Acta Cryst. (1970). B**26**, 274

On the Application of Phase Relationships to Complex Structures
II. Getting a Good Start

BY G. GERMAIN
Laboratoire de Chimie-Physique, Université de Louvain, 39 *Schapenstraat, Louvain, Belgium*

AND P. MAIN AND M. M. WOOLFSON
Physics Department, University of York, England

(*Received* 17 *September* 1969)

The most important element governing the success or otherwise of direct methods is the choice of the starting point from which new phases are developed. This paper describes three ways in which a good start to the phase determining process can be made: (i) Where, early in the process of determining new phases, a phase relationship of low reliability must be used then the phase determined by it is allowed to have alternative values, one of which should be close to the correct value. An application of one variant of this technique is described. (ii) A method is described for finding groups of reflexions very highly interrelated by phase relationships. It is possible to ensure, by using these reflexions as a starting point for phase development, that multiple indications for new phases are soon encountered. (iii) An algorithm is described for generating a schematic route for phase determination which indicates the best phases to use as a starting point for phase development including those which fix the origin and enantiomorph. A formula is derived for estimating the probable standard deviation of a phase, which would be given by a number of different phase relationships, before any phase information is known.

Introduction

Direct methods of solving crystal structures, by their very nature, might seem to be objective mathematical procedures the success of which should be independent of the skill of the user. Nevertheless there still appears to be a 'green fingers' element in operation and sometimes an unsuccessful application of direct methods to a structure by one individual has been followed by a successful application by someone else.

Perhaps the most important element influencing success or failure is the choice of starting point from which is begun the procedure of determining new signs or general phases. The usual techniques require a 'starting-set of phases' (SSP) to which are assigned definite phases or phase symbols and on this base the whole subsequent process depends. For some structures it has been found that there are 'lucky' SSP's which lead to structure solution while the great majority of SSP's do not. One way in which an 'unlucky' SSP may operate for centrosymmetric structures is for there to be an invalid sign relationship in the first few steps of sign determination which leads subsequently to a complete breakdown of the sign-determining process. A method of allowing for such a contingency has been discussed by Germain & Woolfson (1968).

We are now going to describe general approaches to the problem of getting a good starting point. Two types of procedure will be considered:

(1) Whenever a weak link develops in the chain of phase determination, leading to an uncertain new phase or a doubtful new sign, one proceeds thereafter with some alternative values of the new phase or sign.

(2) The strongest possible SSP is chosen so that weak links are eliminated or reduced in number and so that all new phases are determined by strong relationships. Since, for a given structure determination, the number of very strong phase relationships is limited this usually means that one must get multiple indications for new phases as soon as possible.

A combination of procedures (1) and (2) is likely to be useful for dealing with difficult structures. However the computer time required to allow for extreme situations can easily become excessive and solutions should, in the first place, normally be attempted with the minimum of effort by the most straightforward procedures. Only when the simple approach fails should sophisticated and time-consuming techniques be tried.

Forking in non-centrosymmetric structures

Let us consider a situation in which a number of SSP's have been set up in the way described by Germain & Woolfson (1968). Some reflexions, in general three, may be allocated specific phases to fix the origin of the cell, one reflexion may, depending on the space group, be allocated a specific phase or the alternatives ($\pi/4$, $3\pi/4$) to fix the enantiomorph and a number of others can be tried in all possible combinations of ($\pm\pi/4$, $\pm 3\pi/4$), (0, π) or ($\pm\pi/2$) depending on the reflexion type. New phase indications will be given by the well-known tangent formula

$$\tan(\varphi_I)_{\mathbf{h}} \simeq \frac{\sum\limits_{\mathbf{h}'} |E_{\mathbf{h}'}E_{\mathbf{h}-\mathbf{h}'}| \sin(\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'})}{\sum\limits_{\mathbf{h}'} |E_{\mathbf{h}'}E_{\mathbf{h}-\mathbf{h}'}| \cos(\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'})} = \frac{S_{\mathbf{h}}}{C_{\mathbf{h}}}, \quad (1)$$

where $(\varphi_I)_{\mathbf{h}}$ is the indicated value of the phase whose true value is $\varphi_{\mathbf{h}}$.

The strength of the indication can be estimated from

$$\alpha_{\mathbf{h}} = 2\sigma_3\sigma_2^{-3/2}\,|E_{\mathbf{h}}|\,(S_{\mathbf{h}}^2 + C_{\mathbf{h}}^2)^{1/2}\,, \tag{2}$$

where the symbols have been defined by Karle & Karle (1966). If there is only one contributor to the indication for the new phase then we write

$$\alpha_{\mathbf{h}} = \kappa_{\mathbf{h},\mathbf{h}'} = 2\sigma_3\sigma_2^{-3/2}\,|E_{\mathbf{h}}E_{\mathbf{h}'}E_{\mathbf{h}-\mathbf{h}'}|\,. \tag{3}$$

Karle & Karle gave a formula for the variance of a phase indication given by equation (1) in terms of the corresponding value of α. We have calculated the variance, V_α, and standard deviation, σ_α, for a range of values of α and these are given in Table 1.

On the basis of experience it is found that if a newly determined phase is initially within about 45° of its true value then it will in its turn lead to useful new information. The error in the phase will also probably be reduced at some later stage by tangent-formula refinement.

The probability distribution of the phase indication from equation (1) will, for small deviations, be approximately Gaussian and we shall assume in the following discussion that the statistics of the Gaussian distribution may be applied. It can be seen from Table 1 that if $\alpha \simeq 7{\cdot}0$ then 45° represents two standard deviations which means that there is a probability of 0·95 that the indication for the new phase is within 45° of the true value.

If we use the criterion that there must be a probability of 0·95 for the new phase being within 45° of its true value before it is acceptable, how do we proceed when the best new indication corresponds to an α less than 7·0? Let us consider the case $\alpha = 2{\cdot}3$. Since $\sigma_\alpha = 45°$ there is a probability of 0·95 that the true phase, φ_τ, is within 90° of the indicated phase, φ_I. Thus if we take two alternative values for the new phase, $\varphi_I + \sigma_\alpha$ and $\varphi_I - \sigma_\alpha$, there is a 95% probability that one or other of them is within 45° of the true phase. For $\alpha < 2{\cdot}3$ one should have to take three alternative values for the new phase; thus for $\alpha = 1{\cdot}2$, corresponding to $\sigma_\alpha = 67\frac{1}{2}°$ there will be a 95% probability that one of the three values, φ_I, $\varphi_I + 4\sigma_\alpha/3$, $\varphi_I - 4\sigma_\alpha/3$ will be within 45° of φ_τ. For $\alpha < 1{\cdot}2$ one would simply use the technique of allocating the trial phases to different quadrants *i.e.* $\pm\pi/4$, $\pm 3\pi/4$, which gives complete certainty that one of the allocated values is within 45°.

There are of course other criteria which could be used other than those given above. Thus one might wish to restrain the maximum probable error to less than 45° or to take a probability other than 0·95. Table 2 gives the α ranges for the numbers of alternatives required for various combinations of probability and maximum error.

Thus we can see that if we are working to the 0·90 probability limit and to an error limit of 35° then an α equal to 7·8 or more is required for a single alternative to satisfy the chosen limits while with α between 2·6 and 7·8 two alternatives would be necessary.

This provides us with the basis of a branching procedure which, whenever the phase determining process reaches a point where some probability criterion is not satisfied, takes two, three or four alternative routes. There are various ways in which such a process can be operated – for example one can specify that no route should involve more than two 4-branches, two 3-branches and three 2-branches; when this limit is

Table 1. *Variances and standard deviations for a range of* α

α	V_α (radians²)	σ_α (radians)	σ_α (degrees)
0·0	3·290	1·814	103·9
0·5	2·349	1·533	87·8
1·0	1·604	1·267	72·6
1·5	1·091	1·044	59·8
2·0	0·764	0·874	50·1
2·5	0·563	0·750	43·0
3·0	0·437	0·661	37·9
3·5	0·354	0·595	34·1
4·0	0·298	0·546	31·3
5·0	0·227	0·477	27·3
6·0	0·184	0·429	24·6
7·0	0·155	0·394	22·6
8·0	0·134	0·366	21·0
10·0	0·106	0·325	18·6
12·0	0·087	0·295	16·9
14·0	0·074	0·272	15·6
16·0	0·065	0·254	14·6
18·0	0·057	0·239	13·7
20·0	0·051	0·227	13·0

Table 2. α *ranges for various combinations of probability and maximum error*

Probability	0·90				0·95				0·99			
of error less than	30°	35°	40°	45°	30°	35°	40°	45°	30°	35°	40°	45°
No. of alternatives	α	α	α	α	α	α	α	α	α	α	α	α
1												
	10·5	7·8	6·3	5·0	14·7	10·8	8·5	6·8	~25	18·0	14·5	11·2
2												
	3·2	2·6	2·1	1·7$_5$	4·2	3·3	2·7	2·3	6·7	5·1	4·1	3·4
3												
	1·7$_5$	1·3$_5$	1·0	0·6$_5$	2·3	1·8	1·4	1·1$_5$	3·4	2·7	2·2	1·9
4												

reached the straight-through path is followed thereafter. This condition would require up to 1142 routes of phase determination to be explored, a formidable but not impossible task. It is our experience, gained with the quadrant-phase-fixing method that it is rarely necessary to follow through all the possible paths to the end as many soon reveal themselves as implausible by the poor quality of new phase generation.

More complex computer algorithms can be devised – for example alternative routes could be generated in such a way that no route has more than a certain total deviation from the main straight-through path – but in practice simple applications of branching procedures will be found to be the most useful.

A trial of a simple branching technique has been made with a known structure, the photolysis product solved by Karle, Karle & Estlin (1967). Four reflexions were selected to fix the origin and enantiomorph; these reflexions, their $|E|$ values and allocated phases (as given by the published solution) are given in Table 3.

From this beginning it is possible to develop five new phases, each being given by a single relationship. These reflexions, their $|E|$ values and indicated phases are also listed in Table 3 together with the pairs of reflexions which give them. It will be seen that the indicated phases agree quite well with the published phases, given in parentheses, although larger errors would be expected from the standard deviations associated with the phase relationships.

The strategy adopted in this trial was to allow up to two deviations of $\pm 40°$ in the five relationships displayed in Table 3. This gives the 51 possible SSP's which are illustrated in Table 4 where the symbols + and − indicate the sign of the 40° deviation. It should be noted that the deviations are in the *relationships* and not in the phases as given in Table 3. Thus if we follow through set 51 of deviations we find the five derived phases in Table 3 as 40, 170, 100, 170, 80°.

For each of the 51 SSP's there were developed 50 new phases by the application of equation (1) which was also used to refine the phases in the sense of giving them the greatest self-consistency. At this stage there was calculated for each SSP a figure of merit which, in the terminology of equation (1) was

$$Z = \sum_{\mathbf{h}} |E_{\mathbf{h}}|^2 (S_{\mathbf{h}}^2 + C_{\mathbf{h}}^2) . \tag{4}$$

Since the trial was with a known structure it was possible at this stage to look at the values of Z and the corresponding values of $\langle|\Delta\varphi|\rangle$, the mean deviation of phase from the true value. These are listed in order of magnitude of Z in Table 5.

This Table shows a number of extremely interesting features. Firstly pairs of sets, whose set numbers added together equal 52, have very similar values of Z (they usually differ in figures less significant than those given in the Table). An inspection of Table 4 reveals that these correspond to pairs of sets which have deviations in opposite directions. The reason for this is not understood and is probably a feature of the pattern of phase development in this structure. For the pairs of sets with the highest values of Z one set corresponds to a very satisfactory development of phases while the other gives almost random phases.

At this point the 16 sets with the lowest values of Z were discarded, which incidently included set 26, the no-deviation set, and phases were found for another 50 reflexions for each remaining set. The situation at that stage is shown in Table 6.

Another 50 phases were developed for the 14 sets with highest Z in Table 6 and the final Z values (see Table 7) just began to discriminate the good sets from the bad.

In fact the seven good sets were all very similar and could all be developed to give sufficient phases to give a complete solution to the structure. Just as a matter of interest set 26, the no-deviation set, was developed up to the 150 phase stage and gave $\langle|\Delta\varphi|\rangle = 32{\cdot}7°$. It was inferior to some of the other sets but it might have just been good enough to reveal the structure.

While we have carried out preliminary experiments with branching procedures which indicate their high potentiality for coping with complex structures these have not yet been systematically incorporated into any of our phase-determining computer packages. We have, however, put rather more effort into evolving methods of finding favourable SSP's and this work will now be described.

The use of *n*-sets

A phase relationship may be expressed in the form

$$\cos(\varphi_{\mathbf{h}_1} + \varphi_{\mathbf{h}_2} + \varphi_{\mathbf{h}_3}) = 1 - \delta^2 , \tag{5}$$

where $\mathbf{h}_1 + \mathbf{h}_2 + \mathbf{h}_3 = 0$ and δ^2 is, in general, small and has a probability distribution determined by the magnitude of $\kappa_{\mathbf{h},\mathbf{h}'}$. Knowledge of any two phases gives a

Table 3. *E values and allocated phases for the trial structure*

h	k	l	$\lvert E\rvert$	φ	Given by
4	0	7	2·29	90°	Origin and enantiomorph
0	1	15	2·61	270	
5	4	0	2·28	90	
1	0	13	1·69	270	
4	4	13	1·62	0 (342°)	(5 4 0) & (1 0 13)
8	4	6	2·07	90 (92)	(4 0 7) & (4 4 13)
7	4	7	2·04	180 (182)	(1 0 13) & (8 4 6)
6	4	6	2·45	90 (86)	(1 0 13) & (7 4 7)
2	4	13	2·42	0 (342)	(4 0 7) & (6 4 6)

Table 4. *Possible SSP's*

	Set number																																																		
Relationship	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51
1	−	−	−	−	−	−	−	−	−	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	+	+	+	+	+	+	+	+	+
2	−	0	0	0	0	0	0	0	+	−	−	−	−	−	−	−	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	+	+	+	+	+	+	+	−	0	0	0	0	0	0	0	+
3	0	−	0	0	0	0	0	+	0	−	0	0	0	0	0	+	−	−	−	−	−	0	0	0	0	0	0	0	0	0	+	+	+	+	+	−	0	0	0	0	0	+	0	−	0	0	0	0	0	+	0
4	0	0	−	0	0	0	+	0	0	0	−	0	0	0	+	0	−	0	0	0	+	−	−	−	0	0	0	+	+	+	−	0	0	0	+	0	−	0	0	0	+	0	0	0	−	0	0	0	+	0	0
5	0	0	0	−	0	+	0	0	0	0	0	−	0	+	0	0	0	−	0	+	0	−	0	+	−	0	+	−	0	+	0	−	0	+	0	0	0	−	0	+	0	0	0	0	0	−	0	+	0	0	0

probable value for the third. Such a relationship may be expressed in a generalized notation

$$R \equiv \{\varphi_{\mathbf{h}_1}, \varphi_{\mathbf{h}_2}, \varphi_{\mathbf{h}_3}\}\,, \qquad (6)$$

where containing the phases in the curled bracket means that they are related but does not give explicitly the exact form of the relationship.

Now let us consider a pair of relationships such as

$$\begin{aligned} R_1 &\equiv \{\varphi_a, \varphi_b, \varphi_c\}\,, \\ R_2 &\equiv \{\varphi_a, \varphi_b, \varphi_d\}\,. \end{aligned} \qquad (7)$$

It is clear that if φ_a and φ_b are both known then φ_c and φ_d can be found. Other pairs of known phases can give knowledge about the other two, *e.g.* φ_a and φ_c. The feature of the pair of relationships R_1 and R_2 is that they contain two phases in common. If more relationships can be found which also contain this pair of phases then fixing two suitable phases will give phase indications for all the other phases contained in the relationships.

We can express the total information of the relationships (7) in the form

$$R_1 + R_2 \equiv [\varphi_a, \varphi_b]\; \varphi_c, \varphi_d\,, \qquad (8)$$

and we say that φ_c and φ_d are both associated with the 2-set φ_a, φ_b. If there are more relationships associated with the same 2-set we should have

$$R_1 + R_2 + R_3 + \ldots = [\varphi_a, \varphi_b]\; \varphi_c, \varphi_d, \varphi_e, \ldots\,. \qquad (9)$$

Fixing both components of the 2-set or one member of the 2-set and any one other phase leads to indications for all other phases, each given by a single relationship.

There are, of course, three 2-sets associated with each relationship and sets of relationships can be found with interlinked 2-sets of the form shown below:

$$\begin{aligned} R_1 &\equiv \{\varphi_a, \varphi_b, \varphi_c\} \\ R_2 &\equiv \{\varphi_a, \varphi_b, \varphi_d\} \\ R_3 &\equiv \{\varphi_a, \varphi_c, \varphi_e\} \\ R_4 &\equiv \{\varphi_b, \varphi_d, \varphi_f\}\,. \end{aligned} \qquad (10)$$

The total information of relationships (10) may be represented by

$$R_1 + R_2 + R_3 + R_4 \equiv$$
$$[\varphi_a, \varphi_b]\; \varphi_c, \varphi_d \cup [\varphi_a, \varphi_c]\; \varphi_b, \varphi_e \cup [\varphi_b, \varphi_d]\; \varphi_a, \varphi_f \qquad (11)$$

We have borrowed the terminology of set theory to indicate that the individual components on the right-hand side have some duplicated elements. If two phases are fixed, either a pair in a 2-set or one from a 2-set and one phase associated with that 2-set, then all the remaining phases can be found. Finding relationships linked by 2-sets gives a train of phase indications which can ensure the steady development of phases but there is only one relationship for each phase indication and it is unlikely that all the indications will be strong.

Of far more practical application is the exploitation of sets of relationships linked as

$$\begin{aligned} R_1 &\equiv \{\varphi_a, \varphi_b, \varphi_c\} \\ R_2 &\equiv \{\varphi_a, \varphi_d, \varphi_e\} \\ R_3 &\equiv \{\varphi_f, \varphi_b, \varphi_d\} \\ R_4 &\equiv \{\varphi_f, \varphi_c, \varphi_e\} . \end{aligned} \tag{12}$$

We may write

$$\begin{aligned} R_1 + R_2 &\equiv [\varphi_b, \varphi_c]\, \varphi_a + [\varphi_d, \varphi_e] \varphi_a \\ &\equiv [\varphi_b, \varphi_c, \varphi_d, \varphi_e]\, \varphi_a \end{aligned} \tag{13}$$

and φ_a is said to be associated with the 4-set $[\varphi_b, \varphi_c, \varphi_d, \varphi_e]$. If three of the five phases are known, either three from the 4-set or a correctly chosen two from the 4-set and φ_a, then the other two can be found from the pair of relationships.

The other pair of relationships involve the same 4-set and hence we may write

$$R_1 + R_2 + R_3 + R_4 \equiv [\varphi_b, \varphi_c, \varphi_d, \varphi_e]\, \varphi_a, \varphi_f . \tag{14}$$

In this situation a knowledge of three phases leaves three unknown phases and four relationships, and one of the phases is given by a double indication. For example, if φ_b, φ_c and φ_d are known, φ_a may be found from R_1, φ_e from R_2 and then φ_f is given by both R_3 and R_4.

The practical application of 4-sets is quite straightforward. One can frequently find many pairs of relationships with a common 4-set. If k such pairs of relationships are found then they involve $4+k$ phases and there are $2k$ relationships. If three phases are known – for example because they are among those used to fix the origin and enantiomorph – then two more phases are indicated with a single relationship and $k-1$ phases will have double indications.

This general principle for finding highly interrelated phases can be extended. The three relationships

$$\begin{aligned} R_1 &\equiv \{\varphi_a, \varphi_b, \varphi_c\} \\ R_2 &\equiv \{\varphi_a, \varphi_d, \varphi_e\} \\ R_3 &\equiv \{\varphi_a, \varphi_f, \varphi_g\} \end{aligned} \tag{15}$$

can be combined as

$$R_1 + R_2 + R_3 \equiv [\varphi_b, \varphi_c, \varphi_d, \varphi_e, \varphi_f, \varphi_g]\, \varphi_a . \tag{16}$$

If four phases are fixed, all from the 6-set or a correctly chosen three from the 6-set and φ_a, then the remaining three phases are indicated, each by one relationship. However if there are other groups of three relationships giving the same 6-set then multiple phase indications result. As an example we consider in addition to (15) the relationships

Table 5. *Z and* $\langle|\Delta\varphi|\rangle$ *values*

Set	Z	$\langle\lvert\Delta\varphi\rvert\rangle$	Set	Z	$\langle\lvert\Delta\varphi\rvert\rangle$	Set	Z	$\langle\lvert\Delta\varphi\rvert\rangle$
47	2·124 × 10⁵	20·2°	22	1·053 × 10⁵	39·4°	19	0·926 × 10⁵	38·1
5	2·124	80·9	35	1·051	37·0	33	0·926	84·8
40	2·065	22·8	17	1·051	81·6	44	0·845	86·7
12	2·065	88·3	32	1·042	39·8	8	0·845	87·8
24	1·288	78·4	20	1·042	32·7	43	0·828	72·5
28	1·288	22·8	45	1·003	82·1	9	0·828	59·6
14	1·248	77·3	7	1·003	38·6	41	0·740	83·8
38	1·248	25·5	42	1·002	43·9	11	0·740	86·9
6	1·180	79·3	10	1·002	80·1	49	0·736	84·4
46	1·180	27·3	29	1·000	39·0	3	0·736	85·2
34	1·176	76·8	23	1·000	81·8	31	0·671	84·5
18	1·176	32·3	48	0·999	42·3	21	0·671	85·7
51	1·111	88·0	4	0·999	94·0	26	0·529	47·2
27	1·099	47·8	39	0·962	38·9	37	0·521	56·6
25	1·099	80·8	13	0·962	83·0	50	0·521	44·5
1	1·070	91·0	36	0·927	85·8	2	0·518	49·4
30	1·053	32·0	16	0·927	86·7	15	0·517	44·2

Table 6. *Z and* $\langle|\Delta\varphi|\rangle$ *values after discarding of the lowest* 16 *Z-value sets*

Set	Z	$\langle\lvert\Delta\varphi\rvert\rangle$	Set	Z	$\langle\lvert\Delta\varphi\rvert\rangle$	Set	Z	$\langle\lvert\Delta\varphi\rvert\rangle$
45	6·624 × 10⁵	83·6°	40	6·190 × 10⁵	19·1°	34	5·541 × 10⁵	84·1°
28	6·617	16·5	42	6·172	23·9	48	5·290	56·1
17	6·607	83·3	5	6·061	81·3	4	5·156	85·9
7	6·577	17·7	19	6·005	21·2	36	4·828	90·7
10	6·572	82·3	23	6·001	81·4	51	4·806	92·6
24	6·550	81·1	12	5·993	83·0	13	4·710	96·4
35	6·537	18·5	30	5·958	19·1	27	4·566	34·8
14	6·443	82·2	18	5·904	27·0	1	4·539	95·2
38	6·431	17·1	32	5·904	81·6	25	4·348	91·8
46	6·325	17·4	20	5·898	19·5	16	4·317	95·0
47	6·310	18·2	22	5·892	79·5	39	4·131	89·7
6	6·261	80·7	29	5·890	26·8			

$$\begin{array}{ll} R_4\equiv\{\varphi_h\,\varphi_b\,\varphi_e\} & R_7\equiv\{\varphi_i\,\varphi_b\,\varphi_g\} \\ R_5\equiv\{\varphi_h\,\varphi_c\,\varphi_f\} & R_8\equiv\{\varphi_i\,\varphi_c\,\varphi_d\} \\ R_6\equiv\{\varphi_h\,\varphi_d\,\varphi_g\} & R_9\equiv\{\varphi_i\,\varphi_e\,\varphi_f\} \end{array} \qquad (17)$$

which give altogether

$$\sum_{n=1}^{9} R_n\equiv[\varphi_b,\,\varphi_c,\,\varphi_d,\,\varphi_e,\,\varphi_f,\,\varphi_g]\;\varphi_a,\,\varphi_h,\,\varphi_i\,. \qquad (18)$$

If four phases are known then two more will be found from single indications, two from double indications and one from a triple indication.

A general program has been written which, to a limited extent, exploits the n-set principle to find efficient SSP's. The process will be described with reference to data for the photolysis product, listed in Table 8.

(*a*) For each reflexion the quantity

$$\psi_{\mathbf{h}}=\sum_{\mathbf{h}'}\kappa_{\mathbf{h},\mathbf{h}'} \qquad (19)$$

is computed where the summation is over all relationships in which the reflexion is involved.

(*b*) Four reflexions are chosen to fix the origin and enantiomorph in such a way that for the four reflexions $\sum\psi_{\mathbf{h}}$ is as large as possible. For the photolysis product these were found to be reflexions with code numbers 2, 29, 104 and 141.

A search is now made for the m other reflexions, where m is an input parameter of the program, which, with the original four, will give the most efficient SSP. Efficiency is judged by the criterion that the least value of α for a phase determination should be as high as possible.

For the present example we take $m=2$.

(*c*) The first listed reflexion was found which was not one of the four found in (*b*) and had a value of ψ greater than 0·2 $\psi_{\max}$, where $\psi_{\max}$ is the maximum ψ for all the reflexions. This reflexion, that with code number 1 in Table 8, we shall refer to as r_1.

(*d*) All the relationships in which the reflexion is involved are found. There are 36 such relationships for r_1 and the six strongest of these are given below

$$\begin{array}{ll} \varphi_1+\varphi_{70}\;-\varphi_{74}\qquad\simeq 0\,, & \kappa=3{\cdot}20 \\ \varphi_1+\varphi_{141}-\varphi_{144}\qquad\simeq 0\,, & \kappa=3{\cdot}02 \\ \varphi_1-\varphi_{126}-\varphi_{127}+\pi\simeq 0\,, & \kappa=2{\cdot}76 \\ \varphi_1-\varphi_{151}-\varphi_{152}\qquad\simeq 0\,, & \kappa=2{\cdot}70 \\ \varphi_1+\varphi_{109}-\varphi_{113}\qquad\simeq 0\,, & \kappa=2{\cdot}61 \\ \varphi_1+\varphi_{171}-\varphi_{175}\qquad\simeq 0\,, & \kappa=2{\cdot}23 \end{array}$$

From these relationships all the 6300 4-sets with which r_1 is linked are generated. If in any of these 4-sets two members are known [because they are two of the four fixed in (*b*)] then fixing the remaining two will give a double indication for the phase of r_1. In fact none of the 4-sets with which r_1 is associated had two known members so the next reflexion in Table 8 was found satisfying the conditions given in (*c*). This was the reflexion with code number 3, r_3.

One of the 4-sets with which r_3 was associated had two known members. This is derived from

$$\varphi_2-\varphi_3-\varphi_6\simeq 0\,,\quad \kappa=2{\cdot}70$$

and

$$\varphi_3-\varphi_{140}-\varphi_{141}+\pi\simeq 0\,,\quad \kappa=1{\cdot}52$$

giving the 4-set $[\varphi_2,\,\varphi_6,\,\varphi_{140},\,\varphi_{141}]$.

If φ_6 and φ_{140} are fixed then φ_3 is given by a double indication.

(*e*) Starting with the seven phases, φ_2, φ_3, φ_6, φ_{29}, φ_{104}, φ_{140}, φ_{141} the complete set of relationships is explored to see how a development of phases might progress. It is assumed that when a phase is given by two or more relationships the value of α is given by $\alpha_\kappa=\sum\kappa$, an assumption which will be dealt with in the following section. The process is continued until a point is reached where no new phase can be found or until the total value of $\sum\alpha_\kappa$ for all the reflexions which have been found equals $\frac{4}{3}\sum\kappa$ for all the relationships available.

The minimum value of α_κ is noted for the phase development from this set of seven phases.

(*f*) The reflexions following r_3 in Table 8 are explored in sequence. If they satisfy the conditions in (*c*) then they are dealt with in the way described for r_3. An updated record is kept of that pair of reflexions which, with the original four, gives the highest minimum value of α_κ and the value of $(\alpha_\kappa)_{\min}$.

For the application of the method to the photolysis product it was found that fixing φ_{127} and φ_{146} gave a minimum α_κ of 3·0 for the remainder of the phase determination. While the process may seem to be very time consuming there are various time-saving features – for example no schematic phase development is followed if the minimum α_κ is less than the highest minimum previously recorded. The process outlined above took two minutes on an IBM 7090 computer.

If a seven reflexion starting point was permitted then the minimum α_κ was about 4·0. In an attempt to simu-

Table 7. *Final Z values*

Set	Z	$\langle\lvert\Delta\varphi\rvert\rangle$	Set	Z	$\langle\lvert\Delta\varphi\rvert\rangle$	Set	Z	$\langle\lvert\Delta\varphi\rvert\rangle$
40	$1{\cdot}191\times10^6$	18·6°	46	$1{\cdot}191\times10^6$	18·4°	17	$1{\cdot}069\times10^6$	81·4°
7	1·191	18·5	47	1·191	18·5	10	1·062	84·5
35	1·191	18·6	24	1·122	81·1	45	1·062	84·5
38	1·191	18·5	14	1·122	81·1	42	1·018	41·6
28	1·191	18·5	6	1·111	81·4			

late a difficult structure 6-sets were used to find a starting point of ten reflexions which give a minimum α_κ equal to about 6·0.

In the above work it was assumed that α_κ gave an indication of the expected value of α. This would imply that where there were several indications for a given phase they were all completely consistent; since this is rarely true α_κ will normally be an over-estimate of α. We shall consider now how to get a better estimate of α from the individual values of κ.

Table 8. *Data for the photolysis product*

CODE	H	K	L	E	Ø	CODE	H	K	L	E	Ø	CODE	H	K	L	E	Ø
1	0	0	12	2.38	180	2	0	1	15	2.58	270	3	0	2	8	2.04	180
4	0	2	10	1.31	0	5	0	3	5	1.36	270	6	0	3	7	2.11	90
7	0	3	13	1.90	270	8	0	3	14	2.30	90	9	0	4	4	1.51	180
10	0	4	8	2.19	0	11	0	4	11	1.79	0	12	0	4	12	1.44	180
13	0	4	14	1.94	0	14	0	4	17	1.96	180	15	0	4	19	1.45	0
16	0	5	7	2.40	270	17	0	5	16	1.34	270	18	0	5	17	1.42	90
19	0	6	5	1.79	180	20	0	6	10	1.33	180	21	0	6	11	1.35	0
22	0	6	13	1.50	180	23	0	6	16	1.80	0	24	0	6	17	1.53	0
25	0	6	18	1.86	180	26	1	0	4	1.42	180	27	1	0	8	1.59	180
28	1	0	11	1.28	90	29	1	0	13	1.69	270	30	1	1	3	1.35	262
31	1	1	7	1.38	199	32	1	1	14	1.51	178	33	1	2	0	1.46	270
34	1	2	2	1.43	334	35	1	2	12	1.62	66	36	1	2	14	1.46	215
37	1	3	3	1.25	137	38	1	3	12	1.65	197	39	1	4	4	1.41	248
40	1	4	6	1.25	73	41	1	4	9	1.29	146	42	1	4	10	1.31	315
43	1	4	14	1.91	93	44	1	4	18	1.42	295	45	1	5	9	1.43	19
46	1	5	12	1.38	331	47	1	5	13	1.89	222	48	1	5	16	1.28	40
49	1	5	18	1.52	296	50	1	6	0	1.46	270	51	1	6	13	1.53	308
52	1	6	15	1.54	154	53	1	6	17	1.25	47	54	2	1	1	1.91	303
55	2	1	3	1.53	125	56	2	1	13	1.40	110	57	2	1	15	1.38	288
58	2	1	19	1.41	131	59	2	2	1	1.54	348	60	2	2	4	1.38	8
61	2	2	14	1.84	168	62	2	2	18	1.45	16	63	2	3	4	1.30	200
64	2	3	5	1.37	272	65	2	3	6	1.92	260	66	2	3	7	1.49	65
67	2	3	9	1.50	129	68	2	3	12	1.35	314	69	2	3	14	1.28	150
70	2	4	1	2.23	157	71	2	4	5	1.96	351	72	2	4	8	1.73	19
73	2	4	10	2.05	220	74	2	4	13	2.38	341	75	2	4	17	1.86	195
76	2	4	18	1.66	200	77	2	5	7	1.49	259	78	2	5	8	1.30	332
79	2	5	10	1.45	144	80	2	5	12	1.86	55	81	2	5	14	1.27	258
82	2	6	8	2.14	168	83	2	6	9	1.74	26	84	2	6	15	1.65	326
85	3	1	4	1.34	3	86	3	1	12	1.50	36	87	3	1	14	1.30	222
88	3	1	15	1.37	297	89	3	2	7	1.31	17	90	3	3	12	1.67	176
91	3	4	2	1.34	211	92	3	4	3	1.92	324	93	3	4	7	1.30	119
94	3	4	10	1.62	225	95	3	4	14	1.27	57	96	3	4	15	1.27	140
97	3	5	1	1.25	308	98	3	5	4	1.42	65	99	3	5	6	2.13	187
100	3	5	9	1.68	324	101	3	5	16	1.52	180	102	4	0	0	1.80	0
103	4	0	4	1.85	180	104	4	0	7	2.31	90	105	4	0	9	1.26	270
106	4	0	11	1.58	270	107	4	0	12	1.52	180	108	4	1	15	1.25	281
109	4	2	0	2.51	180	110	4	2	2	1.49	332	111	4	2	3	1.36	39
112	4	2	7	1.40	328	113	4	2	12	1.73	338	114	4	2	14	1.37	168
115	4	3	7	1.32	134	116	4	3	9	1.80	12	117	4	4	1	1.36	144
118	4	4	7	1.97	146	119	4	4	8	2.47	16	120	4	4	10	1.25	162
121	4	4	11	1.36	199	122	4	4	13	1.58	342	123	4	5	0	1.60	0
124	4	5	2	1.40	154	125	4	5	4	1.35	210	126	4	5	5	2.10	91
127	4	5	7	2.18	266	128	4	5	8	1.64	266	129	4	5	10	1.42	54
130	4	5	14	1.41	343	131	4	6	3	1.27	321	132	4	6	5	1.78	212
133	4	6	6	1.34	318	134	4	6	8	1.98	152	135	4	6	9	2.02	39
136	4	6	10	1.31	31	137	4	6	12	1.31	208	138	5	0	8	1.46	180
139	5	2	6	1.86	183	140	5	2	8	1.35	322	141	5	4	0	2.27	90
142	5	4	5	1.28	316	143	5	4	7	1.44	215	144	5	4	12	2.21	228
145	5	4	13	1.61	180	146	5	4	14	2.11	29	147	5	4	15	1.45	347
148	5	5	0	1.50	270	149	5	5	1	1.73	110	150	5	5	2	1.55	127
151	5	5	5	1.68	45	152	5	5	7	2.67	234	153	5	5	12	1.67	7
154	5	6	4	1.50	98	155	5	6	11	1.48	145	156	5	6	14	1.32	190
157	6	1	8	1.43	27	158	6	2	0	1.27	180	159	6	2	1	1.37	329
160	6	2	5	1.53	94	161	6	2	7	1.77	265	162	6	3	5	1.30	49
163	6	3	13	1.43	9	164	6	3	14	1.53	115	165	6	4	1	1.91	156
166	6	4	2	1.45	302	167	6	4	4	1.87	199	168	6	4	6	2.40	86
169	6	4	8	1.48	312	170	6	4	12	1.63	144	171	6	5	0	1.93	0
172	6	5	7	1.51	278	173	6	5	10	1.42	328	174	6	5	11	2.07	63
175	6	5	12	1.92	165	176	6	5	13	1.73	260	177	6	6	2	1.47	126
178	6	6	3	1.40	94	179	6	6	4	1.33	333	180	7	3	1	1.35	89
181	7	3	10	1.33	48	182	7	4	5	1.51	36	183	7	4	7	1.98	182
184	7	4	10	1.29	125	185	7	5	3	1.34	263	186	7	5	9	1.44	254
187	7	6	3	1.38	92	188	7	6	5	1.49	238	189	8	0	7	1.72	90
190	8	2	5	1.56	87	191	8	2	7	1.85	289	192	8	3	3	1.45	10
193	8	3	5	1.44	224	194	8	4	1	1.45	111	195	8	4	3	1.43	239
196	8	4	4	1.59	266	197	8	4	6	1.97	92	198	8	5	0	2.97	0
199	8	5	1	1.38	318	200	8	5	2	2.36	187	201	8	6	2	1.49	85

Estimating α in the absence of phase information

The way in which the different components of the tangent formula combine to give an estimate of $\varphi_{\mathbf{h}}$ is shown in the form of a phase-vector diagram in Fig. 1. The amplitudes of the components of the diagram are of length $\kappa_{\mathbf{h},\mathbf{h}'}$ and they have phases $(\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'})$. The resultant has length $\alpha_{\mathbf{h}}$ and its phase, $(\varphi_I)_{\mathbf{h}}$ is the indicated value of the phase whose true value is $\varphi_{\mathbf{h}}$.

Now we can also write

$$\alpha_{\mathbf{h}}=[\{\sum_{\mathbf{h}'}\kappa_{\mathbf{h},\mathbf{h}'}\sin(\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'}-\varphi_{\mathbf{h}})\}^2 +\{\sum_{\mathbf{h}'}\kappa_{\mathbf{h},\mathbf{h}'}\cos(\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'}-\varphi_{\mathbf{h}})\}^2]^{1/2}\,, \quad (20)$$

since this corresponds to a rotation of Fig. 1 through an angle $-\varphi_{\mathbf{h}}$, as is shown in Fig. 2, and clearly leaves the resultant unchanged in magnitude.

Cochran (1955) found the probability distribution for $\varphi_{\mathbf{h}}$ for a fixed $\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'}$ and given value of $\kappa_{\mathbf{h},\mathbf{h}'}$. This is

$$P(\varphi_{\mathbf{h}})=[2\pi I_0(\kappa_{\mathbf{h},\mathbf{h}'})]^{-1}\exp[\kappa_{\mathbf{h},\mathbf{h}'}\cos(\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'}-\varphi_{\mathbf{h}})]\,, \quad (21)$$

where I_0 is a modified Bessel function.

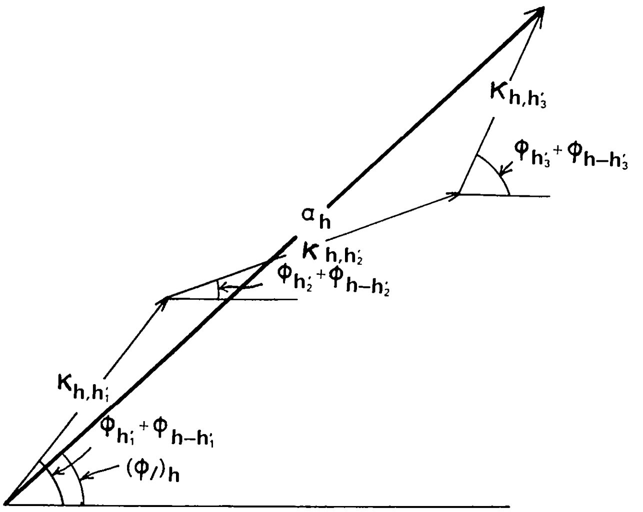

Fig. 1. The composition of different phase indications $\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'}$ with amplitudes $\kappa_{\mathbf{h},\mathbf{h}'}$ to give a resultant of amplitude $\alpha_{\mathbf{h}}$ and phase $(\varphi_I)_{\mathbf{h}}$.

We can use this probability distribution to find the mean value of $\alpha_{\mathbf{h}}^2$ for a given set of $\kappa_{\mathbf{h},\mathbf{h}'}$. Thus, from equation (20) we find

$$\begin{aligned}\alpha^2=&\sum_{\mathbf{h}'}\kappa_{\mathbf{h},\mathbf{h}'}^2\\ &+\sum_{\substack{\mathbf{h}'\\ \mathbf{h}'\neq\mathbf{h}''}}\sum_{\mathbf{h}''}\kappa_{\mathbf{h},\mathbf{h}'}\kappa_{\mathbf{h},\mathbf{h}''}\cos(\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'}-\varphi_{\mathbf{h}})\\ &\qquad\cos(\varphi_{\mathbf{h}''}+\varphi_{\mathbf{h}-\mathbf{h}''}-\varphi_{\mathbf{h}})\\ &+\sum_{\substack{\mathbf{h}'\\ \mathbf{h}'\neq\mathbf{h}''}}\sum_{\mathbf{h}''}\kappa_{\mathbf{h},\mathbf{h}'}\kappa_{\mathbf{h},\mathbf{h}''}\sin(\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'}-\varphi_{\mathbf{h}})\\ &\qquad\sin(\varphi_{\mathbf{h}''}+\varphi_{\mathbf{h}-\mathbf{h}''}-\varphi_{\mathbf{h}})\,. \quad (22)\end{aligned}$$

The expectation value of $\cos(\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'}-\varphi_{\mathbf{h}})$ may be found from (21) and is

$$\langle\cos(\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'})\rangle=\int_0^{2\pi}P(\varphi_{\mathbf{h}})\cos(\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'}-\varphi_{\mathbf{h}'})\,d\varphi_{\mathbf{h}}\,. \quad (23)$$

The exponential term in equation (21) may be expanded as a sum of modified Bessel functions, *viz.*

$$\exp[\kappa_{\mathbf{h},\mathbf{h}'}\cos(\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'}-\varphi_{\mathbf{h}})] =I_0(\kappa_{\mathbf{h},\mathbf{h}'})+2\sum_{m=1}^{\infty}I_m(\kappa_{\mathbf{h},\mathbf{h}'})\cos m(\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'}-\varphi_{\mathbf{h}})\,. \quad (24)$$

Substituting from (24) into (23) and integrating we find

$$\langle\cos(\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'}-\varphi_{\mathbf{h}})\rangle=\frac{I_1(\kappa_{\mathbf{h},\mathbf{h}'})}{I_0(\kappa_{\mathbf{h},\mathbf{h}'})}\,. \quad (25)$$

Similarly it can be found that

$$\langle\sin(\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'}-\varphi_{\mathbf{h}})\rangle=0\,. \quad (26)$$

Putting expectation values into equation (22) we find

$$\langle\alpha_{\mathbf{h}}^2\rangle=\sum_{\mathbf{h}}\kappa_{\mathbf{h},\mathbf{h}'}^2+2\sum_{\substack{\mathbf{h}'\\ \mathbf{h}'\neq\mathbf{h}''}}\sum_{\mathbf{h}''}\kappa_{\mathbf{h},\mathbf{h}'}\kappa_{\mathbf{h},\mathbf{h}''} \times\frac{I_1(\kappa_{\mathbf{h},\mathbf{h}'})}{I_0(\kappa_{\mathbf{h},\mathbf{h}'})}\,\frac{I_1(\kappa_{\mathbf{h},\mathbf{h}''})}{I_0(\kappa_{\mathbf{h},\mathbf{h}''})}\,. \quad (27)$$

This formula is easy to apply to find a value of $\langle\alpha_{\mathbf{h}}^2\rangle^{1/2}$ when a number of κ's are known. For use in a computer for a range of κ up to 5·0 it is found that

$$I_1(\kappa)/I_0(\kappa)\simeq 0\cdot5658\kappa-0\cdot1304\kappa^2+0\cdot0106\kappa^3 \quad (28)$$

gives a maximum error of about 4%. Some values are given in Table 9.

As an example of the application of equation (27) we consider a structure containing 100 equal atoms in

Table 9. *Values of $I_1(\kappa)/I_0(\kappa)$ for various κ*

κ	$I_1(\kappa)/I_0(\kappa)$	Series	κ	$I_1(\kappa)/I_0(\kappa)$	Series
0·0	0·000	0·000	2·5	0·765	0·765
0·5	0·243	0·252	3·0	0·810	0·810
1·0	0·446	0·446	3·5	0·841	0·837
1·5	0·596	0·591	4·0	0·863	0·855
2·0	0·698	0·695	4·5	0·880	0·871
			5·0	0·893	0·894

Table 10. *The first part of the computer output*

```
185   0.00
121   0.00
 63   0.00
 53   0.00
 48   0.00
 41   0.00
 21   0.00
 20   0.00
 18   0.00
 17   0.00
 78   1.22
                  -54     135     1.22
186   1.23
                  101    -104    -1.23
 98   1.25
                   -1     101     1.25
101   1.22
                  -45    -104    -1.22
 45   1.24
                  104     150    -1.24
 37   1.27
                    2     -35    -1.27
 97   1.31
                  -54    -141     1.31
173   1.33
                  -54     135     1.33
 44   1.34
                  103     146     1.34
162   1.36
                  -35    -152     1.36
 15   1.37
                  106     119    -1.37
 96   1.41
                    1      92     1.41
 40   1.42
                  109     139     1.42
180   1.42
                  -54     141     1.42
131   1.48
                    1    -135    -1.48
188   1.50
                  -70     139    -1.50
 49   1.56
                  106     152    -1.56
158   1.63
                   82     119     1.63
 81   1.77
                    2     -70    -1.77
181   1.99
                    2    -182    -1.26       35    -200    -1.23
  4   2.02
                  -10      25     1.29       70      83    -1.23
 91   2.04
                    1     -94    -1.25      -82     139    -1.29
142   2.05
                   29    -119     1.30       54      99    -1.26
199   2.06
                 -103     126     1.30     -104     128    -1.27
137   2.06
                   19     104     1.31      -54     174    -1.26
155   2.07
                   14    -139    -1.31       61     -92     1.27
 93   2.11
                  -43     104    -1.39      141     189     1.23
 95   2.12
                  -82    -139     1.23      104     183    -1.41
 42   2.14
                  -92     104    -1.41     -103     146     1.24
184   2.15
                  -55     152     1.28       92     104     1.39
 46   2.18
                  -90     109     1.40      104     151    -1.30
 24   2.28
                    1      19     1.58      -60      74     1.22
148   2.28
                   -1     153     1.45        2    -147     1.36
153   1.68
                   38     109     1.68
147   2.27
                   52    -109     1.36       88    -198     1.43
140   2.30
                    3    -141    -1.52       27     109     1.31
```

Table 11. *The final portion of the computer output*

```
 118   6.07

  65   5.08

   8   5.25

  56   5.09

  59   5.43

 126   5.58

   1   4.17

  57   6.19


  74   5.71

 123   6.43

 108   6.72

 102   6.51

 189   6.24

 168   6.42

  70   4.60

 161   5.24

  61   4.93

  66   5.29

   3   5.22

 119   4.51

  77   4.77

 134   4.80

  10   5.64

  16   4.75

 157   5.84

  82   5.11

 165   3.92

  72   2.89

  54   0.00
 109   4.86

   6   2.81

 191   0.00
 127   3.63

 198   0.00
 141   0.00
  29   0.00

   2   0.00
 104   0.00
```

```
  54     65    1.75
 -61   -161   -1.56

   8    -82    2.02
 -59    127    1.56

  54     74   -2.24
 157    168    1.69

   1     54    1.54
 -82   -126    1.53

 -70   -109    2.09
 168   -191   -1.66

   1   -127   -2.64

 -70     74    3.06

   6     72   -1.22
  66    119   -1.23
 -82   -127    1.56

 -74    102    1.24

  16   -104   -2.15
-102    198    2.07

   2    102    1.41
  10    127    1.45

 -10    119    2.36
-104    189    1.73

  16   -198   -2.97
  72    165    1.38

  54    127   -2.42

 -72   -104   -2.16

   2   -157    1.58
  66   -198   -1.90

   2    -54   -2.20
  66    127   -1.45

   2    -72   -1.61
 -54    119    1.70

   2     -6    2.69

   2   -127    3.37

   2    -72    1.61
 -54    134    1.37

   2   -127   -2.70

   2     -6   -2.89

  54     72    1.92

   6    165   -1.40
  72   -198    1.78

 -54    127   -2.16

  54   -198   -2.63

   6     54    1.69


  -6    127    2.80

 191    198   -2.81


 104   -198   -3.63
```

```
  56    -65    1.28
 109    191    2.22

  16     59    1.72


  56     70   -1.54


  72    126    1.23
-127   -168   -1.78

-109    165    1.79


  54    168    2.33

 168    168    1.66

 -16    -72   -1.39
  72    127    1.26


 104    168   -3.20

  54    -70   -1.65
 127    189    1.46

   2   -109   -1.96
 -16   -119   -1.80

 -16    127    2.28


  61    161    1.36


 -61   -134    2.12

 127   -157   -1.69

   3   -165   -1.67


   6     77   -1.40


   3    -54   -1.41
 -77   -109   -1.35

-109    119    3.07

 -54     77    1.70

   2    -82   -2.00


  10    109    2.64

   2    -16    3.29

 -54     82    2.38

  54    104    1.53
 -82    198   -2.20

 -72   -109    2.25

  72   -104   -1.85

  54    127   -1.75


-104    191    2.60
```

```
  59     82    1.57


  59    126   -1.51


  57    -70    1.52


  74    123   -1.29


-119    161    1.63


 -74   -157    1.73



  16     82   -1.72
 -77   -119   -1.23


-168    189    2.38

  54    165    1.42


   6    119   -1.58
  16    134   -1.44

 -70    165    1.86


  70   -168    2.23


 -70    104   -3.00

-134    161    1.90

  61   -104   -1.82


 -16    -66   -1.60


  10     54   -1.51




  10    -54   -1.51




-157   -165    1.59

  54    191   -1.23


-165    191    1.83
```

the unit cell and $|E_{\mathbf{h}}|=3{\cdot}0$ with three contributors for which $|E_{\mathbf{h}'}E_{\mathbf{h}-\mathbf{h}'}|$ equals 3·0, 4·0 and 5·0. The corresponding values of κ are 1·8, 2·4 and 3·0 and the value of $\langle\alpha_{\mathbf{h}}^2\rangle^{1/2}$ computed from equation (27) is 6·1 compared with the value 7·2 which would result if the three phase indications agreed precisely.

It is worthwhile to use values of $\langle\alpha_{\mathbf{h}}^2\rangle^{1/2}$ rather than $\sum\kappa_{\mathbf{h},\mathbf{h}'}$ in assessing the probable reliability of a new phase indication for the n-set method. This is particularly true when there are a large number of weak indications as otherwise the probable reliability of the phase indication can be greatly overestimated.

The convergence method

We shall now describe an alternative method of testing the effectiveness of using a particular set of origin and enantiomorph-fixing reflexions and of selecting other reflexions to get the best possible SSP. The method will be described by its application to the photolysis product data listed in Table 8. The origin and enantiomorph combination tested is that used for the n-set example, *i.e.* reflexions with code numbers 2, 29, 104 and 141.

The first step is to work out the value of $\langle\alpha^2\rangle^{1/2}$ for each of the reflexions from the κ's of the relationships in which it is involved. The lowest value of $\langle\alpha^2\rangle^{1/2}$ is sought and in fact it is found that for ten reflexions $\langle\alpha^2\rangle^{1/2}=0$. These reflexions are removed from the data and their computer-program code numbers and values of $\langle\alpha^2\rangle^{1/2}$ are listed, as shown in Table 10. Reflexion 78 is that with the next lowest value of $\langle\alpha^2\rangle^{1/2}$, equal to 1·22. This is listed in Table 10 with the value of $\langle\alpha^2\rangle^{1/2}$ and the single relationship which gives it is indicated by the numbers which follow, which indicate that

$$(\varphi_I)_{78}=-\varphi_{54}+\varphi_{135}\,,\quad \kappa=1{\cdot}22\,.$$

Reflexion 78 is now removed from the list of reflexions and the phase relationship is removed from the list of relationships. The values of $\langle\alpha^2\rangle^{1/2}$ for reflexions 54 and 135 are modified to allow for the removal of the phase relationship and once more the reflexion with the lowest value of $\langle\alpha^2\rangle^{1/2}$ is sought. As before this is listed, removed from the set of reflexions, all relationships in which it is involved removed from the set of relationships and new modified values of $\langle\alpha^2\rangle^{1/2}$ are found. This process is continued with the proviso that the origin and enantiomorph fixing reflexions are never removed from the system but remain to the very end. Table 10 shows a reproduction of the beginning of the computer output for the photolysis-product problem. When a negative κ is shown it implies that there is a π to be incorporated in the phase relationship. Thus, for the first reflexion with a double indication for phase, that with code number 181, we have

$$\varphi_{181}\simeq\varphi_2-\varphi_{182}+\pi\,,\quad \kappa=1{\cdot}26$$

and

$$\varphi_{181}\simeq\varphi_{35}-\varphi_{200}+\pi\,,\quad \kappa=1{\cdot}23$$

giving overall $\langle\alpha^2\rangle_{181}^{1/2}=1{\cdot}99$.

However, the interesting and most useful part of the computer output comes at the end, where the process 'converges' on the four starting reflexions, and this is reproduced in Table 11.

It should be noticed that, from the way this convergence tabulation is produced, the relationships indicating any particular phase involve only those reflexions which have not yet been eliminated from the Table and hence are listed somewhere below. The other characteristic of the convergence process is that at each stage it removes from the system that reflexion which is least well linked to the remainder and thus it tends to enrich the quality of the interrelationships of those which remain.

Let us now examine Table 11 to see what it tells us. Starting from the bottom there are the four reflexions which fix the origin and enantiomorph and for these $\langle\alpha^2\rangle^{1/2}=0$ as must be so since such reflexions cannot be interelated. There follows reflexion 198 which is not linked to any below and then we find reflexion 127

Table 12. *The number of reflexions which must be fixed for different values of* α

Minimum α	Number to fix	Code numbers
2·0	7	2, 29, 54, 104, 141, 191, 198
2·5	7	2, 29, 54, 104, 141, 191, 198
3·0	9	2, 6, 29, 54, 72, 104, 141, 191, 198
3·5	11	2, 6, 29, 54, 72, 99, 104, 141, 149, 191, 198
4·0	14	2, 6, 29, 31, 54, 72, 99, 104, 127, 141, 149, 165, 191, 198

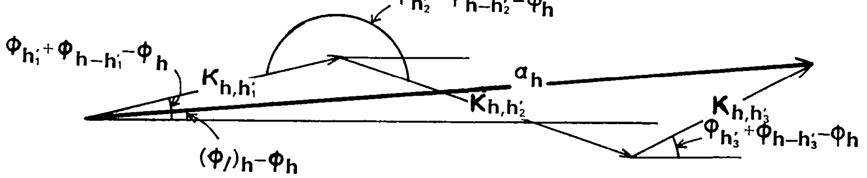

Fig. 2. Changing the phase indications in Fig. 1 to $\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'}-\varphi_{\mathbf{h}}$ leaves the magnitude of the resultant unchanged.

linked to 104 and 198 by a fairly strong phase relationship. The next reflexion 191 has no contributors and it is followed by reflexion 6 with one fairly strong phase indication. We then find the first double indication with a correspondingly large value of $\langle\alpha^2\rangle^{1/2}$, 4·86, and next reflexion 54 with no contributor. From this point onwards there are multiple indications with large values of $\langle\alpha^2\rangle^{1/2}$ and, if the phase development up to reflexion 72 has been satisfactory, it should stand a good chance of progressing successfully thereafter.

If it is decided, on the basis of a judgement of the complexity of the problem, that one should not rely on any α less than, say, 2·5 then it is straightforward from the tabulated values of $\langle\alpha^2\rangle^{1/2}$ to decide which reflexions should be incorporated in the starting set. With this α limit for the present problem one needs a starting set of 7 reflexions and the computer program outputs their code numbers. Of course one does not scan the whole convergence table or one will end up deciding that all the reflexions listed in Table 10 need to be incorporated in the starting set. In practice it is sufficient to examine the first 50 or so steps in the schematic phase development process.

In Table 12 are shown the number of reflexions which need to be fixed for values of α from 2·0 to 4·0 by steps of 0·5 for the photolysis product.

The convergence program is quite rapid; with 200 reflexions, 800 relationships and a medium-speed computer (70,000 operations per second) the whole process takes two minutes. Thus it is possible to run the program on a number of trial sets of origin and enantiomorph fixing reflexions. To assist the judgement of which set is to be preferred one may use as a figure-of merit

$$C=\sum_{r=1}^{100}\langle\alpha^2\rangle_r^{1/2},$$

where the summation is over the first 100 stages of phase development for which phases are not fixed.

It must be appreciated that the convergence mapping is a *schematic* phase-development scheme and the actual progress of phase development will be dictated by *computed* values of α and not the estimated values, $\langle\alpha^2\rangle^{1/2}$. In practice the first steps of an actual phase development usually follow closely the schematic scheme. When deviations do occur it is often due to having calculated values of α which are higher than the corresponding values of $\langle\alpha^2\rangle^{1/2}$ so that the actual path of phase development is better than that indicated in the convergence map.

It is possible to use the principle of convergence mapping in a more elaborate way in which one repeatedly obtains an updated estimate of which new reflexions should be fixed. To illustrate this let us suppose that a starting set is chosen from a convergence map for a structure with space group $P2_12_12_1$ and the map indicates that the four origin and enantiomorph fixing reflexions plus two others will enable some progress to be made with phase development. One of the two extra reflexions might, let us say, be a general one and tried in four quadrants, $(\pm\pi/4, 3\pi/4)$ while the other may be special reflexion $(0, \pi)$ or $(\pi/2, -\pi/2)$. The eight starting points can be used to find ten new phases and the group of ten will, in general, be different for each of the starting points. New convergence maps, one for each of the eight starting points, converging on the sixteen reflexions with known phase may indicate that for some starting points two more reflexions need to be fixed, for others only one or perhaps none at all.

There will obviously be a time penalty to pay for such a scheme but it might be worth while if one is dealing with a difficult structure.

Implementation

The techniques described in this paper are being incorporated in a non-centrosymmetric program package, written in FORTRAN conforming to the A.S.A. standard.

The package will be flexible so that, for example, an '*n*-sets' or 'convergence' routine can be used or can be bypassed. The system will offer an automatic route for structure determination which will not require any intervention on the part of the user although it will be possible for an educated user to exercise judgement and to intervene. For the automatic route, decision-making processes, such as convergence mapping, will be used to optimize as far as possible the probability of finding the correct solution.

The extensive collaboration between our laboratories has been made possible by the generous financial support of the North Atlantic Treaty Organization. Two of us (P.M. & M.M.W.) would like to express our gratitude for continued support by the Science Research Council. The success of our work has been assisted by the excellence of the computer facilities provided both at Louvain and at York, and also at Pittsburgh where one of us (P.M.) spent a period developing the *n*-set method. We have also been encouraged by the interest of Professor M. van Meerssche in our project.

References

COCHRAN, W. (1955). *Acta Cryst.* **8**, 473.

GERMAIN, G. & WOOLFSON, M. M. (1968). *Acta Cryst.* **B24**, 91.

KARLE, J. & KARLE, I. L. (1966). *Acta Cryst.* **21**, 849.

KARLE, I. L., KARLE, J. & ESTLIN, J. A. (1967). *Acta Cryst.* **23**, 494.