_3, \varphi_4)$$

$$\simeq \frac{1}{2\pi I_0(G)} \exp[G \cos(\varphi_1 + \varphi_2 + \varphi_3 - \varphi_4)], \qquad (16)$$

where G is defined according to (13).

If one or more cross magnitudes are large, the conditions under which (15) and (16) have been derived usually occur only at the first stages of practical procedures. Often, besides magnitudes, the phases of some cross reflections are also known. The availability of this information introduces triplets in calculations.

Let us suppose, for example, that φ_7 is also known. Then (15) is replaced by

$$P(\varphi_1 | \varphi_7, \ldots)$$

$$\simeq \frac{1}{L} \exp\left[\frac{2}{\sqrt{N}} R_1 R_4 R_7 \cos(\varphi_1 - \varphi_4 + \varphi_7) - 4C \cos(\varphi_1 + \varphi_2 + \varphi_3 - \varphi_4)\right] \times I_0(R_5 X_5) I_0(R_6 X_6).$$

If φ_6 and φ_7 are known, then

$$P(\varphi_1 | \varphi_6, \varphi_7, \ldots)$$

$$\simeq \frac{1}{L} \exp\left\{\frac{2}{\sqrt{N}} [R_1 R_3 R_6 \cos(\varphi_1 + \varphi_3 - \varphi_6) + R_1 R_4 R_7 \cos(\varphi_1 - \varphi_4 + \varphi_7)] - \frac{2}{N} [R_1 R_2 R_6 R_7 \cos(\varphi_1 - \varphi_2 - \varphi_6 + \varphi_7) - 4C \cos(\varphi_1 + \varphi_2 + \varphi_3 - \varphi_4)]\right\} I_0(R_5 X_5).$$

If φ_5, φ_6, φ_7 are known, then

$$P(\varphi_1 | \varphi_5, \varphi_6, \varphi_7, \ldots)$$

$$\simeq \frac{1}{L} \exp\left\{\frac{2}{\sqrt{N}} [R_1 R_2 R_5 \cos(\varphi_1 + \varphi_2 - \varphi_5) + R_1 R_3 R_6 \cos(\varphi_1 + \varphi_3 - \varphi_6) + R_1 R_4 R_7 \times \cos(\varphi_1 - \varphi_4 + \varphi_7)] - \frac{2}{N} [\mathrm{R}_1 \mathrm{R}_2 \mathrm{R}_6 \mathrm{R}_7 \times \cos(\varphi_1 - \varphi_2 - \varphi_6 + \varphi_7) + R_1 R_3 R_5 R_7 \times \cos(\varphi_1 - \varphi_3 - \varphi_5 + \varphi_7) + R_1 R_4 R_5 R_6 \times \cos(\varphi_1 + \varphi_4 - \varphi_5 - \varphi_6) + 2R_1 R_2 R_3 R_4 \times \cos(\varphi_1 + \varphi_2 + \varphi_3 - \varphi_4)]\right\}$$

$$\simeq \frac{1}{2\pi \mathrm{I}_0(\alpha_1)} \exp[\alpha_1 \cos(\varphi_1 - \theta_1)], \qquad (17)$$

where $\tan\theta_1 = T/Q$,

$$Q = \frac{2}{\sqrt{N}}[R_1 R_2 R_5 \cos(\varphi_2 - \varphi_5) \times \ldots]$$

$$- \frac{2}{N}[R_1 R_2 R_6 R_7 \cos(\varphi_2 + \varphi_6 - \varphi_7) + \ldots],$$

$$T = \frac{2}{\sqrt{N}}[R_1 R_2 R_5 \sin(\varphi_2 - \varphi_5) + \ldots]$$

$$- \frac{2}{N}[R_1 R_2 R_6 R_7 \sin(\varphi_2 + \varphi_6 - \varphi_7) + \ldots],$$

$$\alpha_1 = [Q^2 + T^2]^{1/2}.$$

In practice it often occurs that some cross magnitudes are not measured. Suitable formulae can be found for these cases too, but are not given here for the sake of simplicity. An examination of the various cases leads to the general formula

$$P(\varphi_{\mathbf{h}}|\ldots) \simeq \frac{1}{L}\exp\left\{\frac{2}{\sqrt{N}}R_{\mathbf{h}}\sum_j R_{\mathbf{k}_j}R_{\mathbf{h}+\mathbf{k}_j}\right.$$
$$\times \cos(\varphi_{\mathbf{h}} + \varphi_{\mathbf{k}_j} - \varphi_{\mathbf{h}+\mathbf{k}_j}) - 2\sum_j w_j C_j$$
$$\left.\times \cos(\varphi_{\mathbf{h}} + \varphi_{\mathbf{k}_j} + \varphi_{\mathbf{l}_j} - \varphi_{\mathbf{h}+\mathbf{k}_j+\mathbf{l}_j})\right\}$$
$$\times I_0(R_1 X_1)\ldots I_0(R_p X_p), \qquad (18)$$

where:

(i) $P(\varphi_{\mathbf{h}}|\ldots)$ is the conditional probability distribution of $\varphi_{\mathbf{h}}$ given any number of magnitudes and any set of phases;

(ii) the first summation concerns all the triplets involving $\varphi_{\mathbf{h}}$ with known phases $\varphi_{\mathbf{k}_j}$, $\varphi_{\mathbf{h}+\mathbf{k}_j}$;

(iii) the second summation concerns all the quartets involving $\varphi_{\mathbf{h}}$ and known phases $\varphi_{\mathbf{k}_j}$, $\varphi_{\mathbf{l}_j}$, $\varphi_{\mathbf{h}+\mathbf{k}_j+\mathbf{l}_j}$. The coefficient w_j is 2, 1, 0, −1 according to whether 3, 2, 1, 0 cross magnitudes of the jth quartet are known;

(iv) the magnitudes $R_1, \ldots, R_p$ which appear in the arguments of the I_0 functions correspond to the cross reflections of the quartets in (iii) provided their phases are unknown.

The mode of the distribution (18), the mean value of $\varphi_{\mathbf{h}}$ and the variance can be obtained by numerical methods.

Equation (18) does not deal with triplets and quartets as if they were independent relationships. Furthermore, it exploits all the information available at a given moment. If the *a priori* information available for $\varphi_{\mathbf{h}}$ changes during the phasing procedure, the way in which (18) is applied changes too.

The contribution of the terms I_0 in (18) is always significant, unless $\varphi_{\mathbf{h}}$ is estimated when all the measured reflections are given assigned phase values. However, that does not occur in usual direct procedures where only reflections with magnitudes larger than about 1·4–1·6 can be phased. This observation suggests an important peculiar property of (18): unlike Sayre's method, (18) is able to estimate $\varphi_{\mathbf{h}}$ both by using the available phase information, and by exploiting the information contained in diffraction magnitudes with unknown phases.

Let us suppose for a moment that all the phases of measured cross reflections of the quartets in (18) are known. Then (18) reduces to the Von Mises distribution (17) where

$$\tan\theta_1 = \tau/Q,$$

$$Q = \frac{2}{\sqrt{N}}R_{\mathbf{h}}\sum_j R_{\mathbf{k}_j}R_{\mathbf{h}+\mathbf{k}_j}\cos(\varphi_{\mathbf{h}+\mathbf{k}_j} - \varphi_{\mathbf{k}_j})$$
$$- 2\sum_j w_j C_j \cos(\varphi_{\mathbf{h}+\mathbf{k}_j+\mathbf{l}_j} - \varphi_{\mathbf{k}_j} - \varphi_{\mathbf{l}_j}),$$

$$T = \frac{2}{\sqrt{N}}R_{\mathbf{h}}\sum_j R_{\mathbf{k}_j}R_{\mathbf{h}+\mathbf{k}_j}\sin(\varphi_{\mathbf{h}+\mathbf{k}_j} - \varphi_{\mathbf{k}_j})$$
$$- 2\sum_j w_j C_j \sin(\varphi_{\mathbf{h}+\mathbf{k}_j+\mathbf{l}_j} - \varphi_{\mathbf{k}_j} - \varphi_{\mathbf{l}_j}), \qquad (19)$$

$$\alpha_1 = [Q^2 + T^2]^{1/2}.$$

The conditions under which (19) holds are similar to those assumed for (8). As a consequence, cross magnitudes with unknown phases are not even used in (19). However, in (18) any magnitude is supposed to be known, while (19) takes account, *via* the coefficient w_j, of the number of cross magnitudes really measured for any quartet. This information is not unimportant; in fact, depending on whether a cross reflection is measured or not, it forms or does not form triplets. This kind of balance of triplet and quartet contributions seems able to reduce the 'squaring' effect during phase-refinement procedures. Suppose now that the contribution of the I_0 functions in (18) cannot be neglected. If N is large enough we can expand the functions as

$$I_0(z) \simeq 1 + z^2/4 \simeq \exp(z^2/4) \qquad (20)$$

and then we obtain the following approximate expression:

$$\tan\varphi_{\mathbf{h}} \simeq \left\{2\frac{R_{\mathbf{h}}}{\sqrt{N}}\sum_j R_{\mathbf{k}_j}R_{\mathbf{h}+\mathbf{k}_j}\sin(\varphi_{\mathbf{h}+\mathbf{k}_j} - \varphi_{\mathbf{k}_j})\right.$$
$$- 2\sum_j C_j(w_j - {}^j w_5\,{}^j R_5^2 - {}^j w_6\,{}^j R_6^2 - {}^j w_7\,{}^j R_7^2)$$
$$\left.\times \sin(\varphi_{\mathbf{h}+\mathbf{k}_j+\mathbf{l}_j} - \varphi_{\mathbf{k}_j} - \varphi_{\mathbf{l}_j})\right\} \bigg/$$
$$\left\{2\frac{R_{\mathbf{h}}}{\sqrt{N}}\sum_j R_{\mathbf{k}_j}R_{\mathbf{h}+\mathbf{k}_j}\right.$$
$$\left.\times \cos(\varphi_{\mathbf{h}+\mathbf{k}_j} - \varphi_{\mathbf{k}_j}) - \ldots\right\}, \qquad (21)$$

where ${}^jR_5$, ${}^jR_6$, ${}^jR_7$ are the cross magnitudes of the jth quartet. ${}^jw_5$, ${}^jw_6$, ${}^jw_7$ are always zero except when ${}^jR_5$, ${}^jR_6$, ${}^jR_7$ are known and the phases unknown; in this case they are equal to unity.

It is worth considering the use of (18) and (21). If (20) is used, (18) reduces to a Von Mises ditribution with a maximum given by (21). If N is large enough (tentatively ≥ 300, then (20) strictly holds and (21) gives a good estimation of $\varphi_{\mathbf{h}}$. From a practical point of view (21) is preferred because it is less time-consuming. If N is not too large (21) can overestimate the quartet with respect to the triplet contribution. In accordance with (13) it should then be convenient to replace C_j in (21) by $C_j/(1 + Q'_j)$, where

$$\begin{aligned} Q'_j = {} & (\varepsilon_{\mathbf{h}}\,\varepsilon_{\mathbf{h}_j} + \varepsilon_{\mathbf{l}_j}\,\varepsilon_{\mathbf{h}+\mathbf{k}_j+\mathbf{l}_j})\,{}^jw_5\,{}^j\varepsilon_5 \\ & + (\varepsilon_{\mathbf{h}}\,\varepsilon_{\mathbf{l}_j} + \varepsilon_{\mathbf{k}_j}\,\varepsilon_{\mathbf{h}+\mathbf{k}_j+\mathbf{l}_j})\,{}^jw_6\,{}^j\varepsilon_6 \\ & + (\varepsilon_{\mathbf{h}}\,\varepsilon_{\mathbf{h}+\mathbf{k}_j+\mathbf{l}_j} + \varepsilon_{\mathbf{k}_j}\,\varepsilon_{\mathbf{l}_j})\,{}^jw_7\,{}^j\varepsilon_7]/2N. \qquad (22) \end{aligned}$$

Q'_j can be set to zero if it is found that $Q'_j < 0$.

8. Centrosymmetric space groups: a general formula using triplet and quartet relationships

Let us calculate in centrosymmetric space groups the conditional probability of $\varphi_{\mathbf{h}}$ when both the magnitudes in the first phasing shell of $\Phi = \varphi_{\mathbf{h}} + \varphi_{\mathbf{k}} + \varphi_{\mathbf{l}} - \varphi_{\mathbf{h}+\mathbf{k}+\mathbf{l}}$ and the phases $\varphi_{\mathbf{k}}$, $\varphi_{\mathbf{l}}$, $\varphi_{\mathbf{h}+\mathbf{k}+\mathbf{l}}$ are known. Using the same notation as for (15) we obtain

$$\begin{aligned} P(E_1 | E_2, E_3, E_4, R_5, R_6, R_7) & \\ \simeq \frac{1}{L}\exp\left(-\frac{2}{N}E_1E_2E_3E_4\right)\cosh(R_5X_5) & \\ \times \cosh(R_6X_6)\cosh(R_7X_7), & \end{aligned}$$

where

$$X_5 = \frac{1}{\sqrt{N}}(E_1E_2 + E_3E_4),$$

$$X_6 = \frac{1}{\sqrt{N}}(E_1E_3 + E_2E_4),$$

$$X_7 = \frac{1}{\sqrt{N}}(E_2E_3 + E_1E_4).$$

By means of considerations similar to those developed for the non-centrosymmetrical case we obtain the general formula

$$\begin{aligned} P(E_{\mathbf{h}}|\ldots) \simeq \frac{1}{L}\exp\left\{E_{\mathbf{h}}\left[\frac{1}{\sqrt{N}}\sum_j E_{\mathbf{k}_j}E_{\mathbf{h}+\mathbf{k}_j}\right.\right. & \\ \left.\left. -\frac{1}{N}\sum w_j E_{\mathbf{k}_j}E_{\mathbf{l}_j}E_{\mathbf{h}+\mathbf{k}_j+\mathbf{l}_j}\right]\right\} & \\ \times \cosh(R_1X_1)\ldots\cosh(R_pX_p), \quad (23) & \end{aligned}$$

where:

(i) $P(E_{\mathbf{h}}|\ldots)$ is the conditional probability distribution of $E_{\mathbf{h}}$ given any number of magnitudes and any set of phases;

(ii) the first summation concerns all the triplets involving $E_{\mathbf{h}}$ with reflections $E_{\mathbf{k}_j}$, $E_{\mathbf{h}+\mathbf{k}_j}$ having known phase;

(iii) the second summation concerns all the quartets involving $E_{\mathbf{h}}$ and reflections $E_{\mathbf{k}_j}$, $E_{\mathbf{l}_j}$, $E_{\mathbf{h}+\mathbf{k}_j+\mathbf{l}_j}$ having known phase. The coefficient w_j is 2, 1, 0, -1 depending on whether 3, 2, 1, 0 cross magnitudes of the jth quartet are known.

(iv) the magnitudes $R_1, \ldots, R_p$ which appear in the arguments of the functions cosh correspond to the cross reflections of the quartets in (iii) provided their signs are unknown.

Since $R_{\mathbf{h}}$ is known, we are mostly interested in the sign $S_{\mathbf{h}}$ of $E_{\mathbf{h}}$. Denoting

$$X_5^{\pm} = \frac{1}{\sqrt{N}}(\pm R_1E_2 + E_3E_4),$$

$$X_6^{\pm} = \frac{1}{\sqrt{N}}[\pm R_1E_3 + E_2E_4), \ldots,$$

we obtain

$$\begin{aligned} P(S_{\mathbf{h}} = + | \ldots) \simeq 1/\{1 + \exp(-2Q) & \\ \times [\cosh(R_1X_1^-)]/[\cosh(R_1X_1^+)] & \\ \ldots [\cosh(R_pX_p^-)]/[\cosh(R_pX_p^+)]\}, & \\ (24) & \end{aligned}$$

where

$$\begin{aligned} Q = R_{\mathbf{h}}\left[\frac{1}{\sqrt{N}}\sum_j E_{\mathbf{k}_j}E_{\mathbf{h}+\mathbf{k}_j}\right. & \\ \left. -\frac{1}{N}\sum_j w_j E_{\mathbf{k}_j}E_{\mathbf{l}_j}E_{\mathbf{h}+\mathbf{k}_j+\mathbf{l}_j}\right]. \quad (25) & \end{aligned}$$

If N is sufficiently large, we can expand $\cosh z$ according to

$$\cosh z \simeq \exp(z^2/2)$$

and then (24) reduces to

$$P(S_{\mathbf{h}} = + | \ldots) \simeq 0{\cdot}5 + 0{\cdot}5\tanh G, \qquad (26)$$

where

$$\begin{aligned} G = \frac{1}{\sqrt{N}}\sum_j R_{\mathbf{h}}E_{\mathbf{k}_j}E_{\mathbf{h}+\mathbf{k}_j} - \frac{1}{N}\sum_j R_{\mathbf{h}}E_{\mathbf{k}_j}E_{\mathbf{l}_j}E_{\mathbf{h}+\mathbf{k}_j+\mathbf{l}_j} & \\ \times (w_j - {}^jw_5\,{}^jR_5^2 - {}^jw_6\,{}^jR_6^2 - {}^jw_7\,{}^jR_7^2). \qquad (27) & \end{aligned}$$

The meaning of symbols in (27) is the same as in (21). In accordance with considerations made for (21), if N is not too large then (26) can overestimate the quartet with respect to the triplet contribution. Thus, $C_j =$

$(1/N)\ R_{\mathbf{h}}E_{\mathbf{k}_j}E_{\mathbf{l}_j}E_{\mathbf{h}+\mathbf{k}_j+\mathbf{l}_j}$ in (27) can be replaced by $C_j/(1+Q_j')$, where Q_j' is defined by (22).

9. A comparison with determinantal approaches in small structures

The maximum-determinant rule was first applied by de Rango (1969) to solve the structure of trigonelline hydrate ($P\bar{1}$). Although other applications have been made from time to time, only in recent years have more systematic applications of determinantal approaches to small structures been tried.

Vermin & de Graaf (1978) successfully applied the rule to pyrocalciferol ($P2_1$, $Z = 2$, $N = 86$). Taylor, Woolfson & Main (1978) combined Karle–Hauptman determinants with a magic-integers approach. Several tests on real structures proved that Karle–Hauptman determinants are very discriminating as figures of merit; however, their power is not sufficient to justify the amount of computing time they need.

If the order of the determinant is much smaller than N, triplet and quartet contributions constitute the most important part in the evaluation of a Karle–Hauptman determinant. Let us calculate up to order $1/N$ the joint probability distribution of all the normalized structure factors which appear in a determinant D_n under the hypothesis that the reciprocal vectors are fixed and the positional vectors are the primitive random variables. In short notation we shall denote this distribution by

$$P \equiv P(\{E_{\mathbf{h}_j+\mathbf{k}}\}, \{E_{\mathbf{h}_i-\mathbf{h}_j}\},\quad i,j=1,\ldots,n).$$

We obtain [see Heinermann, Krabbendam & Kroon (1979) and Heinermann, Kroon & Krabbendam (1979) for an expression obtained with a different hypothesis]

$$\begin{aligned} P \simeq \frac{1}{\pi^{n(n-1)/2}} \exp\Bigg\{ & -\sigma^2 + \frac{2}{\sqrt{N}} \sum R_{\mathbf{h}_i+\mathbf{k}} R_{\mathbf{h}_j+\mathbf{k}} R_{\mathbf{h}_i-\mathbf{h}_j} \\ & \times \cos(\varphi_{\mathbf{h}_i+\mathbf{k}} - \varphi_{\mathbf{h}_j+\mathbf{k}} - \varphi_{\mathbf{h}_i-\mathbf{h}_j}) \\ & + \frac{2}{\sqrt{N}} \sum R_{\mathbf{h}_i-\mathbf{h}_j} R_{\mathbf{h}_i-\mathbf{h}_q} R_{\mathbf{h}_j-\mathbf{h}_q} \cos(\varphi_{\mathbf{h}_i-\mathbf{h}_j} - \varphi_{\mathbf{h}_i-\mathbf{h}_q} \\ & + \varphi_{\mathbf{h}_j-\mathbf{h}_q}) - \frac{2}{N} \sum R_{\mathbf{h}_q+\mathbf{k}} R_{\mathbf{h}_i+\mathbf{k}} R_{\mathbf{h}_i-\mathbf{h}_j} R_{\mathbf{h}_j-\mathbf{h}_q} \\ & \times \cos(\varphi_{\mathbf{h}_q+\mathbf{k}} - \varphi_{\mathbf{h}_i+\mathbf{k}} + \varphi_{\mathbf{h}_i-\mathbf{h}_j} + \varphi_{\mathbf{h}_j-\mathbf{h}_q}) \\ & - \frac{2}{N} \sum R_{\mathbf{h}_j-\mathbf{h}_q} R_{\mathbf{h}_i-\mathbf{h}_p} R_{\mathbf{h}_i-\mathbf{h}_q} R_{\mathbf{h}_j-\mathbf{h}_p} \cos(\varphi_{\mathbf{h}_j-\mathbf{h}_q} \\ & + \varphi_{\mathbf{h}_i-\mathbf{h}_p} - \varphi_{\mathbf{h}_i-\mathbf{h}_q} - \varphi_{\mathbf{h}_j-\mathbf{h}_p}) + \ldots \Bigg\}, \end{aligned} \tag{28}$$

where A is the product of the $n(n-1)/2$ magnitudes involved in D_n and σ^2 is the sum of their squares.

The coefficient $-2/N$ of the quartet contributions is due to the fact that only two cross magnitudes per quartet are in the determinant. This means that Karle–Hauptman determinants are in general unable to exploit the information contained in the third cross reflections of the quartets. Unfortunately, they are also unable to exploit the information about the magnitudes of the first two cross reflections when their phases are unknown (the common practice is that of equating a reflection in D_n to zero when its phase is unknown).

If (28) is further developed, an expression similar to (18) can be obtained which also holds when the phases of some of the two cross reflections of the quartets in D_n are unknown. In conclusion, a relationship such as (18) is a more flexible and powerful tool for estimating a phase $\varphi_{\mathbf{h}}$ than a small-order Karle–Hauptman determinant.

References

Allegra, G. & Colombo, A. (1974). *Acta Cryst.* A**30**, 727–729.

Busetta, B., Giacovazzo, C., Burla, M. C., Nunzi, A., Polidori, G. & Viterbo, D. (1980). *Acta Cryst.* A**36**, 68–74.

Collins, D. M. (1975). *Acta Cryst.* A**31**, 388–389.

De Titta, G. T., Edmonds, J. W., Langs, D. A. & Hauptman, H. (1975). *Acta Cryst.* A**31**, 472–479.

Eller, G. von (1973). *Acta Cryst.* A**29**, 63–67.

Fan, H. (1965). *Acta Phys. Sin.* **21**, 1105–1113.

Giacovazzo, C. (1975). *Acta Cryst.* A**31**, 252–259.

Giacovazzo, C. (1976*a*). *Acta Cryst.* A**32**, 91–99.

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

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

Giacovazzo, C. (1977*b*). *Acta Cryst.* A**33**, 944–948.

Gilmore, J. (1977). *Acta Cryst.* A**33**, 712–716.

Gilmore, J. (1979). Personal communication.

Hauptman, H. (1975). *Acta Cryst.* A**31**, 671–679.

Heinermann, J. J. L., Krabbendam, H. & Kroon, J. (1979). *Acta Cryst.* A**35**, 101–105.

Heinermann, J. J. L., Kroon, J. & Krabbendam, H. (1979). *Acta Cryst.* A**35**, 105–107.

Hoppe, W., Gassmann, J. & Zechmeister, K. (1970). In *Crystallographic Computing*, edited by F. R. Ahmed. Copenhagen: Munksgaard.

Krabbendam, H. & Kroon, J. (1971). *Acta Cryst.* A**27**, 48–53.

Putten, N. van der (1979). Personal communication.

Rango, C. de (1969). Thèse. Paris.

Sayre, D. (1952). *Acta Cryst.* **5**, 60.

Schenk, H. (1974). *Acta Cryst.* A**30**, 477–481.

Schenk, H. & van der Putten, N. (1978). International School of Crystallography, Erice, Italy.

Silverton, J. V. (1978). *Acta Cryst.* A**34**, 634.

Simerska, M. (1956). *Czech. J. Phys.* **6**, 1–7.

Simonov, V. I. (1976). In *Crystallographic Computing Techniques*, edited by F. R. Ahmed. Copenhagen: Munksgaard.

Taylor, D. J., Woolfson, M. M. & Main, P. (1978). *Acta Cryst.* A**34**, 870–883.

Vermin, W. J. & de Graaf, R. A. G. (1978). *Acta Cryst.* A**34**, 892–894.

Woolfson, M. M. (1958). *Acta Cryst.* **11**, 277–283.