

## Estimation of Triplets with Interatomic Vectors

BY M. J. KRONENBURG

Laboratory for Crystallography, University of Amsterdam, Nieuwe Achtergracht 166, 1018 WV Amsterdam, The Netherlands

(Received 23 February 1993; accepted 5 May 1993)

### Abstract

This paper investigates the exponential triplet probability distribution and the corresponding expectation values that incorporate structural information of an *a priori* known part of the electron density. The expectation values improve from the Cochran expectation values (no structural information) to the true values (all information known). Combination of the expectation values gives an estimation formula for triplets with structural information. The validity of the resulting estimated triplets is tested with mean absolute errors and correlation coefficients against the true values of the triplets. The single and double Patterson functions may serve to find the structural information *ab initio*. Some problems are mentioned that mean application of the formulae to standard structure-solution methods is still difficult. The development of more general formulae for other multiplets and higher Patterson functions is expected.

### 1. Introduction

The use of an *a priori* uniform independent distribution of atoms in the unit cell, which is an assumption in the derivation of the well known exponential multiplet distributions (Hauptman, 1976), seems to contrast with the structural information known from the Patterson function (Patterson, 1935). From the Patterson function, triplets may be estimated in a completely different way (Hauptman & Karle, 1962; Kroon & Krabbendam, 1970). Various kinds of *a priori* information may be used in the derivation of the exponential triplet probability distribution (Heinerman, 1977). In this paper, the known fragment has a fixed orientation and all other atoms are considered to be randomly distributed. The connection with Lagrange multipliers (Kronenburg, Peschar & Schenk, 1991) is stressed.

In the resulting expectation values, the peaks of the single and double Patterson functions (Sayre, 1953) of the known fragment are present. This suggests the possibility of combining Patterson functions with probabilistic methods.

### 2. Estimation of triplets with structural information

The irreducible cluster integrals for structure factors in space group *P1* were derived in an earlier paper (Kronenburg *et al.*, 1991). It was demonstrated there that the multiplet distributions (Hauptman, 1976) follow from the first irreducible cluster integral  $\varepsilon_1$ , where interatomic correlations between pairs of atoms are taken into account. The second irreducible cluster integral  $\varepsilon_2$  accounts for correlations between triples of atoms. It contributes an extra exponential term to the exponential triplet probability distribution and changes the values of the Lagrange multipliers.

The structure factors  $F_\mu$  in space group *P1* are given by

$$F_\mu = \sum_{\nu=1}^N f_\nu \exp(ik_\mu \cdot \mathbf{r}_\nu), \quad (1)$$

where the  $f_\nu$  are the scattering factors, the  $\mathbf{r}_\nu$  are the atomic positions in the unit cell,  $N$  is the total number of atoms in the unit cell and the  $\mathbf{k}_\mu$  are the reciprocal vectors. For a single crystal, the  $\mathbf{k}_\mu$  are restricted to the reciprocal point lattice:

$$\mathbf{k}_\mu \subset 2\pi(h_1\mathbf{a}^* + h_2\mathbf{b}^* + h_3\mathbf{c}^*), \quad (2)$$

where  $h_1, \dots, h_3$  are integers and  $\mathbf{a}^*, \dots, \mathbf{c}^*$  are the reciprocal-unit-cell vectors. The  $\sigma_n$  values are the power sums over the scattering factors:

$$\sigma_n = \sum_{\nu=1}^N f_\nu^n. \quad (3)$$

The normalized structure factors may be introduced by substituting  $F_\mu = \sigma_2^{1/2} E_\mu$ .

Evaluation of formula (5.2) in the paper by Kronenburg *et al.* (1991), with  $n_{\mu s}$  now nine indices [ $\mu \in 1\dots 3$  is the index over the three reciprocal vectors that form the triplet and  $s \in 1\dots 3$  is the index over a triple of atoms in the known part of the structure], yields the second irreducible cluster integral. This contributes an extra term in the exponential triplet probability distribution [with  $\mu$  and  $\mu'$  elements of  $(1\dots 3)$ ,  $\mu \neq \mu'$ ,  $\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 = \mathbf{0}$ ,

and  $\gamma_1(2x) \approx x$ ]:

$$P(\psi_i | |F_\mu|^{\text{obs}}, \mathbf{k}_\mu) \propto \exp \left( 2\beta_1\beta_2\beta_3 |F_1F_2F_3|^{\text{obs}} \times \left\{ \sigma_3 \cos \psi_i + \sum_{\nu \neq \kappa \neq \lambda} f_\nu f_\kappa f_\lambda \times \cos[\psi_i - \mathbf{k}_\mu \cdot (\mathbf{r}_\kappa - \mathbf{r}_\nu) - \mathbf{k}_\mu \cdot (\mathbf{r}_\lambda - \mathbf{r}_\nu)] \right\} \right), \quad (4)$$

where the triple summation over  $\nu$ ,  $\kappa$  and  $\lambda$  is over the triangles of the  $N_{\text{known}}$  atoms of the known part of the structure.  $|F_\mu|^{\text{obs}}$  are the observed diffraction magnitudes and  $\psi_i = \arg(F_1) + \arg(F_2) + \arg(F_3)$  is the triplet-phase-sum random variable. For the number of known atoms,  $0 \leq N_{\text{known}} \leq N$ . Following the derivation (Kronenburg *et al.*, 1991) of the Lagrange multipliers  $\beta_\mu$ :

$$\beta_\mu = \left\{ \sigma_2 + \sum_{\nu \neq \kappa} f_\nu f_\kappa \cos[\mathbf{k}_\mu \cdot (\mathbf{r}_\kappa - \mathbf{r}_\nu)] \right\}^{-1}, \quad (5)$$

where the double summation over  $\nu$  and  $\kappa$  is over the pairs of the known part of the structure. The result in the paper by Heinerman (1977) for a single known group with known orientation is equivalent. Since the expressions contain only interatomic vectors, the choice of origin is arbitrary. This is in accordance with the fact that the triplet  $\psi_i$  is invariant under origin translation.

The interatomic vectors in (5) must also be present in (4). This means that triangles of interatomic vectors should be known; interatomic vectors that are not correlated have no effect on the exponential triplet distribution. This also follows from the irreducible cluster integrals (Kronenburg *et al.*, 1991). How these triangles are to be found is a different question and is discussed in the last section. When a group of  $N_{\text{known}}$  positions of atoms is known, all pairs and triangles of atoms may be summed separately; they are correlated anyway.

The expectation values  $\langle \cos \psi_i \rangle$  and  $\langle \sin \psi_i \rangle$  follow from (4):

$$\langle \cos \psi_i \rangle = \gamma_1 \left( 2\beta_1\beta_2\beta_3 |F_1F_2F_3|^{\text{obs}} \times \left\{ \sigma_3 + \sum_{\nu \neq \kappa \neq \lambda} f_\nu f_\kappa f_\lambda \times \cos[\mathbf{k}_\mu \cdot (\mathbf{r}_\kappa - \mathbf{r}_\nu) + \mathbf{k}_\mu \cdot (\mathbf{r}_\lambda - \mathbf{r}_\nu)] \right\} \right), \quad (6)$$

$$\langle \sin \psi_i \rangle = \gamma_1 \left( 2\beta_1\beta_2\beta_3 |F_1F_2F_3|^{\text{obs}} \times \left\{ \sum_{\nu \neq \kappa \neq \lambda} f_\nu f_\kappa f_\lambda \right. \right.$$

$$\left. \left. \times \sin[\mathbf{k}_\mu \cdot (\mathbf{r}_\kappa - \mathbf{r}_\nu) + \mathbf{k}_\mu \cdot (\mathbf{r}_\lambda - \mathbf{r}_\nu)] \right\} \right), \quad (7)$$

where  $\gamma_n(x) = I_n(x)/I_0(x)$ , the  $I_n(x)$  are the modified Bessel functions and the  $\beta_\mu$  are given by (5).

Using expectation values (6) and (7), the following estimations of the triplets result:

$$\psi_i^{\text{est}} = \arg(\langle \cos \psi_i \rangle + i \langle \sin \psi_i \rangle) \quad (8)$$

There are two important special cases: the case where no structural information is known and the case where all structural information is known. In the case of the triplet invariant, this structural information consists only of triangles of interatomic vectors.

When no structural information is known, (5) reduces to  $\beta_\mu = \sigma_2^{-1}$  and (4) reduces to the Cochran result, with  $2\sigma_3\sigma_2^{-3}|F_1F_2F_3|^{\text{obs}} \cos \psi_i$  in the exponent. In this case, (6) reduces to the Cochran expectation value  $\langle \cos \psi_i \rangle = \gamma_1(2\sigma_3\sigma_2^{-3}|F_1F_2F_3|^{\text{obs}})$  and (7) reduces to  $\langle \sin \psi_i \rangle = 0$ . It follows that, in this case, (8) reduces to  $\psi_i^{\text{est}} = 0$ ; this also follows from the Cochran distribution, which always has its maximum at  $\psi_i = 0$ .

When all structural information is known, (5) reduces to  $\beta_\mu = (|F_\mu|^{\text{true}})^{-2}$  and, with  $|F_\mu|^{\text{true}} = |F_\mu|^{\text{obs}}$ , (4) reduces to  $2\cos(\psi_i - \psi_i^{\text{true}})$  in the exponent. In this case, (6) reduces to  $\langle \cos \psi_i \rangle = \gamma_1(2\cos \psi_i^{\text{true}})$  and (7) reduces to  $\langle \sin \psi_i \rangle = \gamma_1(2\sin \psi_i^{\text{true}})$ . It follows that, in this case, (8) reduces to  $\psi_i^{\text{est}} \approx \psi_i^{\text{true}}$ .

### 3. Test results and conclusions

In order to verify the validity of the expectation values (6) and (7) and of the estimated values (8), some random structures were generated. The total number of atoms of these random structures is specified by  $N$  and an *a priori* known part of  $N_{\text{known}}$  atoms was selected. The number of terms in the double summation of (5) thus becomes  $2 \binom{N_{\text{known}}}{2}$  and the number of terms in the triple summation of (4) becomes  $6 \binom{N_{\text{known}}}{3}$ , where  $\binom{N_{\text{known}}}{2}$  is the number of pairs and  $\binom{N_{\text{known}}}{3}$  is the number of triangles in the known part. The factors two and six result from the numbers of (combinations of) interatomic vectors present in a pair and in a triangle, respectively.

In Table 1, the mean absolute errors and the correlation coefficients are given for true and expected cosine (Karle, 1972) and sine values of the triplets and for the true and estimated triplet values for a random structure with a varying number of known atoms and varying numbers of strongest triplets. The mean absolute error  $\Delta$  is a measure of the mean distance of the data points (see the figures) from the linear identity line. This means that  $\Delta$  decreases when the expectation values of the triplets are more accurate. The correlation coefficient  $\Gamma$  is a

Table 1. Comparison of true triplets with estimated triplets for total number of atoms  $N = 100$  with different numbers of atoms  $N_{\text{known}}$  included in expressions (6), (7) and (8)

# $T$  is the number of strongest triplets included,  $T_{\text{min}}$  is the underlimit of the triple product  $|E_1 E_2 E_3|/N^{1/2}$ ,  $\Delta = \overline{|x - y|}$  are the mean absolute errors and  $\Gamma = (\overline{xy} - \overline{x}\overline{y})/[(\overline{x^2} - \overline{x}^2)(\overline{y^2} - \overline{y}^2)]^{-1/2}$  are the correlation coefficients, where the  $x$  are the true values and the  $y$  the expectation values or estimated values.

$N_{\text{known}}$	$\Delta_{(6)}$	$\Gamma_{(6)}$	$\Delta_{(7)}$	$\Gamma_{(7)}$	$\Delta_{(8)}$	$\Gamma_{(8)}$	$\Delta_{(6)}$	$\Gamma_{(6)}$	$\Delta_{(7)}$	$\Gamma_{(7)}$	$\Delta_{(8)}$	$\Gamma_{(8)}$
# $T = 50$ , $T_{\text{min}} = 0.682$						# $T = 200$ , $T_{\text{min}} = 0.516$						
0	0.338	-0.123					0.417	0.118				
5	0.337	-0.110	0.498	0.214	0.639	0.217	0.417	0.115	0.560	0.046	0.839	-0.005
10	0.327	0.062	0.501	0.115	0.644	0.117	0.412	0.171	0.569	0.005	0.859	0.041
15	0.350	0.009	0.475	0.323	0.658	0.280	0.408	0.155	0.551	0.174	0.841	0.219
20	0.375	0.053	0.447	0.447	0.696	0.306	0.403	0.223	0.498	0.362	0.803	0.324
30	0.361	0.318	0.385	0.599	0.616	0.608	0.408	0.346	0.438	0.530	0.821	0.413
40	0.363	0.552	0.372	0.675	0.617	0.708	0.338	0.593	0.413	0.602	0.749	0.495
50	0.291	0.560	0.344	0.736	0.547	0.712	0.304	0.687	0.368	0.702	0.664	0.584
70	0.220	0.701	0.226	0.880	0.327	0.879	0.233	0.837	0.250	0.866	0.409	0.796
100	0.187	0.989	0.094	0.995	0.059	0.997	0.170	0.994	0.110	0.996	0.058	0.998
# $T = 500$ , $T_{\text{min}} = 0.423$						# $T = 2000$ , $T_{\text{min}} = 0.300$						
0	0.470	0.158					0.532	0.113				
5	0.470	0.149	0.584	0.072	0.954	0.011	0.532	0.116	0.603	0.035	1.088	0.021
10	0.466	0.193	0.584	0.062	0.961	0.047	0.528	0.153	0.602	0.059	1.090	0.055
15	0.467	0.153	0.573	0.143	0.978	0.100	0.523	0.183	0.593	0.134	1.089	0.119
20	0.447	0.310	0.533	0.319	0.925	0.226	0.513	0.261	0.572	0.249	1.077	0.181
30	0.428	0.400	0.475	0.491	0.941	0.307	0.488	0.364	0.527	0.405	1.071	0.282
40	0.363	0.596	0.422	0.600	0.796	0.492	0.431	0.536	0.466	0.547	0.936	0.430
50	0.327	0.682	0.372	0.691	0.724	0.559	0.383	0.647	0.421	0.632	0.881	0.480
70	0.253	0.828	0.260	0.859	0.444	0.795	0.289	0.815	0.303	0.816	0.612	0.642
100	0.159	0.994	0.121	0.996	0.059	0.998	0.153	0.995	0.128	0.996	0.061	0.999

measure of the spread of the data points along the identity line. This means that  $\Gamma$  increases when positive and negative triples are better distinguished by the expectation values. In general, both  $\Delta$  and  $\Gamma$  are useful for comparison of expectation values with actual values. When the data points are all on the identity line (perfect correlation), then  $\Delta = 0$  and  $\Gamma = 1$  unless all data points are equal.

For  $N_{\text{known}} = 0$ , the expectation values and estimations correspond to the Cochran (1955) expression.

In Table 1, the  $\psi_i^{\text{est}}$  values show a continuously increasing value of the correlation coefficient  $\Gamma$  with increasing *a priori* information. From this it is concluded that the  $\psi_i^{\text{est}}$  values (8) incorporate the *a priori* information of  $N_{\text{known}}$  atoms well. The mean absolute error  $\Delta$  only decreases when already most *a priori* information is present. This means that  $\Delta$  is not a good measure for indicating the correctness of triplet values where *a priori* information is concerned.

The figures show the true and expected or estimated values of an  $N = 100$  atoms random structure, with a varying number of  $N_{\text{known}}$  known atoms.

Fig. 1 gives the expected and true cosines of the triplet values when  $N_{\text{known}} = 0$ , which yields the Cochran result. Expected and true sines of the triplet values for  $N_{\text{known}} = 0$  would yield all data points on the  $x$  axis. This is also the case for estimated and true triplet values from (8).

Figs. 2, 3 and 4 give the results of (6), (7) and (8), respectively, for the same 100-atom random structure, but with  $N_{\text{known}} = 40$  known atoms. In contrast with the cosines in Fig. 2, the sines in Fig. 3 give a more or less symmetric distribution of data points.

Combination of Figs. 2 and 3 with (8) gives Fig. 4. Although most data points in Fig. 4 tend towards the identity line, some triplets are still wrongly estimated. This means that incorporation of partial *a*

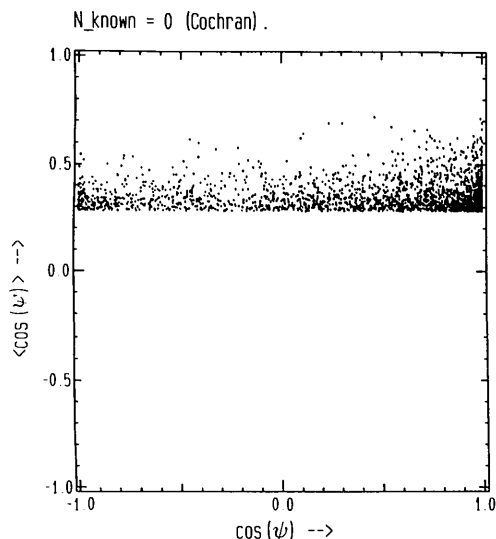


Fig. 1. Equation (6) calculated from random structure,  $N = 100$  and  $N_{\text{known}} = 0$  (Cochran result, see text and Table 1), 2000 strongest triplets.

*priori* information does not guarantee improvement of all estimations, but only an improvement averaged over many estimations. Combination of the information of Figs. 2 and 3 gives a better correlation in Fig. 4. Also, the form of the  $\gamma_1$  functions in (6) and (7) is reflected in Figs. 2 and 3 and not in Fig. 4.

When the positions of all atoms are known ( $N_{\text{known}} = 100$ ), Fig. 5 shows the form of the  $\gamma_1$  function in (6). The figure for the sine values is similar. Combination of cosine and sine values with (8) gives Fig. 6. The fluctuation also follows from the  $\gamma_1$  functions in (6) and (7). This fluctuation, however, is very small and indicates that the change of the combination of (6), (7) and (8) from statistical to exact

identities is tenable. When all structural information concerning the interatomic triangles of the complete structure is present, the statistical identities reduce to almost exact identities.

#### 4. Connection with Pattersons functions

In the preceding sections and in the figures, the *a priori* information was calculated from a part of a randomly generated structure. In this section, the Patterson functions are considered for finding *ab initio* the appropriate structural information for improving estimations of triplet invariants. The single Patterson function, which can be computed

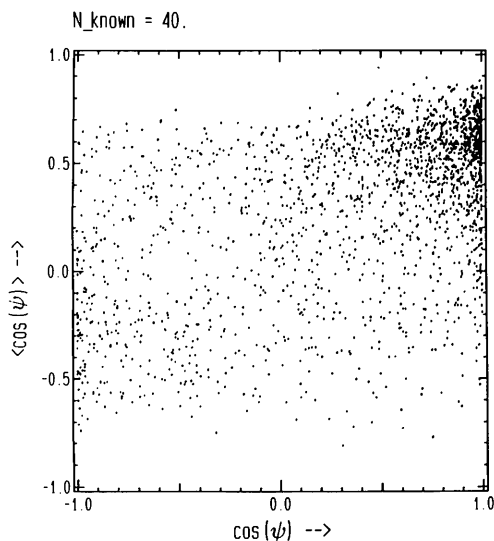


Fig. 2. Equation (6) calculated from random structure,  $N = 100$  and  $N_{\text{known}} = 40$  (see text and Table 1), 2000 strongest triplets.

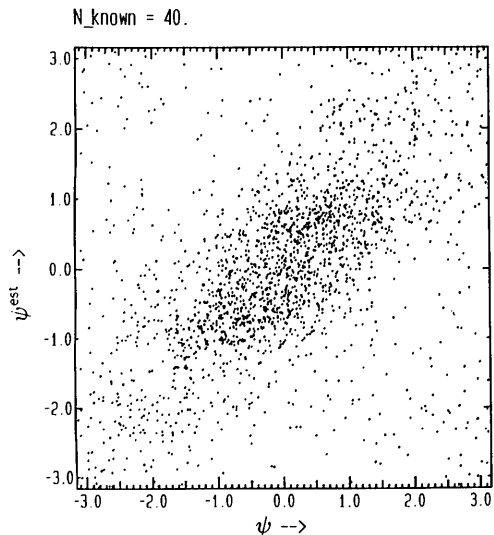


Fig. 4. Equation (8) calculated from random structure,  $N = 100$  and  $N_{\text{known}} = 40$  (see text and Table 1), 2000 strongest triplets.

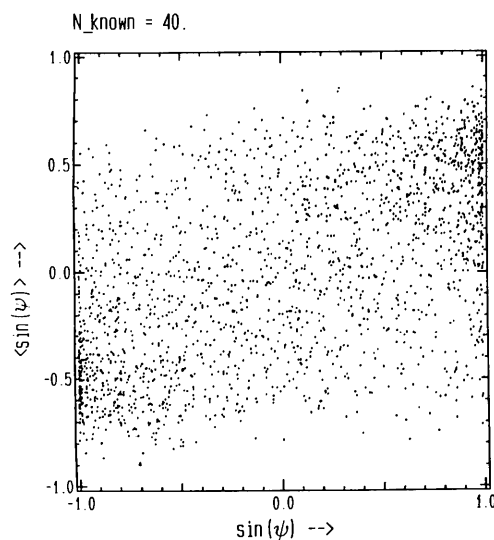


Fig. 3. Equation (7) calculated from random structure,  $N = 100$  and  $N_{\text{known}} = 40$  (see text and Table 1), 2000 strongest triplets.

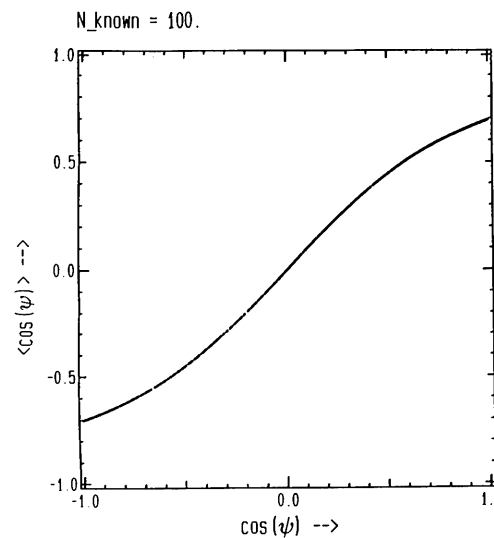


Fig. 5. Equation (6) calculated from random structure,  $N = 100$  and  $N_{\text{known}} = 100$  (see text and Table 1), 2000 strongest triplets.

from observed magnitudes only and may be considered as *a priori* information, contains (possibly overlapping) peaks of interatomic vectors of the complete true structure. However, in the equations of the preceding sections, only interatomic vectors of triangles of the known part of the structure are to be substituted. These triangles are not known *a priori* and involve the true double Patterson function, which can only be computed when the true triplet values are known. The double summation in (5) is also a summation over the peaks in the single Patterson function of the known part of the structure and the triple summation in (4), (6) and (7) is also a summation over the peaks in the double Patterson function of the known part of the structure. In the following, the single and double Patterson functions are considered in connection with the exponential triplet probability distribution (4), the corresponding expectation values, (6) and (7), and the estimated triplets, (8). The following single Patterson function is normalized on the height of the Patterson peaks:

$$\rho_1(\mathbf{u}) = [1/(M+1)] \sum_{\mu=0}^M (|F_{\mu}|^{\text{obs}})^2 \exp(-i\mathbf{k}_{\mu} \cdot \mathbf{u}), \quad (9)$$

where  $M$  is the number of reflections taken into account. For  $\mu=0$ ,  $\mathbf{k}_0=\mathbf{0}$  and  $F_0=\sigma_1$ , by definition. At the origin  $\mathbf{u}=\mathbf{0}$ , a peak of height  $\sigma_2$  is always present.

The single Patterson function does not fully indicate which pairs of interatomic distances form triangles. The  $\binom{N}{2}[\binom{N}{2}-1]$  possible combinations of pairs of interatomic vectors contain only  $2\binom{N}{3}$  pairs that actually form triangles in the true structure. The overlap of peaks in the single Patterson function may be an additional difficulty when finding the unique set of triangles.

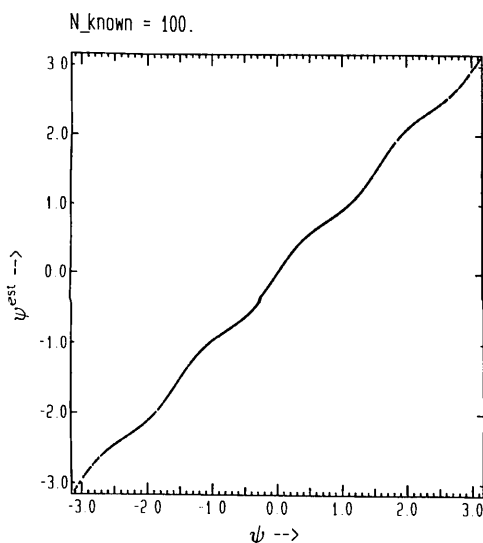


Fig. 6. Equation (8) calculated from random structure,  $N=100$  and  $N_{\text{known}}=100$  (see text and Table 1), 2000 strongest triplets.

The double Patterson function  $\rho_2(\mathbf{u},\mathbf{v})$  (Sayre, 1953), which is a function of two interatomic vectors  $\mathbf{u}$  and  $\mathbf{v}$  and which is normalized on peak heights, has the following form (with  $\mathbf{k}_{\mu_i} + \mathbf{k}_{\mu'_i} + \mathbf{k}_{\mu''_i} = \mathbf{0}$  for each triplet  $i$ ):

$$\rho_2(\mathbf{u},\mathbf{v}) = [1/(T+1)] \sum_{i=0}^T |F_{\mu_i} F_{\mu'_i} F_{\mu''_i}|^{\text{obs}} \times \exp [i(\psi_i - \mathbf{k}_{\mu_i} \cdot \mathbf{u} - \mathbf{k}_{\mu'_i} \cdot \mathbf{v})], \quad (10)$$

where  $T$  is the number of triplets taken into account. The order of the three indices  $\mu_i$ ,  $\mu'_i$  and  $\mu''_i$  that form the triplet is arbitrary because  $\mathbf{u}$  and  $\mathbf{u}-\mathbf{v}$  form the same triangle. Again, for  $i=0$ ,  $\mu_i = \mu'_i = \mu''_i = 0$  and  $F_0 = \sigma_1$ , by definition. At the origin  $\mathbf{u}=\mathbf{v}=\mathbf{0}$ , there is always a peak of height  $\sigma_3$ . Because, for triplets,  $\mathbf{k}_{\mu_i} \neq \mathbf{0}$ ,  $\mathbf{k}_{\mu'_i} \neq \mathbf{0}$  and  $\mathbf{k}_{\mu''_i} \neq \mathbf{0}$ , the sections  $\mathbf{u}=\mathbf{0}$ ,  $\mathbf{v}=\mathbf{0}$  and  $\mathbf{u}=\mathbf{v}$  cannot contain significant peaks when the triplet values  $\psi_i$  have the correct values.

The number of triplets  $T$  is the total number of triplets in the data set of  $M$  reflections.

For each triangle in the actual structure, six peaks in the double Patterson function are present. These peaks are  $(\mathbf{u},\mathbf{v})$ ,  $(\mathbf{v},\mathbf{u})$ ,  $(-\mathbf{u},\mathbf{v}-\mathbf{u})$ ,  $(\mathbf{v}-\mathbf{u},-\mathbf{u})$ ,  $(-\mathbf{v},\mathbf{u}-\mathbf{v})$  and  $(\mathbf{u}-\mathbf{v},-\mathbf{v})$ .

When such a combination of peaks is identified, they may be substituted directly into (6) and (7). There may be overlap in the double Patterson function because of more equally oriented equal triangles in the actual structure. Because of the form of expressions (6) and (7), the peak height can, nevertheless, be substituted in these expressions as a single peak  $f_{\nu} f_{\kappa} f_{\lambda}$ . This may be an important observation for use of the Patterson functions with these formulae. It means that overlap is not an obstacle for application of the double Patterson function to these expressions.

For computation of the double Patterson function (10), some values for the triplets  $\psi_i$  that are not known *a priori* must be substituted. Some starting point is needed, in which only the observed structure-factor magnitudes are present. In earlier work (Hauptman & Karle, 1962), the single Patterson function was used to estimate triplet values. The single Patterson function can be computed from observed structure-factor magnitudes only by (9). The basic assumption for this procedure is the following approximation.

$$\rho_2(\mathbf{u},\mathbf{v}) \approx \sigma_3 \sigma_2^{-1} [\rho_1(\mathbf{u}) \rho_1(\mathbf{v}) \rho_1(\mathbf{u}-\mathbf{v})]^{1/2}. \quad (11)$$

The sections  $\mathbf{u}=\mathbf{0}$ ,  $\mathbf{v}=\mathbf{0}$  and  $\mathbf{u}=\mathbf{v}$  of this estimate of  $\rho_2(\mathbf{u},\mathbf{v})$  contain the single Patterson function. As mentioned, these sections are zero in the double Patterson function with only triplets (10) and should be made zero explicitly after computation of (11).

From this first estimated double Patterson function, a number of highest peaks may be selected

excluding the three sections  $\mathbf{u} = \mathbf{0}$ ,  $\mathbf{v} = \mathbf{0}$  and  $\mathbf{u} = \mathbf{v}$  as mentioned. Some problems, however, need further attention.

The identification of peaks in the (estimated) double Patterson function may be executed in several ways. Using more strong triplets results in more reliably identified peaks. As mentioned, all triplets may be included in the double Patterson function (10). The number of identified peaks is determined by some criterion in the identification method, such as height and sharpness of the peak. This criterion should be strengthened as more peaks are identified, leading to better estimated triplets.

When a number of peaks in the double Patterson function are identified, the corresponding triangles in the true structure may have interatomic vectors in common. By this partial overlap, the number of terms in the expression for the  $\beta_\mu$  values, (5), decreases. In other words, a collection of triangles from the double Patterson function must be assembled into some group of interatomic vectors in order to compute the  $\beta_\mu$  values with (5). It is not yet clear how this assembling of triangles of atoms into partially overlapping triangles can be done. It is expected that for this higher Patterson functions and higher multiplets are needed. Also, the single Patterson function may serve as an indication for this partial overlap of triangles.

Although the true double Patterson function fixes the enantiomorph, the approximated one, (11), does not do so. For this approximated double Patterson function, there are pairs of peaks  $(\mathbf{u}, \mathbf{v})$  and  $(-\mathbf{u}, -\mathbf{v})$ . It may be assumed that by a choice of only one of the two peaks (and the other corresponding five

peaks as mentioned above), the enantiomorph may be fixed by the first choice of the strongest non-origin peak. In the triplet expectation values computed by (8), this choice of enantiomorph will also yield triplet values in which the enantiomorph is fixed up to some accuracy; this also holds for the recomputed double Patterson function (10). From that moment on, of a pair of peaks  $(\mathbf{u}, \mathbf{v})$  and  $(-\mathbf{u}, -\mathbf{v})$ , one will be larger than the other and the largest peak is, of course, to be preferred. In this way, by choice of the first strongest peak, the enantiomorph ambiguity may be solved.

The problems mentioned make application of the formulae presented to standard structure-solution methods difficult. Higher-order Patterson functions (Giacovazzo, 1980; Vaughan, 1958) may also be considered in order to solve these problems.

The author thanks Dr R. Peschar for many useful discussions.

#### References

- COCHRAN, W. (1955). *Acta Cryst.* **8**, 473–478.  
 GIACOVAZZO, C. (1980). *Direct Methods in Crystallography*. London: Academic.  
 HAUPTMAN, H. (1976). *Acta Cryst.* **A32**, 877–882.  
 HAUPTMAN, H. & KARLE, J. (1962). *Acta Cryst.* **15**, 547–551.  
 HEINERMAN, J. J. L. (1977). *Acta Cryst.* **A33**, 100–106.  
 KARLE, J. (1972). *Acta Cryst.* **B28**, 3362–3369.  
 KRONENBURG, M. J., PESCHAR, R. & SCHENK, H. (1991). *Acta Cryst.* **A47**, 469–480.  
 KROON, J. & KRABBENDAM, H. (1970). *Acta Cryst.* **B26**, 312–314.  
 PATTERSON, A. L. (1935). *Z. Kristallogr.* **90**, 517–542.  
 SAYRE, D. (1953). *Acta Cryst.* **6**, 430–432.  
 VAUGHAN, P. A. (1958). *Acta Cryst.* **11**, 111–115.

*Acta Cryst.* (1993). **A49**, 877–880

## Correction for the Dynamical Electron Diffraction Effect in Crystal Structure Analysis

BY SHA BING-DONG, FAN HAI-FU AND LI FANG-HUA

*Institute of Physics, Chinese Academy of Sciences, Beijing 100080, People's Republic of China*

(Received 19 March 1993; accepted 6 May 1993)

### Abstract

A method is proposed to correct for the dynamical electron diffraction effect in crystal structure analysis. A rough structure model is first obtained by conventional structure-analysis methods neglecting the dynamical diffraction effect. From the rough structure model, multislice calculations are used to estimate the crystal thickness through the observed

dynamical diffraction wave amplitudes. With this estimated thickness, the observed diffraction wave amplitudes are calibrated to give a set of fictitious observed kinematic structure-factor magnitudes. Based on such a set of magnitudes, a traditional least-squares procedure is used to refine structural parameters. The reliability of the result is checked by the consistency between the observed dynamical diffraction wave amplitudes and those found from