

Acta Cryst. (1971). B27, 2037

Automation of the non-centrosymmetric symbolic addition. I. Fast determination of the unknown symbols.

By H. SCHENK, *Laboratory for Crystallography, University of Amsterdam, Nieuwe Prinsengracht 126, Amsterdam, The Netherlands*

(Received 3 May 1971)

A method is described for the determination of numerical values of the symbols employed in the non-centrosymmetric symbolic addition. The method is fast and applicable to all space groups. An analogous procedure can be used for centrosymmetric space groups.

Introduction

Recently direct methods based on the Σ_2 formula have been employed successfully in many non-centrosymmetric structure determinations, mostly of space groups $P2_1$ and $P2_12_12_1$. Roughly the methods used can be divided into two groups, one working with symbolic phases (Karle & Karle, 1966) and the other with numerical phases, which are refined by the tangent formula (Germain & Woolfson, 1968; Germain, Main & Woolfson, 1970; Hall, 1970).

The methods from the second group do not permit the use of more than approximately 100 different starting sets of phases, in view of the computational costs. Kaufman & Leiserowitz (1970), for instance use some 800 starting sets but this approach consumes more computer time than our quarterly budget.

The first group of methods offers more elegant opportunities for automation. In this approach a readily obtainable figure of merit for the correctness of a set of trial values for the symbols is essential.

This paper deals with a method for symbol screening, which is very fast and applicable to all space groups.

Background

The method to be presented here enters the symbolic addition procedure (Karle & Karle, 1966) at the moment a basic set of reflexions has received symbolic phases of the form $\varphi_H = a_H + \sum_j a_{H,j} x_j$ (the x_j are the symbols, a_H is a constant, $a_{H,j}$ are integers) which have been obtained using strict acceptance criteria. All other reflexions either fall short of the acceptance criteria or have several differing phase indications φ_H .

In the tangent refinements, Germain & Woolfson (1968)

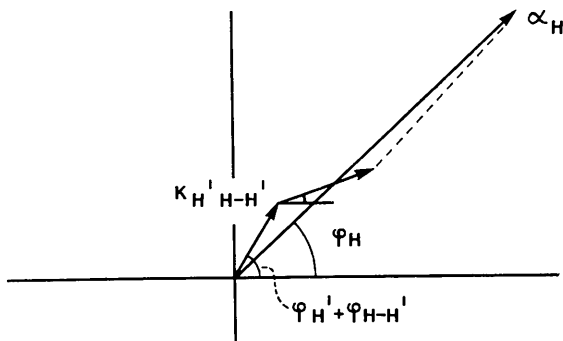


Fig. 1. The composition of different phase indications $\varphi_{H'} + \varphi_{H-H'}$ with amplitudes $K_{H'H-H'} = 2\sigma_3\sigma_2^{-3/2}|E_H E_{H'} E_{H-H'}|$ to give a resultant α_H with phase φ_H .

are handling a figure of merit for the correctness of a set of phases φ_H

$$C = \sum_H \alpha_H^2 = \text{maximum} \quad (1)$$

$$\text{where } \alpha_H = \left\{ \left[\sum_{H'} K_{H'H-H'} \cos(\varphi_{H'} + \varphi_{H-H'}) \right]^2 + \left[\sum_{H'} K_{H'H-H'} \sin(\varphi_{H'} + \varphi_{H-H'}) \right]^2 \right\}^{1/2}$$

(see Fig. 1).

This figure of merit may also be applied to the Σ_2 list, consisting of all reflexions with $|E| > 1.0$ or 1.2 and calculated on the basis of the symbolically phased reflexions. Then C can be calculated as a function of the numerical values for the symbols $x_1 \dots x_n$, but this is a rather time-consuming calculation.

In centrosymmetric symbolic addition the criterion:

$$\sum_H \sum_{H'} |E_H E_{H-H'} E_{H'}| \{1 - S_H \times S_{H'} \times S_{H-H'}\} = \text{minimum}$$

is used (Schenk, 1969). This criterion gives the deviation from complete internal consistency and in fact we found in 80% of our structure determinations by this method that the lowest criterion corresponded to the correct solution. In an additional 10% the second lowest criterion gave the correct solution.

In analogy we proposed the criterion

$$\sum_H \left\{ \left(\sum_{H'} K_{H'H-H'} \right) - \alpha_H \right\} = \text{minimum} \quad (2)$$

(Schenk, 1969) as indicative for correct solutions in non-centrosymmetric symbolic additions. However, (1) and (2) are equally time consuming.

Closely related to (2) is the figure of merit:

$$\sum_H \sum_i \sum_j K_{H_i H-H_i} + K_{H_j H-H_j} - |R_{Hij}| = \text{minimum} \quad (3)$$

$$\text{where } |R_{Hij}| = (K_{H_i H-H_i}^2 + K_{H_j H-H_j}^2 + 2K_{H_i H-H_i} K_{H_j H-H_j} \cos \Delta\varphi_{Hij})^{1/2}$$

and $\Delta\varphi_{Hij} = \varphi_{H_j} + \varphi_{H-H_i} - \varphi_{H_i} - \varphi_{H-H_j}$ (see Fig. 2).

In fact if the number of triplets for each H does not exceed 2 formulae (3) and (2) are identical. Again no gain of computing time is achieved.

The following criterion enables a fast screening:

$$Q = \sum_H \sum_i \sum_j (K_{H_i H-H_i} + K_{H_j H-H_j})^2 - |R_{Hij}|^2 = \text{minimum}. \quad (4)$$

Q can be rewritten as

$$Q = \sum_H \sum_i \sum_j K_{H_i H-H_i} K_{H_j H-H_j} (1 - \cos \Delta\varphi_{Hij}). \quad (5)$$

Many of the terms in this threefold summation have iden-

tical $\Delta\varphi_{Hij}$ and may be taken together by summing their $K_{H_iH-H_i} K_{H_jH-H_j}$'s. This leads to:

$$Q = \sum_i w_i (1 - \cos \Delta\varphi_i) = \text{minimum.} \quad (6)$$

Since $\Delta\varphi_i$ are expressions in x_j the figure of merit Q is a function of the x_j . In the summation some w_i are relatively very weak and can be neglected. If one w_i is dominating, the corresponding $\Delta\varphi_i$ is made equal to $0 \pmod{2\pi}$. This may reduce the number of trial sets of x_j , for which Q has to be calculated.

Programs

Two computer programs were written in ALGOL 60 for the X8-Electrologica computer (2.5 microsecond). The first program generates the $\Delta\varphi_i$ and their weights w_i from a Σ_2 list based on the group of symbolically signed reflexions. The second program calculates Q as a function of the symbolic phases $x_1 \cdots x_n$.

About the computer time involved it can be mentioned that in the case of the sulphur steroid (see later) the present versions of the programs consume 2 minutes for reducing 1500 triplet interactions into approximately 100 $\Delta\varphi_i$'s and 2 minutes for the calculation of Q for 500 sets of trial values for $x_1 \cdots x_4$. It is expected that these speeds can be improved.

Results and discussion

The method has been applied to two crystal structure determinations: the photolysis product of Karle, Karle & Estlin (1967) and a sulphur steroid (van de Ven & Schenk, 1971), both of space group $P2_12_12_1$.

In the first determination 3 symbols were chosen in order to build up a starting set of 42 symbolic phases, which shows no inconsistencies. In Fig. 3, section $x_3 = 750$ through the three-dimensional function $Q(x_1, x_2, x_3)$ with the parameters in fractions multiplied by 1000. The deepest minima of the complete function are all in this section with parameters $x_1 = 250$ or 750 and $x_2 = 250$ or 750 . After tangent refinement the best solution showed an average deviation of 25° from the least-squares phases of the correct solution and in fact one of the three remaining minima refines to the enantiomorph.

The second structure required 4 symbols for the determination of 50 symbolic phases. The section $x_2 = 750$, $x_4 = 83$ of its four-dimensional function Q is given in Fig. 4. The lowest Q value is contained in this section and occurs at $x_1 = x_3 = 0$. Using this set of initial parameters tangent refinement yielded 470 phases. From the subsequent E map the structure could be found.

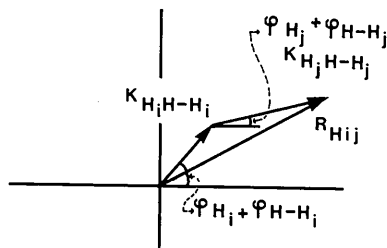


Fig. 2. The composition of two different phase indications to give a resultant with modulus $|R_{Hij}| = (K_{H_iH-H_i}^2 + K_{H_jH-H_j}^2 + 2K_{H_iH-H_i}K_{H_jH-H_j} \cos \Delta\varphi)^{1/2}$ with $\Delta\varphi = \varphi_{H_i} + \varphi_{H-H_j} - \varphi_{H_j} - \varphi_{H-H_i}$

From (5) it can be seen that a reflexion H with many different phase indications influences the weights w_i in (6) more than a reflexion with a small number of different phase indications. To avoid this an alternative scheme is suggested, in which the sum $\sum_i \sum_j K_{H_iH-H_i} K_{H_jH-H_j}$ is rescaled to $\sum_i K_{H_iH-H_i}$. By this procedure (5) becomes

$$Q = \sum_H \sum_i \sum_j K_{H_iH-H_i} K_{H_jH-H_j} \frac{\sum_i K_{H_iH-H_i}}{\sum_i \sum_j K_{H_iH-H_i} K_{H_jH-H_j}} \times (1 - \cos \Delta\varphi_{Hij})$$

which leads to modified values w_i in (6).

An analogous searching procedure can be introduced in the centrosymmetric symbolic addition. Then an analysis of 8 symbols would take less than 1 minute of X-8 computing time. Our present procedure (Schenk, 1969) analyses 8 symbols in approximately 1 hour.

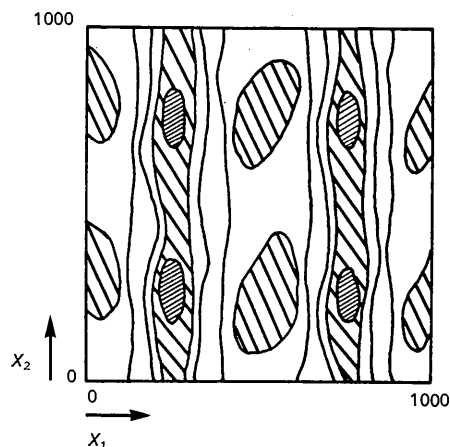


Fig. 3. Section $x_3 = 250$ through the function $Q(x_1, x_2, x_3)$ of the photolysis product. The minima have been shaded, deep minima heavily shaded. The parameters x_i are given in fractions multiplied by 1000.

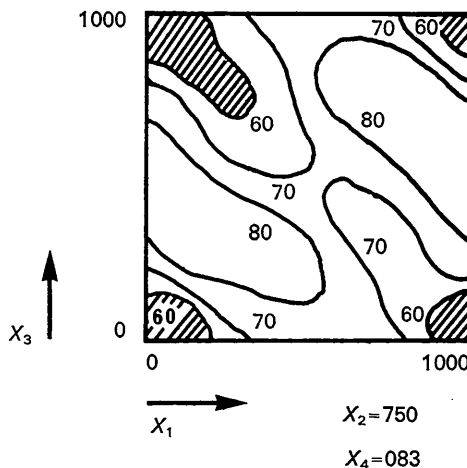


Fig. 4. Section $x_2 = 750$, $x_4 = 83$ through the function $Q(x_1, x_2, x_3, x_4)$ of the sulphur steroid. The parameters are given in fractions multiplied by 1000. Minima in the function have been shaded.

References

- GERMAIN, G. & WOOLFSON, M. M. (1968). *Acta Cryst.* B24, 91.
 GERMAIN, G., MAIN, P. & WOOLFSON, M. M. (1970). *Acta Cryst.* B26, 274.
 HALL, S. R. (1970). In *Crystallographic Computing*. Edited by F. R. AHMED, p. 66. Copenhagen: Munksgaard.

- KARLE, J. & KARLE, I. L. (1966). *Acta Cryst.* 21, 849.
 KARLE, I. L., KARLE, J. & ESTLIN, J. A. (1967). *Acta Cryst.* 23, 494.
 KAUFMAN, H. W. & LEISEROWITZ, L. (1970). *Acta Cryst.* B26, 442.
 SCHENK, H. (1969). *Automation of the Symbolic Addition Method*. Amsterdam: North Holland.
 VEN, C. F. W. VAN DE & SCHENK, H. (1971). To be published.

Acta Cryst. (1971). B27, 2039

Automation of the non-centrosymmetric symbolic addition. II. Use of centric reflexions for symbol determination. By H. SCHENK, *Laboratory for Crystallography, University of Amsterdam, Nieuwe Prinsengracht 126, Amsterdam, The Netherlands*

(Received 3 May 1971)

In non-centrosymmetric symbolic addition, applied to structures with centric projections, a systematic use of the phase indications of the centric reflexions accelerates the determination of numerical values for the symbols and avoids solutions which are unacceptable in projection.

Introduction

In the automation of non-centrosymmetric symbolic addition the systematic use of reflexions, the phases of which are restricted by symmetry operations can be very helpful in the determination of the values of the unknown symbolic phases. The most common symmetry element is a centre of symmetry in projections. In this paper a method for determining the values of unknown symbols from the phase indications of centric reflexions is presented.

Method

A basic set of symbolic phases $\varphi_H = a_H + \sum_i a_{Hi}x_i$, where a_H is a constant, a_{Hi} are integers and x_i are the unknown symbols, is obtained by the symbolic addition procedure (Karle & Karle, 1966) using a high-acceptance criterion. Then for all centric reflexions the symbolic phase indications are calculated. These indications have values restricted to s and $s + \pi$, where s in monoclinic and orthorhombic space

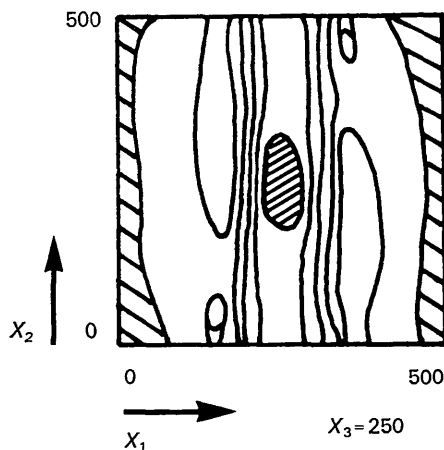


Fig. 1. Section through $x_3 = 250$ of the function $CI(x_1, x_2, x_3)$ of the photolysis product. The parameters are given in fractions multiplied by 1000. The regions with low CI values have been shaded, regions with very low CI values have been heavily shaded.

groups equals either 0 or $\frac{1}{2}\pi$, depending on the space group and the choice of origin. Then for each phase indication it follows that:

$$\varphi_{H'} + \varphi_{H-H'} - s = 0 \pmod{\pi} \quad (1)$$

in which H is a centric reflexion and H' and $H-H'$ belong to the basic set of symbolic phases. With these relations the following figure of merit can be constructed:

$$CI = \sum_{H, H'} K_{H'H-H'} |\sin(\varphi_{H'} + \varphi_{H-H'} - s)| = \text{minimum} \quad (2)$$

which can be rewritten to

$$CI = \sum_i k_i |\sin(b_i + \sum_j a_{ij}x_j)| \quad (3)$$

where

$$\begin{aligned} \varphi_{H'} + \varphi_{H-H'} - s &= a_{H'} + \sum_j a_{H'j}x_j + a_{H-H'} + \sum_j a_{H-H'j}x_j - s \\ &= b_i + \sum_j a_{ij}x_j \end{aligned}$$

and $K_{H'H-H'} = k_i$.

In (3) many terms $(b_i + \sum_j a_{ij}x_j)$ are identical. By summing their k_i 's the number of terms in (3) can be reduced:

$$CI = \sum_m w_m |\sin(b_m + \sum_j a_{mj}x_j)|. \quad (4)$$

Usually CI is dominated by a number of large w_m 's so that the very small w_m may be neglected.

From $b_m + \sum_j a_{mj}x_j = p \pmod{\pi}$ it follows that:

$$CI(x_1 \cdots x_n) = CI(x_1 + \pi \cdots x_n) = CI(x_1 + \pi \cdots x_n + \pi).$$

Thus it is impossible to differentiate between $x_i = q$ and $x_i = q + \pi$. These ambiguities can be solved for instance by calculating the Q value, which is based on all reflexions (Schenk, 1971), for each of these possibilities.

Two other figures of merit suggest themselves for screening the trial sets x_i in centrosymmetric projections:

$$CII = \sum_m w_m \{1 - \cos 2(b_m + \sum_j a_{mj}x_j)\} \quad (5)$$

and

$$CIII = \sum_m w_m (b_m + \sum_j a_{mj}x_j)_{\text{restr}}. \quad (6)$$

In (6) the restriction is $-\frac{1}{2}\pi \leq (b_m + \sum_j a_{mj}x_j)_{\text{restr}} < \frac{1}{2}\pi$.