CASCARANO, G., GIACOVAZZO, C. & LUIĆ, M. (1985*a*). In *Statistical Crystallography*. New York: Adenine Press.
CASCARANO, G., GIACOVAZZO, C. & LUIĆ, M. (1985*b*). *Acta Cryst.* A**41**, 544-551.
CHRISTIDIS, P. C. & RENTZEPERIS, R. J. (1975). *Z. Kristallogr.* **141**, 233-245.
COCHRAN, W. (1955). *Acta Cryst.* **8**, 473-478.
FAN HAI-FU, YAO JIA-XING, MAIN, P. & WOOLFSON, M. M. (1983). *Acta Cryst.* A**39**, 566-569.
GIACOVAZZO, C. (1980). *Direct Methods in Crystallography*. London: Academic Press.
GRAMLICH, V. (1975). *Acta Cryst.* A**31**, S90.
GRAMLICH, V. (1984). *Acta Cryst.* A**40**, 610-616.
HAUPTMAN, H. & KARLE, J. (1953). *Solution of the Phase Problem. I. The Centrosymmetric Crystal. Am. Crystallogr. Soc. Monogr.* No. 3. New York: Polycrystal Book Service.
HAUPTMAN, H. & KARLE, J. (1959). *Acta Cryst.* **12**, 846-850.
ITO, T. & NOWACKI, W. (1974). *Z. Kristallogr.* **139**, 85-102.
JEFFERY, J. W. (1964). *Acta Cryst.* **17**, 776-777.
LIPSON, H. & WOOLFSON, M. M. (1952). *Acta Cryst.* **5**, 680-682.
PONTENAGEL, W. M. G. F. & KRABBENDAM, H. (1983). *Acta Cryst.* A**39**, 333-340.
PRICK, P. A. J., BEURSKENS, P. I. & GOULD, R. O. (1983). *Acta Cryst.* A**39**, 570-576.
ROGERS, D. & WILSON, A. J. C. (1953). *Acta Cryst.* **6**, 439-449.
SAYRE, D. (1952). *Acta Cryst.* **5**, 60-65.

Acta Cryst. (1987). A**43**, 22-29

Figures of Merit in Direct Methods: a New Point of View

BY G. CASCARANO AND C. GIACOVAZZO

Dipartimento Geomineralogico, Università, Campus Universitario, Via Salvemini, 70124 *Bari, Italy*

AND D. VITERBO

Istituto di Chimica Fisica, Università, Corso M. D'Azeglio 48, 10125 *Torino, Italy*

(*Received* 16 *December* 1985; *accepted* 12 *May* 1986)

Abstract

The probabilistic approach has been extensively used for analysing the statistical meanings of some traditional figures of merit. New figures of merit have also been introduced; some exploit one-phase and two-phase structure seminvariants of the first rank, the expected negative and/or enantiomorph-sensitive triplet invariants, and the expected negative and/or enantiomorph-sensitive quartets. Other figures of merit exploit the distribution of statistical parameters connected with PSI(0) and active triplets. A new combined figure of merit is shown to be a powerful tool for selecting the correct solutions among the various sets output by multisolution methods.

1. Introduction

For the determination of very complex structures by direct methods a large initial set of known phases seems to be a basic requirement. This aim can be achieved by introducing a large number of permutable phases which are used to generate different phase sets. Magic integer sequences (Main, 1977) are often employed for phase permutations: a relatively large number of phase sets are thereby created among which the correct solutions have to be found. Figures of merit (FOM) are usually used to screen the set of solutions, prior to computing Fourier transforms (*E* map).

In general, FOM's are functions based on quantities which can be expected to have extreme values for the correct solution. The expectation relies on a probabilistic background and/or on algebraic properties. In the *SIR* program (Nunzi *et al.*, 1984), several low-order structure seminvariants and invariants are estimated by means of representation theory (Giacovazzo, 1977, 1980*a*). Some of them are actively used for phase expansion and refinement, others are only employed to compute the FOM's. Since one or more FOM's are available for each type of structure seminvariant or invariant, the combined figure of merit CPHASE, based on a variety of FOM's, is expected to be effective in finding the correct solution [see Hašek, Schenk, Kiers & Schagen (1985) for some tests of distribution-fitting methods for centrosymmetric structures].

A probabilistic approach is also introduced which enables us to analyse the statistical meanings of some traditionally widely used FOM's. New effective figures of merit are devised which, combined with CPHASE, give rise to a reliable total combined figure of merit CFOM which is expected to be unity for the correct solutions.

2. The combined figure of merit CPHASE

Overbeek & Schenk (1976) first proposed a FOM based on $\sum_1$ relationships. In the *SIR* program the estimates of the one-phase structure seminvariants of

0108-7673/87/010022-08$01.50

first rank *via* their second representation are improved *via* the estimates of a special class of two-phase seminvariants (Burla, Nunzi, Giacovazzo & Polidori, 1981). The phase value φ_H of any one-phase seminvariant is assumed to be distributed according to the von Mises distribution (Cascarano, Giacovazzo, Calabrese, Burla, Nunzi, Polidori & Viterbo, 1984):

$$M(\varphi_1;\theta_1,G_1)\simeq[2\pi I_0(G_1)]^{-1}\exp[G_1\cos(\varphi_1-\theta_1)].$$

where θ_1 is the expected value of φ_1. The structure seminvariants with the largest values of G_1 are actively used in the phase expansion process, while all the seminvariants cooperate to give the following figure of merit:

$$\text{SS1FOM}=\frac{\sum w_1G_1\cos(\varphi_1-\theta_1)}{\sum w_1G_1D_1(G_1)}=\frac{T_1}{B_1}$$

where $D_1(G)=I_1(G)/I_0(G)$ and I_i is the modified Bessel function of order i. The denominator of SS1FOM is the expected value of the numerator; $w_1=0{\cdot}5$ for seminvariant phases actively used, $w_1=1$ for the others. In practice only structure seminvariants with restricted phase values are used; for them $G_1D_1(G_1)$ is replaced by $(G_1/2)\tanh(G_1/2)$ (the same will occur for other invariants or seminvariants with restricted phase values which will later be used in the FOM's).

Two-phase structure seminvariants of first rank are estimated *via* their first representation according to von Mises-like distributions. The most reliable of them are actively used in the phasing procedure, while all of them are used in the figure of merit

$$\text{SS2FOM}=\frac{\sum w_2G_2\cos(\Phi_2-\theta_2)}{\sum w_2G_2D_1(G_2)}=\frac{T_2}{B_2},$$

where θ_2 is the expected value of the structure seminvariant and $\Phi_2=\varphi_{H_1}+\varphi_{H_2}$. $w_2=0{\cdot}5$ for seminvariants actively used, $w_2=1$ for the others.

Since G_1 and G_2 are always positive, SS1FOM and SS2FOM are expected to be positive and unitary for the correct solution.

On request *SIR* estimates triplet phase invariants *via* their second representation [*i.e.* by equation (10) of the paper by Cascarano, Giacovazzo, Camalli, Spagna, Burla, Nunzi & Polidori (1984)]. The estimate of a single triplet involves the use of a large set of diffraction magnitudes whose contribution often allows the identification of negative cosines. The most reliable triplets which are estimated positive, ranked in order of accuracy, define a new convergence map and are actively used in tangent procedures. Triplets whose cosines are estimated negative usually have reliability too low to be actively used. In the *SIR* program they cooperate to define the figure of merit

$$\text{NTREST}=\sum G_3\cos\Phi_3/\sum G_3D_1(G_3)=T_3/B_3$$

where G_3 is positive or negative according to whether $\cos\Phi_3$ is estimated positive or negative respectively. Thus NTREST is expected positive and unitary for the correct solution.

In a direct procedure it is often difficult to define and maintain the enantiomorph and in many cases one ends up with centrosymmetric solutions. Therefore enantiomorph-sensitive figures of merit may play an important role in identifying the correct solution (van der Putten & Schenk, 1979; Pontenagel, 1984). The formula estimating triplets *via* their second representation provides reliable estimates only when $|G_3|$ is large. Therefore, it is able in principle to estimate with high reliability only triplets with phases around 0 or π. In accordance with the theory, enantiomorph-sensitive triplets should present a rather flat distribution ($G_3\simeq 0$), so that they cannot be reliably fixed. On the other hand, they are expected to lie preferably around $\pm\pi/2$ because of the fact that triplets near to 0 or π have been screened by the corresponding large values of $|G_3|$. The above considerations legitimize the use of the following enantiomorph-specific figure of merit:

$$\text{FENTRS}=\sum\sin(|\Phi_3'|)/n_3=T_3'/n_3=\langle\sin|\Phi_3'|\rangle$$

where n_3 is the number of triplets characterized by $|G_3|$ values between 0 and 0·2.

A further effective way of using information contained in weak reflexions is the NQEST figure of merit first proposed by Schenk (1974) (see also De Titta, Edmonds, Langs & Hauptman, 1975):

$$\text{NQEST}=\frac{\sum E_hE_kE_lE_m\cos(\varphi_h+\varphi_k+\varphi_l+\varphi_m)}{\sum|E_hE_kE_lE_m|}$$

based on selected 'negative quartets' with $\mathbf{h}+\mathbf{k}+\mathbf{l}+\mathbf{m}=0$, $|E_h|,|E_k|,|E_l|,|E_m|$ all large and the three unique cross terms $|E_{h+k}|$, $|E_{h+l}|$, $|E_{k+l}|$ all small. In *SIR* (Busetta, Giacovazzo, Burla, Nunzi, Polidori & Viterbo, 1980), negative quartets are estimated *via* their complete first representation, which, in favourable cases, contains more than seven magnitudes. The reliability of each quartet is estimated according to a von Mises distribution (see Giacovazzo, 1980*b*, § 8.3.5) whose parameter G_4 depends on all the magnitudes contained in the first representation. The above considerations suggest the following FOM:

$$\text{NQEST}=\sum G_4\cos\Phi_4/\sum G_4D_1(G_4)=T_4/B_4.$$

G_4 is negative if $\cos\Phi_4$ is estimated negative. Thus NQEST is expected to be unitary for the correct solution.

In our experience G_4 proved to be a useful reliability parameter. Its use requires more computing time than that of a mere threshold on the magnitudes of the cross terms, but it often leads to the identification of a more discriminating subset of quartets. Up to 500 quartets (with largest values of $|G|$) are chosen

Table 1. *The values of* $\langle\Phi\rangle = \sin^{-1}(ENANT)$ *for the 50 solutions output by SIR for CYCLO*

Correct solutions in boldface type.

Set	$\langle\Phi\rangle$	Set	$\langle\Phi\rangle$	Set	$\langle\Phi\rangle$	Set	$\langle\Phi\rangle$	Set	$\langle\Phi\rangle$
1	24	**11**	**40**	21	25	31	33	**41**	**40**
2	27	12	27	22	30	32	31	42	27
3	31	13	28	23	31	33	25	43	25
4	29	14	24	24	27	34	25	44	27
5	37	15	29	25	25	35	33	45	27
6	33	16	25	26	24	36	25	46	31
7	28	17	32	27	24	37	26	47	29
8	29	18	20	28	27	38	32	48	30
9	28	19	25	**29**	**40**	39	32	49	25
10	27	20	20	30	29	40	27	50	27

for NQEST. They include among others: (*a*) quartets with systematically absent cross reflexions; (*b*) quartets with only two cross terms in measurements. In this case $0{\cdot}85G$ is the parameter used in subsequent calculations.

It should be stressed that NQEST and NTREST use different subsets of data; therefore they have to be considered independent. Indeed, each quartet in NQEST involves the pair $\varphi_{\mathbf{h}}+\varphi_{\mathbf{k}}-\varphi_{\mathbf{h}+\mathbf{k}}$ and $\varphi_{\mathbf{l}}+\varphi_{\mathbf{m}}+\varphi_{\mathbf{h}+\mathbf{k}}$ having $|E_{\mathbf{h}+\mathbf{k}}|$ small. On the contrary, NTREST exploits triplets estimated negative in spite of the large $|E_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{h}+\mathbf{k}}|$ values.

The formula estimating quartets *via* their first representation provides reliable estimates for Φ_4 only if $|G_4|$ is large. Therefore, according to the theory, quartets having $|G_4| \simeq 0$ cannot be reliably estimated. On the other hand, they are expected to lie in preference in the regions around $\pm\pi/2$ because quartets lying around 0 or π are expected to be marked by large values of $|G_4|$. Then the following enantiomorph-sensitive FOM arises:

$$\text{FENQUS} = \sum \sin(|\Phi_4'|)/n_4 = T_4'/n_4 = \langle \sin|\Phi_4'|\rangle$$

where the average is taken over the n_4 quartets characterized by $|G_4|$ values between 0 and 0·2 [see van der Putten & Schenk (1979) for other enantiomorph-sensitive figures of merit].

Since (Schenk, 1972) the enantiomorph is easily lost in some space groups (*e.g.* $P1$, $P2_1, \ldots$) and easily maintained in others (*e.g.* $P2_12_12_1, \ldots$), FENTRS and FENQUS are not included in the set of FOM's and are only used as a warning for the user. Solutions for which

$$\text{ENANT} = (T_3' + T_4')/(n_3 + n_4) < \sin 35°$$

are explicitly marked by the program. A numerical example is shown in Table 1, where the value of ENANT for CYCLO ($C_{14}H_{20}N_2O_4$; $P2_1$, $Z=4$; Cerrini, Gavuzzo, Fedeli, Lucente, Pinnen & Zanotti, 1986) is given for the 50 sets output by *SIR*. The second representation formula of triplets played a central role for the solution of this *MULTAN*-resistant crystal structure. The correct sets (bold in Table 1) are marked by the largest values of ENANT; all wrong solutions show a more or less pronounced loss of enantiomorph.

Table 2. *The values of* $\langle\Phi\rangle = \sin^{-1}(ENANT)$ *for the 14 solutions output by SIR for AZET*

p.l.e. stands for 'possible loss of enantiomorph'.

Set	$\langle\Phi\rangle$		Set	$\langle\Phi\rangle$	
1	24	p.l.e.	8	23	p.l.e.
2	21	p.l.e.	9	23	p.l.e.
3	23	p.l.e.	10	29	p.l.e.
4	18	p.l.e.	11	26	p.l.e.
5	23	p.l.e.	12	16	p.l.e.
6	25	p.l.e.	13	18	p.l.e.
7	23	p.l.e.	14	26	p.l.e.

A complementary example is shown in Table 2: ENANT is calculated for 14 sets output by *SIR* for AZET ($C_{21}H_{16}ClNO$; $Pca2_1$, $Z=8$; Colens, Declercq, Germain, Putzeys & Van Meerssche, 1974).

Owing to the well known tendency of this crystal structure to lose the enantiomorph, no satisfactory solution may be expected among the 14 sets. Accordingly the program warns the user that the enantiomorph is probably lost (ENANT calculated from published phases is >35°).

In MULTAN-like programs the combined figure of merit is calculated by means of expressions

$$\sum w\frac{F - F_{\min}}{F_{\max} - F_{\min}} + \sum w\frac{f_{\max} - f}{f_{\max} - f_{\min}} \qquad (1)$$

where the F function symbolizes a FOM which is expected to be maximum for the correct solution, and the f function is used for FOM's which are expected to be minima. w_i are weights which reflect the *a priori* confidence of the user in the various FOM's.

Such a scheme, which works quite well for some traditional figures of merit, is unsuitable for our FOM's. To give an example, in some structures we may estimate very few unreliable one-phase seminvariants and a relatively large number of reliable negative triplets. In other structures the inverse situation may occur. In this view the *a priori* choice of w_i appears to be rather arbitrary. As a further example, suppose that the dispersion of the values of NQEST for the various solutions is large but NQEST is never positive. The influence of NQEST on the combined figure of merit is large but, probabilistically speaking, no solution appears to be reliable. In other words, the dispersion of the values of a given FOM is not an absolute measure of reliability.

Since the various phase relations which concur to define a given FOM are evaluated singly, the effectiveness of each FOM depends on the number of phase relations actually used and on their reliabilities. That suggests the combined figure of merit:

$$\text{CPHASE} = (T_1 + T_2 + T_3 + T_4)/(B_1 + B_2 + B_3 + B_4). \qquad (2)$$

The value of CPHASE for the correct solution is expected to be unity.

It should be stressed that usually NTREST and NQEST are dominant components of CPHASE since the number of available reliable negative triplet and quartet invariants is often larger than the corresponding number of the one- and two-phase seminvariants in SS1FOM and SS2FOM.

3. A revision of some traditional figures of merit

The most widely used FOM's (Declercq, Germain & Woolfson, 1979; Karle & Karle, 1966; Cochran & Douglas, 1957) are:

$$\text{ABSFOM}=\frac{\sum_{\mathbf{h}}\alpha_{\mathbf{h}}-\sum_{\mathbf{h}}(\alpha_{\mathbf{h}})_r}{\sum_{\mathbf{h}}\langle\alpha_{\mathbf{h}}\rangle-\sum_{\mathbf{h}}(\alpha_{\mathbf{h}})_r};$$

which can often be reduced to

$$\text{MABS}=\sum_{\mathbf{h}}\alpha_{\mathbf{h}}\Big/\left(\sum_{\mathbf{h}}\langle\alpha_{\mathbf{h}}\rangle\right); \tag{3}$$

$$\text{RALPHA}=\sum_{\mathbf{h}}|\alpha_{\mathbf{h}}-\langle\alpha_{\mathbf{h}}\rangle|\Big/\sum_{\mathbf{h}}\langle\alpha_{\mathbf{h}}\rangle; \tag{4}$$

$$\text{PSI(0)}=\left(\sum_{\mathbf{h}}\alpha'_{\mathbf{h}}\right)\Big/\left(\sum_{\mathbf{h}}v_{\mathbf{h}}^{1/2}\right). \tag{5}$$

In (3) and (4) $\alpha_{\mathbf{h}}$ is the parameter of the Karle & Karle (1966) distribution estimating $\varphi_{\mathbf{h}}$ when involved in r triplets, *i.e.*

$$\alpha_{\mathbf{h}}=\left[\left(\sum_{j=1}^{r}G_j\cos\theta_j\right)^2+\left(\sum_{j=1}^{r}G_j\sin\theta_j\right)^2\right]^{1/2},$$

$$G_j=2|E_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{h}+\mathbf{k}}|/N^{1/2},\qquad \theta_j=\theta_{\mathbf{k}_j}+\theta_{\mathbf{h}-\mathbf{k}_j},$$

$\langle\alpha_{\mathbf{h}}\rangle$ is the expected value of $\alpha_{\mathbf{h}}$, and $(\alpha_{\mathbf{h}})_r$ is the expected value of $\alpha_{\mathbf{h}}$ when the phases are supposed to be randomly distributed. In *SIR* $\alpha_{\mathbf{h}}$, $\langle\alpha_{\mathbf{h}}\rangle$ and $(\alpha_{\mathbf{h}})_r$ can also be calculated by means of the parameters of the von Mises distributions relative to the second representation of triplets (Cascarano, Giacovazzo, Camalli, Spagna, Burla, Nunzi & Polidori, 1984).

In (5),

$$\alpha'_{\mathbf{h}}=\left[\left(\sum_{j=1}^{r}A_j\cos\theta_j\right)^2+\left(\sum_{j=1}^{r}A_j\sin\theta_j\right)^2\right]^{1/2},$$

where $\mathbf{h}$ is a reciprocal-lattice vector with $|E_{\mathbf{h}}|\simeq 0$,

$$v_{\mathbf{h}}=\sum_{j=1}^{r}A_j^2,\qquad A_j=|E_{\mathbf{k}_j}E_{\mathbf{h}-\mathbf{k}_j}|/N^{1/2}.$$

$\mathbf{k}_j$ and $\mathbf{h}-\mathbf{k}_j$ are indices of strong $|E|$ values for which phases have been determined.

In a recent paper (Cascarano, Giacovazzo, Burla, Nunzi & Polidori, 1984), the distribution of $\alpha_{\mathbf{h}}$ has been derived, assuming that θ_j, $j=1,\ldots,r$, are a random sample of circular variables $\varphi_{\mathbf{k}_j}+\varphi_{\mathbf{h}-\mathbf{k}_j}$ which are independently distributed around the true phase value $\varphi_{\mathbf{h}}$ with probability density function $f_j(\theta_j)$, $j=1,\ldots,r$. Under these conditions,

$$P(\alpha_{\mathbf{h}})=\frac{1}{2\pi}\alpha_{\mathbf{h}}\int_0^{\infty}\int_0^{2\pi}\rho J_0(\rho\alpha_{\mathbf{h}})\prod_{j=1}^{r}C_j(\rho,\psi)\,\mathrm{d}\rho\,\mathrm{d}\psi \tag{6}$$

where J_0 is the Bessel function of order zero, ρ and ψ are two carrying variables associated with $\alpha_{\mathbf{h}}$ and $\theta_{\mathbf{h}}$ and $C_j(\rho,\psi)$ is the jth characteristic function.

Distribution (6) was reduced to simpler expressions in two cases of interest to us:

(*a*) The variables θ_j are distributed around $\varphi_{\mathbf{h}}$ according to the von Mises distribution $M[\theta_j;\varphi_{\mathbf{h}},G_j]$. This situation describes the mathematical conditions presumably satisfied by phases occurring in triplets actively used in the phasing process.

(*b*) The variables θ_j are uniformly and independently distributed in the interval $(0,2\pi)$. This situation describes the mathematical conditions presumably satisfied by the θ_j variables when $|E_{\mathbf{h}}|\simeq 0{\cdot}0$.

By means of those distributions, in §§ 4 and 5 the statistical meanings of (3), (4) and (5) will be analysed and new FOM's will be introduced.

4. A distribution for triplets having $|E_{\mathbf{h}}|=0$

4.1. *Non-centrosymmetric case*

We are interested in the distribution of $\alpha'_{\mathbf{h}}$ when the moduli $|A_j|$ are known and the variables θ_j are random variables independently and isotropically distributed on the circle. Then, according to Cascarano, Giacovazzo, Burla, Nunzi & Polidori (1984),

$$P(\alpha'_{\mathbf{h}})\simeq(2\alpha'_{\mathbf{h}}/v_{\mathbf{h}})\exp(-\alpha'^2_{\mathbf{h}}/v_{\mathbf{h}}). \tag{7}$$

In accordance with the central limit theorem the variable $S=\sum_{\mathbf{h}}\alpha'_{\mathbf{h}}$ will be normally distributed around $\langle S\rangle$, given by

$$\langle S\rangle=\sum_{\mathbf{h}}\langle\alpha'_{\mathbf{h}}\rangle=\left(\sum_{\mathbf{h}}v_{\mathbf{h}}^{1/2}\right)\pi^{1/2}/2=0{\cdot}886\left(\sum_{\mathbf{h}}v_{\mathbf{h}}^{1/2}\right) \tag{8}$$

with variance

$$\sum_{\mathbf{h}}(\langle\alpha'^2_{\mathbf{h}}\rangle-\langle\alpha'_{\mathbf{h}}\rangle^2)=\left(\sum_{\mathbf{h}}v_{\mathbf{h}}\right)-\left(\sum_{\mathbf{h}}v_{\mathbf{h}}\right)\pi/4$$

$$\simeq 0{\cdot}215\left(\sum_{\mathbf{h}}v_{\mathbf{h}}\right). \tag{9}$$

The expectation

$$\sum_{\mathbf{h}}\alpha'_{\mathbf{h}}\Big/\sum_{\mathbf{h}}v_{\mathbf{h}}^{1/2}\equiv\text{PSI(0)}\simeq\pi^{1/2}/2=0{\cdot}886 \tag{10}$$

arises from (8), which corrects the relation PSI(0) = 1 largely quoted in the literature. In its turn (9) justifies the expectation

$$\left[\sum_{\mathbf{h}}(\alpha'_{\mathbf{h}}-\langle\alpha'_{\mathbf{h}}\rangle)^2\right]\Big/\sum_{\mathbf{h}}v_{\mathbf{h}}\simeq 0{\cdot}215. \tag{11}$$

By the change of variable $\psi_{\mathbf{h}}=\alpha'_{\mathbf{h}}/v_{\mathbf{h}}^{1/2}$, (7) becomes

$$P(\psi_{\mathbf{h}})\simeq 2\psi_{\mathbf{h}}\exp(-\psi_{\mathbf{h}}^2). \qquad (12)$$

It is not surprising that the standardized variable $\psi_{\mathbf{h}}$ has the same distribution as the normalized structure factor $|E_{\mathbf{h}}|$ in non-centrosymmetric space groups.

Knowledge of the distribution (12) allows us to use the general statistical procedures aimed at testing if an experimental sample with n elements complies with the expected distribution. In this view we should: (*a*) divide the real line into p mutually exclusive intervals, $A_1,\ldots,A_p$; (*b*) calculate the number of sample values n_i falling into A_i, $i=1,2,\ldots,p$; (*c*) calculate $P(A_i)$ according to (12); (*d*) calculate

$$d^2=\sum_{i=1}^{p}\{[n_i-nP(A_i)]^2/[nP(A_i)]\}$$

and reject the hypothesis if $d^2>C$, where C is obtained from the tables of the chi-squared distribution. If a significance level of α is desired, $C=\chi^2_{p-1,1-\alpha}$.

In a simpler way we could limit ourselves to checking the first moments of the distribution (12), or also, to calculating the percentage of $\psi_{\mathbf{h}}$ with amplitude greater than a given threshold (such procedures are usually preferred for checking, by means of Wilson statistics, the centro- or non-centrosymmetrical nature of the space group). Accordingly, the correct solution in a multisolution approach is expected to satisfy the following relations:

$$\frac{1}{n}\sum_{\mathbf{h}}\left(\frac{\alpha'_{\mathbf{h}}}{v_{\mathbf{h}}^{1/2}}\right)\simeq 0{\cdot}886; \qquad \frac{1}{n}\sum_{\mathbf{h}}\left(\frac{\alpha'^2_{\mathbf{h}}}{v_{\mathbf{h}}}\right)\simeq 1 \qquad (13)$$

$$\frac{1}{n}\sum_{\mathbf{h}}\left|\frac{\alpha'^2_{\mathbf{h}}}{v_{\mathbf{h}}}-1\right|\simeq 0{\cdot}736;$$

$$\frac{1}{n}\sum_{\mathbf{h}}\left(\frac{\alpha'^2_{\mathbf{h}}}{v_{\mathbf{h}}}-\frac{\pi^{1/2}}{2}\right)^2\simeq 0{\cdot}215 \qquad (14)$$

where n is the number of $\mathbf{h}$ reflexions used by the procedure.

If a correlation between $\theta_{\mathbf{h}}=\theta_{\mathbf{k}}+\theta_{\mathbf{h}-\mathbf{k}}$ and $\varphi_{\mathbf{h}}$ exists (as frequently occurs for wrong solutions), the measured moments (13) and (14) will be larger than the expected ones. It may be concluded that the correct solutions should be characterized by the smallest values of the measured moments.

4.2. *The centrosymmetric case*

This case is similar to the one-dimensional random walk, each step of the walk confined to the real axes. Now

$$\alpha'_{\mathbf{h}}=\left(\sum_{j=1}^{r}E_{\mathbf{k}_j}E_{\mathbf{h}-\mathbf{k}_j}\right)\Big/N^{1/2},$$

$$v_{\mathbf{h}}=\left(\sum_{j=1}^{r}|E_{\mathbf{k}_j}E_{\mathbf{h}-\mathbf{k}_j}|^2\right)\Big/N$$

$$P(\alpha'_{\mathbf{h}})\simeq(2\pi v_{\mathbf{h}})^{-1/2}\exp(-\alpha'^2_{\mathbf{h}}/2v_{\mathbf{h}}).$$

According to the central limit theorem the variable

$$S=\sum_{\mathbf{h}}\alpha'_{\mathbf{h}}$$

will be normally distributed around zero with variance equal to

$$\sum_{\mathbf{h}}\langle\alpha'^2_{\mathbf{h}}\rangle=\sum_{\mathbf{h}}v_{\mathbf{h}}.$$

Therefore the expectations

$$\sum_{\mathbf{h}}\alpha'_{\mathbf{h}}\simeq 0 \quad\text{and}\quad \left(\sum_{\mathbf{h}}\langle\alpha'^2_{\mathbf{h}}\rangle\right)\Big/\left(\sum_{\mathbf{h}}v_{\mathbf{h}}\right)\simeq 1 \qquad (15)$$

arise. On considering the distribution

$$P(|\alpha'_{\mathbf{h}}|)=[2/(\pi v_{\mathbf{h}})]^{1/2}\exp(-\alpha'^2_{\mathbf{h}}/2v_{\mathbf{h}}) \qquad (16)$$

the following expectations are derived:

$$\left(\sum_{\mathbf{h}}|\alpha'_{\mathbf{h}}|\right)\Big/\left(\sum_{\mathbf{h}}v_{\mathbf{h}}^{1/2}\right)=\mathrm{PSI}(0)=(2/\pi)^{1/2}=0{\cdot}798 \qquad (17)$$

$$\left(\sum_{\mathbf{h}}(|\alpha'_{\mathbf{h}}|-\langle|\alpha'_{\mathbf{h}}|\rangle)^2\right)\Big/\left(\sum_{\mathbf{h}}v_{\mathbf{h}}\right)\simeq 0{\cdot}363 \qquad (18)$$

where $|\alpha'_{\mathbf{h}}|=0{\cdot}798\,v_{\mathbf{h}}^{1/2}$. Different values are then expected for PSI(0) in centro- and in non-centrosymmetric space groups.

By the change of variable $\psi=|\alpha'_{\mathbf{h}}|/v_{\mathbf{h}}^{1/2}$ (16) changes into

$$P(\psi_{\mathbf{h}})\simeq(2/\pi)^{-1/2}\exp(-\psi_{\mathbf{h}}^2/2) \qquad (19)$$

which coincides with the distribution of the normalized structure factor $|E_{\mathbf{h}}|$ in centrosymmetric space groups. According to § 4.1 the first moments of (19) can be used to find the correct solution; in particular,

$$(1/n)\sum_{\mathbf{h}}(|\alpha'_{\mathbf{h}}|/v_{\mathbf{h}}^{1/2})\simeq 0{\cdot}798;$$

$$(1/n)\sum_{\mathbf{h}}(\alpha'^2_{\mathbf{h}}/v_{\mathbf{h}})\simeq 1 \qquad (20)$$

$$(1/n)\sum_{\mathbf{h}}|(\alpha'^2_{\mathbf{h}}/v_{\mathbf{h}})-1|\simeq 0{\cdot}968;$$

$$(1/n)\sum_{\mathbf{h}}[(|\alpha'_{\mathbf{h}}|/v_{\mathbf{h}}^{1/2})-(2/\pi)^{1/2}]^2\simeq 0{\cdot}363. \qquad (21)$$

5. Some applications of the distribution $P(\alpha)$

5.1. *The non-centrosymmetric case*

If r is sufficiently large, $P(\alpha_{\mathbf{h}})$ is normally distributed according to

$$P(\alpha_{\mathbf{h}})\simeq[(2\pi)^{1/2}\sigma_{\mathbf{h}}]^{-1}\exp[-(\alpha_{\mathbf{h}}-\langle\alpha_{\mathbf{h}}\rangle)^2/2\sigma_{\mathbf{h}}^2] \qquad (22)$$

where

$$\langle\alpha_{\mathbf{h}}\rangle=\sum_{j=1}^{r}G_jD_1(G_j), \qquad (23)$$

$$\sigma_{\mathbf{h}}^2=\tfrac{1}{2}\sum_{j=1}^{r}G_j^2[1+D_2(G_j)-2D_1^2(G_j)]. \qquad (24)$$

According to the central limit theorem the variable $S=\sum_{\mathbf{h}}\alpha_{\mathbf{h}}$ will be normally distributed around $\sum_{\mathbf{h}}\langle\alpha_{\mathbf{h}}\rangle$ with variance $\sum_{\mathbf{h}}\sigma_{\mathbf{h}}^2$. Then the expectation values

$$\left(\sum_{\mathbf{h}}\alpha_{\mathbf{h}}\right)\Big/\left(\sum_{\mathbf{h}}\langle\alpha_{\mathbf{h}}\rangle\right)\simeq 1 \tag{25}$$

$$\sum_{\mathbf{h}}\left(\alpha_{\mathbf{h}}-\langle\alpha_{\mathbf{h}}\rangle\right)^2\Big/\left(\sum_{\mathbf{h}}\sigma_{\mathbf{h}}^2\right)\simeq 1 \tag{26}$$

arise. Equations (25) and (26) may usefully be compared with MABS and RALPHA respectively.

By the change of variable $y_{\mathbf{h}}=(\alpha_{\mathbf{h}}-\langle\alpha_{\mathbf{h}}\rangle)/\sigma_{\mathbf{h}}$, the distribution (22) reduces to

$$P(y_{\mathbf{h}})\simeq(2\pi)^{-1/2}\exp(-y_{\mathbf{h}}^2/2). \tag{27}$$

According to § 4 the first moments of (27) may be used for finding the correct solution: in particular,

$$(1/n)\sum_{\mathbf{h}}|\alpha_{\mathbf{h}}-\langle\alpha_{\mathbf{h}}\rangle|/\sigma_{\mathbf{h}}\simeq 0{\cdot}798 \tag{28}$$

$$(1/n)\sum_{\mathbf{h}}[(\alpha_{\mathbf{h}}-\langle\alpha_{\mathbf{h}}\rangle)/\sigma_{\mathbf{h}}]^2\simeq 1 \tag{29}$$

$$(1/n)\sum_{\mathbf{h}}|[(\alpha_{\mathbf{h}}-\langle\alpha_{\mathbf{h}}\rangle)/\sigma_{\mathbf{h}}]^2-1|\simeq 0{\cdot}968 \tag{30}$$

$$(1/n)\sum_{\mathbf{h}}\{[(\alpha_{\mathbf{h}}-\langle\alpha_{\mathbf{h}}\rangle)/\sigma_{\mathbf{h}}]^2-(2/\pi)\}\simeq 0{\cdot}363 \tag{31}$$

may be used as a statistical measure of the deviation of the experimental from the expected distribution of α so that they are expected to be a minimum for the correct solution.

5.2. *The centrosymmetric case*

If r is sufficiently large, $P(\alpha_{\mathbf{h}})$ is normally distributed according to (22) where

$$\alpha_{\mathbf{h}}=\sum_{j=1}^{r}G_j'\tanh G_j',$$

$$\sigma^2=\sum_{j=1}^{r}G_j'^2(1-\tanh^2 G_j'),$$

$$G_j'=|E_{\mathbf{h}}E_{\mathbf{k}_j}E_{\mathbf{h}-\mathbf{k}_j}|/N^{1/2}.$$

Then the expectations (25)-(26) and (28)-(31) may be used for finding the correct solution.

6. The combined figure of merit CFOM

Let us denote

$$\mathrm{NPS}(0)=(1/q_0)\sum_{\mathbf{h}}\alpha_{\mathbf{h}}'/\sum_{\mathbf{h}}v_{\mathbf{h}}^{1/2},$$

$$\mathrm{NPS}(1)=(1/q_1)(1/n)\sum_{\mathbf{h}}(\alpha_{\mathbf{h}}'^2/v_{\mathbf{h}}),$$

$$\mathrm{NPS}(2)=(1/q_2)(1/n)\sum_{\mathbf{h}}|(\alpha_{\mathbf{h}}'^2/v_{\mathbf{h}})-1|,$$

$$\mathrm{NPS}(3)=(1/q_3)\sum_{\mathbf{h}}[(\alpha_{\mathbf{h}}'/v_{\mathbf{h}}^{1/2})-(\pi^{1/2}/2)]^2,$$

where

$$q_0=0{\cdot}886,\qquad q_1=1,\qquad q_2=0{\cdot}736,\qquad q_3=0{\cdot}215$$

for non-centrosymmetric space groups, and

$$q_0=0{\cdot}798,\qquad q_1=1,\qquad q_2=0{\cdot}968,\qquad q_3=0{\cdot}363$$

for centrosymmetric space groups. NPS(I), $I=0,\ldots,3$ are then all expected to be unity. In accordance with § 4.1 the combined FOM

$$\mathrm{PSCOMB}=\tfrac{1}{4}\left\{\mathrm{NPS}(0)+\sum_{i=1}^{3}[\mathrm{NPS}(I)]^{1/2}\right\}$$

is expected to be a minimum for the correct solution.

In practice this does not always occur, mostly because of a strong correlation between the FOM's based on PSI(0) triplets and the MABS figure of merit. Indeed, wrong solutions marked by small values of MABS (say 0·4-0·8) are very often characterized by small values of PSCOMB [the low consistency among PSI(0) triplets is nothing but the consequence of the low consistency among active triplets]. On the other hand, solutions characterized by large values of MABS (say 0·9-1·3) and small values of PSCOMB are often correct. In accordance with the above observations we assume:

$$\mathrm{PSCOMB}=\frac{1}{4(\mathrm{MABS})}\left\{\mathrm{NPS}(0)+\sum_{i=1}^{3}[\mathrm{NPS}(I)]^{1/2}\right\}.$$

In Table 3 we give for DIOLE ($C_{10}H_{18}O_2$; $I\bar{4}2d$, $Z=16$) a summary of figures of merit output by tangent formula and including PSCOMB. Set 12 is the correct solution; it gives large values of NPS(I), $I=0,\ldots,3$, but the smallest value of PSCOMB.

Let us now define, for centrosymmetric as well for non-centrosymmetric solutions,

$$\mathrm{NALF}(1)=(1/q_1)(1/n)\sum_{\mathbf{h}}|S\alpha_{\mathbf{h}}-\langle\alpha_{\mathbf{h}}\rangle|/\sigma_{\mathbf{h}}$$

$$\mathrm{NALF}(2)=(1/q_2)\left[\sum_{\mathbf{h}}(S\alpha_{\mathbf{h}}-\langle\alpha_{\mathbf{h}}\rangle)^2\Big/\sum_{\mathbf{h}}\sigma_{\mathbf{h}}^2\right]$$

$$\mathrm{NALF}(3)=(1/q_3)(1/n)\sum_{\mathbf{h}}[(S\alpha-\langle\alpha\rangle)/\sigma_{\mathbf{h}}]^2$$

$$\mathrm{NALF}(4)=(1/q_4)(1/n)\sum_{\mathbf{h}}|[(S\alpha_{\mathbf{h}}-\langle\alpha_{\mathbf{h}}\rangle)/\sigma_{\mathbf{h}}]^2-1|$$

where $S=1/\mathrm{MABS}$ and

$$q_1=0{\cdot}798,\qquad q_2=1,\qquad q_3=1,\qquad q_4=0{\cdot}968.$$

Two considerations suggest rescaling (by S) the experimental values of α to the expected ones:

(1) The moments NALF(I), $I=1,\ldots,4$ can easily lose any statistical meaning if no rescaling is made. Indeed, the experimental α values critically depend on the weighting scheme used in the tangent refinement, on the starting set, on the number of triplets *etc.* Thus, correct solutions are often marked by very different sets of α which, however, give rise to similar values of NALF(I).

Table 3. *MABS, NPS(I), I = 0, . . . , 3, and PSCOMB for the 12 sets of phases output by tangent formula for DIOLE*

Set	MABS	NPS(0)	NPS(1)	NPS(2)	NPS(3)	PSCOMB
1	0·51	1·44	1·91	1·75	2·13	2·72
2	0·51	1·46	2·05	2·02	2·72	2·92
3	0·57	1·73	2·93	3·02	4·77	3·25
4	0·52	1·47	2·00	1·87	2·33	2·79
5	0·65	1·91	3·47	3·73	6·19	3·16
6	0·65	1·90	3·43	3·69	6·06	3·11
7	0·66	2·22	4·65	5·16	9·45	3·68
8	0·55	1·62	2·57	2·67	3·98	3·14
9	0·52	1·42	2·05	2·11	3·09	2·89
10	0·66	2·24	4·75	5·30	9·81	3·73
11	0·67	1·93	3·42	3·61	5·76	3·02
12*	0·69	1·74	2·93	3·07	4·86	2·68
†	1·00	2·27	4·97	5·66	10·35	2·52

* Indicates the correct solution.
† The last set corresponds to published phases.

Table 4. *NALF(I), I = 1, . . . , 4, and ALFCOMB for the 12 sets of phases output by tangent formula for TURSCH 10*

Set	NALF(1)	NALF(2)	NALF(3)	NALF(4)	ALFCOMB
1	2·37	7·50	5·68	5·26	2·45
2	2·27	5·82	5·14	4·74	2·28
3*	1·49	2·62	2·26	2·02	1·51
4	2·21	6·04	5·07	4·66	2·27
5	2·17	5·67	4·87	4·45	2·22
6*	1·51	2·71	2·33	2·10	1·53
7	2·24	6·06	5·07	4·65	2·28
8	2·28	6·29	5·26	4·84	2·32
9*	1·51	2·71	2·33	2·09	1·53
10	2·15	5·71	4·81	4·41	2·21
11*	1·33	2·37	1·79	1·59	1·37
12*	1·33	1·91	1·71	1·50	1·31

* Indicate the correct solutions.

(2) MABS is mathematically correlated with NALF(I): if no rescaling is made, large NALF(I) moments should systematically occur for solutions with MABS very different from unity.

NALF(I), $I = 1, \ldots, 4$, are all expected to be unity. In accordance with § 5.1 the combined FOM

$$\text{ALFCOMB} = \tfrac{1}{4}\left\{\text{NALF}(1) + \sum_{I=2}^{4} [\text{NALF}(I)]^{1/2}\right\}$$

is expected to be a minimum for the correct solution. In Table 4 we give for TURSCH 10 ($C_{15}H_{24}O_2$; $P6_322$, $Z = 12$) a summary of figures of merit output by tangent formula and concerning ALFCOMB. The sets 3, 6, 9, 11 and 12 are correct solutions, characterized by the smallest values of ALFCOMB.

In their turn PSCOMB and ALFCOMB may be collated with MABS and CPHASE to produce the overall combined figure of merit

$$\begin{aligned}\text{CFOM} = \left(\sum_{i=1}^{4} w_i\right)^{-1} \{ & w_1 \,\text{DABS} \\ & + w_2 \exp[-(1{\cdot}0 - \text{CPHASE})^{1{\cdot}5}] \\ & + w_3 \exp[-(\text{ALFCOMB} - 1)^{1{\cdot}5}] \\ & + w_4 \exp[-(\text{PSCOMB} - 1)^{1{\cdot}5}]\} \end{aligned} \qquad (32)$$

where DABS = 1 − |MABS − 1|.

It may be noted that:

(1) The larger the difference |MABS − 1| (it is always assumed DABS ≥ 0) the smaller the contribution to CFOM. Such a criterion does not hold for solutions with overcorrelated active triplets (for example, the so-called uranium solutions) or for solutions with small values of MABS.

(2) The contribution of any function F to the overall figure of merit (1) is maximum (=1) when $F = F_{max}$ and minimum (=0) when $F = F_{min}$ even in the case in which $F_{max} \simeq F_{min}$ or when F does not comply with expectations. A similar observation may be made for the function f. In (32) MABS, PSCOMB and ALFCOMB are always compared with their expected values (=1) so that the contribution of each of them to CFOM is maximum (=1) when the experimental and the expected values coincide. The larger the difference between the expected and experimental values, the smaller the contribution to CFOM.

(3) DABS, PSCOMB and ALFCOMB are figures of merit of asymptotical nature (they should hold for large as well as for small structures) while CPHASE is calculated *via* phase relationships whose reliability depends on the structural complexity. So CPHASE is expected to be comparatively more reliable for small than for large structures.

(4) A comparison of the reliability of the various FOM's in CFOM has not been attempted. Thus w_1, w_2, w_3, w_4 in (32) are in practice consequences of our *a posteriori* confidence in the various FOM's. In our experience, for usual structures,

$$w_1 = 0{\cdot}2; \qquad w_2 = 1{\cdot}0; \qquad w_3 = 1{\cdot}0; \qquad w_4 = 1{\cdot}4$$

are sensible weights. The numerical values w_i cannot be directly considered a measure of our confidence in the various FOM's. For example, the choice $w_2 < w_4$ does not imply that CPHASE is expected to be less reliable than PSCOMB: indeed CPHASE is usually spread out on a larger interval [say (−1, 1)] than PSICOMB [say (1, 2)] and relative weights in CFOM take into account this fact too.

In Table 5 the values of various FOM's for the correct solutions are shown for various test structures. They are compared with corresponding values for published phases and with FOM's corresponding to the largest CFOM's relative to incorrect solutions.

7. Concluding remarks

A revision of the statistical meaning of some widely used FOM's is made and new FOM's are described. Each FOM, including the combined CFOM, is expected to be unitary for the correct solution. Table 5 shows that this expected behaviour is frequently obeyed and that CFOM is usually a very discriminant

Table 5. *FOM's for correct solutions for various test structures*

For useful comparison, the largest FOM's for incorrect solutions are given in parentheses (in QUINO no incorrect solution was found): FOM's calculated from published phases are shown in square brackets.

	DABS	CPHASE	ALFCOMB	PSCOMB	CFOM
APAPA	0·16	0·97	0·85	0·86	0·89
	(0·15)	(0·62)	(0·78)	(0·86)	(0·77)
	[0·04]	[1·00]	[0·67]	[0·92]	[0·87]
AX118	0·08	1·00	1·00	0·53	0·74
	(0·22)	(0·79)	(0·04)	(0·41)	(0·30)
	[0·15]	[1·00]	[1·00]	[0·60]	[0·77]
AZET	0·45	0·90	0·99	0·25	0·68
	(0·45)	(0·95)	(0·98)	(0·30)	(0·71)
	[0·27]	[1·00]	[1·00]	[0·93]	[0·96]
CEPHA	0·07	0·97	0·83	0·81	0·87
	(0·18)	(0·24)	(0·17)	(0·44)	(0·33)
	[0·09]	[1·00]	[1·00]	[1·00]	[0·99]
CORTI	0·03	0·86	0·54	0·69	0·66
	(0·06)	(0·53)	(0·30)	(0·46)	(0·44)
	[0·02]	[0·96]	[0·61]	[0·97]	[0·83]
DIOLE	0·05	0·82	0·71	0·65	0·74
	(0·22)	(0·88)	(0·23)	(0·08)	(0·40)
	[0·30]	[0·52]	[0·99]	[1·00]	[0·84]
ERGO	0·05	0·44	0·75	0·82	0·69
	(0·17)	(0·51)	(0·47)	(0·76)	(0·61)
	[0·10]	[0·65]	[0·71]	[0·95]	[0·79]
LITHO	0·11	0·89	0·74	0·94	0·85
	(0·24)	(0·56)	(0·21)	(0·39)	(0·41)
	[0·06]	[0·99]	[0·85]	[1·00]	[0·95]
PG205	0·07	1·00	1·00	0·72	0·84
	(0·11)	(0·85)	(0·25)	(0·47)	(0·42)
	[0·07]	[1·00]	[1·00]	[0·99]	[0·99]
PHOTO	0·04	0·40	0·34	0·29	0·36
	(0·04)	(0·38)	(0·33)	(0·26)	(0·34)
	[0·10]	[0·89]	[0·98]	[0·99]	[0·97]
QUINO	0·07	1·00	1·00	0·96	0·97
	—	—	—	—	—
	[0·10]	[1·00]	[1·00]	[1·00]	[0·99]
TURSCH10	0·15	1·00	0·95	0·92	0·93
	(0·10)	(0·73)	(0·32)	(0·71)	(0·58)
	[0·14]	[1·00]	[0·97]	[0·94]	[0·94]

All the results in the table are obtained by using the second representation formula for triplet invariants (Cascarano, Giacovazzo, Camalli, Spagna, Burla, Nunzi & Polidori, 1984). References for crystal structures tested in this table are not given, for the sake of brevity. The reader is referred to the tape distributed by the crystallographic group of the University of York, England.

figure of merit. Deviations from expectations are mostly generated by sources such as:

(*a*) Poor experimental data and/or imperfect normalization procedure.

(*b*) The phases of the 'correct' solution significantly deviate (20–35°) from the true phases. Then the various FOM's register this situation giving rise to a CFOM markedly different from unity. As an example (see Table 5) we have introduced in the starting set of PHOTO only three magic integers; the mean deviation of the best solution phases from the published ones is 31°, and the related CFOM is 0·36, markedly different from unity.

(*c*) Systematic errors are made in the estimation of the structure invariants and seminvariants. Wrong estimates of the unit-cell contents (then wrong N values are used throughout the procedure) or the presence of structural regularities are possible sources of the errors. In particular, pseudotranslational symmetry is a severe source of errors which may violate the fundamental postulates on which the above statistical analysis is based. For such cases specific FOM's have to be devised.

The above considerations suggest a practical criterion; solutions with CFOM ~ 1 can be considered very promising, solutions with CFOM $\ll 1$ can be considered wrong solutions, unless structural regularities (the presence of which is often suggested by a statistical analysis of the normalization output) may be invoked.

One of us (DV) is indebted to CSI Piemonte for financial support.

References

BURLA, M. C., NUNZI, A., GIACOVAZZO, C. & POLIDORI, G. (1981). *Acta Cryst.* A**37**, 677–684.

BUSETTA, B., GIACOVAZZO, C., BURLA, M. C., NUNZI, A., POLIDORI, G. & VITERBO, D. (1980). *Acta Cryst.* A**36**, 68–74.

CASCARANO, G., GIACOVAZZO, C., BURLA, M. C., NUNZI, A. & POLIDORI, G. (1984). *Acta Cryst.* A**40**, 389–394.

CASCARANO, G., GIACOVAZZO, C., CALABRESE, G., BURLA, M. C., NUNZI, A., POLIDORI, G. & VITERBO, D. (1984). *Z. Kristallogr.* **167**, 37–47.

CASCARANO, G., GIACOVAZZO, C., CAMALLI, M., SPAGNA, R., BURLA, M. C., NUNZI, A. & POLIDORI, G. (1984). *Acta Cryst.* A**40**, 278–283.

CERRINI, S., GAVUZZO, E., FEDELI, W., LUCENTE, G., PINNEN, F. & ZANOTTI, G. (1986). In the press.

COCHRAN, W. & DOUGLAS, A. S. (1957). *Proc. R. Soc. London Ser. A*, **243**, 281.

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.* A**35**, 622–626.

DE TITTA, G. T., EDMONDS, J. W., LANGS, D. A. & HAUPTMAN, H. (1975). *Acta Cryst.* A**31**, 472–479.

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

GIACOVAZZO, C. (1980*a*) *Acta Cryst.* A**36**, 362–372.

GIACOVAZZO, C. (1980*b*) *Direct Methods in Crystallography.* London: Academic Press.

HASEK, J., SCHENK, H., KIERS, C. TH. & SCHAGEN, J. D. (1985). *Acta Cryst.* A**41**, 333–340.

KARLE, J. & KARLE, I. (1966). *Acta Cryst.* **21**, 849–859.

MAIN, P. (1977). *Acta Cryst.* A**33**, 750–757.

NUNZI, A., BURLA, M. C., POLIDORI, G., GIACOVAZZO, C., CASCARANO, G., VITERBO, D., CAMALLI, M., SPAGNA, R. (1984). *Acta Cryst.* A**40**, C425.

OVERBEEK, A. R. & SCHENK, H. (1976). *Proc. K. Ned. Akad. Wet.* B**79**, 341–343.

PONTENAGEL, W. M. G. F. (1984). *Acta Cryst.* A**40**, 314–323.

PUTTEN, N. VAN DER & SCHENK, H. (1979). *Acta Cryst.* A**35**, 381–387.

SCHENK, H. (1972). *Acta Cryst.* A**28**, 412–422.

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