

## Introduction of Orientational Information into the Triple-Product Sign Probability

BY J. KROON AND H. KRABBENDAM

*Laboratorium voor Kristalchemie, Rijksuniversiteit, Utrecht, The Netherlands*

(Received 2 June 1969)

A probability formula for the signs of triple products  $U_h U_{h'} U_{h+h'}$  is derived, in which previous knowledge of molecular orientation is accounted for. For lower-order triple products considerable improvement of sign prediction is obtained when the molecular orientation is determined with reasonable accuracy.

### Introduction

Often, before starting a crystal-structure determination, the conformation of the constituent molecules is known, or at least partly known. Often too the orientations of these known parts can be determined from the Patterson function. However, the next step, determination of the positions of these parts, involves laborious methods of deconvolution. Application of direct methods from this point onward should offer distinct advantages.

It will be shown that information about molecular orientation implies certain knowledge of the double Patterson function  $\int \varrho(\mathbf{r})\varrho(\mathbf{r}+\mathbf{u})\varrho(\mathbf{r}+\mathbf{v})d\mathbf{r}$ . Then using the conventional heavy-atom method, sign probabilities can be found for the Fourier coefficients of the double Patterson synthesis, that is, for the triple products  $F_h F_{h'} F_{h+h'}$ .

### Triple-product calculations

As was pointed out by Sayre (1953) the double Patterson function:

$$T(\mathbf{u}, \mathbf{v}) \equiv \int \varrho(\mathbf{r})\varrho(\mathbf{r}+\mathbf{u})\varrho(\mathbf{r}+\mathbf{v})d\mathbf{r}$$

can be written, for point atoms at rest, as:

$$T(\mathbf{u}, \mathbf{v}) = \frac{1}{V^2} \sum_{\mathbf{h}} \sum_{\mathbf{h}'} \tau_{\mathbf{h}\mathbf{h}'} \exp[-2\pi i(\mathbf{h} \cdot \mathbf{u} + \mathbf{h}' \cdot \mathbf{v})] \quad (1)$$

where

$$\tau_{\mathbf{h}\mathbf{h}'} \equiv U_{\mathbf{h}} U_{\mathbf{h}'} U_{\mathbf{h}+\mathbf{h}'}$$

If we assume for simplicity that the structure consists of  $N$  equal atoms, then:

$$U_{\mathbf{h}} = \frac{1}{N} \sum_j \exp(2\pi i \mathbf{h} \cdot \mathbf{r}_j)$$

and the triple product can be expressed as:

$$\tau_{\mathbf{h}\mathbf{h}'} = \frac{1}{N^3} \sum_i \sum_j \sum_k \exp\{2\pi i[\mathbf{h} \cdot (\mathbf{r}_j - \mathbf{r}_i) + \mathbf{h}' \cdot (\mathbf{r}_k - \mathbf{r}_i)]\} \quad (2)$$

As (2) only involves interatomic vectors, information about molecular orientation can be inserted into this formula. For the sake of argument we will restrict our formulae to the case of space group  $P\bar{1}$  with two

molecules in the unit cell, in which the orientations of the entire molecules are supposed to be determined.

We write:

$$\tau_{\mathbf{h}\mathbf{h}'}^{\bar{}} = \tau_{\mathbf{h}\mathbf{h}'}^M + \tau_{\mathbf{h}\mathbf{h}'}^R \quad (3)$$

where

$$\tau_{\mathbf{h}\mathbf{h}'}^M = \frac{1}{N^2} + \frac{2}{N^3} \sum_{\substack{i \quad j \quad k \\ \text{not } i=j=k}}^{N/2 \quad N/2 \quad N/2} \cos\{2\pi[\mathbf{h} \cdot (\mathbf{r}_j - \mathbf{r}_i) + \mathbf{h}' \cdot (\mathbf{r}_k - \mathbf{r}_i)]\} \quad (4a)$$

contains only pairs of intra-molecular vectors, that is  $i, j$  and  $k$  indicate atoms belonging to the same molecule, and

$$\tau_{\mathbf{h}\mathbf{h}'}^R = \frac{1}{N^3} \sum_i \sum_j \sum_k \exp\{2\pi i[\mathbf{h} \cdot (\mathbf{r}_j - \mathbf{r}_i) + \mathbf{h}' \cdot (\mathbf{r}_k - \mathbf{r}_i)]\} \quad (4b)$$

contains all the other combinations of vectors, that is  $i, j$  and  $k$  indicate atoms not all belonging to the same molecule. The second term of  $\tau_{\mathbf{h}\mathbf{h}'}^{\bar{}}$ , say  $\xi_{\mathbf{h}\mathbf{h}'}$ , can be rewritten as:

$$\begin{aligned} \xi_{\mathbf{h}\mathbf{h}'} &= \frac{2}{N^3} \sum_i \sum_j \sum_k^{N/2 \quad N/2 \quad N/2} \\ &\quad \text{not } i=j=k \\ &\quad (\cos 2\pi \mathbf{h} \cdot \mathbf{r}_j \cos 2\pi \mathbf{h}' \cdot \mathbf{r}_k \cos 2\pi(\mathbf{h} + \mathbf{h}') \cdot \mathbf{r}_i \\ &\quad - \sin 2\pi \mathbf{h} \cdot \mathbf{r}_j \sin 2\pi \mathbf{h}' \cdot \mathbf{r}_k \cos 2\pi(\mathbf{h} + \mathbf{h}') \cdot \mathbf{r}_i \\ &\quad + \sin 2\pi \mathbf{h} \cdot \mathbf{r}_j \cos 2\pi \mathbf{h}' \cdot \mathbf{r}_k \sin 2\pi(\mathbf{h} + \mathbf{h}') \cdot \mathbf{r}_i \\ &\quad + \cos 2\pi \mathbf{h} \cdot \mathbf{r}_j \sin 2\pi \mathbf{h}' \cdot \mathbf{r}_k \sin 2\pi(\mathbf{h} + \mathbf{h}') \cdot \mathbf{r}_i) \quad (5) \end{aligned}$$

where  $i, j$  and  $k$  indicate atoms belonging to the same molecule. Since according to (4a),  $\xi_{\mathbf{h}\mathbf{h}'}$  is independent of the choice of origin and since the molecular orientation is known, the atomic coordinates relative to some chosen origin can be inserted into formula (5). This formula was actually used for the computation.

Another known part of  $\tau_{\mathbf{h}\mathbf{h}'}$ , which, however, is of no use for sign-probability calculations, will be dealt with in the Appendix.

### Sign probability

With reference to equation (3) the problem arises as to the probability  $P(\mathbf{h}\mathbf{h}')$  that the sign of  $\tau_{\mathbf{h}\mathbf{h}'}$  equals the sign of  $\tau_{\mathbf{h}\mathbf{h}'}^M$ . A similar problem about the probability

that the observed structure factor  $F_h$  has the same sign as its known part occurred in the heavy-atom technique. Woolfson (1956) derived an expression for this probability, assuming the probability distribution of the structure factors to be Gaussian.

As the probability distribution of the triple products is Gaussian too, providing that the constituent structure factors are independent (Kitaigorodskii, 1961), we can apply Woolfson's formula to our case of triple products. It reads, in a modified form:

$$P_+(\mathbf{h}, \mathbf{h}') = \frac{1}{2} + \frac{1}{2} \tanh \left( \frac{\tau_{hh'}^M |\tau_{hh'}|}{\langle (\tau_{hh'}^R)^2 \rangle} \right). \quad (6)$$

We will derive the denominator in the same way as Wilson for  $\langle F_h^2 \rangle$  (Wilson, 1949). In the expression

$$\tau_{hh'}^2 = \frac{1}{N^6} \sum_{i,j} \sum_{k,l} \sum_{m,n} \exp \{ 2\pi i [\mathbf{h} \cdot (\mathbf{r}_j - \mathbf{r}_i + \mathbf{r}_l - \mathbf{r}_m) + \mathbf{h}' \cdot (\mathbf{r}_k - \mathbf{r}_i + \mathbf{r}_l - \mathbf{r}_n)] \} \quad (7)$$

for large  $N$  only terms with  $(i=l) \cap (j=m) \cap (k=n)$  need to be considered. Therefore

$$\langle \tau_{hh'}^2 \rangle = \frac{N^3}{N^6} = \frac{1}{N^3}. \quad (8)$$

If no orientational parameters are known, formula (4) reduces to

$$\tau_{hh'}^M = \frac{1}{N^2}$$

and

$$\langle (\tau_{hh'}^R)^2 \rangle = (N^3 - N)/N^6$$

so that according to (6)

$$P_+(\mathbf{h}\mathbf{h}') = \frac{1}{2} + \frac{1}{2} \tanh \left[ \frac{N^4}{N^3 - N} |U_h U_{h'} U_{h+h'}| \right] \quad (9)$$

which approaches, for a reasonably large value of  $N$ , to the well-known formula developed by Cochran & Woolfson (1955):

$$P_+(\mathbf{h}\mathbf{h}') = \frac{1}{2} + \frac{1}{2} \tanh [N |U_h U_{h'} U_{h+h'}|]. \quad (10)$$

If we include the orientational information, developed in the preceding section, then:

$$\langle (\tau_{hh'}^R)^2 \rangle = \frac{N^3 - 2(\frac{1}{2}N)^3}{N^6}$$

and

$$P_+(\mathbf{h}\mathbf{h}') = \frac{1}{2} + \frac{1}{2} \tanh \left[ \frac{N^6 |U_h U_{h'} U_{h+h'}| (\xi_{hh'} + 1/N^2)}{N^3 - 2(\frac{1}{2}N)^3} \right], \quad (11)$$

where  $\xi_{hh'}$  represents the orientational contribution as given by (5), for the special case of two molecules in the unit cell ( $P\bar{1}$ ).

Naturally the expression for  $\xi_{hh'}$  can be extended to other space groups and to the case of partial orientational information. Also equation (11) can be modified to be applicable if the positions of some atoms are known (e.g. in the heavy-atom case).

### Example

Equation (11) was tried out with the aid of an example taken from the literature. We chose the (010) projection of cyclohexane-1,2-dione (Mossel & Romers, 1964). The orientation of the molecule was found by inspection of the (010) Patterson synthesis. Using the scaling factor and temperature factor found by Wilson's method ( $B=1 \text{ \AA}^2$ ,  $c_{\text{Wilson}}/c_{\text{least squares}}=1.11$ ), we calculated unitary structure factors, of which only those with values exceeding 0.15 were included in the analysis. Our information regarding molecular orientation was then introduced into equation (11) and probabilities  $P_+^0(\mathbf{h}\mathbf{h}')$  were calculated for the triple products concerned. These probabilities were compared with probabilities  $P_+(\mathbf{h}\mathbf{h}')$  obtained by application of (10) by means of an improvement index:

$$I_n = \frac{1}{n} \sum_{\mathbf{h}} \sum_{\mathbf{h}'} [P_+^0(\mathbf{h}\mathbf{h}') - P_+(\mathbf{h}\mathbf{h}')] s_{hh'} \quad (12)$$

where  $s_{hh'}$  is the sign of each triple product as has been found from the structure analysis, and  $n$  the total number of triple products involved. Using the triple products ( $|U|s > 0.15$ ) produced within the Cu-sphere,  $I_{335}$  was calculated to be 0.11 with the final orientation, and it was calculated to be only 0.01 with the orientation found by Patterson analysis; a marked difference indeed. However, a pronounced improvement was obtained when only terms up to  $\theta=30^\circ$  were included, giving values  $I_{22}$  of 0.16 and 0.12 for the precise and rough orientation of the molecule respectively. The reason for this effect is clear: the calculated values of the high-order triple products are more sensitive to errors in molecular orientation than those of the low-order triple products.

Table 1 shows the discrepancies between the coordinates of the atoms within one molecule as obtained from Patterson analysis, and the coordinates obtained from the outcome of the structure refinement, both relative to the same origin.

Table 1. *Orientalional parameters of the cyclohexane-1,2-dione molecule in the (001) projection*

Parameters obtained by Patterson analysis (first and third column) and structure refinement (second and fourth column).

$$a=6.65, b=6.21, c=6.87 \text{ \AA}, \beta=99^\circ 49', P2_1, Z=2.$$

	$x(\text{Patt.})$	$x(\text{L.S.})$	$z(\text{Patt.})$	$z(\text{L.S.})$
C(1)	0.446	0.427	0.441	0.447
C(2)	0.515	0.517	0.246	0.256
C(3)	0.374	0.356	0.065	0.072
C(4)	0.159	0.173	0.071	0.060
C(5)	0.082	0.100	0.258	0.254
C(6)	0.227	0.198	0.441	0.423
O(1)	0.562	0.538	0.588	0.606
O(2)	0.035	0.090	-0.085	-0.094

Table 2 gives a survey of the probabilities  $P_+^0$  (obtained from equation (11) with the rough orientation),  $P_+^S$  (obtained from equation (11) with the precise

Table 2. *Actual signs and sign probabilities for large triple products of low order*

$P_+$ : probability obtained from formula (10);  
 $P_{+^0}$ : probability obtained from formula (11) with rough orientation;  
 $P_{+^s}$ : probability obtained from formula (11) with correct orientation.

$h k l$	$U$	$h k l$	$U$	$h k l$	$U$	$s(\mathbf{hh}')$	$P_+(\mathbf{hh}')$	$P_{+^0}(\mathbf{hh}')$	$P_{+^s}(\mathbf{hh}')$
$\bar{4} 0 1$	0.47	3 0 0	0.42	$\bar{1} 0 1$	0.34	+	0.90	0.99	1.00
$\bar{4} 0 1$	0.47	3 0 1	0.18	$\bar{1} 0 2$	0.18	+	0.62	0.59	0.60
$\bar{4} 0 1$	0.47	3 0 2	0.29	$\bar{1} 0 3$	0.18	-	0.69	0.19	0.29
$\bar{3} 0 1$	0.27	0 0 2	0.17	$\bar{3} 0 3$	0.17	-	0.56	0.60	0.41
$\bar{3} 0 1$	0.27	2 0 0	0.24	$\bar{1} 0 1$	0.34	+	0.67	1.00	0.97
$\bar{3} 0 1$	0.27	3 0 1	0.18	0 0 2	0.17	+	0.56	0.56	0.52
$\bar{3} 0 2$	0.28	2 0 0	0.24	$\bar{1} 0 2$	0.18	+	0.59	0.59	0.66
$\bar{3} 0 2$	0.28	3 0 0	0.42	0 0 2	0.17	+	0.56	0.56	0.63
$\bar{3} 0 3$	0.17	2 0 0	0.24	$\bar{1} 0 3$	0.18	-	0.56	0.60	0.39
$\bar{2} 0 2$	0.28	$\bar{1} 0 1$	0.34	$\bar{3} 0 3$	0.17	+	0.63	0.41	0.82
$\bar{2} 0 2$	0.28	2 0 0	0.24	0 0 2	0.17	-	0.59	0.43	0.39
$\bar{2} 0 2$	0.28	3 0 1	0.18	1 0 3	0.26	+	0.52	0.52	0.55
$\bar{1} 0 1$	0.34	$\bar{1} 0 2$	0.18	$\bar{2} 0 3$	0.41	+	0.69	1.00	0.99
$\bar{1} 0 1$	0.34	0 0 2	0.17	$\bar{1} 0 3$	0.18	+	0.58	0.74	0.63
$\bar{1} 0 1$	0.34	1 0 5	0.59	0 0 6	0.70	+	0.99	1.00	1.00
$\bar{1} 0 1$	0.34	3 0 2	0.29	2 0 3	0.40	+	0.78	1.00	0.98
$\bar{1} 0 1$	0.34	4 0 0	0.21	3 0 1	0.18	-	0.60	0.54	0.45
$\bar{1} 0 2$	0.18	3 0 1	0.18	2 0 3	0.40	-	0.60	0.64	0.58
$\bar{1} 0 2$	0.18	4 0 0	0.21	3 0 2	0.29	+	0.59	0.56	0.57
$\bar{1} 0 3$	0.18	2 0 0	0.24	1 0 3	0.26	-	0.59	0.24	0.35
$\bar{1} 0 3$	0.18	3 0 0	0.42	2 0 3	0.40	-	0.72	0.02	0.01
0 0 2	0.17	3 0 0	0.42	3 0 2	0.29	-	0.66	0.35	0.34

orientation) and  $P_+$  [from equation (10)], for the lower-order triple products ( $\theta$  up to  $30^\circ$ ). It is seen from this Table that there is a great increase in reliable signs for triple products after introduction of structural information with the aid of formula (11). It is evident that great care must be taken to determine the molecular orientation as accurately as possible.

We are indebted to Professor A.F. Peerdeman for valuable discussions and we wish to thank Dr Roeli Olthof-Hazekamp for critical reading of the manuscript.

### APPENDIX

In the case of equal point atoms at rest the three sections  $\mathbf{u}=\mathbf{0}$ ,  $\mathbf{v}=\mathbf{0}$  and  $\mathbf{u}=\mathbf{v}$  of the double Patterson map are in fact ordinary Patterson syntheses. This becomes manifest by writing formula (2) as:

$$\tau_{\mathbf{hh}'} = \frac{1}{N^2} + \frac{1}{N^3} \left\{ \sum_{i \neq j \neq k \neq i}^N \sum^N \sum^N \exp \{ 2\pi i [\mathbf{h} \cdot (\mathbf{r}_j - \mathbf{r}_i) + \mathbf{h}' \cdot (\mathbf{r}_k - \mathbf{r}_i)] \} \right. \\
+ \sum_{i \neq k}^N \sum^N \exp [2\pi i \mathbf{h}' \cdot (\mathbf{r}_k - \mathbf{r}_i)] \\
+ \sum_{i \neq j}^N \sum^N \exp [2\pi i (\mathbf{h} + \mathbf{h}') \cdot (\mathbf{r}_j - \mathbf{r}_i)] \\
\left. + \sum_{i \neq j}^N \sum^N \exp [2\pi i \mathbf{h} \cdot (\mathbf{r}_j - \mathbf{r}_i)] \right\}$$

$$= -2/N^2 + \frac{1}{N^3} \sum_{i \neq j \neq k \neq i}^N \sum^N \sum^N \exp [2\pi i \mathbf{h} \cdot (\mathbf{r}_j - \mathbf{r}_i) + \mathbf{h}' \cdot (\mathbf{r}_k - \mathbf{r}_i)] \\
+ \frac{1}{N} (U_{\mathbf{h}}^2 + U_{\mathbf{h}'}^2 + U_{\mathbf{h}+\mathbf{h}'}^2). \quad (13)$$

It should be noted, however, that the second term of formula (13) is strongly correlated with the last term. For if there is a large peak at  $(\mathbf{u}_1, \mathbf{0})$  and a large peak at  $(\mathbf{0}, \mathbf{v}_1)$  in the double Patterson map, it will be fairly probable that there will be a considerable peak at  $(\mathbf{u}_1, \mathbf{v}_1)$ . Thus there will be a strong correlation between the contributions from large peaks in these zero hypersections and those from the unknown part of the double Patterson. This correlation prohibits the use of formula (13) for deriving a sign probability formula.

### References

- COCHRAN, W. & WOOLFSON, M. M. (1955). *Acta Cryst.* **8**, 1.  
KITAIGORODSKII, A. I. (1961). *The Theory of Crystal Structure Analysis*. New York: Consultants Bureau.  
MOSEL, A. & ROMERS, C. (1964). *Acta Cryst.* **17**, 1217.  
SAYRE, D. (1953). *Acta Cryst.* **6**, 430.  
WILSON, A. J. C. (1949). *Acta Cryst.* **2**, 318.  
WOOLFSON, M. M. (1956). *Acta Cryst.* **9**, 804.