

## Direct Phase Determination of Triple Products from Bijvoet Inequalities. III. An Improved Double Patterson Approach

BY W. M. G. F. PONTENAGEL,\* H. KRABBENDAM AND J. KROON

*Laboratorium voor Kristal- en Structuurchemie, Rijksuniversiteit Utrecht, Padualaan 9,  
3584 CH Utrecht, The Netherlands*

(Received 11 November 1985; accepted 11 June 1986)

### Abstract

The scope of the application of Bijvoet differences to the direct determination of triple-product phases is extended to the general case of an arbitrary number of non-identical anomalous scatterers. Apart from an improved procedure for the estimation of the hyperaxis contributions of the double Patterson function to its Fourier coefficients, an additional estimate of spatial contributions is described. For small and medium-sized structures the adopted approach is expected to lead to accurate results. The application to the structure determination of calcium tri lactate trihydrate is given as an example.

### Introduction

Kroon, Spek & Krabbendam (1977) proposed that a complex double Patterson function (with triple products as Fourier coefficients) could be looked on as being a 'structure' containing a number of 'anomalous scatterers', of which the one with the largest scattering power is situated at the origin. Therefore the problem of finding *a priori* triple-product phases from Bijvoet inequalities is analogous to the problem of finding structure-factor phases from the same experimental data when not all anomalous scatterers are localized. Following the method introduced simultaneously by Peerdeman & Bijvoet (1956), Ramachandran & Raman (1956) and Okaya & Pepinsky (1956), Kroon, Spek & Krabbendam (1977) obtained the following expression:

$$\sin \bar{\varphi}_{\mathbf{h},\mathbf{k}} = \frac{|\tau_{\mathbf{h},\mathbf{k}}|^2 - |\tau_{-\mathbf{h},-\mathbf{k}}|^2}{4\tau''_{\mathbf{h},\mathbf{k}}[\frac{1}{2}(|\tau_{\mathbf{h},\mathbf{k}}|^2 + |\tau_{-\mathbf{h},-\mathbf{k}}|^2) - |\tau''_{\mathbf{h},\mathbf{k}}|^2]^{1/2}} \quad (1)$$

where  $\bar{\varphi}_{\mathbf{h},\mathbf{k}} = \frac{1}{2}(\varphi_{\mathbf{h},\mathbf{k}} - \varphi_{-\mathbf{h},-\mathbf{k}})$ ,  $\varphi_{\mathbf{h},\mathbf{k}}$  is the phase of the triple product  $\tau_{\mathbf{h},\mathbf{k}} = F_{\mathbf{h}}F_{\mathbf{k}}F_{-\mathbf{h}-\mathbf{k}}$  and  $\tau''_{\mathbf{h},\mathbf{k}}$  is a contribution to the imaginary part of  $\tau_{\mathbf{h},\mathbf{k}}$  originating from anomalous scattering. To obtain satisfactory results a scaling procedure appeared to be necessary in (1) in case  $\tau''_{\mathbf{h},\mathbf{k}}$  is estimated only from the origin contribution of the double Patterson function. In a subsequent paper, Heinerman, Krabbendam, Kroon & Spek (1978) showed that better results could be obtained

if more characteristics of the double Patterson function were taken into account. Their estimation of double Patterson hyperaxis contributions was based on the resemblance of the hypersections with the known Patterson function.

In this paper we shall report (1) more general expressions for the estimation of hyperaxis contributions, (2) a procedure for the estimation of spatial contributions of the double Patterson function to  $\tau''_{\mathbf{h},\mathbf{k}}$  and (3) the application of the results to the X-ray data of a *P1* structure.

### Theory

For a *P1* structure containing anomalous scatterers, the normalized structure factor can be written as

$$E_{\mathbf{h}} = \alpha_{\mathbf{h}}^{-1/2} \sum_{i=1}^N |f_i(\mathbf{h})| \exp [2\pi i \delta_i(\mathbf{h})] \exp (2\pi i \mathbf{h} \cdot \mathbf{r}_i) \quad (2)$$

where  $|f_i(\mathbf{h})| \exp [2\pi i \delta_i(\mathbf{h})] = f_i^0(\mathbf{h}) + f_i' + i f_i''$  is the complex scattering factor of the *i*th atom (of which  $\mathbf{r}_i$  is the position vector), *N* is the total number of atoms in the unit cell and

$$\alpha_{\mathbf{h}} = \sum_{i=1}^N |f_i(\mathbf{h})|^2.$$

The symbols *h*, *k* and *h + k* will be used hereafter to denote either  $\mathbf{h}$ ,  $\mathbf{k}$  and  $\mathbf{h} + \mathbf{k}$  or  $-\mathbf{h}$ ,  $-\mathbf{k}$  and  $-\mathbf{h} - \mathbf{k}$  respectively. After defining

$$a_{ijm} \equiv (\alpha_{\mathbf{h}} \alpha_{\mathbf{k}} \alpha_{\mathbf{h}+\mathbf{k}})^{-1/2} |f_i(\mathbf{h}) f_j(\mathbf{k}) f_m(\mathbf{h} + \mathbf{k})| \times \cos \{2\pi [\delta_i(\mathbf{h}) + \delta_j(\mathbf{k}) + \delta_m(\mathbf{h} + \mathbf{k})]\} \quad (3)$$

and analogously  $b_{ijm}$  with the cosine replaced by a sine, we obtain for a normalized triple product:

$$\tau_{\mathbf{h},\mathbf{k}} \equiv E_{\mathbf{h}} E_{\mathbf{k}} E_{-\mathbf{h}-\mathbf{k}} = O_{\mathbf{R}} + i O_{\mathbf{I}} + S_{\mathbf{R}} + i S_{\mathbf{I}} \quad (4)$$

where

$$O_{\mathbf{R}} = \sum_{i=1}^N a_{iii}, \quad O_{\mathbf{I}} = \sum_{i=1}^N b_{iii}$$

$$S_{\mathbf{R}} = \sum_{i=1}^N \sum_{j=1}^N \sum_{\substack{m=1 \\ \text{not } i=j=m}}^N a_{ijm} \exp \{2\pi i [\mathbf{h} \cdot (\mathbf{r}_i - \mathbf{r}_m) + \mathbf{k} \cdot (\mathbf{r}_j - \mathbf{r}_m)]\}$$

\* Present address: DSM Research, Postbus 18, 6160 MD Geleen, The Netherlands.

and

$$S_I = \sum_{i=1}^N \sum_{\substack{j=1 \\ \text{not } i=j=m}}^N \sum_{m=1}^N b_{ijm} \exp \{2\pi i[\mathbf{h} \cdot (\mathbf{r}_i - \mathbf{r}_m) + \mathbf{k} \cdot (\mathbf{r}_j - \mathbf{r}_m)]\}.$$

Using the same definitions we easily obtain for the Friedel-related triple product

$$\tau_{-h,-k} \equiv E_{-h} E_{-k} E_{h+k} = O_R + iO_I + S_R^* + iS_I^*. \quad (5)$$

If the imaginary part of  $S_I$  is neglected [ $\text{Im}(S_I) = 0$ ], it follows from (4) and (5) that  $\tau_{h,k} - \tau_{-h,-k}^* = 2i[O_I + \text{Re}(S_I)]$ , which is a purely imaginary quantity. Therefore  $\tau_{h,k}$  and  $\tau_{-h,-k}^*$  will form a triangle in the Argand diagram with one side parallel to the imaginary axis. Under these conditions (1) can be applied with  $\tau_{h,k}'' = O_I + \text{Re}(S_I)$  leading to two possible values for  $\varphi_{h,k}$ ; the  $\varphi$ ,  $(\pi - \varphi)$  ambiguity may be resolved by choosing the angle nearest to zero. In the sequel we will show how to obtain satisfactory estimates for  $\tau_{h,k}''$ .

Recently Hauptman (1982) and Giacovazzo (1983) also integrated anomalous diffraction with the techniques of direct methods. Following the suggestion of Heinerman *et al.* (1978), they first derived the joint probability distribution (j.p.d.) of the six normalized structure factors  $E_h$ ,  $E_k$ ,  $E_{-h-k}$ ,  $E_{-h}$ ,  $E_{-k}$  and  $E_{h+k}$ . After fixing the magnitudes of the structure factors and integrating with respect to all the phases, subject to the condition  $\varphi_h + \varphi_k + \varphi_{-h-k} \equiv \varphi_{h,k}$ , they both obtained an expression for the conditional probability distribution of the triple-product phase  $\varphi_{h,k}$ , given the magnitudes of the six structure factors mentioned before. Both authors arrive at the same von Mises distribution of which the mode can be anywhere between zero and  $2\pi$ . This is in sharp contrast to the results of our approach which can only yield phase indications in the first or fourth quadrant of the Argand diagram.

It remains to be seen whether the approximations made in both cases account for the differences, or whether the use of six individual structure-factor amplitudes instead of two triple-product amplitudes is an essential improvement, as suggested by Hauptman.

#### Estimation of $\text{Re}(S_I)$

From (4) we obtain

$$\text{Re}(S_I) = H_h + H_k + H_{h+k} + S_{hk} \quad (6)$$

where

$$H_h = \sum_i \sum_j' b_{ij} \cos [2\pi \mathbf{h} \cdot (\mathbf{r}_i - \mathbf{r}_j)]$$

$$H_k = \sum_i \sum_j' b_{ij} \cos [2\pi \mathbf{k} \cdot (\mathbf{r}_i - \mathbf{r}_j)]$$

$$H_{h+k} = \sum_i \sum_j' b_{ij} \cos [2\pi (\mathbf{h} + \mathbf{k}) \cdot (\mathbf{r}_i - \mathbf{r}_j)]$$

$$S_{hk} = \sum_i \sum_j \sum_m' b_{ijm} \cos \{2\pi [\mathbf{h} \cdot (\mathbf{r}_i - \mathbf{r}_m) + \mathbf{k} \cdot (\mathbf{r}_j - \mathbf{r}_m)]\}$$

and where

$$\sum_i \sum_j' = \sum_{i=1}^N \sum_{\substack{j=1 \\ i \neq j}}^N$$

and

$$\sum_i \sum_j \sum_m' = \sum_{i=1}^N \sum_{\substack{j=1 \\ i \neq j \wedge i \neq m \wedge j \neq m}}^N \sum_{m=1}^N.$$

#### (A) Hyperaxis contributions

Following Heinerman *et al.* (1978),  $H_h$ ,  $H_k$  and  $H_{h+k}$  in (6) can be estimated from the Fourier coefficients of an origin-removed Patterson function from which the effects of anomalous dispersion are eliminated. For example,

$$H_h \approx \frac{1}{2} A_h (|E_h|^2 + |E_{-h}|^2 - 2) \quad (7)$$

where  $A_h$  is a scaling constant which can be determined by a least-squares analysis.  $H_k$  and  $H_{h+k}$  are estimated in the same way. The calculations are given in Appendix I and the result is

$$A_h = \frac{\sum_i \sum_j' d_{ij}^{(h)} b_{ij}}{\sum_i \sum_j' [d_{ij}^{(h)}]^2} \quad (8)$$

where  $d_{ij}^{(h)} = \alpha_h^{-1} |f_i(h) f_j(h)| \cos \{2\pi [\delta_i(h) - \delta_j(h)]\}$ .

For  $A_k$  and  $A_{h+k}$  we find analogous expressions with  $d_{ij}^{(h)}$  replaced by  $d_{ij}^{(k)}$  and  $d_{ij}^{(h+k)}$  respectively and  $b_{ij}$  replaced by  $b_{ji}$  and  $b_{ij}$  respectively. It must be noted that, contrary to the results of Heinerman *et al.* (1978), three distinct scaling factors are needed for each triple product, one for each hyperaxis. The main reason for this difference is that in the current approach it is possible to include non-identical anomalous scatterers besides non-identical lighter atoms. This implies the estimation of a complete sum of cosine terms instead of a few individual cosines and ultimately leads to scaling factors which depend, not only on the contents of the unit cell, but also on  $|\mathbf{h}|$ ,  $|\mathbf{k}|$  and  $|\mathbf{h} + \mathbf{k}|$ .

#### (B) Spatial contributions

The threefold summation  $S_{hk}$  in (6) corresponds to contributions from spatial peaks in the double Patterson function. In all previous work this part of  $\tau_{h,k}''$  was neglected, but in Appendix II it is shown that  $S_{hk}$  can also be estimated. The procedure is based on the fact that  $\langle |E_h E_k E_{-h-k}| \cos \varphi_{h,k} \rangle$  and  $\langle |E_{-h} E_{-k} E_{h+k}| \cos \varphi_{-h,-k} \rangle$  are known if both triple-product phases are assumed to be distributed according to the Cochran distribution (Cochran, 1955). The

result obtained in Appendix II is

$$\begin{aligned}
 S_{hk} \approx & B_{hk} \left[ \frac{1}{2} |E_h E_k E_{-h-k}| I_1(\kappa_1) / I_0(\kappa_1) \right. \\
 & + \frac{1}{2} |E_{-h} E_{-k} E_{h+k}| I_1(\kappa_2) / I_0(\kappa_2) \\
 & - O_R - \frac{1}{2} C_h (|E_h|^2 + |E_{-h}|^2 - 2) \\
 & - \frac{1}{2} C_k (|E_k|^2 + |E_{-k}|^2 - 2) \\
 & \left. - \frac{1}{2} C_{h+k} (|E_{h+k}|^2 + |E_{-h-k}|^2 - 2) \right] \quad (9)
 \end{aligned}$$

where  $I_1$  and  $I_0$  are modified Bessel functions of order one and zero respectively.  $\kappa_1$  and  $\kappa_2$  are the  $\kappa$  values of the two Friedel-related triple products,  $C_h$ ,  $C_k$  and  $C_{h+k}$  are least-squares scaling constants obtained in the same way as  $A_h$ ,  $A_k$  and  $A_{h+k}$  in the previous section and  $B_{hk}$  is a scaling constant obtained *via* a least-squares analysis on (9), which results in

$$B_{hk} = \frac{\sum_i \sum_j \sum_m' b_{ijm} a_{ijm}}{\sum_i \sum_j \sum_m' a_{ijm}^2}. \quad (10)$$

The expressions for  $C_h$ ,  $C_k$  and  $C_{h+k}$  can be obtained from Appendix II.

#### Reliability of triple-product phases from sine invariants

Unique triple-product phase indications may be obtained from the approximate values of the sine invariants by choosing the phase angle nearest to zero. Qualitatively this approach may be seen as a combination of the  $\sin \varphi$  formula with the Cochran distribution function. The reliability of the resulting phase indication is of special interest. Apart from the more obvious sources of errors like non-exact experimental data and the statistical nature of the determination of  $\tau''_{h,k}$ , the quality of the sine invariants will be determined by the validity of the approximation  $\text{Im}(S_I) = 0$ . On the basis of the analogy between  $S_I$  and  $S_R$  we deduce from  $b_{ijm} \ll a_{ijm}$  that the conditions for the triangle construction will be reasonably well fulfilled if  $\text{Im}(S_R)$  is small. This indicates that the best results will be obtained for those triple products for which the Cochran distribution predicts a phase angle near to zero, the more so as in those cases the ambiguity will be resolved with more confidence. The usual  $\kappa$  values ( $2\sigma_3\sigma_2^{-3/2}|EEE|$ ) may therefore be used to rank the phase indications according to their reliability, provided the estimation of  $\text{Re}(S_I)$  is included in the phase calculation.

#### Practical results and discussion

For the structure determination of calcium trisectate trihydrate ( $\text{CaC}_9\text{H}_{16}\text{O}_9 \cdot 3\text{H}_2\text{O}$ ; space group  $P1$  with  $a = 5.823$ ,  $b = 8.295$ ,  $c = 8.975$  Å;  $\alpha = 78.39$ ,  $\beta = 73.61$  and  $\gamma = 75.19^\circ$ ,  $Z = 1$ ), the full details of which will be published elsewhere (Kanters, Jansma & Pontenagel, 1986), triple-product phases were estimated

according to the procedure given in this paper. After  $L_p$  correction of all 2982 observed intensities within the Cu  $K\alpha$  limiting sphere (1545 of the  $hkl$  type and 1437 of the type  $\bar{h}\bar{k}l$ ), a scale and temperature factor were obtained from a Wilson plot and the complete set of  $F$  values was normalized to  $E$  values. Real and imaginary parts of anomalous-dispersion corrections for  $\text{Ca}^{2+}$  and O were taken from *International Tables for X-ray Crystallography* (1974):  $f'_{\text{Ca}} = 0.3$ ,  $f''_{\text{Ca}} = 1.4$ ,  $f'_\text{O} = 0$ ,  $f''_\text{O} = 0.1$ . The 188 largest  $E$  values of the  $hkl$  type were transferred to the *MULTAN78* routine *SIGTWO* (Main, Hull, Lessinger, Germain, Declercq & Woolfson, 1978) resulting in 1112 pairs of Friedel-related triple products with  $\kappa$  values ranging from 1.58 to 6.03. Subsequently, three different sets of  $\sin \varphi_{h,k}^{\text{calc}}$  values were derived with the  $\sin \varphi$  formula:

- case 1:  $\tau''_{h,k}$  estimated from origin contributions of the double Patterson only; *i.e.*  $\tau''_{h,k} = O_I$ ;
- case 2:  $\tau''_{h,k}$  estimated from origin and hyperaxis contributions; *i.e.*  $\tau''_{h,k} = O_I + H_h + H_k + H_{h+k}$ ;
- case 3:  $\tau''_{h,k}$  estimated from origin, hyperaxis and spatial contributions; *i.e.*  $\tau''_{h,k} = O_I + H_h + H_k + H_{h+k} + S_{hk}$ .

Sines calculated larger than 1 and smaller than  $-1$  were set equal to 1 and  $-1$  respectively.  $\varphi_{h,k}^{\text{calc}}$  were obtained by choosing the phase angle nearest to zero. In all calculations the 'observed' normalized structure factors were used. In those cases where  $|E_{-h}|$  was not available from experiment,  $|E_h|$  was used instead of  $|E_{-h}|$ . All calculated triple-product phases were compared with their  $\varphi_{h,k}^{\text{true}}$  values which were obtained by averaging  $\varphi_{h,k}^{\text{true}}$  and  $-\varphi_{-h,-k}^{\text{true}}$ . The latter values were calculated from the final atomic coordinates as obtained from the structure determination (in the correct enantiomorph).

To show the influence of the different estimates for  $\text{Re}(S_I)$  we refer to Table 1 where it can be seen that employing extra information about the double Patterson function does lead to more accurate triple-product phases.

In case 1 most of the absolute values of  $\sin \varphi_{h,k}^{\text{calc}}$  appeared to be larger than 1, leading to an excessive number of phases at  $90$  and  $270^\circ$ . The accuracy of these results leaves much to be desired, although the majority of the phases are calculated in the correct quadrant of the Argand diagram. This already indicates that an extra estimate for  $\text{Re}(S_I)$ , which can be considered as an individual scaling for each sine invariant, will lead to better triple-product phases. Comparison of the results of case 2 with those of case 1 shows that this is indeed the case. The extra estimates for  $S_{hk}$  in case 3 clearly improve the  $\varphi_{h,k}^{\text{calc}}$  values even more, as can be seen from Table 1.

The triple-product phases are calculated slightly too close to zero on the average:  $\langle |\varphi_{h,k}^{\text{calc}}| \rangle = 17^\circ$  while  $\langle |\varphi_{h,k}^{\text{true}}| \rangle = 24^\circ$ . From (1) we see that such a bias may

be caused by the application of a wrong scale and/or temperature factor.

The results of Table 1 were obtained by averaging over all 1112 triple products. In order to show that the Cochran  $\kappa$  value can be used to rank the calculated phases according to their reliability, we refer to Table 2 where it can be seen that triple products with larger  $\kappa$  values are calculated more accurately on the average if the estimates of hyperaxis and spatial contributions to  $\tau''_{h,k}$  are taken into account.

### Concluding remarks

The use of 'Bijvoet differences' between Friedel-related triple products leads to estimates for the sine invariants. By combining these results with conventional direct methods (*i.e.* the Cochran distribution) we are able to obtain unique phase indications for triple products anywhere between  $-\frac{1}{2}\pi$  and  $+\frac{1}{2}\pi$ . The restriction to the right-hand side of the Argand diagram is a direct consequence of the fact that application of conventional direct methods to triple products leads to positive estimates for the cosine invariants. Anticipating future developments, we therefore expect that application of our procedure to quartet and higher-order invariants will lead to unique phase indications anywhere between 0 and  $2\pi$ , provided the relevant cross terms are taken into account.

As to the reliability of the results, we have to stress that the present procedure is of an algebraic nature, in which the assumption  $\text{Im}(\mathbf{S}_I) = 0$  and the derived estimations of  $\text{Re}(\mathbf{S}_I)$  play a central role. In order to obtain an estimate of the variances, one ought to calculate the j.p.d.  $P(\mathbf{S}_I, \mathbf{S}_R)$  but such a calculation is outside the scope of this paper.

Meanwhile it seems reasonable to expect that the results will be more accurate if (i)  $N$  is small, (ii) a small number of powerful anomalous scatterers is present and (iii) the Bijvoet differences are large.

As a rule we expect our method to fail for large structures, not only because the  $\varphi$ ,  $(\pi - \varphi)$  ambiguity will not be resolvable satisfactorily ( $O_R/|\mathbf{S}_R| \rightarrow 0$ ), but also because the approximation  $\text{Im}(\mathbf{S}_I) = 0$  will be less adequate ( $O_I/|\mathbf{S}_I| \rightarrow 0$ ). Preliminary calculations with the X-ray data of the protein ferredoxin (Sieker, Adman & Jensen, 1972) confirm these ideas. For small and medium-sized structures, however, the procedure leads to accurate triple-product phases even if observed X-ray data are used. Although the mean error of  $23^\circ$  for the test structure (see Table 1) does not seem to be an improvement compared with the case where all triple-product phases are assumed to be zero ( $\langle |\bar{\varphi}_{h,k}^{\text{true}}| \rangle = 24^\circ$ ), we stress that the current phase indications are extremely important for two reasons:

(1) A subsequent determination of structure-factor phases will be greatly facilitated as no false symmetry is expected to be introduced. The systematic errors

Table 1. Comparison of the three sets of  $\bar{\varphi}_{h,k}^{\text{calc}}$  values

Case 1:  $\text{Re}(\mathbf{S}_I) = 0$ ; case 2:  $\text{Re}(\mathbf{S}_I) = H_h + H_k + H_{h+k}$ ; case 3:  
 $\text{Re}(\mathbf{S}_I) = H_h + H_k + H_{h+k} + S_{hk}$ .

	Case 1	Case 2	Case 3
$\langle  \Delta\bar{\varphi}_{h,k}  \rangle^*$ ( $^\circ$ )	67	31	23
$\langle  \bar{\varphi}_{h,k}^{\text{calc}}  \rangle$ ( $^\circ$ )	75	31	17

$$* \langle |\Delta\bar{\varphi}_{h,k}| \rangle = \langle |\bar{\varphi}_{h,k}^{\text{calc}} - \bar{\varphi}_{h,k}^{\text{true}}| \rangle.$$

of the all-zero estimates for the triple-product phases, as normally obtained from the Cochran distribution, are replaced by random errors if the phase indications are obtained from the measured Bijvoet differences.

(2) Most of the triple-product phases correspond to the correct enantiomorph thus breaking the twofold ambiguity in the solution of the phase problem, *i.e.* the structure determination will eventually lead directly to the correct enantiomorph, thus serving as an extra test to confirm the absolute configuration.

## APPENDIX I

### Estimation of hyperaxis contributions to $\text{Re}(\mathbf{S}_I)$

From (2) we obtain

$$|E_h|^2 - 1 = \alpha_h^{-1} \sum_i \sum_j' |f_i(h)f_j(h)| \times \cos \{2\pi[\delta_i(h) - \delta_j(h) + \mathbf{h} \cdot (\mathbf{r}_i - \mathbf{r}_j)]\} \quad (\text{I.1})$$

and

$$|E_{-h}|^2 - 1 = \alpha_h^{-1} \sum_i \sum_j' |f_i(h)f_j(h)| \times \cos \{2\pi[\delta_i(h) - \delta_j(h) - \mathbf{h} \cdot (\mathbf{r}_i - \mathbf{r}_j)]\}. \quad (\text{I.2})$$

So

$$\frac{1}{2}(|E_h|^2 + |E_{-h}|^2) - 1 = \sum_i \sum_j' d_{ij}^{(h)} \cos [2\pi\mathbf{h} \cdot (\mathbf{r}_i - \mathbf{r}_j)] \quad (\text{I.3})$$

where  $d_{ij}^{(h)} = \alpha_h^{-1} |f_i(h)f_j(h)| \cos \{2\pi[\delta_i(h) - \delta_j(h)]\}$ .

The right-hand side of (I.3) is analogous to  $H_h$  in (6) of the main text. Therefore we can estimate

$$H_h \approx A_h \left[ \frac{1}{2}(|E_h|^2 + |E_{-h}|^2) - 1 \right]. \quad (\text{I.4})$$

The scaling constant  $A_h$  is determined by minimizing the expected value of the squared difference between the left- and right-hand sides of (I.4):

$$\frac{\delta}{\delta A_h} \left\langle \left\{ \sum_i \sum_j' [b_{ij} - A_h d_{ij}^{(h)}] \cos [2\pi\mathbf{h} \cdot (\mathbf{r}_i - \mathbf{r}_j)] \right\}^2 \right\rangle = 0 \quad (\text{I.5})$$

where the averages are performed with respect to the

Table 2. Comparison of triple-product phases ( $^\circ$ ) for two different sets of  $\kappa$  values

	Case 1			Case 2		Case 3	
	$\langle  \bar{\varphi}_{\mathbf{h},\mathbf{k}}^{\text{true}}  \rangle$	$\langle  \Delta\bar{\varphi}_{\mathbf{h},\mathbf{k}}  \rangle$	$\langle  \bar{\varphi}_{\mathbf{h},\mathbf{k}}^{\text{calc}}  \rangle$	$\langle  \Delta\bar{\varphi}_{\mathbf{h},\mathbf{k}}  \rangle$	$\langle  \bar{\varphi}_{\mathbf{h},\mathbf{k}}^{\text{calc}}  \rangle$	$\langle  \Delta\bar{\varphi}_{\mathbf{h},\mathbf{k}}  \rangle$	$\langle  \bar{\varphi}_{\mathbf{h},\mathbf{k}}^{\text{calc}}  \rangle$
$3.2 < \kappa < 3.6$	22	64	74	27	28	20	14
$5.2 < \kappa < 6.0$	16	70	75	24	25	15	12

atomic position vectors:

$$\frac{\delta}{\delta A_h} \left\{ \sum_i \sum_j \sum_l \sum_m [b_{ij} b_{ilm} - A_h b_{ij} d_{lm}^{(h)} - A_h b_{ilm} d_{ij}^{(h)} + A_h^2 d_{ij}^{(h)} d_{lm}^{(h)}] \times \frac{1}{2} [\langle \cos 2\pi \mathbf{h} \cdot (\mathbf{r}_i - \mathbf{r}_j + \mathbf{r}_l - \mathbf{r}_m) \rangle + \langle \cos 2\pi \mathbf{h} \cdot (\mathbf{r}_i - \mathbf{r}_j - \mathbf{r}_l + \mathbf{r}_m) \rangle] \right\} = 0. \quad (\text{I.6})$$

The two averages in (I.6) only lead to permitted non-zero contributions for  $i=m$  and  $j=l$  and for  $i=l$  and  $j=m$  respectively. Therefore we obtain

$$\frac{1}{2} \frac{\delta}{\delta A_h} \left\{ \sum_i \sum_j [b_{ij} b_{ij} + b_{ij} b_{ji} - A_h b_{ij} d_{ij}^{(h)} - A_h b_{ij} d_{ji}^{(h)} - A_h b_{ji} d_{ij}^{(h)} + A_h^2 d_{ij}^{(h)} d_{ij}^{(h)} + A_h^2 d_{ji}^{(h)} d_{ji}^{(h)}] \right\} = 0. \quad (\text{I.7})$$

As  $d_{ij}^{(h)} = d_{ji}^{(h)}$  this leads to

$$\frac{1}{2} \frac{\delta}{\delta A_h} \left\{ \sum_i \sum_j [b_{ij}^2 + b_{ij} b_{ji}] - 4A_h \sum_i \sum_j b_{ij} d_{ij}^{(h)} + 2A_h^2 \sum_i \sum_j [d_{ij}^{(h)}]^2 \right\} = 0. \quad (\text{I.8})$$

Finally we obtain for the best value of the scaling constant

$$A_h = \frac{\sum_i \sum_j b_{ij} d_{ij}^{(h)}}{\sum_i \sum_j [d_{ij}^{(h)}]^2}. \quad (\text{I.9})$$

## APPENDIX II

### Estimation of spatial contributions to $\text{Re}(S_r)$

From (4) and (5) we obtain

$$\begin{aligned} & \frac{1}{2} |E_{\mathbf{h}} E_{\mathbf{k}} E_{-\mathbf{h}-\mathbf{k}}| \cos \varphi_{\mathbf{h},\mathbf{k}} + \frac{1}{2} |E_{-\mathbf{h}} E_{-\mathbf{k}} E_{\mathbf{h}+\mathbf{k}}| \cos \varphi_{-\mathbf{h},-\mathbf{k}} \\ &= \sum_i a_{iii} + \sum_{\substack{i,j,m \\ \text{not } i=j=m}} a_{ijm} \cos 2\pi [\mathbf{h} \cdot (\mathbf{r}_i - \mathbf{r}_m) + \mathbf{k} \cdot (\mathbf{r}_j - \mathbf{r}_m)] \\ &= \sum_i a_{iii} + \sum_i \sum_j a_{ijj} \cos 2\pi \mathbf{h} \cdot (\mathbf{r}_i - \mathbf{r}_j) \\ & \quad + \sum_i \sum_j a_{iji} \cos 2\pi \mathbf{k} \cdot (\mathbf{r}_i - \mathbf{r}_j) \\ & \quad + \sum_i \sum_j a_{ijj} \cos 2\pi (\mathbf{h} + \mathbf{k}) \cdot (\mathbf{r}_i - \mathbf{r}_j) \\ & \quad + \sum_i \sum_j \sum_m a_{ijm} \cos 2\pi [\mathbf{h} \cdot (\mathbf{r}_i - \mathbf{r}_m) + \mathbf{k} \cdot (\mathbf{r}_j - \mathbf{r}_m)]. \end{aligned} \quad (\text{II.1})$$

$\sum_i a_{iii} = O_R$  and is known from the contents of the unit cell. The three double summations can be written as  $\frac{1}{2} C_h (|E_{\mathbf{h}}|^2 + |E_{-\mathbf{h}}|^2 - 2)$ ,  $\frac{1}{2} C_k (|E_{\mathbf{k}}|^2 + |E_{-\mathbf{k}}|^2 - 2)$  and  $\frac{1}{2} C_{h+k} (|E_{\mathbf{h}+\mathbf{k}}|^2 + |E_{-\mathbf{h}-\mathbf{k}}|^2 - 2)$ , respectively, after application of the procedure given in Appendix I. Since the effects of anomalous dispersion are small, we assume that both  $\varphi_{\mathbf{h},\mathbf{k}}$  and  $\varphi_{-\mathbf{h},-\mathbf{k}}$  are distributed according to the Cochran distribution (Cochran, 1955); after replacing the left-hand side of (II.1) by its expected value we obtain an approximate expression for the threefold summation:

$$\begin{aligned} & \sum_i \sum_j \sum_m a_{ijm} \cos 2\pi [\mathbf{h} \cdot (\mathbf{r}_i - \mathbf{r}_m) + \mathbf{k} \cdot (\mathbf{r}_j - \mathbf{r}_m)] \\ &= \frac{1}{2} |E_{\mathbf{h}} E_{\mathbf{k}} E_{-\mathbf{h}-\mathbf{k}}| I_1(\kappa_1) / I_0(\kappa_1) \\ & \quad + \frac{1}{2} |E_{-\mathbf{h}} E_{-\mathbf{k}} E_{\mathbf{h}+\mathbf{k}}| I_1(\kappa_2) / I_0(\kappa_2) \\ & \quad - O_R - \frac{1}{2} C_h (|E_{\mathbf{h}}|^2 + |E_{-\mathbf{h}}|^2 - 2) \\ & \quad - \frac{1}{2} C_k (|E_{\mathbf{k}}|^2 + |E_{-\mathbf{k}}|^2 - 2) \\ & \quad - \frac{1}{2} C_{h+k} (|E_{\mathbf{h}+\mathbf{k}}|^2 + |E_{-\mathbf{h}-\mathbf{k}}|^2 - 2) \end{aligned} \quad (\text{II.2})$$

where  $I_1$  and  $I_0$  are modified Bessel functions of order one and zero respectively,  $\kappa_1$  and  $\kappa_2$  are the  $\kappa$  values of the two Friedel-related triple products, and

$$C_h = \frac{\sum_i \sum_j a_{ijj} d_{ij}^{(h)}}{\sum_i \sum_j [d_{ij}^{(h)}]^2}.$$

$C_k$  and  $C_{h+k}$  are analogous to  $C_h$  with  $a_{ijj}$  replaced by  $a_{jji}$  and  $a_{ijj}$  respectively and  $d_{ij}^{(h)}$  replaced by  $d_{ij}^{(k)}$  and  $d_{ij}^{(h+k)}$  respectively.

With (II.2) we are able to obtain an estimate for  $S_{hk}$  in (6) of the main text:

$$S_{hk} \approx B_{hk} [\text{'right-hand side of (II.2)}]. \quad (\text{II.3})$$

Analogously to the procedure given in Appendix I,  $B_{hk}$  can be found by minimizing

$$\left\{ \sum_i \sum_j \sum_m [b_{ijm} - B_{hk} a_{ijm}] \cos 2\pi [\mathbf{h} \cdot (\mathbf{r}_i - \mathbf{r}_m) + \mathbf{k} \cdot (\mathbf{r}_j - \mathbf{r}_m)] \right\}^2. \quad (\text{II.4})$$

The calculations are straightforward and the result is

$$B_{hk} = \frac{\sum_i \sum_j \sum_m b_{ijm} a_{ijm}}{\sum_i \sum_j \sum_m a_{ijm}^2} \quad (\text{II.5})$$

in which all double summations over the atomic labels are neglected.

## References

- COCHRAN, W. (1955). *Acta Cryst.* **8**, 473-478.  
 GIACOVAZZO, C. (1983). *Acta Cryst.* **A39**, 585-592.  
 HAUPTMAN, H. (1982). *Acta Cryst.* **A38**, 632-641.  
 HEINERMAN, J. J. L., KRABBENDAM, H., KROON, J. & SPEK, A. L. (1978). *Acta Cryst.* **A34**, 447-450.  
*International Tables for X-ray Crystallography* (1974). Vol. IV. Birmingham: Kynoch Press. (Present distributor D. Reidel, Dordrecht.)  
 KANTERS, J. A., JANSMA, S. & PONTENAGEL, W. M. G. F. (1986). In preparation.  
 KROON, J., SPEK, A. L. & KRABBENDAM, H. (1977). *Acta Cryst.* **A33**, 382-385.  
 MAIN, P., HULL, S. E., LESSINGER, L., GERMAIN, G., DECLERCQ, J.-P. & WOOLFSON, M. M. (1978). *MULTAN78. A System of Computer Programs for the Automatic Solution of Crystal Structures from X-ray Diffraction Data*. Univs. of York, England, and Louvain, Belgium.  
 OKAYA, Y. & PEPINSKY, R. (1956). *Phys. Rev.* **103**, 1645-1647.  
 PEERDEMAN, A. F. & BIJVOET, J. M. (1956). *Proc. K. Ned. Akad. Wet.* **B59**, 312-313.  
 RAMACHANDRAN, G. N. & RAMAN, S. (1956). *Curr. Sci.* **25**, 348-351.  
 SIEKER, L. C., ADMAN, E. T. & JENSEN, L. H. (1972). *Nature (London)*, **235**, 40-42.

*Acta Cryst.* (1987). **A43**, 60-64

**REDUC: An Automated Procedure for the Determination of Structure-Factor Phases from the Estimated Values of Structure (Sem)invariant Phase Combinations**

BY W. M. G. F. PONTENAGEL,\* H. KRABBENDAM AND J. KROON

*Laboratorium voor Kristal- en Structuurchemie, Rijksuniversiteit Utrecht, Padualaan 8, 3584 CH Utrecht, The Netherlands*

(Received 11 November 1986; accepted 11 June 1986)

**Abstract**

An algorithm is described to determine structure-factor phases from the estimated phases of (sem)invariants of any order, in a mixed mode if desired. The method essentially consists of a reduction of a redundant set of linear equations by successive elimination of unknowns. The main result of the procedure is a set of mutually independent equations in which structure-factor phases are expressed as linear combinations of a limited number of unknowns (among which a suitable set of origin-defining phases must exist), and (sem)invariant phases of which (reliable) estimates are available. Application to an unknown  $P2_1$  structure using triple products only, followed by a tangent refinement, is given. From further tests it appeared that the method is preeminently expedient as a troubleshooter in intricate structure determinations.

**Introduction**

For small and medium-sized structures, the number of seminvariant phases ( $\Phi$ 's) which can be determined by direct methods with a reasonable accuracy outpaces by far the number of structure-factor phases ( $\varphi$ 's) to be determined for the calculation of

an interpretable  $E$  map. As seminvariants can actually be seen as linear combinations of structure-factor phases, a matrix inversion may be thought of as an appropriate procedure to determine structure-factor phases. For centrosymmetric problems this approach was first introduced by Cochran & Douglas (1955) and later amended by Vand & Pepinsky (1956). In essence that technique consists of two parts: (i) select a set of mutually independent relations from which the  $\varphi$ 's can be obtained *via* a matrix inversion and (ii) employ the relations that have not been used in a sensible way as a check on the plausibility of the obtained solution.

Fortier, Fronckowiak, Smith, Hauptman & De Titta (1978) and Fortier, De Titta, Fronckowiak, Smith & Hauptman (1979) described a procedure which resembles the matrix-inversion method, but differs from it in the fact that the distinction between the steps (i) and (ii) has been removed. It consists of the reduction of a redundant set of special linear equations by successive elimination of unknowns. Starting with an arbitrary set of  $M$  seminvariants we find the following symbolic representation:

$$\sum_k \varphi_k \equiv \Phi_s \quad (s = 1, 2, \dots, M). \quad (1)$$

Successive elimination of structure-factor phases and repeated substitutions in all previously obtained expressions leads in centrosymmetric space groups

\* Present address: DSM Research, Postbus 18, 6160 MD Geleen, The Netherlands.