The Method of Representations of Structure Seminvariants.
The Strengthening of Triplet Relationships

By G. Cascarano and C. Giacovazzo

Istituto di Mineralogia e Petrografia, Università, 70121 Bari, Italy

M. Camalli and R. Spagna

Laboratorio di Strutturistica Chemica G. Giacomello, CP10 00016 Monterotondo Stazione, Roma, Italy

M. C. Burla and A. Nunzi

Istituto di Mineralogia dell' Università, 06100 Perugia, Italy

and G. Polidori

Centro di Calcolo Elettronico dell' Università, 06100 Perugia, Italy

Abstract

A probabilistic approach is described which is able to estimate triplet cosine invariants *via* their second representation. The experimental tests on structures of different complexity show that triplets with positive cosine are estimated with an accuracy better than *via* Cochran's distribution. Furthermore, in favourable cases, a large percentage of negative triplets can be found.

1. Introduction

The triplet relationship

$$\Phi_3 = \varphi_{h_1} + \varphi_{h_2} + \varphi_{h_3} \simeq 0 \ (\text{mod} \ 2\pi), \ h_1 + h_2 + h_3 = 0 \quad (1)$$

is the most widely used phase relationship for solving crystal structures. The occurrence of a few 'bad' triplets [*i.e.* triplets for which (1) is violated] in the early stages of a phase-extension procedure can prevent the crystallographer from solving the crystal structure. Furthermore, the use of a relatively large number of bad triplets in a refinement process can lead to wrong results even if the set of starting phases is relatively accurate (the structural information is lost under refinement). The probability of finding the correct solution is enhanced if the bad triplet relationships are recognized: then they may be excluded from the structure-solving process or suitably used.

Traditionally, Φ_3 is estimated by the Cochran (1955) formula (denoted by P_3 from now on):

$$P_3 \cong [2\pi I_0(C)]^{-1} \exp (C \cos \Phi_3), \quad (2)$$

where I_0 is the modified Bessel function of order zero,

$$C = 2\sigma_3\sigma_2^{-3/2}|E_{h_1}E_{h_2}E_{h_3}|, \qquad \sigma_n = \sum_{j=1}^{N} Z_j^n,$$

Z_j is the atomic number of the jth atom of the structure, N is the number of atoms in the unit cell.

Cochran's formula estimates Φ_3 by exploiting only the information contained in the three moduli $|E_{h_1}|, |E_{h_2}|, |E_{h_3}|$. The assumption that the information contained in all the reciprocal space could be used in order to improve Cochran's estimate of Φ_3 inspired a great deal of work. In this view several formulae were proposed by various authors: the connections of some of them with the present theory are discussed in the Appendix.*

The theory of representations (Giacovazzo, 1977a, b; see Hauptman, 1975, 1977, for a related concept) gave the authors new insight into the problem of recognizing good and bad triplets. We describe here some theoretical results arising from such a theory and practical tests on several crystal structures.

2. The estimation of Φ_3 via its second representation

In accordance with Giacovazzo (1977b), the second representation $\{\psi\}_2$ of the triplet Φ_3 is the collection of special quintets

$$\psi_2 = \Phi_3 + \varphi_k - \varphi_k, \quad (3)$$

where k is a free vector in the reciprocal space. The basis magnitudes of any ψ_2 are $(R = |E|)$

$$R_{h_1}, R_{h_2}, R_{h_3}, R_k$$

and the cross magnitudes are

$$R_{\mathbf{h}_1 \pm \mathbf{k}}, \quad R_{\mathbf{h}_2 \pm \mathbf{k}}, \quad R_{\mathbf{h}_3 \pm \mathbf{k}}.$$

The collection of the basis and cross magnitudes of the various quintets ψ_2 is $\{B\}_2$, which is called the second phasing shell of Φ_3:

$$\{B\}_2 = \{R_{\mathbf{h}_1}, R_{\mathbf{h}_2}, R_{\mathbf{h}_3}, R_{\mathbf{k}}, R_{\mathbf{h}_1 \pm \mathbf{k}}, R_{\mathbf{h}_2 \pm \mathbf{k}}, R_{\mathbf{h}_3 \pm \mathbf{k}}\}.$$

Such considerations suggest for $P1$ and $P\bar{1}$ the study of the ten-variate probability distribution

$$P(E_{\mathbf{h}_1}, E_{\mathbf{h}_2}, E_{\mathbf{h}_3}, E_{\mathbf{k}}, E_{\mathbf{h}_1+\mathbf{k}}, E_{\mathbf{h}_1-\mathbf{k}},$$
$$E_{\mathbf{h}_2+\mathbf{k}}, E_{\mathbf{h}_2-\mathbf{k}}, E_{\mathbf{h}_3+\mathbf{k}}, E_{\mathbf{h}_3-\mathbf{k}}), \tag{4}$$

from which the conclusive conditional probability distribution

$$P(\Phi_3 | 10 \text{ moduli}) \tag{5}$$

is obtained. Equations (4) and (5) may be calculated by means of techniques described in a recent monograph (Giacovazzo, 1980). For the sake of brevity we do not give here any details of these techniques.

Since $\mathbf{k}$ is a free vector a formula can be found which provides the conditional probability distribution of Φ_3 given the basis and cross moduli of any set of quintets ψ_2. We will denote such a probability by $P_{10}(\Phi_3)$ in order to emphasize the fact that the formula explores reciprocal space by means of a ten-node figure. Three nodes (*i.e.* $\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3$) are fixed while $\mathbf{k}$ varies: the remaining seven nodes sweep out reciprocal space.

If the space group has symmetry higher than triclinic, any quintet (3) depends on more than ten magnitudes. Indeed, (3) gives rise to m equivalent quintets:

$$\psi_2 = \Phi_3 + \varphi_{\mathbf{kR}_i} - \varphi_{\mathbf{kR}_i}, \quad i = 1, 2, \dots, m,$$

where m is the number of symmetry operators $\mathbf{C}_j \equiv (\mathbf{R}_j, \mathbf{T}_j)$ not related by a centre of symmetry ($\mathbf{R}_j$ is the rotational part, $\mathbf{T}_j$ the translational part of the symmetry operator). In conclusion, any quintet ψ_2 depends on, in addition to the basis magnitudes $R_{\mathbf{h}_1}, R_{\mathbf{h}_2}, R_{\mathbf{h}_3}, R_{\mathbf{k}}$, the $6m$ cross magnitudes

$$R_{\mathbf{h}_1+\mathbf{kR}_1}, \ R_{\mathbf{h}_1-\mathbf{kR}_1}, \ R_{\mathbf{h}_2+\mathbf{kR}_1}, \ R_{\mathbf{h}_2-\mathbf{kR}_1}, \ R_{\mathbf{h}_3+\mathbf{kR}_1}, \ R_{\mathbf{h}_3-\mathbf{kR}_1}$$
$$R_{\mathbf{h}_1+\mathbf{kR}_2}, \ R_{\mathbf{h}_1-\mathbf{kR}_2}, \ R_{\mathbf{h}_2+\mathbf{kR}_2}, \ R_{\mathbf{h}_2-\mathbf{kR}_2}, \ R_{\mathbf{h}_3+\mathbf{kR}_2}, \ R_{\mathbf{h}_3-\mathbf{kR}_2}$$
$$\vdots \qquad \cdots \qquad \vdots$$
$$R_{\mathbf{h}_1+\mathbf{kR}_m}, \ R_{\mathbf{h}_1-\mathbf{kR}_m}, \ R_{\mathbf{h}_2+\mathbf{kR}_m}, \ R_{\mathbf{h}_2-\mathbf{kR}_m}, \ R_{\mathbf{h}_3+\mathbf{kR}_m}, \ R_{\mathbf{h}_3-\mathbf{kR}_m}.$$
$$\tag{6}$$

In this case we may write

$$\{B\}_2 = \{R_{\mathbf{h}_1}, R_{\mathbf{h}_2}, R_{\mathbf{h}_3}, R_{\mathbf{k}}, R_{\mathbf{h}_1 \pm \mathbf{kR}_i}, R_{\mathbf{h}_2 \pm \mathbf{kR}_i}, R_{\mathbf{h}_3 \pm \mathbf{kR}_i},$$
$$i = 1, 2, \dots, m\}.$$

The conditional probability $P(\Phi_3 | \{B\}_2)$ now provides an estimate for Φ_3 in any space group. Since its exact expression is not easy to obtain, we prefer to introduce a simple approximation of $P(\Phi_3 | \{B\}_2)$ which may be derived as a proper combination of the various ten-variate distributions:

$$P(\Phi_3 | R_{\mathbf{h}_1}, R_{\mathbf{h}_2}, R_{\mathbf{h}_3}, R_{\mathbf{k}}, R_{\mathbf{h}_1 \pm \mathbf{kR}_i}, R_{\mathbf{h}_2 \pm \mathbf{kR}_i}, R_{\mathbf{h}_3 \pm \mathbf{kR}_i}). \tag{7}$$

All the calculations described in this paper refer to such a formula which will be shortly denoted by P_{10}. Of course, $\mathbf{k}$ is a free vector in our procedure and varies over the asymmetric region of reciprocal space.

3. The estimation of Φ_3 in the centrosymmetric space groups

In the centrosymmetric space groups $P_{10}(\Phi_3)$ may be replaced by P_{10}^+ which gives the conditional probability that the triplet is positive. We obtain

$$P_{10}^+ \simeq \tfrac{1}{2} + \tfrac{1}{2} \tanh G, \tag{8}$$

where

$$G = C(1 + Q),$$

$$Q = \sum_{\mathbf{k}} \left[\frac{\sum\limits_{i=1}^{m'} A_{\mathbf{k},i} / N}{1 + (\varepsilon_{\mathbf{h}_1} \varepsilon_{\mathbf{h}_2} \varepsilon_{\mathbf{h}_3} + \sum\limits_{i=1}^{m'} B_{\mathbf{k},i})/2N} \right], \tag{9}$$

$$A_{\mathbf{k},i} = \varepsilon_{\mathbf{k}} [\varepsilon_{\mathbf{h}_1+\mathbf{kR}_i}(\varepsilon_{\mathbf{h}_2-\mathbf{kR}_i} + \varepsilon_{\mathbf{h}_3-\mathbf{kR}_i})$$
$$+ \varepsilon_{\mathbf{h}_2+\mathbf{kR}_i}(\varepsilon_{\mathbf{h}_1-\mathbf{kR}_i} + \varepsilon_{\mathbf{h}_3-\mathbf{kR}_i})$$
$$+ \varepsilon_{\mathbf{h}_3+\mathbf{kR}_i}(\varepsilon_{\mathbf{h}_1-\mathbf{kR}_i} + \varepsilon_{\mathbf{h}_2-\mathbf{kR}_i})],$$

$$B_{\mathbf{k},i} = \varepsilon_{\mathbf{h}_1}[\varepsilon_{\mathbf{k}}(\varepsilon_{\mathbf{h}_1+\mathbf{kR}_i} + \varepsilon_{\mathbf{h}_1-\mathbf{kR}_i})$$
$$+ \varepsilon_{\mathbf{h}_2+\mathbf{kR}_i}\varepsilon_{\mathbf{h}_3-\mathbf{kR}_i} + \varepsilon_{\mathbf{h}_2-\mathbf{kR}_i}\varepsilon_{\mathbf{h}_3+\mathbf{kR}_i}]$$
$$+ \varepsilon_{\mathbf{h}_2}[\varepsilon_{\mathbf{k}}(\varepsilon_{\mathbf{h}_2+\mathbf{kR}_i} + \varepsilon_{\mathbf{h}_2-\mathbf{kR}_i})$$
$$+ \varepsilon_{\mathbf{h}_1+\mathbf{kR}_i}\varepsilon_{\mathbf{h}_3-\mathbf{kR}_i} + \varepsilon_{\mathbf{h}_1-\mathbf{kR}_i}\varepsilon_{\mathbf{h}_3+\mathbf{kR}_i}]$$
$$+ \varepsilon_{\mathbf{h}_3}[\varepsilon_{\mathbf{k}}(\varepsilon_{\mathbf{h}_3+\mathbf{kR}_i} + \varepsilon_{\mathbf{h}_3-\mathbf{kR}_i})$$
$$+ \varepsilon_{\mathbf{h}_1+\mathbf{kR}_i}\varepsilon_{\mathbf{h}_2-\mathbf{kR}_i} + \varepsilon_{\mathbf{h}_1-\mathbf{kR}_i}\varepsilon_{\mathbf{h}_2+\mathbf{kR}_i}];$$

$\varepsilon = |E|^2 - 1$, $(\varepsilon_{\mathbf{h}_1}\varepsilon_{\mathbf{h}_2}\varepsilon_{\mathbf{h}_3} + \sum_{i=1}^{m} B_{\mathbf{k},i})$ is assumed to be zero if it is experimentally negative. The prime to the summation warns the reader that precautions have to be taken in order to avoid duplications in the contributions (see § 6).

G may be positive or negative. In particular, if $G < 0$ the triplet is estimated negative.

The accuracy with which the value of Φ_3 is estimated by (8) strongly depends on $\varepsilon_{\mathbf{k}}$. Thus, in practice, only a subset of reciprocal space (the reflections $\mathbf{k}$ with large values of ε) may be used for estimating Φ_3.

4. The estimation of Φ_3 in non-centrosymmetric space groups

$\cos \Phi_3$ is estimated *via* the von Mises distribution

$$P_{10}(\Phi_3) \simeq \frac{1}{2\pi I_0(G)} \exp(G \cos \Phi_3), \tag{10}$$

where $G = 2C(1 + Q)$ and Q is defined by (9).

The distribution (10) is unimodal and takes its maximum at $\Phi = 0$ if $G > 0$, at $\Phi = \pi$ if $G < 0$. Therefore, it is able in principle to estimate with high reliability only triplets with phase values around 0 or π. In accordance with the theory, enantiomorph-sensitive triplets should present a rather flat distribution ($G \simeq 0$), so that they cannot be reliably fixed. This is the most important limitation of the present theory: however, its use may have important applications in direct methods, both during the tangent procedures and for recognizing the correct solution in multisolution approaches (Camalli, Cascarano, Giacovazzo, Spagna & Viterbo, 1984).

It is expected that the theory will be more useful for centrosymmetric than for non-centrosymmetric space groups. Indeed, the formulas depend on the ε factors: according to Wilson's distributions, $\langle |\varepsilon| \rangle$ is 0.968 for centrosymmetric crystal structures and 0.736 for non-centrosymmetric ones.

The larger dispersion of the ε factors around zero for centrosymmetric space groups favours larger contributions to the centrosymmetric formula.

5. The theory of representations and the quadrupoles

From the phasing magnitudes of the second representation a number of quadrupoles may be constructed, each contributing to the estimation of Φ_3. For example, from the magnitudes on the ith line of the matrix (6) the following six quadrupoles arise (of type I):

$$
\left\{
\begin{array}{l}
\varphi_{h_1} + \varphi_{h_2} + \varphi_{h_3} \\
-\varphi_{h_1} - \varphi_{kR_i} + \varphi_{h_1 + kR_i} \\
-\varphi_{h_2} + \varphi_{kR_i} + \varphi_{h_2 - kR_i} \\
-\varphi_{h_3} - \varphi_{h_1 + kR_i} - \varphi_{h_2 - kR_i}
\end{array}
\right.
\quad
\left\{
\begin{array}{l}
\varphi_{h_1} + \varphi_{h_2} + \varphi_{h_3} \\
-\varphi_{h_1} - \varphi_{kR_i} + \varphi_{h_1 + kR_i} \\
-\varphi_{h_2} - \varphi_{h_1 + kR_i} - \varphi_{h_3 - kR_i} \\
-\varphi_{h_3} + \varphi_{kR_i} + \varphi_{h_3 - kR_i}
\end{array}
\right.
$$

$$
\left\{
\begin{array}{l}
\varphi_{h_1} + \varphi_{h_2} + \varphi_{h_3} \\
-\varphi_{h_1} + \varphi_{kR_i} + \varphi_{h_1 - kR_i} \\
-\varphi_{h_2} - \varphi_{kR_i} + \varphi_{h_2 + kR_i} \\
-\varphi_{h_3} - \varphi_{h_1 - kR_i} - \varphi_{h_2 + kR_i}
\end{array}
\right.
\quad
\left\{
\begin{array}{l}
\varphi_{h_1} + \varphi_{h_2} + \varphi_{h_3} \\
-\varphi_{h_1} - \varphi_{h_2 + kR_i} - \varphi_{h_3 - kR_i} \\
-\varphi_{h_2} - \varphi_{kR_i} + \varphi_{h_2 + kR_i} \\
-\varphi_{h_3} + \varphi_{kR_i} + \varphi_{h_3 - kR_i}
\end{array}
\right.
\quad (11)
$$

$$
\left\{
\begin{array}{l}
\varphi_{h_1} + \varphi_{h_2} + \varphi_{h_3} \\
-\varphi_{h_1} + \varphi_{kR_i} + \varphi_{h_1 - kR_i} \\
-\varphi_{h_2} - \varphi_{h_1 - kR_i} - \varphi_{h_3 + kR_i} \\
-\varphi_{h_3} - \varphi_{kR_i} + \varphi_{h_3 + kR_i}
\end{array}
\right.
\quad
\left\{
\begin{array}{l}
\varphi_{h_1} + \varphi_{h_2} + \varphi_{h_3} \\
-\varphi_{h_1} - \varphi_{h_2 - kR_i} - \varphi_{h_3 - kR_i} \\
-\varphi_{h_2} + \varphi_{kR_i} + \varphi_{h_2 - kR_i} \\
-\varphi_{h_3} - \varphi_{kR_i} + \varphi_{h_3 + kR_i}.
\end{array}
\right.
$$

Each of the above quadrupoles gives a well recognizable contribution to (8). For example,

$$
\varepsilon_k \varepsilon_{h_1 + kR_i} \varepsilon_{h_2 - kR_i}, \qquad \varepsilon_k \varepsilon_{h_1 + kR_i} \varepsilon_{h_3 - kR_i},
$$
$$
\varepsilon_k \varepsilon_{h_1 - kR_i} \varepsilon_{h_2 + kR_i} \cdots
$$

are, in order, the contributions to A arising from the quadrupoles (11). Varying R_i over the m symmetry operators and k over the asymmetric region of the reciprocal space leads, by a probabilistic approach, to our formulas (8) and (10).

The question now is if (11) are or are not the only quadrupoles exploitable by the second representa-

tion. The form

$$
\begin{array}{l}
\varphi_{h_1} + \varphi_{h_2} + \varphi_{h_3} \\
-\varphi_{h_1} - \varphi_{kR_i} + \varphi_{h_1 + kR_i} \\
-\varphi_{h_2} + \varphi_{kR_j} + \varphi_{h_2 - kR_j} \\
-\varphi_{h_3} - \varphi_{(h_1 + kR_i)R_p} - \varphi_{(h_2 - kR_j)R_s}
\end{array}
\quad (12)
$$

is a quadrupole too (type II), provided k, R_p and R_s are suitable matrices for which

$$
h_1 R_p + h_2 R_s + h_3 + kR_i R_p - kR_j R_s = 0. \quad (13)
$$

The quadrupole (12) is structurally different from quadrupoles (11) because: (a) it involves cross terms from two lines of the matrix (6); (b) the sum of the four triplets in (12) is no longer strictly equal to zero. Indeed, (12) may be written as

$$
\left\{
\begin{array}{l}
\varphi_{h_1} + \varphi_{h_2} + \varphi_{h_3} \\
-\varphi_{h_1} - \varphi_k + \varphi_{h_1 + kR_i} + 2\pi kT_i \\
-\varphi_{h_2} + \varphi_k + \varphi_{h_2 - kR_j} - 2\pi kT_j \\
-\varphi_{h_3} - \varphi_{h_1 + kR_i} - \varphi_{h_2 - kR_j} \\
\quad + 2\pi[(h_1 + kR_i)T_p + (h_2 - kR_j)T_s],
\end{array}
\right.
$$

from which the sum Δ of the four triplets is easily found:

$$
\Delta = 2\pi[h_1 T_p + h_2 T_s + k(T_i - T_j + kR_i T_p - kR_j T_s)].
$$

Examples of this kind of quadrupole were given by Viterbo & Woolfson (1973). These authors called 'consistent' the quadrupoles for which $\Delta = 0$, the others inconsistent.

In the symmorphic space groups inconsistent quadrupoles cannot exist, but quadrupoles with terms lying in two lines of the matrix (6) do occur.

The above considerations suggest the following.

(a) The algebraic theory of representations is able to take into account quadrupoles of types I and II (or consistent and inconsistent quadrupoles). The contributions of the various quadrupoles to (8) and (10) are organized per quintets. This seems to introduce an efficient exploitation of their properties (see § 6).

(b) Our approach is applied in the present paper only to quadrupoles I. The reasons for this limitation are only practical: the quadrupoles II only occur when condition (13) is verified (special relations between k and the h_i's are required), so their percentage is usually small, especially for space groups of low symmetry, and their search is time consuming. Incidentally, their contribution may be relevant for making easier the direct solution of some crystal structures. We defer to a further paper the task of evaluating quadrupoles II and their relevance for structure solution.

Table 1. *Abbreviations, references, space groups and formulas for the sixteen test structures*

N is the number of non-hydrogen atoms in the unit cell.

	References	Space group	Formula	Z	N
KENNA	Shakked & Kennard (1977)	$P\bar{1}$	$C_{30}H_{37}NO_5$	2	72
GRA4	York (private communication)	$P\bar{1}$	$C_{30}H_{22}N_2O_4$	2	72
EVIT	Schenk, Kops, van der Putten & Bode (1978)	$P2_1/c$	$C_{21}H_{30}O_2$	4	92
PG205	York (private communication)	$P2_1/n$	$C_{19}H_{22}N_2O$	8	176
QUINO	Wallwork & Powell (1980)	$R\bar{3}$	C_6O_2	54	432
PMM1	Goldberg (1975)	$Pnma$	$C_{17}H_{24}O_7$	8	192
REGOL	Goldberg & Rezmovitz (1978)	$Pbca$	$C_{20}H_{23}NO_6$	8	216
ROSA	Wong (1978)	$Pbca$	$C_{21}H_{19}NO$	8	184
BCD	York (private communication)	$P2_1$	$(C_6H_{10}O_5)_7.12H_2O$	2	178
HEPTA	Beurskens, Beurskens & van den Hark (1976)	$P2_1$	$C_{13}H_{18}O_9$	4	88
FACTS	York (private communication)	$P2_12_12_1$	$C_{45}H_{57}N_7O_{12}$	4	256
ERGO	Hull, Leban, Main, White & Woolfson (1976)	$P2_12_12_1$	$C_{28}H_{44}O$	8	232
AZET	Colens, Declercq, Germain, Putzeys & Van Meerssche (1974)	$Pca2_1$	$C_{21}H_{16}ClNO$	8	192
APAPA	Suck, Manor & Saenger (1976)	$P4_12_12$	$C_{30}H_{35}N_{15}O_{16}P_2$	8	504
DIOLE	York (private communication)	$I\bar{4}2d$	$C_{10}H_{18}O_2$	16	192
TUR10	York (private communication)	$P6_322$	$C_{15}H_{24}O_2$	12	204

6. Experimental results

For our tests a computer program has been implemented which can estimate triple-phase relationships in all the space groups according to P_3 and P_{10}. The program seeks triplets among the largest 400 $|E|$'s and stores the 3000 most reliable ones. The running vector **k** is allowed to vary over a limited number of reflections (over the $|E|$'s larger than a given threshold E_t). For convenience the experimental G values are rescaled on C values.

Throughout the procedure checks are made in order to avoid duplications of the contributions to P_{10} [we emphasized that by inserting the prime to the summation symbol in (9)] and the use of special quintets [for which (8) and (10) do not hold]. Thus, a vector $\mathbf{kR}_j$ is not used in (7) if one of the following relations occurs.

$\mathbf{kR}_j = \pm\mathbf{kR}_i$, where $\mathbf{kR}_i$ has been previously used.

$(\mathbf{h}_i \pm \mathbf{kR}_j)$ is symmetry equivalent to one of the $\mathbf{h}_p$'s or to **k**.

$|E_{\mathbf{h}_i\pm\mathbf{kR}_j}| \geq |E_\mathbf{k}| \geq E_t$. Such a condition is introduced because the running vector **k** will become in different moments both **k** and $-\mathbf{h}_i \mp \mathbf{kR}_j$, giving rise to two quintets providing identical contributions to (8) and (10). By means of the above condition we chose one of them.

The reliability of the P_3 and P_{10} formulas were tested on known structures of different complexity. Table 1 shows their references and the most relevant features.

In Table 2 we give for QUINO and for various values of ARG the number of triplet relationships (nr) having the argument of the hyperbolic tangent larger than ARG. According to P_3 all the triplets are always estimated positive, according to P_{10} some of them are estimated negative. The number of wrongly estimated triplets (nw) is given in parentheses for each nr. In practice, in Table 2 the triplet relationships are ranked in decreasing order of reliability according to P_3 and P_{10}. The following may be observed.

Table 2. *QUINO: number of triplet relationships (nr) and number of wrongly estimated triplet relationships (nw) according to P_3 and P_{10}*

The running vector **k** is allowed to vary on the largest 150 $|E|$'s.

ARG	P_3		Positive estimated triplets P_{10}		Negative estimated triplets P_{10}	
	(nr)	(nw)	(nr)	(nw)	(nr)	(nw)
0·0			2824	(279)	176	(29)
0·8	3000	(426)	2148	(53)	12	(0)
1·4	935	(63)	1258	(5)	2	(0)
2·0	325	(9)	583	(1)		
2·4	180	(1)	312	(0)		
3·4	50	(1)	48	(0)		
5·0	10	(0)	2	(0)		

(*a*) According to P_3 the first error occurs at 33 in the list; according to P_{10} the first error occurs at 577 in the list. There are 63 errors among the most reliable 935 triplets estimated positive by P_3 and only 5 errors among the most reliable 1258 triplets estimated positive by P_{10}. Thus, positive triplets are selected with much more effectiveness by P_{10} than by P_3.

(*b*) A relevant number of triplets is estimated negative by P_{10} and a large percentage of them is really negative. The first error in the list of the negative triplets is the 33rd relationship.

It would take too much space to give a table such as 2 for each crystal structure; thus, for the sake of brevity, we collect in Table 3 the most useful figures concerning our tests on centrosymmetric structures. For each crystal structure the running vector **k** is allowed to vary over the 70, 100, 150 largest $|E|$'s in order to check the influence of the allowed ranges on the accuracy of the formula. For each range:

(*a*) the triplets are ranked in decreasing order of reliability according to P_{10} and P_3; in the table, the order number (nfe) of the first error in the lists is given. The entry for P_3 is given in parentheses;

Table 3. *Useful figures concerning the tests on centrosymmetric structures*

P_{10} is used by allowing the running vector **k** to vary over the largest 70, 100, 150 $|E|$'s. See text for the meaning of the symbols.

	Positive estimated triplets $P_{10}(P_3)$ (nfe)	Negative estimated triplets (nr)	(n)
KENNA	1634 (555)	190	(31)
	1714	198	(34)
	1682	198	(30)
GRA4	1761 (831)	36	(16)
	1981	38	(17)
	2101	40	(17)
EVIT	1501 (157)	219	(42)
	1617	225	(35)
	1589	226	(29)
PG205	135 (77)	239	(105)
	155	234	(96)
	180	213	(81)
QUINO	478 (33)	239	(50)
	521	233	(45)
	577	176	(29)
PMM1	1418 (289)	191	(45)
	1307	187	(41)
	1296	171	(23)
REGOL	765 (66)	221	(40)
	924	196	(33)
	754	179	(26)
ROSA	29 (24)	70	(31)
	31	41	(23)
	29	20	(11)

Table 4. *Useful figures concerning the tests on non-centrosymmetric structures*

P_{10} is used by allowing the running vector **k** to vary over the largest 70, 100, 150 $|E|$'s. See text for the meaning of the symbols.

	First 100 positive estimated triplets $P_{10}(P_3)$ $\langle|\varphi_3|\rangle$	First 100 positive estimated triplets $P_{10}(P_3)$ necos	Negative estimated triplets nest	$\langle|\varphi_3|\rangle$
BCD	38 (45)	4 (9)	5	94
	36	2	8	118
	37	4	11	85
HEPTA	20 (24)	0 (1)	0	—
	20	0	0	—
	21	0	0	—
FACTS	48 (56)	11 (15)	12	95
	47	11	18	103
	48	12	24	109
ERGO	38 (46)	7 (10)	34	98
	37	6	34	90
	36	7	35	91
AZET	30 (37)	2 (6)	103	97
	32	3	108	103
	30	2	78	107
APAPA	38 (51)	5 (14)	3	77
	41	7	4	96
	39	7	2	73
DIOLE	41 (57)	3 (15)	0	—
	41	3	0	—
	42	3	0	—
TUR10	37 (42)	7 (9)	64	109
	40	6	42	112
	36	5	12	124

(b) the total number (nr) of the triplets estimated negative by P_{10} is given together with the number of errors ne (in parentheses).

Table 3 confirms the trend of Table 2. Even for ROSA, where P_{10} and P_3 seem to rank the positive triplets with the same accuracy, P_{10} is really more effective. Indeed, a complete view of the experimental results would show, for example, that among the 250 most reliable triplets estimated positive, there are 31 errors if estimated by P_3, and only 10 errors if estimated by P_{10}.

Table 4 collects some useful figures concerning our tests for non-centrosymmetric structures. Again, the running vector **k** is allowed to vary over the 70, 100, 150 largest $|E|$'s. For each crystal structure and for each range Table 4 shows:

(a) the average values $\langle|\Phi_3|\rangle$ observed for the most reliable 100 triplet relationships estimated positive according to P_{10} (to P_3 in parentheses);

(b) the number of triplets (necos) with negative cosine among the most reliable 100 triplets estimated positive by P_{10} (by P_3 in parentheses);

(c) the number (nest) of triplets whose cosine is estimated negative by P_{10} and the observed average value $\langle|\Phi_3|\rangle$.

A short analysis of Table 4 shows that P_{10} is more effective than P_3 for selecting sets of triplets with Φ_3 closely distributed about the value of zero. Furthermore, P_{10} often selects a number of triplets with Φ_3 distributed far from zero: thus they can be eliminated

Table 5. *Useful figures concerning the application of P_{10} and MDKS formulas to the five non-centrosymmetric structures*

The meanings of the symbols are the same as in Table 4.

	First 100 positive estimated triplets P_{10} (MDKS) $\langle|\varphi_3|\rangle$	First 100 positive estimated triplets P_{10} (MDKS) necos	Negative estimated triplets P_{10} (MDKS) nest	P_{10} (MDKS) $\langle|\varphi_3|\rangle$
BCD	38 (47)	4 (16)	5 (632)	94 (64)
HEPTA	20 (25)	0 (4)	0 (381)	— (43)
FACTS	48 (41)	11 (11)	12 (665)	95 (61)
ERGO	38 (51)	7 (16)	34 (597)	98 (62)
AZET	30 (24)	2 (1)	103 (529)	97 (68)

from the tangent formula procedures. For the cases in which negative triplets are not found by P_{10} it may be shown that the negative triplets are in large part subsets of the sets characterized by small positive values of G. For example, for DIOLE, of the 460 triplets having $G \le 0.4$ according to P_{10}, 246 have negative cosine.

An important practical conclusion can be drawn from Tables 2–4. The effectiveness of the method does not critically depend on the range within which the running vector **k** is allowed to vary. Thus, a short range may be chosen for practical application without relevant damage. Only a few minutes of computing

time (4–7) are needed to estimate 3000 triplets selected by P_{10} on an IBM 370/158 machine when k is allowed to vary over the largest 70 $|E|$'s.

7. Conclusions

The illustrative examples computed by (8) and (10) indicate that these formulas can provide more useful phase information than P_3. The information is of varied sort.

The triplets estimated positive by P_{10}, ranked in a new order of accuracy, define a new convergence map and can actively be used in the tangent procedures. It should also be stressed that the integration of our formulas with the random approaches of phases (Declercq, Germain & Woolfson, 1979) is very easy and can facilitate the convergence from random phases to the correct solution.

The triplets whose cosines are estimated negative by P_{10} often are not sufficiently accurate to be actively used in tangent procedures. However, they can be successfully exploited as a powerful figure of merit for finding out the correct solution in multisolution procedures (Camalli *et al.*, 1984). We stress the point that such a figure of merit is statistically independent of that using negative quartets.

References

BEURSKENS, T., BEURSKENS, G. & VAN DEN HARK, TH. E. M. (1976). *Cryst. Struct. Commun.* 5, 241–246.
CAMALLI, M., CASCARANO, G., GIACOVAZZO, C., SPAGNA, R. & VITERBO, D. (1984). In preparation.
COCHRAN, W. (1955). *Acta Cryst.* 8, 473–478.
COLENS, A., DECLERCQ, J. P., GERMAIN, G., PUTZEYS, J. P. & VAN MEERSSCHE, M. (1974). *Cryst. Struct. Commun.* 3, 119–122.
DECLERCQ, J. P., GERMAIN, G. & WOOLFSON, M. M. (1979). *Acta Cryst.* A35, 622–626.
GIACOVAZZO, C. (1977a). *Acta Cryst.* A33, 527–531.
GIACOVAZZO, C. (1977b). *Acta Cryst.* A33, 933–944.
GIACOVAZZO, C. (1980). *Direct Methods in Crystallography.* London: Academic Press.
GOLDBERG, I. (1975). *Acta Cryst.* B31, 2592–2600.
GOLDBERG, I. & RESMOVITZ, H. (1978). *Acta Cryst.* B34, 2894–2896.
HAUPTMAN, H. (1975). *Acta Cryst.* A31, 680–687.
HAUPTMAN, H. (1977). Am. Crystallogr. Assoc. Michigan State Meet., 7–12 August 1977. Abstract H3.
HULL, S. E., LEBAN, I., MAIN, P., WHITE, P. S. & WOOLFSON, M. M. (1976). *Acta Cryst.* B32, 2374–2381.
SCHENK, H., KOPS, R. T., VAN DER PUTTEN, N. & BODE, J. (1978). *Acta Cryst.* B34, 505–507.
SHAKKED, Z. & KENNARD, O. (1977). *Acta Cryst.* B33, 516–522.
SUCK, D., MANOR, P. C. & SAENGER, W. (1976). *Acta Cryst.* B32, 1727–1737.
VITERBO, D. & WOOLFSON, M. M. (1973). *Acta Cryst.* A29, 205–208.
WALLWORK, S. C. & POWELL, H. M. (1980). *J. Chem. Soc. Perkin Trans. 2*, pp. 641–646.
WONG, R. Y. (1978). *Acta Cryst.* B34, 3482–3484.

The Two-Wave X-ray Field Calculated by Means of Integral-Equation Methods

By JOHANNES BREMER

Division of X-ray Physics and Crystallography, The Norwegian Institute of Technology, The University of Trondheim, 7034 Trondheim-NTH, Norway

Abstract

The problem of calculating the two-wave X-ray field on the basis of the Takagi–Taupin equations is discussed for the general case of curved lattice planes. A two-dimensional integral equation which incorporates the nature of the incoming radiation, the form of the crystal/vacuum boundary, and the curvature of the structure, is deduced. Analytical solutions for the symmetrical Laue case with incoming plane waves are obtained directly for perfect crystals by means of iteration. The same method permits a simple derivation of the narrow-wave Laue and Bragg cases. Modulated wave fronts are discussed, and it is shown that a cut-off in the width of an incoming plane wave leads to lateral oscillations which are superimposed on the *Pendellösung* fringes. Bragg and Laue shadow fields are obtained. The influence of a non-zero kernel is discussed and a numerical procedure for calculating wave amplitudes in curved crystals is presented.

1. Introduction

One important problem in the theory of X-ray diffraction is to describe wave propagation in general three-dimensional structures which are not crystalline in the traditional sense. It is well known that the Takagi–Taupin method (Takagi, 1969; Taupin, 1964) permits the dynamical X-ray wave-field in both perfect and slightly imperfect crystals to be calculated. A more general way of handling the same problem for statistically distributed defects has recently been put forward by Kato (1980). A variety of situations exists, however,