the average effect of these defects over the whole crystal. There will be perfect regions within the crystal, but defects such as dislocations and stacking faults do not normally dissect a crystal into discrete misoriented domains. The term 'ellipsoidal perfect domains' has been used above for convenience, but it should not be imagined that the authors believe in such objects *per se*. The expression (5) is a simple way of allowing the average perfect dimension along **N** to vary with crystal orientation.

It is not known whether the sample of Cr–Cl boracite is exceptional in its defect structure for boracites, or for crystals grown by vapour transport. The form of a needle is, however, very unusual (Schmid, 1973). It is intended to continue with measurements of the kind described here, on boracites and other materials, in combination if possible with study of the defects by X-ray topography.

We should like to thank Dr G. S. Pawley for illuminating discussions and Dr S. E. B. Gould for her help in data reduction. One of us (RJN) gratefully acknowledges the financial support of a Science Research Council Research Fellowship.

References

ARNDT, U. W. & WILLIS, B. T. M. (1966). *Single Crystal Diffractometry*. Cambridge Univ. Press.
BECKER, P. J. & COPPENS, P. (1974*a*). *Acta Cryst.* A**30**, 129–147.
BECKER, P. J. & COPPENS, P. (1974*b*). *Acta Cryst.* A**30**, 148–153.
COOPER, M. J. & ROUSE, K. D. (1970). *Acta Cryst.* A**26**, 214–223.
COPPENS, P. & HAMILTON, W. C. (1970). *Acta Cryst.* A**26**, 71–83.
ITO, T., MORIMOTO, N. & SADANAGA, R. (1951). *Acta Cryst.* **4**, 310–316.
NELMES, R. J. (1969). *Acta Cryst.* A**25**, 523–526.
NELMES, R. J. (1974). *J. Phys. C. Solid State Phys.* **7**, 3840–3854.
NELMES, R. J. & THORNLEY, F. R. (1974). *J. Phys. C. Solid State Phys.* **7**, 3855–3874.
SCHMID, H. (1965). *J. Phys. Chem. Solids* **26**, 973–988.
SCHMID, H. (1969). *Growth Cryst.* **7**, 25–52.
SCHMID, H. (1973). Private communication.
STEWART, J. M. (1972). *The X-RAY System – Version of June* 1972. Technical Report TR-192, Computer Science Center, University of Maryland.
ZACHARIASEN, W. H. (1967). *Acta Cryst.* **23**, 558–564.

Acta Cryst. (1974). A**30**, 757

Structure Factor Algebra in the Probabilistic Procedures for Phase Determination. III

BY C. GIACOVAZZO

Istituto di Mineralogia e Petrografia, Università di Bari, Italy

(*Received* 11 *March* 1974; *accepted* 19 *May* 1974)

An investigation has been carried out on the use of normalized, quasi-normalized and pseudo-normalized structure factors in the probabilistic procedures for phase assignment. A new statistical formula has been established for centrosymmetric space groups.

Introduction

Several ways of normalizing structure factors are used in the procedures for the solution of the phase problem by direct methods. In part I of this paper (Giacovazzo, 1974*a*) we have recalled the definitions of the normalized structure factor $E_{\mathbf{h}}$, the quasi-normalized structure factor $\mathscr{E}_{\mathbf{h}}$ and a pseudo-normalized structure factor $E'_{\mathbf{h}}$, this last advised by Karle & Karle (1966). $E_{\mathbf{h}}$ ensures always that the mean-square $\langle E_{\mathbf{h}}^2\rangle=1$ with the consequent simplicity in some distribution functions: the quasi-normalized structure factor $\mathscr{E}_{\mathbf{h}}$ warrants greater simplicity in the derivation of the algebraic relations.

In the automatic procedures for the calculation of crystal structure invariants (Hauptman, Fisher, Hancock & Norton, 1969), quasi-normalized structure factors are preferred: in actual symbolic-addition procedures or in the multisolution approach the use of E or E' is a personal decision. Giacovazzo (1974*a*, *b*) has shown that the statistical interactions among $E_{\mathbf{h}}, E_{\mathbf{k}}, E_{\mathbf{h}\pm\mathbf{k}}$ are not simple, but depend on the space groups and on the parity of the vectors $\mathbf{h},\mathbf{k},\mathbf{h}\pm\mathbf{k}$.

In fact, in centrosymmetric crystals the formula is

$$P_+(E_{\mathbf{h}})=\tfrac{1}{2}+\tfrac{1}{2}\tanh\left[\frac{E_{\mathbf{h}}}{N^{1/2}}\,W_{\mathbf{h},\mathbf{k}}E_{\mathbf{k}}E_{\mathbf{h}\pm\mathbf{k}}\right], \quad (1)$$

where

$$W_{\mathbf{h},\mathbf{k}}=\frac{1}{m}\,\frac{\langle\xi(\mathbf{h})\xi(\mathbf{k})\xi(\mathbf{h}\pm\mathbf{k})\rangle}{\sqrt{p_{\mathbf{h}}p_{\mathbf{k}}p_{\mathbf{h}\pm\mathbf{k}}}};$$

m is the order of the space group, ξ its trigonometric structure factor. If $E_{\mathbf{h}}$ is a non-centrosymmetric reflexion, we can write

$$P(\varphi_{\mathbf{h}})=\exp\{G_{\mathbf{h},\mathbf{k}}\cos(\varphi_{\mathbf{h}}-\varphi_{\mathbf{k}}-\varphi_{\mathbf{h}-\mathbf{k}})\}/[2\pi I_0(G_{\mathbf{h},\mathbf{k}})], \quad (2)$$

where

$$G_{\mathbf{h},\mathbf{k}}=\frac{1}{m}\,\frac{\langle\xi(-\mathbf{h})\xi(\mathbf{k})\xi(\mathbf{h}-\mathbf{k})\rangle}{\sqrt{p_{\mathbf{h}}p_{\mathbf{k}}p_{\mathbf{h}-\mathbf{k}}}}\,\frac{2}{\sqrt{N}}\,|E_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{h}-\mathbf{k}}|\,.$$

If $E_{\mathbf{h}}, E_{\mathbf{k}}, E_{\mathbf{h}-\mathbf{k}}$ are general reflexions, equation (1) reduces to the well known Cochran & Woolson (1955) formula and relation (2) to the classic Cochran (1956) formula. In general, however, the quantity $W_{\mathbf{h},\mathbf{k}}|E_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{h}\pm\mathbf{k}}|$ is equal neither to $|E_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{h}\pm\mathbf{k}}|$ nor to $|\mathscr{E}_{\mathbf{h}}\mathscr{E}_{\mathbf{k}}\mathscr{E}_{\mathbf{h}\pm\mathbf{k}}|$; in the same way $G_{\mathbf{h},\mathbf{k}}$ coincides neither with $2/\sqrt{N}\,|E_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{h}-\mathbf{k}}|$ nor with $2/\sqrt{N}\,|\mathscr{E}_{\mathbf{h}}\mathscr{E}_{\mathbf{k}}\mathscr{E}_{\mathbf{h}-\mathbf{k}}|$. This fact requires a further search for the role of E, $\mathscr{E}$ and E' in the procedures which assign phase values by the Cochran–Woolfson or tangent formulas.

A new phase relation in centrosymmetric crystals

In a recent monograph (Giacovazzo, 1974*c*) we have suggested a way of generalizing in all space groups the probability distribution $P(E_{\mathbf{k}}, E_{\mathbf{h}+\mathbf{k}})$ of the pair $E_{\mathbf{k}}, E_{\mathbf{h}+\mathbf{k}}$, derived by Hauptman & Karle (1958) for $P\bar{1}$, when $\mathbf{h}$ is fixed and $\mathbf{k}$ varies through all reciprocal space. The result, for a centrosymmetric space group of order m, is

$$\begin{aligned}P(E_{\mathbf{k}}, E_{\mathbf{h}+\mathbf{k}}) &= \frac{1}{2\pi\lambda^{1/2}}\exp[-\tfrac{1}{2}\mathbf{E}\bar{\lambda}^{-1}\mathbf{E}]\\ &\times\Big\{1+\frac{1}{t^{3/2}}\Big[\sum_1^t{}_j\,\frac{{}^j\lambda_{30}}{3!0!}H_3(E_{\mathbf{k}})\\ &+\sum_1^t{}_j\,\frac{{}^j\lambda_{03}}{0!3!}H_3(E_{\mathbf{h}+\mathbf{k}})+\sum_1^t{}_j\,\frac{{}^j\lambda_{12}}{1!2!}H_1(E_{\mathbf{k}})H_2(E_{\mathbf{h}+\mathbf{k}})\\ &+\sum_1^t{}_j\,\frac{{}^j\lambda_{21}}{2!1!}H_2(E_{\mathbf{k}})H_1(E_{\mathbf{h}+\mathbf{k}})\Big]\\ &+\frac{1}{t^2}\Big[\sum_1^t{}_j\,\frac{{}^j\lambda_{40}}{4!0!}H_4(E_{\mathbf{k}})\\ &+\sum_1^t{}_j\,\frac{{}^j\lambda_{31}}{3!1!}H_3(E_{\mathbf{k}})H_1(E_{\mathbf{h}+\mathbf{k}})\\ &+\sum_1^t{}_j\,\frac{{}^j\lambda_{22}}{2!2!}H_2(E_{\mathbf{k}})H_2(E_{\mathbf{h}+\mathbf{k}})+\dots\Big]\,,\end{aligned}\tag{3}$$

where

$$\mathbf{E}=(E_{\mathbf{k}}, E_{\mathbf{h}+\mathbf{k}}),\quad \bar{\lambda}=\begin{bmatrix}1 & \frac{1}{t}\sum_1^t{}_j\,{}^j\lambda_{11}\\ \frac{1}{t}\sum_1^t{}_j\,{}^j\lambda_{11} & 1\end{bmatrix},$$

$\bar{\lambda}^{-1}$ the inverse matrix.

Moreover, λ is the determinant of $\bar{\lambda}$, ${}^j\lambda_{rs}$ are the standardized cumulants of the distribution, t is the number of atoms in the asymmetric region of the cell H_v is the Hermite polynomial defined by the equation

$$H_v(x)=(-1)^v\exp[\tfrac{1}{2}x^2]\frac{\mathrm{d}^v}{\mathrm{d}x^v}\exp[-\tfrac{1}{2}x^2]\,.\tag{4}$$

By using linearization theory (Bertaut, 1959) we have obtained the following relations:

$$\frac{1}{t}\sum_1^t{}_j\,{}^j\lambda_{11}=\frac{1}{t}\sum_1^t{}_j\,\frac{1}{m}\langle\xi_j(\mathbf{h}+\mathbf{k})\xi_j(\mathbf{k})\rangle=\frac{1}{\sqrt{N}}\mathscr{E}_{\mathbf{h}};$$

$$\frac{1}{t^2}\sum_1^t{}_j\,\frac{{}^j\lambda_{40}}{4!0!}\geq-\frac{1}{8N};$$

$$\frac{1}{t^2}\sum_1^t{}_j\,\frac{{}^j\lambda_{31}}{3!1!}\simeq\frac{1}{6N^{3/2}}\{2\sum_1^t{}_j\,\mathscr{E}_{\mathbf{h}\mathbf{C}_q}-(2m+3)\mathscr{E}_{\mathbf{h}}\};$$

$$\frac{1}{t^2}\sum_1^t{}_j\,\frac{{}^j\lambda_{22}}{2!2!}=-\frac{1}{2N}-\frac{1}{4N^{3/2}}\mathscr{E}_{2\mathbf{h}}\dots,$$

where $\mathbf{C}_s$ is the s-symmetry operation ($\mathbf{R}_s$ rotation component, $\mathbf{T}_s$ translation component) of the space group.

If we neglect in (3) all the terms bar the first, we obtain the expression

$$\begin{aligned}P(E_{\mathbf{k}}, E_{\mathbf{h}+\mathbf{k}})&=\frac{1}{2\pi(1-\mathscr{E}_{\mathbf{h}}^2/N)^{1/2}}\\ &\times\exp\Big\{-\frac{1}{2(1-\mathscr{E}_{\mathbf{h}}^2/N)}\,(E_k^2-\frac{2}{\sqrt{N}}\mathscr{E}_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{h}+\mathbf{k}}+E_{\mathbf{h}+\mathbf{k}}^2)\Big\}\,.\end{aligned}\tag{5}$$

It is easy to derive from (5) the relation [see Tsoucaris (1970) for $P\bar{1}$]

$$P_+(E_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{h}+\mathbf{k}})=\tfrac{1}{2}+\tfrac{1}{2}\tanh\Big[\frac{1}{\sqrt{N}(1-\mathscr{E}_{\mathbf{h}}^2/N)}\,|\mathscr{E}_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{h}+\mathbf{k}}|\Big];\tag{6}$$

if we neglect in (3) terms of order $1/N$ and higher, equation (6) is valid in all the space groups. By a property of the Hermite polynomials we may write

$$\begin{aligned}\sum_0^\infty{}_v\,\frac{H_v(x)H_v(y)}{v!}t^v&=\frac{1}{\sqrt{1-t^2}}\\ &\times\exp\Big[-\frac{1}{2(1-t^2)}(t^2x^2+t^2y^2-2txy)\Big],\ (|t|<1).\end{aligned}$$

Consequently the result is

$$\begin{aligned}\sum_0^\infty{}_v\,\frac{H_v(E_{\mathbf{k}})H_v(E_{\mathbf{h}+\mathbf{k}})}{v!}\varrho^v&=\frac{1}{\sqrt{1-\varrho^2}}\\ &\times\exp\Big[-\frac{E_{\mathbf{k}}^2+E_{\mathbf{h}+\mathbf{k}}^2-2\varrho E_{\mathbf{k}}E_{\mathbf{h}+\mathbf{k}}}{2(1-\varrho^2)}\Big]\\ &\times\exp[\tfrac{1}{2}(E_{\mathbf{k}}^2+E_{\mathbf{h}+\mathbf{k}}^2)]\,,\end{aligned}\tag{7}$$

where $\varrho=\mathscr{E}_{\mathbf{h}}/N$. Equation (7) gives easily the relation

$$\begin{aligned}P(E_{\mathbf{k}}, E_{\mathbf{h}+\mathbf{k}})&=\frac{1}{2\pi}\sum_0^\infty{}_v\,\frac{1}{v!}H_v(E_{\mathbf{k}})H_v(E_{\mathbf{h}+\mathbf{k}})\\ &\times\exp[-(E_{\mathbf{k}}^2+E_{\mathbf{h}+\mathbf{k}}^2)/2]\varrho^v\,.\end{aligned}\tag{8}$$

If we take relation (4) into account, from (8) we deduce

$$P(E_{\mathbf{k}}, E_{\mathbf{h}+\mathbf{k}}) = \sum_{0}^{\infty}{}_{\nu} \frac{1}{\nu!} \Phi^{(\nu+1)}(E_{\mathbf{k}}) \Phi^{(\nu+1)}(E_{\mathbf{h}+\mathbf{k}}) \varrho^{\nu}, \quad (9)$$

where

$$\Phi^{(\nu+1)}(E) = \frac{\mathrm{d}^{\nu+1}}{\mathrm{d}E^{\nu+1}} \Phi(E)$$

and

$$\Phi(E) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{E} \exp(-t^2/2)\mathrm{d}t .$$

Let us integrate (9) term by term: we obtain

$$\int_{-\infty}^{E_1}\int_{-\infty}^{E_2} P(E_{\mathbf{k}}, E_{\mathbf{h}+\mathbf{k}})\mathrm{d}E_{\mathbf{k}}\mathrm{d}E_{\mathbf{h}+\mathbf{k}} = \sum_{0}^{\infty}{}_{\nu} \frac{\Phi^{(\nu)}(E_1)\Phi^{(\nu)}(E_2)}{\nu!} \varrho^{\nu}, \quad (10)$$

which becomes, when $E_1 = E_2 = 0$,

$$\int_{-\infty}^{0}\int_{-\infty}^{0} P(E_{\mathbf{k}}, E_{\mathbf{h}+\mathbf{k}})\mathrm{d}E_{\mathbf{k}}\mathrm{d}E_{\mathbf{h}+\mathbf{k}} = \sum_{0}^{\infty}{}_{\nu} \frac{[\Phi^{(\nu)}(0)]^2}{\nu!} \varrho^{\nu}. \quad (11)$$

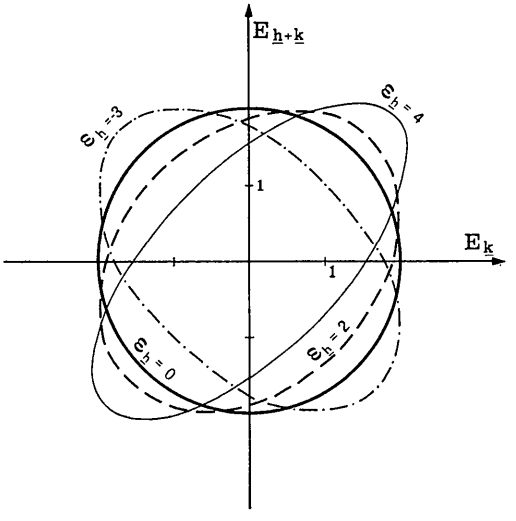

Fig. 1. The curves correspond to $N=40$.

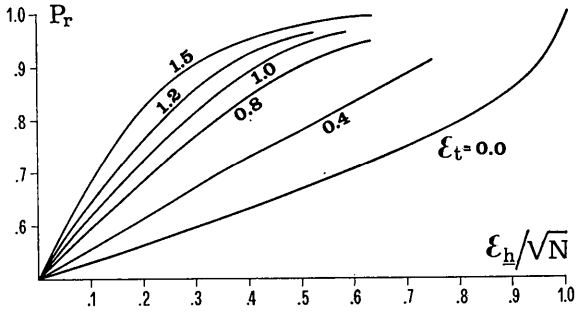

Fig. 2. Percentage of positive products $E_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{h}+\mathbf{k}}$ when a threshold value $\mathscr{E}_t$ is fixed.

The evaluation of the right-hand side of (11) may be made by comparing (9) to (5) when $E_{\mathbf{k}} = E_{\mathbf{h}+\mathbf{k}} = 0$: we obtain

$$\sum_{0}^{\infty}{}_{\nu} \frac{[\Phi^{(\nu+1)}(0)]^2}{\nu!} \varrho^{\nu} = \frac{1}{2\pi\sqrt{1-\varrho^2}} . \quad (12)$$

By integrating (12) in respect to ϱ we deduce

$$\int_0^{\varrho} \sum_{0}^{\infty}{}_{\nu} \frac{[\Phi^{(\nu+1)}(0)]^2}{\nu!} \varrho^{\nu}\mathrm{d}\varrho = \sum_{1}^{\infty}{}_{\nu} \frac{[\Phi^{(\nu)}(0)]^2}{\nu!} \varrho^{\nu} = \frac{1}{2\pi} \arcsin \varrho .$$

We may write finally equation (11) in the form

$$\int_{-\infty}^{0}\int_{-\infty}^{0} P(E_{\mathbf{k}}, E_{\mathbf{h}+\mathbf{k}})\mathrm{d}E_{\mathbf{k}}\mathrm{d}E_{\mathbf{h}+\mathbf{k}} = \tfrac{1}{4} + \frac{1}{2\pi} \arcsin\left(\frac{\mathscr{E}_{\mathbf{h}}}{\sqrt{N}}\right) . \quad (13)$$

The left-hand side of (13) represents the population of the normalized structure factors which lie in the third quadrant of the plane $(E_{\mathbf{k}}, E_{\mathbf{h}+\mathbf{k}})$. By the symmetry properties of equation (5) the left-hand side of (13) also represents the population contained in the first quadrant, while each of the second and fourth quadrants contains $\frac{1}{4} - 1/2\pi \arcsin(\mathscr{E}_{\mathbf{h}}/\sqrt{N})$. In other words, if $\mathscr{E}_{\mathbf{h}}$ is positive, the population of the first and third quadrants is larger than $\frac{1}{2}$: if $\mathscr{E}_{\mathbf{h}}$ is negative the second and fourth quadrants are more crowded than other two.

Relation (13) may be visualized by introducing the ellipse of concentration of the distribution (5):

$$E_{\mathbf{k}}^2 - \frac{2}{\sqrt{N}} \mathscr{E}_{\mathbf{h}} E_{\mathbf{k}} E_{\mathbf{h}+\mathbf{k}} + E_{\mathbf{h}+\mathbf{k}}^2 = 4(1 - \mathscr{E}_{\mathbf{h}}^2/N) .$$

As is well known, a uniform distribution over the area enclosed by this ellipse has the same first and second-order moments as the distribution (5). In Fig. 1 some ellipses of concentration are shown: the didactic value of the figure is related to its simplicity.

We can conclude that the percentage of positive products $E_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{h}+\mathbf{k}}$, when $\mathbf{k}$ varies over all reciprocal space, is equal to $\frac{1}{2} + 1/\pi \arcsin(\mathscr{E}_{\mathbf{h}}/\sqrt{N})$. If $\mathscr{E}_{\mathbf{h}} = 0$, the percentage is equal to $\frac{1}{2}$: if $\mathscr{E}_{\mathbf{h}} = \sqrt{N}$, all the products $E_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{h}+\mathbf{k}}$ are positive.

This last result is in contrast with the Cochran–Woolfson (1955) formula

$$P_+(E_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{h}+\mathbf{k}}) = \tfrac{1}{2} + \tfrac{1}{2} \tanh[|E_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{h}+\mathbf{k}}|/\sqrt{N}] . \quad (14)$$

Klug (1958), however, showed that equation (14) gives an understimate of the probability when the structure factors involved are large. Our result seems more accurate and agrees well with that of Tsoucaris (1970).

A statistical meaning of the $\mathscr{E}$ threshold

In any method of direct phase determination reliable phase assignments are generally associated with the

larger $\mathscr{E}$ values. By the use of the multiple-sign relations, however, the probability formulae permit assignment of reliable phase values to reflexions with small $\mathscr{E}$ also [see, for an extreme case, Dewar (1970)].

Fourier synthesis resulting from the larger-$\mathscr{E}$ coefficients will be affected by an amplitude-termination effect (Bürgi & Dunitz, 1971), particularly severe for structures containing some regularities. On the other hand, it may be advantageous to break off the procedure for the phase assignment at an early stage because of the very rapid propagation of the errors when incorrect signs have been assigned [see, for $P\bar{1}$, Mo, Hjortas & Svinning (1973)].

Hence, in the automatic procedures for phase determination, a limited number of phases is assigned by fixing an $\mathscr{E}$ threshold ($\mathscr{E}_t = 0{\cdot}8$–$1{\cdot}4$). However, it may be difficult to judge *a priori* the worthiness of a given threshold if no absolute figure of merit based on statistical criteria is known. Equation (9) permits us to state a relation between $\mathscr{E}_t$ and the percentage of positive products $E_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{h}+\mathbf{k}}$ that we must expect.

In fact, we deduce from (9) and (10) that the percentage of positive products, when $\mathscr{E}_{\mathbf{h}}$ is fixed and $\mathbf{k}$ varies with all the factors $|E_{\mathbf{k}}| > \mathscr{E}_t$, $|E_{\mathbf{h}+\mathbf{k}}| > \mathscr{E}_t$, is equal to

$$P_r = \left\{\sum_0^\infty{}_\nu \frac{[\Phi^{(\nu)}(-\mathscr{E}_t)]^2}{\nu!}|\varrho|^\nu\right\} \Big/ \left\{\sum_0^\infty{}_\nu \frac{\Phi^{(\nu)}(-\mathscr{E}_t)]^2}{\nu!}|\varrho|^\nu + \sum_0^\infty{}_\nu \frac{[\Phi^{(\nu)}(-\mathscr{E}_t)]^2}{\nu!}(-|\varrho|)^\nu\right\}. \quad (15)$$

Owing to the fast convergence of the series involved in (15), P_r is easily computable. In Fig. 2 we have plotted some curves, each corresponding to a single value of $\mathscr{E}_t$. So, for a structure with $N=100$, the expected percentage of positive products $\mathscr{E}_{\mathbf{h}}\mathscr{E}_{\mathbf{k}}\mathscr{E}_{\mathbf{h}+\mathbf{k}}$, when $\mathscr{E}_{\mathbf{h}} = 1{\cdot}5$ and $\mathscr{E}_t = 1{\cdot}5$, is 0·76: this percentage comes down to 0·70 when $\mathscr{E}_t = 1{\cdot}2$ and to 0·67 when $\mathscr{E}_t = 1$.

In conclusion, any choice of $\mathscr{E}_t$ corresponds to an expected percentage of positive products $E_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{h}+\mathbf{k}}$; for example, if we fix $\mathscr{E}_t = 1{\cdot}2$ for a structure with $N=50$, we must expect, for $\mathscr{E}_{\mathbf{h}} = 1{\cdot}40$, a value $P_r = 0{\cdot}64$.

A suitable choice of $\mathscr{E}_t$, therefore, must take N into account; so it seems reasonable to choose fairly large $\mathscr{E}_t$ for large structures, but for small structures we can lower the threshold. The occurrence of pseudo-symmetry or hyper-symmetry in the crystal structure, a reduced extension of the diffraction data, or accidental or systematic errors in the $|\mathscr{E}|$ values, can cause a lack of agreement between theory and practice.

A very important role in defining a 'good' value of $\mathscr{E}_t$ is played by the order of symmetry of the actual space group. In fact, in space groups of high symmetry multiple-sign relations are frequent from the early stages of the phase-assignment process. This condition, as is well known, makes the crystal structure solution much easier.

In space groups of low symmetry multiple-sign indications are comparatively less frequent and the error propagation can be more rapid: higher threshold values $\mathscr{E}_t$ may be suitable in these cases. These considerations explain the results of Mo, Hjortas & Svinning (1973) for $P\bar{1}$ group.

The non-centrosymmetric case

By analysis slightly modified in comparison with that suggested by Hauptman (1972) we obtain the relation

$$P(\varphi_{\mathbf{k}},\varphi_{\mathbf{h}+\mathbf{k}}|A) = \frac{1}{4\pi^2 I_0(A)} \times \exp\{A\cos(\varphi_{\mathbf{h}}+\varphi_{\mathbf{k}}-\varphi_{\mathbf{h}+\mathbf{k}})\}, \quad (16)$$

where

$$A = \frac{2|\mathscr{E}_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{h}+\mathbf{k}}|}{\sqrt{N(1-\mathscr{E}_{\mathbf{h}}^2/N)}}.$$

We emphasize once more that $\mathbf{k}$ and $\mathbf{h}+\mathbf{k}$ are general reflexions whereas $\mathbf{h}$ may be a general or a special reflexion. In accordance with the results obtained in part II (Giacovazzo, 1974*b*) of this series, the variance of the distribution (16) is bound to $\mathscr{E}_{\mathbf{h}}$ and not to $E_{\mathbf{h}}$.

Conclusions

In this paper we have considered a special type of distribution $P(E_{\mathbf{h}}, E_{\mathbf{k}}, E_{\mathbf{h}+\mathbf{k}})$, which corresponds to a random variable $\mathbf{k}$ when $\mathbf{h}$ is fixed. $E_{\mathbf{k}}$ and $E_{\mathbf{h}+\mathbf{k}}$, therefore, play the role of general reflexions ($\mathscr{E}_{\mathbf{k}} = E_{\mathbf{k}}, \mathscr{E}_{\mathbf{h}+\mathbf{k}} = E_{\mathbf{h}+\mathbf{k}}$), while $E_{\mathbf{h}}$ may be a general or a special reflexion. Under these conditions equation (13) tells us that the percentage of the positive products $E_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{h}+\mathbf{k}}$ is bound, when $\mathbf{k}$ varies over all the reciprocal space, to the quasi-normalized structure factor $\mathscr{E}_{\mathbf{h}}$ and not to $E_{\mathbf{h}}$. This is not a surprising result, after the considerations made in parts I and II of this series.

From a more general point of view, however, in the probabilistic procedures for the phase assignment, one should not use $\mathscr{E}$ factors instead of E.

In fact, a measure of the reliability of the phase relation $\varphi_{\mathbf{h}} \simeq \varphi_{\mathbf{k}} + \varphi_{\mathbf{h}-\mathbf{k}}$ depends, in the non-centrosymmetric as well as in centrosymmetric space group, on the quantities

$$\frac{1}{m}\frac{\langle\xi(-\mathbf{h})\xi(\mathbf{k})\xi(\mathbf{h}-\mathbf{k})\rangle}{\sqrt{p_{\mathbf{h}}p_{\mathbf{k}}p_{\mathbf{h}-\mathbf{k}}}}|E_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{h}+\mathbf{k}}|. \quad (16)$$

If we limit ourself to considering merely the product $\mathscr{E}_{\mathbf{h}}\mathscr{E}_{\mathbf{k}}\mathscr{E}_{\mathbf{h}+\mathbf{k}}$, we can overestimate the reliability of the phase relationship; consideration of the product $E_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{h}+\mathbf{k}}$ alone may lead us to an underestimate.

However, if one does not wish to spend calculation time in the evaluation of the quantity (16), the use of E factors seems more suitable, according to the principles usually adopted for proper weighting.

No theoretical justification has been made, on the other hand, for the use of the pseudo-nolrmaized structure factors E' proposed by Karle & Karle (1966).

References

BERTAUT, E. F. (1959). *Acta Cryst.* **12**, 541–549.
BÜRGI, H. B. & DUNITZ, J. D. (1971). *Acta Cryst.* A**27**, 117–119.
COCHRAN, W. (1956). *Acta Cryst.* **8**, 473–478.
COCHRAN, W. & WOOLFSON, M. M. (1955). *Acta Cryst.* **8**, 1–12.
DEWAR, R. B. K. (1970). In *Crystallographic Computing*, edited by F. R. AHMED. Copenhagen: Munksgaard.
GIACOVAZZO, C. (1974*a*). *Acta Cryst.* A**30**, 626–630.
GIACOVAZZO, C. (1974*b*). *Acta Cryst.* A**30**, 631–634.
GIACOVAZZO, C. (1974*c*). *I Metodi Diretti Nella Soluzione delle Strutture Cristalline.* Milano: Tamburini.
HAUPTMAN, H. (1972). *Z. Kristallogr.* **135**, 1–17.
HAUPTMAN, H., FISHER, J., HANCOCK, H. & NORTON, D. A. (1969). *Acta Cryst.* B**25**, 811–814.
HAUPTMAN, H. & KARLE, J. (1958). *Acta Cryst.* **11**, 149–157.
KARLE, J. & KARLE, I. L. (1966). *Acta Cryst.* **21**, 849–859.
KLUG, A. (1958). *Acta Cryst.* **11**, 515–543.
MO, F., HJORTAS, J. & SVINNING, T. (1973). *Acta Cryst.* A**29**, 358–362.
TSOUCARIS, G. (1970). *Acta Cryst.* A**26**, 492–501.

Acta Cryst. (1974). A**30**, 761

The Phase Problem and its Implications in the Least-squares Refinement of Crystal Structures

BY A. D. RAE

Department of Physical Chemistry, The University of New South Wales, Kensington, N.S.W. 2033, *Australia*

(*Received* 16 *April* 1974; *accepted* 6 *May* 1974)

It is found that covariance problems encountered in pseudosymmetric crystal structures are caused by an incorrect use of the least-squares refinement procedure. Rederivation of the least-squares equations for the situation in which the residual need not have the phase angle α or $\alpha+\pi$, where α is the phase angle associated with a trial structure, reveals that the minimization of the component at $\pi/2$ to α has been omitted from the least-squares equations. Inclusion of the extra terms associated with this minimization reveals that it should now be possible to refine a centrosymmetric crystal in an non-centrosymmetric space group. It is also shown that the use of weights derived from counting statistics alone is incorrect and, with a correct weighting scheme, $\sum w\Delta^2/(n-m)$ should reduce to one in a single cycle. The weighting scheme is re-evaluated for further refinement cycles.

Introduction

For the least-squares refinement of single-crystal structures from X-ray diffraction data it has been customary (*International Tables for X-ray Crystallography*, 1959) to minimize the sum of the n weighted squared residuals $S_1=\sum_h w_h(|F_o|_h-|F_c|_h)^2$, where $w_h^{-1}=\text{var}\,(|F_o|_h)$, by solving the equations

$$\sum_j \sum_h w_h \left(\frac{\partial \Delta}{\partial u_i}\right)_{0h} \left(\frac{\partial \Delta}{\partial u_j}\right)_{0h} [u_j-(u_j)_0] = -\sum_h w_h \Delta_{0h} \left(\frac{\partial \Delta}{\partial u_i}\right)_{0h}; \quad i=1 \text{ to } m,$$

where $\Delta=|F_o|-|F_c|$, the difference between the measured and calculated amplitudes. Throughout this paper the subscript h implies the hth observable and the subscript 0 implies evaluation with parameters $(u_j)_0$ of a trial structure. Now for $\Delta=|F_o|-|F_c|$ and $\tan\alpha_0=(B_c/A_c)_0$ we use $(\partial\Delta/\partial u_i)_{0h}=-(\partial|F_c|/\partial u_i)_{0h}=-[\cos\alpha_0(\partial A_c/\partial u_i)_0+\sin\alpha_0(\partial B_c/\partial u_i)_0]_h$.

The application of this procedure reveals two apparent faults. Firstly the assumption that $w_h^{-1}=\text{var}\,(|F_o|_h)$ because $(F_c)_h$ is without error does not produce the expected result that $\sum_h w_h\Delta_h^2=n-m$ in a single cycle. Secondly, in pseudosymmetric structures the apparent variances of parameters u_j are usually in excess of calculated variances (Rae, 1973). This suggests that both the weighting scheme and the actual least-squares equations are at fault. Investigation shows that this hypothesis is indeed true and that the situation may be remedied.

Theory

The least-squares equations

Account should be taken of the fact that we are dealing with quantities $F=A+iB$ that do not have a fixed phase. If $(F_o)_0$ is an initial estimate of the phased quantity F_o for which only the magnitude $|F_o|$ has been experimentally determined, then it is found that a different set of least-squares equations are generated by