

- DORFMAN, L., FURLENMEIER A., HUEBNER, C. F., LUCAS, R., MACPHILLAMY, H. B., MUELLER, J. M., SCHLITTLER, E., SCHWYZER, R. & ST. ANDRÉ, A. F. (1954). *Helv. Chim. Acta*, **37**, 59.
- DORFMAN, L., HUEBNER, C. F., MACPHILLAMY, H. B., SCHLITTLER, E. & ST. ANDRÉ, A. F. (1953). *Experientia*, **9**, 368.
- HAUPTMAN, H. & KARLE, J. (1956). *Acta Cryst.* **9**, 45.
- JOHNSON, C. K. (1965). *ORTEP*, ORNL-3794, Oak Ridge National Laboratory, Oak Ridge, Tennessee.
- KARLE, I. L. & KARLE, J. (1964). *Acta Cryst.* **17**, 835.
- KARLE, I. L. & KARLE, J. (1966b). *Acta Cryst.* **21**, 860.
- KARLE, J. (1968). *Acta Cryst.* In the press.
- KARLE, J. & HAUPTMAN, H. (1956). *Acta Cryst.* **9**, 635.
- KARLE, J. & KARLE, I. L. (1966a). *Acta Cryst.* **21**, 849.
- MÜLLER, J. M., SCHLITTLER, E. & BEIN, H. J. (1952). *Experientia*, **8**, 338.
- NORDMAN, C. E. & KUMRA, S. K. (1965). *J. Amer. Chem. Soc.* **87**, 2059.
- ROSE, H. A. (1954). *Anal. Chem.* **26**, 1245.
- WOODWARD, R. B., BADER, F. E., BICKEL, F., FRYE, A. J. & KIERSTEAD, R. W. (1958). *Tetrahedron*, **2**, 1.

*Acta Cryst.* (1968). B**24**, 91

## On the Application of Phase Relationships to Complex Structures

BY G. GERMAIN

*Laboratoire de Chimie-Physique, 39 Schapenstraat, Université de Louvain, Belgium*

AND M. M. WOOLFSON

*Physics Department, University of York, England*

(Received 22 May 1967)

For complex centrosymmetric structures the symbolic addition method has two weak links; the first is in the initial few steps when single sign relationships must be trusted and the second is in the acceptance of relationships between sign symbols. A systematic multi-solution symbolic addition procedure is described which greatly reduces the risk of taking a wrong step and offers the possibility of solving structures with up to 400 atoms in the unit cell. A computer program is described which incorporates some of the proposals. Several alternative procedures are also described for extending the application of phase-determining techniques to non-centrosymmetric structures and the result of applying one of these methods is given. The basis of these procedures is that initial phase allocations in the correct quadrant ( $\pi/4$ ,  $3\pi/4$ ,  $5\pi/4$ ,  $7\pi/4$ ) to a few reflexions is sufficient to derive new phases and to refine them by the usual tangent formula.

### Introduction

For centrosymmetric structures the point has now been reached where, for equal-atom structures with no special features, direct methods are at least as effective as other methods of solving crystal structures. After many years in which various types of sign-determining formula have been proposed it has become clear that the simplest one of all, the triple-product sign relationship (t.p.s.r.) is the most powerful and the most useful. The method of using these relationships which has proved to be very powerful is based on that proposed by Zachariasen (1952) in his contribution to the trio of pioneer papers by himself, Sayre (1952) and Cochran (1952). Zachariasen showed that if inequality relationships enabled the signs of some structure factors to be related and if signs, where unknown, were represented by letter symbols then subsequent use of the relationship

$$s(\mathbf{h}) \approx s\left\{ \sum_{\mathbf{h}'} s(\mathbf{h}' ) s(\mathbf{h} - \mathbf{h}') \right\} \quad (1)$$

could lead to a knowledge of signs (perhaps in terms of symbols) for more reflexions and could also often

enable the signs corresponding to the symbols to be determined. It is assumed in (1) that all the signs are those of large unitary structure factors.

It was suggested by Woolfson (1961) that a 'hit-or-miss' version of Zachariasen's method could be tried even when inequality relationships were not available. If a number of the strongest t.p.s.r.'s were accepted as inviolable then further progress could be made by the use of relationship (1).

However, the main credit for demonstrating the full power of this type of method must go to Karle & Karle, who have developed a systematic approach to the use of sign symbols which they call the 'symbolic addition' method. These workers and their associates have solved a number of fairly complex structures (e.g. Karle & Karle, 1964a; Karle, Karle, Owen & Hoard, 1965) and their process is now widely used by crystallographers everywhere.

The process of using symbols has also been extended to the solution of non-centrosymmetric structures. The applications which have been made (Karle & Karle, 1964b; Karle & Karle, 1966a) have relied to some extent on the presence of centrosymmetric projections

so that when some initial origin-fixing phases were chosen as 0,  $\pi$  or  $\pi/2$  the phases of other projection reflexions would be known to be restricted to similar special values. Some general phases (usually three) were represented by letter symbols and the determination of new phases was carried forward by means of the relationship

$$\varphi_{\mathbf{h}} = \langle \varphi_{\mathbf{h}'} + \varphi_{\mathbf{h}-\mathbf{h}'} \rangle_{\mathbf{h}'}, \quad (2)$$

where the average is taken over these values of  $\mathbf{h}'$  for which the phases on the right-hand side are known and for which the associated  $U$ 's are large. This relationship must be used with great caution; if, for example, two pairs of phases are available on the right hand side,  $(-20^\circ, 30^\circ)$  and  $(-170^\circ, -160^\circ)$ , then the two sums of phases are  $10^\circ$  and  $-330^\circ$  and the average  $-160^\circ$ . These values are in fact only  $20^\circ$  apart, for  $-330^\circ$  is equivalent to  $30^\circ$ , and the reasonable average to take is  $20^\circ$ . This demonstrates that when this type of averaging process is carried out in terms of symbols great care must be taken – especially in deducing relationships between the phase symbols.

Finally Karle & Karle perform a cyclic refinement of phases by means of a special application of Sayre's equation applied to the normalized structure factors, the  $E$ 's ( $E=U/\sqrt{N}$  if there are  $N$  equal atoms in the unit cell). This is

$$\varphi_{\mathbf{h}} = \text{phase of } \sum_{\mathbf{h}'} E_{\mathbf{h}'} E_{\mathbf{h}-\mathbf{h}'}, \quad (3)$$

where the summation includes all the  $E$ 's whose phases have been determined. This relationship can be expressed in the more convenient form

$$\tan \varphi_{\mathbf{h}} = \frac{\sum_{\mathbf{h}'} |E_{\mathbf{h}'} E_{\mathbf{h}-\mathbf{h}'}| \sin(\varphi_{\mathbf{h}'} + \varphi_{\mathbf{h}-\mathbf{h}'})}{\sum_{\mathbf{h}'} |E_{\mathbf{h}'} E_{\mathbf{h}-\mathbf{h}'}| \cos(\varphi_{\mathbf{h}'} + \varphi_{\mathbf{h}-\mathbf{h}'})}. \quad (4)$$

Equation (4) emphasizes another hazard in applying equation (2) – that it makes no provision for compounding the probabilities of separate indications of phase. Equation (2) tends to suggest that a phase indication of  $0^\circ$  together with one of  $180^\circ$  would indicate a probable phase of  $90^\circ$  (or  $270^\circ$ ). In fact, if two phase indications have equal weight and differ by  $180^\circ$ , then equation (4) gives  $\tan \varphi_{\mathbf{h}} = 0/0$ , showing that all phase angles are equally probable. A full treatment of the probabilities indicated by equation (4) has been given by Cochran (1955).

It is the purpose of the present paper to demonstrate how these methods can be extended to structures of greater complexity than hitherto and, in the case of non-centrosymmetric structures, how the whole process of phase determination can be made more systematic.

### Centrosymmetric structures

The greatest weakness of the symbolic addition method is in the early stages of its application when, in order to progress, one must rely on single sign relationships to determine new signs (usually in terms of symbols).

For simple and moderately complex structures the first few steps might be reasonably sure but it is clear that for more complex structures the whole process might well come to grief at this stage. However once this barrier is penetrated, and new signs are being found on the basis of two or more t.p.s.r.'s, one may progress far more confidently. If for example the same sign (or sign symbol) allocation is suggested by two t.p.s.r.'s with probabilities  $p_1$  and  $p_2$  then the overall probability that this sign is correct is

$$P_+ = \frac{p_1 p_2}{1 - p_1 - p_2 + 2p_1 p_2}. \quad (5)$$

Thus for  $p_1 = p_2 = 0.80$

$$P_+ = 0.941$$

and for  $p_1 = p_2 = 0.85$

$$P_+ = 0.970$$

so that even a pair of quite weak sign relationships in unison can give a strong indication of sign.

Tollin (1962), in considering the Cochran & Douglas (1955) matrix-inversion method of using t.p.s.r.'s, has considered the probability of obtaining the correct solution when one allows for different numbers of failures in the basic set of sign relationships. This type of consideration is also relevant to the symbolic-addition method. Let us have a look at the early stages of the process and try to assess the risks involved. The probability for a single t.p.s.r., assuming  $N$  equal atoms per unit cell, is given by

$$P_{\mathbf{h},\mathbf{h}'} = \frac{1}{2} + \frac{1}{2} \tanh \left( \frac{|E_{\mathbf{h}} E_{\mathbf{h}'} E_{\mathbf{h}-\mathbf{h}'}|}{N^{\frac{1}{2}}} \right) \quad (6)$$

(Cochran & Woolfson, 1955). The distribution of  $E$ 's does not depend on the structural complexity and in the early stages one is usually dealing with three  $E$ 's, the magnitude of whose product is about 20. Thus the sort of probability one is concerned with is of the order

$$P_{\max} = \frac{1}{2} + \frac{1}{2} \tanh \left( \frac{20}{N^{\frac{1}{2}}} \right)$$

In Table 1  $P_{\max}$  is shown as a function of  $N$ .

Table 1. Relation of  $P_{\max}$  to  $N$

$N$	$P_{\max}$
25	0.9997
50	0.996
100	0.982
200	0.945
400	0.881

The number of steps one must take relying on single t.p.s.r.'s depends on the structural complexity and also on the symmetry number of the space group. The more complex the structure, the larger is the number of points in the reciprocal lattice and, for a given number of known signs, the less is one likely to find double indications for a new sign. However the symmetry number is important here; for the simplest case,  $P\bar{1}$ , only two

points in the reciprocal lattice are determined for each new sign whereas for an orthorhombic space group eight points are determined. Let us consider a situation where the first ten steps are to be taken with single t.p.s.r.'s with probabilities given by Table 1. The probability that one can get through this series of steps without error is  $(P_{\max})^{10}$  and is shown in the first column of Table 2.

Table 2. *Probability of errors in first ten steps*

$N$	Probability of no error in 10 steps	Probability of one or less error in 10 steps	Probability of two or less errors in 10 steps
25	0.998	1.000	1.000
50	0.962	1.000	1.000
100	0.834	0.987	1.000
200	0.568	0.899	0.986
400	0.282	0.663	0.894

One can see that this probability sharply diminishes with increasing  $N$  and for  $N > 100$  there is a large probability of failure. This suggests a method of progress which gives more protection against failure – that is we allow the possibility of up to one or even two failures in the first ten steps. The degree of extra protection this offers is shown in the second and third columns of Table 2.

The process for allowing one error is now described. In the first step one assumes that the relationship succeeds ( $s$ ) or fails ( $f$ ). From  $f$  one takes the next nine steps as all successful. From  $s$  we allow the next relationship to succeed or fail and so on. One ends up having taken eleven different routes, one of which corresponds to no failures and the other ten to the different ways of having one failure.

For the larger values of  $N$  it is possible that more than ten steps would have to be made with single t.p.s.r.'s. In Table 3 are shown the probabilities that there will be two or less failures for various numbers of steps for  $N=400$ .

Table 3. *Probability of two or less failures*

Number of steps	Number of routes	Probability of failures $\leq 2$
10	56	0.894
12	79	0.837
14	106	0.773
16	137	0.705
18	172	0.637
20	211	0.569

The total number of routes which must be explored does not increase alarmingly with the number of steps which must be taken and, even for twenty steps, the probability that one of the routes is correct is higher than one half. The number of routes which can be tackled will depend on the available computer; for example for  $N=400$  and twenty steps one can cover all cases with less than or equal to three failures with 1351 routes giving a 0.723 chance of netting the correct set of steps.

Karle & Karle (1966*b*) have pointed out that often, as the development of signs proceeds, relationships between sign symbols become apparent. They also urge a great deal of caution in accepting such relationships and it is our experience that this can be a source of error even when the sign symbols are apparently strongly linked.

The way we have developed of avoiding pitfalls when deriving relationships between symbols is best explained as part of the description of a program which we have written to carry through the symbolic addition procedure. This program, written in FORTRAN IV for an IBM 360/40 computer with a 16K store, is written in five parts which can be run together or successively after examination of each partial output.

*Part I* reads in the input data in the form

- (a) a title card
- (b) space group type
- (c) list of  $U$ 's or  $E$ 's with their indices (up to 250)
- (d) list of triple products (up to 1000).

The item (d) is output by a previous program which takes in the largest 250  $U$ 's or  $E$ 's and punches out the triple products which relate them.

This part of the program fixes the origin by assigning positive signs to some reflexions and allocates symbols to six other reflexions. The arbitrary signs and symbols are allocated to the reflexions which occur most frequently in the t.p.s.r.'s (within the limitation imposed by the rules for selecting origin-fixing reflexions).

*Part II* goes through the process of building up the number of reflexions to which signs or symbols are allocated by the use of relationship (1). We shall not explain the details of this process except to say that the right-hand side of (1) is calculated for all the reflexions of unknown sign each time a sign or symbol is allocated and this is only allocated to the reflexion for which the new sign (or symbol) indication has the highest probability.

*Part III* searches for relationships between the symbols. If for a t.p.s.r. each of the three members has associated a sign, a symbol or a product of symbols then either the product is identically unity or a relationship between the symbols results. For six symbols –  $a, b, c, d, e, f$  – there can be up to 63 such relationships. That this is so can be seen by considering the symbols as binary bits; there are 63 arrangements of six 1's or 0's excluding all zeros. The relationships are listed in order of probability – this latter quantity being measured by the frequency of occurrence.

*Part IV* solves the relationships between the sign symbols by a similar process to that used by Cochran & Douglas (1955). In the simplest case if a set of six strong non-linearly dependent relationships can be found and are assumed all to hold then a unique set of signs can be allocated to the symbols. In practice we also consider the possibility that one of the rela-

tionships may fail and this gives six more sets of signs for the symbols – making seven sets in all.

In the event that not all the symbols can be solved for, the total number of sets of signs for the symbols, allowing one of the relationships to fail, is  $2^{6-m}(m+1)$ , where  $m$  is the number of symbols for which a solution is found.

*Part V* generates the signs of each reflexion for each of the sets of signs for the symbols. At this stage if a reflexion has associated with it more than one set of symbols indicating its sign, and if these indications oppose one another then, if the overall indication of sign is unsatisfactory, the sign of this reflexion will be given as indeterminate.

Three figures-of-merit are generated for each output set of signs. One counts the number of sign relationships which hold, the second the sum of the probabilities of the sign relationships which hold and the third

$$\sum s_1 s_2 s_3 |E_1 E_2 E_3| .$$

This program has been tried with a number of structures and has always been successful. The largest structure tried has been myo-inositol (Rabinowitz & Kraut, 1964) which contained 96 atoms in the unit cell with space group  $P2_1/c$ . The solution with the highest figures of merit was that for which all the relationships between the sign symbols held, and of the 250  $E$ 's being handled the signs of 240 were correctly determined, eight were undetermined because they did not occur in t.p.s.r.'s, and two signs were incorrect.

Other known structures which have been solved are cyclodecane-1, 6-dione ( $P2_1/c, Z=2, N=24$ ), tetraphenyl-naphthalene ( $P2_1/c, Z=2, N=68$ ) and guanidinium acetate (*Prima*,  $Z=4, N=32$ ). These had previously been solved by various workers at Louvain by inequalities, by hand application of the symbolic addition method and by the Patterson function respectively and will all be published in due course.

However, one structure, the previously unknown structure of one of the tetramers of t-butylfluoroacetylene ( $P2_1/c, Z=2, N=56$ ) was solved after a hand application of the symbolic addition procedure had failed, as had other attempts at solution. It turned out in this case that the correct solution involved the failure of a reasonably well-indicated relationship between symbols. The first seven equations between the symbols output by the computer were:

(1)	$df = -1$	Found 33 times
(2)	$acef = +1$	26
(3)	$ef = -1$	9
(4)	$ad = +1$	6
(5)	$de = +1$	6
(6)	$aef = -1$	6
(7)	$ab = -1$	5

The relationships above, less (5), could be solved for all the sign symbols but, in fact, (4) did not hold. The interesting point is that, although at first inspection the

failure of (4) would seem to lead to more failures of sign relationships, in fact the corresponding set of signs had the *highest* figures of merit. This structure is being refined by G.S.D. King, Union Carbide European Research Associates, Brussels and will be published in due course.

### Non-centrosymmetric structures

The difficulties and dangers of using relationship (2) have already been mentioned and it would certainly be an advantage to avoid its use altogether if this were possible. We have tried a procedure where relationship (4) is the only one used and it was certainly most successful on a simple trial structure. In developing this approach we were guided by the well-known fact that for centrosymmetric structures a recognizable correct structure is revealed by a Fourier synthesis with correct signs and randomly wrong structure amplitudes. Carrying over this experience to non-centrosymmetric structures it suggests that it may be enough to allocate the signs of the real and imaginary parts of some structure factors in order to progress with a phase-determining procedure.

The example we tried was *cis*-di(nitrosomethane) (Germain, Piret & Van Meerssche, 1963), space group  $P2_12_12_1$  with  $Z=4$  and  $N=24$ . The rules for selecting initial phases for non-centrosymmetric space groups is that three phases, the parity of whose phases are non-linearly dependent, may be chosen arbitrarily and one other may be designated as in the range 0 to  $\pm\pi$  (or 0 to  $-\pi$ ) which fixes the enantiomorph one is dealing with. We followed the convenient course of choosing three prime-projection reflexions arbitrarily; this gives the advantage that other prime-projection reflexions are also known with a single ambiguity either as 0 or  $\pi$  or as  $+\pi/2$  or  $-\pi/2$ . For four other reflexions, which had been found to link most frequently with others by triple product phase relationships, various combinations of initial phases were assigned as follows

$$\begin{aligned} & \frac{\pi}{4}, \frac{3\pi}{4} \text{ (enantiomorph fixing)} \\ & \frac{\pi}{4}, \frac{3\pi}{4}, \frac{-\pi}{4}, \frac{-3\pi}{4} \\ & \frac{\pi}{4}, \frac{3\pi}{4}, \frac{-\pi}{4}, \frac{-3\pi}{4} \\ & \frac{\pi}{4}, \frac{3\pi}{4}, \frac{-\pi}{4}, \frac{-3\pi}{4} . \end{aligned}$$

For each of the 128 initial sets of phases new phases were derived by an iterative process. For this structure the criterion of acceptance was as follows:

1st iteration	25 best-determined phases accepted – all others are rejected.
2nd iteration	50 best-determined phases accepted – all others rejected.
3rd iteration	75 accepted <i>etc.</i>

4th iteration 100 accepted *etc.*

5th iteration all phases accepted.

By 'best-determined phases' we mean those with the highest values of

$$Z_h = \left[ \sum_{h'} |E_{h'} E_{h-h'}| \sin(\varphi_{h'} + \varphi_{h-h'}) \right]^2 + \left[ \sum_{h'} |E_{h'} E_{h-h'}| \cos(\varphi_{h'} + \varphi_{h-h'}) \right]^2, \quad (7)$$

which is the square of the modulus of the right hand side of equation (3).

The number of iterations (five) and the numbers of accepted phases (25, 50, 75 *etc.*) are input parameters of the program and can be varied to match the complexity of the structure.

At the end of the last iteration a figure-of-merit criterion is calculated which is

$$C = \sum_h Z_h. \quad (8)$$

It was found that the set of phases with the highest figure of merit turned out to be very close to the correct refined phases as is shown in Table 4.

This structure was a very simple one but an extension of the same principle to more complex structures seems straightforward. However if one has  $n$  general starting reflexions in addition to the three origin-fixing and the one enantiomorph fixing reflexion the number of initial sets of phases increases by a factor of four for each unit increase of  $n$ . This is shown in Table 5.

Table 5. Increase of number of initial sets of phases with increase in number ( $n$ ) of starting reflexions

Total starting reflexions $4+n$	Number of initial sets $2 \times 4^n$
7	128
8	512
9	2048
10	8192

A possible way of increasing the number of starting reflexions while keeping the number of initial sets small is to make use of the substantialization of sign sequences idea proposed by Woolfson (1954). In the simplest form of this the following sixteen sequences of seven signs were used:

Table 4. Comparison of phase angles obtained by the program described with published phase angles for *cis-di(nitrosomethane)*

			Published Program					Published Program					Published Program	
$h$	$k$	$l$	$\varphi$	$\varphi$	$h$	$k$	$l$	$\varphi$	$\varphi$	$h$	$k$	$l$	$\varphi$	$\varphi$
0	3	5	270°	270°	2	11	3	246°	245°	4	9	4	207°	200°
2	0	5	90	90	2	9	1	111	120	4	9	0	180	180
1	4	0	270	270	0	7	5	90	90	2	5	0	360	360
1	1	5	32	15	2	7	1	167	175	2	9	2	160	175
5	5	1	319	330	3	12	1	355	5	2	9	4	150	175
6	1	1	57	60	2	14	2	173	175	7	4	0	90	90
5	2	4	102	105	3	8	3	176	185	2	6	0	180	180
7	2	1	185	180	3	4	3	357	355	0	8	0	180	180
2	3	4	43	30	5	0	1	270	270	1	12	2	276	275
2	10	2	346	350	1	14	1	348	340	1	10	4	66	55
4	10	2	168	185	5	8	1	66	65	3	6	4	343	350
0	12	0	360	360	3	8	5	171	155	5	8	0	270	270
2	4	2	283	275	3	8	1	180	180	1	2	4	282	270
4	6	2	332	335	5	2	1	34	30	1	8	0	90	90
2	6	4	172	165	3	11	2	93	90	3	0	2	360	360
4	10	4	225	240	3	7	4	292	285	3	10	2	222	215
2	4	4	84	60	3	5	4	261	275	5	6	2	238	265
2	2	0	360	360	3	11	0	90	90	5	0	4	180	180
4	10	0	180	180	1	3	0	270	270	1	0	4	180	180
6	4	0	360	360	1	13	2	69	70	5	2	0	90	90
4	9	1	251	250	3	9	2	69	100	1	12	0	90	90
6	3	1	115	105	7	1	2	262	255	1	14	2	235	260
6	7	3	308	300	3	7	0	270	270	5	4	0	90	90
4	7	1	198	205	3	3	4	90	85	3	10	0	90	90
0	5	5	270	270	0	4	5	180	180	0	8	2	360	360
0	13	3	270	270	0	12	3	180	180	0	6	4	180	180
4	3	1	12	35	4	6	4	32	25	5	3	3	34	30
4	5	1	154	150	6	0	3	270	270	3	5	3	73	65
6	9	1	207	225	2	12	3	64	80	1	11	3	320	340
2	15	1	34	35	2	6	1	244	260	3	11	1	97	110
0	15	1	270	270	2	7	0	180	180	1	7	5	148	165
4	13	1	40	45	4	2	4	115	135	3	1	1	274	250
4	11	1	147	150	2	13	2	335	350	1	5	5	235	210
4	7	3	287	295	4	7	2	151	160	5	3	1	257	260
2	9	5	123	100	4	1	4	27	20	3	7	1	281	285
4	9	3	329	305	4	3	4	330	345					

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
(i)	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-
(ii)	-	-	-	+	+	+	+	+	-	-	-	-	+	+	+	+
(iii)	+	+	-	-	+	+	-	-	+	+	-	-	+	+	-	-
(iv)	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
(v)	+	-	-	+	+	-	-	+	-	+	+	-	-	+	+	-
(vi)	+	+	-	-	-	-	+	+	-	-	+	+	+	+	-	-
(vii)	-	+	-	+	+	-	+	-	+	-	+	-	-	+	-	+

No matter what sequence of seven signs is chosen, one of the above sixteen will have at least six of the seven correct. The penalty one suffers for having sixteen instead of 128 sequences, which is necessary to ensure one sequence completely correct, is the probability of one sign being wrong.

We can use a pair of signs to represent the quadrant of a phase  $\varphi$ , the signs being those of  $\sin \varphi$  and  $\cos \varphi$  respectively. Thus the row (i) could represent the sign of  $\cos \varphi$  for the enantiomorph fixing reflexion (we can fix  $\sin \varphi$  as positive and so restrict the phase to  $0-\pi$ ) and the other six rows, taken in pairs, can represent the phases of three other reflexions. If one of the signs is wrong then it means that either  $\cos \varphi$  or  $\sin \varphi$  has the wrong sign while the other member of the pair is correct and this will put  $\varphi$  in a neighbouring quadrant to the correct one.

There are various ways one could use this principle and indeed one could use larger substantialization sequences (Good, 1954). For example it is possible to guarantee 14 out of 15 correct signs in 2048 sequences and this would deal with the case  $n=7$ . For higher-symmetry non-centrosymmetric space groups the choice of origin terms often ensures that certain projection reflexions have phases  $0, \pi$  or  $\pm \pi/2$  and in such cases a single row of the substantialized sequences is sufficient to deal with this reflexion.

We believe that by means of this technique of developing phases from a number of initial sets so that relationship (4) can be used immediately, coupled with the use of substantialized sequences and searching figure-of-merit tests it should be possible to solve non-centrosymmetric structures with  $N=200$  for  $Z=4$ ,  $N=150$  for  $Z=2$  and  $N=100$  for  $Z=1$ . These estimates may well be raised by a factor of two when computers capable of  $10^7$  operations per second become available.

### Conclusions

While it cannot be said that the phase problem has been solved, what can certainly be said is that, for moderately complex structures, direct methods offer a high probability of reaching the correct solution. This probability decreases as the complexity of the structure increases. One can increase the probabilities by

allowing for more possible breakdowns of relationships early in the processes, but inevitably this leads to a great increase in the total effort and the limiting factor will often be the speed and size of the available computers. What is certain is that progress can be made in applying phase-determining techniques to the solution of crystal structures of greater complexity than hitherto only if one uses systematic, searching, multi-solution methods for, with a complex structure, it is unlikely that the correct solution can be reached by a straightforward path.

All the suggestions put forward in this paper can be applied with existing computers and it is intended that, when programs implementing these ideas have been written and tested, they will be made available for general use. The limited multi-solution program described in this paper is available now together with complete operating instructions.

This work was begun when we were both in the Physics department of the University of York and we are most grateful to the Science Research Council for financial support of this project. In addition we are grateful for first-class computer facilities both at the Harwell Atlas installation and also at the Louvain Computing Center. One of us (G.G.) is indebted to Prof. M. Van Meerssche for his interest in the progress of this work, and to G. S. D. King, J. P. Putzeis and G. Evrard for providing trial data. We are also pleased to acknowledge some useful suggestions made by Prof. W. Cochran.

### References

- COCHRAN, W. (1952). *Acta Cryst.* **5**, 65.  
 COCHRAN, W. (1955). *Acta Cryst.* **8**, 473.  
 COCHRAN, W. & DOUGLAS, A. S. (1955). *Proc. Roy. Soc. A*, **227**, 486.  
 COCHRAN, W. & WOOLFSON, M. M. (1955). *Acta Cryst.* **8**, 1.  
 GERMAIN, G., PIRET, P. & VAN MEERSSCHE, M. (1963). *Acta Cryst.* **16**, 109.  
 GOOD, I. J. (1954). *Acta Cryst.* **7**, 603.  
 KARLE, I. L. & KARLE, J. (1964a). *Acta Cryst.* **17**, 1356.  
 KARLE, I. L. & KARLE, J. (1964b). *Acta Cryst.* **17**, 835.  
 KARLE, I. L. & KARLE, J. (1966a). *Acta Cryst.* **21**, 860.  
 KARLE, I. L. & KARLE, J. (1966b). *Acta Cryst.* **21**, 849.  
 KARLE, I. L., KARLE, J., OWEN, T. B. & HOARD, J. L. (1965). *Acta Cryst.* **18**, 345.  
 RABINOWITZ, I. N. & KRAUT, J. (1964). *Acta Cryst.* **17**, 159.  
 SAYRE, D. (1952). *Acta Cryst.* **5**, 60.  
 TOLLIN, P. (1962). Ph. D. Thesis, Univ. of Cambridge.  
 WOOLFSON, M. M. (1954). *Acta Cryst.* **7**, 65.  
 WOOLFSON, M. M. (1961). *Direct Methods in Crystallography*. Oxford Univ. Press.  
 ZACHARIASEN, W. H. (1952). *Acta Cryst.* **5**, 68.