HECHT, H.-J., SOBEK, H., HAAG, T., PFEIFER, O. & VAN PEE, K. H. (1994). *Nature (London) Struct. Biol.* **1**, 532–537.
KOSTREWA, D., GRANZIN, J., STOCK, D., CHOE, H.-W., LABAHN, J. & SAENGER, W. (1992). *J. Mol. Biol.* **226**, 209–226.
KRETSINGER, R. H. & NOCKOLDS, C. E. (1973). *J. Biol. Chem.* **248**, 3313–3326.
LABAHN, J., GRANZIN, J., SCHLUCKEBIER, G., ROBINSON, D. P., JACK, W. E., SCHILDKRAUT, I. & SAENGER, W. (1995). *Proc. Natl Acad. Sci. USA*. In press.
MATTEVI, A., OBMOLOVA, G., SCHULZE, E., KALK, K. H., WESTPHAL, A. H., DE KOK, A. & HOL, W. G. J. (1992). *Science*, **255**, 1544–1550.
RICHARDSON, J. S. & RICHARDSON, D. C. (1985). *Methods in Enzymology*, Vol. 115B, edited by H. W. WYCKOFF, C. H. W. HIRS & S. N. TIMASHEFF, pp. 189–206. Orlando: Academic Press.
SUBRAMANIAN, V. & HALL, S. R. (1982). *Acta Cryst.* A**38**, 577–590.
ZANOTTI, G., SCAPIN, G., SPADON, P., VEERKAMP, J. H. & SACCHETTINI, J. C. (1992). *J. Biol. Chem.* **267**, 18541–18550.
ZHANG, K. Y. J. & MAIN, P. (1990). *Acta Cryst.* A**46**, 41–46.

*Acta Cryst.* (1995). A**51**, 820–825

# The Joint Probability Distribution of Any Set of Phases Given Any Set of Diffraction Magnitudes. IV. The Active Use of Psi-Zero Triplets

BY G. CASCARANO AND C. GIACOVAZZO

*Istituto di Ricerca per lo Sviluppo di Metodologie Cristallografiche CNR, c/o Dipartimento Geomineralogico, Campus Universitario, 70124 Bari, Italy*

(*Received* 5 *December* 1994; *accepted* 14 *March* 1995)

## Abstract

In some recent papers [Giacovazzo, Burla & Cascarano (1992). *Acta Cryst.* A**48**, 901–906; Burla, Cascarano & Giacovazzo (1992). *Acta Cryst.* A**48**, 906–912; Cascarano, Giacovazzo, Moliterni & Polidori (1994). *Acta Cryst.* A**50**, 22–27], the method of the joint probability distribution of structure factors has been used to define a function that is frequently a maximum for the correct structure. Such a function was the basis for a modified tangent formula using $P10$ (negative and positive) triplet estimates and negative quartet estimates, which proved more efficient than the classical tangent formula of Karle & Hauptman [*Acta Cryst.* (1956), **9**, 635–651]. The method is here combined with a recent formulation [Giacovazzo (1993). *Z. Kristallogr.* **206**, 161–171], which suggests the supplementary active use in the phasing process of psi-zero triplets. Experimental tests prove the higher efficiency of the method and justify the default active use of psi-zero relationships in *SIR*92.

## Symbols and notation

Symbols and notation are the same as in the following papers; Giacovazzo, Burla & Cascarano (1992); Burla, Cascarano & Giacovazzo (1992); Cascarano, Giacovazzo, Moliterni & Polidori (1994); from now on these are referred to as papers I, II and III, respectively.

## Introduction

In papers I and II of this series, the conditional joint probability distribution of $n$ phases given $p$ ($p \geq n$) moduli was studied. Large values of $n$ and $p$ are allowed; *e.g.* $n$ may be the number of strong reflections that are usually phased by a modern direct procedure and $p$ may be the number of measured reflections. The resulting distribution is of exponential type and contains triplet and quartet contributions. Contrary to any expectations, the distribution is not maximized by the correct set of phases as one would expect for sufficiently large values of $n$ and $p$. Accordingly, the combined use of triplets and quartets proved of limited usefulness for practical applications.

In paper III, the failure of the distribution was ascribed to the strong correlation between triplets and positive estimated quartets. A modified expansion of the distribution was then proposed that neglects the contribution of the positive estimated quartets and retains terms arising from triplets and negative quartets only. The distribution is often maximized by the correct solution. Accordingly, an efficient tangent formula was described actively using triplet [estimated positive or negative by the $P10$ formula (Cascarano, Giacovazzo, Camalli, Spagna, Burla, Nunzi & Polidori, 1984)] and negative quartet relationships.

The main guidelines of the three papers may be summarized as follows: if a suitable function may be found that is maximum for the correct solution, then a tangent formula may be designed that is expected to be more efficient than the traditional tangent formula of Karle & Hauptman (1956). The function to maximize is in practice a figure of merit (FOM), a tool for recognizing the correct solution among numerous trial solutions.

This idea was first formulated in a recent paper by Giacovazzo (1993) where a new method is proposed that actively uses the information contained in the psi-zero triplets in order to drive phases towards the correct




ISSN 0108-7673 

values. For reader usefulness, we summarize here Giacovazzo's main guidelines and results:

(*a*) The Cochran (1952) criterion

$$S = \int_V \rho^3(\mathbf{r})\,d\mathbf{r} = \sum_{\mathbf{h},\mathbf{k}_j} G_{\mathbf{h},\mathbf{k}_j} \cos(\varphi_{\mathbf{h}} - \varphi_{\mathbf{k}_j} - \varphi_{\mathbf{h}-\mathbf{k}_j}) = \max \quad (1)$$

is inadequate for the solution of complex structures (Altomare, Cascarano & Giacovazzo, 1992).

(*b*) Equation (1) may also be written as

$$S = \sum_{\mathbf{h}} \alpha_{\mathbf{h}} \cos(\varphi_{\mathbf{h}} - \theta_{\mathbf{h}}) = \max,$$

where $\theta_{\mathbf{h}}$ and $\alpha_{\mathbf{h}}$ are calculated in accordance with the tangent formula (Karle & Hauptman, 1956):

$$\tan\theta_{\mathbf{h}} = \frac{\sum_j G_{\mathbf{h},\mathbf{k}_j} \sin(\varphi_{\mathbf{k}_j} + \varphi_{\mathbf{h}-\mathbf{k}_j})}{\sum_j G_{\mathbf{h},\mathbf{k}_j} \cos(\varphi_{\mathbf{k}_j} + \varphi_{\mathbf{h}-\mathbf{k}_j})} = \frac{T_{\mathbf{h}}}{B_{\mathbf{h}}} \quad (2)$$

with $\alpha_{\mathbf{h}} = (T_{\mathbf{h}}^2 + B_{\mathbf{h}}^2)^{1/2}$. No matter what the prior distribution of the $\varphi$'s may be, the tangent formula will work to make $\varphi_{\mathbf{h}} = \theta_{\mathbf{h}}$ and $\sum_{\mathbf{h}} \alpha_{\mathbf{h}} = \max$.

(*c*) In accordance with (*b*), the application of the tangent formula will lead to various local maxima for the figure of merit

$$\mathrm{MABS} = \sum_{\mathbf{h}} \alpha_{\mathbf{h}} \Big/ \sum_{\mathbf{h}} \langle\alpha_{\mathbf{h}}\rangle, \quad (3)$$

among which the correct solution may be found by application of FOM's more discriminating than MABS. Why then drive phases towards maxima of MABS and not towards maxima of more powerful FOM's? This would increase the ratio of the number of correct solutions to the number of trials.

(*d*) The maximization of the psi-zero FOM was chosen as an additional pivot of the phasing process. Let $A'_{\mathbf{h}_w,\mathbf{k}_j} = 2|E_{\mathbf{k}_j} E_{\mathbf{h}_w-\mathbf{k}_j}|N^{-1/2}$ and suppose that $|E_{\mathbf{k}_j}|$ and $|E_{\mathbf{h}_w-\mathbf{k}_j}|$ are strong while $|E_{\mathbf{h}_w}|$ is weak (the subscript $w$ stands for weak). As soon as the phases $\varphi_{\mathbf{k}_j}$ and $\varphi_{\mathbf{h}_w-\mathbf{k}_j}$ become available during the refinement process, the following quantities should be calculated:

$$\alpha'_{\mathbf{h}_w} \sin\theta'_{\mathbf{h}_w} = \sum_{\mathbf{k}_j} A'_{\mathbf{h}_w,\mathbf{k}_j} \sin(\varphi_{\mathbf{k}_j} + \varphi_{\mathbf{h}_w-\mathbf{k}_j})$$

$$\alpha'_{\mathbf{h}_w} \cos\theta'_{\mathbf{h}_w} = \sum_{\mathbf{k}_j} A'_{\mathbf{h}_w,\mathbf{k}_j} \cos(\varphi_{\mathbf{k}_j} + \varphi_{\mathbf{h}_w-\mathbf{k}_j}),$$

from which

$$\tan\theta'_{\mathbf{h}_w} = \frac{\sum_{\mathbf{k}_j} A'_{\mathbf{h}_w,\mathbf{k}_j} \sin(\varphi_{\mathbf{k}_j} + \varphi_{\mathbf{h}_w-\mathbf{k}_j})}{\sum_{\mathbf{k}_j} A'_{\mathbf{h}_w,\mathbf{k}_j} \cos(\varphi_{\mathbf{k}_j} + \varphi_{\mathbf{h}_w-\mathbf{k}_j})} = \frac{T'_{\mathbf{h}_w}}{B'_{\mathbf{h}_w}} \quad (4)$$

and

$$\alpha'_{\mathbf{h}_w} = (T'^2_{\mathbf{h}_w} + B'^2_{\mathbf{h}_w})^{1/2}. \quad (5)$$

$\theta'_{\mathbf{h}_w}$ is not an estimate of $\varphi_{\mathbf{h}_w}$ ($\varphi_{\mathbf{h}_w}$ itself could be meaningless with $|E_{\mathbf{h}_w}| = 0$) but only a carrying variable useful later on in the procedure. Similarly, $\alpha'_{\mathbf{h}_w}$ is not a measure of reliability for the relation $\theta'_{\mathbf{h}_w} \simeq \varphi_{\mathbf{h}_w}$ but only a measure of the consistency (for the running trial) of the variables $(\varphi_{\mathbf{k}_j} + \varphi_{\mathbf{h}_w-\mathbf{k}_j})$.

(*e*) The criterion

$$\sum_{\mathbf{h}_w} \alpha'_{\mathbf{h}_w} = \min \quad (6)$$

was suggested: the analytical conditions are found through the equation

$$\delta\Big(\sum_{\mathbf{h}_w} \alpha'_{\mathbf{h}_w}\Big) \Big/ \delta\varphi_{\mathbf{k}_q} = 0. \quad (7)$$

It should be noted that in (7) the minimum of $\sum_{\mathbf{h}_w} \alpha'_{\mathbf{h}_w}$ is searched *via* variations of the phases of strong reflections. Relation (7) looks for the $\varphi_{\mathbf{k}_q}$ value that minimizes $\sum_{\mathbf{h}_w} \alpha'_{\mathbf{h}_w}$. The combination of (7) with the classical tangent formula provides the modified tangent formula

$$\begin{aligned}\tan\theta_{\mathbf{k}_q} = &\Big[\sum_{\mathbf{h}} G_{\mathbf{h},\mathbf{k}_q} \sin\Big(\varphi_{\mathbf{h}} - \varphi_{\mathbf{h}-\mathbf{k}_q}\Big) \\ &- \sum_{\mathbf{h}_w} A'_{\mathbf{h}_w,\mathbf{k}_q} \sin\Big(\theta'_{\mathbf{h}_w} - \varphi_{\mathbf{h}_w-\mathbf{k}_w}\Big)\Big] \\ &\times \Big[\sum_{\mathbf{h}} G_{\mathbf{h},\mathbf{k}_q} \cos\Big(\varphi_{\mathbf{h}} - \varphi_{\mathbf{h}-\mathbf{k}_q}\Big) \\ &- \sum_{\mathbf{h}_w} A'_{\mathbf{h}_w,\mathbf{k}_q} \cos\Big(\theta'_{\mathbf{h}_w} - \varphi_{\mathbf{h}_w-\mathbf{k}_q}\Big)\Big]^{-1} \\ = &\, T''_{\mathbf{k}_q}/B''_{\mathbf{k}_q}. \end{aligned} \quad (8)$$

By analogy with the Karle & Hauptman tangent formula, the parameter

$$\alpha''_{\mathbf{k}_q} = \Big(T''^2_{\mathbf{k}_q} + B''^2_{\mathbf{k}_q}\Big)^{1/2} \quad (9)$$

is defined.

(*f*) The terms $A'$ in (8) are not concentration parameters of von Mises distributions (like the $G$'s). Therefore, additional considerations are needed in order to balance the terms $A'$ with respect to the terms $G$.

The above conclusions proposed by Giacovazzo suggest that the results of paper III could be integrated with the procedure suggested by (6)–(9) provided a function is identified, containing positive and negative triplets, negative quartets and psi-zero contributions, which is a maximum for the correct solution. The scope of this paper is to propose such a new function, to investigate its theoretical and practical implications, to devise a strategy for the phasing process and to describe the first applications. The results presented here justify the default active use of psi-zero triplets in the package *SIR*92 (Altomare, Cascarano, Giacovazzo, Guagliardi, Burla, Polidori & Camalli, 1994).

Table 1. *Code name, space group and crystallochemical data for test structures*

| Structure code | Space group | Molecular formula | $Z$ |
|---|---|---|---|
| AMIDE | $Pbc2$ | $C_7H_9N_3O_2$ | 8 |
| APAPA | $P4_12_12$ | $C_{30}H_{37}N_{15}O_{16}P_2\cdot 6H_2O$ | 8 |
| AZET | $Pca2_1$ | $C_{21}H_{16}ClNO$ | 8 |
| BED | $I4$ | $C_{26}H_{26}N_4O_4$ | 8 |
| BOBBY | $P2_13$ | $Na^+\cdot Ca^{2+}\cdot N(CH_2CO_2)_3^{3-}$ | 4 |
| CEPHAL | $C2$ | $C_{18}H_{21}NO_4$ | 8 |
| CIME | $Cc$ | $C_{10}H_{18}N_6OS$ | 4 |
| CUIMID | $P3_221$ | $C_6H_8N_4ClCu$ | 6 |
| DIAM | $P4_2/n$ | $C_{14}H_{20}O$ | 8 |
| DIOLE | $I\bar{4}2d$ | $C_{10}H_{18}O_2$ | 16 |
| ERGO | $P2_12_12_1$ | $C_{28}H_{44}O$ | 8 |
| ERICA | $P2_1$ | $C_{37}H_{43}FeO_4P$ | 2 |
| FEGAS | $P6_3/mmc$ | $Fe_2Ga_2S_5$ | 2 |
| GIAC | $P2_1/c$ | $C_{17}H_{17}NO_2S$ | 4 |
| GOLDMAN2 | $Cc$ | $C_{28}H_{16}$ | 8 |
| GRA4 | $P\bar{1}$ | $C_{30}H_{22}N_2O_4$ | 2 |
| HOV1 | $C2/m$ | $Pr_{14}Ni_{16}Si_{11}$ | 4 |
| INOS | $P2_1/n$ | $C_6H_{12}O_6\cdot H_2O$ | 8 |
| LOGANIN | $P2_12_12_1$ | $C_{17}H_{26}O_{10}$ | 4 |
| MBH2 | $P1$ | $C_{15}H_{24}O_3$ | 3 |
| MGHEX | $P3_1$ | $C_{48}H_{68}N_{12}O_{12}Mg\cdot 2ClO_4\cdot 4CH_3CN$ | 3 |
| MUNICH1 | $C2$ | $C_{20}H_{16}$ | 8 |
| NEWQB | $P\bar{1}$ | $C_{24}H_{20}N_2O_5$ | 4 |
| NO55 | $Fdd2$ | $C_{20}H_{24}N_4$ | 16 |
| PGE2 | $P1$ | $C_{20}H_{32}O_5$ | 1 |
| POCRO | $B112/m$ | $K_2Se_{16}Cr_{10}$ | 1 |
| QUINOL | $R\bar{3}$ | $C_6H_6O_2$ | 54 |
| RIFOLO | $P2_1$ | $C_{39}H_{49}NO_{13}\cdot CH_3OH\cdot H_2O$ | 2 |
| SALEX | $P\bar{3}$ | $K_{3.86}Na_{5.30}H_3O^+_{0.84}\cdot Fe_6^{3+}\cdot\cdot$ $[O_2(SO_4)_{12}]\cdot 17.08H_2O$ | 1 |
| SCHWARZ | $P1$ | $C_{46}H_{70}O_{27}$ | 1 |
| SELENID | $P2_1$ | $C_{22}H_{28}O_2Se$ | 2 |
| SKN1 | $P3_1$ | $C_7H_{16}ClNO_4$ | 3 |
| SUOA | $P2_12_12_1$ | $C_{28}H_{38}O_{19}$ | 4 |
| TPH | $C222_1$ | $C_{24}H_{20}N_2$ | 12 |
| TUR10 | $P6_322$ | $C_{15}H_{24}O_2$ | 12 |
| WINTER2 | $P2_1$ | $C_{52}H_{83}N_{11}O_{16}\cdot 3CH_2Cl_2$ | 2 |

References for test structures are not given for the sake of brevity. The reader can find them in paper III.

## A function to maximize

In paper III, tests were made on the 36 crystal structures quoted in Table 1. It was shown that the maximum of

$$S''' = \sum_{\text{triplets}} T_{ijl}\cos t_{ijl} + \sum_{\substack{\text{neg. est.}\\ \text{quartets}}} Q''_{ijlm}\cos q_{ijlm}$$
$$= S + SQN'' \qquad (10)$$

tends to characterize the correct solution with much higher frequency than the maximum of $S$. The result was obtained by using a large number of negative quartets. For example, for the eight structures crystallizing in a symmorphic space group (BED, CEPHAL, GRA4, MBH2, NEWQB, PGE2, QUINOL, SCHWARZ), the number of quartets involved in (10) varied (see Table III.2) from the minimum value of 4464 for CEPHAL to the maximum value of 70 136 for QUINOL. The calculation of such a large number of quartets as well as their active use in the phasing process are time consuming and are not advisable for routine structure determination. *SIR*92 only stores 750 negative quartets.

Table 2. *For the 36 test structures, the trial solutions produced by default runs of SIR92 are ranked in order of $S'''$ and $S^{IV}$*

NORD($S'''$) and NORD($S^{IV}$) are the order numbers of the correct solution. If this is not found by a default run of *SIR*92, then NORD refers to the published structure and is given in parentheses. In the last two columns, $Y$ denotes the success of the phasing procedure, $N$ denotes a failure.

| Structure code | NORD($S'''$) | NORD($S^{IV}$) | $P10+NQ$ | $P10+NQ+PSI0$ |
|---|---|---|---|---|
| AMIDE | 1 | 1 | *Y* | *Y* |
| APAPA | 1 | 1 | *Y* | *Y* |
| AZET | 1 | 1 | *Y* | *Y* |
| BED | 7 | 1 | *Y* | *Y* |
| BOBBY | 1 | 1 | *Y* | *Y* |
| CEPHAL | 7 | 1 | *Y* | *Y* |
| CIME | 1 | 1 | *Y* | *Y* |
| CUIMID | 1 | 1 | *Y* | *Y* |
| DIAM | 1 | 1 | *Y* | *Y* |
| DIOLE | 1 | 1 | *Y* | *Y* |
| ERGO | 7 | 1 | *Y* | *Y* |
| ERICA | 1 | 1 | *Y* | *Y* |
| FEGAS | 1 | 1 | *Y* | *Y* |
| GIAC | 1 | 1 | *Y* | *Y* |
| GOLDMAN2 | 1 | 1 | *Y* | *Y* |
| GRA4 | 3 | 4 | *Y* | *Y* |
| HOV1 | 1 | 1 | *Y* | *Y* |
| INOS | 1 | 1 | *Y* | *Y* |
| LOGANIN | 1 | 1 | *Y* | *Y* |
| MBH2 | 25 | 1 | *Y* | *Y* |
| MGHEX | (5) | (1) | *N* | *N* |
| MUNICH1 | (12) | 1 | *N* | *Y* |
| NEWQB | 1 | 1 | *Y* | *Y* |
| NO55 | 2 | 1 | *Y* | *Y* |
| PGE2 | 1 | 1 | *Y* | *Y* |
| POCRO | 1 | 1 | *Y* | *Y* |
| QUINOL | 1 | 1 | *Y* | *Y* |
| RIFOLO | 1 | 1 | *Y* | *Y* |
| SALEX | 1 | 1 | *Y* | *Y* |
| SCHWARZ | (25) | 1 | *N* | *Y* |
| SELENID | 20 | 1 | *Y* | *Y* |
| SKN1 | 1 | 1 | *Y* | *Y* |
| SUOA | (1) | (1) | *N* | *N* |
| TPH | 1 | 1 | *Y* | *Y* |
| TUR10 | 1 | 1 | *Y* | *Y* |
| WINTER2 | (24) | (1) | *N* | *N* |

Thus, it is of non-negligible interest to check if the maximum of $S'''$ will characterize with high frequency the correct solution even when the number of negative quartets is relatively small (*i.e.* only a maximum of 750 as for *SIR*92). With this in view, we used a pre-release of *SIR*92 (only triplets and negative quartets employed) for ranking according to $S'''$ the trial solutions produced for all 36 test structures. In Table 2, we give the order number [NORD($S'''$)] of the correct solution. If the solution is not found then NORD($S'''$) refers to $S'''$ as calculated for the published structure and is given in parentheses. If NORD($S'''$) for the published structure is large, it cannot be expected that $S'''$ is a maximum for the trial corresponding to the correct solution possibly found in a non-default run. We see that NORD($S'''$) is often different from unity: in these cases, the tangent formula (III.6) perversely works towards false solutions and the correct solution may sometimes be found because of the constraints dictated by the starting set of phases. If psi-

zero triplets are taken into account, a possible extension of (10) should be

$$S^{IV} = \sum_{\text{triplets}} T_{ijl}\cos t_{ijl} + \sum_{\substack{\text{neg. est.}\\ \text{quartets}}} Q''_{ijlm}\cos q_{ijlm} - \sum_{\mathbf{h}_w}\alpha'_{\mathbf{h}_w}$$
$$\simeq \tfrac{1}{3}\sum_{\mathbf{h}}\alpha_{\mathbf{h}} - \sum_{\mathbf{h}_w}\alpha'_{\mathbf{h}_w}. \qquad (11)$$

Since $\alpha'_{\mathbf{h}_w}$ does not contain the term $R_{\mathbf{h}_w}$ (while $R_{\mathbf{h}}$ is present in $\alpha_{\mathbf{h}}$), we replace (11) by

$$S^{IV} = \tfrac{1}{3}\sum_{\mathbf{h}}\alpha_{\mathbf{h}}/R_{\mathbf{h}} - \sum_{\mathbf{h}_w}\alpha'_{\mathbf{h}_w}.$$

We rank the trial solutions produced by *SIR*92 in order of $S^{IV}$ and we obtain the third column of Table 2, where NORD($S^{IV}$) is the order number of the correct solution. $S^{IV}$ is a much more efficient figure of merit: for all the structures but GRA4, the maximum value of $S^{IV}$ corresponds to the correct solution. Consequently, a tangent formula based on the combined use of $P10$ triplet estimates, negative quartet estimates and psi-zero triplets is expected to be in principle more efficient than (III.6).

### The new tangent formula

The preceding sections suggest that (8) may be the basis for a new tangent formula provided a sound use is made of the terms $A'_{\mathbf{h}_w,\mathbf{k}_q}$. The problem to be solved is how to combine them with the probabilistic terms $G_{\mathbf{h},\mathbf{k}_q}$. In order to obtain some insight into the problem, we selected three from the set of 36 test structures and we calculated, for typical trial solutions and for the true structure, the distribution of the parameter

$$z_{\mathbf{h}_w} = \left|\alpha'_{\mathbf{h}_w} - \langle\alpha'_{\mathbf{h}_w}\rangle\right|/\sigma_{\alpha'_{\mathbf{h}_w}},$$

where $\langle\alpha'_{\mathbf{h}_w}\rangle = (\pi^{1/2}/2)\sigma_{\alpha_{\mathbf{h}_w}}$ is the expected value of $a'_{\mathbf{h}_w}$ in non-centrosymmetric space groups and $\sigma^2_{\alpha'_{\mathbf{h}_w}} = \sum_{j=1}^{r} A'^2_{\mathbf{h}_w,\mathbf{k}_q}$ is the variance (Cascarano, Giacovazzo & Viterbo, 1987).

The outcome is shown in Fig. 1 for the trial solutions and in Fig. 2 for the true one. It is seen that the number of weak reflections with $z_{\mathbf{h}_w}$ larger than a given threshold, for example with $z_{\mathbf{h}_w} > 2$, is a non-negligible percentage for all the trial solutions while it is relatively small when the correct solutions are considered. An efficient phasing procedure should contribute to make the number of weak reflections with $\alpha'_{\mathbf{h}_w}$ much larger than $\langle\alpha'_{\mathbf{h}_w}\rangle$ a small percentage of the number of weak reflections. $A'_{\mathbf{h}_w,\mathbf{k}_q}$ will efficiently contribute to modify the phase value $\theta_{\mathbf{k}_q}$ suggested by $P10$ and by negative quartet estimates only if the ratio $\alpha'_{\mathbf{h}_w}/\sigma_{\alpha'_{\mathbf{h}_w}}$ is large enough or, in other words, if $\alpha'_{\mathbf{h}_w}$ is remarkably larger than its expected value. In this case, the psi-zero terms have a beneficial feedback effect on the values of the phases of the strong reflections contributing to them. We then propose the following weighting scheme for $A'_{\mathbf{h}_w,\mathbf{k}_q}$: the weight $P$ is defined by

$$P(A'_{\mathbf{h}_w,\mathbf{k}_q}) = D_1(z_{\mathbf{h}_w}) \quad \text{when } \alpha'_{\mathbf{h}_w} > \langle\alpha'_{\mathbf{h}_w}\rangle$$
$$P(A'_{\mathbf{h}_w,\mathbf{k}_q}) = 0 \quad \text{otherwise.}$$

$D_1(x) = I_1(x)/I_0(x)$ is the ratio of the two modified Bessel functions of order 1 and 0, respectively.

The proposed tangent formula is therefore

$$\tan\theta_{\mathbf{k}_q} = \left[\sum_{\mathbf{h}} G_{\mathbf{h},\mathbf{k}_q}\sin(\varphi_{\mathbf{h}} - \varphi_{\mathbf{h}-\mathbf{k}_q}) - \sum_{\mathbf{h}_w} P_{\mathbf{h}_w} A'_{\mathbf{h}_w,\mathbf{k}_q}\sin(\theta'_{\mathbf{h}_w} - \varphi_{\mathbf{h}_w-\mathbf{k}_q})\right] \times \left[\sum_{\mathbf{h}} G_{\mathbf{h},\mathbf{k}_q}\cos(\varphi_{\mathbf{h}} - \varphi_{\mathbf{h}-\mathbf{k}_q}) - \sum_{\mathbf{h}_w} P_{\mathbf{h}_w} A'_{\mathbf{h}_w,\mathbf{k}_q}\cos(\theta'_{\mathbf{h}_w} - \varphi_{\mathbf{h}_w-\mathbf{k}_q})\right]^{-1} = T''_{\mathbf{k}_q}/B''_{\mathbf{k}_q} \qquad (12)$$

with reliability parameter

$$\alpha''_{\mathbf{k}_q} = \left(T''^2_{\mathbf{k}_q} + B''^2_{\mathbf{k}_q}\right)^{1/2}.$$

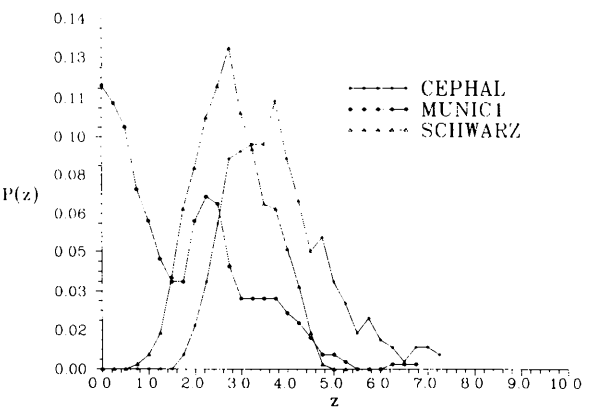


Fig. 1. Distribution of $z$ for typical trial solutions.

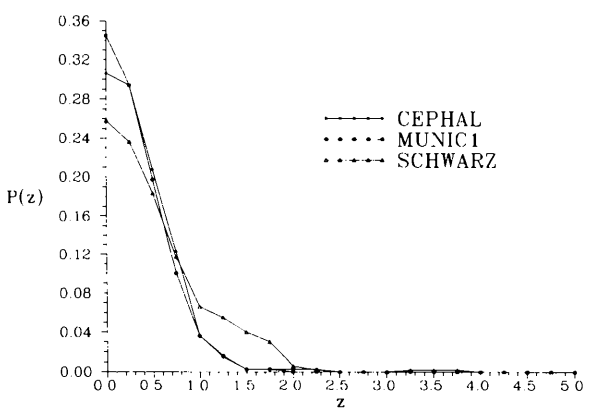


Fig. 2. Distribution of $z$ using true phases.

## Applications

In paper III, a pre-release of *SIR*92 was applied in a default setting (five symbolic phases in the starting set to permute according to the magic-integers approach) to the 36 crystal structures quoted in Table 1 by using in an active way triplets estimated positive and negative by the *P*10 formula and negative quartets. The result is shown in the fourth column of Table 2: *Y* means that the solution was found, *N* that no solution was obtained by the default run. Only five structures, MGHEX, MUNICH1, SCHWARZ, SUOA and WINTER2 were not solved.

*SIR*92 has now been modified in order to allow the active use of the psi-zero triplets both when the magic-integer-permutation method is applied and when the random-starting-set technique (Baggio, Woolfson, Declercq & Germain, 1978) is used. Equation (12) is applied in the phasing process only when the number of determined phases is sufficiently large, that is when a statistically meaningful set of psi-zero triplets is involved in the calculations. The outcome is shown in the last column of Table 2: only MGHEX, SUOA and WINTER2 are still unsolved while MUNICH1 and SCHWARZ are now routinely solved. It is worthwhile noting that MGHEX, SUOA and WINTER2 are easily solved in 200 trials if (12) is combined with the random-starting-set technique.

This success convinced us to modify the default run of *SIR*92 by including the active use of psi-zero triplets.

## Against the loss of enantiomorph

We have seen in the preceding section that in some cases (MUNICH1 and SCHWARZ) the active use of the psi-zero triplets makes the difference between success and failure. Some questions arise:

(1) How does the active use of the psi-zero relationships influence the FOM's? As a general trend, the use of (12) will drive $\alpha'_{\mathbf{h}_w}$ towards $\langle\alpha'_{\mathbf{h}_w}\rangle$ for each $\mathbf{h}'_w$. When this ideal situation is attained, the psi-zero FOM in *SIR*92 (say PSICOMB) will be unity. At the same time, $\alpha_{\mathbf{h}}$ calculated for strong reflections will tend towards $\langle\alpha_{\mathbf{h}}\rangle$ for any $\mathbf{h}$. Thus, the general effect of the active use of psi-zero triplets in the phasing process is the improvement of the FOM's, both for the correct and for the incorrect solutions. This is confirmed in Table 3 where *SIR*92 FOM's for SCHWARZ are shown for the ten solution with the highest values of the combined FOM (CFOM). While in Table 3(*a*) (passive use of psi-zero triplets) no solution is found, in Table 3(*b*) the first four sets are correct solutions. Sets five to ten are false solutions but they have values of CFOM larger than in Table 3(*a*). It may be concluded that (12) always works towards optimizing the FOM's, and sometimes it works so well that the correct solution is found.

(2) Does the active use of psi-zero triplets help against the loss of enantiomorph? In the last column of Table 3(*a*), the average phase value $\langle\Phi_e\rangle$ is given for triplets that are expected to be enantiomorph sensitive (triplets with reliability parameter $G \simeq 0$ when evaluated by the *P*10 formula). In the ideal case, $\langle\Phi_e\rangle$ should be close to 90°: when the enantiomorph is lost for a given trial, the values in column 7 are expected to be close to zero. *SIR*92 signals these events by outputting the values of $\langle\Phi_e\rangle$ when this is lower than 30°. While, in Table 3(*b*), no solution, wrong or correct, loses the enantiomorph, several sets in Table 3(*a*) do.

Table 3. *Comparison of two different phase-determination tests on SCHWARZ*

(*a*) FOM's for ten trial solutions of SCHWARZ with the highest value of CFOM as obtained by *SIR*92 by the active use of triplets and negative quartet relationships

| Set | MABS | ALCOMB | PSCOMB | CPHASE | CFOM | $\langle\Phi_e\rangle$ |
|---|---|---|---|---|---|---|
| 1 | 1.323 | 0.528 | 0.125 | 0.304 | 0.444 | 29.0 |
| 2 | 1.377 | 0.517 | 0.074 | 0.189 | 0.395 | |
| 3 | 1.339 | 0.490 | 0.049 | 0.108 | 0.348 | 29.2 |
| 4 | 1.423 | 0.472 | 0.036 | 0.073 | 0.323 | |
| 5 | 1.393 | 0.465 | 0.046 | 0.083 | 0.322 | |
| 6 | 1.402 | 0.462 | 0.046 | 0.086 | 0.322 | 27.1 |
| 7 | 1.410 | 0.470 | 0.038 | 0.068 | 0.320 | |
| 8 | 1.400 | 0.430 | 0.043 | 0.059 | 0.292 | 24.1 |
| 9 | 1.419 | 0.441 | 0.029 | 0.041 | 0.292 | 27.6 |
| 10 | 1.426 | 0.438 | 0.026 | 0.042 | 0.290 | 27.4 |

(*b*) FOM's for ten trial solutions of SCHWARZ with highest value of CFOM as obtained by *SIR*92 by use of (12)

| Set | MABS | ALCOMB | PSCOMB | CPHASE | CFOM |
|---|---|---|---|---|---|
| 1 | 1.296 | 0.857 | 0.655 | 1.000 | 0.911 |
| 2 | 1.298 | 0.853 | 0.655 | 1.000 | 0.908 |
| 3 | 1.291 | 0.838 | 0.638 | 1.000 | 0.898 |
| 4 | 1.255 | 0.790 | 0.631 | 0.999 | 0.868 |
| 5 | 1.309 | 0.665 | 0.200 | 0.502 | 0.604 |
| 6 | 1.241 | 0.611 | 0.270 | 0.570 | 0.596 |
| 7 | 1.299 | 0.659 | 0.216 | 0.483 | 0.593 |
| 8 | 1.159 | 0.672 | 0.471 | 0.790 | 0.591 |
| 9 | 1.294 | 0.615 | 0.233 | 0.543 | 0.588 |
| 10 | 1.295 | 0.608 | 0.221 | 0.539 | 0.582 |

(3) When both (III.6) and (12) solve the structure, are the phases determined by (12) more accurate than those defined by (III.6)? The analysis of our tests suggests that, in most cases, the phase sets have equivalent quality unless, because of the space group or of some specific pseudocentrosymmetry, the structure shows an evident tendency to lose the enantiomorph. In this case, the active use of the psi-zero triplets curbs this tendency: indeed, just those strong reflections that mainly contribute to fix $\theta'_{\mathbf{h}_w}$ and to generate too high values for $\alpha'_{\mathbf{h}_w}$ are subject in (12) to non-negligible feedback.

An instructive test in this sense is AZET, which has been intensely studied by Lessinger (1976) because of its instability under tangent refinement. Cycles of traditional tangent refinement (Cochran estimates) modify the true phases to values with an average phase error of 36° (partial loss of enantiomorph). If *P*10 triplet estimates and negative quartets are actively used, the final average error is 26° while the additional use of the psi-zero triplets curbs the average error to 21°.

(4) Is the active use of psi-zero triplets always advisable? Often, structures suffering by pseudotranslational symmetry show a bad psi-zero FOM, even for the correct solution. In this case, a phasing process that tries to improve the psi-zero FOM will hinder rather than favour the crystal structure solution. If structure factors are renormalized and triplet reliabilities are re-estimated by using the information on pseudotranslational symmetry as prior (Cascarano, Giacovazzo & Luić, 1987), then the psi-zero FOM is usually better and use of (12) works fine. In *SIR*92, in order to avoid the psi-zero contribution to (12) overcoming the contributions of the strong triplets and the negative quartets, we use a maximum of NLAR/3 weak reflections to construct psi-zero triplets (NLAR is the number of reflections used for constructing strong triplets). In this way, the active use of psi-zero triplets does not hinder the crystal structure solution also for CIME, CUIMID, ERICA, FEGAS, HOV1 and POCRO, which suffer from pseudotranslational symmetry.

## Concluding remarks

The psi-zero triplets have thus far been considered as a tool for calculating a powerful figure of merit (Cochran & Douglas, 1957) for recognizing the correct phase set among numerous trial solutions. In a recent paper, Giacovazzo (1993) proposed their active use in the phasing process: in this paper, we describe the theoretical background necessary for a reasonable active use of the psi-zero triplets and the first applications of the method.

The authors thank Miss C. Chiarella for technical support.

### References

ALTOMARE, A., CASCARANO, G. & GIACOVAZZO, C. (1992). *Acta Cryst.* A**48**, 30–36.
ALTOMARE, A., CASCARANO, G., GIACOVAZZO, C., GUAGLIARDI, A., BURLA, M. C., POLIDORI, G. & CAMALLI, M. (1994). *J. Appl. Cryst.* **27**, 435.
BAGGIO, R., WOOLFSON, M. M., DECLERCQ, J.-P. & GERMAIN, G. (1978). *Acta Cryst.* A**34**, 883–892.
BURLA, M. C., CASCARANO, G. & GIACOVAZZO, C. (1992). *Acta Cryst.* A**48**, 906–912.
CASCARANO, G., GIACOVAZZO, C., CAMALLI, M., SPAGNA, R., BURLA, M. C., NUNZI, A. & POLIDORI, G. (1984). *Acta Cryst.* A**40**, 278–283.
CASCARANO, G., GIACOVAZZO, C. & LUIĆ, M. (1987). *Acta Cryst.* A**43**, 14–22.
CASCARANO, G., GIACOVAZZO, C., MOLITERNI, A. G. G. & POLIDORI, G. (1994). *Acta Cryst.* A**50**, 22–27.
CASCARANO, G., GIACOVAZZO, C. & VITERBO, D. (1987). *Acta Cryst.* A**47**, 22–29.
COCHRAN, W. (1952). *Acta Cryst.* **5**, 65–67.
COCHRAN, W. & DOUGLAS, A. S. (1957). *Proc. R. Soc. London Ser. A*, **243**, 281.
GIACOVAZZO, C. (1993). *Z. Kristallogr.* **206**, 161–171.
GIACOVAZZO, C., BURLA, M. C. & CASCARANO, G. (1992). *Acta Cryst.* A**48**, 901–906.
KARLE, J. & HAUPTMAN, H. (1956). *Acta Cryst.* **9**, 635–651.
LESSINGER, L. (1976). *Acta Cryst.* A**32**, 538–550.

*Acta Cryst.* (1995). A**51**, 825–830

# X-ray Diffraction on Fibonacci Superlattices

BY P. MIKULÍK, V. HOLÝ AND J. KUBĚNA

*Department of Solid State Physics, Faculty of Science, Masaryk University, Kotlářská* 2, 611 37 *Brno, Czech Republic*

AND K. PLOOG

*Paul-Drude-Institut für Festkörperelektronik, Hausvogteiplatz* 5–7, 1086 *Berlin, Germany*

(*Received* 15 *July* 1994; *accepted* 3 *March* 1995)

## Abstract

The exact diffraction curve of the Fibonacci superlattice is calculated using the semi-kinematical approximation of dynamical X-ray diffraction. The properties of the discrete Fourier transform of quasiperiodically arranged layers are employed to derive explicit approximate formulae for the diffracted intensity and the angular positions of peaks. The exact and approximate curves are compared by a numerical simulation and a good agreement is found. The measurement of the diffraction curve was performed on the generalized Fibonacci superlattice built by stacked Fibonacci generations. This superlattice belongs to the same class of local isomorphism as the Fibonacci superlattice if both are infinitely thick. The explicit approximate formulae enabled the fitting of the structural parameters of the superlattice even in the low-resolution experimental set-up when the fitting of the whole measured diffraction curve was not possible.